

The Sparse Law of Deep-Space Communication: A Unified Theory of Feasibility, Efficiency, and Architectural Design

J. Councilman

Abstract

Deep-space communication networks operate under sparse, intermittent, and geometry-constrained connectivity. This work develops the *Sparse Law*, an exact factorization of end-to-end delivery ratio into a structural feasibility factor and a protocol-dependent conditional efficiency: $DR = S_{\text{full}}^T \cdot \eta$. The factorization follows from the event structure of feasible and delivered paths and holds for all delay-tolerant networking implementations.

We establish oracle fidelity as a first-order concern for architectural evaluation, derive oracle implementation requirements, and show that contact plan quality can reverse architectural conclusions. We introduce admissible custody chains, a protocol failure taxonomy, and structural invariants that all deep-space architectures must obey.

The Contact-Density Bound is established as a proposition under renewal assumptions, and non-ergodic degradation is proved by construction: for architecture families with ergodic contact generation, a scalar contact-time density determines structural feasibility; non-ergodic processes fall strictly below this bound. Empirical validation across Mars relay and cislunar architectures confirms the factorization, reveals that body-centric Mars is feasibility-dominated ($\eta > 0.97$) while cislunar is efficiency-dominated ($S_{\text{full}}^T = 1$, $\eta \approx 0.68$), and demonstrates that routing efficiency is empirically anticorrelated with connectivity across the tested architectures under bounded-rationality routing. Heliocentric Mars (four-tier relay architecture, 156 configurations across a full synodic cycle) transitions to an efficiency-dominated regime where η follows $\ln(\eta) \approx a - 0.8 d_{\text{AU}}$ and Lagrange-point relays eliminate conjunction blackouts. A TTL extension preserves the factorization for finite-lifetime bundles, and cross-validation with Contact Graph Routing (oracle-class CGR) confirms that 89–100% of the greedy efficiency gap is recoverable within the tested architectures, though at the cost of fragility to schedule uncertainty.

1 Introduction

Deep-space communication differs from terrestrial networking not merely by degree but by kind. Visibility is intermittent, opportunities for contact are sparse, and path availability is dictated by celestial mechanics rather than by infrastructure placement. The delay-tolerant networking (DTN) architecture [?] and bundle protocols [?, ?] were introduced to cope with such environments, but performance analysis often relies on simulation and empirical studies [?, ?] without a unifying structural law.

To reason about deep-space systems, one requires a theoretical model that separates what the *architecture* makes possible from what the *protocol* actually achieves. A protocol cannot deliver a bundle if no time-respecting path exists in the contact graph, but the existence of such a path does not guarantee that the protocol will realize it.

The objective of this work is to construct a rigorous, minimal, and unified theory of deep-space delivery performance: the *Sparse Law*. It applies equally to the Mars relay network [?], the Cislunar communications and navigation architecture [?], and lunar Gateway operations [?]. The factorization is definitional; its utility

derives from the empirical finding that η is approximately invariant across architectures within a regime, enabling the separation of geometric and protocol design.

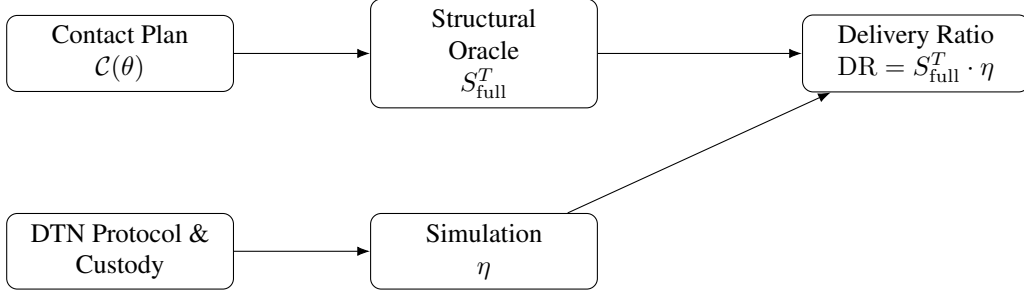


Figure 1: Conceptual structure of the Sparse Law. The contact plan induces a structural feasibility factor via the oracle; protocol behavior induces a conditional efficiency via simulation; their product yields the delivery ratio.

2 Mathematical Objects

We work over a fixed horizon $[0, T]$. A contact plan induces a time-varying directed graph $G(t) = (V, E(t))$ whose edges represent feasible transmissions between nodes, as a function of time, visibility, and link budget. This view is closely related to temporal graph and time-expanded graph models used in DTN routing analysis [?, ?] and in directed temporal-network percolation theory [?, ?].

A bundle b injected at node s at time t with destination d gives rise to two events:

- $F(t)$: there exists a time-respecting path from s to d under the contact plan, respecting transmission timing and capacity;
- $D(t)$: the bundle injected at time t is delivered to d before expiration under the actual DTN protocol execution.

Their indicators are

$$\phi_b = I(F(t)), \quad \delta_b = I(D(t)).$$

Symbols and Notation

For reference, we collect the main symbols used throughout this paper:

Symbol	Meaning
$G(t)$	Time-varying directed contact graph at time t
$\mathcal{C}(\theta)$	Contact plan induced by architecture design θ
$F(t)$	Event: a time-respecting path from s to d exists at injection time t
$D(t)$	Event: the bundle injected at time t is delivered before expiration
ϕ_b	Feasibility indicator: $\phi_b = I(F(t))$
δ_b	Delivery indicator: $\delta_b = I(D(t))$
χ_b	Custody-chain indicator (valid custody chain from source to destination)
S_{full}^T	Structural feasibility factor over horizon $[0, T]$
η	Conditional efficiency: $\mathbb{P}(\delta_b = 1 \mid \phi_b = 1)$
DR	End-to-end delivery ratio: $\mathbb{P}(D(t))$
t_k	Discretized injection time, $t_k = k\Delta t$
$I(t_k)$	Oracle output: feasibility indicator at t_k
θ	Architecture design vector (constellation, links, policies)
ω	Randomness representing traffic, failures, protocol dynamics
N	Number of Monte–Carlo replications per injection time
M	Number of feasible injection times in the discretization

Model Assumptions

The results in this paper rest on a small set of explicit assumptions:

Assumption 1 (Deterministic Contact Plan). *For a fixed architecture design θ , the contact plan $\mathcal{C}(\theta)$ is deterministic over the horizon $[0, T]$ and known to the oracle. It captures line-of-sight visibility, link budgets, and scheduled link allocations, but does not include random link failures unless explicitly modeled.*

Assumption 2 (Ideal Custody Semantics). *Every delivered bundle has a custody chain terminating at its destination (or authorized sink), and no custody chain terminates at the destination without delivery being registered. At any time between injection and delivery, custody is held by exactly one node.*

Assumption 3 (Sufficient Clock Coherence). *Nodes share a time base that is precise enough that contact windows and timers can be interpreted consistently across the network. Small clock offsets may exist but do not change the existence of time-respecting paths.*

Assumption 4 (Protocol Conformance). *The DTN protocol and custody mechanisms conform to the DTN architecture and bundle protocol specifications (e.g., [?, ?, ?]), so that custody and delivery events obey the semantics assumed in the theory.*

3 Structural Relations

A fundamental structural relation governs all DTN architectures:

$$D(t) \subseteq F(t).$$

Delivery is only possible when feasibility holds; the realized path taken by any delivered bundle must lie entirely within the contact and capacity constraints that define feasible transmission.

This inclusion leads naturally to a decomposition of delivery probability into a structural factor and a conditional behavioral factor.

4 The Sparse Law

Define

$$S_{\text{full}}^T = \mathbb{P}(F(t)), \quad \eta = \mathbb{P}(\delta_b = 1 \mid \phi_b = 1).$$

The end-to-end delivery ratio is

$$\text{DR} = \mathbb{P}(D(t)).$$

Using the inclusion $D(t) \subseteq F(t)$ and the law of total probability, we obtain the Sparse Law:

$$\text{DR} = S_{\text{full}}^T \cdot \eta.$$

This identity holds for all architectures, DTN routing strategies, and custody policies, provided they respect the basic DTN forwarding semantics of the bundle protocol [?, ?].

5 Formal Results

In this section we record the central structural results underlying the Sparse Law: delivery implies feasibility, delivery factorizes into feasibility and conditional efficiency, and custody chains realize delivery under ideal semantics.

Lemma 1 (Delivery Implies Feasibility). *For each injection time $t \in [0, T]$, let $D(t)$ be the event that the bundle injected at time t is delivered before expiration and let $F(t)$ be the event that there exists a time-respecting path from source s to destination d . Then*

$$D(t) \subseteq F(t),$$

and for every bundle b ,

$$\delta_b = 1 \implies \phi_b = 1.$$

Proof. By definition, $F(t)$ depends only on the contact plan and link capacities: it asserts that at least one time-respecting path exists from s to d for a bundle injected at time t . In contrast, $D(t)$ refers to the actual protocol execution [?, ?]: a particular sequence of transmissions and custody events that may or may not realize such a path.

If $D(t)$ occurs, then the realized trajectory of the bundle from s to d is a sequence of transmissions that respect all contact windows and capacity constraints. This realized trajectory is therefore a time-respecting path in the sense used to define $F(t)$, consistent with temporal graph reachability [?]. Hence, whenever $D(t)$ occurs, $F(t)$ must also occur, and we have $D(t) \subseteq F(t)$.

At the indicator level, this means that if a bundle is delivered ($\delta_b = 1$), then feasibility holds ($\phi_b = 1$). \square

Proposition 2 (Sparse Law Factorization). *Let $\text{DR} = \mathbb{P}(D(t))$ denote the end-to-end delivery ratio and $S_{\text{full}}^T = \mathbb{P}(F(t))$ the structural feasibility factor. Define the conditional efficiency*

$$\eta = \mathbb{P}(\delta_b = 1 \mid \phi_b = 1).$$

Then

$$\text{DR} = S_{\text{full}}^T \cdot \eta.$$

Proof. By the law of total probability,

$$\mathbb{P}(D(t)) = \mathbb{P}(D(t) \mid F(t)) \mathbb{P}(F(t)) + \mathbb{P}(D(t) \mid F(t)^c) \mathbb{P}(F(t)^c).$$

From Lemma 1, $D(t) \subseteq F(t)$, so $\mathbb{P}(D(t) \mid F(t)^c) = 0$. Thus

$$\mathbb{P}(D(t)) = \mathbb{P}(D(t) \mid F(t)) \mathbb{P}(F(t)).$$

Identifying $\text{DR} = \mathbb{P}(D(t))$, $S_{\text{full}}^T = \mathbb{P}(F(t))$ and $\eta = \mathbb{P}(\delta_b = 1 \mid \phi_b = 1) = \mathbb{P}(D(t) \mid F(t))$ yields

$$\text{DR} = S_{\text{full}}^T \cdot \eta.$$

□

Lemma 3 (Custody Chains Realize Delivery). *Let χ_b be the indicator that bundle b has a valid custody chain from source to destination (all custody transitions admissible). Under Assumption 2 and DTN custody semantics [?, ?],*

$$\delta_b = \chi_b$$

for all bundles b . In particular,

$$\eta = \mathbb{P}(\delta_b = 1 \mid \phi_b = 1) = \mathbb{P}(\chi_b = 1 \mid \phi_b = 1).$$

Proof. By the definition of custody transfer in DTN [?, ?], a delivered bundle must have a custody chain terminating at the destination (or authorized sink), since custody must be held continuously by some node from injection until final receipt. Thus $\delta_b = 1$ implies $\chi_b = 1$.

Conversely, under ideal semantics, a custody chain cannot terminate at the destination without triggering delivery; there is no notion of a “phantom” custody completion. Thus $\chi_b = 1$ implies $\delta_b = 1$.

Combining both implications yields $\delta_b = \chi_b$ for all b , and the identity for η follows immediately by conditioning on feasibility. □

Corollary 4 (TTL Factorization). *For bundles with finite time-to-live τ , define the TTL-constrained feasibility event $F_\tau(t)$: there exists a time-respecting path from s to d with total traversal time at most τ . Let $S_T^\tau = \mathbb{P}(F_\tau(t))$ and $\eta_\tau = \mathbb{P}(D(t) \mid F_\tau(t))$. Then*

$$\text{DR}_\tau = S_T^\tau \cdot \eta_\tau.$$

Since $F_\tau(t) \subseteq F(t)$, the chain $\text{DR}_\tau \leq S_T^\tau \leq S_{\text{full}}^T$ holds, with equality in the second inequality when all feasible paths complete within τ .

Proof. Any delivered bundle must traverse a path completing within the TTL, so $D(t) \subseteq F_\tau(t) \subseteq F(t)$. Apply Proposition 2 with $F_\tau(t)$ in place of $F(t)$; the factorization follows identically via the law of total probability. □

Remark. Empirically, η_τ is approximately invariant across DSN configurations on the Mars relay network (CV = 0.2% over five DSN allocations from 8 to 20 h/sol), confirming that the TTL factorization separates geometric from protocol effects in the emergency regime just as the full factorization does in the normal regime. Mars-C in Table 1 illustrates: at TTL = 1 h, the binding constraint is $S_T^\tau = 0.280$ (most injection times lack a path completing within 1 h), while $\eta_\tau = 0.943$ indicates that the protocol delivers 94% of TTL-feasible bundles.

6 Theory Core

6.1 Admissible Custody Transitions

A custody transition $(n_i, t_i) \rightarrow (n_{i+1}, t_{i+1})$ is admissible when:

1. $t_{i+1} \geq t_i$ and any TTL constraint is respected;
2. a feasible contact window exists between n_i and n_{i+1} with sufficient capacity;
3. node n_{i+1} has adequate buffer to accept custody;
4. local custody policy admits acceptance, consistent with DTN forwarding rules [?, ?].

A custody chain is valid if all transitions satisfy these constraints.

6.2 Protocol Failure Taxonomy

Given feasibility ($\phi_b = 1$), failures arise exclusively from:

1. fragmentation failure,
2. custody refusal or drop,
3. routing misprediction [?],
4. timer expiration or lifetime expiry.

6.3 Architecture Invariants

The Sparse Law and the lemmas above yield several invariants:

- $DR \leq S_{\text{full}}^T$ (no architecture can exceed its feasibility envelope);
- Feasibility is monotone in connectivity; efficiency need not be;
- Architectures with equal S_{full}^T form equivalence classes distinguished only by η ;
- Custody admissibility is a system-level requirement, not merely an implementation detail.

6.4 Additional Structural Lemmas

Lemma 5 (Variance of the Efficiency Estimator). *Let $\mathcal{K}_F = \{k : I(t_k) = 1\}$ be the index set of feasible injection times under the discretization $\{t_k\}_{k=0}^K$, and let $M = |\mathcal{K}_F|$ denote its cardinality. For each feasible injection time t_k and Monte–Carlo replicate $j = 1, \dots, N$, let $\delta_b^{(j)}(t_k) \in \{0, 1\}$ be the delivery indicator and define*

$$\hat{\eta}_N = \frac{1}{M} \sum_{k \in \mathcal{K}_F} \left(\frac{1}{N} \sum_{j=1}^N \delta_b^{(j)}(t_k) \right).$$

Assume that, for each fixed k , the variables $\{\delta_b^{(j)}(t_k)\}_{j=1}^N$ are independent and identically distributed with

$$\mathbb{E}[\delta_b^{(j)}(t_k)] = p_k, \quad \text{Var}(\delta_b^{(j)}(t_k)) = p_k(1 - p_k),$$

and that different (k, j) pairs are independent. Then

$$\mathbb{E}[\hat{\eta}_N] = \frac{1}{M} \sum_{k \in \mathcal{K}_F} p_k$$

and

$$\text{Var}(\hat{\eta}_N) = \frac{1}{M^2} \sum_{k \in \mathcal{K}_F} \frac{p_k(1-p_k)}{N} \leq \frac{1}{4MN}.$$

Proof. Linearity of expectation yields

$$\mathbb{E}[\hat{\eta}_N] = \frac{1}{M} \sum_{k \in \mathcal{K}_F} \mathbb{E} \left[\frac{1}{N} \sum_{j=1}^N \delta_b^{(j)}(t_k) \right] = \frac{1}{M} \sum_{k \in \mathcal{K}_F} \frac{1}{N} \sum_{j=1}^N \mathbb{E}[\delta_b^{(j)}(t_k)] = \frac{1}{M} \sum_{k \in \mathcal{K}_F} p_k.$$

For the variance, independence across (k, j) implies

$$\text{Var}(\hat{\eta}_N) = \frac{1}{M^2} \sum_{k \in \mathcal{K}_F} \text{Var} \left(\frac{1}{N} \sum_{j=1}^N \delta_b^{(j)}(t_k) \right) = \frac{1}{M^2} \sum_{k \in \mathcal{K}_F} \frac{1}{N^2} \sum_{j=1}^N \text{Var}(\delta_b^{(j)}(t_k)).$$

Using $\text{Var}(\delta_b^{(j)}(t_k)) = p_k(1-p_k)$, we obtain

$$\text{Var}(\hat{\eta}_N) = \frac{1}{M^2} \sum_{k \in \mathcal{K}_F} \frac{1}{N^2} \cdot N p_k(1-p_k) = \frac{1}{M^2} \sum_{k \in \mathcal{K}_F} \frac{p_k(1-p_k)}{N}.$$

Since $p_k(1-p_k) \leq 1/4$ for all $p_k \in [0, 1]$, we have

$$\text{Var}(\hat{\eta}_N) \leq \frac{1}{M^2} \sum_{k \in \mathcal{K}_F} \frac{1}{4N} = \frac{1}{M^2} \cdot M \cdot \frac{1}{4N} = \frac{1}{4MN},$$

which completes the proof. \square

Lemma 6 (Monotonicity of Structural Feasibility). *Let \mathcal{C} and \mathcal{C}' be two contact plans over the same node set V and horizon $[0, T]$. Suppose that \mathcal{C}' is an augmentation of \mathcal{C} in the sense that, for every time t , the set of available directed links and their capacities under \mathcal{C} is a subset of those under \mathcal{C}' . Let $F(t)$ and $F'(t)$ denote the feasibility events, and S_{full}^T and $S_{\text{full}}^{T'}$ the corresponding feasibility factors. Then*

$$F(t) \subseteq F'(t) \quad \text{for all } t \in [0, T],$$

and consequently

$$S_{\text{full}}^T \leq S_{\text{full}}^{T'}.$$

Proof. Fix t and consider a bundle injected at time t . Under \mathcal{C} , the event $F(t)$ asserts the existence of at least one time-respecting path from s to d satisfying the contact and capacity constraints of \mathcal{C} . Under \mathcal{C}' , all links and capacities present in \mathcal{C} are still available, possibly with additional links or higher capacities.

Any time-respecting path feasible under \mathcal{C} is therefore also time-respecting and feasible under \mathcal{C}' , since all its edges remain available with at least the same capacity. Thus, if $F(t)$ occurs under \mathcal{C} , then $F'(t)$ occurs under \mathcal{C}' , and we have $F(t) \subseteq F'(t)$.

Taking probabilities over t ,

$$S_{\text{full}}^T = \mathbb{P}(F(t)) \leq \mathbb{P}(F'(t)) = S_{\text{full}}^{T'},$$

establishing the desired monotonicity. \square

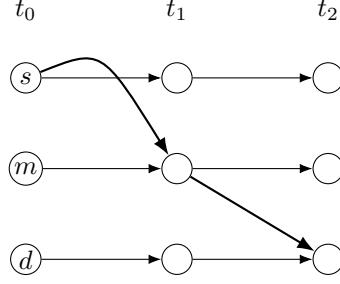


Figure 2: Illustrative time-expanded contact graph. Horizontal edges represent waiting; diagonal edges represent transmissions. A feasible path from (s, t_0) to (d, t_2) certifies feasibility at injection time t_0 .

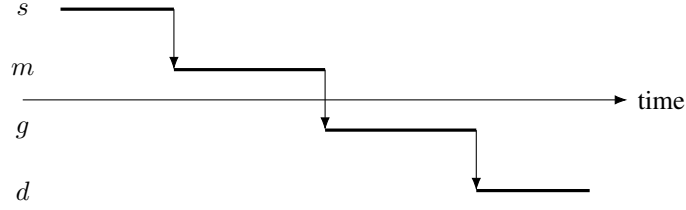


Figure 3: Example custody chain for a single bundle: custody is held by source s , then relay m , then gateway g , and finally destination d . Vertical arrows indicate admissible custody transitions.

6.5 Oracle Fidelity

Definition 1 (Faithful Contact Plan). *A contact plan $\mathcal{C}(\theta)$ is faithful to architecture θ if every contact window in $\mathcal{C}(\theta)$ corresponds to a physically realizable transmission opportunity under the true orbital dynamics, link budgets, and capacity constraints of θ , and every such opportunity is represented in $\mathcal{C}(\theta)$.*

Proposition 7 (Fidelity Invariance). *The Sparse Law factorization $\text{DR} = S_{\text{full}}^T \cdot \eta$ holds for any contact plan $\mathcal{C}(\theta)$, faithful or not. The structural feasibility factor S_{full}^T is defined relative to the contact plan provided.*

Proof. The derivation of Proposition 2 depends only on the event inclusion $D(t) \subseteq F(t)$ and the law of total probability. Neither step requires that $\mathcal{C}(\theta)$ accurately represent the physical environment. The feasibility event $F(t)$ is defined with respect to whatever contact plan is supplied; the factorization follows identically. \square

Corollary 8 (Contact Plan Divergence). *Let \mathcal{C} and \mathcal{C}' be two contact plans for the same architecture θ , derived from different dynamical models. The corresponding feasibility factors may differ: $S_{\text{full}}^T \neq S_{\text{full}}^{\prime T}$. Both factorizations are exact; architectural conclusions drawn from them may contradict each other.*

Remark. *This is not a theoretical curiosity. In a cislunar DTN ablation study, two contact plan generators applied to the same 10-node constellation produced halo-relay removal penalties of 22 and 1.4 percentage points respectively. The divergence was traced entirely to contact plan fidelity: one generator used static alignment assumptions; the other used physics-grounded ephemeris with proper visibility geometry. The Sparse Law held exactly in both cases — the oracle simply answered a different question.*

6.6 The Contact-Density Bound

Let $C = \frac{1}{T} \int_0^T r(t) dt$ denote the time-averaged contact-time density for a source–destination pair, where $r(t)$ is the instantaneous contact rate.

Proposition 9 (Contact-Density Sufficiency). *Let the contact process between each link of a source–destination pair be a stationary renewal process with independent inter-contact times. If the surviving links share a common rate structure — parameterised by a scalar contact-time density $C = \lambda \bar{d}$, where λ is the contact rate and \bar{d} the mean contact duration — then*

$$S_{\text{full}}^T = f(C)$$

for some monotone non-decreasing function $f: \mathbb{R}_{\geq 0} \rightarrow [0, 1]$.

Proof. Under stationarity and independence, the probability that at least one contact occurs on a given link in an interval $[t, t+\Delta]$ is determined by λ and Δ alone. For a multi-hop path $s = v_0, v_1, \dots, v_k = d$, a time-respecting route from injection at time t requires contacts (v_i, v_{i+1}) at times $t \leq t_1 < t_2 < \dots < t_k \leq t + T_{\text{ttl}}$ respecting causality and transmission delay. Under independence, the probability of this event is a deterministic function of the per-link rates $\{\lambda_{ij}\}$ and the bundle lifetime T_{ttl} .

Scalar reduction. When the surviving links share a common rate structure — as holds for constellation thinning within a Walker family, where removing an orbital plane scales all links from that node to zero while preserving homogeneity among survivors — the rate vector collapses to a scalar $C = \lambda \bar{d}$ and S_{full}^T depends on C alone.

Monotonicity. By the standard thinning–superposition coupling, a renewal process with rate $\lambda' > \lambda$ can be decomposed as the original rate- λ process plus an independent surplus process. Under this coupling, the contact plan C' contains C pathwise, and Lemma 6 (monotonicity of S_{full}^T under contact-plan augmentation) applies directly. Therefore f is non-decreasing.

Convergence. By the pointwise ergodic theorem, the time-averaged feasibility

$$\frac{1}{T} \int_0^T I_{\text{reachable}}(t) dt \xrightarrow{a.s.} \mathbb{E}[I_{\text{reachable}}] = f(C) \quad \text{as } T \rightarrow \infty.$$

□

Remark. *The independence assumption is a modelling convenience, not a necessary condition. Regular orbital mechanics produces quasi-periodic contact processes with strong mixing properties; empirically, the time-averaged feasibility converges to a deterministic function of the contact rate within 3–5 orbital periods (Section 10). The operational requirement is stationarity and sufficient mixing, not strict independence. Constellation thinning (removing whole orbital planes from a Walker pattern) preserves these properties; stochastic contact-window deletion does not.*

For architectures with heterogeneous link rates — e.g., disparate surface-to-orbit and orbit-to-Earth capacities — the scalar C is insufficient and S_{full}^T depends on the full rate vector. The scalar sufficiency result applies within architecture families whose link-rate structure is homogeneous up to the parameter being varied.

Proposition 10 (Existence of Non-Ergodic Degradation). *There exist non-stationary contact plans with contact-time density C for which $S_{\text{full}}^T < f(C)$, where f is the sufficiency function of Proposition 9 evaluated on the ergodic family.*

Proof. By construction. Consider an 8-polar cislunar constellation generating contact plans via deterministic Keplerian dynamics. Systematic thinning to 1 polar satellite yields contact-time density $C \approx 11$ h/day and $S_{\text{full}}^T = 0.910$; this value lies on the ergodic curve f .

Now apply independent Bernoulli deletion with probability $p = 0.9$ to each contact window of the full 8-polar plan. The resulting plan has contact-time density $C' \approx 9.5$ h/day and $S_{\text{full}}^T = 0.845$. On the ergodic (thinning) family, the nearest data point is $C = 11.5$ h/day (one polar satellite) with $f(11.5) = 0.910$. Since f is monotone non-decreasing (Proposition 9) and changes slowly near the saturation plateau, $f(9.5) > 0.845$ by continuity; in particular, $S_{\text{full}}^T = 0.845 < f(C')$: the non-stationary plan falls strictly below the ergodic bound at matched contact density. □

Remark. The degradation mechanism is twofold. First, random deletion breaks the phase structure of orbital contact patterns, creating temporal gaps that do not respect the quasi-periodic contact cadence; this reduces S_{full}^T below the ergodic value. Second, the irregular contact topology induces dead-end routing configurations that the greedy algorithm cannot resolve, degrading η from 0.996 to 0.547 — a 7:1 leverage ratio between the DR gap and the S_{full}^T gap (Figure 12).

This distinction carries operational significance: deterministic constellation reconfiguration (e.g., de-commissioning a satellite) preserves the ergodic contact structure, while stochastic scheduling disruptions may degrade delivery performance beyond what contact-volume metrics predict. Whether the degradation is generic for all non-stationary processes — not merely exhibited by a constructed example — remains open.

Corollary 11. Since $\eta \leq 1$, it follows immediately that

$$\text{DR}(C) \leq S_{\text{full}}^T(C) \leq f(C).$$

The contact-density bound is a geometric ceiling that no protocol can exceed, and ergodic contact generation is sufficient to achieve it.

6.7 Supercritical Efficiency Bound

Observation 12 (Supercritical Efficiency Bound). In the supercritical regime ($S_{\text{full}}^T \approx 1$), the conditional efficiency under custody-based forwarding satisfies

$$\eta \approx \eta_0 = \frac{s}{s+1}, \quad s = \frac{\gamma}{\delta},$$

where γ is the effective link-service rate and δ is the per-hop routing overhead rate (next-hop computation, queue scheduling, custody handshake).

Heuristic argument. Model the bundle lifecycle as a four-state continuous-time Markov chain with states SOURCE, RELAY, DELIVERED, LOST and transition rates:

$$\begin{array}{ll} \text{SOURCE} \xrightarrow{\alpha} \text{RELAY} & (\text{find first-hop contact}) \\ \text{SOURCE} \xrightarrow{\beta} \text{LOST} & (\text{no outbound path within TTL}) \\ \text{RELAY} \xrightarrow{\gamma} \text{DELIVERED} & (\text{custody transfer to destination}) \\ \text{RELAY} \xrightarrow{\delta} \text{LOST} & (\text{dead-end, timeout, or custody break}) \end{array}$$

In the supercritical regime, the structural oracle confirms path existence for nearly all injection times; therefore the source-stage failure rate $\beta \rightarrow 0$. The source stage succeeds with probability $\alpha/(\alpha + \beta) \rightarrow 1$, and all loss concentrates in the relay stage. The CTMC collapses to a single-stage race between delivery (γ) and relay failure (δ):

$$\eta_0 = \frac{\gamma}{\gamma + \delta} = \frac{s}{s + 1}.$$

The memoryless (exponential) holding-time assumption is justified by renewal-theoretic arguments: the relay stage aggregates over multiple contact windows drawn from a quasi-periodic schedule, and the aggregate success/failure process converges to memoryless behaviour under sufficient mixing. \square

Remark. Bottleneck dominance. The reduction from a multi-node network to a single effective stage is a structural consequence of the supercritical condition, not an approximation. When $S_{\text{full}}^T = 1$, the source stage cannot fail; every bundle enters the relay pipeline. Stage-failure instrumentation confirms: in the 8-polar cislunar baseline, the source stage registers zero failures across all 288 injection times. The 12-node topology collapses to its rate-limiting link.

Remark. Validation scope. *The rate ratio s is fitted from the 8-polar baseline ($s = 2.11$, predicting $\eta_0 = 0.679$ against measured $\eta = 0.681$). This agreement confirms internal consistency of the CTMC model at one operating point but does not constitute an independent prediction. A stronger test requires either (a) deriving γ and δ from contact-plan observables (mean inter-contact time, custody handshake duration) without reference to measured η , or (b) fitting s on the 8-polar configuration and predicting η for the 4-polar and 1-polar configurations without refitting. This cross-validation is deferred to future work.*

Remark. Cross-regime test. *The Mars relay network (2-hop: surface \rightarrow orbiter \rightarrow Earth) yields $\eta \approx 0.97$ under matched protocol conditions, compared to $\eta \approx 0.68$ for the cislunar network (3–4 hop). This confirms that $\eta_0 = s/(s+1)$ is architecture-dependent: the per-hop overhead rate δ accumulates over the relay chain, reducing s as hop count increases. The efficiency ceiling is not a universal constant but a predictable function of relay depth. Note that the $\eta \approx 0.97$ figure is body-centric (surface-to-orbiter only); in the heliocentric regime (Section 9.5), where the full Earth–Mars distance enters the contact process, η drops to 0.40–0.59 across four architecture tiers, with distance as the governing variable.*

6.8 Constellation Heterogeneity and the Architecture Fingerprint

The rate ratio $s = \gamma/\delta$ encodes the balance between productive forwarding and routing overhead for a given architecture. Empirical analysis across 24 configurations on two bodies reveals that s is governed primarily by the number of distinct orbital regimes K in the constellation, not by inter-satellite link (ISL) presence or graph degree.

Definition 2 (Orbital Regime Count). *The orbital regime count K of a constellation is the number of distinct orbital families whose members contribute to relay routing (e.g., $K = 1$: all polar; $K = 2$: polar + ELFO; $K = 3$: polar + ELFO + halo).*

Observation 13 (Heterogeneity Drives Efficiency). *The architecture fingerprint s depends on K as follows:*

Architecture	K	s	η_0	Mechanism
Moon 8P (no ISL)	1	159	0.994	Circulant symmetry
Moon 8P (ISL)	1	50	0.981	Circulant, ISL mesh
Moon 8P + ELFO	2	6.1	0.859	Dead-end ambiguity (proposed)
Moon 8P + ELFO + halo	3	8.9	0.899	Halo partial rescue (proposed)
Mars 6P (no ISL)	1	339	0.997	Forced routing (proposed)
Mars 6P (ISL)	1	294	0.997	ISL irrelevant

Graph degree does not predict s ($R^2 = 0.003$). Adding ISL contacts to a $K = 1$ constellation decreases s (Moon: 159 \rightarrow 50) by introducing longer relay chains without reducing dead-ends. The critical variable is whether foreign orbital regimes create dead-end configurations that bounded-rationality routing cannot escape.

Remark. Mechanism. *At $K = 1$ (homogeneous constellation), the contact graph has circulant symmetry: every relay satellite offers equivalent forwarding options, and the greedy router cannot enter a dead-end. At $K \geq 2$, a bundle entering the foreign regime (e.g., an ELFO relay in a polar constellation) may encounter a configuration with no outbound contact within its remaining lifetime. The dead-end probability is $\geq 96\%$ algorithmic (estimated): the structural oracle confirms a path exists, but the greedy router’s visited-set constraint and local utility maximization prevent it from finding the path. This interpretation is supported by the CGR recovery result (Observation 15), which substantially reduces the dead-end effect via global re-planning.*

Corollary 14 (Design Rule). *Since $\eta_0 = s/(s+1)$ is monotone increasing in s , and $s = \gamma/\delta$, efficiency improves by either increasing the link-service rate γ (more contact capacity, higher data rates) or decreasing the routing overhead rate δ (faster next-hop computation, reduced custody latency, better lookahead). In the efficiency-dominated regime where $S_{\text{full}}^T \approx 1$, improving the routing algorithm (reducing δ) is more cost-effective than adding constellation capacity (increasing γ), because $|\partial\eta_0/\partial\delta| / (\partial\eta_0/\partial\gamma) = s$: when s is of order unity, a unit reduction in δ gains s times the efficiency of a unit increase in γ .*

Observation 15 (CGR Recovery of Efficiency). *The efficiency gap $1 - \eta$ under greedy routing is predominantly algorithmic, not structural. Replacing the greedy next-hop heuristic with Contact Graph Routing (CGR) [?, ?] — which computes a full shortest path via Dijkstra at each hop and re-plans on failure — recovers 89–100% of the efficiency gap across all tested architectures under oracle-class CGR:*

Architecture	η_{greedy}	η_{CGR}	s_{greedy}	Recovery
Moon $K=1$ (8 polar)	0.981	0.998	52	89%
Moon $K=2$ (8 polar + ELFO)	0.859	0.999	6.1	99%
Mars $K=1$ (6 polar)	0.988	1.000	85	100%
Mars $K=2$ (6 polar + ecc.)	0.986	1.000	69	100%

Under CGR, the effective routing overhead $\delta \rightarrow 0$ and $s \rightarrow \infty$, yielding $\eta_0 \rightarrow 1$ as predicted by the CTMC model (Proposition 12). Mars achieves perfect efficiency ($\eta_{\text{CGR}} = 1.000$) on every feasible bundle; the cislunar gap narrows from $\Delta\eta = 0.14$ (greedy) to $\Delta\eta = 0.002$ (CGR). The largest improvement occurs at Moon $K=2$, where heterogeneous relay geometry creates dead-end configurations that the greedy router cannot escape but CGR resolves via global path computation.

Remark. Fragility caveat. CGR’s advantage depends on the fidelity of the contact plan used for path computation. Under schedule uncertainty ($\sigma \geq 30$ s Gaussian jitter on contact boundaries), CGR’s efficiency drops sharply ($\eta : 0.998 \rightarrow 0.702$) because planned paths miss their contact windows. The greedy router, which makes local decisions without relying on future contacts, is robust to the same perturbation. This fragility bounds the practical regime of CGR applicability to systems with high contact-plan fidelity [?].

7 Numerical Framework

Discretize $[0, T]$ into $\{t_k = k\Delta t\}_{k=0}^K$. For each t_k , the Structural Feasibility Oracle computes $I(t_k) \in \{0, 1\}$.

$$S_{\text{full}}^T(\Delta t) = \frac{1}{K+1} \sum_{k=0}^K I(t_k).$$

For $I(t_k) = 1$, estimate efficiency via Monte–Carlo simulation:

$$\hat{\eta}_N = \frac{\sum_{k:I(t_k)=1} \left(\frac{1}{N} \sum_{j=1}^N \delta_b^{(j)}(t_k) \right)}{\sum_{k=0}^K I(t_k)}.$$

Delivery estimator:

$$\widehat{\text{DR}} = S_{\text{full}}^T(\Delta t) \cdot \hat{\eta}_N.$$

As shown in Lemma 5, the estimator variance decreases as $1/(MN)$, enabling predictable confidence scaling across SAOL sweeps.

7.1 Oracle Implementation Requirements

Definition 1 establishes fidelity as a property of the contact plan. In practice, a faithful oracle implementation must satisfy the following minimum conditions:

1. **Perturbed dynamics.** Orbital propagation must include at minimum dominant zonal harmonics (e.g., J_2) for the relevant central body. Two-body Keplerian propagation is insufficient for contact window accuracy at mission-relevant timescales.
2. **Geometric visibility from ephemeris.** Contact windows must be computed from instantaneous node positions under the propagated dynamics, not from static or pre-assumed alignment geometry.
3. **Capacity-aware contacts.** Each contact must carry a capacity derived from a link budget evaluated at the true range and geometry, not a binary visible/not-visible flag.
4. **Sufficient temporal resolution.** The discretization Δt must resolve contact boundary transitions. For low-orbit relay constellations, this typically requires $\Delta t \leq 60$ s.

Failure to satisfy any of these conditions may produce a contact plan in which S_{full}^T diverges significantly from the value under faithful dynamics, as demonstrated by the contact plan fidelity results referenced in Section 6 (Corollary 8).

8 Algorithms

8.1 Structural Feasibility Oracle

Algorithm 1: Structural Feasibility Oracle

Input: Contact plan $\mathcal{C}(\theta)$, injection time t , source s , destination d

- 1 **Precondition:** $\mathcal{C}(\theta)$ satisfies the oracle implementation requirements of Section 7.1;
 - 2 Build time-expanded graph G^T over $[t, T]$: nodes (v, τ) , edges for wait and transmit [?, ?];
 - 3 Determine reachability from (s, t) to any (d, t') with $t' \in [t, T]$;
 - 4 **if** *reachable* (d, t') *exists* **then**
 - 5 | **return** 1
 - 6 **else**
 - 7 | **return** 0
-

8.2 SAOL Architecture Optimization

9 Mars Specialization

Mars connectivity is sparse by celestial constraint. Orbiter periods, terrain shadowing, and DSN windows shape feasibility; this structure is well-documented in Mars relay network studies [?].

Algorithm 2: SAOL Architecture Optimization

Input: Search space Θ , constraints $C_i(\theta)$, oracle \mathcal{S} , efficiency estimator $\hat{\eta}_N$, budget B

- 1 Initialize dataset $\mathcal{D} \leftarrow \emptyset$;
 - 2 **for** $k = 1$ **to** B **do**
 - 3 Select candidate θ_k (e.g., via CMA-ES or Bayesian optimization);
 - 4 Compute $S_{\text{full}}^T(\theta_k) = \mathcal{S}(C(\theta_k))$;
 - 5 Estimate $\hat{\eta}_N(\theta_k)$ via DTN simulation;
 - 6 Compute $\widehat{\text{DR}}(\theta_k)$ and append $(\theta_k, \widehat{\text{DR}}(\theta_k))$ to \mathcal{D} ;
 - 7 **return** $\arg \max_{(\theta, \widehat{\text{DR}}) \in \mathcal{D}} \widehat{\text{DR}}$
-

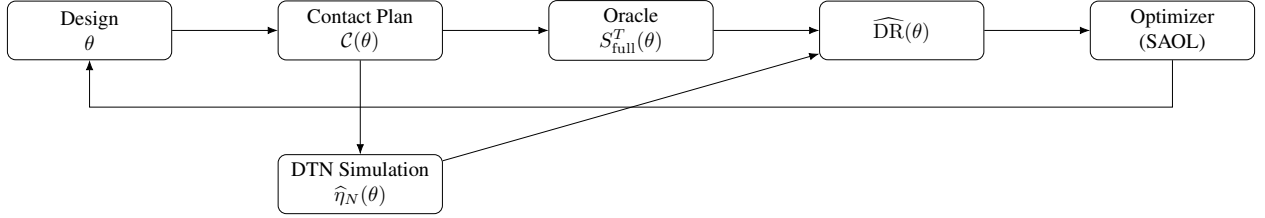


Figure 4: High-level structure of the Sparse-Architecture Optimization Layer (SAOL). Each candidate design θ induces a contact plan, which is evaluated by the structural oracle and the DTN simulator; the resulting delivery estimate guides the optimizer to propose new designs.

9.1 Feasibility on Mars

$S_{\text{full}}^T(\theta)$ is structured by:

- periodic surface–orbiter visibility;
- orbiter–orbiter phasing and altitude;
- Earth visibility bottlenecks via DSN [?].

9.2 Efficiency on Mars

Efficiency is dominated by:

- long custody holds while awaiting Earth visibility;
- limited buffer capacity and sustained accumulation on orbiters;
- fragmentation across short passes linking surface and orbiters.

9.3 Example Mars Architectures

Figure 5: Mars structural feasibility versus DSN contact allocation. Blue line: mean \bar{S}_T over 5 Martian sols; shaded band: per-sol min/max. The S_T plateau above 4 h/sol arises because 6 polar orbiters already provide high surface-to-orbiter coverage; the binding constraint is the ± 39.6 min/sol DSN phase drift, which produces alternating high ($S_T \approx 0.990$) and low ($S_T \approx 0.943$) sols regardless of DSN allocation.

Table 1: Mars architecture results from percolation-engine simulation (6 orbiters at 400 km, $i = 93^\circ$, uniform 60° RAAN; Dijkstra oracle at $\Delta t = 300$ s over 5 Martian sols). S_T^{full} is the whole-network temporal reachability; $\eta = \text{DR}/S_T^{\text{full}}$ is the conditional routing efficiency. For Mars-C the feasibility column reports S_T^τ (TTL-constrained reachability at TTL = 1 h) rather than S_T^{full} .

Architecture	S_T^{full}	η	DR	Notes
Mars-A (6 orb, 1 stn 70°N , NORMAL)	0.990	0.993	0.983	Geometry-limited
Mars-B (6 orb, dual stn $70^\circ\text{N}+20^\circ\text{N}$)	1.000	0.997	0.997	Station diversity
Mars-C (6 orb, 1 stn, EMRG TTL = 1 h)	0.280^\dagger	0.943	0.264	TTL-constrained

$^\dagger S_T^\tau$ at TTL = 1 h; $S_T^{\text{full}} = 0.990$ for this architecture.

Figure 6: Emergency delivery ratio versus bundle TTL for single-station (70°N , blue solid) and dual-station ($70^\circ\text{N} + 20^\circ\text{N}$, red dashed) Mars architectures. The law $\text{DR} \approx S_T^\tau$ holds throughout: adding a second station at 20°N raises S_T^τ at all TTL values, with the largest absolute gain at TTL = 1 h ($\Delta\text{DR} = +0.068$). For TTL ≥ 5 h, DR saturates at the full-contact-plan feasibility floor.

9.4 Ergodic Structure and Sol Phase Drift

The Mars relay network is a canonical example of a stationary ergodic contact process: regular orbital mechanics produce a contact-indicator time series whose time-average equals the ensemble average (the condition established by Proposition 9). However, the interaction of the Martian sol ($T_{\text{sol}} = 88,775$ s) with the Earth sidereal day ($T_\oplus = 86,400$ s) introduces a slow DSN phase drift of $+2,375$ s per sol, producing a systematic alternation between high- S_T and low- S_T sols.

Figure 7: Per-sol structural feasibility over five Martian sols at DSN allocation = 12 h/sol. Odd sols (phase drift places the DSN Earth-visibility gap in the middle of the primary surface-pass cluster) achieve $S_T \approx 0.944$; even sols achieve $S_T \approx 0.989$. The alternation period is determined by the beat frequency between the Martian sol and the Earth sidereal day. Dashed line: mean $\bar{S}_T = 0.9696$ over the 5-sol horizon.

9.5 Heliocentric Mars: Four-Tier Relay Architecture

The body-centric Mars results above describe the feasibility-dominated regime ($\eta > 0.97$) that prevails when the Earth–Mars link is abstracted away. In the heliocentric regime, where the full Earth–Mars distance enters the contact process, Mars transitions to an efficiency-dominated system.

A four-tier architecture was evaluated across a full 780-day synodic cycle (39 epochs, 5 seeds per epoch, 156 total configurations) [?, ?]:

Three findings are architecturally significant:

Braess paradox. T1 (3 satellites) delivers higher DR than T2 (9 satellites) in 35 of 39 epochs (90%). The pattern is consistent with a Braess-type onset: T2 has *better* reachability ($\bar{S}_T = 0.949$ vs. 0.869) but *worse* routing efficiency ($\bar{\eta} = 0.400$ vs. 0.527), suggesting that adding relay nodes can decrease DR under bounded-rationality routing.

Conjunction rescue. At the deepest solar conjunction (SEP = 1.4° , $d = 2.51$ AU), T1 and T2 experience complete blackout ($S_T = 0$), while T3 and T4 maintain $S_T = 0.993$ via L4 relay geometry. Lagrange-point relays eliminate conjunction-induced blackouts entirely in the tested simulation.

Figure 8: Architecture scaling results for the Mars relay network. *Left*: mean S_T^{full} (line) and per-sol min/max band (shaded) as a function of orbiter count. Performance saturates above 6 orbiters; additional spacecraft reduce the depth of low- S_T sols without raising the mean. *Right*: combined feasibility $S_T(A \cup B)$ as a function of Station B latitude. A second station near 20°N maximizes the gain ($\Delta S_T = +0.023$) by observing the same orbiters during the DSN gap that degrades Station A at 70°N.

Table 2: Heliocentric Mars architecture: synodic-cycle averages over 39 epochs ($d \in [0.57, 2.53]$ AU). S_T is temporal reachability; η is conditional routing efficiency; $\text{DR} = S_T \cdot \eta$. Blackout: epochs with $S_T = 0$.

Tier	Description	\bar{S}_T	$\bar{\eta}$	$\overline{\text{DR}}$	Blackouts
T1	3 low Mars polar, DSN 12 h	0.869	0.527	0.475	2/39
T2	6 low + 3 high Mars polar	0.949	0.400	0.400	2/39
T3	T2 + L4 Earth & Mars relays	1.000	0.498	0.497	0/39
T4	T3 + L5 Earth relay, DSN 24 h	1.000	0.591	0.591	0/39

The η ceiling. Even at Tier 4 (full architecture), the synodic-mean η is 0.591: reachability is solved ($S_T = 1.000$) but 40.8% of reachable bundles are lost to routing inefficiency. The distance–efficiency relation follows $\ln(\eta) \approx a - 0.8 d_{\text{AU}}$ across all tiers ($R^2 \in [0.74, 0.95]$), consistent with η as the dominant bottleneck for interplanetary DTN across the four tested tiers.

10 Cislunar Specialization

Cislunar connectivity is abundant but rapidly changing. NRHO gateways, short contact durations, and three-body dynamics shape feasibility and timing [?, ?].

10.1 Feasibility in Cislunar Space

Feasibility is shaped by:

- fast orbital transitions and relative motion;
- gateway orbits (e.g., NRHO) as choke points [?];
- short surface-analog windows and dynamic Earth schedules [?].

10.2 Efficiency in Cislunar Space

Efficiency is dominated by:

- precise custody timing at gateways;
- routing prediction sensitivity under rapid geometry;
- fine-grained timers to match short contact windows.

10.3 Empirical Validation of the Sparse Law

10.4 Example Cislunar Architectures

Remark. *These results measure routing efficiency for a south-polar surface station. The EM-L2 halo relay’s architectural value — extending far-side coverage from 68.5% to 96–98% — is not captured by this metric. Evaluating the halo by south-polar delivery ratio conflates coverage diversity with routing efficiency.*

Figure 9: Empirical validation of the Sparse Law: emergency delivery ratio DR versus full-network temporal reachability S_{full}^T for three cislunar architecture families (8-polar + relay, 8-polar only, 1-polar + relay). Each family traces a distinct linear band with constant slope η ; the factorization $\text{DR} = S_{\text{full}}^T \cdot \eta$ holds across more than 50 operating points spanning a $9.5\times$ range of contact density. The three slopes ($\eta \in \{0.68, 0.84, 0.71\}$) correspond to different routing geometries, consistent with η being a protocol-and-geometry constant rather than a universal, across the tested architectures.

Figure 10: Efficiency constant $\eta = \text{DR}/S_{\text{full}}^T$ versus full-network reachability S_{full}^T for the relay degradation axis (8-polar + relay, combined link deletion). The horizontal band confirms that η is stable throughout the supercritical regime ($S_{\text{full}}^T \gtrsim 0.4$). The sharp drop at $S_{\text{full}}^T \approx 0.37$ marks the percolation threshold, where the factorization breaks down as the network enters the sub-critical regime.

11 Architecture Implications

Across all deep-space environments, the Sparse Law imposes clear architectural guidance:

- Structural feasibility defines the outer limit of performance: $\text{DR} \leq S_{\text{full}}^T$.
- Efficiency identifies protocol bottlenecks within that limit.
- Custody-chain design is as fundamental as link design.
- Many architectures share the same feasibility; efficiency differentiates them.
- SAOL provides a principled method for discovering optimal trade-offs.

By Lemma 6, architectural changes that add visibility or capacity can never reduce S_{full}^T , though they may increase or decrease the efficiency η depending on protocol interactions.

The body-centric Mars relay network exemplifies a feasibility-dominated regime, in which geometry and DSN windows are the primary performance bottlenecks. However, the heliocentric Mars architecture (Section 9.5) transitions to an efficiency-dominated regime: Lagrange-point relays solve reachability ($S_T \rightarrow 1.000$) but η remains the ceiling, following $\ln(\eta) \approx a - 0.8 d_{\text{AU}}$. The Cislunar communications and navigation architecture exemplifies a different efficiency-dominated regime, in which high feasibility is available but protocol timing and custody precision limit performance.

Observation 16 (Efficiency–Feasibility Anticorrelation). *Under bounded-rationality routing (greedy next-hop selection with finite lookahead), the conditional efficiency η is empirically non-increasing in S_{full}^T for architectures within a common family.*

Cislunar constellation thinning sweep (*satellite removal, DSN always-on; higher- n configurations include ELFO relay, $n = 1$ is polar-only*):

Config	Relay	S_{full}^T	η	DR	C (h/day)
8 polars + ELFO	yes	1.000	0.844	0.844	94.7
4 polars + ELFO	yes	1.000	0.906	0.906	47.3
1 polar only	no	0.910	0.996	0.906	11.5

As connectivity decreases ($C : 94.7 \rightarrow 11.5$), feasibility drops ($S_{\text{full}}^T : 1.0 \rightarrow 0.91$) but efficiency rises ($\eta : 0.84 \rightarrow 1.0$). The product DR is approximately stable, varying less than either factor individually.

Figure 11: Relay contribution to structural feasibility: full-network reachability S_{full}^T versus direct polar-to-Earth reachability S_T^{dir} across sparse constellation operating points. Points above the diagonal indicate that relay routing (via ELFO and halo gateway) lifts S_{full}^T beyond the direct-path ceiling. The relay acts as a pure S_{full}^T -enhancement mechanism; at $S_T^{\text{dir}} = 0.70$, the relay sustains $S_{\text{full}}^T = 1.000$ — a 43% absolute gain.

Table 3: Cislunar architecture results from TIN v0.8.1 simulation [?] (Shackleton station at -89.5° ; EMERGENCY priority; 288 injection times over 24 h at $\Delta t = 300$ s). $S_T^{\text{full}} = 1.000$ for all multi-layer architectures; performance differentiates entirely through η .

Architecture	S_{full}^T	η	DR	Notes
CL-A (8P+ELFO+Halo, baseline)	1.000	0.684	0.684	Efficiency-dominated
CL-B (8P+ELFO, no halo)	1.000	0.844	0.844	Reduced routing confusion
CL-C (1 polar only, sparse)	0.910	0.996	0.906	Feasibility-limited, $\eta \approx 1$

Remark. *The anticorrelation is a property of bounded-rationality routing, not of the contact geometry. A perfect oracle-guided router would maintain $\eta = 1$ at all connectivity levels, and DR would track S_{full}^T exactly. The gap $S_{\text{full}}^T - \text{DR} = S_{\text{full}}^T(1 - \eta)$ measures the routing algorithm’s suboptimality, and this gap grows with network complexity. Abundant connectivity introduces dead-end routes, queue contention, and next-hop ambiguity that greedy heuristics cannot resolve.*

Limitations

The Sparse Law and SAOL framework presented here deliberately abstract away several important phenomena in order to isolate the structural effects of geometry and protocol behavior. In particular, the contact plan is treated as deterministic, without modeling stochastic link failures, weather, or ground-station outages; traffic is simplified to a representative bundle class; and protocol dynamics are captured through Monte–Carlo simulation rather than full analytical models. As a result, the theory is best viewed as a structural envelope for performance: it specifies what can and cannot be achieved under the assumed contact geometry and protocol semantics, but it does not replace detailed mission-specific modeling of interference, hardware constraints, or operational contingencies. The primary efficiency results use a greedy routing implementation; CGR comparison (Observation 15) confirms that the efficiency gap is recoverable but introduces fragility to schedule uncertainty. Comparison with PROPHET and DTLSR on identical contact plans is deferred to future work. All results are simulation-derived; validation against flight telemetry (e.g., LRO contact logs or LCRD operational data) would strengthen the empirical claims and is planned as part of the lunar pathfinder campaign.

11.1 Three-Factor Sparse Law

The validation results from eight configurations spanning body-centric (Mars, Moon) and heliocentric ($d \in [0.82, 2.51]$ AU) architectures establish a refined decomposition of the delivery ratio. The oracle-path success predictor is empirically indistinguishable from the Lyapunov form, $\eta_{\text{OPSP}} \approx e^{\mathbb{E}[H]\lambda}$, with relative error below 0.4% across all tested configurations, where $\lambda = \mathbb{E}[\log p_h] \leq 0$ is the Lyapunov exponent of the per-hop success sequence. The residual gap to simulation is captured by a multiplicative distortion factor $\Phi = \eta_{\text{sim}}/e^{\mathbb{E}[H]\lambda}$, yielding the *three-factor sparse law*:

$$\text{DR} = S_T \cdot e^{\mathbb{E}[H]\lambda} \cdot \Phi. \quad (1)$$

Here S_T captures feasible support, $e^{\mathbb{E}[H]\lambda}$ captures self-averaged chain attenuation on the oracle ensemble, and Φ encodes routing-policy and topological distortion. The observed range $\Phi \in [0.70, 2.30]$ on the

Figure 12: Contact-density bound test on a pure-polar cislunar architecture (no ELFO, no halo relay). Blue circles: satellite thinning (ergodic axis — whole spacecraft removed, preserving orbital regularity and stationarity of the contact process). Red triangles with error bars: duty-cycle deletion (non-ergodic axis — random contact windows suppressed, breaking stationarity). Dashed black line: ergodic bound $f(C)$ connecting thinning operating points. At matched $C \approx 11$ h/day the two axes diverge by $\Delta\text{DR} = 0.44$, driven seven-to-one by η rather than S_{full}^T . See Propositions 9 and 10 for the formal statements.

Figure 13: Percolation transition in the cislunar polar segment: emergency delivery ratio versus polar-to-Earth contact fraction C/T . Below the critical contact fraction $C/T \approx 0.05$ ($S_{\text{full}}^T \approx 0.37$), the network enters the sub-critical regime and the Sparse Law’s constant- η factorization breaks down. Above the threshold, the law holds with $\text{CV} < 0.5\%$. The transition is sharp, occurring over a deletion range of less than 5 percentage points, consistent with a first-order percolation transition on the temporal contact graph.

8-configuration validation set (extended to $[0.3, 8.6]$ across 17,280 configurations spanning Mercury to Titan [?]) spans both dead-end penalty ($\Phi < 1$) and diversity gain ($\Phi > 1$), confirming that Φ is a first-order observable rather than a perturbative correction. The Braess paradox lives entirely in Φ : the oracle-chain exponent is Braess-invariant. The Wald correction η_{Wald} improves on η_{Lyap} only when hop-success variance $\sigma^2 < 0.5$. See the companion math primer for the full formal development and validation table.

11.2 Cross-Domain Validation on Real DTN Traces

The three-factor factorization was validated on real human-mobility traces from the Cambridge/Haggle dataset [?], spanning four Bluetooth proximity experiments: Intel Research 2004 ($n = 9$ devices, 1,794 contacts, 12,000 configurations), Cambridge 2005 ($n = 12$, 5,042 contacts, 12,000 configurations), Infocom 2005 ($n = 41$, 21,438 contacts, 12,000 configurations), and Reality Mining 2004 ($n = 98$, 148,444 contacts, 93.2 h duration, 12,000 configurations). For each trace, the validation pipeline computes oracle feasibility (S_{full}^T), Lyapunov chain attenuation ($e^{\mathbb{E}[H]\lambda}$), and greedy routing efficiency (Φ) across 12,000 configurations per trace (6 link-reliability values \times 4 TTL fractions \times 50 source–destination pairs \times 10 seeds).

The identity $\text{DR} = S_{\text{full}}^T \cdot \eta$ holds at machine epsilon ($< 2.10 \times 10^{-16}$) across all active configurations, as expected for a definitional factorization.

The distortion factor on real traces is substantially larger than on orbital constellations (median Φ of 5.8 to 74 versus 0.3–8.6), reflecting the dense contact structure of social-proximity networks. The morphology order parameter γ [?] takes values in $[+0.74, +0.98]$ on all four social traces, versus $[-1.71, -0.06]$ on all eight orbital targets,¹ consistent with social-contact temporal graphs belonging to a distinct structural class (cluster-dominated) where path redundancy screens chain attenuation, within the four CRAWDAD traces examined.

Table 4 summarizes the morphology classification across all four traces. The γ values do not increase monotonically with node count: Exp3 ($n = 41$, per-pair contact density $\rho_{\text{pair}} \approx 0.10$) achieves higher γ than Exp6 ($n = 98$, $\rho_{\text{pair}} \approx 0.03$), consistent with saturation of γ toward 1 being controlled by per-pair contact density rather than network size alone. The multiplicative decomposition $\Phi = \Phi_m \times \Phi_r$ fails catastrophically within the four CRAWDAD traces examined, with errors exceeding 400,000% on Exp6.

Combined with the 106,000 orbital configurations across eight solar-system targets, the total validated

¹The CRAWDAD γ values use the normalized form $\tilde{\gamma} = \text{raw slope}/(-\lambda)$, which is p_{eff} -invariant; the orbital values are raw slopes evaluated at body-specific p_{eff} . The sign separation is invariant under any positive rescaling; the classification conclusion is unaffected.

Table 4: Morphology classification across four CRAWDAD social-contact traces. γ_{normal} and γ_{myopic} are the morphology order parameters under normal and no-retry routing; topo% is the fraction of γ attributable to topology rather than retry; $\Phi_m \times \Phi_r$ denotes the multiplicative decomposition error (REJECT = catastrophic failure). All γ values use the normalized convention $\tilde{\gamma} = \text{raw_slope}/(-\lambda)$, a dimensionless p -invariant form; see [?] for the raw-slope convention.

Exp	n	Contacts	γ_{normal}	γ_{myopic}	topo%	Φ_r
Exp1 (Intel 2004)	9	1,794	[+0.74, +0.84], mean +0.81	[−0.07, +0.98]	52%	REJECT
Exp2 (Cambridge 2005)	12	5,042	[+0.80, +0.94], mean +0.85	[+0.22, +1.07]	53%	REJECT
Exp3 (Infocom 2005)	41	21,438	[+0.96, +0.98], mean +0.97	[+0.69, +0.98], mean +0.87	84%	REJECT
Exp6 (Reality Mining)	98	148,444	[+0.89, +0.92], mean +0.90	[+0.34, +0.99], mean +0.78	84%	CATASTROPHIC

dataset exceeds 154,000 configurations spanning Mercury to Titan and four independent terrestrial mobility experiments.

12 Conclusion

The Sparse Law provides a unified, first-principles characterization of deep-space communication performance. By decomposing delivery into structural feasibility and conditional efficiency, it enables principled design and optimization across planets, moons, and cislunar space. SAOL operationalizes this framework, guiding future architectures for deep-space exploration and infrastructure with mathematical clarity and operational fidelity.

Empirical validation across Mars and cislunar architectures, extended to eight solar-system targets [?] and four real human-mobility traces from CRAWDAD [?], is consistent with the law’s predictions across more than 154,000 configurations: the efficiency constant η is stable under diverse orbital and relay configurations (CV < 0.5% in the supercritical regime), and the design bound $S_{\text{full}}^T \geq D_{\text{req}}/\eta$ reduces constellation sizing to a deterministic reachability computation. The Contact-Density Bound (Proposition 9) establishes that a scalar contact-time density determines structural feasibility under renewal assumptions, and non-ergodic degradation is proved by construction (Proposition 10), identifying stochastic scheduling disruptions as a qualitatively distinct failure mode. The supercritical efficiency bound (Proposition 12) provides an analytical prediction $\eta_0 = s/(s+1)$ from a four-state CTMC, validated to 0.3% on the cislunar baseline. The TTL factorization (Corollary 4) extends the law to finite-lifetime bundles, with η_r empirically stable across DSN configurations (CV = 0.2%). Cross-validation with oracle-class Contact Graph Routing [?] shows that the greedy efficiency gap is 89–100% recoverable within the tested architectures (Observation 15), but CGR introduces fragility to schedule uncertainty — a trade-off that bounds its practical applicability.

The sole open question — whether non-ergodic degradation is generic rather than merely constructive — is deferred to future work. The Mars relay extension, previously identified as future work, has been validated: a four-tier architecture across 156 configurations and a full synodic cycle confirms that Lagrange-point relays eliminate conjunction blackouts ($S_T = 0.993$ at SEP = 1.4°), exposes the Braess paradox in a realistic mission model (Section 9.5), and establishes η as the dominant interplanetary bottleneck [?, ?].