Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Quantifying Redundancy in Directed Acyclic Transmission Networks with Conserved Flow: A Hadith Case Study

Art Saiti

artsaiti@proton.me

Independent Researcher

Research Article

Keywords:

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11763730/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Quantifying Redundancy in Directed Acyclic Transmission Networks with Conserved Flow: A Hadith Case Study

Art Saiti[1*]

[1*]Independent Researcher, Tetovo, North Macedonia.

Corresponding author(s). E-mail(s): artsaiti@proton.me;

Abstract

When information passes through a network from sources to a destination, node centralities rank nodes by importance, but do not measure their actual contribution to what was transmitted or how far the contributions of several nodes overlap. We quantify both with a conserved flow on a directed acyclic network. Material enters at the sources, is divided among successors in proportion to recorded transmissions, and is absorbed at a single sink; the flow is the source–sink random-walk flow underlying Newman's betweenness, adapted to directed acyclic graphs, and is solved exactly in one forward pass. Removing a set of nodes deletes the material that passed through them, and the share of delivered flow lost is the set's severance.

We show that the severance of a single node is its flux discounted by downstream leakage, so single-node removal only reproduces a flow-based ranking, whereas the redundancy of a set, the sum of its members' severances minus their joint severance, is the delivered material that passed through more than one of them. Hadith, the reports of the Prophet Muhammad, are well suited to test the method: because of their religious importance, each report was transmitted with its isnad, the chain of narrators through which it reached its compiler, giving a record that is acyclic, ordered by generation and ends at a single known terminus. In Sahih al-Bukhari and Sahih Muslim, redundancy separates relays from aggregators: the material carried by both Shu'ba and his transmitter Ghundar is 4.3% of Muslim's delivered flow (95% bootstrap interval 3.7–5.1%) against 2.1% (1.8–2.6%) in Bukhari, and Muslim's three leading narrators overlap more than Bukhari's. Both collections narrow sharply at the same two generations, and the results hold under alternative modeling choices.

Keywords: knowledge transmission, network vulnerability, node removal, redundancy, flow centrality, absorbing Markov chains, hadith, isnad

1 Introduction

Ḥadīth are the reported sayings and actions of the Prophet Muḥammad. Each report has two parts: the *matn*, its content, and the *isnād*, the chain of narrators that recorded it. Initially, reports circulated orally and began to be written out in the eighth century (although they existed earlier in private notes), and were compiled in the ninth, with the Ṣaḥīḥs of al-Bukhārī and Muslim among the most prominent. Due to their religious significance, Muslim scholars developed an elaborate criticism of their transmission: they assessed the reliability of each narrator (*'ilm al-rijāl*), whether teacher and student could have met, and whether a report was corroborated by independent chains [1].

The isnād makes hadith unusually well suited to network analysis. Most historical networks must be reconstructed from indirect or fragmentary evidence, while here every report carries its own record of who received it from whom. The record is ordered in time and the biographical literature assigns narrators to generations (*ṭabaqāt*) by their distance from the Prophet, and since transmission goes from earlier to later generations, the network is directed and acyclic. Each collection ends at a single known terminus, its compiler, so every chain has the same final destination. The material is extensive: by traditional accounts some six hundred thousand reports were known, and this paper deals with the digitized part of it. The tradition also grades the reliability of each narrator, which could in principle enter the model. We leave it out deliberately: the reports analyzed here were included precisely because their narrators were judged reliable, so weighting by that judgment would be circular (§7.1).

These properties make qualitative questions posed by the tradition answerable in a quantitative manner, namely: How much of a collection passed through a given narrator? How much would be lost if he were removed? Do two prominent narrators carry independent bodies of material, or largely the same material twice? The last is the critics' own question of independent attestation, and a ranking of single narrators cannot answer it.

The same questions arise wherever material passes forward in time from recorded sources to a known destination: in academic genealogies and manuscript traditions, in the libraries a software application depends on, and, with some adaptation, in the spread of news (§6).

The main contribution of this paper is a formal model of the isnād as a transmission network that answers these questions, built so that it transfers to similar records generally.

- We make explicit the model of transmission that a measure of narrator importance presupposes, classifying it using Borgatti's typology of network flows [2] (§2).
- We treat a collection as a conserved flow from its earliest narrators to its compiler, solved exactly in one pass (§4). Removing a set of narrators destroys the material they carried; we show that for a single narrator this loss reduces to his share of the flow, and that the redundancy of a set is exactly the material that passed through more than one of its members (§4.2).
- We apply the model to Ṣaḥīḥ al-Bukhārī and Ṣaḥīḥ Muslim, with a bootstrap over reports to assess stability (§5). Redundancy separates narrators who relay one

teacher's material from those who gather many teachers', and both collections narrow sharply at the same two generations, a corpus-wide counterpart of the common link.

Section 6 discusses the implications for hadith studies and for transmission networks generally.

1.1 Previous Work

Şentürk [3] was the first to reconstruct the hadith transmission network as a whole, from the seventh to the sixteenth century, claiming in the process that it represents the most detailed existing social network in human history.

Saeed et al. [4] apply graph centrality measures to isnād networks, and their results confirm the traditional narrative — that Abū Hurayra, for instance, transmitted an exceptionally large number of reports. Alam and Schneider [5] apply comparable methods to Ṣaḥīḥ al-Bukhārī and reach similar conclusions: out-degree places Abū Hurayra and Anas ibn Mālik first, PageRank places Shu'ba first and al-Zuhrī second, and betweenness places Sufyān ibn 'Uyayna first.

These findings are limited by what the measures count. Degree counts a narrator's links, or if weighted, the transmissions along them, but does not measure how much of the corpus flowed through him to the collector. A narrator with many students may be inconsequential to the collection if each link carried by him represents a single tradition, while a narrator with few students may be vital if no other route reaches the collector outside of him. Because a centrality score is not conserved, removing a narrator and recomputing it yields a new ranking rather than an amount lost; as such, it is not possible to simply determine what amount of material is lost when a narrator is removed.

One reason for this difficulty lies in the kind of network the isnād corpus represents. Transmission runs forward in time, so the network is directed and acyclic: no report goes back to a previous generation. Most centrality measures exist for networks without this property, and some depend on its absence. Eigenvector centrality presupposes a recurrent structure in which influence can circulate, but the adjacency matrix of an acyclic network is nilpotent and has no non-zero eigenvalue; PageRank remains defined only because its teleportation term supplies the recurrence the network lacks. Acyclic, time-ordered networks have been studied mainly through bibliometric data, whose structure Price described early [6]. Main path analysis traces the dominant line of development through a citation network by counting the source-to-sink paths through each link [7, 8]; random-graph models and reductions exist specifically for directed acyclic networks [9, 10]; and work on temporal networks restricts influence to time-respecting paths, along which something can travel only forward [11, 12]. These approaches share the premise we adopt that in a time-ordered network the direction of time is a fundamental aspect of the structure.

The isnād is an unusually complete instance of this type. Its ordering is given explicitly by the generation labels of the biographical literature rather than inferred from dates with some exceptions pertaining to same-generation links (§3.2); the record states which narrator heard from which and every chain in a collection ends at a

single known terminus, the compiler. Together these make the information flow along it exactly solvable (§4). Citation networks, by contrast, have no common terminus, and papers written at nearly the same time can cite one another.

Schacht [13] noticed that the isnāds of a tradition frequently converge on one narrator before fanning out again, and Juynboll [14] built the observation into a systematic method for locating that *common link* or *madār al-isnād*, the pivot on which a chain turns. This is a topic of academic contention [15, 16] since it has some serious historical and religious implications, and the paper does not make any claims supporting or rejecting its veracity, but being able to quantify the amount of material passing through such narrators provides a metric that directly corresponds to this topic.

Aldoboni et al. [17] represent chains as paths in a narrator graph and train a model to detect fabricated ones, classifying corruptions drawn from classical terminology such as narrator substitution (*taṣḥīf*) and chain splicing (*idrāj*). Their formalization is a learning task on individual chains and does not model the transmission network itself or how a collection depends on its narrators, which is the subject of this paper.

Work on isnād networks nonetheless remains sparse, and formal models of them are still quite rare. A particular constraint is the state of digitization: the available sources are community-maintained, and no authoritative index or database resolves narrator identities across them (§7.1).

2 The isnād as a flow process

Borgatti [2] argues that a centrality measure is the expected value of a node outcome: how often traffic passes a node, or how soon it arrives under the flow process its formula assumes, and he classifies such processes along two axes. The *trajectory* is the kind of route traffic may follow: a walk is any sequence of adjacent nodes, a trail repeats no link, a path repeats no node, and a geodesic is a shortest path. The *mechanism* is how traffic passes between nodes: by *transfer*, when it moves and is in one place at a given time, or by *duplication*, when it is copied and the sender keeps it, either to one neighbor at a time (serial) or to all at once (parallel).

Trajectory.

Transmission runs forward in time, so the network is acyclic once same-generation links are set aside (§3.2). In an acyclic graph every walk is a path. Walks, trails and paths therefore coincide, and only the distinction between geodesics and other paths remains. Transmission in our case does not follow geodesics: the preference for elevated chains (*ʿuluww*) shaped which teachers a student sought, not a route chosen through some predetermined network. We therefore model isnād transmission to follow paths.

Mechanism.

A narrator who transmits a report still remembers it, so in reality the process the record describes is duplication, serial or parallel according to whether a report was given to one student or to a group. The model instead treats it as a direct transaction: a narrator's material divides in different ratios among his students, and each unit follows one route. This is contingent on the question we are asking. We ask how the

delivered collection is allocated across the routes that carried it, and how much of it depends on a given narrator. Only a conserved flow answers that; under duplication, copies multiply at every branching, and counts of copies through different narrators partition nothing. Of course, the model does not claim that a teacher forgot what he taught or transmitted, but assuming each delivered unit can be attributed to one recorded route greatly simplifies the modeling.

Destination.

Every chain in a collection ends at its compiler. Borgatti imposed sources and targets on all his simulated processes so that they could be compared with formulas that assume them; in an isnād collection that structure is given by the record itself.

The model therefore can be said to sit in Borgatti's [paths, transfer] categorization, his mooch process, for which he found no standard measure. Acyclicity merges it with the [walks, transfer] cell of Newman's random-walk betweenness [18], which is the flow solved in §4. On an acyclic network that flow gives in closed form the passage frequencies Borgatti estimated by simulation. The duplication process the record describes is kept as a benchmark, the route count, which differs from the flux only in dropping the normalization (§5).

3 Data and model

3.1 From chains to a weighted digraph

Data.

The chains are those of Ṣaḥīḥ al-Bukhārī and Ṣaḥīḥ Muslim published by sunnah.com [19] in digital format, which number 7,563 and 3,033 reports. The data contains 7,277 and 7,368 entries, each with its transmission chain, together with the narrators' identities and their ṭabaqa. Every narrator carries a unique identifier in this index, shared by the two collections. Reports that the compilers repeat under different chains are kept, since each repetition is a unique transmission, and the object of study is the network of narrators rather than the set of distinct hadith or their content. This is also why Muslim has fewer reports than Bukhārī but a similar number of entries. We take the digitized record as given and do not assess its accuracy.

The 129 and 83 entries whose chain names fewer than two narrators are discarded, leaving 7,148 and 7,285 chains.

Orientation.

Sources record chains collector-first. Let $g(v)$ denote the *ṭabaqa* rank of v. For sake of simplicity, every chain is reversed and written early-to-late.

Let V be the narrators appearing in some chain, $n = |V|$. We define the multiplicity a_{uv} as the number of times u is recorded transmitting directly to v, and the terminal count τ_v as the number of chains ending at v, material the collector received from v without a further intermediary.

3.2 Vertical and lateral edges

An edge is *vertical* if $g(u) \neq g(v)$, oriented from the lower rank, and *lateral* if $g(u) = g(v)$. Only vertical edges are retained.

A ṭabaqa is defined in part by whom a narrator is recorded as having met, so adjacent ranks overlap in absolute chronology and a same-rank pair carries no ordering evidence. Death dates would in principle disambiguate, but are missing for a large share of V. As a form of compromise we therefore treat generation as the sole ordering evidence and exclude lateral edges, reporting their mass share λ as a diagnostic. This is the more conservative choice: an excluded edge shortens the modeled chain by a pre-specified amount, whereas a misoriented one reverses a transmission and corrupts it.

Vertical edges with $g(v) - g(u) > \gamma_{\max} = 4$ are discarded as implausible; the mass so removed is likewise reported.

3.3 The absorbing collector

Adjoin Ω at index n, $N = n + 1$, and set $\tilde{A}_{vu} = A_{vu}$, $\tilde{A}_{\Omega u} = \tau_u$, $\tilde{A}_{u\Omega} = 0$. The empty column at Ω makes it absorbing rather than recurrent.

With $c_u = \sum_v \tilde{A}_{vu}$, call u *live* if $c_u > 0$. The transmission operator is the column-normalization

$$P_{vu} = \tilde{A}_{vu}/c_u \;\; (c_u > 0), \qquad P_{vu} = 0 \;\; (c_u = 0), \qquad \text{so} \qquad \sum_v P_{vu} = \mathbf{1}\{u \text{ live}\}. \quad (1)$$

Nodes that are neither live nor Ω are *orphans*: narrators whose onward edges were removed by the filters of §3.2. Write $\mathcal{D} = \{u : c_u = 0\}$ for the dead columns, orphans together with Ω.

Remark 1 The denominator includes τ_u, so absorption competes with the student edges for the same mass. Omitting the Ω row would force a narrator's entire output down whatever onward edges survive, inflating minor students in proportion to τ_u.

3.4 Source

Roots are narrators with output but no recorded teacher in the graph, $\mathcal{R} = \{u : (A^\top \mathbf{1})_u > 0,\ (A\mathbf{1})_u = 0\}$, and $S = \sum_{u \in \mathcal{R}} s_u e_u$ with $S_\Omega = 0$.

We use a uniform seeding approach, $s_u \equiv 1$. The alternative, $s_u = (A^\top \mathbf{1})_u$, reintroduces the volume each root is recorded as narrating, and the flux then largely reproduces that volume: the Companions with the most attributed traditions dominate by construction, which is what existing centrality studies already report [5]. Uniform seeding instead holds the sources equal and lets the network decide, so that ρ measures what the structure of transmission did with the material rather than how much entered at each root. Consequently ρ is a share of the modeled flow, not of the collection's traditions.

4 The flux equation

Conservation at each node — what v holds equals what its teachers sent plus what was injected — is

$$\boxed{(I - P)\,\rho = S} \tag{2}$$

Proposition 1 (Well-posedness) Suppose g is defined at both endpoints of every retained edge. Order $\tilde{V} = V \cup \{\Omega\}$ by g, ties broken arbitrarily, with Ω last. Then P is strictly lower triangular, thus nilpotent with $P^m = 0$ for m the greatest number of *nodes* on any path. Then $I - P$ is unit triangular, $\det(I - P) = 1$, and

$$\rho = \mathcal{N}S = \sum_{k \geq 0} P^k S, \qquad \mathcal{N} = (I - P)^{-1}, \tag{3}$$

is a finite sum, computable by one pass of forward substitution.

Proof Every retained edge satisfies $g(\text{teacher}) < g(\text{student})$, since same-rank edges are excluded (§3.2); edges into Ω point to the last index; column Ω vanishes. So $P_{vu} \neq 0$ implies $u \prec v$, and a strictly triangular matrix is nilpotent. □

$\mathcal{N}$ is a discrete Green's function: $\mathcal{N}_{vu}$ is the response at v to a unit injected at u, and (3) expands it as a sum over paths weighted by the product of splitting fractions. The expansion is the Neumann series of $(I - P)^{-1}$, which in general converges only under a condition on the spectral radius of P and here simply terminates, because P is nilpotent. In particular, the Ω row of $\mathcal{N}$ gives, for each narrator, the fraction of his material that reaches the collector.

Relation to Newman's random-walk betweenness.

The flux is the directed form of the quantity behind Newman's random-walk betweenness [18]. Newman lets a walker leave a source s and move to a uniformly chosen neighbor, with transition matrix $M = AD^{-1}$, until it is absorbed at a target t; the expected number of visits to each node is then $(I - M_t)^{-1}e_s$, with row and column t removed, which is (3) with P in place of M_t. Summed over the pairs formed by the roots and the collector, his node currents give the flux,

$$\rho_v \;=\; \sum_{s \in \mathcal{R}} s_s \, I_v^{(s\Omega)}.$$

Three differences each make our case much simpler. Newman's network is undirected, so he counts net crossings, read as current in a network of unit resistors [18, 20]; here the edges are directed and weighted by recorded multiplicity, so the walk is used directly. The network is acyclic, so no unit returns, net and gross crossings coincide, and the system is solved by forward substitution (Proposition 1) rather than by inverting a Laplacian. Since every source shares one target, the contributions of all roots add up in a single solve instead of being averaged over all pairs. Where Newman's cancellation gives a node in a cul-de-sac zero betweenness, material that enters an orphan

here is stranded, counted as loss in (5), and excluded from severance by the delivery factor $\mathcal{N}_{\Omega v}$ in (8).

More directly, read as a walk, a unit's chance of taking a channel is the share of recorded material that channel carries; the walk is an accounting device imposed by the model, not a claim that narrators chose their students at random.

4.1 Conservation

Define the edge flux $f_{vu} = P_{vu}\rho_u$. Summing over destinations gives the outflow $\rho_u \mathbf{1}\{u \text{ live}\}$ by (1); summing over sources gives the inflow $(P\rho)_v = \rho_v - S_v$. Hence

$$\sum_u f_{vu} + S_v \;=\; \rho_v \;=\; \sum_w f_{wv} \qquad (v \text{ live}), \tag{4}$$

Kirchhoff's current law. Left-multiplying (2) by $\mathbf{1}^\top$ and using $\mathbf{1}^\top P = \mathbf{1}^\top - \mathbb{1}_{\mathcal{D}}^\top$, the total $\mathbf{1}^\top\rho$ cancels and

$$\rho_\Omega + \sum_{u \text{ orphan}} \rho_u \;=\; \|S\|_1. \tag{5}$$

Interior flows telescope and only dead columns survive. Since every chain ends at the collector (§3), orphans arise only from the filters of §3.2: a narrator whose onward links were all lateral or over the generation cap, and who never ends a chain, keeps what reaches him. Material entering an orphan is therefore lost, and $\|S\|_1 - \rho_\Omega$ measures the loss directly. Because it is measured rather than absorbed silently, (5) doubles as a validity check on the preprocessing: a large orphan share indicates that the record, not the transmission, is shaping ρ. Both (4) and (5) are asserted numerically as invariants.

The same operator also describes how far material travels: an adjoint solve gives the number of narrators a unit crosses on its way to the collector, and the terms $P^k S$ of (3) split the flux by number of hops. Neither enters the results beyond the diagnostic $\bar{\ell} = \|\rho\|_1/\|S\|_1$ of Table 1, the mean number of nodes a unit crosses; both are set out in Appendix A.

In hadith terminology, depth resembles the elevation (*'uluww*) of a chain: elevated (*'ālī*) chains are those passing through the fewest narrators and judged to be less prone to corruption. Care should be taken to not confuse the two. Elevation belongs to a single chain while depth here is an average over all the material a narrator passed on, mixed across reports. Measuring elevation in this context is of no use since it applies to single chains only and provides no information about the general network structure, while depth can be analyzed from the perspective of the average number of hops a hadith needs to reach its destination. This somewhat implies the reading that the lower the average hops per narrator the more trustworthy they are, something we do not claim; depth is therefore reported as a diagnostic and can be revisited in future work.

4.2 Ablation

Conservation makes counterfactual removal well posed. For $K \subseteq V$ let $D_K = I - \sum_{u\in K} e_u e_u^\top$ zero the coordinates in K. Removing K deletes the narrators together

with every channel through them,

$$P_K = D_K P D_K, \qquad S_K = D_K S, \tag{6}$$

so that K neither receives, transmits, nor is seeded. Mass is destroyed rather than rerouted: a predecessor of a removed narrator has its column sum fall below unity, and material entering that channel is lost. The alternative would assert that a student would have received from another teacher what he is recorded as receiving from this one, which puts us in probabilistic territory outside of the paper's scope. P_K is strictly triangular whenever P is, so the ablated system $(I - P_K)\rho^{(K)} = S_K$ inherits the one-pass solve of Proposition 1 at one forward substitution per counterfactual, and

$$\sigma(K) \;=\; 1 - \frac{\rho_\Omega^{(K)}}{\rho_\Omega} \tag{7}$$

is the fraction of the delivered flow severed. Since removal destroys rather than reroutes, $\sigma(K)$ measures dependence on the recorded routes, and is an upper bound on the material genuinely at risk wherever corroborating parallels (*mutāba'āt*) existed but went unrecorded.

Two aggregates are reported. The *solo* effect $\sigma(\{u\})$ ranks narrators individually; the *joint* effect $\sigma(K)$ for a set K measures their combined indispensability.

The solo effect contains no information beyond the flux. Because the network is acyclic, no unit of material passes a narrator twice, and because removal does not reroute, the material that does not pass v is unaffected by his removal. Removing v therefore destroys exactly the delivered share of what passed through him:

$$\sigma(\{v\}) \;=\; \frac{\rho_v \mathcal{N}_{\Omega v}}{\rho_\Omega}, \tag{8}$$

where $\mathcal{N}_{\Omega v}$, the Ω row of the fundamental matrix, is the fraction of v's material that reaches the collector. Solo severance is flux discounted by leakage, and ranks narrators almost exactly as ρ does (§5).

The joint effect is where new information lies. Summing (8) over K counts each delivered unit once for every member of K it passes, whereas $\sigma(K)$ counts it once if it passes any. Their difference,

$$\sum_{u \in K} \sigma(\{u\}) - \sigma(K) \;=\; \frac{1}{\rho_\Omega} \sum_{\text{delivered units}} \big(j - 1\big)^+, \tag{9}$$

with j the number of members of K on a unit's route, is the *redundancy* among them: the delivered material that travels through more than one of the narrators, counted once for each additional member it passes. For a pair $u \prec v$ it is simply the delivered material that passes through both, $\rho_u \mathcal{N}_{vu} \mathcal{N}_{\Omega v} / \rho_\Omega$, so that a relay who transmits almost nothing but his teacher's material is almost wholly redundant with him. Set-level measures exist for other centralities [21]; conservation gives this one a direct

reading in terms of routes. Appendix B restates these identities in terms of a pipe network.

Relation to vitality.

Severance is a vitality index in the sense of Koschützki et al. [22]: a node or set is scored by the change in some network-wide quantity after its removal, here the delivered flow. Its closest relative is max-flow vitality, the reduction in the maximum flow between a fixed source and sink when an element is removed [23]. The two share the source–sink structure — with its roots joined to a single super-source, an isnād collection is an s–t network — and differ in what happens after removal. Max-flow vitality recomputes the optimal flow, so material that used a removed narrator is rerouted through whatever capacity remains. Severance keeps every surviving split fixed, so that material is destroyed: the two embody opposite assumptions about substitution, as much as capacity allows there and none here. A practical difference follows. Ausiello et al. are concerned with computing the vitality of every element efficiently, since each removal can change the optimal flow throughout the network. Without rerouting, the solo severances of all narrators follow at once from ρ and the Ω row of $\mathcal{N}$ (8), one forward and one backward substitution in all; only sets of narrators need a solve of their own.

4.3 Uncertainty

Every quantity above is computed from the traditions a compiler happened to include.

Using a non-parametric bootstrap [24], we can estimate the influence of the particular chosen chains on the results.

Each replicate draws, with replacement, as many traditions as the collection contains; chains are kept intact, so none is split or recombined. The graph, the filters of §3, the operator and the flux are then rebuilt exactly as for the observed data, and the statistics recomputed. Here $B = 1000$ replicates are used per collection and 95% percentile intervals are reported.

This approach can seemingly imply that other chains could have entered but did not. Rather the selection is based on the perceived authenticity of narrators and as such the results are not indicative of uncertainty as much as the stability of a corpus.

Two choices need stating.

- *Re-ranking.* In each replicate the top k narrators are chosen afresh by that replicate's flux, with the eligibility rule used for the observed data. Variation in who belongs to the top set is therefore part of the variation reported, and makes intervals at small k wide and asymmetric. A fixed pair of named narrators is tracked separately.
- *Comparing collections.* The two collections are resampled independently, and the interval of a difference is taken over the differences of replicates paired by index.

We do not compare against randomly rewired networks. Degree-preserving rewiring would create teacher–student links that the record does not contain, many of them chronologically or geographically impossible; more to the point, the claims below are

comparisons between two observed collections, not departures from some random baseline. Resampling whole traditions answers the question those claims raise: whether they would survive a different selection of the same kind of material.

5 Results

5.1 Corpus and diagnostics

We apply the model to Ṣaḥīḥ al-Bukhārī and Ṣaḥīḥ Muslim; Table 1 reports the graphs and the preprocessing steps of §3.2. Roots are predominantly Companions — 162 of 193 and 189 of 256 at ṭabaqa 1, the remainder spread thinly across ṭabaqāt 2–9. The overwhelming majority of these are narrators whose incoming edges the filters removed rather than narrators the record leaves without a teacher; in Bukhārī this holds for 29 of the 31. Under uniform seeding these inject at a Companion's rate; seeding only the Companions changes none of the conclusions reached (§5.5). Conservation (5) shows 93.6% and 92.8% of injected material delivered. The model's mean chain length $\bar{\ell}$, computed under out-strength seeding so that it is comparable with the recorded mean (Appendix A.1), falls short of it by 8.4% and 20.4%, the larger shortfall in the more heavily filtered collection. Figure 1 shows the resulting network, with node size proportional to flux carried by a single narrator.

Table 1 Corpus and preprocessing diagnostics. Stranded mass is orphan loss in (5); λ, the $\gamma_{\max}$ share and transmissions with a narrator of unknown ṭabaqa are removed before the operator is built and do not enter the conservation identity. Chain lengths are narrator counts, $\bar{\ell}$ computed under out-strength seeding for comparability.

	Bukhārī	Muslim
Chains used	7,148	7,285
Narrators	1,557	1,924
Vertical edges	4,472	7,005
Roots (ṭabaqa 1 / all)	162 / 193	189 / 256
Lateral mass share λ	5.10%	11.64%
Discarded, $g(v) - g(u) > \gamma_{\max}$	1.26%	1.61%
Missing ṭabaqa	0.09%	0.10%
Delivered to Ω	93.61%	92.77%
Stranded in orphans	6.39%	7.23%
Mean chain, recorded	5.33	6.37
Mean chain, model $\bar{\ell}$	4.88	5.07

5.2 Single narrators

Al-Zuhrī leads both collections (Tables 2 and 3), carrying 15.2% of the flow in Bukhārī and 10.4% in Muslim, with severance of 15.9% and 11.1%; Shu'ba ibn al-Ḥajjāj is

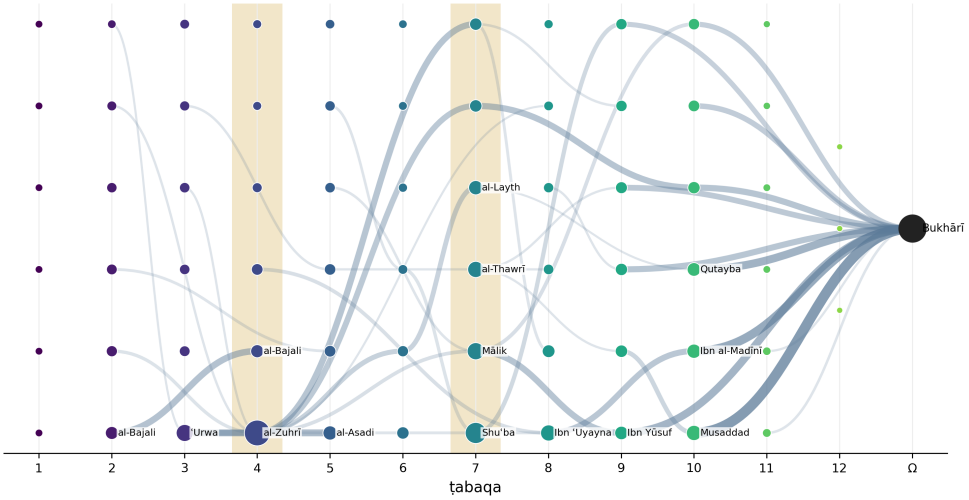

Fig. 1 The flow of transmission, with the size of the circles representing total flux carried.

second in both. Given his particular standing, it of use to know that tradition credits Al-Zuhrī with the first systematic compilation of hadith, undertaken at the request of the caliph Umar ibn Abd al-Azīz [1]. Five further narrators are common to both top twelve — Mālik, al-Thawrī, Ibn ʻUyayna, ʻUrwa ibn al-Zubayr and al-Layth ibn Saʻd — so seven of twelve are shared between the two corpora. Where the lists diverge they do so late: Muslim's includes four narrators at ṭabaqa 9, among them Ibn Wahb, who carries 7.7% of Muslim's flow and ranks third, with no counterpart in Bukhārī. The collections agree on the second-century pivots and part company at the compiler's own generation, which the construction predicts, since the final rank is each compiler's own *shuyūkh*.

Table 4 compares solo severance with the node rankings used in earlier work, over the sixty heaviest eligible narrators of Bukhārī. σ correlates with weighted degree at Spearman $r_s = 0.66$, with PageRank at 0.64 and with shortest-path betweenness at 0.46, while its correlation with flux is 0.998, as (8) requires: the residual reflects only differences in leakage between narrators. Single-narrator vulnerability thus agrees with the rankings on who matters and differs on how much. Agreement is weakest for shortest-path betweenness, whose geodesic assumption [2] favors elevated (*ʻālī*) chains for their brevity rather than for what they carry. Eigenvector centrality is undefined (§1.1); standard implementations, which iterate on $A + I$, return a vector piled onto the final generation.

The route count is a node-level version of the duplication benchmark of §2: the number of distinct routes through the narrator from any root to the compiler along recorded links, with no splitting and no conservation. Because it counts paths through the graph, it includes routes assembled from links attested in different chains, and it multiplies wherever the network branches. It agrees with σ less than any other measure except betweenness ($r_s = 0.48$, six of the top ten), so the choice of transfer over duplication visibly shapes the ranking: σ measures how much of the delivered

Table 2 Ablation, Ṣaḥīḥ al-Bukhārī. *Solo* is the fraction of delivered flow severed by removing the narrator alone; *cumulative* by removing him together with all those above him. Ranks are ṭabaqāt in the numbering of Ibn Ḥajar. *Flux* is $\rho_v / \|S\|_1$, a share of injected material, whereas solo and cumulative are shares of delivered flow, which is why solo can exceed flux.

k	Narrator	Ṭabaqa	Flux	Solo	Cumulative
1	al-Zuhrī	4	0.152	0.159	0.159
2	Shuʻba	7	0.099	0.106	0.264
3	Mālik	7	0.062	0.067	0.314
4	ʻUrwa ibn al-Zubayr	3	0.061	0.064	0.346
5	al-Thawrī	7	0.054	0.058	0.402
6	Ibn ʻUyayna	8	0.054	0.057	0.445
7	al-Layth ibn Saʻd	7	0.041	0.043	0.457
8	Hishām ibn ʻUrwa	5	0.040	0.042	0.464
9	Ismāʻīl ibn Abī Khālid	4	0.032	0.033	0.484
10	Qays ibn Abī Ḥāzim	2	0.032	0.032	0.486
11	Ibn al-Mubārak	8	0.031	0.032	0.496
12	Yaḥyā al-Qaṭṭān	9	0.029	0.031	0.508

Table 3 Ablation, Ṣaḥīḥ Muslim; columns as in Table 2.

k	Narrator	Ṭabaqa	Flux	Solo	Cumulative
1	al-Zuhrī	4	0.104	0.111	0.111
2	Shuʻba	7	0.087	0.093	0.205
3	Ibn Wahb	9	0.077	0.084	0.256
4	Ibn ʻUyayna	8	0.056	0.060	0.300
5	Ibn Jaʻfar (Ghundar)	9	0.040	0.043	0.300
6	Mālik	7	0.036	0.039	0.326
7	ʻUrwa ibn al-Zubayr	3	0.035	0.037	0.346
8	al-Thawrī	7	0.035	0.038	0.383
9	Yūnus ibn Yazīd	7	0.031	0.034	0.383
10	Wakīʻ	9	0.030	0.032	0.402
11	ʻAbd al-Razzāq	9	0.029	0.031	0.413
12	al-Layth ibn Saʻd	7	0.028	0.030	0.433

material depends on a narrator, the route count how many ways through the network run through him. A study of whether individual reports survive the removal of a narrator, rather than how much of the collection depends on him, should therefore use reachability, not flow.

5.3 Redundancy among the leading narrators

Removing the five heaviest eligible narrators together severs 40.2% of Bukhārī's delivered flow and 30.0% of Muslim's, against solo sums of 45.2% and 39.1% (Tables 2 and 3). Table 5 reports the redundancy of the top k relative to their solo sum, with 95% bootstrap intervals, and Figure 2 plots it.

Table 4 Solo severance σ against node rankings, Ṣaḥīḥ al-Bukhārī. Spearman correlation over the sixty heaviest eligible narrators (roots, and narrators who transfer half or more of their material directly to the collector, excluded), and the number of the top ten by σ that are also in the top ten by each measure. Because the pool is selected by flux, its range is restricted, and correlations may be lower than they would be across all narrators. Betweenness is computed on the narrator graph without the collector node, so routes ending at the compiler are not counted.

Measure	Spearman r_s	Top-10 overlap
Weighted degree	0.66	8/10
Shortest-path betweenness	0.46	7/10
PageRank	0.64	7/10
Eigenvector	not defined (acyclic)	
Route count (duplication)	0.48	6/10
Flux ρ	0.998	9/10

Relative redundancy rises with k in both collections, since a larger set contains more pairs whose routes intersect. Muslim's point estimate is the higher at every $k \geq 3$, but the difference is clearly distinguishable from resampling variation only among the top three (0.111 against 0.050; difference 0.058, interval 0.013–0.116) and the top six and seven (at $k = 6$, 0.241 against 0.127; 0.112, 0.017–0.143). At $k = 5$ Muslim's redundancy is roughly double Bukhārī's (0.231 against 0.110), but the lower end of the interval of the difference lies close to zero (0.109; 0.005–0.166), so the result there is borderline. At $k = 4$ and from $k = 8$ onward the interval of the difference includes zero, and by $k = 12$ the two collections are level (0.313 against 0.297). The contrast is therefore a property of the very top of the ranking, not of the collections as wholes, and beyond the top three it also depends on how the sources are seeded (§5.5).

The mechanism is principally relaying, in which a large share of the material carried by a node has been supplied by another. By (9), the redundancy of a pair is the delivered material that passes through both. Shu'ba and Muḥammad ibn Ja'far (Ghundar), who draws 99% of his material from Shu'ba, his teacher, jointly carry 4.3% of Muslim's delivered flow (95% interval 3.7–5.1%) against 2.1% of Bukhārī's (1.8–2.6%); the intervals do not overlap. In Muslim the shared material equals Ghundar's entire solo severance of 4.3% (Table 3): once Shu'ba is removed, removing Ghundar destroys almost nothing further, and the cumulative column shows Yūnus ibn Yazīd and al-Zuhrī have the same student–teacher dynamic. Shu'ba himself is the opposite case, drawing on a large pool of teachers of whom the largest supplies only a tenth of his material. A node ranking places relay and aggregator side by side; only the joint effect separates them.

Table 5 Redundancy among the k heaviest narrators relative to their solo sum, $\big(\sum_{u\in K}\sigma(\{u\}) - \sigma(K)\big)/\sum_{u\in K}\sigma(\{u\})$, with 95% bootstrap percentile intervals ($B = 1000$, top k re-ranked in each replicate). The difference is Muslim minus Bukhārī (bootstrap median and interval); * marks intervals that exclude zero. Rows $k = 1, 2$ are omitted: redundancy is zero at $k = 1$ by definition and negligible at $k = 2$, since al-Zuhrī and Shu'ba share almost no material.

	Bukhārī		Muslim		Difference	
k	Obs.	95% int.	Obs.	95% int.	Median	95% int.
3	0.050	[0.001, 0.091]	0.111	[0.093, 0.125]	0.058*	[0.013, 0.116]
4	0.122	[0.036, 0.133]	0.137	[0.119, 0.154]	0.029	[−0.004, 0.110]
5	0.110	[0.063, 0.144]	0.231	[0.118, 0.246]	0.109*	[0.005, 0.166]
6	0.127	[0.112, 0.177]	0.241	[0.147, 0.266]	0.112*	[0.017, 0.143]
7	0.172	[0.153, 0.192]	0.258	[0.180, 0.311]	0.074*	[0.013, 0.145]
8	0.219	[0.173, 0.232]	0.240	[0.214, 0.323]	0.045	[−0.004, 0.126]
9	0.229	[0.204, 0.260]	0.287	[0.224, 0.330]	0.044	[−0.014, 0.106]
10	0.263	[0.229, 0.282]	0.293	[0.233, 0.335]	0.033	[−0.031, 0.089]
11	0.283	[0.248, 0.307]	0.311	[0.245, 0.351]	0.021	[−0.037, 0.080]
12	0.297	[0.265, 0.323]	0.313	[0.261, 0.355]	0.011	[−0.041, 0.070]

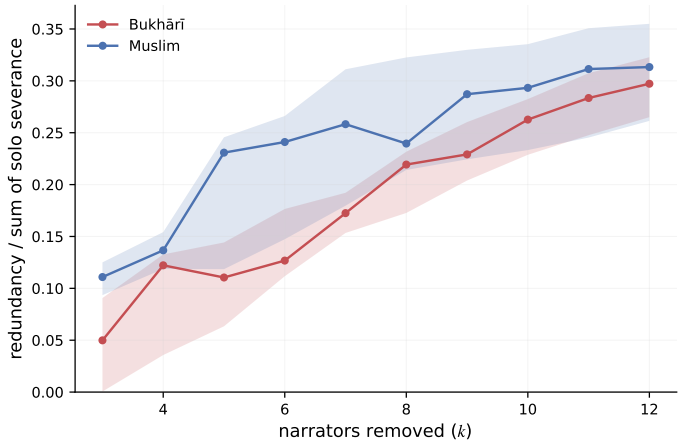

Fig. 2 Redundancy relative to the solo sum among the k heaviest narrators, with 95% bootstrap bands ($B = 1000$).

5.4 Generational bottlenecks

The second finding concerns generations rather than individuals. Within each ṭabaqa, the Gini coefficient of flux measures how unevenly that generation's share of the collection is divided among its members: 0 if every narrator of the generation carries the same amount, almost 1 if a few carry nearly all the information. Table 6 and Figure 3 give the profile.

The Companion layer is near-uniform (0.156, 0.149) due to the inherent seeding condition. Every root receives a unit seed and nothing else, and transmission among

Table 6 Gini coefficient of flux within each ṭabaqa (Ibn Ḥajar's numbering), with cohort sizes n.

	Bukhārī		Muslim	
Ṭabaqa	n	Gini	n	Gini
1	192	0.156	222	0.149
2	97	0.655	124	0.667
3	212	0.739	326	0.696
4	105	0.837	173	0.810
5	100	0.750	137	0.705
6	158	0.741	258	0.707
7	120	0.821	180	0.809
8	124	0.786	123	0.724
9	197	0.763	145	0.757
10	201	0.766	207	0.800
11	46	0.751	28	0.667

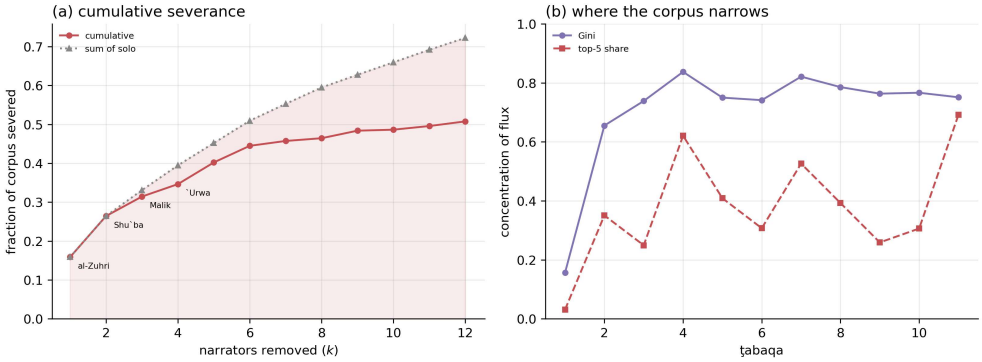

Fig. 3 The effect of ablation in Bukhārī as well as the Gini coefficient of narrators across generations.

the Companions (*riwāyat al-ṣaḥābī `an al-ṣaḥābī*) is largely same-ṭabaqa and therefore discarded (§3.2).

The cohort's flux thus takes two values, one for the 162 (189) roots and zero for the 30 (33) Companions whose onward links the filters removed, and the Gini coefficient equals the zero share, 30/192 and 33/222. Nothing in the construction forces what follows. In both collections concentration rises to a peak at ṭabaqa 4 — 0.837 in Bukhārī, 0.810 in Muslim — where al-Zuhrī alone takes 40% and 31% of his cohort's flux, falls, and peaks again at ṭabaqa 7 (0.821, 0.809), the generation of Shu'ba, Mālik and al-Thawrī. Material that entered evenly across some two hundred Companions passes, twice, through a handful of narrators before fanning out again to the compilers' own teachers.

In both collections ṭabaqa 4 exceeds both neighbors, by 0.09 to 0.11, and ṭabaqa 7 exceeds ṭabaqa 6, by 0.08 and 0.10. In Muslim ṭabaqa 7 also stands clearly above

ṭabaqa 8 (0.809 against 0.724); in Bukhārī the drop is small (0.821 against 0.786), so concentration there remains high from ṭabaqa 7 through ṭabaqa 8.

The profiles diverge late. Muslim shows a smaller third rise at ṭabaqa 10 (0.800, against 0.757 at ṭabaqa 9), in cohorts of 207 and 145 narrators, whereas Bukhārī's ṭabaqa 10 (0.766) is level with its ṭabaqa 9 (0.763); ṭabaqa 11 is small in both collections (46 and 28 narrators) and is not interpreted. Close to the compiler's own teachers, Muslim's material thus passes through a somewhat narrower set of authorities than Bukhārī's, which fits the divergence of the two top-twelve lists at the compiler's end of the chain.

That two collections compiled from different *shuyūkh* narrow at the same two earlier ṭabaqāt means the pattern is not an artifact of either compiler's selection alone. It represents the counterpart, at the level of a whole collection, of the convergence Schacht and Juynboll observed in the isnāds of single traditions [13, 14]: a *madār* not of one report but of the corpus. The two bottlenecks also fall at generations the tradition associates with changes in how hadith were kept: ṭabaqa 4 with the first systematic recording of reports, in which al-Zuhrī is a central figure, and ṭabaqa 7 with the first organized compilations, among them Mālik's *Muwaṭṭa'*.

Whether the narrowing reflects transmission itself or the later preference of compilers for isnāds running through recognized authorities is the question the common-link debate leaves open [15, 16], and the model cannot settle it. What it establishes is that both Ṣaḥīḥs, compipled separately, share it, and how severe each bottleneck is.

5.5 Sensitivity to modeling choices

Table 7 repeats the headline quantities under four alternatives to the baseline: seeding only the Companions, so that roots created by the filters inject nothing; seeding each root by its out-strength; and generation caps $\gamma_{\max} = 3$ and 5. The flux ranking changes little: over the sixty heaviest narrators of the baseline its rank correlation with the baseline is 0.92–0.99, and 0.81 and 0.83 under out-strength seeding, which reintroduces the recorded volume of each root. Al-Zuhrī and Shu'ba lead both collections in every variant.

The relay result holds throughout: the material shared by Shu'ba and Ghundar is 4.0–5.0% of Muslim's delivered flow and 2.0–2.3% of Bukhārī's. Muslim's higher redundancy among the top three also holds in every variant. Among the top five and seven it holds except under out-strength seeding, which reverses it, in line with the borderline intervals at those k (Table 5). Both bottlenecks persist: ṭabaqa 4 exceeds both neighbors, and ṭabaqa 7 exceeds ṭabaqa 6, in every variant and both collections, by at least 0.04. Seeding only the Companions, which removes the roots created by the filters (§5), changes none of these conclusions.

6 Discussion

Vulnerability of a transmission network.

Read as a robustness analysis, the results describe how much of each collection rests on its transmitters. A single narrator, al-Zuhrī, severs 15.9% of Bukhārī's delivered flow and 11.1% of Muslim's; the twelve heaviest together sever 50.8% and 43.3% (Tables 2

Table 7 Headline quantities under alternative modeling choices. r_s: Spearman correlation of flux with the baseline over the baseline's sixty heaviest non-root narrators. Solo: severance of the leading narrator (al-Zuhrī in every variant). Pair: material shared by Shu'ba and Ghundar, as a share of delivered flow. R_k: redundancy of the top k relative to their solo sum. Δ: differences in Gini coefficient between ṭabaqāt. Baseline: uniform seeding, $\gamma_{\max} = 4$.

	Variant	r_s	Solo	Pair	R_3	R_5	$\Delta_{4,3}$	$\Delta_{4,5}$	$\Delta_{7,6}$
Bukhārī	Baseline	1.000	0.159	0.021	0.050	0.110	0.099	0.087	0.080
	Companions only	0.927	0.184	0.023	0.084	0.116	0.072	0.071	0.066
	Out-strength	0.813	0.167	0.023	0.046	0.145	0.053	0.065	0.110
	$\gamma_{\max} = 3$	0.928	0.143	0.022	0.001	0.113	0.094	0.118	0.043
	$\gamma_{\max} = 5$	0.982	0.164	0.020	0.051	0.142	0.098	0.086	0.069
Muslim	Baseline	1.000	0.111	0.043	0.111	0.231	0.115	0.105	0.102
	Companions only	0.941	0.143	0.050	0.119	0.185	0.066	0.069	0.116
	Out-strength	0.831	0.110	0.044	0.117	0.134	0.060	0.073	0.123
	$\gamma_{\max} = 3$	0.921	0.102	0.043	0.123	0.212	0.114	0.137	0.068
	$\gamma_{\max} = 5$	0.991	0.110	0.040	0.113	0.227	0.118	0.105	0.101

and 3). These are upper bounds on dependence on the recorded routes, since severed material is destroyed rather than rerouted (§4.2).

The two collections are fragile in different ways. Bukhārī's flow rests more on its leading narrators (the top five jointly sever 40.2% against 30.0%), but these carry largely complementary material; Muslim's rests less on any one of them, but more of what its leading narrators carry is shared, both among the top three (relative redundancy 0.111 against 0.050; Table 5) and in teacher–relay pairs such as Shu'ba and Ghundar. Connectivity-based robustness [25] cannot draw this distinction, since it registers only whether routes remain, not how much of the material that actually arrived each removal destroys.

Where the information lies.

Single-node severance agrees closely with flux and only moderately with the other rankings (Table 4). The identities of §4.2 show that the first agreement is forced: severance is flux discounted by leakage, so it ranks narrators almost exactly as flux does. What a conserved flow adds lies at the level of sets, where redundancy (9) has an exact reading as material carried more than once. Generations are the limiting case: since lateral links are excluded, a route meets each ṭabaqa at most once, so the narrators of a generation share no material and their severances add up exactly. For transmission networks generally, this suggests that node-removal studies are most informative when they remove sets.

Implications for hadith studies.

Three results bear on questions the classical tradition poses.

Redundancy can make the concept of independent routes measurable, meaning material that reaches the compiler only through a teacher and his relay is less independently attested than its two narrators' separate standings might suggest. The relay structure separates two roles that rankings conflate, the aggregator who draws on many teachers and the relay who passes on one teacher's material.

The generational bottlenecks bear on the common-link debate as mentioned before, since they measure where, and how sharply, the corpus narrows.

Applied to further collections, the same measures would allow a broader comparison of hadith collections, again with the main obstacle being digitalization.

Beyond the isnād.

The construction requires only that transmission be recorded, directed and acyclic, with known sources and a single known terminus. Other religious traditions, such as Buddhist textual lineages, share this structure; bibliometric applications such as academic lineages are also natural fits; and so is the dependency graph of a software application, in which libraries reference one another until they reach the application itself. In the case of software, severance measures how much of an application rests on a given library, and redundancy separates a library that merely repackages another from one that combines many. Other domains can also be analyzed, with specific caveats. News can be modeled the same way as transmission of information from a large station or aggregator, to smaller aggregators such as portals and finally the absorbing audience. One could then ask which nodes are the most critical to the flow of information, and which are simply copying from larger sources. We leave this to future work.

Limitations.

- *Source.* The chains and narrator identities are those of a community-maintained digital edition, not a critical one, and the results describe that record as it stands. Errors in narrator identification or generation labels would propagate into the flux; Muslim, the more heavily filtered collection, is the more exposed. Between 6% and 7% of injected material is stranded at narrators whose onward links the filters removed.
- *Ordering.* Generation is the only ordering evidence. Same-ṭabaqa links (5.1% of transmissions in Bukhārī, 11.6% in Muslim) are excluded, and links spanning more than $\gamma_{\max} = 4$ ṭabaqāt (1.3%, 1.6%) are discarded; the main results hold for $\gamma_{\max} = 3$ and 5 (§5.5).
- *Seeding.* Uniform seeding gives every root equal weight, including the minority of roots outside the Companion generation that the filters create; out-strength seeding would instead reproduce the recorded volume of each Companion. Seeding only the Companions changes none of the conclusions, and out-strength seeding affects only the redundancy of the top five and seven (§5.5).
- *Pooling.* The operator is first-order: each narrator's output is divided in fixed proportions regardless of which teacher supplied it. Node and edge fluxes reproduce the recorded totals exactly, but redundancy between narrators who are not directly linked is computed over routes that may combine links attested in different chains, and can differ from the overlap the chains themselves record [26, 27]. The teacher–relay pairs on which our interpretation rests, such as Shu'ba and Ghundar, are directly linked, and the relay draws almost all his material from the teacher, so pooling has little room to act there; validating set-level redundancy against the chains is left to future work (§7.1).

- *Scope.* Two collections are compared. The bootstrap measures dependence on the selection of traditions within each, not on the choice of collections.

7 Conclusion

Treating hadith transmission as a conserved transfer flow, in Borgatti's sense, on a directed acyclic network whose ordering and terminus the isnād itself supplies, makes the flow exact in one pass and node removal well posed. Severance of single narrators reduces to flux; redundancy among sets of narrators does not, and it is there that the flow adds to standard centrality. In Ṣaḥīḥ Muslim the leading narrators overlap more than in Ṣaḥīḥ al-Bukhārī among the top three, and relays such as Ghundar, whose material is almost entirely Shu'ba's account for a larger share of the flow. Both collections narrow sharply at the generation of al-Zuhrī and at that of Shu'ba, Mālik and al-Thawrī, and Muslim shows a smaller third rise close to his own teachers. The tools are not original, from Newman's random-walk flow and vitality measures to the multiplier models of input–output economics (Appendix C); the contribution lies in their adaptation to a historical transmission record and in what they show about two of its central collections. To our knowledge, this is the first treatment of an isnād corpus as a conserved flow, and the first to give an exact quantitative form to a concept of classical hadith criticism, the independence of routes. It complements chain-level approaches such as that of Aldoboni et al. [17], which judge the authenticity of individual chains, by measuring how much of a collection rests on its narrators. The same construction applies wherever material passes forward in time from recorded sources to a known destination.

7.1 Future work

The principal constraint is the state of the sources. The corpus is community-maintained and narrator identities are not resolved across the different digitized sources. Because of the peculiarities of Arabic naming conventions, Wirahman [28] resolves narrator identities probabilistically, which falls short of the certainty about a narrator's identity that hadith criticism requires.

An authoritative database, linked to the biographical dictionaries rather than to any one collection, would allow the flow to be solved on the union of several works, with a collector node for each. Solved per tradition on that union, concentration of the flux on one node would become a general test for the common link, applied across the corpus rather than to traditions selected in advance.

Two extensions concern the model itself. The first carries the subject of each report through the model. Counting transmissions by subject gives one matrix $A^{(c)}$ per subject c, with $\sum_c A^{(c)} = A$, and with terminal counts and sources counted the same way, an operator $P^{(c)}$ solved, like P, in one pass.

Its results can be of interest because a narrator may pass a quarter of his material to one student and yet all of that material pertains to a topic like fasting. Two narrators carrying the same total flux, one across many chapters and one within a single chapter, can then be told apart, as can the subjects severed by removals. This presupposes a

standard subject scheme and a rule for reports bearing on several subjects, which can be analyzed more in depth in future iterations.

The second lets the model remember where material came from. At present, a narrator's incoming material is pooled before it is passed on. The model therefore records how much Ghundar receives from Shu'ba and from his other teachers, but not how those origins are allocated among his outgoing transmissions. Using incoming teacher–student links as memory states, in the vein of memory networks [26, 27], would preserve that provenance, and because the expanded network remains acyclic, the same one-pass solution applies. Comparing the results with the present ones, or with overlap counted directly from the recorded chains, would show how much the pooling matters.

Formalizing the terminology of hadith criticism.

Redundancy gives an exact form to one notion of the critical tradition, the independence of routes, and others get the same treatment. Ibn Ḥajar classifies a report by the fewest narrators at any level of its transmission — *gharīb* with one, *'azīz* with two, *mashhūr* with more, and *mutawātir* when their number precludes collusion — which the subgraph of a single tradition gives directly as its narrowest generation. Corroboration by a parallel chain through the same Companion (*mutāba'a*) or through another (*shāhid*) is the independence of routes at the level of a single report, and the elevation of a chain (*'uluww*) is related to, though not identical with, the depth of Appendix A.1. Judgments imported from the biographical literature would enter instead as weights on edges: the grading of narrators (*ta'dīl*, *jarḥ*), meaningful in collections not already selected for their narrators' reliability, and the competing conditions for accepting a link, an established meeting (*liqā'*) or contemporaneity alone (*mu'āṣara*), commonly attributed to al-Bukhārī and to Muslim, once death dates are resolved. The latter may also bear on the higher share of same-generation links in Muslim (Table 1). Formalizing these notions would not replace the critics' judgments, but it would make them comparable across collections and testable against existing records.

Declarations

Availability of data and materials.

The code is available at the repository https://github.com/saitiart/hadith-research/. The data was collected from https://sunnah.com/. The data is publicly available.

Competing interests.

The author declares that they have no competing interests.

Funding.

No funding was received for this paper.

Authors' contributions.

The author conceived the study, developed the model, carried out the analysis and wrote the manuscript.

Use of AI tools.

The author used Claude Opus 5.5 (Anthropic) and Kimi 3.1 (Moonshot AI) to review the manuscript, suggest and draft revisions to some passages, and assist in writing the plotting, bootstrapping and sensitivity-analysis code. The model, its formulation and implementation, the data processing, and the interpretation of the results are the author's own. The author reviewed and verified all AI-assisted text and code and takes full responsibility for the content of the paper.

Appendix A Depth and hop decomposition

A.1 Depth

The adjoint system reads column u of P rather than row v:

$$(I - P^{\top})\,\nu = \mathbf{1} \qquad \Longleftrightarrow \qquad \nu_u = 1 + \sum_v P_{vu}\,\nu_v, \tag{A1}$$

so $\nu = \mathcal{N}^{\top}\mathbf{1}$ counts the *nodes* remaining from u to the collector, averaged over onward channels by their splitting fractions; $\nu_\Omega = 1$ because column Ω is empty, and hops are $\nu - 1$. Orphans receive $\nu = 1$ by the same mechanism despite never reaching Ω, and must be masked before ν is reported.

Corollary 1 (Duality) $\rho^{\top}(I - P^{\top})\nu$ evaluates to $\|\rho\|_1$ grouping right and to $\nu^{\top}S$ grouping left, so $\bar{\ell} := \|\rho\|_1/\|S\|_1 = \nu^{\top}S/\|S\|_1$.

Under out-strength seeding and with no edges discarded, $\bar{\ell}$ coincides with the mean chain length of the corpus, since every unit then follows a recorded chain end to end. Excluded edges and orphan loss both shorten the modeled path, so with filtering $\bar{\ell}$ underestimates the empirical mean; Table 1 reports the gap as a calibration diagnostic. In input–output terms, $\nu_u = (\mathbf{1}^{\top}\mathcal{N})_u$ is the column multiplier of u and $\bar{\ell}$ the aggregate multiplier (Appendix C).

A.2 Hop decomposition

Let $R_k = P^k S$, so that $(R_k)_v$ is the material reaching v in exactly k hops and $\rho = \sum_k R_k$; this is the Neumann series (3) read term by term, with at most m terms however many paths the network contains. In the structural path analysis of social accounting matrices [29], the influence of each path is scaled by a path multiplier Δ_p/Δ that accounts for feedback through adjacent circuits. Here every principal submatrix of $I - P$ is unit triangular, so every path multiplier equals one and

$$\mathcal{N}_{vu} \;=\; \sum_{p\,:\,u\to v} \;\prod_{(u'\to v')\in p} P_{v'u'}.$$

Weighting each term by its hop index gives the mean path length of the material reaching v,

$$\bar{h}_v \;=\; \frac{\sum_k k\,(R_k)_v}{\sum_k (R_k)_v} \;=\; \mathbb{E}\big[\text{path length} \;\big|\; \text{material reaches } v\big]. \tag{A2}$$

Appendix B A hydraulic reading

Stating (2)–(9) in the language of fluid flow could allow for a more intuitive understanding of the mechanism.

Read each node as a junction and each retained edge as a pipe. The convention of (1) fits naturally within a plumbing analogy: column u is what u sends downstream, row v is what v receives, and P_{vu} is the fraction of u's throughput carried by the pipe to v. Each junction is a fixed-ratio splitter: whatever arrives is divided among the outgoing pipes in the same proportions, whatever the state of the pipes downstream may be. Every live junction passes on all it receives, so its column sums to one; an orphan is a dead end whose column is empty, and water reaching it goes no further. Ω is the drain into which the delivered water runs.

The forward solve.

S is the water poured in at the taps. $\mathcal{N}$ describes what the system does to a single unit: $\mathcal{N}_{vu}$ is the share of one unit poured in at u that ever passes junction v, summed over every route it can take, a finite total because no water flows back. Multiply by what was actually poured and $\rho = \mathcal{N}S$ is what a flow meter at each junction would read.

Timing the flow.

Let us pour the water in as a single pulse, and let each pipe measure one tick of the clock. $R_k = P^k S$ is the water passing each junction at tick k, so a meter with a clock records the hop decomposition of §A.2, and $\rho = \sum_k R_k$ is its total over time. The pulse clears the system within m ticks, because no route is longer than the longest chain. $\bar{h}_v$ in (A2) is the mean time for the water to reach v. An elevated chain is a pipe that bypasses intermediate junctions and directly heads to the destination, so water taking it arrives early and lowers that mean. The total in transit, $\|R_k\|_1$, stays at $\|S\|_1$ while nothing has reached the drain and falls as water runs out.

The adjoint solve.

Stand at a tap with the water off and ask how far a unit poured in there would get before it drains or comes to a dead end: that is ν_u, worked out by (A1) backwards from the drain. Where a junction splits, ν_u averages the onward routes by how much each carries, so a junction feeding one short and one long pipe falls in between, and one whose routes run into dead ends falls lower. Like $\mathcal{N}$, it describes the pipes alone: both are fixed by the recorded topology and unchanged by the seeding, whereas ρ depends on both, and one backward substitution serves every seeding one might later try.

The identity.

The corollary says the same total can be counted at the meters or at the taps: reading every meter gives $\|\rho\|_1$, weighting each tap by the water poured into it gives $\nu^\top S$, and $\bar{\ell}$ divides either by $\|S\|_1$ to give the mean number of junctions a unit crosses. What it measures is depth, not branching. Long relayed chains keep water in transit for many hops and $\bar{\ell}$ is large; short chains empty at once and it approaches its floor of 2, reached only when every tap empties after a single pipe. An unbranched two-hop chain already gives $\nu = 3$. Splitting moves $\bar{\ell}$ only through the average it induces over routes of unequal length, and may lower it as easily as raise it.

Capping junctions.

Ablation caps a set of junctions K (§4.2): nothing is poured in there, and water that arrives there spills out rather than going on. Because the splitters upstream are fixed-ratio, closing a junction does not push its share into the neighboring pipes, as it would in real life; the share that was headed for K is simply lost, and we can imagine a cracked pipe leaking water at that point. This is the primary difference from max-flow vitality, which lets the water find an alternative path. Severance (7) is the drop in the reading at the drain. For a single junction the drop is easy to read off: no water passes a junction twice, so capping v loses exactly the water that passed v and would have reached the drain, the meter reading ρ_v times the fraction $\mathcal{N}_{\Omega v}$ of v's water that drains, which is (8).

Counting water twice.

Cap several junctions one at a time and add up the losses, and any water that ran through two of them is counted twice. Capping them together counts it once. The difference is the redundancy (9): the water that passed more than one of the capped junctions. Two junctions in series on the same line share their water, two on parallel lines share none. In the worked example of Appendix C, v_2 and v_4 are partly in series, and capping both loses $\frac{1}{4}$ less than the sum of capping each; v_2 and v_3 feed v_4 through parallel pipes, and their losses add exactly. The distinction between relay and aggregator is the same one. A relay is a junction fed almost entirely by one pipe from a single upstream junction: cap the upstream junction and the relay runs dry, so capping it as well loses almost nothing more. An aggregator is fed by many pipes, and capping any one of them only lowers its flow. One idealization matters here more than anywhere else. Each junction is a well-mixed tank: water arriving from different pipes mixes before it is split, so the record of which incoming pipe feeds which outgoing one is lost. Totals at every meter are unaffected by this, but the water two junctions share depends on it, which is why redundancy is the quantity most exposed to the pooling assumption.

Appendix C A worked example

C.1 Corpus

Four narrators with $g(v_1) = 0$, $g(v_2) = g(v_3) = 1$, $g(v_4) = 2$. Eight chains, shown canonically (early-to-late) after the orientation convention of §3:

h_1: $v_1 \to v_2 \to v_4$	h_3: $v_1 \to v_2$	h_5: $v_1 \to v_3 \to v_4$	h_7: $v_1 \to v_4$
h_2: $v_1 \to v_2 \to v_4$	h_4: $v_1 \to v_2$	h_6: $v_1 \to v_3 \to v_4$	h_8: $v_1 \to v_4$

Chains h_7, h_8 skip generation 1 and are the elevated isnāds; h_3, h_4 are direct citations by the collector. The multiplicities of §3 are $a_{v_1 v_2} = 4$, $a_{v_1 v_3} = a_{v_1 v_4} = 2$, $a_{v_2 v_4} = 2$, $a_{v_3 v_4} = 2$, and the terminal counts $\tau_{v_2} = 2$ (from h_3, h_4) and $\tau_{v_4} = 6$ (from $h_1, h_2, h_5, h_6, h_7, h_8$). Every edge joins distinct generations, so no edge is lateral in the sense of §3.2 and $\lambda = 0$; the skip edge $v_1 \to v_4$ spans two generations and survives $\gamma_{\max} = 4$.

C.2 Operators

Writing $A_{vu} = a_{uv}$ and adjoining the absorbing collector of §3.3:

$$A = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \end{array} \begin{array}{c} \begin{array}{cccc} v_1 & v_2 & v_3 & v_4 \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 2 & 2 & 2 & 0 \end{pmatrix} \end{array}, \qquad \tilde{A} = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ \Omega \end{array} \begin{array}{c} \begin{array}{ccccc} v_1 & v_2 & v_3 & v_4 & \Omega \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 0 & 2 & 0 & 6 & 0 \end{pmatrix} \end{array}$$

The bottom row is the terminal counts and the empty column at Ω is what makes it absorbing. Column sums are $c = (8, 4, 2, 6, 0)$; note $c_{v_2} = 2 + \tau_{v_2} = 4$, so the two direct citations enter the denominator and compete with the student edge, as the remark of §3.3 requires. Normalizing by (1),

$$P = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ \Omega \end{array} \begin{array}{c} \begin{array}{ccccc} v_1 & v_2 & v_3 & v_4 & \Omega \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 1 & 0 \end{pmatrix} \end{array}$$

with column sums $(1, 1, 1, 1, 0)$, confirming $\sum_v P_{vu} = \mathbf{1}\{u \text{ live}\}$. Ordering $v_1 \prec v_2 \prec v_3 \prec v_4 \prec \Omega$ by generation makes P strictly lower triangular, which is the hypothesis of Proposition 1. Only v_1 has zero in-strength and positive out-strength, so $\mathcal{R} = \{v_1\}$ and, under uniform seeding, $S = e_{v_1}$ with $\|S\|_1 = 1$.

C.3 Evolution

Because P is strictly triangular, Proposition 1 applies and the Green's function (3) is the finite sum $\mathcal{N} = (I - P)^{-1} = I + P + P^2 + P^3$:

$$\mathcal{N} = \begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ \Omega \end{array} \begin{array}{c} \begin{array}{ccccc} v_1 & v_2 & v_3 & v_4 & \Omega \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 1 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & 1 & 0 & 0 \\ \frac{3}{4} & \frac{1}{2} & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} \end{array}$$

It is unit triangular with $\det(\mathcal{N}) = 1$, as the proposition requires. Read a column: $\mathcal{N}_{vu}$ is the share of one unit injected at u that ever crosses v, summed over every route. Column v_2 says that half of what enters at v_2 reaches v_4, the other half being absorbed directly, while all of it reaches Ω by one route or the other. The Ω row is the delivery fraction from each node, here uniformly 1 because the network has no orphans; in the corpora of §5 the corresponding figures average 93.6% and 92.8%.

Everything below is one reading of this matrix. The flux is $\rho = \mathcal{N}S$, here simply column v_1; the depth vector is $\nu = \mathcal{N}^\top \mathbf{1}$, the column sums; and $\bar{\ell}$ is the sum of the seeded column. That $\mathcal{N}$ and ν do not depend on S is the distinction drawn in Appendix B: they rate the recorded topology, whereas ρ also depends on what was poured in.

Multiplying $\mathcal{N}$ by S at once would hide how the material got there, so we instead take the series a term at a time. The kth term of (3) is $R_k = P^k S$, the material that has traveled exactly k hops, and its rows are the hop decomposition of §A.2:

k	v_1	v_2	v_3	v_4	Ω	$\lVert R_k \rVert_1$
0	1	·	·	·	·	1
1	·	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	·	1
2	·	·	·	$\frac{1}{2}$	$\frac{1}{2}$	1
3	·	·	·	·	$\frac{1}{2}$	$\frac{1}{2}$
ρ	1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{3}{4}$	1	$\frac{7}{2}$

Row $k = 0$ is the injection S itself. At $k = 1$ the source splits three ways in the proportions of column v_1 of P, the $\frac{1}{4}$ arriving at v_4 being the elevated chains, which reach the last generation in a single hop. At $k = 2$ that material has drained into Ω, alongside the $\frac{1}{4}$ absorbed directly from v_2; meanwhile v_4 receives $\frac{1}{2} \cdot \frac{1}{2}$ from v_2 and $1 \cdot \frac{1}{4}$ from v_3. Only that last $\frac{1}{2}$ remains in transit, and it is absorbed at $k = 3$, so $R_4 = 0$ and the series terminates as nilpotence guarantees.

The column totals are the flux, $\rho = \sum_k R_k$, and they reproduce column v_1 of $\mathcal{N}$ as they must. The final column tracks $\lVert R_k \rVert_1$, the material still in transit after k hops: it holds at $\lVert S \rVert_1 = 1$ while everything is still moving and falls to zero at $k = m$, the profile described in Appendix B. Its weighted average is $\bar{\ell} = \frac{7}{2}$.

Both invariants hold. The flux satisfies (2), $(I-P)\rho = S$, row by row: row v_4 asserts that what v_4 holds equals what his three teachers sent, $\frac{3}{4} - \left[\frac{1}{4}(1) + \frac{1}{2}(\frac{1}{2}) + 1(\frac{1}{4})\right] = 0$, and row Ω does the same for the collector, $1 - \left[\frac{1}{2}(\frac{1}{2}) + 1(\frac{3}{4})\right] = 0$. Conservation (5) holds as $\rho_\Omega = 1 = \|S\|_1$, the Ω row of $\mathcal{N}$ having already shown that delivery is complete here.

C.4 Depth

Solving the adjoint system (A1) backwards from the collector gives $\nu_\Omega = 1$, $\nu_{v_4} = 1 + 1(1) = 2$, $\nu_{v_3} = 1 + 1(2) = 3$, $\nu_{v_2} = 1 + \frac{1}{2}(2) + \frac{1}{2}(1) = \frac{5}{2}$, and $\nu_{v_1} = 1 + \frac{1}{2}(\frac{5}{2}) + \frac{1}{4}(3) + \frac{1}{4}(2) = \frac{7}{2}$.

	g	$\rho_v/\|S\|_1$	gen. share	ν_v	$\bar{h}_v$
v_1	0	1.000	1.000	3.500	0.00
v_2	1	0.500	0.667	2.500	1.00
v_3	1	0.250	0.333	3.000	1.00
v_4	2	0.750	1.000	2.000	1.67

By (A2), $\bar{h}_{v_4} = \left(1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{2}\right)/\frac{3}{4} = \frac{5}{3}$: a third of v_4's material arrived by the elevated route, which lowers his depth index accordingly. Note $\nu_{v_3} > \nu_{v_2}$ although v_2 carries twice v_3's flux — v_3's only outlet is a two-step path, whereas v_2 discharges half of his material directly to the collector. Depth and volume are independent axes.

The duality of Corollary 1 is exact: $\|\rho\|_1 = \frac{7}{2} = \nu^\top S = \nu_{v_1}$, so $\bar{\ell} = \frac{7}{2}$ nodes, i.e. $\frac{5}{2}$ hops. Counting directly, h_1, h_2, h_5, h_6 contribute three hops each (including the terminal step) and h_3, h_4, h_7, h_8 two, giving $20/8 = \frac{5}{2}$. With a single root, uniform and out-strength seeding coincide up to scale, so the identity applies exactly here; agreement is then a test of whether generation filter is not dropping edges asymmetrically.

C.5 Edge flux and Kirchhoff's law

The edge flux of §4 is $f_{vu} = P_{vu}\rho_u$: the fraction of u's throughput carried by the pipe to v, times what u actually holds.

channel	P_{vu}	ρ_u	f_{vu}
$v_4 \to \Omega$	1	3/4	3/4
$v_1 \to v_2$	1/2	1	1/2
$v_1 \to v_3$	1/4	1	1/4
$v_1 \to v_4$	1/4	1	1/4
$v_2 \to v_4$	1/2	1/2	1/4
$v_2 \to \Omega$	1/2	1/2	1/4
$v_3 \to v_4$	1	1/4	1/4

Equation (4) reads these two ways at once. Its left member, $\sum_u f_{vu} + S_v$, sums the column of the table over channels ending at v and adds whatever was injected there;

its right member, $\sum_w f_{wv}$, sums the channels leaving v; and both equal ρ_v. Nothing accumulates at a junction.

At v_4: inflow $\frac{1}{4}+\frac{1}{4}+\frac{1}{4}$ from v_1, v_2, v_3, no injection, giving $\frac{3}{4}=\rho_{v_4}$, which leaves again in a single channel $v_4 \to \Omega$ of the same size. At v_2: inflow $\frac{1}{2}$ from v_1, matched by $\frac{1}{4}+\frac{1}{4}$ leaving down the two channels $v_2 \to v_4$ and $v_2 \to \Omega$ — the split that $c_{v_2}=4$ produced. At the source, the injection $S_{v_1}=1$ is the whole of the inflow and leaves as $\frac{1}{2}+\frac{1}{4}+\frac{1}{4}$.

Summing (4) over all live nodes is what telescopes to (5): every interior f_{vu} appears once as an outflow and once as an inflow and cancels, leaving only what enters at the taps and what reaches the dead columns. Here that is $\frac{1}{4}+\frac{3}{4}=1$ arriving at Ω against $\|S\|_1=1$ injected, with no orphan term.

C.6 Ablation

Solo severances follow from (6)–(7), each by a single forward substitution on the projected system $(I-P_K)\rho^{(K)}=S_K$: removing v_2 leaves $\rho_\Omega^{(K)}=\frac{1}{2}$ and so $\sigma(\{v_2\})=\frac{1}{2}$; likewise $\sigma(\{v_3\})=\frac{1}{4}$ and $\sigma(\{v_4\})=\frac{3}{4}$. Two pairs show the difference between overlapping and disjoint routes.

Removing v_2 and v_4 together severs everything, $\sigma=1$, against a solo sum of $\frac{1}{2}+\frac{3}{4}=\frac{5}{4}$. The redundancy $\sum_{u\in K}\sigma(\{u\})-\sigma(K)=\frac{1}{4}$ is the material counted twice when the two are ranked separately, since half of v_2's onward output passes through v_4. Removing v_2 and v_3 together severs $\frac{3}{4}$, exactly their solo sum $\frac{1}{2}+\frac{1}{4}$, and the redundancy is zero: their routes into v_4 are parallel and neither carries material the other supplied. This is the contrast of §4.2 at minimum scale, and neither pair is distinguishable by flux alone, since ρ ranks v_4 above v_2 above v_3 in both cases.

Note also that ablation destroys rather than reroutes: with v_4 removed, v_2's column sum falls to $\frac{1}{2}$ and the material entering the $v_2 \to v_4$ channel is lost rather than pushed down $v_2 \to \Omega$.

C.7 Relation to input–output analysis

The flux equation has the form of the accounting multiplier model of a social accounting matrix [29], in which the incomes y_n of the endogenous accounts satisfy $y_n = A_n y_n + x$, so that $y_n = (I-A_n)^{-1}x = M_a x$. Here A_n holds each account's expenditure shares by column, x the exogenous injections and M_a the accounting multipliers. The correspondence is direct and needs no transposition: ρ stands for y_n, P for A_n, S for x and $\mathcal{N}$ for M_a, with material moving from teacher to student as payments move from spending to receiving account. In the multiplier model the columns of A_n sum to less than one because part of each account's spending leaks to the exogenous accounts, and total leakages equal total injections. The collector plays the role of those accounts: restricted to the narrators, the columns of P fall short of one by exactly the share each narrator hands to the compiler, and (5) is the balance of injections and leakages, with delivery and orphan loss as the two leaks.

References

[1] Brown, J.A.C.: Hadith: Muhammad's Legacy in the Medieval and Modern World. Oneworld, Oxford (2009)

[2] Borgatti, S.P.: Centrality and network flow. Social Networks **27**(1), 55–71 (2005) https://doi.org/10.1016/j.socnet.2004.11.008

[3] Şentürk, R.: Narrative Social Structure: Anatomy of the Hadith Transmission Network, 610-1505. Stanford University Press, Stanford, CA (2005). https://books.google.mk/books?id=WaMbabqdxQMC

[4] Saeed, S., Yousuf, S., Khan, F., Rajput, Q.: Social network analysis of hadith narrators. Journal of King Saud University - Computer and Information Sciences **34**(6, Part B), 3766–3774 (2022) https://doi.org/10.1016/j.jksuci.2021.01.019

[5] Alam, T., Schneider, J.: Social Network Analysis of Hadith Narrators from Sahih Bukhari (2021). https://arxiv.org/abs/2102.02009

[6] Solla Price, D.J.: Networks of scientific papers. Science **149**(3683), 510–515 (1965) https://doi.org/10.1126/science.149.3683.510

[7] Hummon, N.P., Doreian, P.: Connectivity in a citation network: The development of DNA theory. Social Networks **11**(1), 39–63 (1989) https://doi.org/10.1016/0378-8733(89)90017-8

[8] Batagelj, V.: Efficient Algorithms for Citation Network Analysis. arXiv:cs/0309023 (2003)

[9] Karrer, B., Newman, M.E.J.: Random graph models for directed acyclic networks. Physical Review E **80**, 046110 (2009) https://doi.org/10.1103/PhysRevE.80.046110

[10] Clough, J.R., Gollings, J., Loach, T.V., Evans, T.S.: Transitive reduction of citation networks. Journal of Complex Networks **3**(2), 189–203 (2015) https://doi.org/10.1093/comnet/cnu039

[11] Kempe, D., Kleinberg, J., Kumar, A.: Connectivity and inference problems for temporal networks. Journal of Computer and System Sciences **64**(4), 820–842 (2002) https://doi.org/10.1006/jcss.2002.1829

[12] Holme, P., Saramäki, J.: Temporal networks. Physics Reports **519**(3), 97–125 (2012) https://doi.org/10.1016/j.physrep.2012.03.001

[13] Schacht, J.: The Origins of Muhammadan Jurisprudence. Clarendon Press, Oxford (1950)

[14] Juynboll, G.H.A.: Muslim Tradition: Studies in Chronology, Provenance and Authorship of Early Hadith. Cambridge Studies in Islamic Civilization. Cambridge University Press, Cambridge (2008). https://books.google.mk/books?id=oLpG68cO7fMC

[15] Motzki, H.: Die Anfänge der Islamischen Jurisprudenz: Ihre Entwicklung in Mekka Bis zur Mitte des 2./8. Jahrhunderts. Abhandlungen für die Kunde des Morgenlandes, vol. Bd. 50,Nr. 2. Deutsche Morgenländische Gesellschaft, Stuttgart (1991)

[16] Schoeler, G.: Charakter und Authentie der Muslimischen Überlieferung Über Das Leben Mohammeds. Studien zur Geschichte und Kultur des islamischen Orients. De Gruyter, Berlin (2014)

[17] Aldoboni, G., Alam, M.T., Nassar, L., Karray, F.: Isnadguard: Detecting fabricated chains of narration in hadith transmission networks. In: The 6th Muslims in ML (MusIML) Workshop at ICML 2026 (2026). https://openreview.net/forum?id=zCYlbhCwiA

[18] Newman, M.E.J.: A measure of betweenness centrality based on random walks. Social Networks **27**(1), 39–54 (2005)

[19] Sunnah.com: Sunnah.com: The Hadith of the Prophet Muhammad at Your Fingertips. https://sunnah.com. Accessed: 2026-Aug-16 (2026)

[20] Brandes, U., Fleischer, D.: Centrality measures based on current flow. In: STACS 2005: 22nd Annual Symposium on Theoretical Aspects of Computer Science. Lecture Notes in Computer Science, vol. 3404, pp. 533–544. Springer, Berlin (2005). https://doi.org/10.1007/978-3-540-31856-9_44

[21] Everett, M.G., Borgatti, S.P.: The centrality of groups and classes. The Journal of Mathematical Sociology **23**(3), 181–201 (1999) https://doi.org/10.1080/0022250X.1999.9990219

[22] Koschützki, D., Lehmann, K.A., Peeters, L., Richter, S., Tenfelde-Podehl, D., Zlotowski, O.: Centrality indices. In: Brandes, U., Erlebach, T. (eds.) Network Analysis: Methodological Foundations. Lecture Notes in Computer Science, vol. 3418, pp. 16–61. Springer, Berlin (2005). https://doi.org/10.1007/978-3-540-31955-9_3

[23] Ausiello, G., Franciosa, P.G., Lari, I., Ribichini, A.: Max flow vitality in general and st-planar graphs. Networks **74**(1), 70–78 (2019) https://doi.org/10.1002/net.21878

[24] Efron, B., Tibshirani, R.J.: An Introduction to the Bootstrap. Chapman & Hall, New York (1993)

[25] Albert, R., Jeong, H., Barabási, A.-L.: Error and attack tolerance of complex

networks. Nature **406**(6794), 378–382 (2000) https://doi.org/10.1038/35019019

[26] Rosvall, M., Esquivel, A.V., Lancichinetti, A., West, J.D., Lambiotte, R.: Memory in network flows and its effects on spreading dynamics and community detection. Nature Communications **5**, 4630 (2014) https://doi.org/10.1038/ncomms5630

[27] Lambiotte, R., Rosvall, M., Scholtes, I.: From networks to optimal higher-order models of complex systems. Nature Physics **15**(4), 313–320 (2019) https://doi.org/10.1038/s41567-019-0459-y

[28] Wirahman, T.: Linking Hadith Narrator Identities Across Heterogeneous Arabic Biographical Databases: A Multi-Signal Entity Resolution Pipeline (2026). https://arxiv.org/abs/2607.05424

[29] Defourny, J., Thorbecke, E.: Structural path analysis and multiplier decomposition within a social accounting matrix framework. The Economic Journal **94**(373), 111–136 (1984) https://doi.org/10.2307/2232220