Three Entropy Measures of Double-Entry Bookkeeping Graph Classification Structure*

Pierre Jinghong Liang

Accounting AI Research Lab
Carnegie Mellon University

April 2026
Original Draft: March 2025

Abstract

In this position paper, three laws are proposed to describe the conventional double-entry system of bookkeeping: (1) the balance law; (2) the conservation law; (3) the linearity law. The three laws, and their associated algebraic identities, are then expressed in the bookkeeping language of journals and ledgers, and equivalently in the mathematical language of a directed and weighted graph (or network). The three laws are shown to impose a certain recognizable classification structure on the records generated by conventional double-entry bookkeeping. Within the graph classification framework, we develop computational tools, inspired by Claude Shannon's entropy concept, to capture the graph properties of the classified bookkeeping records. These measures are: Balance-sheet Node-entropy, Transaction Edge-entropy, and Bookkeeping Graph (Laplacian) entropy. The relative entropy (or the Kullback-Leibler divergence) defined on these entropy metrics is retrieved as the quantity of new knowledge contained in the current bookkeeping records relative to past records. The graph-theoretic mathematical representation we construct here offers a framework for some recent computational experiments which deploy related entropy and other metrics computed on bookkeeping graphs constructed using financial statement data from publicly traded companies. Our position is that quantifying the bookkeeping graph structure has the potential to offer insights into accounting information transmission, processing, and final end-uses.

Keywords: double entry bookkeeping, adjacency matrix, bookkeeping graphs, Laplacian matrix, Shannon entropy, von Neumann graph entropy.

*This manuscript grew from a draft prepared specifically for a presentation at Columbia University in March 2025, which benefited from in-depth discussions with Shyam Sunder. The major additional content in this 2026 version is the axiomatic treatment of both the node-entropy and the edge-entropy. Overally the work reported here emerged from reading classic research on the double-entry bookkeeping by Yuji Ijiri and John Butterworth, on accounting classification by Shyam Sunder, on applied information theory by Henri Theil and Baruch Lev, as well as bookkeeping structure work by Anil Arya, John Fellingham, Jon Glover, Haijin Lin, Brian Mittendorf, Doron Nissim, Stephen Penman, Doug Schroeder, and Gilbert Strang. It is also influenced by sponsored research projects conducted at Carnegie Mellon University (CMU) and by teaching courses titled "Business Language Analytics" (BLA) also at CMU. We are grateful for the inspirations gained from industry professionals, students, and research collaborators Leman Akoglu, Christos Faloutsos, Nan Li, Jane Pyo, and Gaoqing Zhang in these experiences.

"All accounting requires us to classify or partition various objects, events, and transactions into a finite set of mutually exclusive and collectively exhaustive categories. That is the essence of accounting."

Shyam Sunder
[Sunder, 2023], page 11.

1 Introduction

Double-entry bookkeeping remains a foundation of the financial infrastructure in any modern organization. As recognized long ago, double-entry bookkeeping is conceptually founded on the mathematical structures of the incidence matrix and directed graphs. By recording each transaction in at least two accounts, the double-entry system links all accounts of an entity together and in the process impose laws that govern the relation among inflows and outflows of accounts and their resulting balances. Such laws can be represented as properties of an incidence matrix or, equivalently, as properties of a directed graph. One such law is succinctly summarized by a famous theorem from Leonhard Euler, according to Professor John Fellingham [2018].[1] For a modern treatment of the historical importance of the double-entry bookkeeping, see Basu and Waymire [2008, 2021].

Beyond its mathematical elegance praised by the likes of Professor Arthur Cayley [Cayley, 1894], double-entry bookkeeping's sustained, unchallenged, and widespread use is evidence of its ultimate usefulness. Among accounting scholars, Professor Yuji Ijiri has long written extensively on its mathematical structure [Ijiri, 1967, 1975, 1993]. John Butterworth [1972] envisioned its use as an economic function in providing information: the structure can be thought of as part of an information source for an economic decision-making purpose. In Arya et al. [2000b], a statistical inference problem was formulated to assess the role of double-entry bookkeeping structure.[2]. See Fellingham [2017] for a more complete treatment. More recently, its graph/network structure has been utilized in pattern recognition and anomaly detection tasks leveraging modern graph-mining computational tools, see Liang et al. [2021] and Liang [2023].

In this position paper, we study these bookkeeping laws themselves and their immediate implications on the classifications in the records generated by double-entry bookkeeping. We note that this paper focuses on the *structure* of the classifications created by a double-entry recordkeeping rule. That is, the rest of the paper studies the laws imposed on the numerical values of each classification and how they evolve over time. A different research question, which we do not

[1]On page 1, Professor Fellingham states that "One way to describe a general result is a famous theorem from Leonhard Euler: The number of nodes minus one plus the number of enclosed regions equals the number of arcs (see, for example, Trudeau, 1993). Another way is to use accounting words: The number of T-accounts minus one plus the number of loops equals the number of journal entries. There is also a linear algebraic expression about the matrix underlying the system: The dimension of the row space plus the dimension of the null space equals the number of columns in the matrix."

[2]See related work in Arya et al. [2000c] and Arya et al. [2004] and its extension into quadruple-entry bookkeeping by Li et al. [2019].

pursue here, could focus on how the specific classification works in practice and how it should work. For example, for two potential transactions, what are the conditions under which they are booked in the same pair of accounts or in two different pairs? Here issues like uniformity and diversity come into play [Sunder, 1983].[3] Three laws are considered to describe a conventional double-entry bookkeeping system:

- **the balance law**: classifications of account balances obey *Assets = Liabilities + Equities* relation (with *Net Assets* replacing *Equities* for non-profit organizations).
- **the conservation law**: each classified account balance and flows overtime obeys *Ending balance = Beginning balance + inflows - outflows* relation.
- **the linearity law**: aggregated balances or flows obey additive relations such as *current assets = cash + marketable securities + accounts receivable + ...* relation.

The first part of the paper explicates the three laws as properties of the bookkeeping graph that uses nodes and edges between two nodes to represent the relations among account classifications. We choose to focus on these laws because they are universal to all conventional double-entry bookkeeping systems and the resulting classifications in the data they generate. Important classifications such as income statement versus balance sheets are included in the scope to the extent that balance sheet statement must obey the balance law (or the fundamental accounting identity), and income statement reflects a linear aggregation of income statement accounts. Under these three laws, additional structure or laws may be constructed to further delineate the distinction between, and the importance of, balance-sheet and income-statement classifications.[4]

The paper's second part is on developing the theoretically-consistent summary measures of the various classifications inherent in accounting records generated by a conventional double-entry bookkeeping system. Commonly used summary statistics from financial records include earnings and book-value. While these summary statistics have been demonstrated to be important and informative in practice, they are partial descriptions of the full set of records generated by firm activities. They are the end-result of the accounting data structure; they are not the direct measures of the *articulated* financial statement structure itself.

We seek to develop a set of comprehensive summary measures of the classification structure afforded by the double-entry bookkeeping laws. In the language of signal processing, we consider *the bookkeeping structure itself as part of the message* and to accomplish this task, we rely on an applied area in graph theory on measuring graph/network structure and complexity. Like many modern interdisciplinary sciences, it shares a theoretical foundation in *Information Theory* originated in Shannon [1948] in the 1940's. Three measures are proposed.

[3]Specifically, as Sunder [1983] pointed out much earlier, accounting rules (like revenue recognition or leases) can be "viewed as a scheme of classification. ... A scheme of classification can be based on two separate criteria: (1) those transactions which show certain similarities are placed in the same class; and (2) those which show dissimilarities are placed in the different classes. ... The fundamental problem arises from a conflict between these two criteria for classification." (page 103). See Sunder [1983] for a detailed discussion.

[4]Ijiri [1993] emphasizes the balance sheet versus income statement as a core idea behind double-entry bookkeeping.

- **Balance-sheet Node-entropy**: building on the classic work by Theil [1969], and Lev [1968, 1970] in which Shannon's entropy concept is applied to financial statement data, we extend the idea to a two-way joint-classification of the common-size balance-sheet where an asset classification is matched with a liabilities/equity classification. The matching procedure is inspired by the economic principle in the financial statement reformulation in work by Professor Stephen Penman [2013]. As a result, Balance-sheet Node-entropy deploys a joint-classification of an economic resources of both its use and its source. In short, the summary measure captures the classifications of underlying stock/node values in distinct account combinations.

- **Transaction Edge-entropy**: this second summary measure captures the classifications generated by linkages or transactions between account-pairs as opposed to each individual accounts themselves. The intuitive idea is that even if the relative balance of each account remains the same across time or entities, the directions and amounts of the inflows and outflows could have been quite different. The Edge-based entropy measure is designed to capture the classifications of these inflows and outflows of the entire bookkeeping records into a single-dimensional summary measure, also based on Shannon's entropy.

- **Bookkeeping Graph-entropy**: this third summary measure attempts to capture the entire bookkeeping graph classification of *BOTH* node- and edge-values. Exploiting the eigenvalue/vector properties of the Laplacian matrix constructed from the adjacency matrix of a directed and weighted graph, the von-Neumann-entropy-inspired graph entropy measures the complexity of the graph by using Shannon's entropy formula to summarize the eigenvalues that capture the various sub-graph structures.

As a whole, the paper constructs a graph-theoretic framework for bookkeeping records including aggregate financial statements. As such, it offers graph-theoretic constructs supporting empirical uses of financial statement data. For example, within the framework, each of the three measures can be viewed as an application of measuring graph structure using *Graph Information Functionals* of certain bookkeeping graph properties called *graph invariants*. Viewed this way, it provides a unified theoretical foundation underlying recent computational experiments conducted using related measures based on bookkeeping graphs constructed using publicly available financial statements issued by companies with their shares traded in the US stock exchanges. Our position is that quantifying the bookkeeping graph structure has the potential to offer insights into accounting information transmission, processing, and final end-uses.

2 Conventional Double-entry Bookkeeping

2.1 A Bookkeeping Example

The core operating device to implement double-entry bookkeeping is a journal entry. While modern-day journal entries are implemented as a part of a much more sophisticated management information system, the elementary building blocks of a journal entry remain conceptually unchanged since its invention in the middle-ages.[5] These conceptual building blocks are:

- entry date
- (at least two) accounts
- dollar amount assigned to each account
- debit and credit assignment to accounts

Figure 1 provides three examples of simple journal entries taught in countless introductory accounting classes throughout the world.

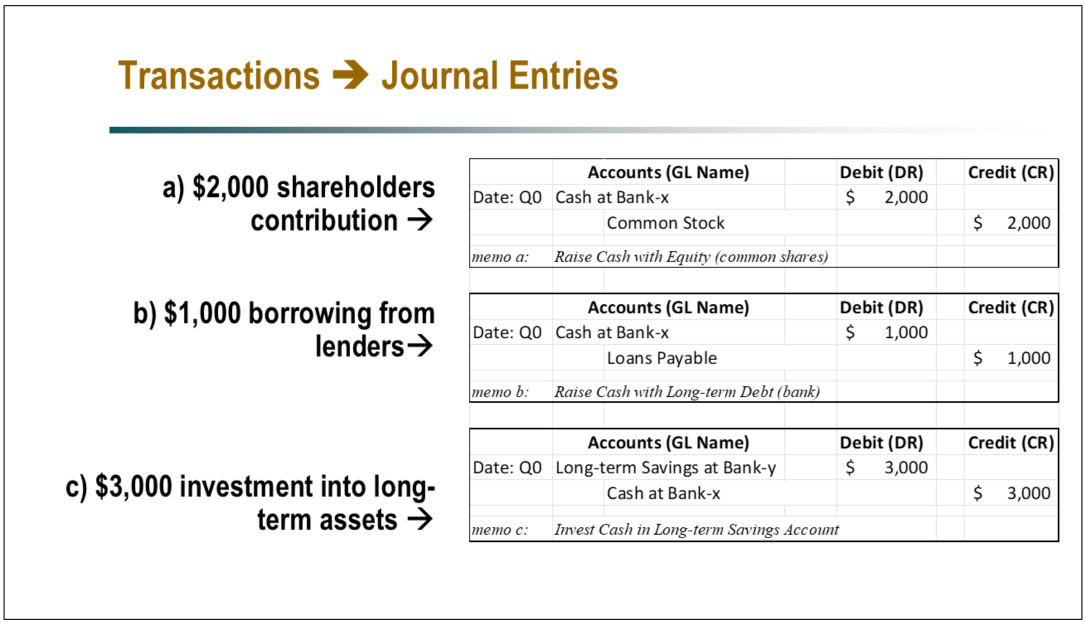

Figure 1: Simple Journal Entries Example

Another essential part of the conventional bookkeeping system is a Chart of Accounts (COA) that lists all the accounts available to be used in journal entries. In addition, the COA also provides account-level information about each account as to its membership in a hierarchy of account classification system. As a result, each account has certain classification features such as balance-sheet accounts (assets, liability, equity) or income-statement accounts (revenue, expenses, and gains or losses). Figure 2 provides an example of a COA. The bookkeeping process proceeds from journal entries to posting to general (and sub-) ledgers and at the end of the reporting period, building trial balances and producing financial statements.

[5] Over the last 500+ years, double-entry bookkeeping received deep interests from accountants and non-accountants alike. They include mathematicians, logicians, philosophers, physicists, sociologists, and of course, economists and accountants. One personal hero here is Professor Yuji Ijiri of Carnegie Mellon, who wrote extensively on the subject. For those interested, the conceptual distinction between so-called classificational versus causal journal entries are discussed in the classic Yuji Ijiri text [Ijiri, 1975]

Chart of Accounts (COA) at the General Ledger (GL) Level							
GL Account Number	GL Account Name	GL Short Name	BS/IS	Accout Type	GL FSLI Caption	FSLI Classification	Notes
1xx01	Cash at Bank-xxx	Cash	BS	Assets	Cash and Cash Equivalents	Current Assets	
1xx02	Accounts Receivable	AR	BS	Assets	Accounts Receivable	Current Assets	
1xx03	Inventory	INV	BS	Assets	Inventory	Current Assets	
1xx04	Long-term Savings Account at Bank-y	LTSecurities	BS	Assets	Long-term Investment	Non-current Assets	
1xx05	Property, Plants and Equipments - Net	NPPE	BS	Assets	Net PPE	Non-current Assets	
1xx06	Intangible Assets (including goodwill)	Intangibles	BS	Assets	Intangible Assets (including goodwill)	Non-current Assets	
2xx01	Accounts Payable	AP	BS	Liabilities	Accounts Payables	Current Liabilities	
2xx02	Wage payable	WP	BS	Liabilities	Other Short-term Payables	Current Liabilities	
2xx03	Deferred Income Taxes	DTL	BS	Liabilities	Deferred Tax Liabilities	Non-current Liabilities	
2xx04	Loans Payable	Loans	BS	Liabilities	Long-term Debt	Non-current Liabilities	
3xx01	Common Stock	CS	BS	Equity	Common Stock and APIC	Contributed Equity	
3xx02	Common Stock - ESO	CS-ESO	BS	Equity	Common Stock - ESO	Contributed Equity	
3xx03	Treasury Stock	TS	BS	Equity	Treasury Stock	Contributed Equity	
3xx04	Retained Earnings	RE	BS	Equity	Retained Earnings / Deficits	Earned Equity	
3xx05	AFS Unrealized Gains/Losses	AFS-URG/L	BS	Equity	Accumulated Other Comprehensive Income (AOCI)	AOCI	
4xx01	Interest Revenue	Rev-Int	I/S	Revenue	Interest Revenue	Non-operating Income	
4xx02	Sales Revenue	Sales	I/S	Revenue	Sales Revenue	Operating Revenue	
5xx01	Cost of Goods Sold	COGS	I/S	Expenses	Cost of Goods Sold	Operating Expenses	
5xx02	Sales, General and Admin	SGA	I/S	Expenses	Sales, General and Admin	Operating Expenses	
5xx03	Interest Expense	Int Exp	I/S	Expenses	Interest Expenses	Non-operating Income	
5xx04	Income Tax Expenses	IncTax	I/S	Expenses	Income Tax Expenses	Income Tax Expense	
6xx01	Common Dividend	DIV		DIV	Common Dividend	Common Dividends	

Figure 2: Chart of Accounts Example

2.2 Record-keeping Task: a definition

Now we construct a minimal environment in which record-keeping can be defined formally and double-entry bookkeeping, as a specialized record-keeping system, can be defined formally with features available for analysis.

Perception and Classification Drawing from Sunder [2023], we begin our formalization with the inherent need for perception in either analog or digitalized form. Classification emerges from converting the perceptions received from sensors into discrete, digitalized civilized construct. By converting what an entity perceives around itself over time into its recorded financial history, according to Sunder [2023], modern accounting at its core, maintains the root function of classification inherent in the perceptual development in the long history of human and societal endeavors.

Entity and its Recorded History We assume there exists a well-defined entity. Following Coase [1937], Simon [1952] and Sunder [1996], we posit that a set of contracts can be subsumed as a well-defined *entity*. The entity concept[6] is a legal fiction that allows us to focus our attention

[6] The notion of entity is ubiquitous in studies of accounting theory. It appears in the report of the American Institute's special committee on research (1958): "Postulate A-3: Entities (including identification of the entity) Economic activity is carried through specific units or entities. Any report on the activity must identify clearly the particular unit of entity involved" (see Sewell Bray [1966] p.39-40). In Ijiri [1967], the entity is more precisely defined as an object vested with control: "accounting does not exist for a simple collection of individuals such as a mob unless such a collection is considered to constitute an entity" (p.69). The reference to the original set of contracts is made explicit is Mattessich [1995]: "A-5. *Economic Entity:* There is some economic entity (...) represented by a specific accounting system. Such entity consists of economic subjects and economic objects and can enter into contracts."

away from the complexity of the original environment toward a single subject of record-keeping: an artificial "firm."

While we consider the subject of record-keeping as the suitably reduced accounting entity, we consider the objects of record-keeping those actions and outcomes that the entity perceives or experiences during the course of its existence. For example, we consider an entity which is fully aware of its own well-defined control over resources and actions. How these resources are used and their consequences, such as past, current, or future transactions, are candidate items to be recorded and accounted for.[7] For this entity, we formalize its recorded history using the following definitions.

Let time be discrete and be indexed by $t \in \{0, 1, 2, ..., T\}$

Definition 1 Entity History and Quantification *An entity $\mathcal{E}$'s (unrecorded) history $\mathcal{H}$ is described by a sequence of events denoted by z_t at time t where each z_t allows at least one quantification which we denote by $\chi_t \in (R)$ generated by a quantification $\mathcal{Q}$ function: $\mathcal{Q} : Z_t \to (R)$*

$$z_t \in Z_t \equiv \{z_t^1, z_t^2, z_t^3, ...\} \quad and \quad \chi_t = \mathcal{Q}(z_t) \in (R) \tag{1}$$

That $\mathcal{H}$ is considered *unrecorded* can be motivated by the external technical constraint or internal human processing capacity that prevent each event z_t in $\mathcal{H}$ being costlessly memorized and processed by the entity. Since the purpose is to focus on record-keeping, we refrain from modeling the precise function that produces these events (as they would require recording as well). Standard information economics approach would begin a state-act-outcome formulation [Marschak and Miyasawa, 1968].[8] For this reason, the term *event* is used instead of a set of distinct terms such as *actions* (deliberate choice), *states of nature* (a choice by nature), and *outcome* (a deterministic function of actions and states).

Now we formalize the recording of the (unrecorded) entity history. We do not formalize a demand for this recording activities and treat it as universally accepted that such recording is necessary so the focus is on the structure of such recording (e.g., single-entry or double-entry, etc.). We begin with the elements of any recording: the alphabet for labeling and monetary units of numerical counting.

Definition 2 Alphabet and Record-keeping Rule *An Alphabet $\mathcal{L}$ is set of a finite number of alphanumerical strings with finite length with a generic element labeled as l:*

$$l \in \mathcal{L} \equiv \{l^1, l^2, l^3, ..., l^m\} \tag{3}$$

[7] Ijiri [1967] identified the importance of the control axiom and it is an absolutely necessary element to build a reasonable accounting theory.

[8] In that case, the entity would have control over a sequence of actions available from which the entity must choose one action denoted a

$$a \in A \equiv \{a^1, a^2, a^3, ...\} \tag{2}$$

and combined with a state-of-the-world $s \in S$, a consequence function $C : S \times A \to X$ would convert its past actions a into outcomes. Then by endowing the entity with a preference over outcomes, a model of decision making allows an analyst to recover a sequence of the equilibrium outcome based on maximizing the suitably constructed utility function implied by the entity's preference, as shown formally in Marschak and Miyasawa [1968].

A Record-keeping Rule $\mathcal{R}$ is a function: $\mathcal{R}$ which assigns a label $l \in \mathcal{L}$ to an event z_t:

$$\mathcal{R} : Z_t \to \mathcal{L} \tag{4}$$

We call the assigned label R_t of event z_t, i.e., $R_t = \mathcal{R}(z_t)$ as the record of z_t under rule $\mathcal{R}$.

Applying Record-keeping Rule $\mathcal{R}$ to an (unrecorded) history $\mathcal{H}$ generates a recorded history $\mathcal{RH}$ of the entity. The classification role of record-keeping becomes immediate: two possible events z_t' and z_t'' are classified into the same category if and only if they are assigned the same label $l^k, k \in M \equiv \{1, 2, 3, ..., m\}$.

Definition 3 Recorded History of an Entity *An entity's recorded history $\mathcal{RH}$ of its history $\mathcal{H}$ is the sequence of records R_t's of individual events z_t in the history $\mathcal{H}$:*

$$\mathcal{RH}\left(\{z_1, z_2, z_3, ...\}|\mathcal{R}\right) \equiv \{R_1, R_2, R_3, ...\} \tag{5}$$

Among all possible record-keeping rules, we focus on a class of rules that uses an *ordered-pair* of labels $\{l_i, l_j\}$ along with a numerical monetary count q_t to record each and every individual event z_t.

Chart of Accounts and Double-entry-bookkeeping

Definition 4 Chart of Accounts *A Chart of Accounts (COA) $\mathcal{COA}$ is set of a finite number of alphanumerical account names with finite length with a generic element labeled as acct:*

$$acct \in \mathcal{COA} \equiv \{acct^1, acct^2, acct^3, ..., acct^m\} \tag{6}$$

Definition 5 Double-Entry Record-keeping Rules *A Double-Entry Record-keeping Rule using a Chart of Accounts is a Record-keeping Rule $\mathcal{R}$ that uses two distinct elements in $\mathcal{COA}$ as an ordered-pair and a monetary count $q_t(\chi_t)$ as the record of every event z_t:*

$$\mathcal{R} : Z_t \to \mathcal{COA}^2 \times \mathbb{R} \tag{7}$$

Using a Double-entry Record-keeping Rule, a record of an event z_t has the following content:

$$R_t = \mathcal{R}(z_t) = \{acct_i, acct_j, q_t\}$$

Compound journal entries are not explicitly treated here but are implicitly treated by the abstraction that all such entries can be reconstructed with a sequence of simple journal entries in definition (7).

Recall, we noted earlier that this paper focuses on the *structure* of the classifications created by a double-entry record-keeping rule. That is, the rest of the paper studies the laws imposed on the numerical values of $acct_i$'s and how each $acct_i$ evolves over time as the entity experiences more events z_t's. We do not investigate the function of $\mathcal{R}(\cdot)$ itself other than its double-entry structure.[9] Instead, we study the mathematical restrictions on the records $(\{R(z_1), R(z_2), R(z_3), ..., R(z_n)\})$ themselves, under whichever specific functional form (US GAAP, IFRS, etc.) of $\mathcal{R}(\cdot)$ within the class defined by Definition (7).

3 Three Laws of Double-entry Bookkeeping

This section describes the laws that govern the properties of an entity's recorded (financial) history following the Double-entry Record-keeping Rule defined in definition (7). Specifically, these three laws dictate all quantities assigned to each $acct_i$ must sum up, govern how each $acct_i$'s balance evolves over time, and how aggregate accounts are constructed linearly.

3.1 Three Laws in words and in identities

3.1.1 Law #1: the balance law

In words, this law is the most familiar to all accountants and accounting students. There are many ways this law manifests with the most well-known version as the Balance-sheet Identity:

$$Assets = Liabilities + Equity.$$

This law imposes a restriction on the stock values (*at a point in time*) in the data generated by the double-entry bookkeeping system.

3.1.2 Law #2: the conservation law

In words, this law is also very familiar to all accountants and accounting students. For each account, this law simply dictates that the ending amount must evolve from the beginning amount with all inflows to the account added and all outflows from the account subtracted:

$$\textit{Ending balance} = \textit{Beginning balance} + \textit{all inflows} - \textit{all outflows}.$$

Working at every T-account level, this law imposes a restriction on the stocks (*at points of time*) and flows (*over a period of time*) values in data generated by the double-entry bookkeeping system.

[9]For example, a property of the $\mathcal{R}$ rule can be recognition: for some events z_t, there is no recognition: $\mathcal{R} = \emptyset$ such as how US GAAP treats certain unrealized gains and losses for some securities. Generally, we may be interested in the question that for two potential events z'_t and z''_t, what are the conditions under which $R(z'_t) = R(z''_t)$ or $R(z'_t) \neq R(z''_t)$? Two different sales contracts may be recognized differently regarding revenue recognition under some conditions.

3.1.3 Law #3: the linearity law

In words, this law is, again, familiar to all accountants and accounting students. Accounts can be aggregated into desirable aggregates by simply combining their stocks and flows. For example, for stock values, we can combine *cash*, *accounts receivable*, and a few others into a group with a higher-level account labeled *Current Assets*.

$$\textit{Current Assets Balance} = \textit{Cash balance} + \textit{Accounts Receivable Balance} + \ldots$$

For flow values, we can combine revenue and expenses, plus a few others, into a new account labeled *Income (Summary)*:

$$\textit{Net Income} = \textit{Revenue - Expenses} + \textit{Gains} \ldots$$

or combine cash collected from customers, payments to vendors, plus a few others into a new account labeled *Cash Flows from Operating Activities*.

$$\begin{aligned}\textit{Cash Flows from Operating Activities} = \;& \textit{Cash collected from customers}\\ & - \textit{Cash payments to vendors} +/- \ldots\end{aligned}$$

This law imposes a restriction on the stock and flow values in data generated by the double-entry bookkeeping system whenever some account aggregation is used.

3.2 Bookkeeping in Weighted Directed Graph Form

It is well-known that Double-entry bookkeeping has a directed graph (or network) representation, as shown in work as early as Charnes et al. [1963] and Mattessich [1964]. We call this representation of double-entry bookkeeping the Directed Weighted Graph (or network). In this graph or network, a node represents a T-account and an edge represents an entry linking two accounts/nodes. A third well-known representation of double-entry bookkeeping, the incidence matrix form, is contained in Appendix 6.3.

Figure 3 offers a simple illustration. Notice within this example, journal entries on page 4 are represented by the graph in Figure 3 consisting of four nodes representing a subset of accounts listed in the COA in Figure 2 (nodes *Cash* and *Common Stock*, etc.) and three edges (denoting transactions -a, -b, and -c). Conceptually, the graph/network representations of the same double-entry bookkeeping system are mere representations. That is, they do not add or subtract any information from what is already contained in the journal entries themselves. Any value will need to come from using specific representations when facing specific tasks in practice. In the research reported in this document, it turns out, graph representation has led to summary measures of classification that are compatible with modern information science tools. Below, we formalize the graph/network structure before using information theory to construct summary measures of bookkeeping classification.

Let bookkeeping graph be denoted $G = \{V, E, W, X\}$, where

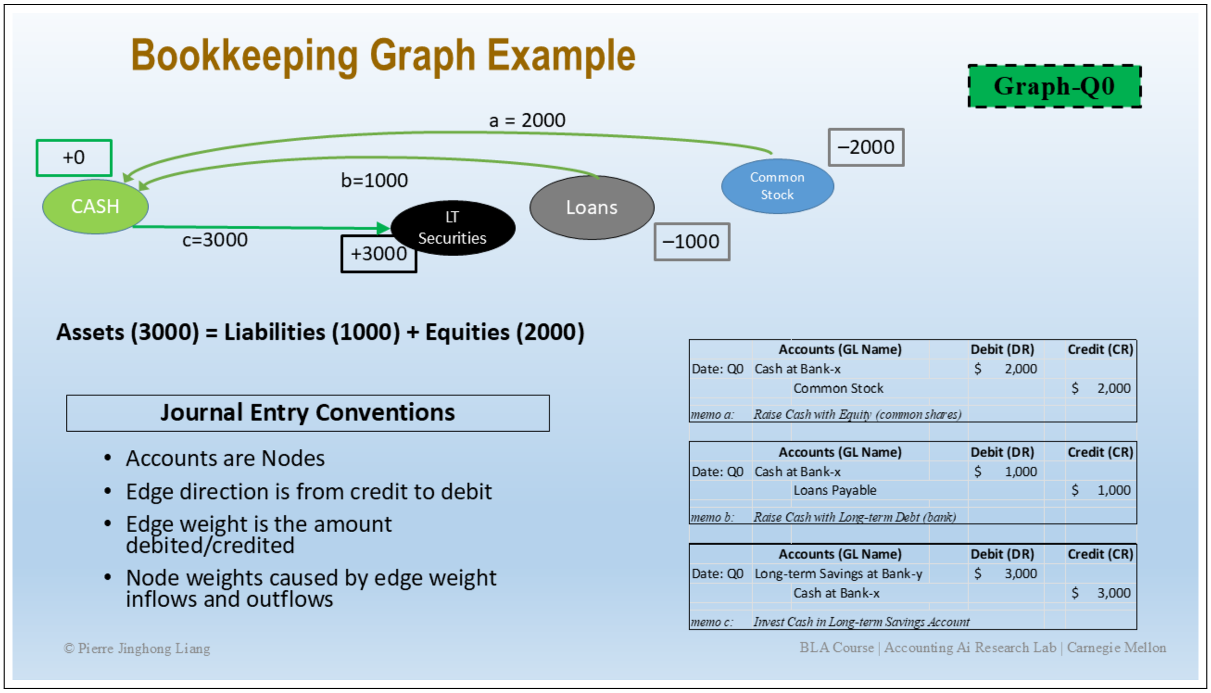

Figure 3: Directed Weighted Bookkeeping Graph Example

- V is the set of m nodes or vertices representing distinct accounts. Let $M = \{1, 2, ..., m\}$;

$$V = \{v_1, v_2, ...v_m\}$$

For bookkeeping graphs, the set of vertices is simply the Chart of Accounts: $V = \mathcal{COA}$.

- E is the set of edges, represented by an ordered-pair of nodes/vertices $e_{ij} \equiv \{v_i, v_j\}$, an edge from node v_i to node v_j, The set E collects all the edges that are present in the graph.

$$E = \{e_{ij} | i, j \in M, \text{ nodes } i \text{ and } j \text{ are connected from } i \text{ to } j\}$$

An edge between two accounts means a journal entry is indeed recorded between the two accounts following the convention of debit being an inflow to account-j and credit being an outflow from account-i. For example, an event z_t interpreted "*raising cash by issuing debt*" transaction would be recorded using the ordered pair $\{acct_i = Loans\ Payable,\ accout_j = Cash\}$. We denote the cardinality of the set E to be n or $|E| = n$.

- W is a $m \times m$ non-negative-valued Adjacency Matrix where the ij-th entry corresponds to the (dollar) weight of edge-ij if the edge exists. If the edge does not exist (i.e., $e_{kl} \notin E$), then $w_{kl} = 0$. The Adjacency Matrix for the bookkeeping graph in Figure 3 is given below:

$$W = \begin{array}{c} \\ Cash \\ LT{-}Sec \\ Loans \\ ... \\ Equity \end{array} \begin{array}{c} \begin{array}{ccccc} Cash & LT{-}Sec & Loans & ... & Equity \end{array} \\ \left[\begin{array}{ccccc} 0 & 3000 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ 1000 & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & 0 & \dots \\ 2000 & 0 & 0 & \dots & 0 \end{array}\right] \end{array}$$

For bookkeeping graphs, the weight of an edge representing the record of event z_t is the quantitative (monetary) value q_t: $w_{ij} = q_t$ such that $\mathcal{R}(z_t) = \{acct_i, acct_j, qt\}$. For example, $w_{31} = \$1000$ encodes the journal entry b crediting the *Loans Payable* account and debiting *Cash* totals to \$1000. Notice, this matrix always has zeros in the diagonal cells (i.e., no self-edges or no debiting and crediting of the same account) and generally asymmetric: w_{ij} need not be the same as w_{ji}.

- $X \in \mathbb{R}^m$ is a $m \times 1$ real-value vector where the i-th entry corresponds to the (dollar) weight of node-i. So, using superscript T for the transpose operation, the transpose of X is

$$X^T = \{x_1, x_2, ..., x_m\}$$

Bookkeeping graphs use x_i to represent the *change* in the account balance. For example, $x_i = -\$412$ means that change in account-i is a (net) credit of \$412

3.3 Three Laws Expressed in Graphs

The balance law This law manifests itself in the key graph property: all node-weights in vector X add to zero. For $i \in M$ and a given G:

$$\sum_{i=1}^{m} x_i = 0$$

Following the convention of using negative node-weights to represent positive (credit) values in liabilities and equities accounts, for a three-account system, the law would lead to:

Assets - Liabilities - Equity = 0.

The conservation law This law manifests itself in the another key property contained in the directed-weighted graph: node-weight is equal to the sum of weights of all incoming edges minus the sum of weights of all outgoing edges: As a result, for each account, the change in account balance (the motion over time) become, for $i, j, k \in M$ and a given G:

$$X_i^{End} - X_i^{Beg} \equiv x_i \ = \sum_{j=1}^{m} w_{ji} - \sum_{k=1}^{m} w_{ik}$$

The linearity law This law manifests itself in an aggregation operation on the graph: to aggregate the accounts into a new account, a new node is created by merging a few existing nodes. As a result, for the new combined account, evolution of the balance becomes:

$$\underset{1\times m}{g^T} \cdot \underset{m\times 1}{X^{end}} = \underset{1\times m}{g^T} \cdot \underset{m\times 1}{X^{beg}} + \underset{1\times m}{1^T \cdot} \underset{m\times m}{W \cdot} \underset{m\times 1}{g} - \underset{1\times m}{g^T} \cdot \underset{m\times m}{W \cdot} \underset{m\times 1}{1}$$

For example, to aggregate all *current assets* accounts, we use a $m \times 1$ vector, denoted g of all zeros except 1's in all positions where individual *current assets* accounts reside. The $m \times 1$ vector 1 represents a vectors of all 1's in each cell.

An extended example based on the three-transaction Figure 3, with details in the Appendix 6.1, shows under suitable aggregations of *net operating assets* nodes and *operating revenue and expense* edges, the conservation law relation of the *Net Operating Assets* node is equivalent to the Free Cash Flow identity generated by the Reformulated Financial Statements in Penman [2013] (page 244):

$$\textit{Change in Net Operating Assets} = \textit{Operating Income} - \textit{Free Cash Flows}$$

where the weight of the *Net Operating Assets* node is the sum of the weights of all operating assets and all operating liabilities nodes. And similarly, the weight of the *Net Operating income* edge is the sum of the directed weights of all operating revenue and all operating expense edges.

4 Quantifying Bookkeeping Graph Classification

Working independently and together, three bookkeeping laws generate a set of accounting records with a double-entry classification system which allows information retrieval. In fact, the most familiar retrieval is a set of a few summary statistics, among them the most prominent being 1) an income figure such as net income (or net income before extra-ordinary items, etc.) and 2) a balance-sheet figure such as total assets or total shareholders' equity (or total equity excluding non-controlling interests).

While these summary statistics are important and prove quite informative when linked to evidence from outside the bookkeeping system such as stock price, trading volume (around announcement dates), they are rather partial descriptions of the full set of records generated by firm activities. In the framework of double-entry bookkeeping graphs under the three laws, these figures are summaries of certain aspects of larger graph. These summary figures are the end-result of the bookkeeping graph structure, they are not the direct measures of the *articulated* financial statement structure itself.

This section develops systematic summary measures of the classification structure afforded by the double-entry bookkeeping laws. In our quantification exercise, we consider, at the general level, *the weighted directed graph itself as the message*. Our current goal is to develop summaries that capture the structure of graph somehow. To accomplish this task, we appeal to information theory and its application to the applied problem of measuring information content in graphs. Shortly after

Shannon's path breaking work, the entropy concept was quickly adapted to measure graph structure and complexity in the graph analysis field, among the growing list of fields information theory overlaps: physics, probability, mathematics, complexity, and even financial portfolio (Kelly bets). Not surprisingly, it was applied to accounting by Theil [1969] and Lev [1968, 1970], although not fully incorporating double entry bookkeeping. In this paper, we reapply the basic idea of entropy but to the bookkeeping graph in a way that formally treats the graph as the full signal/message. Considering specific features of bookkeeping graph, we develop a sequence of parameterized graph measures that attempt to measure the information contained in the graph classification. We take three broad steps, moving from node-classification to edge-classification, and finally to graph classification (combining nodes and edges). For each, we propose a set of graph information quantity measures of the data which obey the three bookkeeping laws. They are

- Balance-sheet Node-entropy
- Transaction Edge-entropy
- Bookkeeping Graph-entropy

Common features of these measures are that they are (1) summary (thus uni-dimensional) measures of a set of bookkeeping figures of a higher-dimension; (2) internal measure of information contained in the bookkeeping classification structure (without referring to quantification outside the records); (3) all based on information theory originally developed by Shannon [1948].

4.1 Graph Theory Preliminaries

In order to formally derive the summary measures, we first introduce a few necessary terms used in graph theory. Reinhard [2017] is an excellent introduction to graph theory.

Graph isomorphism and Graph Invariants Informally, two graphs are *isomorphic* with equal number of nodes if a relabeling (or mapping) of the node identities of one graph generates the other graph (i.e., preserving the adjacent nodes) and vice-versa. For example, the three graphs in Figure 4 are isomorphic. Notice between graphs A and B, node-labels are also preserved (in which case they are also automorphic) while isomorphism between graphs A and C requires a node-relabeling shown in the table within the figure.

A *Graph Invariant* is a mapping that takes graph as an argument and assigns equal values to all isomorphic graphs.

Definition 6 *Let $\mathcal{G}^{iso}(G)$ be the set of all graphs isomorphic of graph G. A Graph Invariant $s(G) \in C$ of graph G is a mapping $S : \mathcal{G}^{iso}(G) \to C$ such that*

$$s(G') = s(G'') \quad \forall \quad G', G'' \in \mathcal{G}^{iso}(G)\}$$

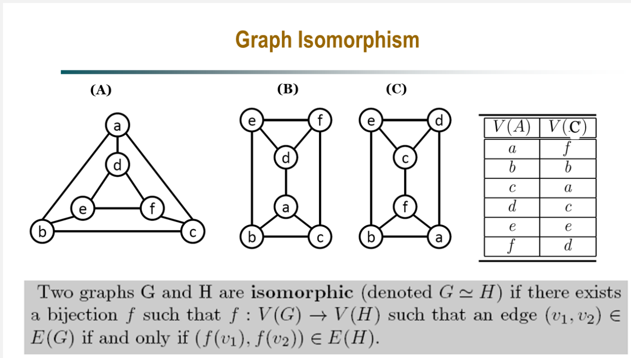

Figure 4: Three Isomorphic Graphs

The total number of vertices and the total number of edges of a graph are two simple graph invariants. The greatest number of pairwise adjacent vertices is another.

Graph Information Functional and Graph Entropy Measures

Definition 7 *A Graph Information Functional f is a function that assigns an element from an abstract set S of graph invariants to a positive real number: $f : S \to \mathbb{R}_+$*

Now we are ready to define a class of Graph Entropy measures using the idea of *Graph Information Functionals*.

Definition 8 ***Graph Entropy*** *Let abstract set S contains K distinct elements of graph invariants for a given graph G. The Entropy H of Graph G based on functional f is given by:*

$$H(G|f) = -\sum_{k=1}^{K} \frac{f(s_k)}{\sum_{k=1}^{K} f(s_k)} log_2 \frac{f(s_k)}{\sum_{k=1}^{K} f(s_k)} \tag{8}$$

Mowshowitz [1968] formalized the application of Shannon's entropy concept to finite graphs using the idea of graph invariants which generalizes Shannon's equation to cases beyond probabilities to any system structure.[10] See Dehmer [2008] for an expanded discussion of a large class of graph information functionals. See Appendix 7.1 for the original definition of Shannon's Entropy, the associated relative entropy (i.e., K-L Divergence).

Three Specific Functionals for Bookkeeping Graphs

1. Let the set of graph invariants be the set of node-weights: for a given graph, let $S = V$ and the graph information functional be the absolute values of the node-weights:

   $$f(v_i) = |x_i|, \forall i \in N.$$

   Under this graph functional, we call the resulting graph entropy: **Balance-sheet Node-entropy**.

[10] Consider any system having N elements partitions into k classes according to an equivalence criterion α (that is, α determines with two of the N elements are equivalent). With this classification scheme, let N_i be the number of elements classified by α as class-i, $i \in \{1, 2, ..., k\}$. Then the entropy equation applied to the fractions $p_i = N_i/N$'s becomes the information content of the α-classification system. See page 160-161 of Bonchev and Rouvray [2003] for more details.

2. Let the set of graph invariants be the set of edge-weights: for a given graph, let $S = E$ and the graph information functional be the edge-weights:

$$f(e_{ij}) = w_{ij}, \forall i, j \in M.$$

So, the total number of non-zero functional values equal to n. Under this graph functional, we call the resulting graph entropy: **Transaction Edge-entropy**.

3. Let the set of graph invariants $S = \Lambda$ where Λ is the set of positive eigenvalue of the normalized graph Laplacian matrix (to be defined later). The graph information functional be the eigenvalues themselves:

$$f(\lambda_i) = \lambda_i, \forall\ i \in \{1, 2, 3.., |\Lambda|\}.$$

Under this graph functional, we call the resulting graph entropy: **Bookkeeping Graph-entropy**.

Now we operationalize these graph-information-functionals-based entropy measures in the context of bookkeeping graphs created using public financial statements.

4.2 Summarizing Nodes Classification

Now we connect our Balance-sheet Node-entropy measure to classic work by Professor Henri Theil [1969]. In his pioneering work in applying Shannon's Information Theory to financial statement analysis using relative entropy, a measure derived from Shannon's entropy to estimate the new information contained in the current balance sheet. In the original application by Theil [1969] use common-sized balance figures, treated as probabilities, to calculate the K-L divergence from one year-end to the following year-end. Professor Theil terms his measure *Expected Information* of the message in the current balance sheet from the previous balance sheet, with the definition below (see Theil [1969], page 461):

Definition 9 ***Expected Information [due to Theil [1969]]*** *Let* $p = \{p_1, p_2, ..., p_n\}$ *and* $q = \{q_1, q_2, ..., q_n\}$ *be the prior and the posterior probabilities of* $x = \{x_1, x_2, ..., x_n\}$*. The* Expected Information *of the message which transforms the prior to posterior is defined as*

$$I(q : p) = \sum_{i=1}^{n} q(x) log \frac{q(x)}{p(x)} = E_q log \frac{q(X)}{p(X)}$$

According to Theil [1969], one can interpret adapting the entropy measures to financial statement setting naturally. For example, Theil's *Expected Information* measure, equivalent to the K-L divergence between the current balance sheet fractions to previous fractions, tells "how far we have to travel" from the fractions of the first year to those of the second. (page. 463).

Theil [1969]'s work was followed by accounting scholars such as Lev [1968], Lev [1970], and Itami [1977]. The deeper connection between entropy and accounting information is systematically explored in a statement of *Fundamental Theorem of Accounting* by Fellingham [2017] and Fellingham et al. [2019]. A recent paper by Li et al. [2024] reports statistical properties of the Theil [1969]'s measure (applied to assets balances) based on large sample evidence. The measure was shown to behave well in empirical applications including analyst behavior and executive compensation.

In our framework, Theil [1969] measure is a first-step in summarizing bookkeeping classification because it only concerns a list of asset accounts, or node-values in graph terms. Next we develop a general version of Theil [1969] measure by exploiting the economic meaning of the balance law.

Balance-sheet Node-entropy In our application, we exploit the accounting meaning of the balance law. The common interpretation of a classified balance sheet is that the assets represent *uses or deployment* of entity's scarce resources and the liabilities and equities represent *funding sources* of the same resources. Thus, the uses and sources must be equal at all times. Classified balance sheets add additional information by listing various categories of uses and sources. Among many dimensions to categorize (or classify), liquidity of assets and maturities of liabilities and equities are very common. Thus, we have the familiar fundamental accounting identity in the classified manner:

$$\begin{aligned} \textit{Short-term Assets} + \textit{Long-term Assets} &= \textit{Short-term Liabilities} \\ &+ \textit{Long-term Liabilities} + \textit{Equities}. \end{aligned}$$

While one can unambiguously interpret the identity as that *all* assets are funded by *all* liabilities and equities, it is less clear that the balance sheet provides unambiguous use-fund pairs information at the classified account level. Consider the numerical example in Figure 1. While one can say that $3,000 long-term security is entirely funded by $1,000 loans and $2,000 equity, things would be different if transaction-c is changed to $2,700, leaving $300 in the cash account. We may entertain a possible two-way classification of the balance-sheet identity. Consider the following two potential two-way classifications:

Comparison Box: Two-way Classification Approaches

	Loans	Equity	Total
Cash	100	200	300
LT-Security	900	1,800	2,700
Total	1,000	2,000	$3,000

Table 1: Proportional

	Loans	Equity	Total
Cash	300	0	300
LT-Security	700	2,000	2,700
Total	1,000	2,000	$3,000

Table 2: Liquidity-Maturity Match

Our generalization of Theil [1969]'s *Expected Information* measure is to apply the entropy concept (thus the K-L divergence measure) to the *common-sized version* of the two-way classification table. Notice entropy measure on Asset-only classification (or Liability-Equity only classification) is just a special case of the two-way classification case.

For a given balance sheet (or a set of nodes that obey bookkeeping law #1), define a two-way

Node-Value	Liab-1	Liab-2	Liab-3	Liab-4	Equity	Total
Asset-1						Asset-1 sub-total
Asset-2						Asset-2 sub-total
Asset-3						Asset-3 sub-total
Asset-4						Asset-4 sub-total
Total	Liab-1 sub-total	Liab-2 sub-total	Liab-3 sub-total	Liab-4 sub-total	Equity sub-total	TA or Total L+E

Table 3: Two-way classification of Balance Sheet Example

classification as a matrix, denoted Classified Balance-Sheet matrix *CBSM*, where each row's entries add up to one and only one line-item in the asset classification and each column's entries add up to one and only one line-item in the liabilities/equities classification. Table 3 shows a four-asset-four-liabilities-one-equity example. Since all the row totals and all the column totals are 1, common-sized version of the two-way classification of Balance-sheet corresponds to a probability distribution. Denote Common-sized version of *CBSM* as a joint probability distribution p for the current year and q for the past year. Now we can define Balance-sheet Node-entropy and its associated Relative Entropy.

Definition 10 Balance-sheet Node-entropy and Relative Entropy *Let p (resp. q) be the joint distribution derived from common-sized two-way classified balance sheet of the current (resp. past) year. The **Balance-sheet Node-entropy** is defined as Shannon Entropy of joint distribution p and the **Balance-sheet Node-Relative-entropy** is defined as the K-L divergence of joint distribution p from distribution q.*

Notice Balance-sheet Node-entropy and Relative-entropy subsume the Theil [1969]'s Balance sheet *Expected Information* and the Accounting Entropy measure studied in Li et al. [2024] as simply reducing the joint distribution to the marginal distribution (of assets only or liabilities/equities only). Following Cover and Thomas [2006], this entropy measure is the only measure that satisfy the three axioms. The statement and proof is given in Li et al. [2024] and reproduced in Section 7.2 in Appendix B.

The open question is how to construct the joint distribution and how the construction may affect the property of the resulting entropy and relative entropy measures. While there may be many approaches to proceed, we consider two specific approaches and use them as examples to demonstrate the flexibility and adaptivity of the entropy measure. The two approaches are:

- **Proportional Approach**: as illustrated in Table 1, this approach assigns $dollars of each asset line-item (row) into each liabilities/equities categories (columns) in relative proportion to the liabilities/equities total. Since all information is already contained in the marginal, it can be shown that the entropy measure is monotonic to the sum of the entropy of marginal distribution of assets-only and the entropy of that of the liabilities/equities.

- **Economic Matching Approach**: inspired by the classification described in Penman [2013], this approach attempts to capture potential relation among firm's investment (i.e., use of assets) and financing (i.e., source of funding) decisions. The idea is that *operating assets* (e.g., inventory) are funded by *operating liabilities* (e.g., accounts payable) first and *financial assets* (e.g., securities) are funded by *financial obligations* (e.g., bank borrowing and debts) first, as illustrated in Table 2.

The updated version of Li et al. [2024] reports the properties and empirical results of the Economic-Matching approach, in addition to the marginal approach.

4.3 Summarizing Edges Classification

The next summary measure attempts to capture classification information contained in **edges** in the bookkeeping graph. Notice the node-value distribution X contains node classification. While the node-values are constructed using inflows and outflows to each node (by #2 the conservation law), the node-entropy measure is nevertheless not completely expressive with respect to edges. That is, for many different combinations of edge-weights/values in the adjacency matrix, the node-weight/value distribution would remain the same and thus the node-entropy measures would remain the same even though the bookkeeping graphs have been altered. Driven by the existence of the null space of the incidence matrix representing the double-entry bookkeeping structure, the incomplete expressivity of node-weights of the underlying edge-weights is well-known and well-studied, see Arya et al. [2000b] for a detailed discussion.

To construct the Transaction Edge-entropy, we normalize the adjacency matrix W. In this all edge-weight entries into the matrix remain positive and the direction of each edge is represented by the *positioning* of cell. Recall, debiting account-j and crediting account-i by \$10 is accomplished by setting $w_{ij} = \$10$, the value of the cell located at the i^{th} row and the j^{th} column of the adjacency matrix W. Now we define the common-sized Adjacency Matrix W^{cs} defined where each cell entry is common-sized as follows:

$$w_{ij}^{cs} = \frac{w_{ij}}{\sum_i \sum_j w_{ij}}$$

Since all edge-weights are positive, all the common-sized edge values are positive and less than one and sum up to one. So, the collection of the common-sized weights defines a suitable probability distribution p of the edge-weights for the current year. And let q be the same for the past year. Now we can define Transaction Edge-entropy and its associated Relative Entropy.

Definition 11 Transaction Edge-entropy and Relative Entropy *Let p (resp. q) be the joint distribution derived from common-sized adjacency matrix of the current (resp. past) year. The **Transaction Edge-entropy** is defined as Shannon Entropy of joint distribution p and the*

Transaction Edge-Relative-entropy *is defined as the K-L divergence of joint distribution p from distribution q.*

Edge-Value	acct-1	acct-2	acct-3	...	acct-m	Total
acct-1	0	w_{12}	w_{13}	...	w_{1m}	account-1 total outflows
acct-2	w_{21}	0	w_{23}	...	w_{2m}	account-2 total outflows
acct-3	w_{31}	w_{32}	0	...	w_{3m}	account-3 total outflows
...	...	...	...	0	...	... total outflows
acct-m	w_{m1}	w_{m2}	w_{m3}	...	0	Asset-m total outflows
Total	acct-1 inflows	acct-2 inflows	acct-3 inflows	... inflows	acct-m inflows	Total in- or out-flows

Table 4: Transaction Edge-entropy using Adjacency Matrix Encoding

Unlike Balance-sheet Node-entropy, the axiomatic foundation for this entropy measure is not a straightforward application of Cover and Thomas [2006]. Adapting the underlying axiomatic foundation to our bookkeeping graph edges case, the statement and proof are given in Section 7.3 in Appendix B.

Small Sample Use-case: a clustering experiment Consider 9 companies with their balance sheet, income statement, statement of cash flows, and comprehensive income statement publicly available. Using these statements, one can construct an aggregated bookkeeping graph such as Figure 5 below.

These 9 directed weighted graphs are then converted into their corresponding adjacency matrices and pair-wise distance between these matrices are calculated. Based on the pair-wise 36 (9 x 8 / 2) distances, a simple clustering exercise is performed to group firms with closer distance together. The result is shown below in this knowledge graph labeled as Figure 6 below. As intuitive two retailers have graphs that are closer; so are those among the three technology companies; and among the four banks. Details of this small sample clustering experiment are provided in Appendix 6.2.

A recent work-in-process by a team at Carnegie Mellon reports findings from a large-sample study of empirical graphs such as Figure 5. These findings include the properties of the graphs and the large sample properties of the Transaction Edge-entropy and its associated K-L divergence measures. Results on applications to capital market return-based research questions are also reported.

4.4 Summarizing Graph Classification

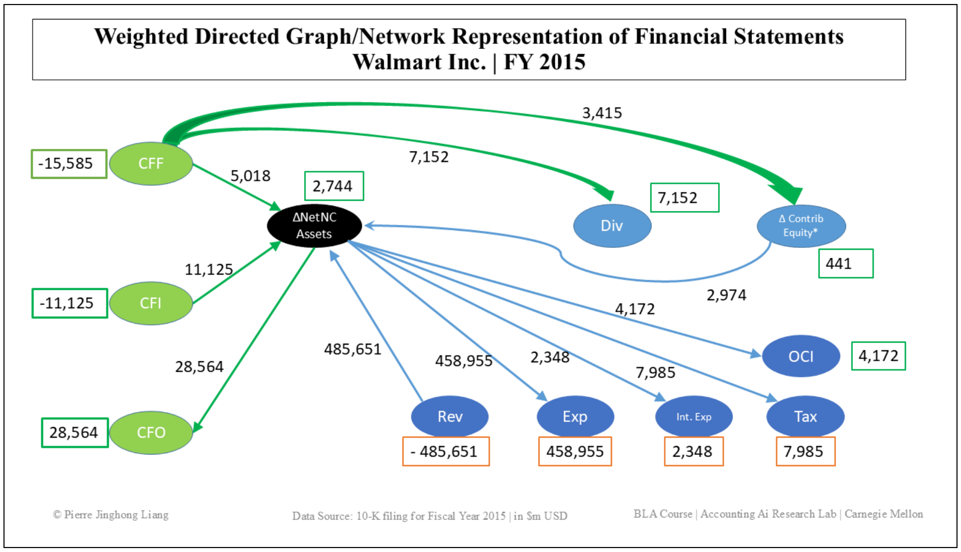

Figure 5: An Example Bookkeeping Graph for Walmart FY2015

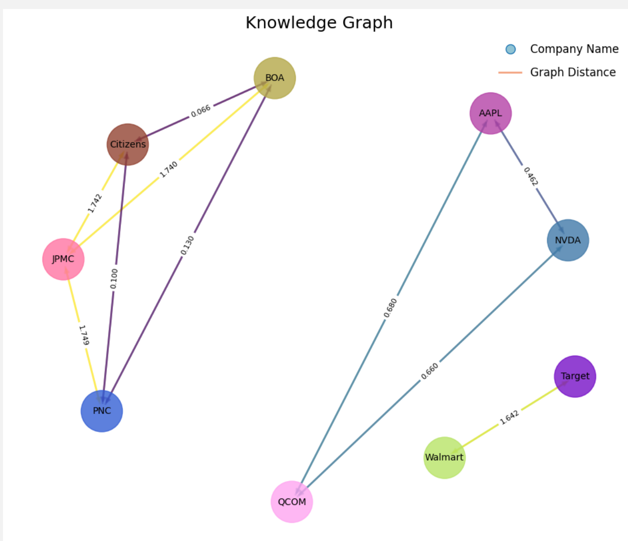

Figure 6: A Knowledge Graph based on Pair-wise Graph-Distance based on Edge-entropy

The final summary measure attempts to capture classification information contained in **both** nodes and edges in the bookkeeping graph. Notice while node-weight vector X contains all information encoded into the node-weights; they are reproducible by the edge classification encoded in the Adjacency matrix W. While Edge-entropy measure is constructed using all information necessary to construct the bookkeeping graph, it is nevertheless order-invariant. That is, if we switch positions of any pair of entries in the adjacency matrix the Edge-entropy measures would remain the same even though the bookkeeping graphs have been altered.

To address this expressivity issue, we deploy the Laplacian Matrix representation of weighted and directed graph. Generally, the Laplacian L of any graph G with m nodes is defined as

$$L = ID - A \tag{9}$$

where I is a $m \times m$ identity matrix, D is a $m \times 1$ vector of node-values (equal to X in our case), and A is a $m \times m$ adjacency matrix (equal to W in our case). Originally proposed in the context of physical circuit laws by Kirchhoff [1847], the Laplacian matrix became the central object of focus in studying the spectral properties of graph as in Chung [1997].[11]

Node-inflow +Edge Values	acct-1	acct-2	acct-3	...	acct-m	Total
acct-1	$\sum_{j=1}^{m} w_{j1}$	$-w_{12}$	$-w_{13}$	...	$-w_{1m}$	account-1 total outflows
acct-2	$-w_{21}$	$\sum_{j=1}^{m} w_{j2}$	$-w_{23}$	...	$-w_{2m}$	account-2 total outflows
acct-3	$-w_{31}$	$-w_{32}$	$\sum_{j=1}^{m} w_{j3}$	...	$-w_{3m}$	account-3 total outflows
...	...	...	...	$\sum_{j=1}^{m} w_{j...}$	...	... total outflows
acct-m	$-w_{m1}$	$-w_{m2}$	$-w_{m3}$	...	$\sum_{j=1}^{m} w_{jm}$	Asset-m total outflows
Total	acct-1 inflows	acct-2 inflows	acct-3 inflows	... inflows	acct-m inflows	Total in- or out-flows

Table 5: Bookkeeping Graph-entropy using Graph Laplacian Encoding

For undirected unweighted graphs (i.e., symmetric adjacency matrix), L is symmetric and positive semi-definite. So, its eigen-value and eigen-vectors are well-understood. Similarly to non-graph signal processing principles, the eigen-value decomposition of the Laplacian provides the foundation to graph signal processing as any signal contained in the undirected unweighted graph can be decomposed into $m-1$ components corresponding to the $m-1$ eigen-value and eigen-vector pairs, just like wavelets in regular signal processing such as music notes.

For the Laplacians of directed and weighted graphs, such as our bookkeeping graphs, such decomposition is possible but challenging as L matrix is now asymmetric and two versions of D vector are possible: D_{in} and D_{out} representing the total inflows to or outflows from each node. Table 5 is an example of the bookkeeping graph Laplacian constructed using the adjacency matrix W. In this example, the diagonal of the Laplacian is occupied by total inflows to each node.

Graph Laplacian Preserves Graph Structure The first two entropy-based summary measures surely extract information from the graph. But they capture only a part of the graph's topological features. Node-entropy is a knowledge retrieval tool that focuses only on node-value in the graph, ignoring potentially different edge-values that contribute to the node-value. Similarly, Edge-entropy is a knowledge retrieval tool that focuses only on edge-value in the graph, ignoring the relative position of the edge between two distinct nodes. In other words, the two measures do capture *Graph Structure* but do not capture it in a comprehensive way; the measure's expressivity

[11]The matrix name is a tribute to the French scholar Pierre-Simon Laplace (1749 - 1827), who made important contributions to the development of engineering, mathematics, statistics, physics, astronomy, and philosophy, in works such as the five-volume *Mécanique céleste* (Celestial Mechanics) (1799–1825).

is not rich enough. Graph Laplacian, in contrast, captures the structure in a systematic way. To see this, consider the following numerical examples in Figures 7 and 8. Here four four-node unweighted undirected graphs are used to illustrate that neither node-entropy nor edge-entropy alone capture graph structure. In these examples, the node degree vectors of both fully-connected graph (Full) and circle graph (Circle), although different in graph structure, are shown to have equal distribution of node-values thus sharing the same node-entropy. Similarly, the edge degree adjacency matrices of both path graph (Path) and star graph (Star), although different in graph structure, are shown to have equal distribution of edge-values thus sharing the same edge-entropy.

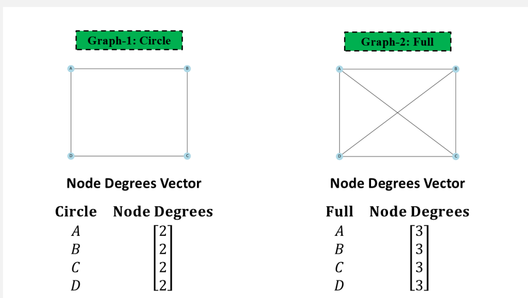

Figure 7: Full- & Circle-graphs Sharing Node-entropy

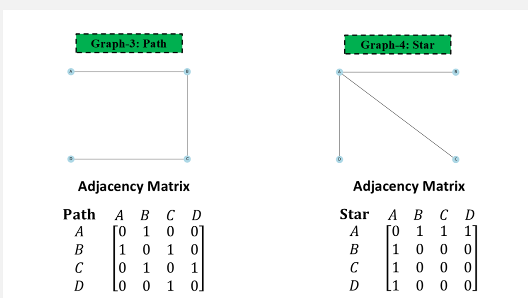

Figure 8: Path- & Star-graphs Sharing Edge-entropy

As a well-developed concept graph spectral analysis [Chung, 1997], graph Laplacian matrix as defined in Equation (9) preserves the graph structure in the context of spectral analysis. One key step to derive the eigenvalue (and the associated eigen vectors) of the normalized version of matrix L. Figure 9 shows the regular Laplacian matrices of the four four-node graphs used in the illustration earlier. Notice the Laplacian incorporate both the node-degree vector and the adjacency matrix in one matrix. In a physics context, von Neumann showed that the quantum structure is captured by the trace of the eigenvalues of a normalized version of the Laplacian. The Shannon entropy of the eigenvalue distribution in the graph context is thus called von Neumann graph entropy in his honor.

Challenges in Analyzing Bookkeeping Graph Laplacian While much of the graph signal processing has been developed to process and extract information contained in graph form [Ortega, 2022], the processing of directed weighted graphs is an active area of modern research. One key challenge is to address the difficulty in retrieving and interpreting the eigen-values derived from a Laplacian matrix that is asymmetric. In Bookkeeping graph case, the asymmetry of the Laplacian comes from the inherent directed nature of the edges in any bookkeeping graph: from credit to debit. With the asymmetric, the eigen-values are no longer guaranteed to be real number because L is not always positive semi-definite. Many solutions have been proposed in applied areas with use-cases in physics, medicine, biology, modern machine learning, and graph data-mining.

A research team at Carnegie Mellon is now actively developing solutions to (1) construct and

Graph-1: Circle — Graph Laplacian Matrix (Path)

$$\begin{array}{c|cccc} & A & B & C & D \\ \hline A & 2 & -1 & 0 & -1 \\ B & -1 & 2 & -1 & 0 \\ C & 0 & -1 & 2 & -1 \\ D & 0 & 0 & -1 & 2 \end{array}$$

Graph-2: Full — Graph Laplacian Matrix (Star)

$$\begin{array}{c|cccc} & A & B & C & D \\ \hline A & 3 & -1 & -1 & -1 \\ B & -1 & 3 & -1 & -1 \\ C & -1 & -1 & 3 & -1 \\ D & -1 & -1 & -1 & 3 \end{array}$$

Graph-3: Path — Graph Laplacian Matrix (Path)

$$\begin{array}{c|cccc} & A & B & C & D \\ \hline A & 1 & -1 & 0 & 0 \\ B & -1 & 2 & -1 & 0 \\ C & 0 & -1 & 2 & -1 \\ D & 0 & 0 & -1 & 1 \end{array}$$

Graph-4: Star — Graph Laplacian Matrix (Star)

$$\begin{array}{c|cccc} & A & B & C & D \\ \hline A & 3 & -1 & -1 & -1 \\ B & -1 & 1 & 0 & 0 \\ C & -1 & 0 & 1 & 0 \\ D & -1 & 0 & 0 & 1 \end{array}$$

Figure 9: Laplacian Matrices for the Four Example Four-node Graphs

transform the appropriate Laplacian for directed and weighted bookkeeping graphs that preserve directionality; (2) overcome technical challenge to reduce computational complexity to produce eigenvalue and vectors with a promising path being converting the Bookkeeping Graph Laplacian into a Hermitian matrix while preserving the directionality and relative edge weights; (3) analyze the eigenvalue and vectors to compute graph-entropy and its associated K-L divergence; and (4) apply the graph-entropy metrics in answering questions in realistic empirical settings.

5 Conclusions

In this paper, three laws are proposed to describe the conventional double-entry system of bookkeeping. Inspired by Claude Shannon's entropy concept, the paper proposes three different summary measures of records produced by a conventional double-entry bookkeeping system. These measures are: Balance-sheet Node-entropy, Bookkeeping Graph Edge-entropy, and Bookkeeping Graph Laplacian-entropy. These measures have deep roots in information theory and its application to study graph complexity.

While real progress is made only when using these measures make real contributions in empirical work, we wish to add to the long line of high praises of the enduring beauty and almost magical depth of the 500-year-old invention that is Double-entry Bookkeeping. By providing records in a graph form (before graph theory was invented in the 1700's), double-entry bookkeeping continues to offer layers after layers of structural features that both serve society well and provide ample opportunities for scholarly inquires armed with increasingly modern computational tools.

In this light, bookkeeping is indeed worthy defending today as much as when Professor Henry Rand Hatfield [1924] made his famous defense. On a personal note, we consider it worthy to pursue the one subject that had attracted Professor Yuji Ijiri's "attention and fascination" for forty plus years. At the 500th anniversary celebration of Pacioli's original text, Ijiri [1993] fondly stated: "[d]ouble-entry bookkeeping, especially its mathematical structure, has been the author's life-long mystery" (page 266). We continue to share the fascination indeed.

References

Arya, A., Fellingham, J., Schroeder, D., 2000a. Accounting information, aggregation, and discriminant analysis. Management Science 46, 790–806.

Arya, A., Fellingham, J. C., Glover, J. C., Schroeder, D. A., Strang, G., 2000b. Inferring transactions from financial statements. Contemporary Accounting Research 17, 366–385.

Arya, A., Fellingham, J. C., Mittendorf, B., Schroeder, D. A., 2004. Reconciling financial information at varied levels of aggregation. Contemporary Accounting Research 21, 303–324.

Arya, A., Fellingham, J. C., Schroeder, D. A., 2000c. Estimating transactions given balance sheets and an income statement. Issues in Accounting Education 15, 393–411.

Basu, S., Waymire, G. B., 2008. Accounting is an evolved economic institution. Foundations and Trends® in Accounting 2, 1–174.

Basu, S., Waymire, G. B., 2021. The evolution of double-entry bookkeeping. Available at SSRN 3093303 .

Bonchev, D., Rouvray, D. H., 2003. Complexity: Introduction and fundamentals. CRC Press.

Butterworth, J. E., 1972. The accounting system as an information function. Journal of Accounting Research pp. 1–27.

Cayley, A., 1894. The Principles of Book-keeping by Double Entry. Cambridge University Press.

Charnes, A., Cooper, W. W., Ijiri, Y., 1963. Breakeven budgeting and programming to goals. Journal of Accounting Research pp. 16–43.

Chung, F. R., 1997. Spectral graph theory, vol. 92. American Mathematical Soc.

Coase, R., 1937. The nature of the firm. Economica 4, 386–405.

Cover, T. M., Thomas, J. A., 2006. Elements of information theory. Wiley-Interscience.

Dehmer, M., 2008. Information processing in complex networks: Graph entropy and information functionals. Applied Mathematics and Computation 201, 82–94.

Fellingham, J., 2017. Accounting: An Information Science. Unpublished Manuscript.

Fellingham, J., 2018. The double entry system of accounting. Accounting, Economics, and Law: A Convivium 8.

Fellingham, J. C., Lin, H., Schroeder, D., 2019. Entropy and accounting. Available at SSRN .

Hatfield, H. R., 1924. Historical defense of bookkeeping. Journal of Accountancy 37, 1.

Ijiri, Y., 1967. The Foundation of Accounting Measurement. Prentice Hall.

Ijiri, Y., 1975. Theory of accounting measurement. Amer Accounting Assn.

Ijiri, Y., 1993. The beauty of double-entry bookkeeping and its impact on the nature of accounting information. Economie Notes by Monte dei Paschi di Siena 22, 265–285.

Itami, H., 1977. The concept of the amount of information in information evaluation. Hitotsubashi journal of commerce and management 12, 33–43.

Kirchhoff, G., 1847. Ueber die auflösung der gleichungen, auf welche man bei der untersuchung der linearen vertheilung galvanischer ströme geführt wird (about the resolution of the equations to which one is led when investigating the linear distribution of galvanic currents). Annalen der Physik 148, 497–508.

Kullback, S., Leibler, R. A., 1951. On information and sufficiency. The Annals of Mathematical Dtatistics 22, 79–86.

Lev, B., 1968. The aggregation problem in financial statements: An informational approach. Journal of Accounting Research pp. 247–261.

Lev, B., 1970. The informational approach to aggregation in financial statements: Extensions. Journal of Accounting Research pp. 78–94.

Li, J., Liang, P. J., Hwang, H., 2019. Inter-entity bookkeeping networks: Representations and applications. Tech. rep., Carnegie Mellon.

Li, N., Liang, P. J., Zhang, G., 2024. Revisiting accounting entropy: Measuring information of accounting classification. working paper .

Liang, P. J., 2023. Bookkeeping graphs: Computational theory and applications. Foundations and Trends® in Accounting 17, 77–172.

Liang, P. J., Wang, A., Akoglu, L., Faloutsos, C., 2021. Pattern recognition and anomaly detection in bookkeeping data. Asian Bureau of Finance and Economic Research Working Papers .

Marschak, J., Miyasawa, K., 1968. Economic comparability of information systems. International Economic Review 9, 137–174.

Marshak, J., 1959. Remarks on the economics of information (1959). In: *Economic Information, Decision, and Prediction: Selected Essays: Volume II*, Springer, pp. 91–117.

Mattessich, R. V., 1964. Accounting and Analytical Methods. Irwin, Inc.: Homewood, IL.

Mattessich, R. V., 1995. Critique of Accounting. Quorum Books.

Mowshowitz, A., 1968. Entropy and the complexity of graphs: I. an index of the relative complexity of a graph. The bulletin of mathematical biophysics 30, 175–204.

Ortega, A., 2022. Introduction to graph signal processing. Cambridge University Press.

Penman, S. H., 2013. Financial statement analysis and security valuation, vol. 5. McGraw-Hill New York.

Reinhard, D., 2017. Graph theory. Springer, New York.

Sewell Bray, F., 1966. Accounting Postulates and Principles. Modern Accounting Theory, Prentice-Hall, ed. Backer, Morton.

Shannon, C. E., 1948. A mathematical theory of communication. The Bell System Technical Journal 27, 379–423.

Simon, H. A., 1952. A comparison of organisation theories. The Review of Economic Studies 20, 40–48.

Sunder, S., 1983. Limits to information. Accounting Research: Theoretical and Applied pp. 101–110.

Sunder, S., 1996. Theory of Accounting and Control. South-Western College Pub.

Sunder, S., 2023. Digitization, perception, cognition, and categorization: Consequences for accounting. Available at SSRN 4402108 .

Theil, H., 1969. On the use of information theory concepts in the analysis of financial statements. Management Science 15, 459–480.

Transact- ions #	Memo	Dr/Cr		Accounts	$Dollars	
h	sell LT asset for cash	Dr.	Cash Cr.	 LT-securities	$2,900	 $2,900
$i1$	take delivery of inventory	Dr.	Inventory Cr.	 A/P	$900	 $900
$i2$	pay cash to vendors	Dr.	A/P Cr.	 Cash	$900	 $900
j	pay cash for PPE	Dr.	NPPE Cr.	 Cash	$2,000	 $2,000
$n1$	incur Labor expenses	Dr.	SGA Cr.	 Wage Payable	$3,500	 $3,500
$n2$	pay cash to employees	Dr.	Wage Payable Cr.	 Cash	$2,800	 $2,800
o	Depreciation to inventory	Dr.	Inventory Cr.	 NPPE	$400	 $400
p	Depreciation to SGA	Dr.	SGA Cr.	 NPPE	$250	 $250
$q1$	Credit sales	Dr.	A/R Cr.	 Revenue	$10,00	 $10,00
$q2$	pay cash to vendors	Dr.	Cash Cr.	 A/R	$8,000	 $8,000
o	Product costing	Dr.	COGS Cr.	 Inventory	$1,100	 $1,100
$d1$	incur interest expenses	Dr.	Int Exp Cr.	 Loans Payable	$20	 $20
$d2$	pay cash to lender	Dr.	Loans Payable Cr.	 Cash	$20	 $20

Table 6: Bookkeeping Graph Example: Operating Transactions

6 Appendix A: Examples

6.1 Expanded Example of Graph Representation of Double-entry bookkeeping

In this appendix, we expand the three-transactions example in Figure 3 by adding a host of operating transactions. Then bookkeeping graphs are constructed based on these transactions. Linear aggregations are applied to generated aggregated balances such as net operating assets from individual asset accounts and operating income after tax from individual revenue and expense flows. A few accounting identities are verified and Financial statements are prepared.

Step-1: Adding Operating transactions Table 6 contains the operating transaction examples and their associated journal entry. Refers to Figure 2 for the Chart of Accounts information used in the journal entries.

Step-2: Constructing Bookkeeping Graph Following the graphing convention explained in Figure 3, the operating transactions listed in Table 6 are converted into the bookkeeping graph below:

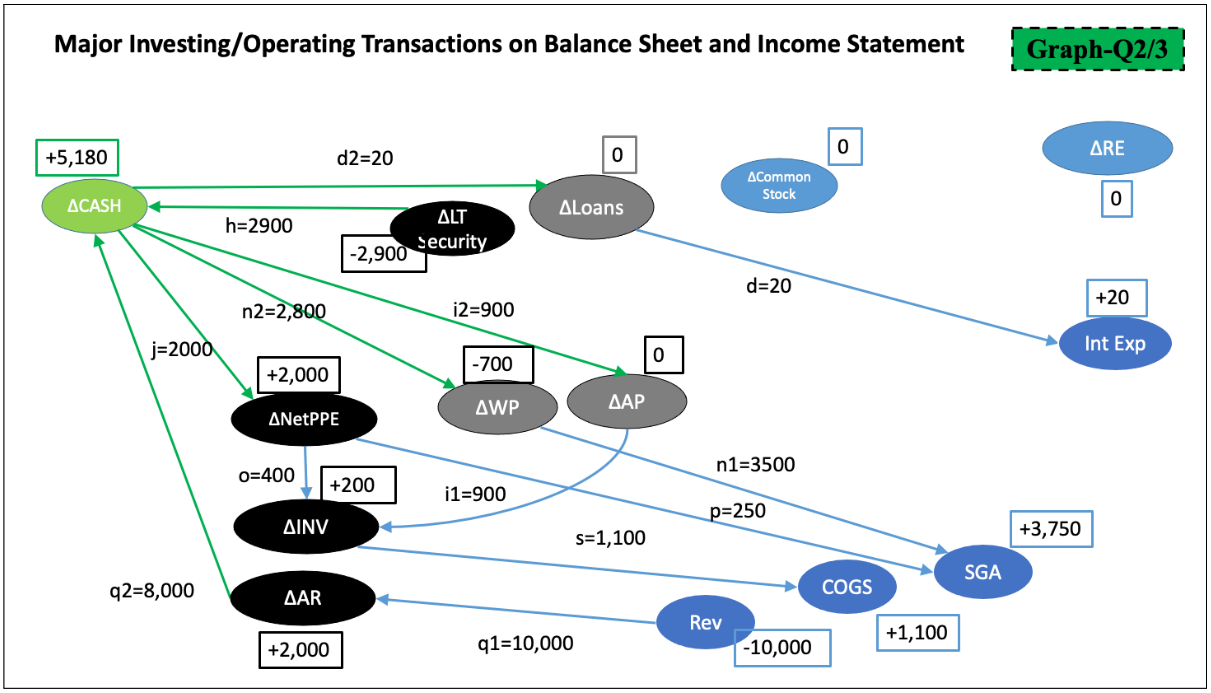

Figure 10: Example of Graph with Operating Transaction Journal Entries

Step-3: Merging and un-merging nodes Now we make three linear aggregation assumptions:

- Net Financial Assets (NFA):

$$Net\ Financial\ Assets = LT\text{-}security - Loans$$

- Net Operating Assets (NOA):

$$Net\ Operating\ Assets = Accounts\ Receivable + Inventory + NetPPE$$

$$- Accounts\ Payable - Wage\ Payable$$

- Operating Income (OI):

$$Operating\ Income = Rev - COGS - SGA$$

We unmerge the *Cash* node into three cash nodes representing *Cash flows from operating activities* (CFO), *Cash flows from investing activities* (CFO), and *Cash flows from financing activities* (CFF).

$$Change\ in\ Cash = CFO + CFI + CFF$$

With these four aggregation and disaggregation procedures, the new simplified Bookkeeping Graph becomes:

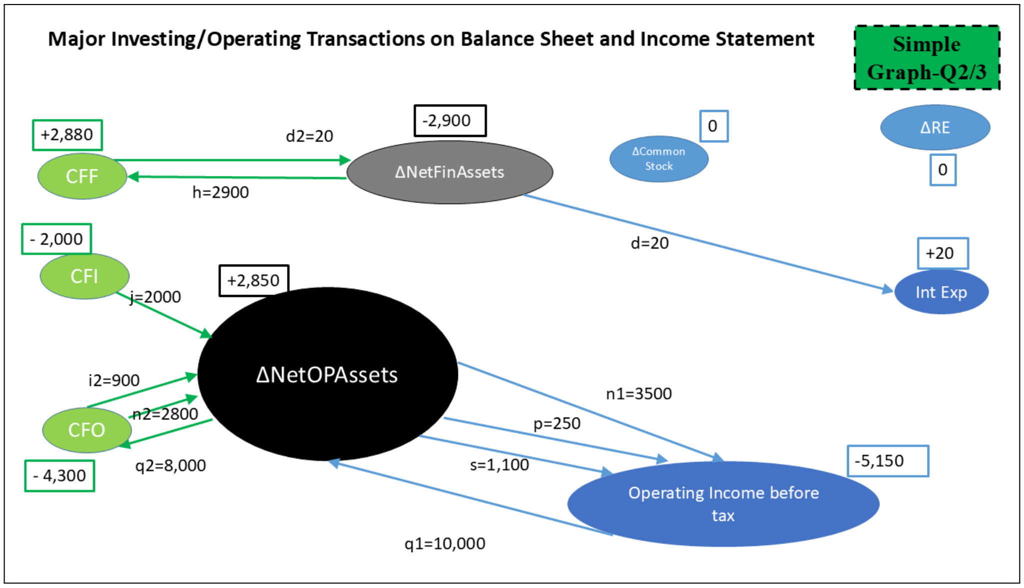

Figure 11: Example of Simple Graph with Operating Transaction Journal Entries

Step-4: Deduce aggregated accounting identities Now verify the conservation law for the aggregated node $\Delta NetOPAssets$:

$$\begin{aligned} \Delta NetOPAssets\ \$2,850 &= Net\ Inflows\ from\ Operating\ Income\ \$5,150 \\ &-(Net\ Outflows\ to\ CFO\ \$4,300 - Net\ Inflows\ to\ CFI\ \$2,000) \end{aligned}$$

Step-5: Financial Statements Income Statement and Balance Sheet are given in Figure 12.

Example: Financial Statements

Income Statement	
Sales Revenue	$10,000
Cost of Goods Sold	1,100
Gross Margin	8,900
Operating Expenses	
SGA	3,750
Operating Income	5,150
Interest Expense	20
Pre-tax Income	5,130
Net Income	5,130

Balance Sheets	
Cash (8000+2900-2900-2800-20)	5,180
Accounts Receivable	2,000
Inventory (900+400-1100)	200
NPPE (2000-400-250)	1,350
LT Security (3000-2900)	100
Total Assets	8,830
Accounts Payable (900-900)	0
Wage Payable (3500-2800)	700
Loans Payable	1,000
Total Liabilities	1,700
Common Stock	2,000
Retained Earnings	5,130
Total Liabilities and Equity	8,830

Figure 12: Example of Financial Statements

6.2 Edge-entropy Small Sample Experiment

FY2015 data of nine companies from their balance sheet, income statement, statement of cash flows, and comprehensive income statements are retrieved from Compustat (North America Annual). Using these statements, one can construct an aggregated bookkeeping graph such as Figure 5 using the three laws of the bookkeeping along with the following assumptions.

- Total equity is broken into three nodes: Retained earnings, AOCI, and contributed equity;
- Net Non-Cash Assets node is computed as *Total Assets − Cash* - *Total Liabilities*;
- change in Retained Earnings other than net income are considered cash dividend;
- Cash flows from financing activities other than debt-financing are linked to contributed equity;
- all income statement items (including OCI) originate or are destined to Net Non-Cash Assets.
- any equity-based expenses are considered to pass-through Net Non-Cash Assets.

In generating the knowledge graph, common-sized adjacency matrices are computed for each of the nine companies using total assets and pair-wise Euclidean distance is calculated among the nine common-sized matrices. The numerical results are given below. To generate the cluster, an ad hoc cut-off and 1.8 is used (i.e., firms are considered graph-close if the distance is less than 1.8, thus a connection in the knowledge graph is drawn). Converting the data frame below into a knowledge graph gives rise to Figure 6

Pair-wise Graph Distance	JPMC	BOA	PNC	Citizens	Target	Walmart	NVDA	QCOM	AAPL
JPMC		1.73992	1.74911	1.74171	2.83391	3.83425	2.176	2.17385	2.26347
BOA	1.73992		0.129668	0.065502	3.32185	3.70006	1.93656	1.91483	2.01737
PNC	1.74911	0.129668		0.0995119	3.32502	3.70266	1.94153	1.88796	2.01896
Citizens	1.74171	0.065502	0.0995119		3.32735	3.70582	1.93872	1.91157	2.0256
Target	2.83391	3.32185	3.32502	3.32735		1.64237	2.14972	2.42337	2.1256
Walmart	3.83425	3.70006	3.70266	3.70582	1.64237		2.43256	2.76366	2.39153
NVDA	2.176	1.93656	1.94153	1.93872	2.14972	2.43256		0.659527	0.462408
QCOM	2.17385	1.91483	1.88796	1.91157	2.42337	2.76366	0.659527		0.680238
AAPL	2.26347	2.01737	2.01896	2.0256	2.1256	2.39153	0.462408	0.680238	

Figure 13: Pair-wise Graph Distance Knowledge Graph

6.3 Double-entry Bookkeeping in Matrix Form

The modern interest in the matrix representation of double-entry bookkeeping was evident in Butterworth [1972] and Ijiri [1975]. Here we follow the treatment and illustration by Arya et al. [2000b]. The core of this representation is distilled in the creation of a transformation matrix, where each cell takes on only one of three values $\{-1, 0, +1\}$[12]. Under this matrix representation, the accounting process using a double entry rule is re-imagined as transforming a vector representing a group of transactions into another vector representing a set of "T-accounts" balances. We adopt the same notations used in Arya et al. [2000b] and introduce the following.

- Vector $y = (y_1, y_2, ...y_n)$ is an $n \times 1$ column vector, denoting n transaction amounts.
- Vector $x = (x_1, x_2, ..., x_m)$ is an $m \times 1$ column vector, denoting changes in balances of m T-accounts; $m \leq n$ due to accounting aggregation.
- Transformation matrix A is an $m \times n$ matrix representing double-entry bookkeeping process. Each row represents an account, and each column represents a transaction. The accounts are connected by the double-entry bookkeeping process of the transaction. Each column has a pair of non-zero entries: debit to an account is denoted by +1, and credit to an account is denoted by -1. Each transaction in the vector y causes the changes to account balances in vector x through multiplying the debit (credit) account by +1 (-1) and the transactions amount. Therefore, matrix A is the transformation matrix that aggregates transactions in y into balances in x, i.e.,

$$\Rightarrow \boxed{\underset{m\times 1}{x_{end}} = \underset{m\times 1}{x_{beg}} + \underset{m\times n}{A} \cdot \underset{n\times 1}{y}}$$

We use the same simple numerical example as in Arya et al. [2000b] p.369, where there are seven transactions and six accounts, and the matrix representation of double-entry bookkeeping is given by

$$\begin{bmatrix} \text{Cash} & 10 \\ \text{Inventory} & 0 \\ PPE & 0 \\ \text{Sales} & 0 \\ COGS & 0 \\ G\&A & 0 \end{bmatrix} + A \cdot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 6 \\ 10 \\ 5 \\ 3 \end{bmatrix} \qquad \begin{array}{l} x_0 : \text{opening balances} \\ x_1 : \text{ending balances} \\ y : \text{Transactions} \\ x = x_1 - x_0 \text{ account change} \\ A : \text{double entry matrix} \end{array}$$

[12] A matrix with this property is called an incidence matrix, related to a class of matrices known as TUM: totally unimodular matrices

$$A = \left|\begin{array}{ccccccc} \multicolumn{4}{c}{(y_1 - y_4)} & \multicolumn{3}{c}{(y_5 - y_7)} \\ \multicolumn{4}{c}{\text{Cash Transactions}} & \multicolumn{3}{c}{\text{Deferrals \& Accruals}} \\ \multicolumn{4}{c}{\leftarrow - - - - - - \rightarrow} & \multicolumn{3}{c}{\leftarrow - - - - \rightarrow} \\ -1 & -1 & -1 & +1 & 0 & 0 & 0 \\ +1 & 0 & 0 & 0 & -1 & +1 & 0 \\ 0 & +1 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 & 0 & 0 & +1 \end{array}\right|$$

Notice within this example, the journal entry of the cash sales is represented by column-4 of the A-matrix and row-4 of the transaction vector y.

The January 15, 20XX {Dr. Cash; Cr. Sales} journal entry is represented by

$$\{A(Col = 4) = [+1, 0, 0, -1, 0, 0]' \text{ and } y_4\}$$

Matrix representation is an elegant recreation of the double-entry core of bookkeeping and has received much interests even after the shift toward the information paradigm in accounting research. Starting with Butterworth [1972] and Ijiri [1975], later works are represented by Arya et al. [2000c], Arya et al. [2000b], Arya et al. [2000a], Arya et al. [2004] and Li et al. [2019].

7 Appendix B: Mathematical Definitions and Technical Companions

7.1 Shannon's Entropy and Relative Entropy

Shannon's Entropy and its generalization We appeal to the Information Theory, built upon the fundamental entropy ideas pioneered by Shannon [1948], in search of such an internal summary measure of accounting classification. While Shannon's original idea ushered in a broader and deeper Information Theory, we focus on two basic ideas and measures as the foundation to construct our internal summary measure of accounting information: Claude Shannon's Entropy [Shannon, 1948] and Relative Entropy or Kullback-Leibler divergence [Kullback and Leibler, 1951].

Shannon's Entropy

Definition 12 (Entropy) *Let X be a discrete random variable with alphabet $\mathcal{X}$ and probability mass function $p(x) = Pr\{X = x\}, x \in \mathcal{X}$. The entropy $H(X)$ of X is*

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) log_b \, p(x)$$

where b is the base of logarithm.

If the set of signals x's consists of n potential messages that have equal probabilities $p_x = 1/n$, the Entropy of the information source is $H(X) = -n(1/n)log(1/n) = log(n)$ which is increasing in parameter n. Intuitively, one can interpret that the larger the n, the more is the underlying uncertainty. Equivalently, the larger the n, the more precise is the signal serve as an information instrument [Marshak, 1959]. From an operational perspective, the larger the n, it takes more resources to transmit the signal X in expectation.

Relative Entropy: Kullback-Leibler Divergence

Definition 13 (Relative Entropy) *Let $p(x)$ and $q(x)$ are two probability functions of random variable X with alphabet $\mathcal{X}$, the Kullback-Leibler (K-L) Divergence from probability function $q(x)$ to $p(x)$ is defined as*

$$D(q||p) = \sum_{X \in \mathcal{X}} q(x) log \frac{q(x)}{p(x)} = E_q log \frac{q(X)}{p(X)}$$

K-L divergence is always non-negative but is not a true distance measure (it violates symmetry and triangle inequality). It does offer a measure of the inefficiency of assuming distribution q when distribution p is true. Under the true distribution p, $H(p)$ is the average description length of variable X, but if one instead uses q by error, $H(p) + D(p||q)$ is the average length instead. When applying the K-L divergence to quantify the movement from prior probability distribution to the posterior, the K-L divergence captures the new information content contained in the posterior as compared to the prior distribution.

7.2 Axioms for Node-entropy measure

Lemma 14 (Cover and Thomas [2006]) *Consider a positive $m \times 1$ vector $X = [x_1, x_2, ...x_m]'$ where $0 < x_i < +\infty$ and $m \in \{2, 3, 4, ...\} \subset \mathbb{N}$. Now consider a sequence of symmetric functions $J_m(X)$ indexed by m, satisfying the following three properties:*

1. *Normalization: $J_2(x, x) = 2x(1 - log(2x))$*
2. *Continuity: $J_2(x, y - x)$ is continuous in x for all $\{x, y \mid x < y\}$*
3. *Grouping:*

$$J_m(x_1, x_2, x_3, x_4...x_m) = J_{m-1}(x_1 + x_2, x_3, x_4...x_m) + (x_1 + x_2) J_2 \left(\frac{x_1}{x_1 + x_2}, \frac{x_1}{x_1 + x_2} \right)$$

Then J_m must be in the form of:

$$J_m(x_1, x_2, x_3, x_4...x_m) = -\sum_{i=1}^{m} x_i log x_i = K H_m(P) - K log K$$

where $P = \frac{X}{K}$, $K = \sum x_i > 0$, *and* $H_m(P)$ *is Shannon's entropy of* P.

This lemma generalizes the entropy measure to cases beyond probability distribution to include all positive vectors.[13] Notice if we re-scale X into $P = [p_1, p_2, ...p_m]' = \frac{X}{\sum x_i}$, then it is straightforward to show:

$$-\sum_{i=1}^{m} p_i log p_i \equiv H_m(P) = \frac{J_m(X) + \sum x_i log \sum x_i}{\sum x_i}$$

Consider a competing function $J_m^* = X'X$ (sum of squares) which is a typical summary statistics of total variation (or total information to be learned) in standard statistical analysis based on linear models. The information-theory-based entropy measure J_m is a summary statistics of total variation $J_m(X) = X' log X$, monotonic to $X'X$ for positive X and now shown to obey the grouping axiom while $X'X$ does not.

7.3 Axioms for Edge-entropy measure

Let $\mathcal{G} = (V, E, W)$ denote a directed weighted graph on m nodes, with:

- $V = \{v_1, v_2, \ldots, v_m\}$ the set of nodes, interpreted as accounts in the chart of accounts;
- $E \subseteq V \times V$ the set of directed edges, with e_{ij} representing a journal-entry flow crediting account i and debiting account j;
- $W = (w_{ij}) \in \mathbb{R}_{\geq 0}^{m \times m}$ the weighted adjacency matrix, with $w_{ij} \geq 0$ the cumulative dollar weight of flow from i to j, and $w_{ii} = 0$ for all i (no self-transactions on atomic accounts).

Denote by $\mathcal{W}_m^0$ the set of $m \times m$ non-negative matrices with zero diagonal. Total edge mass is $S(W) = \sum_{i \neq j} w_{ij}$, assumed positive. The normalized edge distribution is $\pi_{ij} = w_{ij}/S(W)$ for $i \neq j$.

7.3.1 Node-Relabeling and the Domain of Edge Entropy

The chart-of-accounts ordering is an artifact of recordkeeping practice, not informational content: a balance sheet that lists Cash before Receivables conveys the same information as one that lists them in the reverse order. The natural object of analysis is therefore not the matrix W itself but its equivalence class under *simultaneous row-column permutation*—the operation that relabels nodes consistently in both the inflow (column) and outflow (row) dimensions of the adjacency matrix.

Formally, define the equivalence relation $\sim$ on $\mathcal{W}_m^0$ by

$$W \sim W' \quad \iff \quad \exists\, \sigma \in S_m \text{ such that } w'_{ij} = w_{\sigma(i),\sigma(j)} \text{ for all } i, j.$$

[13] The Cover and Thomas [2006] text proves the theorem above for the case where $\sum x_i = 1$. The lemma above is a corollary with a slight linear rescaling manipulation.

The quotient $\widetilde{\mathcal{W}}_m^0 := \mathcal{W}_m^0 / \sim$ is the space of node-labeled-up-to-relabeling adjacency matrices.

This is the right domain on which to define entropy measures of the bookkeeping graph for two reasons. First, simultaneous row-column permutation respects the directed structure of journal-entry flows: it preserves which entries lie above the diagonal, which lie below, and which are bilateral. Arbitrary permutation of the $m(m-1)$ off-diagonal cells—which would let one swap w_{ij} with w_{ji} freely—would destroy this directionality and is therefore not an admissible symmetry on the bookkeeping graph. Second, simultaneous row-column permutation is precisely the symmetry generated by node-relabeling: it quotients out the arbitrary chart-of-accounts ordering while preserving every other structural feature.

We therefore frame all subsequent results as statements about functions on $\widetilde{\mathcal{W}}_m^0$, treating node-relabeling invariance as a property of the domain rather than an axiom.

Definition 15 (Shannon edge entropy) *For $W \in \mathcal{W}_m^0$ with $S(W) > 0$, the Shannon edge entropy is*

$$H_m^{edge}(W) \;=\; -\sum_{i \neq j} \pi_{ij} \log \pi_{ij}, \tag{10}$$

where the sum is over ordered pairs with $i \neq j$.

By construction, H_m^{edge} is a function on the quotient $\widetilde{\mathcal{W}}_m^0$: it depends only on the multiset of edge weights, not on the row-column labeling. We do *not* claim invariance under arbitrary permutation of the $m(m-1)$ off-diagonal cells, which would be a stronger and accounting-inadmissible symmetry. The relevant invariance is exactly node-relabeling, built into the domain.

7.3.2 The Node-Merge Operation and the Grouping Identity

The natural grouping operation in the adjacency-matrix setting is not pairwise merging of edges (which lacks accounting interpretation) but merging of two nodes into one, with the incident edges aggregated accordingly.

Definition 16 (Node-merge, zero-diagonal) *For $W \in \mathcal{W}_m^0$ and $a, b \in \{1, \dots, m\}$ with $a \neq b$, the* node-merge *operation $W \mapsto W^{(a,b)}$ produces a matrix in $\mathcal{W}_{m-1}^0$ as follows. Let c denote the merged node. Then:*

$$w_{cj}^{(a,b)} = w_{aj} + w_{bj} \quad \text{for } j \notin \{a, b\}, \tag{11}$$

$$w_{ic}^{(a,b)} = w_{ia} + w_{ib} \quad \text{for } i \notin \{a, b\}, \tag{12}$$

$$w_{ij}^{(a,b)} = w_{ij} \quad \text{for } i, j \notin \{a, b\}, \tag{13}$$

$$w_{cc}^{(a,b)} = 0. \tag{14}$$

The bilateral mass $w_{ab} + w_{ba}$ is eliminated; total edge mass becomes $S' = S(W) - w_{ab} - w_{ba}$.

The zero-diagonal convention $w_{cc}^{(a,b)} = 0$ corresponds to the *intercompany elimination* practice in consolidated financial reporting: transactions between entities that become a single reporting unit post-merger vanish from the consolidated statements.

Define, for the merge $(a, b) \to c$:

$$\begin{aligned} \tau &= \pi_{ab} + \pi_{ba} && \text{(bilateral mass fraction)}, \\ p_j^{\text{out}} &= \pi_{aj} + \pi_{bj} && \text{(row-merge mass for external node } j), \\ p_i^{\text{in}} &= \pi_{ia} + \pi_{ib} && \text{(column-merge mass for external node } i). \end{aligned}$$

We denote by $H_2(x, 1-x) = -x \log_2 x - (1-x)\log_2(1-x)$ the binary Shannon entropy.

Theorem 17 (Node-merge grouping identity, zero-diagonal) *For any $W \in \mathcal{W}_m^0$ with $S(W) > 0$ and any pair $a \neq b$,*

$$\begin{aligned} H_m^{edge}(W) \;=\; &(1-\tau)\, H_{m-1}^{edge}(W^{(a,b)}) \;+\; H_2(\tau, 1-\tau) \\ &+ \sum_{j \notin \{a,b\}} p_j^{out}\, H_2\!\left(\tfrac{\pi_{aj}}{p_j^{out}}, \tfrac{\pi_{bj}}{p_j^{out}}\right) \\ &+ \sum_{i \notin \{a,b\}} p_i^{in}\, H_2\!\left(\tfrac{\pi_{ia}}{p_i^{in}}, \tfrac{\pi_{ib}}{p_i^{in}}\right) \\ &+ \tau\, H_2\!\left(\tfrac{\pi_{ab}}{\tau}, \tfrac{\pi_{ba}}{\tau}\right). \end{aligned} \tag{15}$$

Proof. Partition the $m(m-1)$ off-diagonal cells of W into four disjoint classes:

1. *Untouched external cells*: (i, j) with $i, j \notin \{a, b\}$ and $i \neq j$. These pass through the merge unchanged.

2. *Row-merge pairs*: for each $j \notin \{a, b\}$, the pair $\{(a, j), (b, j)\}$ collapses to a single cell (c, j) of mass p_j^{out}.

3. *Column-merge pairs*: for each $i \notin \{a, b\}$, the pair $\{(i, a), (i, b)\}$ collapses to a single cell (i, c) of mass p_i^{in}.

4. *Bilateral pair*: $\{(a, b), (b, a)\}$, which is *removed* from the distribution under the zero-diagonal convention.

For any partition of a probability distribution $\{\pi_k\}$ into groups $G_1, \ldots, G_K$ with group masses $P_\ell = \sum_{k \in G_\ell} \pi_k$, the standard grouping identity gives

$$-\sum_k \pi_k \log \pi_k \;=\; -\sum_\ell P_\ell \log P_\ell \;+\; \sum_\ell P_\ell \left(-\sum_{k \in G_\ell} \tfrac{\pi_k}{P_\ell} \log \tfrac{\pi_k}{P_\ell} \right). \tag{16}$$

Apply (16) with the four classes above as groups. The first term on the right-hand side of (16) is

$$-\sum_{i,j\notin\{a,b\},i\neq j}\pi_{ij}\log\pi_{ij} \;-\; \sum_{j\notin\{a,b\}}p_j^{\text{out}}\log p_j^{\text{out}} \;-\; \sum_{i\notin\{a,b\}}p_i^{\text{in}}\log p_i^{\text{in}} \;-\; \tau\log\tau.$$

The within-group entropy terms for row-merge, column-merge, and bilateral classes are binary entropies multiplied by their respective masses (classes with singleton groups contribute zero).

Using the post-merge renormalization $\pi_{ij}^{(a,b)} = \pi_{ij}/(1-\tau)$ for external cells and $\pi_{cj}^{(a,b)} = p_j^{\text{out}}/(1-\tau)$, $\pi_{ic}^{(a,b)} = p_i^{\text{in}}/(1-\tau)$, one can verify

$$H_{m-1}^{\text{edge}}(W^{(a,b)}) \;=\; \frac{1}{1-\tau}\left[-\sum_{i,j\notin\{a,b\},i\neq j}\pi_{ij}\log\pi_{ij} - \sum_j p_j^{\text{out}}\log p_j^{\text{out}} - \sum_i p_i^{\text{in}}\log p_i^{\text{in}}\right] + \log(1-\tau).$$

Multiplying by $(1-\tau)$ and adding $-\tau\log\tau = H_2(\tau, 1-\tau) - (1-\tau)\log(1-\tau) + \tau\log\tau - \tau\log\tau$ gives, after simplification, the claimed identity. (The algebra is straightforward; the key step is the bookkeeping of the $\log(1-\tau)$ terms.) ■

7.3.3 Axiomatic Characterization

We now consider an arbitrary sequence of functions $\{J_m\}_{m\geq 2}$, where each $J_m : \widetilde{\mathcal{W}}_m^0 \to \mathbb{R}$ assigns a real number to each node-relabeling-equivalence class of zero-diagonal adjacency matrices, and ask: under what axioms must J_m coincide with H_m^{edge}?

Because the domain $\widetilde{\mathcal{W}}_m^0$ is the quotient under simultaneous row-column permutation, node-relabeling invariance is built into the formulation rather than imposed as a separate axiom. This parallels the Faddeev (1956) treatment of the classical case, in which entropy is defined on the space of probability distributions on m atoms (rather than on ordered tuples in $\mathbb{R}^m$), so that symmetry is a property of the domain rather than an axiom about the function. The Cover and Thomas (2006) presentation derives symmetry from the grouping axiom; we adopt the cleaner Faddeev convention because it isolates more sharply the axiom that does the substantive work—the grouping axiom—from the structural choice of domain.

The Axioms Denote by $J_2^{\text{bin}}(x) := J_2(\pi)$ where π is the two-edge distribution with probabilities x and $1-x$. (On $\widetilde{\mathcal{W}}_2^0$, the two off-diagonal cells are exchanged by node-relabeling, so $J_2^{\text{bin}}(x) = J_2^{\text{bin}}(1-x)$ automatically.)

Axiom 18 (Normalization) $J_2^{bin}(1/2) = 1$.

Axiom 19 (Continuity) *J_2^{bin} is continuous on* $(0,1)$.

Axiom 20 (Node-merge grouping) *For any $W \in \mathcal{W}_m^0$ and any pair $a \neq b$, the identity* (15) *holds, with a* single *binary information function J_2^{bin} appearing in all four binary-entropy roles on the right-hand side: bilateral-elimination, row-split, column-split, and the recursive base.*

Three remarks on the axioms.

First, the axiomatization is structurally parallel to Faddeev's: normalization, continuity, and a single grouping axiom. Cover and Thomas's monotonicity axiom is replaced by node-merge grouping, which is stronger.

Second, no symmetry axiom appears in the list. Node-relabeling invariance is built into the domain (Section 3.2); no further symmetry between row-cells and column-cells, or between π_{ij} and π_{ji}, is imposed. This is essential: imposing such a symmetry on the matrix would conflate the credit side with the debit side at the matrix-cell level, foreclosing on debit-credit duality as a substantive constraint on the entropy measure.

Third, the single-function clause of Axiom 20 is stated for definiteness. In principle, one could admit four distinct binary functions—$\phi_{\text{bil}}, \phi_{\text{row}}, \phi_{\text{col}}, \phi_{\text{base}}$—in the four roles and still obtain a well-defined grouping rule. One might therefore expect the single-function clause to carry substantive axiomatic content, specifically encoding debit-credit duality at the axiom level. Proposition **??** below establishes that this expectation is wrong: under Axioms 18 and 19 and the weakened Axiom 20 admitting four continuous symmetric normalized functions, node-relabeling invariance alone forces the four functions to coincide. The single-function clause of Axiom 20 is therefore redundant: no separate axiomatic commitment to duality is needed.

A Sufficient-Determination Lemma The following lemma separates two tasks in the uniqueness argument: pinning down the functional form of J_2^{bin}, and propagating that determination to J_m on the full domain $\widetilde{\mathcal{W}}_m^0$. It is valid for any J_2^{bin}, not just H_2.

Lemma 21 (Sufficiency of J_2^{bin}) *Suppose $\{J_m\}_{m\geq 2}$ satisfies Axioms 18, 19, 20, and suppose $J_2^{bin} : (0,1) \to \mathbb{R}$ is determined. Then $J_m(W)$ is determined for every $m \geq 2$ and every $W \in \widetilde{\mathcal{W}}_m^0$ with $S(W) > 0$.*

Proof. Induction on m.

Base ($m = 2$). For $W \in \widetilde{\mathcal{W}}_2^0$, the edge distribution is $(x, 1-x)$ for some $x \in (0,1)$, and $J_2(W) = J_2^{\text{bin}}(x)$ by the definition of J_2^{bin}. Determined.

Inductive step ($m > 2$). Fix $W \in \widetilde{\mathcal{W}}_m^0$ with $S(W) > 0$, and choose any pair $a \neq b$. By Axiom 20:

$$J_m(W) = (1-\tau)\, J_{m-1}(W^{(a,b)}) + J_2^{\text{bin}}(\tau) + \sum_{j\notin\{a,b\}} p_j^{\text{out}}\, J_2^{\text{bin}}\left(\tfrac{\pi_{aj}}{p_j^{\text{out}}}\right) + \sum_{i\notin\{a,b\}} p_i^{\text{in}}\, J_2^{\text{bin}}\left(\tfrac{\pi_{ia}}{p_i^{\text{in}}}\right) + \tau\, J_2^{\text{bin}}\left(\tfrac{\pi_{ab}}{\tau}\right).$$

By the inductive hypothesis, $J_{m-1}(W^{(a,b)})$ is determined. Every evaluation of J_2^{bin} on the right-hand side is at an explicit argument in $(0,1)$ and is therefore determined by hypothesis. Hence $J_m(W)$ is determined. ■

Remark 22 (Consistency of Axiom 20) *Axiom 20 prescribes a value for $J_m(W)$ for each choice of pair (a,b). For Axiom 20 to be satisfiable at all, all pair choices must prescribe the same value.*

This consistency is guaranteed once $J_2^{bin} = H_2$ by Theorem 17, which establishes that H_m^{edge} satisfies the identity for every (a, b). The Sufficiency Lemma therefore determines $J_m(W)$ consistently: the right-hand side of Axiom 20, evaluated with $J_2^{bin} = H_2$ and with $J_{m-1} = H_{m-1}^{edge}$ by induction, equals $H_m^{edge}(W)$ independent of the pair choice.

Theorem 23 (Uniqueness of edge entropy) *Suppose $\{J_m\}_{m\geq 2}$ satisfies Axioms 18–20. Then for every $m \geq 2$ and every $W \in \mathcal{W}_m^0$ with $S(W) > 0$,*

$$J_m(W) \;=\; H_m^{edge}(W).$$

Proof. The proof proceeds in three steps: (i) derive a functional equation on J_2^{bin} by applying Axiom 20 to a three-node family in two distinct ways; (ii) solve the functional equation using Axioms 18 and 19; (iii) apply Lemma 21 to extend the determination to all of $\widetilde{\mathcal{W}}_m^0$.

Step 1: Functional equation from $m = 3$. We derive the functional equation on the direction-symmetric 3-node family $\pi_{ij} = \pi_{ji}$. The restriction to this family is a dimensional convenience, not an accounting commitment: the family's 2-parameter freedom matches the argument dimensionality of the Faddeev equation, and its direction-symmetry causes the recursive-base and within-bilateral arguments to land at 1/2, where normalization immediately fixes them. The accounting content—debit-credit duality—does not enter at this step; it enters through the domain's node-relabeling invariance.Lemma 21 establishes that the determination of J_2^{bin} via the symmetric-family equation extends to J_m on all $W \in \widetilde{\mathcal{W}}_m^0$, symmetric or not.

Consider the three-node family parameterized by $(a, b, c) \in \mathbb{R}^3_{\geq 0}$ with $2a + 2b + 2c = 1$, defined by

$$\pi_{12} = \pi_{21} = a, \quad \pi_{13} = \pi_{31} = b, \quad \pi_{23} = \pi_{32} = c.$$

Merging $\{2, 3\}$: $\tau = 2c$, $p_1^{\text{out}} = a + b$, $p_1^{\text{in}} = a + b$. The post-merge graph has two edges with probabilities $(a + b)/(a + b) = 1/2$ each after renormalization (by symmetry $p_1^{\text{out}} = p_1^{\text{in}}$ and $1 - 2c = 2(a + b)$), so $J_2^{\text{post}} = J_2^{\text{bin}}(1/2) = 1$ by Axiom 18. The bilateral split is $c/(2c) = 1/2$, giving binary $J_2^{\text{bin}}(1/2) = 1$.

Applying Axiom 20:

$$J_6(\pi) = H_2(2c, 1 - 2c) \cdot \mathbf{1} + (1 - 2c) \cdot 1 + 2(a + b)\, J_2^{\text{bin}}\Big(\tfrac{a}{a+b}\Big) + 2c \cdot 1, \tag{17}$$

where H_2 here denotes the *known* binary function induced by Axiom 20—which, by the single-function requirement, coincides with J_2^{bin}. Write $\phi := J_2^{\text{bin}}$ (treating ϕ as a function of a single argument in $(0, 1)$, with $\phi(x) = \phi(1 - x)$ by symmetry).

Merging $\{1, 2\}$: $\tau' = 2a$, analogous computation gives

$$J_6(\pi) = \phi(2a) + (1 - 2a) \cdot 1 + 2(b + c)\, \phi\Big(\tfrac{b}{b+c}\Big) + 2a \cdot 1. \tag{18}$$

Equating (17) and (18) and simplifying (using $a + b + c = 1/2$):

$$\phi(2c) + (1 - 2c)\,\phi\Big(\tfrac{a}{a+b}\Big) \;=\; \phi(2a) + (1 - 2a)\,\phi\Big(\tfrac{b}{b+c}\Big)\,. \tag{19}$$

Substitute $x := 2a, y := 2c$, so $a = x/2, c = y/2, b = (1 - x - y)/2$. Then $a/(a + b) = x/(1 - y)$ and $b/(b + c) = (1 - x - y)/(1 - x)$. Using $\phi(1 - z) = \phi(z)$, $\phi((1 - x - y)/(1 - x)) = \phi(y/(1 - x))$. Equation (19) becomes

$$\phi(y) + (1 - y)\,\phi\Big(\tfrac{x}{1-y}\Big) \;=\; \phi(x) + (1 - x)\,\phi\Big(\tfrac{y}{1-x}\Big)\,, \tag{20}$$

valid for all (x, y) with $x, y > 0$ and $x + y < 1$. As (a, b, c) range over the open 2-simplex, the substitutions $x = 2a, y = 2c$ range over the full open triangle $\{(x, y) : x, y > 0, x + y < 1\}$, which is the standard domain of Faddeev's equation.

Step 2: Solve (20)*.* Equation (20) is Faddeev's (1956) fundamental functional equation for binary entropy. The unique continuous solution satisfying $\phi(1/2) = 1$ and $\phi(x) = \phi(1 - x)$ is

$$\phi(x) \;=\; -x \log_2 x - (1 - x) \log_2(1 - x) \;=\; H_2(x, 1 - x).$$

A standard reference is Aczél and Daróczy (1975), Chapter 3. The argument reduces (20) to the Cauchy equation $f(xy) = f(x) + f(y)$ via the substitution $f(x) = -\phi(x) - \phi(1 - x)$ on the domain where both arguments are in $(0, 1)$, together with continuity.

Step 3: Extend to $m \geq 2$ via the Sufficiency Lemma. By Step 2, $J_2^{\text{bin}} = H_2$ on $(0, 1)$. Apply Lemma 21: $J_m(W)$ is determined for every $m \geq 2$ and every $W \in \widetilde{\mathcal{W}}_m^0$.

To identify the determined value as $H_m^{\text{edge}}(W)$, proceed by induction. For $m = 2$: $J_2(W) = J_2^{\text{bin}}(\pi_{12}) = H_2(\pi_{12}) = H_2^{\text{edge}}(W)$ directly. For $m > 2$: the inductive hypothesis $J_{m-1} = H_{m-1}^{\text{edge}}$ together with $J_2^{\text{bin}} = H_2$ makes the right-hand side of Axiom 20's identity coincide arithmetically with the right-hand side of Theorem 17's identity applied to $H_m^{\text{edge}}(W)$. By the consistency of Axiom 20 (guaranteed by Theorem 17 as satisfiability witness), either side equals $H_m^{\text{edge}}(W)$. Hence $J_m(W) = H_m^{\text{edge}}(W)$ for all m and all W. This completes the proof. ■