# A Network Analysis of Spinoza's *Ethics*

Michelle Dong & Adrien Wright

April 2024

## 1 Introduction

Baruch Spinoza was, in the words of the scholar Antonio Negri, an "anomaly." He worked as a philosopher in the late 17th century in Amsterdam. Like many European philosophers of this time (his contemporaries included Descartes, Liebniz, and Hobbes), he engaged with philosophical questions about the use of reason and the source of political power. He wrote his most famous text, the *Ethics*, in the system of geometric proof, where axioms and definitions are used to support deductive movement between propositions. However, this is where the general similarities between Spinoza and the common philosophy of his time end. Although he engaged with his philosophical environment on its own terms– constructing relationships between common concepts like "freedom," "substance," "Reason," and "God"– he nonetheless built a system which differed very radically from his contemporaries. Indeed, his views were so anomalous that he was excommunicated by his Jewish community in Amsterdam, lived in fear of political persecution by Dutch Monarchists, and ended up in an unmarked mass grave. So what is this anomalous system?

Spinoza reused common terms from Enlightenment philosophy, but his meanings are not always the most obvious. For example: what is freedom? His contemporary Hobbes gives an answer: a free person is "he, that in those things, which by his strength and wit he is able to do, is not hindred to doe what he has a will to" (Hobbes, Chapter XXI). This definition is, on its surface, very standard to our understanding of what it means to be free: a capacity to act. For Spinoza, on the other hand, a being is "free" when it "exists solely from the necessity of its own nature" (Spinoza trans. Shirley, 31). This definition is far less intuitive; the relevance of existence and necessity to the concept of freedom may not be immediately clear to a reader of the *Ethics* opening to the first page for the first time. We apply the tools of network science to this noteworthy text to develop a deeper understanding of the system that Spinoza builds. Rather than connect concepts in the anomalous text of the book, we build a network of citation between statements. This reveals the structure of the book's proof, from which we analyze centrality characteristics of the network, detect communities, and study the distribution of various graph attributes. In performing this broad survey, we identify interesting characteristics of the proof structure of the text; this quantitative approach opens several questions for future scholarship.

# 2 Background

## 2.1 Literature Review

In choosing Ethics as the basis of our network, we looked at a variety of existing works. In particular, we wouldn't have as much success creating our networks without the works of Torin Doppelt's, *Digital Spinozism.* This project digitalizes the framework of Spinoza's Ethics and incorporates visualizations of the network on an interactive page. Doppelt catalogues citations in the *Ethics* as well as many other texts from Spinoza's corpus. Doppelt spoke with us about the important distinction between relevance in the proof structure and relevance to the text's content. He identified that many "important" nodes, widely recognizable by Spinoza scholars, may be unused or uncited, because they terminate arguments or are justified in the prose rather than with deductive reasoning. He also told us a little about the history of studying Spinoza's proof, including attempts to document citations by Jonathan Bennett, R.F. Tredwell, and more. The *Digital Spinozism* project includes visualizations through different hierarchical, physics-based, and geometric graph layouts, providing an insightful entry point to the structure of the text outside of its written chronology. He also shares an adjacency matrix for citations in the text, as well as displays some centrality measures for each node. (In his work, Spinoza values knowledge which operates outside of a framework of temporal duration, and prizes the ability to consider ideas in their entirety.) In all, Doppelt's project and his generous dialogue with us provided a firm mathematical as well as philosophical foundation to approach the work.

Alain Badiou is among the Spinoza scholars who advocate for a serious consideration of the text's geometric proof structure. *In What is a Proof in Spinoza's Ethics?*, He looks at what he names the "skeleton" of proofs. The skeleton of a proof essentially takes a proposition and traces all the ancestors that allows for the proposition to hold, all the way back to existing definitions and axioms that don't have any ancestors; we could understand this as a backwards tree, where the proposition is the root and we are creating the tree of parent existing proofs, with axioms and definitions as the leaves, Figure 1.

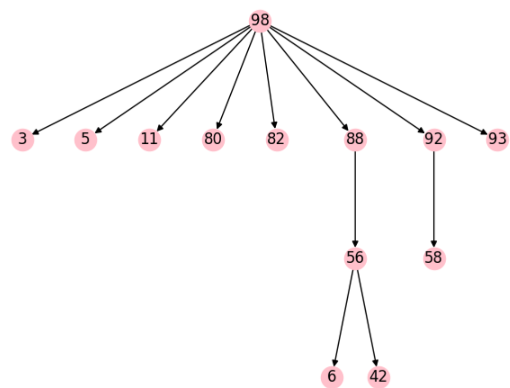


Figure 1: An example of one of our skeletons: the skeleton for node 98 which corresponds to Proposition 1P28P

The skeleton serves as a measure for the complexity of a proof and “a sort of spectral analysis of the entire ontology of Spinoza” (46, Badiou). For our project, we will replicate this method to explore the skeleton of proofs. Compared to centrality measures which only capture one “layer” of interaction, between the node and its parent, the skeleton allows us to chase interactions further to see more of the indirect influences of a node throughout the whole network. This will be addressed in sections 4.3.3 and 4.3.4.

# 3 Methods

## 3.1 Building the Network

We construct a directed network on the proof structure of Spinoza's *Ethics*. The nodes of this network are what we characterize as distinct “statements:” pieces of text that Spinoza separates from one another by particular formatting in the text. These nodes have many attributes (their “type,” the part of the book that contains them, or the primary statement to which they are connected, for example); and many different relationships to one another (the statements cite each other for support, and reference many of the same terms and concepts). We will explicate some relevant attributes shortly. In this network, statement i has a directed edge pointing towards statement j when statement i cites statement j. For example, IP7d states: “Substance cannot be produced by anything else (Cor.Pr.6), and is therefore self-caused (causa sui); that is (Def.1), its essence necessarily involves existence; that is, existence belongs to its nature” (Spinoza trans.

Shirley, 34). The statement IP7d has two directed out-edges: one pointing to IP6c, and one pointing to ID1D. The overall network has 898 nodes and 1421 edges. The data underlying the network is from Torin Doppelt's *Digital Spinozism* prject; for textual interpretation, we referenced Samuel Shirley's English translation.

Nodes are named with the roman numeral for the "part" of the book they are from, the primary statement to which they are attached, and their "type" (this will be discussed in section 3.2 and Table 1). IP7d, then, refers to the proof of Proposition 7 in Part I. Spinoza uses a different and sometimes inconsistent system of reference; we studied the nodes using these names and also a numerical ID corresponding to the order they appear in the text. We are indebted to Torin Doppelt for his careful work documenting these often inconsistent references.

We made many careful choices about how to construct the network to best preserve the structure of the text. Maintaining the integrity of these atomic statements was one such choice. We also decided that the direction of an edge (u,v) should encode the predicate (u references v), since the context of our project was to produce a pseudo-citation network on the text. Here, we diverged from Badiou's skeletons: Badiou considers a genealogical interpretation of the text, in which a directed edge (u,v) exists when (u is referenced by v). This network is thus the reverse of ours. Badiou also groups secondary statements with their primary statements. Our choices about network construction reflect our aim to contextualize the proof structure as a citation-network-like object, acknowledging and differentiating the functions of various statement types to reveal the delicate interplay of roles in the text.

### 3.2 Areas of Inquiry

In this project, we aimed to complete a broad network analysis on the proof structure, which meant engaging with a wide variety of topics such as centrality measurement, community detection, distributions of various metrics, and modelling. We offer an overview of the information about the proof structure which network methods can provide; we hope to show the value of treating the text's structural qualities with consideration, and open new questions for Spinoza scholars.

We consider both "primary" and "secondary" statements of various types, which Table 1 identifies. Primary statements are those which stand on their own in the text; secondary statements are those which attach to some primary statements. For example, a proposition is considered a primary statement, whereas the proposition's proof is considered a secondary statement.

| Primary | Secondary |
|---|---|
| D = Definition | e = Explication |
| A = Axiom | d = Proof |
| P = Proposition | c = Corollary |
| F = Preface | a = Additional Proof |
| L = Lemma | s = Scholium |
| X = Appendix | |
| S = Postulate | |

Table 1: "Types" of statements by category (Primary vs. Secondary).

We consider primary and secondary statements independently because they can reference entirely independent sets of propositions for entirely independent reasons. Those familiar with the conventions of logical proofwriting may be more or less familiar with many of these types. Spinoza's use of the statement types, however - much like his citation style - is relatively idiosyncratic. Many of the types he uses seem synonymous, such as "explication," "proof," and "additional proof", but he treats each slightly differently (explications attach to definitions wheras proofs attach to propositions, and additional proofs use different concepts than their standard proofs). An overview of the different types is discussed in Beth Lord's illuminating companion text, *Spinoza's Ethics* (2010), particularly the introduction. We will not address these differences too deeply here, but some distinguishing features of the types will nonetheless emerge in our analysis of centrality.

The analysis of centrality measures took a survey of various centrality measures using the Python package networkx. We initially planned to study eigenvector centrality in addition to the other measures explored, but this measure fails to converge, which is a problem that sometimes occurs on directed networks. We didn't known which measure would best indicate importance, so we found the correlation between these measures using RStudio to select a "benchmark measure" which hopefully would indicate several interpretations of centrality.

Next, we examined the results of several community-sorting algorithms and discussed the possible relevance of these detectable community groups. Next, we investigated the distribution of various attributes of the graph, including degree distribution, community size, and "skeleton uses." We measured skeleton uses using networkx's dfs_tree function which creates a tree equivalent to our idea of the skeleton.

```
arr = [] #array to store every skeleton
for node in G:
    graph = nx.dfs_tree(G, node) #Networkx func. that
    #creates the equiv. of a skeleton
    arr.append(graph)
```

This is of particular significance to our treatment of the proof as a citation-network-like object. Some models of citation networks, such as Price's model and (a particular instantiation of it) the Barabasi-Albert model, obey power laws for degree distribution with alpha values $2 \leq \alpha \leq 3$ for a large number of nodes. Although this paper does not attempt an in-depth fit of these distributions, we comment on the appearance of power-law-like behaviors and their possible interpretation. Finally, we compare the network to a random model based on Price's model to further explore structural qualities of the text.

# 4 Results

## 4.1 Centrality

What makes a statement important? Network science provides tools for identifying important nodes within networks. A variety of centrality measures identify different characteristics which might make a statement important in the proof network. We considered the following measures, each using the relevant networkx method:

(1) **ID** of a node i, referred to in shorthand as "id," is a node value assigned based on the order a node appears in the text. This is not a centrality measure per se but allows us to track when nodes are "earlier" or "later" in the text, as well as allows us to identify particular nodes for discussion.

(2) **In-Degree Centrality** of a node i, referred to in shorthand as "ideg," is the number of nodes which connect to node i in proportion to the rest of the network. In context, this is how many other statements are cited by node i proportional to the possible number of citations.

(3) **Out-Degree Centrality** of a node i, referred to in shorthand as "odeg," is the proportion of nodes to which i connects to in proportion to the rest of the network. In context, this is how many other statements cite to node i compared to the possible number of citations.

(4) **Betweenness Centrality** of a node i, referred to in shorthand as "btwn," is the proportion of shortest paths between nodes on the network which pass through i. In context, this is how many "efficient proofs," or short lines of citation, use i, in proportion to the total number of short paths.

(5) **Closeness Centrality** of a node i, referred to in shorthand as "clos," is the reciprocal of the sum of i's distance to all other nodes. In context, this expresses the "relevance" of node i to other nodes, and is higher when more short lines of proof exist between other nodes (networkx counts "in" paths).

(6) **Katz Centrality** of a node i, referred to in shorthand as "katz," is an iterative measure which weights a node based on the importance of its neighbors. In context, a node will have high Katz centrality if it uses many

other “important” nodes to support its claims. *NB. Katz Centrality adds “self-importance” to all nodes at each iteration, which its simpler counterpart, Eigenvector Centrality, does not. We chose not to use Eigenvector Centrality because importance can “drain” out of directed networks using this measure, which happens here; networkx fails to converge this measure on the network after 100 iterations.*

We consider all these centrality measures because it is not immediately evident which conceptual understanding of importance will most accurately identify central nodes on the network. Instead, we consider the correlation between each of these measures to identify a benchmark centrality measure which may indicate multiple instances of centrality. Figure 2 shows the correlation between each measure.

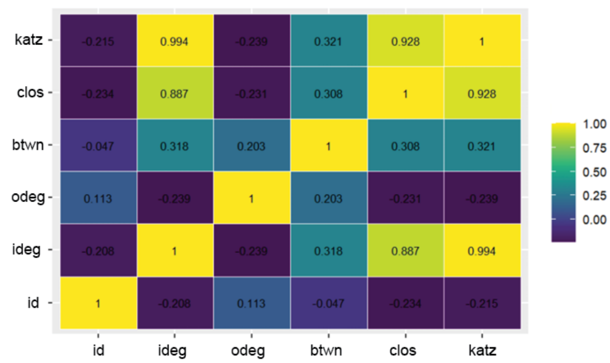


Figure 2: Heatmap demonstrating the correlation between vectors of values for each centrality measure.

Several interesting characteristics emerge in the correlation coefficients between the relevant centrality measures. The first point of interest is that ID, or chronological order of the nodes in the text, is negatively associated with most of the centrality measures. This means that nodes which are later in the text tend in general to be associated with lower centrality values for measures like in-degree, betweenness, closeness, and Katz centrality. The correlation with out-degree, on the other hand, is low but still positive. It makes sense that the number of other statements a statement cites would not change very strongly over the chronological arrangement of the text. Overall, this also indicates that earlier statements are in general associated with higher centrality scores by the measures we consider than are later statements.

#### 4.1.1 Node-Level Anaylsis

Next, we attempt to identify a benchmark centrality measure as discussed above. There are no universal indicators of centrality; each centrality measure has at least one other measure with which it is negatively correlated. Nonetheless, one measure stands out: Katz centrality is highly correlated with closeness centrality (with a coefficient of 0.928) and in-degree centrality (with a coefficient of 0.994). A node which cites other important statements is thus in general associated with short distances to other nodes in the proof structure, and with a high proportion of citations from other nodes. The top 10 nodes by Katz centrality are shown in Table 2.

| Rank | Name | ID | katz | ideg | odeg | clos | btwn |
|---|---|---|---|---|---|---|---|
| 1 | 3P11s | 344 | 1.000 | 1.000 | 0.588 | 1.000 | 0.877 |
| 2 | 3P7P | 333 | 0.974 | 0.926 | 0.000 | 0.000 | 0.795 |
| 3 | 3P13s | 350 | 0.886 | 0.815 | 0.000 | 0.000 | 0.693 |
| 4 | 3P3P | 324 | 0.866 | 0.8.15 | 0.000 | 0.000 | 0.714 |
| 5 | 2P17P | 219 | 0.796 | 0.593 | 0.000 | 0.000 | 0.912 |
| 6 | 1P15P | 58 | 0.787 | 0.666 | 0.000 | 0.000 | 0.664 |
| 7 | 3P27P | 387 | 0.752 | 0.630 | 0.000 | 0.000 | 0.505 |
| 8 | 2P11c | 174 | 0.711 | 0.593 | 0.000 | 0.000 | 0.5194 |
| 9 | 3P28P | 397 | 0.698 | 0.556 | 0.000 | 0.000 | 0.505 |
| 10 | 2P13P | 179 | 0.685 | 0.556 | 0.000 | 0.000 | 0.462 |

Table 2: Top 10 ranked nodes by Katz centrality, shown with other centrality measures for those nodes. Each centrality measure has been scaled by a constant such that the highest possible score for each measure is 1.000.

Interestingly, these nodes share some sites of thematic overlap: they explore how the mind relates to the body, how we are impacted by emotions, and how we strive or "endeavor" to preserve our being via a Spinozist construct called the "conatus." This overlap becomes evident while considering these terms in the (translated) text of each high scoring statement. The word "mind" appears in (1, 4, 5, 8, 10), The word "body" appears in (1, 5, 10), the word "emotion" or reference to a specific Spinozist emotion appear in (1,3,7,9), and some conjugation of the word "endeavor" appears in (2,3,9). This is not a complete redundancy between the propositions, but the commonality and points of overlap suggest the relevance of these themes. Spinoza worked in the shadow of the philosophical giant who preceded him, Reneé Descartes, whose theory of the mind/body relationship Spinoza vehemently rejected. Although many of these central nodes are not necessarily the most widely recognizable (with the notable exception of 3P7P, in which he famously defines the "conatus"), their centrality in the proof structure makes sense when we consider their relevance to debates about the mind/body problem. Again, since this network does not account for the text of the propositions, the kinds of " importance" it encodes are not necessarily Spinoza's flashiest arguments. Instead, centrality measurement brings to the forefront those nodes which are actually most *supported* or *supporting* via citation in the text. This debate about the mind/body problem was extremely relevant to Spinoza's philosophical climate, and his stance ran counter to a widely popular interpretation of his time. So, it makes sense that the supportive infrastructure of the text would centralize those nodes for which his readers would most demand a strong proof.

Among these top 10 nodes, the positive correlation between Katz centrality and in-degree centrality seems to intuitively hold, as does the negative correlation between Katz centrality and out-degree. Zero out-degree is common for nodes of the "Proposition" type, since the support for propositions is provided in their associated proofs (d) or additional proofs (a), rather than in the propositions themselves. The majority of nodes in the top ten for Katz centrality are propositions.

However, the strong positive correlation between Katz centrality and closeness centrality given on the complete vectors of centrality measures is not echoed in this subset of centrality measures. In fact, even though 3P11s is the top ranked node for both Katz and closeness (as well as in-degree), every other node in the list has a scaled closeness centrality of 0.000.

### 4.1.2 Type-Level Analysis

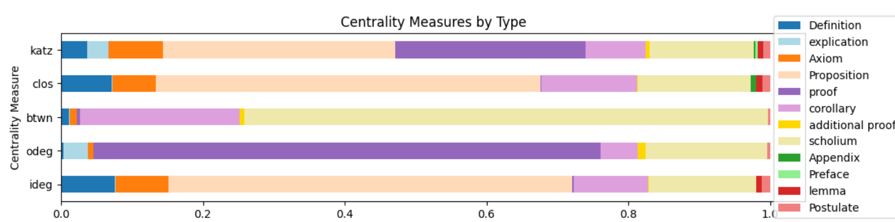


Figure 3: Proportion of total centrality scores occupied by each statement type

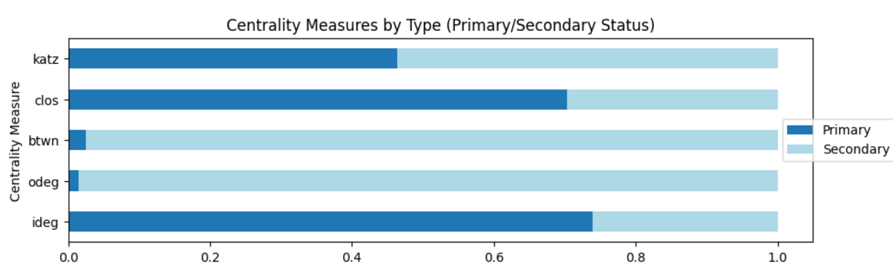


Figure 4: Proportion of total centrality scores occupied by each statement category (primary vs. secondary

An analysis of centrality measures by type provides a more global view of what kinds of statements are integral to the proof structure in what ways. Many of the measures are dominated, unsurprisingy by Propositions, which are the backbone of the text; interestingly, the next most dominant groups are secondary (proofs, scholia, and corollaries). Figure 4 highlights the relative kinds of importance of the two categores. Primary porpositons dominate Katz centrality, closeness centrality, and in-degree centrality; they have important neighbors in dense subproofs and are often referenced. Secondary statements fulfill a complementary role; they "connect the dots" between other statements via betweenness and hold the vast majority of citations via out-degree.

### 4.2 Community Detection

Given that we are studying the structure of a proof network, we were curious what the community structures the citations admit. Does the author build statements to form an "ultimate proof" such that we observe a giant community, or are the proofs more distinct objects and we would observe more smaller communities? Since the book is divided into 5 parts, we were also curious to see if this was a logical division; if the partition corresponds to groupings in the proof, or exists for reasons outside the proof structure.

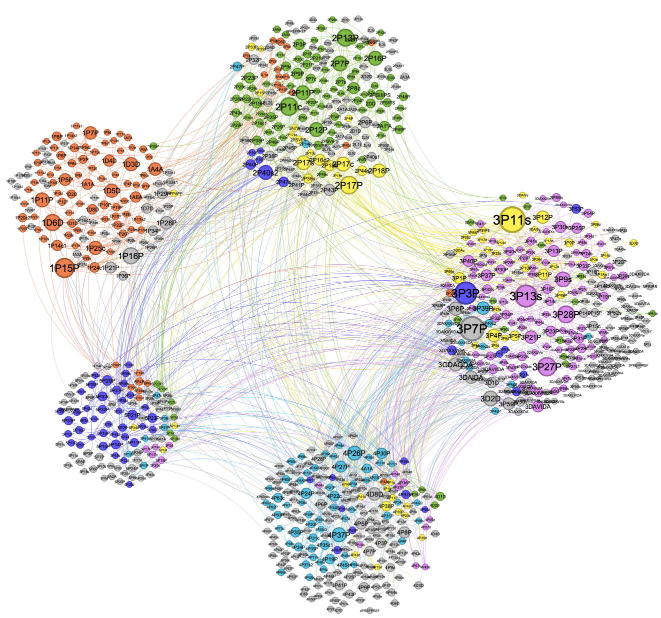


Figure 5: Network clusters represents the different parts of the book, colored by louvain communities, and node sized by in-degree.

Looking at Figure 5, we can observe the largest 6 communities that resulted from the Louvain community detection algorithm. From the visualization, we see that the larger communities are pretty large in size. This indicates that the statements perhaps build off of each other either directly or directly, or that they share common neighbors. In looking at the overall clusters, there is a clear "main" color for each of the clusters. This indicates that the partition of the book does play as role in the community structure and this makes sense especially when considering that the parts of the books do have an overall theme.

To quantitatively determine the quality of the split and whether the author have used a different partition, we compared modularity scores. In particular, the modularity score if communities were determined solely based on the parts of the book, against the modularity score for communities based off of the *greedy_modularity_communities*, *louvain_communities*, and *asyn_lpa_communities* algorithm implemented by NetworkX.

| Community detection algorithm | Modularity Score | number of communities |
|---|---|---|
| *greedy_modularity_communities* | 0.6725180401943106 | about 158 |
| *louvain_communities* | 0.6703023561823478 | about 160 |
| *asyn_lpa_communities* | 0.31550716333513457 | about 459 |
| community by parts | 0.41979932063582304 | 5 |

Table 3: Comparing the results of various community partitions.

Judging based on modularity scores (given in Table 3) our partition would be "better" if we split up the statements based on *greedy_modularity_communities* and *louvain_communities*. However, we can also notice that with just 5 communities, the original partition does considerably well, especially compared to the *asyn_lpa_communities* algorithm. To give Spinoza additional credit, it wouldn't have made much sense to split the book into more than 150 parts for the sake of something called "modularity". It seems that Spinoza's partition, which he justifies with the different subject matter the sections addresse, does correspond to at least some measure of association between the statements of that part in the proof. The fact that this score is not higher indicates that significant connections between the parts exist. This makes sense, since he continually recapitulates and refers to some consistent concepts throughout the text.

### 4.3 Are any Behaviors Power Law-like?

Even though we have a small network, we were curious if the network might show any power law-like behaviors: That is, whether trends in the distribution of some attributes might indicate the existence of "hubs" or a high number of low-stakes nodes. We suspect that to build the proofs throughout the entire book, Spinoza inherently relies on a couple of fundamental statements.

This was an observation Badiou echoed in his, where he saw that certain statements are used repeatedly: "we noticed that Definitions 3 and 5 are used repeatedly: four times in the skeleton of one proposition!" (46, Badiou).

### 4.3.1 Degree

We began by looking that the in and out degree distribution of our statements. The in-degree of node $i$ tells us how many times the statement $i$ has been cited by other nodes and can be an indication for the "fundamentalness" of the statement. The out-degree of node $i$ tells us how many nodes the statement $i$ cites and can be an indication for the complexity of the proof. Figure 6 shows our in and out-degree distributions and the results of our power law fit.

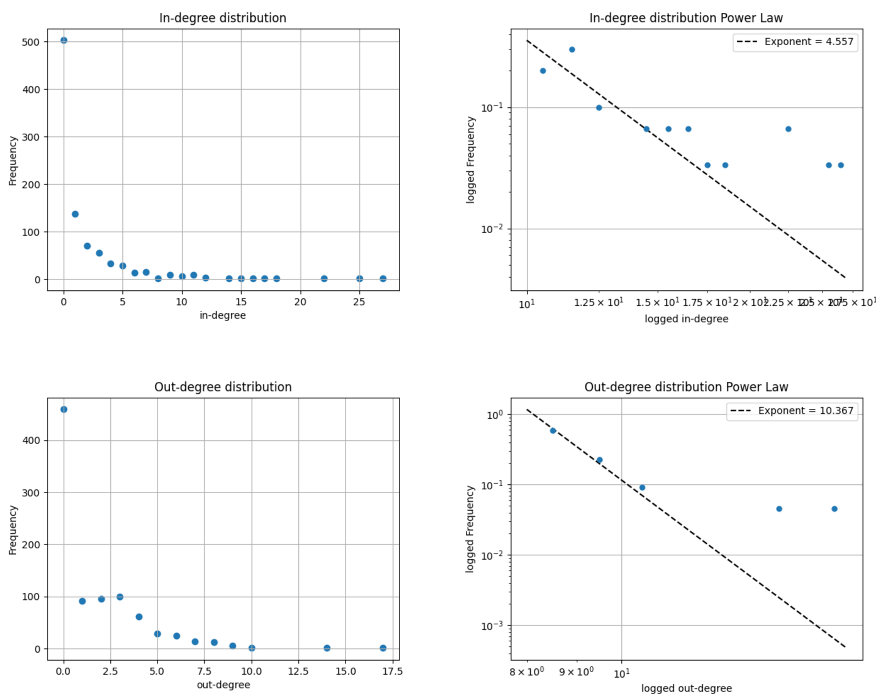


Figure 6: Plots of in and out-degree distribution with their corresponding power law fit

The resulting alpha values for the power law fit are not within the range $2 < \alpha < 3$. The range of values for both in and out-degree are very close to each other, and we don't have a large tail distribution. We suspect that this might be due to the fact that this is a proof network where nodes build off of each other. Essentially, even if node $a$ is a very fundamental statement, if node $b$ is a more refined version of node $a$, later nodes will call on node $b$, there will be a more refined version of node $b$, and so on. By looking at just in and out degrees, we aren't capturing the indirect influences of a node throughout the whole network.

### 4.3.2 Community Size

The next aspect we turn towards for power law behavior was in community sizes, particularly the community sizes that result from the *greedy_modularity_communities*, *louvain_communities*, and *asyn_lpa_communities* algorithms. From looking at the visualization for community detection from section 4.2, we suspected that there were some large communities and many more small communities. We observe that this was the case, but attempting o fit the size distributions to a power law curve once again does not admit an alpha value in the range $2 < \alpha < 3$. Figure 7 demonstrates this. Our large communities aren't dramatically bigger in size than the smaller communities, and the distributions were also more evenly spread than with a natural power law.

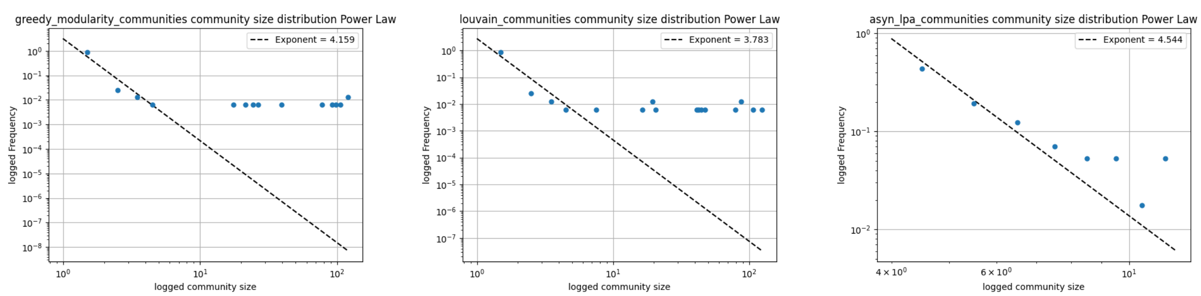


Figure 7: Power law fit of community size distributions from *greedy_modularity_communities*, *louvain_communities*, and *asyn_lpa_communities* algorithms

### 4.3.3 Skeletons

As mentioned previously, the skeleton of a statement is essentially a tree of all the "ancestors" of the statement that allows for the statement to hold. Compared to just the out-degrees which gives us the statement's "parents", the skeleton allows us to get a sense of all statements that has an influence. The skeleton also can serve as a measure for the complexity of a proof, in particular the number of nodes in the skeleton and the depth of the skeleton. The number of nodes in the skeleton gives us a sense of statements the proof relies on and the depth of the skeleton indicates the levels of proof necessary to make this point.

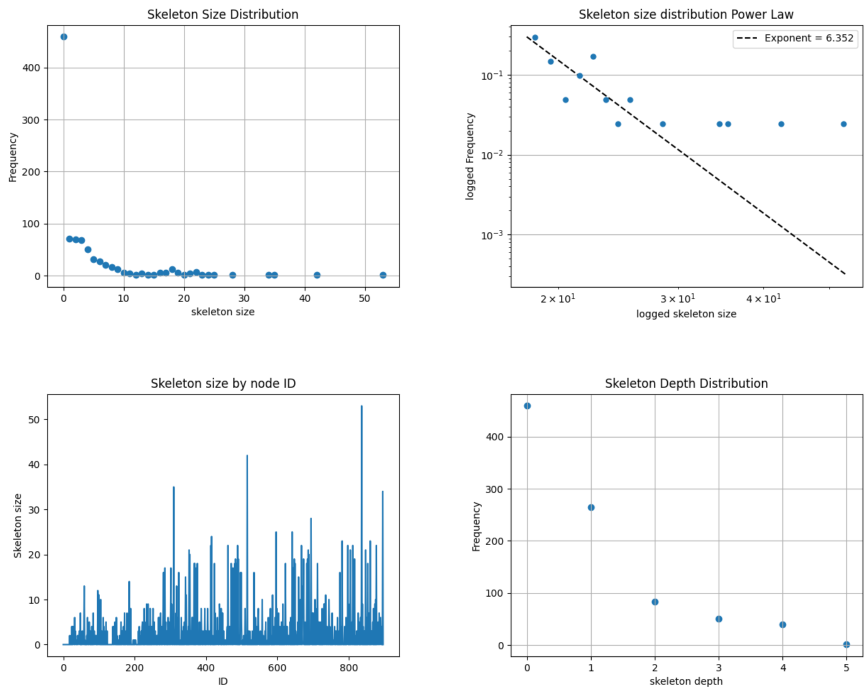


Figure 8: Distribution of skeleton sizes in our network, the resulting power law fit, and the distribution of skeleton depths

Looking that Figure 8, we can get a good sense of the proofs in our network. In particular we can observe that there are many stand along proofs, with 0 skeleton size and depth, which are probably definitions and axioms. We can also notice that skeletons are bigger (not plotted but is relatively longer as well) for proofs at the end of the book, which is reasonable.

#### 4.3.4 Skeleton Uses

The in-degree distribution does not account for the context that nodes build off of each other and the indirect influence doesn't get counted. We tried to see if measuring "skeletons" as Badiou described would account for this dynamic. We define "skeleton use" for a node $i$ to be the total number of times the node $i$ shows up every skeleton. The measure, skeleton use, allows us to capture the influence of the node both directly and indirectly.

Using the analogy of node $a$ and node $b$ from previously, node $a$'s skeleton use will increase every time node $b$ is used. If node $c$ cites node $b$, node $a$'s skeleton use will also increase when node $c$ is cited.

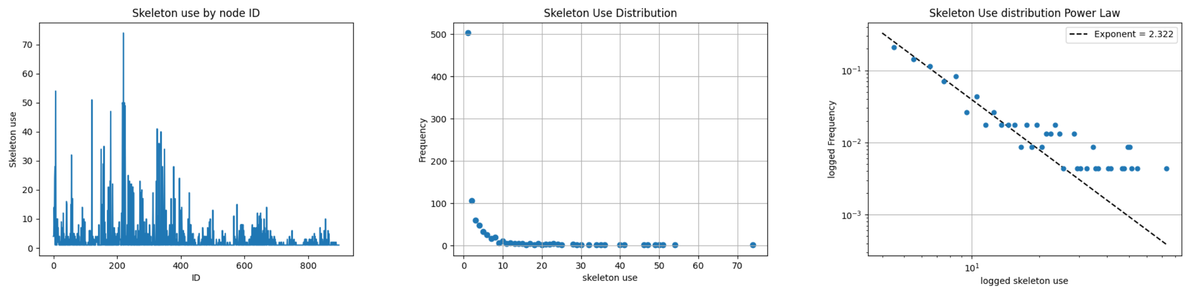


Figure 9: Distribution of skeleton uses in our network and the resulting power law fit

Figure 9, shows the distribution of skeleton uses in our network and the resulting power law fit. Here, we finally observe a natural power law distribution with $2 < \alpha < 3$. This is significant because we can understand the proof network to rely on a couple of fundamental statements, where the proofs are more connected than just separate.

### 4.4 Comparing the Network to Price's Model

Because we posit that the proof structure is comparable to a network of citations, we were interested in seeing if we could modify Price's model for citation networks to echo our proof structure. We began by choosing our values for $a$ and $c$ based on out network. $a$ is a tuning parameter; in the traditional Price's model, c gives a uniform out-degree to each node. However, we modified this model to allow for variable out-degree. Since our maximum out-degree is 17, we choose $c = 17$. Then, the model gives the ratio between $c$ and $a$ is the proportion of edges that are added based on popularity versus randomly. We chose the tuning parameter $a = 35$[1]. The edges would be slightly more random because our network doesn't really have large hubs.

Then, we add to our model the condition that each edge is added with probability $P = 0.10$. This is because while the maximum out-degree is 17, we don't want every node to have an out-degree of 17. In our network, the average out-degree is about 1.58, and thus, if we implement $P = 0.10$, our expected out-degree of each node would be 1.7.

Taking the randomly generated network, we removed "unrealistic" edges. First, we removed the fulled connected edges in our first $c$ nodes because they impose that early statements can cite later statements and also because the fully connectedness ends up messing with our skeletons. Then, we removed all edges that are "unrealistic" for the proof. An edge is considered as unrealistic if

[1] We also tried choosing $a$ based on our power law exponents using this property: $p_k = P(deg(v) = k) \sim k^{-2-a/c}$ , but the results are about the same.

the edge directly connects two nodes that are already indirectly connected, i.e., already justified through that line of proof.

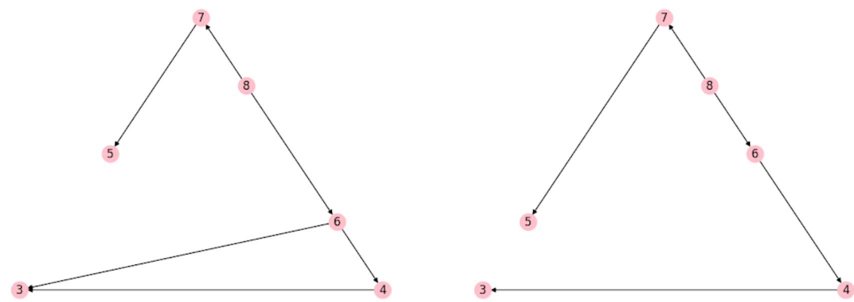


Figure 10: An example of edge removal of an “unrealistic” edge.

In Figure 10, the edge from 6 to 3 is considered unrealistic because 6 can reach 3 indirectly through 4. This edge removal makes simplifications in that it is possible that node 6 uses a different aspect of node 3 that node 4 doesn’t use. But this edge removal process also tones down the ”popularity” property (where there are many nodes connected to an earlier node even though there are indirect paths), from the Price’s model that doesn’t occur in our network.

Figure 11 gives a visualization for the different properties of a random network we produced using this altered Price’s model.

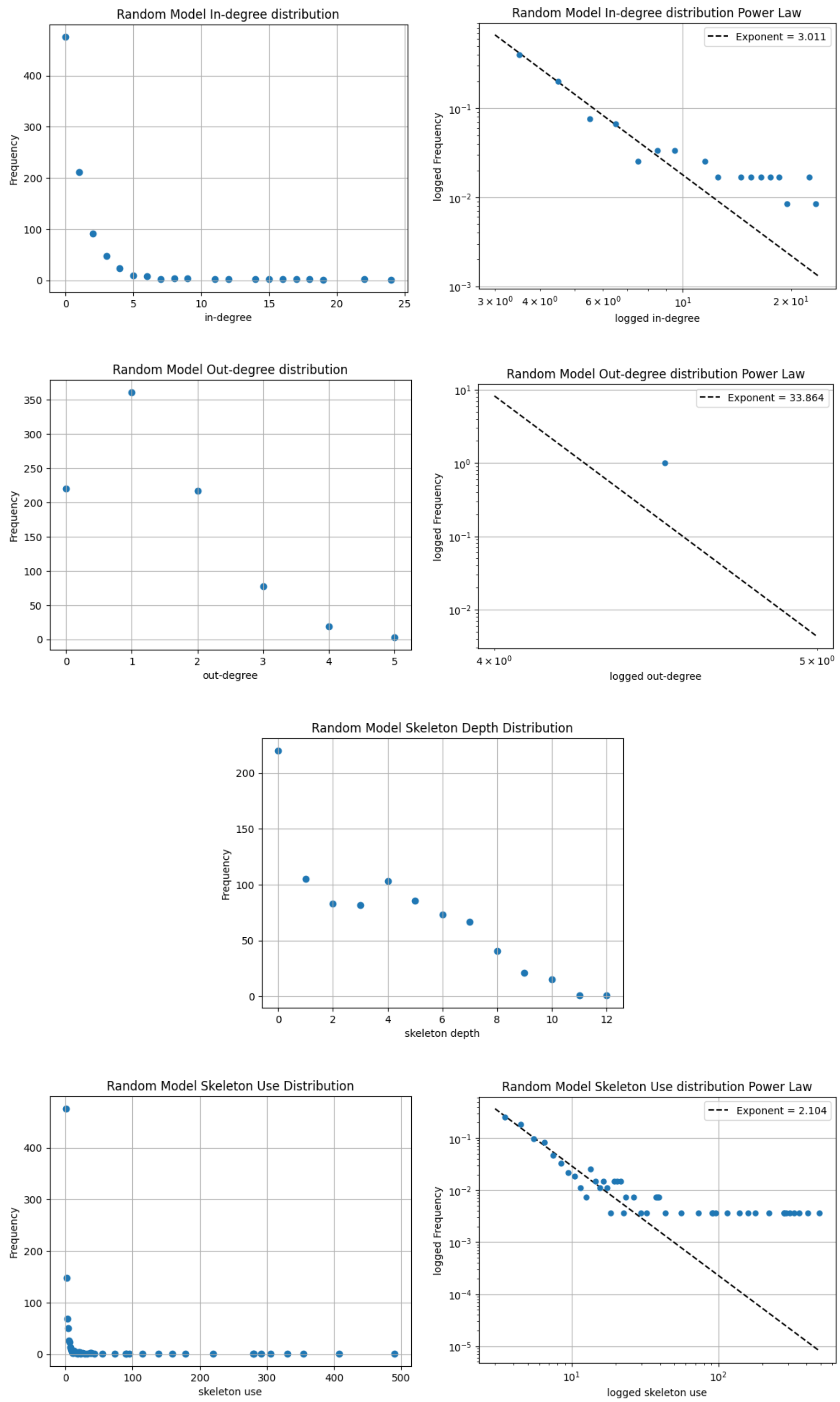


Figure 11: Plots of in and out-degree distribution with their corresponding power law fit; skeleton depth distribution; and skeleton use distribution with power law fit for a Random Network created using altered Prices Model

Comparing each visualization to the corresponding plot of our original network, we can observe that there are definitely similarities in terms of general trends. Still, however, we can notice if we look closer that some of the scales are slightly off and that the fitted power law exponents are quite different. We are definitely not able to capture all aspects of Spinoza's proof structure, which is pretty reasonable.

Though probably not significant in terms of the value that is added, more work could look into the calibration of parameters for this model: $c$, $a$ and $P$, or a completely different model that replicates Spinoza's proof structure better.

## 5 Conclusion

This project is intended to suggest potential areas of exploration for both Spinoza scholars and mathematicians interested in applying network science to logical proofs. Several possibilities for future work stand out. Firstly, although we interpret the network as citation-network-like, others who consider this question such as Alain Badiou understand dynamics in the network to more closely reflect genealogical behavior. An analysis of the reverse of our graph would help explore this alternate interpretation. Furthermore, we were interested in exploring power law-like behavior in order to flag the possibility of hub nodes on the network. A more detailed attempt to best fit the relevant distributions, especially one which examines possibilities beyond just power law behavior, could reveal more interesting aspects of the network. Finally, we might hope to put this mathematical approach in conversation with Spinozist philosophers. One of the results of this inquiry which most stood out was the difference in which nodes were important in *proof* compared to which nodes are most widely recognizable in the world of philosophy. To develop a finer picture of this point, we might want to ask Spinoza scholars which statements they consider to be the most "important" in the text, and analyze the role of these nodes as they exist embedded in the proof. This enrichment of the project's context could also involve a text analysis which encodes proximity of word uses, as in Andrew Beveridge's analysis of Game of Thrones in "Network of Thrones."

Many surprising characteristics emerged in the network from a diverse range of network science tools. Indeed, a quantitative analysis of the proof structure reveals many characteristics which would not have otherwise been evident. Centrality measurement does not necessarily promote those nodes which are well-known; instead, it reveals a possible motivation for proof structure, in which citation legitimizes the most actively controversial of Spinoza's claims. Then, we compared the strength and usefulness of Spinoza's natural partition of the text into its five parts, particularly in sharp relief to various community detection algorithms. Then, we attempted to detect power law-like behavior among various networks of the text, and uncovered meaningful trends in statements' accumulated influence, measured with "skeleton uses." Finally, we adjusted a popular model for citation networks, Price's model, to reflect behaviors we identified in Spinoza's network. Spinoza structured the book with intention and care; we advocate that this structure is incredibly relevant to a deep understanding of the text.

## 6 Works Cited


Badiou, Alain. 'What Is a Proof in Spinoza's Ethics?', in Dimitris Vardoulakis (ed.), *Spinoza Now* (Minneapolis, MN, 2011; online edn, Minnesota Scholarship Online, 24 Aug. 2015), https://doi.org/ 10.5749/minnesota/9780816672806.003.0002.

Beveridge, Andrew. *Network of Thrones,* 14 Jan. 2020, networkofthrones.wordpress.com/.

Doppelt, Torin. *Digital Spinozism*, digital.spinozism.org/.

Hobbes, Thomas. *Leviathan*, www.gutenberg.org/files/3207/3207-h/3207-h.htm.

Hsu, Chih-Ling. "Random Graph Models." *An Explorer of Things*, chih-ling-hsu.github.io/2020/05/15/Graph-Models.

Lord, Beth. *Spinoza's Ethics: An Edinburgh Philosophical Guide (Edinburgh Philosophical Guides Series).* Edinburgh University Press, 2010.

Negri, Antonio. *The Savage Anomaly: The Power of Spinoza's Metaphysics and Politics.* University of Minnesota Press, 2016.

Spinoza, Baruch. *The Ethics; Treatise on the Emendation of the Intellect; Selected Letters.* Translated by Samuel Shirley, Hackett Publishing Company, Inc.