Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025 ASEE
Paper ID #48858

Are Engineering Degrees Really More Complex? Characterizing the Complexities of Academic Programs by Discipline

Prof. Gregory L. Heileman, The University of Arizona

Gregory (Greg) L. Heileman currently serves as the Associate Vice Provost for Academic Administration and Professor of Electrical and Computer Engineering at the University of Arizona, where he is responsible for facilitating collaboration across campus t

Prof. Chaouki T Abdallah, Georgia Institute of Technology

Chaouki T. Abdallah started his college education at the Ecole SupÃ©rieure d'IngÃ©nieurs de Beyrouth - UniversitÃ© Saint-Joseph in Beirut, Lebanon, but finished his undergraduate studies at Youngstown State University, with a Bachelors of Engineering degr

Kristina A Manasil, The University of Arizona

Kristi Manasil is a second-year PhD student in the College of Information Science at the University of Arizona. She received her bachelor's degree in Computer Science from the University of Arizona. Her areas of focus are data visualization, machine learning, learning analytics and educational data mining.

Melika Akbarsharifi, The University of Arizona

Melika Akbarsharifi is a PhD candidate in Software Engineering at the University of Arizona, studying under Professor Gregory L. Heileman. Her research at the Curricular Analytics Lab focuses on developing advanced computational methods and machine learning approaches to optimize educational pathways and student success.

Roxana Akbarsharifi, The University of Arizona

Roxana Akbarsharifi is a PhD student in Software Engineering at the University of Arizona. Her research focuses on educational analytics and developing tools to improve student outcomes and support academic success. Her research interests include software engineering, machine learning, data analytics, and data visualization, with an emphasis on applying these technologies to solve educational challenges and enable data-driven decision making in higher education.

Are Engineering Degrees Really More Complex? Characterizing the Complexities of Academic Programs by Discipline

Gregory L. Heileman,[†] Chaouki T. Abdallah,[•] Kristi Manasil,[⋆]
Melika Akbarsharifi,[†] Roxana Akbarsharifi,[†] and Ahmad Slim[‡]

heileman@arizona.edu, chaouki.abdallah@lau.edu.lb, kmanasil@arizona.edu,
akbarsharifi@arizona.edu, roxana@arizona.edu, and ahslim@unm.edu

[†]Department of Electrical & Computer Engineering
[⋆]College of Information Science
University of Arizona

[•]Department of Electrical & Computer Engineering
Lebanese American University

[‡]Department of Electrical & Computer Engineering
University of New Mexico

Abstract

In this paper we consider how curricular complexity varies according to academic discipline, both broadly across the wide variety of fields in higher education, and more narrowly within the engineering discipline and specific engineering sub-disciplines. This work involved collecting and analyzing all of the undergraduate curricula associated with all of the undergraduate programs at thirty different universities. The universities involved in this study included a broad cross-section of R1 and R2 institutions, as well as two HBCUs, located in sixteen different states. We first describe the curricular complexity metric used in this study. Next, we consider all curricula across all of the institutions involved in this study, and we show that most curricula are distributed at the lower end of the complexity scale with relatively fewer at the higher end of this scale; moreover, this distribution has a long right-skewed tail. Furthermore, many of the high-complexity programs in the tail correspond to engineering programs. Given the long tail associated with the distribution of data in our study, the average curricular complexity value is easily skewed by a small

number of highly complex curricula. Thus, any statements made regarding the average curricular complexity value at a particular institution, or across any collection of institutions, when all disciplines are considered, should be considered highly unreliable. However, if we instead disaggregate programs according to discipline, the long tail behavior is significantly moderated, and interesting distributions emerge. By characterizing the features of these distributions, it becomes possible to make quantitative and comparative statements about the complexities of particular disciplines, including engineering. By further disaggregating according to engineering sub-disciplines, we obtain distributions that resemble well known distributions, such as the gamma distribution. This provides a means, for the first time, to meaningfully compare and contrast the curricular complexity of the engineering field to those in other fields, as well as the complexities of the sub-disciplines within the engineering field, e.g., civil engineering, chemical engineering, electrical engineering, etc.

Introduction

A fundamental question related to student success is, are some academic programs inherently more difficult to complete than others? If so, then additional questions immediately follow. For instance, can we quantify the differences between academic program difficulties? The ability to do so would allow us to formalize common beliefs that are often stated without firm factual bases. For example, it is not uncommon for students to hear advice regarding the difficultly of engineering programs; such as, engineering programs are the hardest programs on campus to finish on time. In this paper we provide a framework for addressing these questions that uses a formal method for measuring program curricular complexity, applied to a large corpus of curricular data collected from thirty different universities. We demonstrate that it is indeed possible to quantify the complexity differences that exist between the different academic program in our data set. Furthermore, we show that it is possible to characterize these differences in statistically meaningful ways; that is, in a manner that we believe should prove useful in guiding curricular design and reform efforts aimed at facilitating student success.

Background

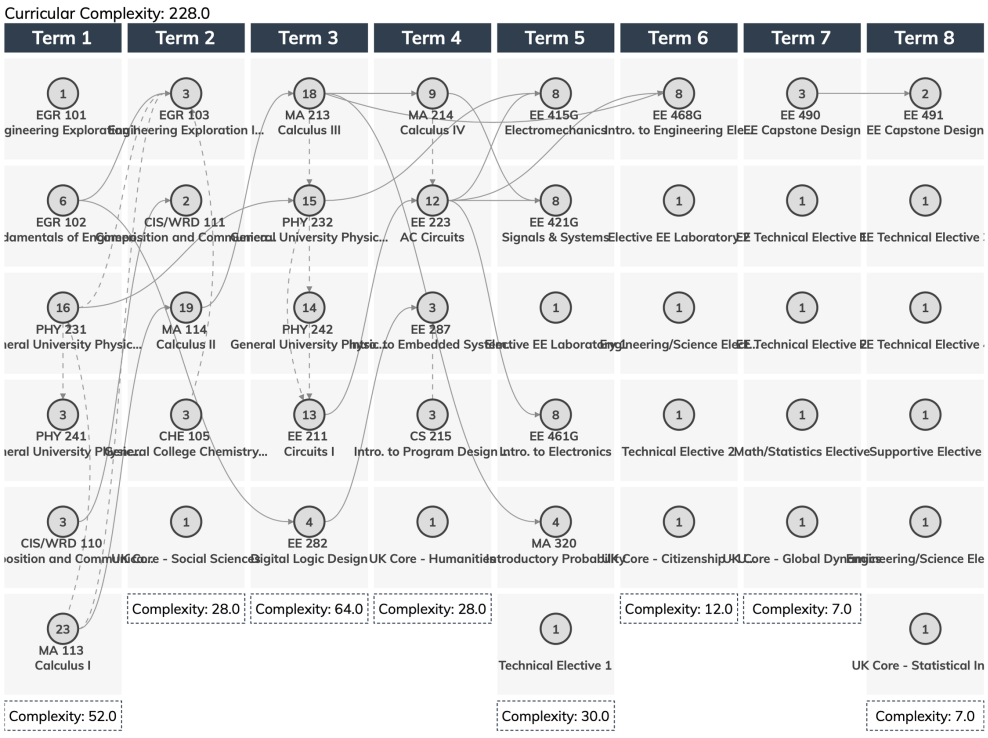
For this study, a data set consisting of all the undergraduate curricula at thirty one different universities across the United States was collected and analyzed. These institutions are all members of the Undergraduate Education at Research Universities (UERU) organization, the entity that managed this study. The universities involved in this study are all public, with the exception of one private institution, and all have Pell grant-recipient student populations that account for at least 30% of their overall student populations. The institutional types also included flagship, land-grant, urban-serving, HBCU, HSI, R1, and R2 universities. Each university participant uploaded the curricula associated with each of their undergraduate academic programs to the website http://CurricularAnalytics.org. The total number of curricula collected, across all institutions (accounting for degree concentrations/emphases) was 3,830.

In this study, a *curriculum* refers to the set of courses (along with the corresponding set of course prerequisites) that, if successfully completed, would allow a student to earn the degree associated with the curriculum. An example electrical engineering curriculum is provided in Figure 1 (a). This curriculum is represented as a graph, where the vertices are the required courses in the curriculum, and the directed edges (arrows) between the vertices correspond to prerequisites. That is, the course on the source end of the directed edge is a prerequisite for the course on the destination end. Directed edges drawn as dashed lines correspond to co-requisites. The complexity of each curriculum was computed using a unitless graph-theoretic metric imposed by the pre- and co-requisite relationships between the courses in a curriculum. This metric, referred to as *structural complexity*, involves two factors. First, each course c in a curriculum is assigned a *blocking factor* which is simply the number of courses a student is precluded from taking, due to pre- and co-requisite constraints, until they have successfully completed course c. In Figure 1 (b), the blocking factor of the Calculus I course is 15. The second factor, called the *delay factor* is determined by the longest pathway in the graph that includes course c. In Figure 1 (b), the delay factor of the Calculus I course is 8. The structural complexity of a course c is determined by adding its blocking and delay factors, and the structural complexity (or more simply complexity) of a curriculum is determined by summing all the course complexities in a curriculum. In Figure 1 (b), the complexity of the Calculus I course is 23, and the complexity of the entire curriculum is 228. It can be shown that under certain mild assumptions, on average, curricula with higher complexity take more time (i.e., are more difficult) for students to complete.[1]

In order to consider how the complexities of academic programs vary across and within different fields of study, each academic program considered in this study was classified by the universities involved in the study using Classification of Instructional Programs (CIP) codes. Every postsecondary school in the U.S. receiving federal student financial aid is required to match their academic programs to CIP codes, and to periodically report specific program data to the federal government.[2]

The CIP framework is a taxonomic coding scheme for classifying instructional programs organized around three levels:

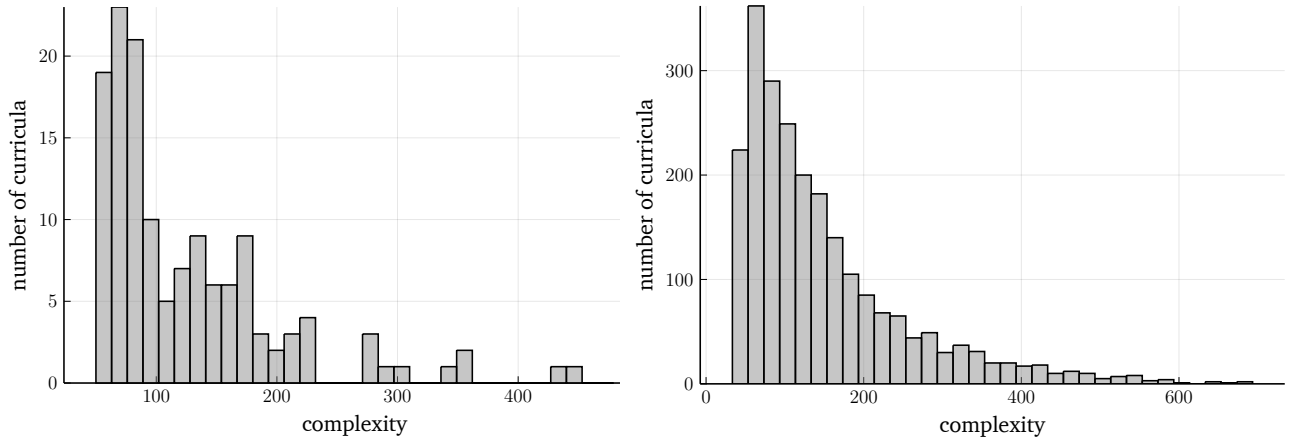
- **Two-digit series:** represents the most general grouping of programs. The standard format for representing a two-digit CIP is as ##., where ## is a number between 01 and 99. E.g., 04. = Architecture and Related Services; 13. = Education; 14. = Engineering; and 26. = Biological and Biomedical Sciences.
- **Four-digit series:** represents an intermediate grouping of programs that have comparable content and objectives. The standard format for representing a four-digit CIP is as ##.##., where ## is a number between 01 and 99. E.g., 04.04. = Landscape Architecture, Environmental Design; 13.03. = Education, Curriculum and Instruction; 14.09. = Engineering, Computer Engineering; and 26.03. = Biology, Botany/Plant Biology.
- **Six-digit series:** represents a detailed grouping of specific instructional programs. The standard format for representing a four-digit CIP is as ##.####, where ## is a



(a)

(b)

Figure 1: **(a)** An example electrical engineering program curriculum, organized as a degree plan over eight terms. The courses in the curriculum are shown as vertices, and the prerequisites are shown as directed edges. **(b)** Highlighting the Calculus I course in this curriculum shows that Calculus I blocks 15 other courses in the curriculum (shown in green), and the longest path in the curriculum that includes Calculus I (shown as a blue dashed line) has length 8.



(a) (b)

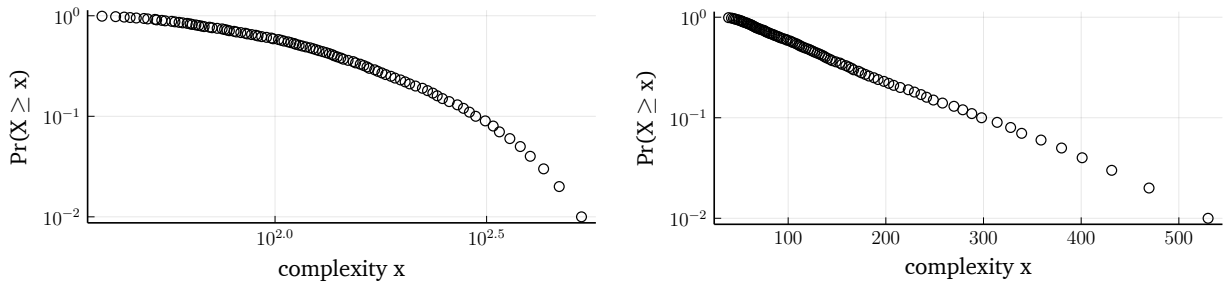
Figure 2: Histogram of program complexities for (a) all of the undergraduate programs at the University of Arizona, and (b) all of undergraduate programs at all institutions involved in the study.

number between 01 and 99, and #### is a number between 0000 and 9999. E.g., 04.0403 = Landscape Architecture, Environmental Design, Sustainable Design/Architecture; 13.0301 = Education, Curriculum and Instruction, Curriculum and Instruction; 14.0903 = Engineering, Computer Engineering, Computer Software Engineering; and 26.0307 = Biology, Botany/Plant Biology, Plant Physiology.

Below, we use the term *field of study*, or more simply *field*, synonymously with the those areas of study identified by the two-digit CIP series, and we use the term *discipline* to refer to the areas of study identified by the four-digit CIP series. Finally, we use the term *sub-discipline* to refer to those areas of study identified by the six-digit CIP series That is, a field refers to a broad area of study encompassing many disciplines, and disciplines may contain many sub-disciplines. For instance, the engineering field contains the chemical engineering, civil engineering, computer engineering, etc. disciplines, and the computer engineering discipline contains sub-disciplines such as computer hardware engineering, computer software engineering, etc.

Complexity Distribution – All Curricula

Next we consider the various complexity distributions that emerge by treating the complexity values in the data set as a random variable. First, we consider the distribution of complexity values across all academic programs. A very similar distribution of program curricular complexities is found at each of the institution involved in this study. Specifically, universities tend to offer many programs at the lower end of the complexity scale, and relatively fewer programs with higher complexity. A typical example from the University of Arizona is shown in the histogram provided in Figure 2 (a), where there are numerous programs with complexities in the 50–250 range, and far fewer with complexities above 250.



(a) (b)

Figure 3: The empirical distribution (as a complementary CDF) $Pr(X \geq x)$ for the entire data set plotted on (a) a doubly logarithmic scale, and (b) a semi-logarithmic logarithmic scale.

Figure 2 (b) shows the complexity histogram created by combining all of the curricular data from all of the institutions involved in this study. This histogram resembles the *power-law distribution*, an important probability distribution that has attracted significant scientific interest, given its interesting properties, and its ubiquitous nature, arising in so many different areas of study, including physics, biology, computer science, and the social sciences, just to name a few.[3,4] The power law distribution is characterized by a long tail, i.e., most samples will cluster around a smaller value, with occasional "black swan" events leading to samples in the tail of the distribution that are far away from this cluster. The density function of a power law distribution has the form

$$p(x) = Cx^{-\alpha},$$

with exponent $\alpha > 1$ (referred to as the scale parameter), where C is a constant, and it is assumed $x > x_{min} > 0$; that is, the power law relationship only holds above the value x_{min}. An important and unique characteristic of the power law distribution is its *scale invariance*; that is, scaling x by a constant factor a will not change the fundamental shape of the distribution,

$$p(ax) = C(ax)^{-\alpha} = a^{-\alpha}p(x) \propto x^{-\alpha}.$$

This agrees with the observation that the complexity distributions at all of the individual universities resemble to some extent Figure 2 (b).

Given that

$$\log(p(x)) = \alpha \log x + C,$$

where C is a constant. It follows that one way of detecting the possible presence of a power law distribution is to plot samples drawn from a distribution on a log-log scale. If the these samples fall on a line, there is evidence (a necessary but not sufficient condition) of the power law distribution. Furthermore, the slope of this line provides an estimate of the value α. A log-log plot of the entire curricular data set is shown in Figure 3 (a), and a best fit line to these data points has slope -1.08. However, we should mention that the detection of the power law in empirical data is a notoriously difficult problem, due to many factors that influence real-world data sets.[5] For instance, the fact that the data points in Figure 3 (a)

deviate from a line in the extreme tail of the empirical distribution may be attributed to the fact that undergraduate curricula in the United States are limited in scale. In particular, undergraduate curricula must fit into a four-year time-frame, and are generally limited to 120 credit hours. Thus, the complexity values in the curricular data set are naturally truncated at values above roughly 700. Indeed, the shape of Figure 3 (a) is similar to those of truncated power law distributions.[6,7]

Another possible distribution that may explain this data set is the *exponential distribution*, given by

$$p(x) = \lambda e^{-\lambda x},$$

where $\lambda > 0$ is referred to as the rate parameter. Because

$$\log(p(x)) = -\lambda x + C,$$

data drawn from an exponential distribution should appear as a straight line on a semi-logarithmic plot. In Figure 3 (b) we show the entire curricular data set plotted on a semi-logarithmic scale, and we observe that the data points fall on a nearly straight line.

Thus, there is evidence the curricular data is either distributed according to a truncated power law distribution or an exponential distribution. For the purposes of this research, either distribution leads to the same conclusion. Specifically, both the mean of the exponential distribution, and the mean of the power law distribution (if it exists), are highly sensitive to outliers or extreme values, meaning even a single very large value can significantly shift the calculated mean, making it more susceptible to changes in the data compared to other measures like the median. That is, due to the skewed nature of these distributions, the mean is heavily influenced by extreme values, leading to larger a difference between the mean and median. For this reason, the median is often considered a more robust measure of central tendency in these cases, as it is less affected by outliers.

Why does this matter? Given the longs tails associated with the distributions described above, the average curricular complexity value is easily skewed by a small number of highly complex curricula. Thus, any statements made regarding the average curricular complexity value at a particular institution, or across any collection of institutions, when all disciplines are considered, should be considered highly unreliable. Thus, such comparisons are ill-advised. However, we will show in the next section that comparisons between programs within a given discipline do make sense, and therefore they can be useful in guiding curricular design and reform efforts.

Complexity Distribution – By Field

There are 36 unique two-digit CIP codes associated with the academic programs in this study, each corresponding to a different academic field, as discussed earlier. Let us treat the curricular complexities of all programs sharing a given two-digit CIP code as a random variable. This will allow us to estimate the probability density functions associated with the curricular complexities in each field. These estimates provide a principled means for comparing the complexities of academic programs, across different fields, as well as among

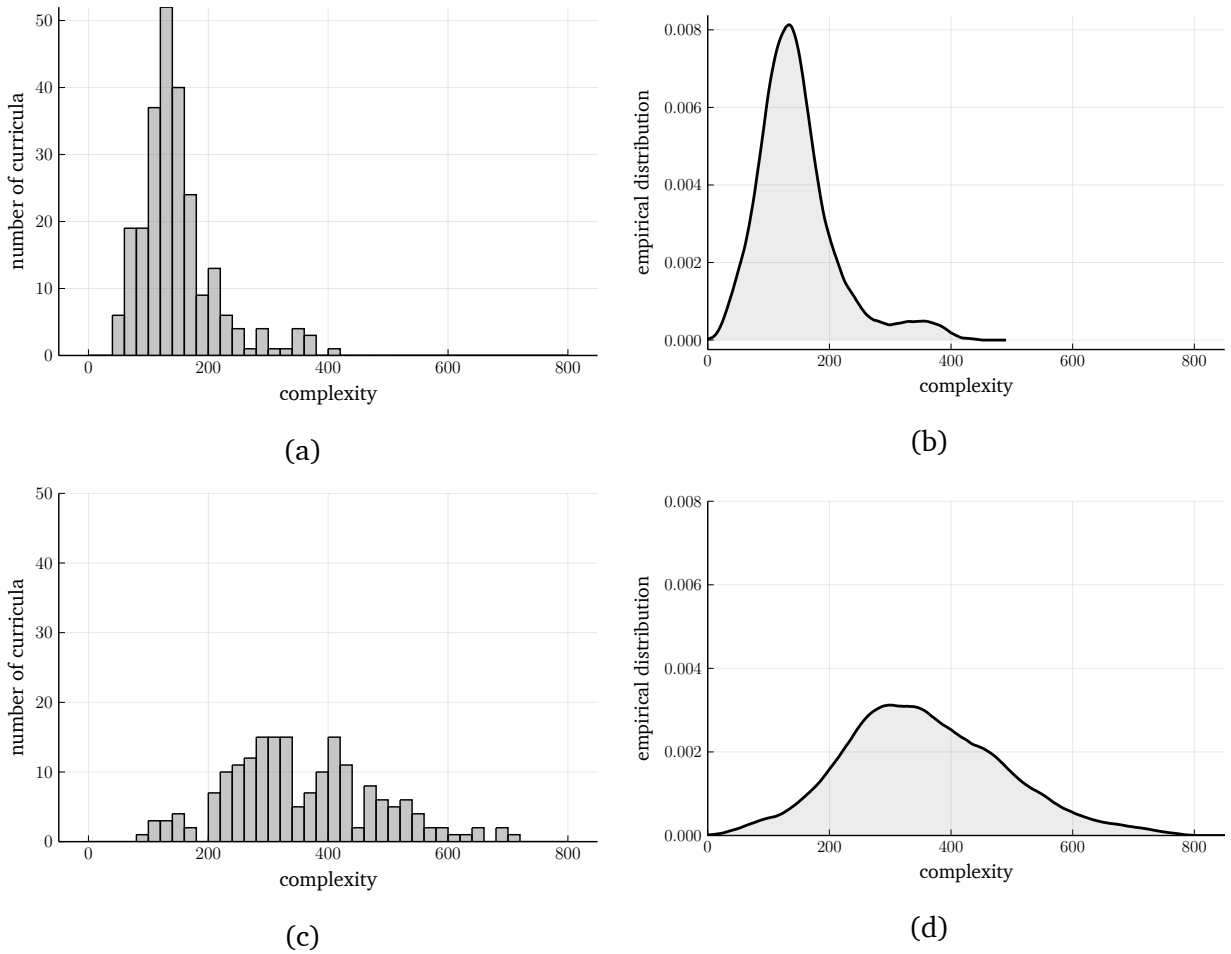


Figure 4: For the field of business, i.e., two-digit CIP 52., (a) the histogram of business program complexities across all institutions, and (b) the KDE computed using this histogram. For the field of engineering, i.e., two-digit CIP 14., (c) the histogram of engineering program complexities across all institutions, and (d) the KDE computed using this histogram.

the various disciplines in a given field. The ability to quantify these distributions will allow us to answer questions such as,"Is my engineering program more complex than the engineering programs at other similar institutions?" and "From a curricular perspective, is engineering a more complex discipline than biology, and if so, by how much?"

In Figures 4 (a) and (c) we show the complexity distributions for all of the business and engineering programs in our study, identified by the two-digit CIP series 52, and 14., respectively. The qualitative complexity differences between these two fields is easy to see. In order to better visualize these differences, we used kernel density estimation (KDE) techniques to estimate the probability density functions associated with these data sets. This methodology uses kernel functions as weights in order to create a smoothed versions of the histograms shown in Figures 4 (a) and (c). Specifically, let $x_1, \ldots, x_n$ denote i.i.d. samples drawn from some unknown probability density f. Then, the *kernel density estimator* for f

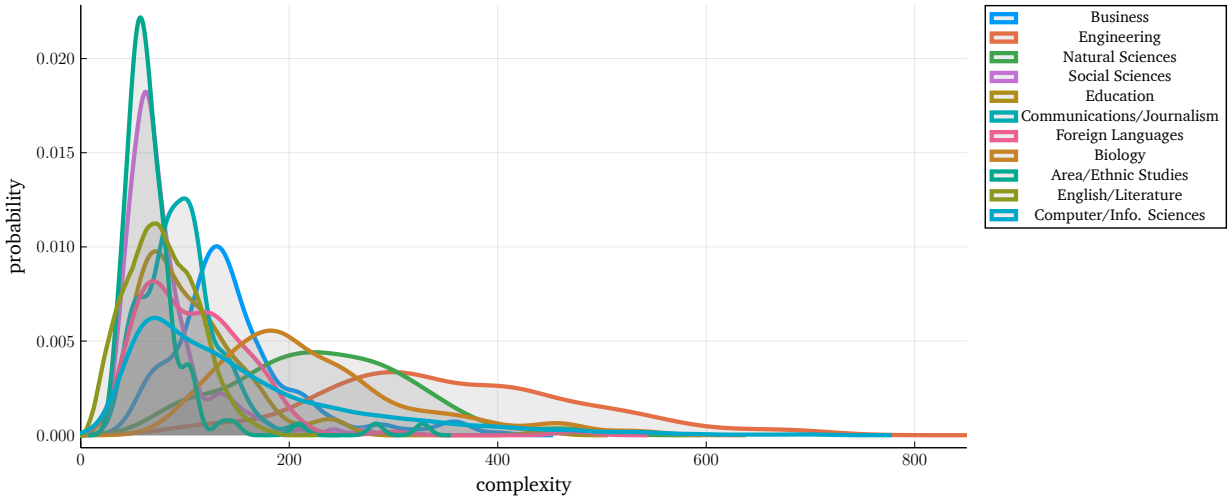


Figure 5: The KDEs for the complexity distributions associated with eleven different fields in the data set.

is given by

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K_n \left(\frac{x - x_i}{h} \right),$$

where K is a non-negative function known as the *kernel*, and h is a smoothing parameter known as the *width*. The KDE for the fields of business and engineering are shown in Figures 4 (b) and (d), allowing us to clearly see the differences between the means and variances in these two empirical distributions.

We see in Figures 4 (b) and (d) that business programs are more tightly clustered around 170, while engineering programs tend to cluster around 325. That is, the engineering field as a whole is inherently more complex than the business field. It also important to note that engineering programs have much larger variability, as compared to the business programs, as well as among all of the other fields identified by their two-digit CIP codes. To visualize this, we selected nine additional fields from the data set, and plotted them all together, along with the business and engineering fields, as shown in Figure 5. This figure clearly shows the stark complexity differences between the curricula in particular disciplines. Notice that many disciplines have prominent peaks in their complexity distributions, but the distribution for engineering is quite flat, demonstrating a large variance across this field. Other fields with large variance include natural sciences, biology, and the health professions. For those fields with large variances, we suspect the variability may be attributable to the particular disciplines that comprise the field. For the engineering discipline, this is investigated in more detail in the following section. Figure 5 also makes clearly evident the underlying structures that leads to the long-tailed distribution of data shown in Figure 2 (b), and also explains a possible generative model for the power law distribution when considering all fields. Specifically, the power law distribution often emerges when mixing data sources having a range of variances.[8]

As mentioned previously, it is not advisable to compare curricular complexities when all dis-

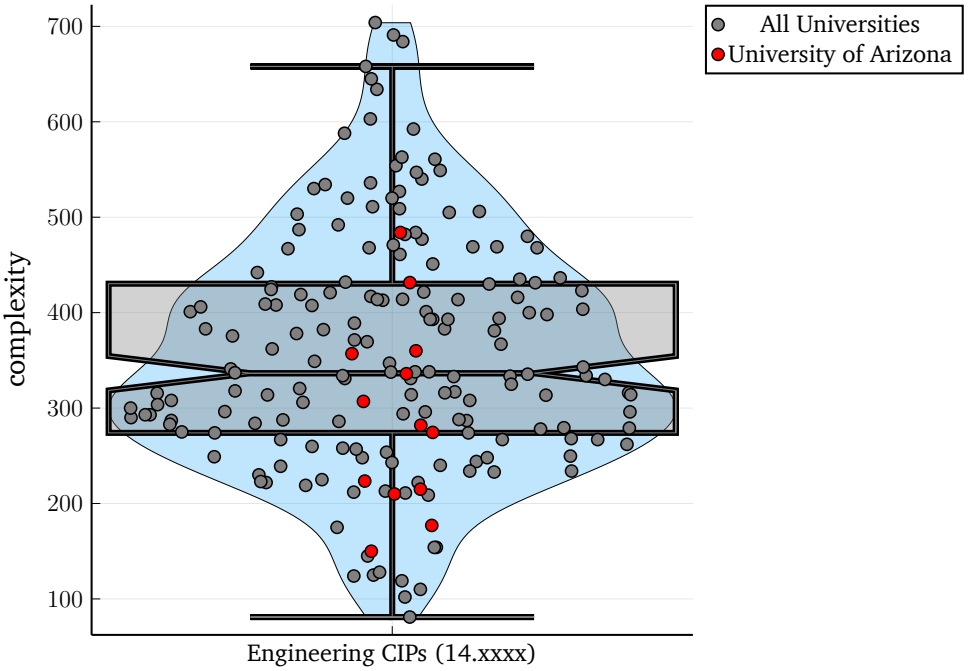


Figure 6: Box scatter plot for the engineering field. The red dots correspond to programs at the University of Arizona, including Aerospace Engineering, complexity = 431; Electrical & Computer Engineering, complexity = 336; Industrial Engineering, complexity = 210; Mechanical Engineering, complexity = 357; Mining Engineering, complexity = 223; Optical Sciences & Engineering, complexity = 274; Systems Engineering, complexity = 177; Architectural Engineering, complexity = 360; Applied Physics, complexity = 215; Biosystems Engineering, complexity = 150; Chemical Engineering, complexity = 484; Civil Engineering, complexity = 307; Environmental Engineering, complexity = 282.

ciplines are considered together. However, because long-tails are not nearly as prevalent in the field-specific empirical distributions shown here, we believe highly useful comparisons can be made at the discipline level. An example of such a comparison is provided in Figure 6.

The plot in this figure should be interpreted as follows. First the box-and-whisker diagram (also know as a boxplot) summarizes the data within a given field of study (two-digit CIP category), in this case engineering. The box itself spans the lower (Q1) and upper (Q3) quartiles of the data set, and the distance between these two is known as the *interquartile range* (IQR). The line in the middle of the box denotes the median value of the data set, which is roughly 340. Thus, 50% of the data lies below this line, with complexity scores less than 340, and the other 50% lies above the line, with complexity scores greater than 340. Furthermore, the data points below the box constitute 25% of the data set corresponding to the lowest complexity scores, and the data points above the box constitute the 25% of the data corresponding to the highest complexity scores. Thus, the box itself contains the middle 50% of the complexity scores in the data set. The whiskers extending from the box show range of the data, from minimum to maximum complexity scores, where

the data points above the maximum value are considered outliers. More specifically, the whiskers extend to the farthest data point that is within 1.5 times the IQR, with data points outside this range considered outliers. Finally the notch in the box indicates the most likely values of the median value. The size of the notch is directly proportional to the IQR, and inversely proportional to the square root of the number of samples in the data set. The notch itself provides an approximate 95% confidence interval for the median of the entire population of programs with this two-digit CIP. That is, it provides a rough estimate of the confidence we should have when using the sample median as an estimate of the population median. Thus, the notch in this plot is useful for comparing the samples drawn from different fields of study. If the notches from two different fields of study do *not* overlap on the complexity axis, it is an indication the median complexities values for these fields of study are different.

Next, the blue outlined shape on this plot is an empirical probability distribution function obtained by applying kernel density estimation techniques to the sample data. If this shape is cut in half along the central axis of the boxplot, and then the left half is rotated (clockwise) by 90 degrees, you will obtain the shape of the empirical probability distribution function for the data set shown in Figure 4 (d). This portion of the diagram is useful for determining the number of modes that might exist the distribution of the actual population. For instance, in the plot shown in Figure 6 appears to have two modes, one with a peak at approximately 300 complexity points, and another at approximately 400 complexity points.

Finally, the data points themselves are plotted, using a function that randomly scatters them about the central axis of the box plot according the empirical probability distribution function. If these points were not scattered in this fashion, they would plot one on top of the other along the central axis. In other words the placement of these data points along the horizontal dimension has no meaning, other than to make them visible. Specifically, by scattering them, it is easy to compare the red data points, corresponding to the programs at a given institution, to those belonging to the other institutions in this study. We contend this plot will provide useful information curriculum designers and curriculum committees, as they work to modify and improve their curricula.

Complexity Distribution – Engineering Disciplines

In order to better understand the large complexity variance in our data set for the engineering field, we next consider the complexities of disciplines within the engineering field, according to their four-digit CIP codes. Figure 7 shows the KDEs associated with the six different engineering disciplines, constructed in the same manner as described above. From this figure, it is easy to see how the combination of engineering discipline complexities produces the widely dispersed complexity distribution for the entire engineering field shown in Figure 4 (d). Furthermore, in this data set the specific engineering disciplines are much more defined. For instance, the systems/industrial engineering programs tend to cluster around a particular complexity value that is quite different from that of the chemical engineering programs. Similarly, the civil engineering programs appear to be less complex than the mechanical engineering programs. Each of these engineering disciplines had 15–25

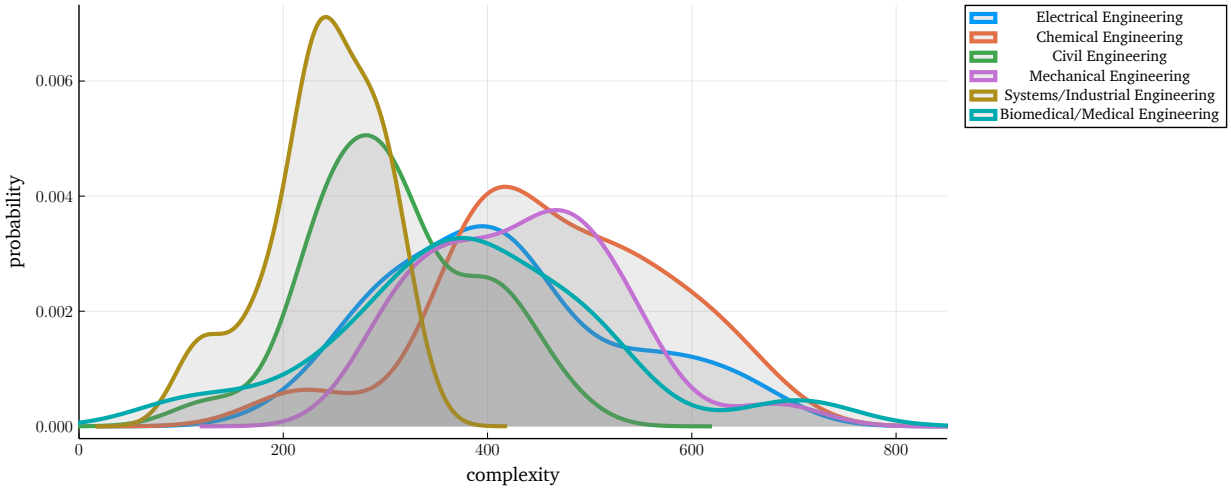


Figure 7: The KDEs for the complexity distributions associated with six different engineering disciplines in the data set.

programs in our data set. Thus, at the moment we are reluctant to say more about these particular distributions of engineering disciplines. Our intention is to collect additional data for the purpose of performing more detailed discipline-specific analyses. Finally, it is worth noting that the discipline specific distributions in engineering have shapes that more closely resemble a Gaussian distribution. Thus, additional data collection may allow us to better quantify these distributions, allowing us to answer question about the means and variances of the complexities of civil engineering programs, chemical engineering programs, etc.

Discussion

In this paper we provided a methodology for quantifying the complexity of academic curricula on field-specific and discipline-specific bases. We also demonstrated the care that should be taken when comparing program complexities. Specifically, comparisons between the curricular complexities of academic programs make little sense unless academic fields are taken into account. This was demonstrated by showing the collection of curricular complexities across all programs in this study resemble a truncated power law or an exponential distribution (or perhaps some combination of the two); that is, a distribution with statistics that are highly sensitive to outliers.

We also showed that reliable within-field comparisons across institutions are possible, and we provided a box scatter plot demonstrating how the complexities of the engineering programs at one institution compare to those at all of the other institutions in the study.

We showed that according to a reliable curricular complexity metric, engineering programs tend to be among the most complex academic programs on a campus, and the complexity variability for the engineering field is larger than that of any other academic field. We showed the large complexity variability in the engineering field is likely attributable to the

complexities of the disciplines within the engineering field. Furthermore, the distributions of the engineering disciplines tend to have shapes with a Gaussian appearance. If the parameters of these discipline-level distributions can be determined through additional data collection, it would lead to very powerful capabilities with respect to the study of academic programs. For instance, it would allow for meaningful comparisons between the complexity statistics of the same discipline at different institutions.

It is interesting to note the similarities that exist between the programs from the same disciplines at the different institutions in our study. One explanation for why academic fields and disciplines across universities tend to have more in common, with regards to their curricular structure, than do the collection of curricula within a single institution is provided by Lombardi.[9] Specifically, he notes that the faculty at American universities organize themselves according to *guilds* defined by their academic specialties. "Moreover, within each university, each faculty guild serves as the local branch of a national guild of the same specialty. For example, all professors in a university history department belong to the same national guild, even though the local university employs them. The national guild establishes the intellectual standards for their work; the local university deals with their employment and work assignments.[9]"

References

[1] Gregory L. Heileman, Chaouki T. Abdallah, Ahmad Slim, and Michael Hickman. Curricular analytics: A framework for quantifying the impact of curricular reforms and pedagogical innovations. *www.arXiv.org*, arXiv:1811.09676 [cs.CY], 2018. `arxiv.org/abs/1811.09676`.

[2] The Classification of Instructional Programs. National Center for Education Statistics, `https://nces.ed.gov/ipeds/cipcode`, 2020.

[3] Aaron Clauset, Cosma Rohilla Shaliz, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703, 2009.

[4] Yogesh Virkar and Aaron Clauset. Power-law distributions in binned empirical data. *The Annals of Applied Statistics*, 8(1):89–1193, 2014.

[5] Anna D. Broido and Aaron Clauset. Scale-free networks are rare. *Nature Communications*, 10(1017):270–277, 2019. `https://doi.org/10.1038/s41467-019008746-5`.

[6] Inmaculada B. Aban, Mark M. Meerschaert, and Anna K. Panorska. Parameter estimation for the truncated pareto distribution. *Journal of the American Statistical Association*, 101(473):270–277, 2006.

[7] Stephen M. Burroughs and Sarah F. Tebbens. Upper-truncated power laws in natural systems. *Pure and Applied Geophysics*, 158:741–757, 2001.

[8] M. Patriarca, E. Heinsalu, L. Marzola, A. Chakraborti, and K. Kaski. Power laws in statistical mixtures. In *Proceedings of the European Conference on Complex Systems*, pages 271–282. Springer Proceedings in Complexity, 2016.

[9] John V. Lombardi. *How Universities Work*. Johns Hopkins University Press, Baltimore, MD, 2013.