Extending curricular analytics to analyze undergraduate physics programs

John Hansen, Amanda Nemeth, and John Stewart*

Department of Physics and Astronomy, West Virginia University, Morgantown, West Virginia 26506, USA

(Received 24 February 2024; accepted 11 October 2024; published 8 November 2024)

Curricular analytics (CA) is a quantitative method that analyzes the sequence of courses (curriculum) that students in an undergraduate academic program must complete to fulfill the requirements of the program. The main hypothesis of CA is that the less complex a curriculum is, the more likely it is that students complete the program. This study compares the curricular complexity of undergraduate physics programs at 60 institutions in the United States. The institutions were divided into three tiers based on national rankings of the physics graduate program, and the means of each tier were compared. No significant difference between the means of each tier was found, indicating that there is not a relationship between program curricular complexity and program ranking. Further analysis focused on the physics, chemistry, and mathematics courses, defined as the core courses of the curriculum. Significant differences in the number of required core courses and the complexity per core course were measured between the tiers; both were measured as large effects. Programs with the highest rankings required fewer core courses while having a higher complexity per core course. These institutions have more strict prerequisite requirements than lower ranking programs. This study also showed complexity was quantitatively related to curricular flexibility operationalized as the number of available eight-semester degree plans. The number of available degree plans exponentially decreased with increasing core complexity per course. Modifications to a curriculum at one institution were analyzed; a similar relationship between the number of available degree plans and increasing complexity per core course was found.

DOI: 10.1103/PhysRevPhysEducRes.20.020143

I. INTRODUCTION

Physics education research (PER) has, since its inception, investigated critical issues of importance to the physics community. This research has explored the science of learning and teaching physics [1] as well as the role that physics programs and classes have in the promotion of equity, diversity, and inclusion of their students [2,3]. More recently, PER has begun to explore the retention of physics majors to degree completion [4,5]. Retention is often one of the primary goals of physics programs because the number of physics majors can determine the amount of economic support received by a program from its university; in the case of smaller programs, it could potentially determine the survival of the program itself [6]. Curricular analytics (CA) [7] is a quantitative method developed to explore the pathways students traverse as they complete academic programs. Curricular complexity is a measure of how constraining the required courses and their prerequisite structure are to student progression through the program. The central hypothesis of CA is that, as a program's curricular complexity is decreased, the student completion rate of the program will increase. As such, CA is a method that can inform the restructuring of program requirements and curriculum to improve student retention.

Physics curricula, the required courses and prerequisite relations in a physics degree, are superficially independent of issues of diversity and inclusion; however, this study will show that the complexity of the curriculum affects a program's flexibility in allowing students to complete the curriculum in less than 8 semesters or 4 years. For most institutions, the 4-year degree plans of physical science and engineering students assume a student is ready to enroll in Calculus 1 their first semester; these students are considered "math ready." A student's initial mathematics class is generally determined by their standardized test scores (ACT or SAT), often supplemented by a mathematics placement test. A number of recent studies have shown that prior preparation measured by standardized test score or conceptual physics pretest score of introductory physics students differs by demographic group [8–10]. As such, students without access to advanced high school course offerings may require more flexible curricula than students with more enriched high school backgrounds for successful graduation in 4 years. Often students from historically

*Contact author: jcstewart1@mail.wvu.edu

marginalized communities have less access to advanced high school coursework than other students [11]. Additional factors beyond academic preparation such as parental support can also influence success in college physics [12]. Equity is dependent on an institution identifying ways in which it can support timely graduation of science, technology, engineering, and mathematics (STEM) students who have been underserved in high school.

Further, because CA examines the curricular structure, it does not require the input of student or class outcome data to characterize complexity. As such, it may form an important theoretical tool to allow departments to explore hypothetical changes to physics programs to improve retention.

A. Research questions

This study seeks to answer three research questions:

RQ1: How do curricular metrics vary for U.S. physics programs? Is there a correlation between program ranking and curricular metrics?

RQ2: How are curricular metrics related to other measures of curricular flexibility?

RQ3: How much do curricular changes affect curricular flexibility?

In this study, we restrict our analysis to information provided by a program's curricular structure. This ignores a second important constraint on student course taking; when and how often a course is offered. We will explore the effects of course offering patterns in a future work. Further, this work also ignores the instructional structures and choices within the courses which can also influence a student's progression to degree; this is also a topic for future research.

B. Results of prior research

Curricular analytics represents a new research strand within PER studying the structure of physics curricula which aims to understand the features of physics programmatic decisions such as the courses required, the prerequisites of those courses, and how often the courses are offered on the ultimate success of physics students measured by the rate of obtaining physics degrees.

1. Physics retention

Studies of physics major retention have examined the factors influencing the intention to continue in the physics major, factors that predict retention of physics majors, and factors related to the retention of students through a physics course sequence. Many factors influencing retention were not curricular in nature.

Stiles-Clark and MacLeod surveyed physics majors after completion of the introductory physics course sequence about the factors that had been important to their decision to continue as physics majors [13]. Students reported that instructional quality, career opportunities, and an interest in physics were primary motivators for persisting in the physics major.

Stewart *et al.* examined physics major retention to degree at two institutions with undergraduate populations with varying levels of academic preparation [5]. The factors predicting leaving college and leaving the physics major while staying in college differed and were different at the two institutions. For the less selective institution, high school GPA was the most important predictor of physics majors leaving college without a degree, while math readiness was the most important predictor of changing major while staying in college. At the more selective institution, only composite ACT scores were significant in predicting retention. Aiken *et al.* used machine learning techniques to explore the factors which were most predictive of physics majors successfully completing a physics degree [4]. Enrolling in modern physics and an engineering course emerged as the most predictive factors.

Using network analysis, Zwolak *et al.* showed that a student's academic and social integration in his or her introductory physics course was important to that student's persistence to the second course in the introductory sequence [14]. This supported work by Forsman *et al.* who used both academic and social networks and complexity science to explore retention [15].

2. Prior studies using curricular analytics

Heileman *et al.* [7] introduced CA for analyzing the structure of a program's curriculum. Other methods that are similar in scope and design have been used to explore student progression through degree programs and are a growing area of STEM education research [16].

CA has been used to understand the structural effects of successful curricular innovations. Klingbeil and Bourne [17] introduced a curricular modification designed to aid the progression of incoming engineering students through the Calculus 1 and 2 sequence. Many students who enter the university are not ready to enroll in Calculus 1. These students require several semesters to complete additional mathematics classes before they can enroll in their first engineering course. This is a common problem in physics and engineering programs where students must complete the introductory calculus sequence before entering their program-centered classes. While maintaining ABET standards in the engineering program at the university, an introductory Engineering Mathematics (EGR 101) course was introduced. This course focused on hands-on approaches to the most important mathematics methods that are used in engineering courses. Successful completion of EGR 101 allowed students to advance to program-centered engineering courses such as the introductory physics sequence, engineering mechanics and statics, and computer programming sequences before completing the traditional calculus prerequisites for these courses.

This change nearly doubled graduation rates while narrowly improving the average GPA. Students from historically marginalized communities, including women and minorities, experienced the largest increase in graduation rate. This change reduced the effect of the introductory calculus sequence, allowing students to take the introductory calculus sequence at the same time as their program centered courses. Heileman *et al.* [7,18] showed that this change reduced the curricular complexity for students unprepared to take Calculus 1 upon entering the program, supporting their argument that less complex curricula lead to increased graduation rates.

These types of curricular changes and their effects were investigated by Slim *et al.* [19]. In their study, Markov decision processes were used to quantify the relationship between curricular complexity and graduation rate and were then used to model how curricular changes affect graduation rates. Decreasing the complexity of the curriculum increased graduation rates.

To further support the benefit of less complex curricula, a study compared the curricular complexity of electrical engineering programs [20]. Program ranking was taken from the U.S. News rankings of engineering programs, where, for example, a school ranked 5th is considered to have a higher ranking than a school ranked 95th. Programs with higher rankings had less complex curricular structures than schools with lower ranking. This implies that higher-ranking schools had less complex paths to completion of an electrical engineering degree than lower-ranking programs. A similar study compared curricular complexity and ranking within computer science programs finding similar results [21]. The relationship between complexity and ranking in disciplines other than electrical engineering and computer science has yet to be established.

Other studies have applied CA to analyze the curricular complexity of transfer student pathways to degree completion, with the result that transfer student pathways are more complex than standard program pathways [22,23]. Similarly, one study looked at the complexity of the suggested path of study that an institution advises students to take and found that the actual paths that students followed to degree completion were less complex than the suggested path [24]. Other applications of CA include the use of the curricular complexity as a variable in a geometric probabilistic model that was used to predict graduation rates of students in different academic programs [25].

II. METHODS

A. Sample

Curricular complexity was compared across three tiers of physics programs in the United States. Following prior studies in electrical engineering and computer science, these tiers were selected using program rankings from the 2022 U.S. News and World Report College Rankings [26] for graduate physics programs; we hypothesized that graduate rankings would largely mirror undergraduate rankings. Each of the programs in the ranking offers a doctoral degree; there are 188 programs in the ranking. Twenty schools were randomly selected from the first two deciles (programs ranked 1 to 38) to make up the upper tier. The middle tier consisted of 20 schools randomly selected from the fourth and fifth deciles of the rankings and included schools ranked from 75 to 113. The lower tier was made up of 20 schools from the ninth and tenth deciles, which included schools ranked from 150 to 188. In the random sampling within each tier, if an institution was selected that did not have a clear, publicly available, delineation of the requirements to complete their undergraduate physics program, a different institution was randomly selected. Institutions that operate on a quarter system were also excluded from the sampling as it was unclear how to modify the complexity of a program in a quarter system to be comparable to a program in a semester system.

Most programs included in the analysis have several different degree tracks available to physics majors. In each case, the degree track that was suggested for students planning to continue their physics education in graduate school was selected. The program requirements consist of a set of required classes that all students must take and then a number of physics or mathematics electives with a list of course offerings that fulfill the elective. To maintain consistency, similar courses were selected for each program's elective requirements when possible.

To explore the effects of modifying curricula in a controlled manner, we focused on the physics curriculum of a single university from the second tier. A recent study explored the physics retention patterns of this program [5]; the university is referenced as institution 1 in this prior study. The institution is a large public land-grant with an overall undergraduate population of 20 500. The general undergraduate demographic composition in Fall 2019 was 82% White, 4% Black or African American, 4% Hispanic/Latino, 4% nonresident alien, 4% two or more races, with other groups 2% or less. The 25th to 75th percentile range of ACT composite scores was 21 (59%) to 27 (85%) for the 25th percentile to the 75th percentile of students scores [27]. About 31% of undergraduate students met the eligibility requirements for Pell grants. This institution will be referenced as Middle Tier Public University (MTPU) in this study.

B. Curricular analytics

The primary metric of CA is the overall curricular complexity. This is composed of two components: a structural component and an instructional component.

The instructional properties consist of the instructor quality, course support services such as tutoring and office hours, and any other property of the instruction. The instructional complexity component of a curriculum's complexity is more difficult to quantify than the structural

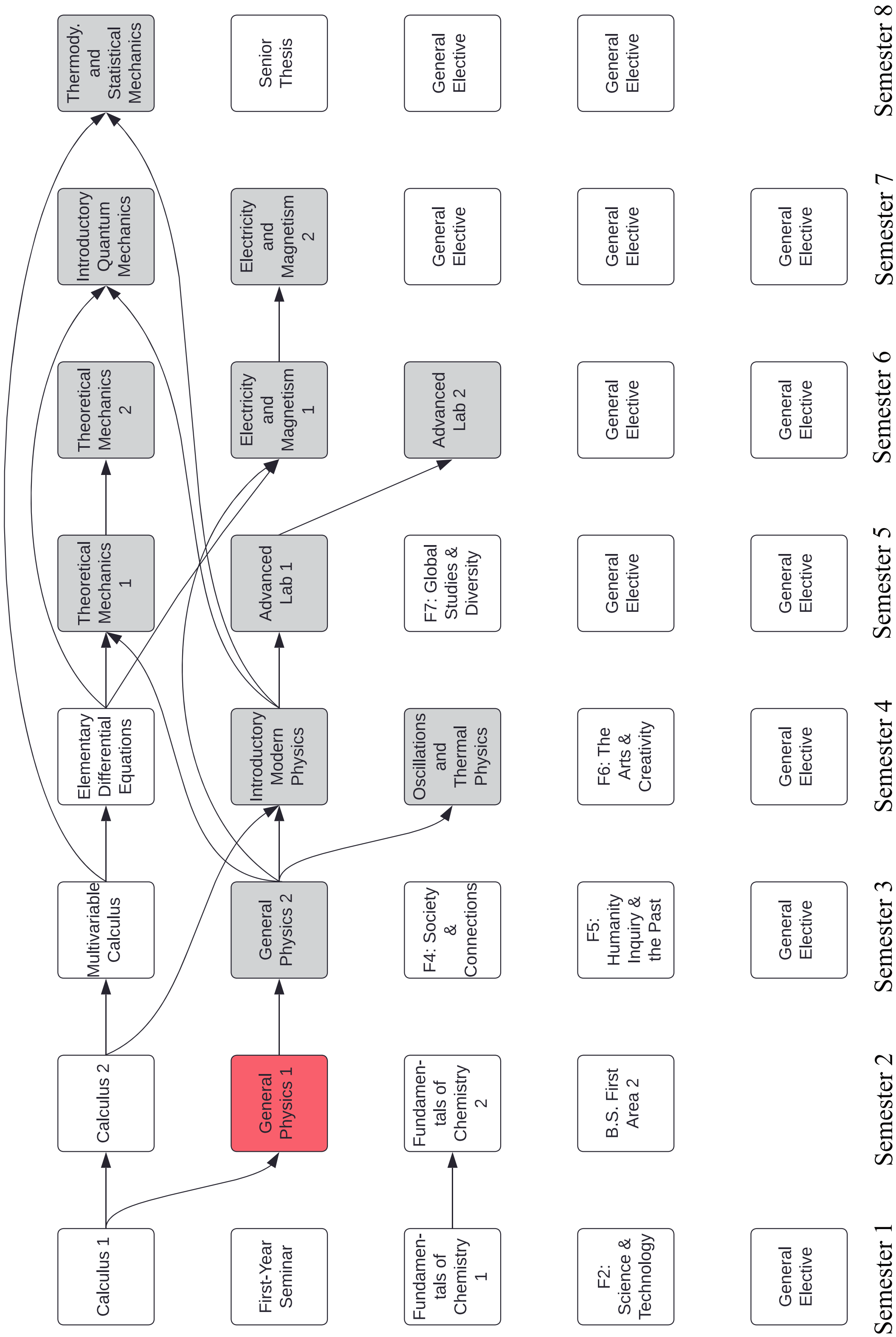

FIG. 1. Example degree plan for a mid-tier institution. The corresponding curriculum graph is shown in Fig. 1 in the Supplemental Material [28]. General Physics 1 is shaded red and the courses it blocks are shaded gray. The number of gray courses is the blocking factor of General Physics 1. The delay factor of General Physics 1 can be found by counting the number of courses in its longest path, in this case 6. The university divides general education requirements into seven categories labeled F1 to F7 in the figure.

complexity. Instructional characteristics are qualitative in nature; it is challenging to consistently quantify their effects on student outcomes. Heileman *et al.* suggest the use of course grades or pass and fail rates as an approximation to the instructional complexity [7]; however, this is far from a complete measure of instructional complexity.

The structural complexity of a curriculum is quantified by examining the prerequisite structures of the curriculum. This prerequisite structure is visualized by using a directed acyclic graph, where individual courses are nodes and the edges connecting nodes are prerequisite or corequisite requirements. This is called a curriculum graph. An example of a curriculum graph is shown in Fig. 1 in Supplemental Material [28]. This curriculum graph, which was generated at the CA website [29], captures the structural relations but is not intuitively presented. For example, the farthest right column has all general electives. This is because the curriculum graph does not capture information about which courses should be taken each semester. Figure 1 shows a degree plan, a plan showing in which of the eight semesters each course in the curriculum should be taken consistent with the constraints represented in the curriculum graph. There are generally many degree plans consistent with each curriculum graph. Each degree plan would contain the same nodes and edges as the curriculum graph.

Heileman *et al.* defined five characteristics of a program's structure: the delay factor, the degrees of freedom, the blocking factor, the reachability factor, and the centrality factor. In the present work, only the delay and blocking factors, which are required to calculate the structural complexity, and the centrality factor are discussed. For a full treatment of each factor, refer to Heileman *et al.* [7].

Required courses in a curriculum are generally part of a required course sequence, where each course in the sequence must be completed before advancing to the next course in the sequence. Some courses may be part of several sequences. The delay factor, d_n, of a course n is defined as the number of courses (or nodes on the curriculum graph) that are included in the longest sequence that contains course n. For example, in Fig. 1, the delay factor of General Physics 1 would be 6 resulting from the path traversing nodes, Calculus 1, General Physics 1, General Physics 2, Introductory Modern Physics, Advanced Lab 1, and Advanced Lab 2. Often the longest sequence includes courses that act as gateway courses; courses that are a prerequisite course to many other required courses. This longest course sequence is called the "longest path" in CA.

The blocking factor, b_n, of course n is the number of courses or nodes for which n is a prerequisite or equivalently the total number of courses that follow after n in all the course sequences that include n. For example, the blocking factor of General Physics 1 in Fig. 1 is 11; the classes blocked by General Physics 1 are shaded in gray in the figure. In graph theory, the blocking factor is the number of descendants of the class in the curriculum graph.

The individual course complexity v_n is the sum of the course delay and blocking factor, $v_n = d_n + b_n$. The structural complexity of a program, α_c, is the sum of the course complexity of each course in the curriculum [Eq. (1)].

$$\alpha_c = \sum_n v_n = \sum_n (d_n + b_n). \tag{1}$$

The course centrality factor identifies courses that have several important prerequisites that are also prerequisite for many required courses. The course centrality factor attempts to measure how critical the progression through a course is for the completion of a curriculum; for a course n, it is calculated by summing the length of all continuous complete directed paths p that contain the course n. A path is complete if it cannot be extended by traversing an additional edge. For a path to be included in the summation, course n must be an interior node to p. The course with the highest centrality factor is the course with the highest average path length passing though it. Although the course centrality factor plays no direct role in calculating the structural complexity, it gives information as to what courses are especially crucial to successful program outcomes and as such is of interest to student retention research studies.

In this study, we consider only the structural properties of physics curricula and reserve the instructional component for future studies. For brevity, the structural complexity will be called the complexity in what follows.

In this work, most (structural) complexities were calculated using the CA website [29]. Additional metrics were calculated using software developed for this project, discussed in the following section.

III. RESULTS

A. Curricular analytics across multiple institutions

Figure 2 shows boxplots of several curricular metrics divided by program tier. The (structural) complexity, C, is shown in Fig. 2(a). The shaded boxes of each tier represent the 25% to 75% range of that tier, also known as the interquartile range (IQR). The dark horizontal line is the median of the tier. The full range of the data points, excluding outliers, in each tier is contained between the tips of the two vertical lines, called whiskers. These whiskers contain the lower 25% and the upper 25% of scores in the data. Outliers are identified by dots shown outside the whiskers. Outliers are determined by adding $1.5 \times \text{IQR}$ to the 75% score or subtracting it from the 25% score [30]. There was little difference between the mean complexity of the tiers, with only a 12 complexity point difference between the lower and middle tiers, 7 points between the lower and upper tiers, and 19 points between the middle and lower tiers. The means of the complexity of each tier were compared using analysis of variance (ANOVA),

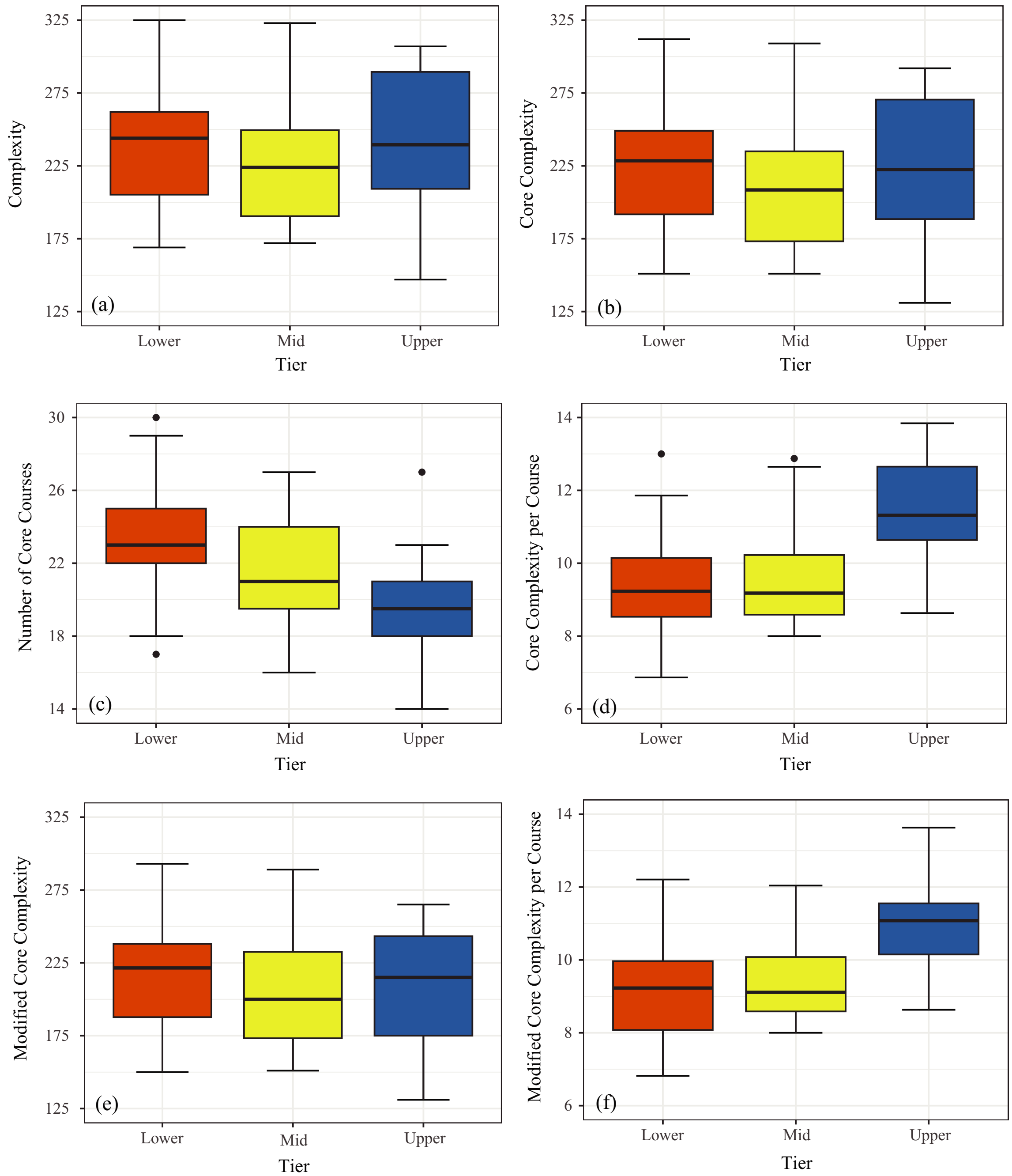

FIG. 2. Distribution of curricular metrics for physics programs with different rankings. (a) shows curricular complexity by tier, (b) core complexity by tier, (c) the number of core courses by tier, (d) core complexity per course by tier, (e) modified core complexity by tier, and (f) modified core complexity per course by tier.

showing the differences in means were not significant at the $p = 0.05$ level ($p = 0.355$). An additional statistical description of the ANOVA analysis is presented in Table I. For each metric presented, ANOVA was performed separating the 60 institutions studied into three tiers. As such, the present study does not replicate the results of Heileman *et al.* [7] for electrical engineering and computer science [21].

Complexity has obvious flaws when one wishes a more nuanced comparison of curricula. Complexity increases through two very different mechanisms. First, a program that requires more physics, mathematics, and chemistry classes will have a higher complexity than one requiring fewer such classes. Second, programs with more restrictive prerequisite structures will generally have a higher complexity than programs with less restrictive structures. All

TABLE I. ANOVA results for the full and core curricula. For each metric, ANOVA was performed separating the 60 institutions studied into three tiers. The table presents the F statistic, its p value, and eta-squared η^2 as a measure of effect size. The effect size criteria for η^2 are that 0.01 is a small effect, 0.06 a medium effect, and 0.14 a large effect [31]. Metrics indicated by an * indicated a violation of normality. In these instances, a Kruskal-Wallis test was conducted to verify the significance of the reported results.

Metric	$M \pm SD$ (lower)	$M \pm SD$ (mid)	$M \pm SD$ (upper)	F	p	η^2
		Full curricula				
Complexity (C)	235 ± 40.7	223 ± 38.5	242 ± 48.8	1.06	0.355	0.036
Complexity per course	6.22 ± 1.10	5.90 ± 1.02	6.53 ± 1.09	1.74	0.185	0.058
Modified complexity per course	6.07 ± 1.01	5.80 ± 1.02	6.20 ± 1.09	0.99	0.380	0.033
Modified complexity	230 ± 37.3	219 ± 35.5	230 ± 38.1	0.54	0.586	0.019
Number of courses*	37.9 ± 1.04	37.9 ± 2.30	37.0 ± 2.58	1.24	0.297	0.042
		Core curricula				
Core complexity (C^C)	221 ± 42.4	207 ± 40.4	225 ± 49.4	0.96	0.391	0.032
Number of core courses (N^C)	23.4 ± 3.08	21.6 ± 3.28	19.7 ± 2.96	7.26	0.002	0.203
Core complexity per course* ($\overline{C^C}$)	9.46 ± 1.52	9.61 ± 1.35	11.4 ± 1.56	10.67	<0.001	0.272
Modified core complexity (C^{CM})	215 ± 38.8	203 ± 37.5	212 ± 38.7	0.57	0.570	0.020
Modified core complexity per course ($\overline{C^{CM}}$)	9.23 ± 1.45	9.43 ± 1.12	10.8 ± 1.29	8.90	<0.001	0.238
Maximum path length*	6.80 ± 0.95	6.55 ± 0.83	7.65 ± 1.14	6.93	0.002	0.196
Modified maximum path length*	6.50 ± 0.69	6.40 ± 0.60	7.00 ± 0.73	4.57	0.015	0.138

programs investigated require a total of approximately 120 credit hours or approximately 40 classes. Adding a required class to the physics degree (usually a physics or mathematics class) removes a general education elective which has been modeled as having no prerequisite and adds a class that generally has a chain of mathematics and physics prerequisites, increasing complexity. To focus on the mathematics and science part of the curriculum, Fig. 2(b) shows the complexity of only the mathematics, physics, and chemistry classes; we call this the core complexity C^C. This removes the effect of general education electives which are generally of lesser importance for the difficulty of progression of physics students through the degree.

The core complexity shows the same general pattern as the complexity with little difference between the tiers. The core complexity can be calculated at the CA website by removing the general elective courses from the curriculum. The core complexity can be separated into the two effects: the number of core courses, N^C, and the average core complexity per course, $\overline{C^C}$, where $\overline{C^C} = C^C/N^C$. Figure 2(c) shows the number of core courses and Fig. 2(d) shows core complexity per course separated by program tier. Both figures show differences between the tiers. These differences are statistically significant, see Table I. Both are large effects ($\eta^2 = 0.203$ and $\eta^2 = 0.272$, respectively). The effect size criteria for η^2 are that 0.01 is a small effect, 0.06 a medium effect, and 0.14 a large effect [31].

One of the central decisions physics departments make about their curricula is the relation of Calculus 1 and Physics 1. Two common choices are to make Calculus 1 a prerequisite for Physics 1 or to make Calculus 1 a corequisite for Physics 1. Both choices produce the same curricular complexity because corequisites and prerequisites are treated the same in the longest path calculation. This is clearly unacceptable for evaluating physics programs. To repair this, we suggest the longest path be calculated by counting corequisite links as zero length. We call this definition of the longest path, the "modified longest path." Because students can either take the courses as prerequisite or corequisite, the length of the actual path taken will be between the modified longest path and the original longest path. Figure 2(e) presents the modified core complexity, C^{CM}, using the modified longest path definition, while Fig. 2(f) presents the core complexity per course with the modified definition, $\overline{C^{CM}}$. The modified core complexity per course was also significantly higher for upper tier programs.

Table I shows the result of applying an ANOVA analysis to the curricular metrics discussed above. The top panel investigates the full curriculum including general education courses; the bottom panel investigates the core curriculum. Using the full curricula, none of the quantities are statistically significantly different, while using the core curricula, the number of core courses, the core complexity per course, and the modified core complexity per course all are significantly different as suggested in Fig. 2. The significant differences are all large effects. This suggests the inclusion of general education electives in the calculation of the complexity serves to obscure some important differences between curricula. Both the number of core courses and core complexity per course had significant differences between the lower and upper tiers ($p = 0.02$ and $p < 0.001$, respectively) while the core complexity per course was significantly different between the mid and upper tiers ($p < 0.001$). All are large effects. Some metrics presented in Table I violated the assumption normality required by ANOVA; these have been marked with an asterisk (*). In these cases, a Kruskal-Wallis test was also conducted to verify significance yielding no differences in significance results from those of ANOVA. See Supplemental Material for more information on this statistical test [28].

Our initial finding that complexity does not differ by program tier resulted from two effects that obscured differences in programs: (i) the effect of general educational electives tending to mute differences between curricula and (ii) the effect of the number of core courses and the core complexity per course cancelling.

The statistical results in Table I strongly indicate that including general education classes in the analysis of curricular complexity prevents the identification of important differences between programs. This may at first seem unintuitive. The following presents a simple model that shows one way general education courses could mute differences in complexity. Each general elective has a complexity of one in our model; as such the core complexity is the complexity minus the number of general education courses. Because all programs have approximately 40 total courses, the inclusion of general education courses in the complexity tends to partially obscure differences in the science and mathematics components of the program. For example, if two curricula with the same number of total courses (N^T) contain $N_1^C < N_2^C$ core classes resulting in core complexity $C_1^C < C_2^C$ (in general, complexity grows with the number of courses), then the overall complexity would be $C_1 = C_1^C + (N^T - N_1^C)$ and $C_2 = C_2^C + (N^T - N_2^C)$. A larger number $(N^T - N_1^C)$ gets added to the smaller complexity reducing the difference between the two complexities, C_1 and C_2.

With the modified definition of the longest path, a program with a modified longest path length of greater than eight cannot be completed in eight semesters. Figure 3 shows the distribution of modified longest paths across the institutions. One program in the upper tier had one modified longest path of length 9. This program was included in the eight longest path bar in Fig. 3. A program with a modified longest path of eight has at least eight classes whose failure guarantees that the degree cannot be completed in eight semesters. The four programs with modified longest paths of eight had 12 to 14 courses on eight-semester longest paths giving students many courses whose failure guaranteed not finishing the degree in 4 years. Many programs in the lower and middle tiers did not have a modified longest path longer than six classes providing students more flexibility if a class failed. The distribution of modified longest paths in the upper tier was dramatically different than the lower and middle tier with many more institutions with modified longest paths of length 7 or greater. This is consistent with the higher modified complexity per course of the upper tier institutions indicating more restrictive curricula.

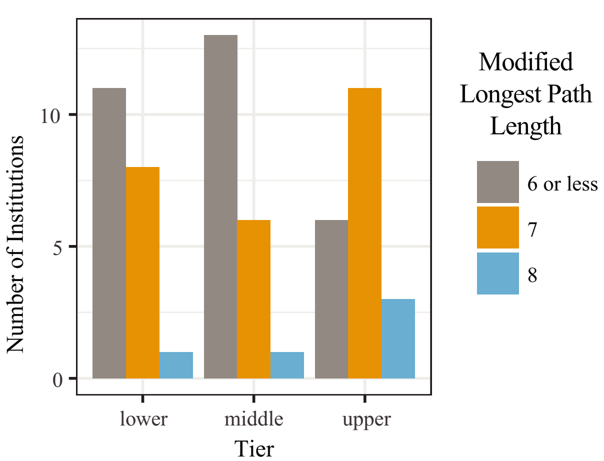

FIG. 3. The number of programs with modified levels of longest path length by tier.

Beyond differences between the tiers, Fig. 2 also shows the broad range of program metrics within tiers and across tiers with a range of core courses from 14 to 30, a range of complexity from 150 to 325, and core complexity per course from 7 to 14. This indicates there are substantially different curricular models implemented across physics programs in the United States. Some programs are substantially more course intensive or have substantially more restrictive prerequisite structures than other programs at the same level.

The central course is the most important course to progress to and to successfully complete for matriculation through the curriculum. For 38 institutions, the central course was Calculus 2; General University Physics 2 was the central course for 17 institutions; and General University Physics 3 was the central course for 5 institutions. General University Physics 2 was the introductory, calculus-based, electricity and magnetism course. General University Physics 3 had a different description at each of the five institutions where it was the central course. At each institution, it had some coverage of wave mechanics and introductory quantum physics; at two of the institutions, it had some coverage of relativity. One of the institutions also included basic thermodynamics as part of its description.

B. Flexibility and delay

Central to CA metrics are two concepts: the flexibility of the program and the degree to which an academic mis-step delays graduation. The introduction provides evidence that complexity is related to measures of student success, but neither the original CA paper [7] nor the supporting studies cited in the Introduction provide direct evidence that the complexity is related to curricular flexibility or potential for delay. Both the delay factor and the blocking factor represent heuristics that intuitively should be related to degree flexibility and risk of delay to graduation.

In this section, we provide direct quantitative evidence that curricular metrics are related to the flexibility of a degree program or the likelihood of a delay of graduation. We further address a primary flaw in CA. While CA asserts that decreasing complexity leads to improved student success, it does not provide guidelines about what magnitude of change in complexity is important. To do this, degree flexibility is operationalized as the number of distinct degree plans, DP, which can be completed in eight or fewer semesters. This is calculated from the curriculum graph with some additional assumptions. A degree plan specifies which classes are taken in which semester. The curriculum graph provides the information needed to compute all available degree plans. Only the core science and mathematics classes were included in the degree plans. We further required that at least one class be taken each semester and that no semester contained more than 15 credit hours. Because the simulation was restricted to mathematics and science classes, semesters containing more the 15 credits were considered to be too difficult for most students. Calculations were performed assuming each course was offered every semester. This is a serious limitation which will be explored in future work. Software was constructed to enumerate all degree plans. The computer run time for each institution was restricted to 1 week on a modern computer; this allowed the calculation of the number of core degree plans of 47 of the 60 institutions. Of the original 20 institutions in each tier, 14 lower tier, 16 middle tier, and 17 upper tier institutions could be calculated. Those which could not be enumerated generally had a variety of required courses in the first semester which led to a combinatorial explosion.

The number of degree plans calculated covered a very broad range of up to 1.3×10^9. To understand the range of degree plans identified, it is helpful to understand the number of degree plans unconstrained by prerequisite requirements. This is size of the space the software searches to identify the number of degree plans fitting the constraints. The total number of degree plans for N^C core courses taken in eight semesters is shown in Eq. (2),

$$\text{Degree Plans} = \sum_{\{sems\}} \frac{N^C!}{n_1!n_2!n_3!n_4!n_5!n_6!n_7!n_8!}, \quad (2)$$

where n_i is the number of courses in semester i and $\{sems\}$ is the set of all sequences $(n_1, n_2, n_3, n_4, n_5, n_6, n_7, n_8)$ such than $\sum_i n_i = N^C$ and $1 \le n_i \le 5$ assuming all classes are 3 credit hour courses. There is not a closed form solution to the sum, but the general size can be estimated by examining one of the terms. The number of core courses ranges from 14 to 30. For 16 core courses, if 2 courses are taken each semester, there are $16!/(2!)^8 = 8 \times 10^{10}$ unconstrained degree plans. For 24 courses, if 3 are taken each semester there are $24!/(3!)^8 = 4 \times 10^{17}$ degree plans. Because these would be only one term in the sum, the actual number of unconstrained degree plans is much larger. The prerequisite requirements dramatically reduce the number of valid degree plans, but the number of combinations remaining can still be quite large. The software created for this project performs a depth first search on the curriculum graph first generating all possible first semester course combinations, then for each of these all valid second semester combinations until a degree plan is constructed. The software is available at github [32]. To our knowledge, this is the first implementation of software to simulate available degree plans. Future versions should substantially optimize the process allowing computation in shorter times.

For students, as they progress through their physics degree, they rapidly reduce the number of degree plans still available as they take classes. For programs with substantially more degree plans, students are likely to have more choices of the classes that can be taken and the order of classes that can be taken. This can allow them the flexibility to recover from failing a class or from not being able to take a class as scheduled.

1. Flexibility

Figure 4(a) shows the number of degree plans plotted against the core complexity on a logarithmic scale. All plots in Fig. 4 include both a linear regression line and a smoothing curve. The fraction of the variance explained by each regression, R^2, and the correlation, r, of the logarithm of the dependent variable with the independent variable are presented in each figure. The smoothing curve was calculated using the default algorithm of "geom_smooth" in "ggplot" in the R software system [33]. For the number of data points in each plot in this work, the locally estimated scatterplot smoothing (loess) algorithm was the default which fits regression lines to a subset of the data (75% in this case). There is little relation between core complexity and the number of degree plans. This results from the competition of an exponential increase in the number of degree plans with the number of core courses and an exponential decrease in the number of degree plans with increased core complexity per course as shown in Figs. 4(b) and 4(c). Figure 4(d) plots the degree plans against the modified core complexity per course; the regression line in this plot is a better fit to the smoothing curve than the regression line in Fig. 4(c). While the core complexity is not related to the number of available degree plans, both the number of core courses and core complexity per course are.

The increasing relation of the number of classes with degree plans is not particularly useful. While it is true that there are more semester combinations in a curriculum containing more courses, the semesters themselves are more difficult likely to mitigate any positive effect. As such, increasing the number of courses to increase flexibility is not a sound curricular design principle.

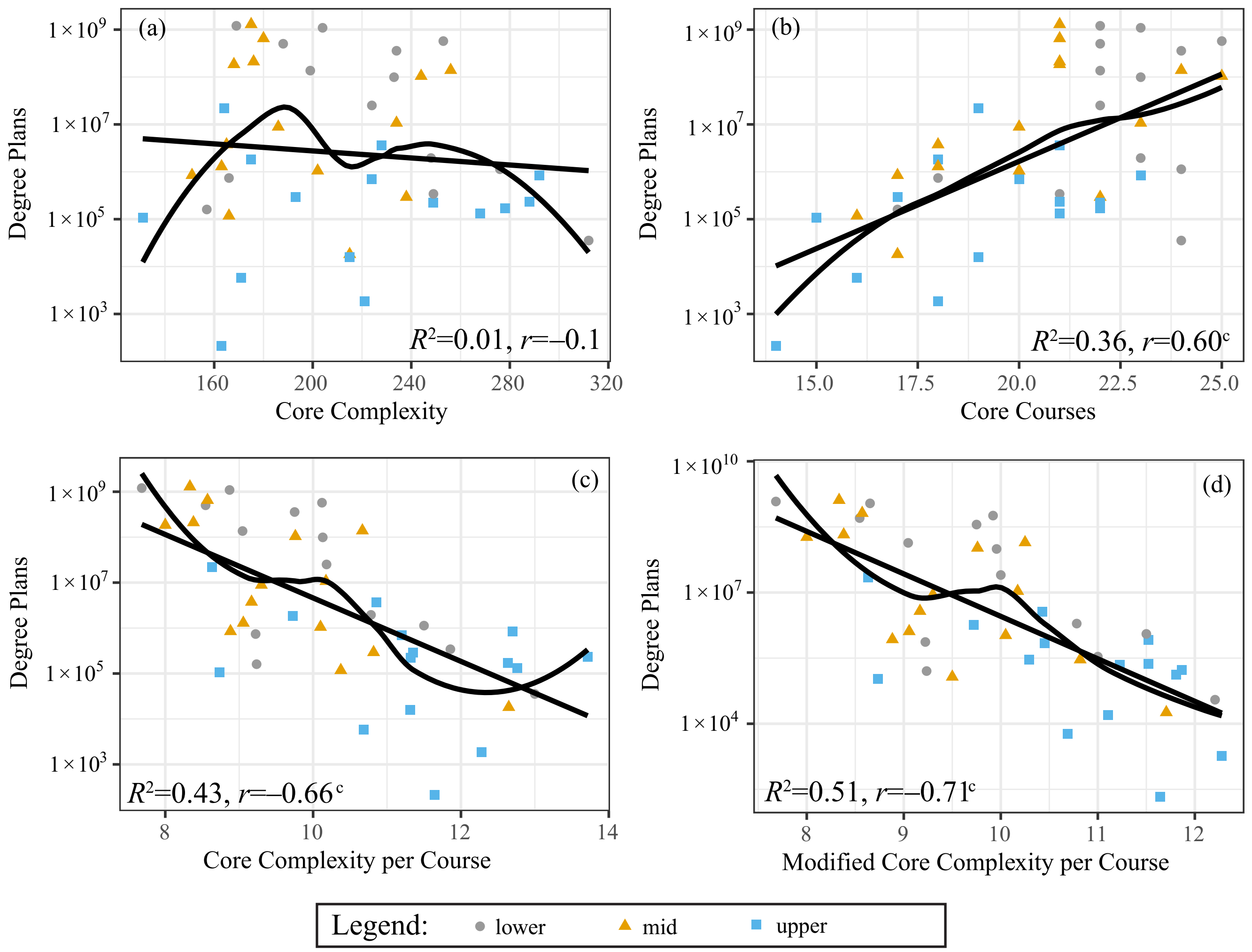

FIG. 4. Degree plans plotted against curricular metrics. All plots are on a logarithmic scale. (a) shows the number of degree plans plotted against the core complexity, (b) the number of degree plans plotted against the number of core courses, (c) the number of degree plans plotted against the core complexity per course, and (d) the number of degree plans plotted against the modified core complexity per course. A linear regression line and a smoothing curve is plotted. The variance explained by each regression line, R^2, is included on each figure. Each figure also includes the correlation coefficient r between the independent and dependent variable as well as its significance level denoted by a superscript: a denotes $p < 0.05$, b denotes $p < 0.01$, and c denotes $p < 0.001$.

Conversely, the exponentially decreasing number of degree plans with increasing curricular complexity per course is something departments have no simple metric to characterize. This relation can be quantitatively explored using linear regression fitting the natural logarithm of the number of degree plans against the modified core complexity per course in Eq. (3),

$$\ln(DP) = \beta_0 + \beta_1 \cdot \overline{C^{CM}}, \tag{3}$$

where β_0 is the intercept and β_1 is the slope. The regression model was a statistically significant ($p < 0.001$) improvement over the null model containing only the intercept with $\beta_0 = 37.2$ and $\beta_1 = -2.24$. Exponentiating Eq. (3) yields Eq. (4).

$$DP = e^{\beta_0} \cdot e^{\beta_1 \cdot \overline{C^{CM}}}. \tag{4}$$

For each one unit increase in $\overline{C^{CM}}$, the number of available degree plans is multiplied by $e^{-2.24} = 0.11$; the number of available degree plans is reduced by 89%.

While the results which follow provide evidence that the modified core complexity per course is a better metric for characterizing curricula, this measure is currently not supported at the CA website. As such, it would be challenging for physics programs to calculate. For greater usability, results using core complexity without the modified definition of the longest path are also presented. The core complexity per course, $\overline{C^C}$, was more weakly related to DP with $\beta_1 = -1.61$ and $e^{-1.61} = 0.20$; a one unit increase in core complexity per course decreased the available degree plans by 80%.

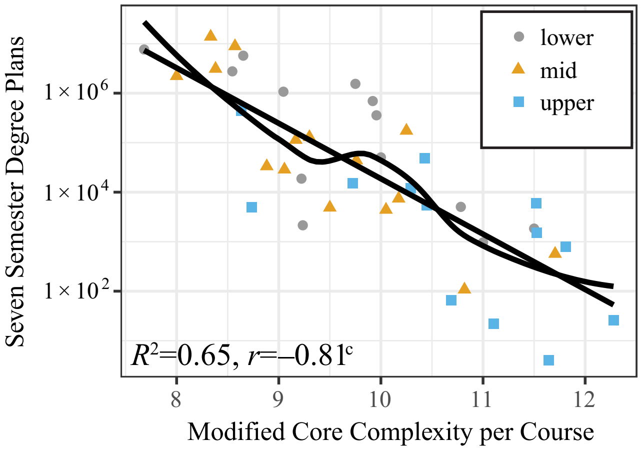

FIG. 5. Seven-semester degree plans plotted against the modified curricular complexity per course on a logarithmic scale.

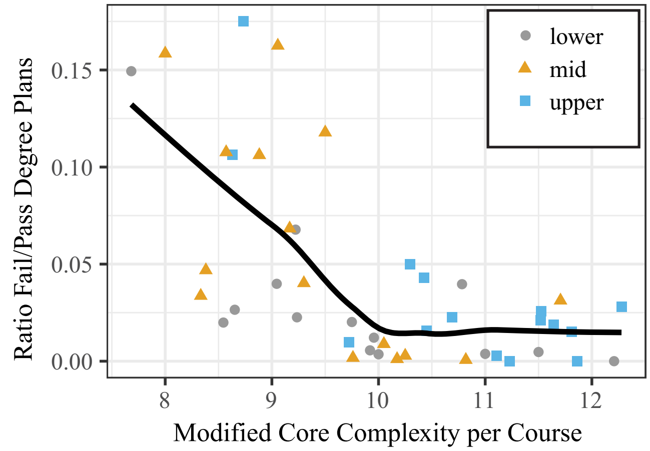

FIG. 6. The ratio of the number of degree plans available after failing Calculus 2 to the plans available if Calculus 2 is passed plotted against the modified core complexity per course.

2. Delay

Both the blocking and delay components of complexity should be related to the general effect of failing a class on a student's progress through a curriculum. Failing a class, particularly a required chemistry, mathematics, or physics class, may delay graduation by the semester required to retake the class. Students needing additional mathematics courses before entering Calculus 1 may experience similar delays. One measure of the ability of a curriculum to accommodate a semester delay is the number of degree plans requiring less than eight semesters to complete. Most institutions studied had few simulated degree plans requiring less than seven semesters; as such, we focus on the number of available seven-semester degree plans, DP_7. Figure 5 plots DP_7 against $\overline{C^{CM}}$. As did the total number of degree plans, DP_7 decreases exponentially with increasing $\overline{C^{CM}}$. The regression coefficient β_1 fitting $\ln(DP_7) = \beta_0 + \beta_1 \cdot \overline{C^{CM}}$ is commensurate to that found in the previous section, $\beta_1 = -2.58$. As such, the number of seven-semester degree plans decreases by $e^{\beta_1} = 0.08$ or 92% for each unit increase in $\overline{C^{CM}}$.

Most programs had few degree plans requiring 6 or fewer semesters, particularly compared to the 7- and 8-semester degree plans, with 28 of the 47 calculated programs having no 6 or fewer semester degree plans. The modified longest path sets the minimum time to degree. As such, this result is consistent with Fig. 3 which shows 30 of the 60 programs studied with minimum modified longest path of 7 higher. Most programs with modified longest path 6 or smaller had some, but not many, six-semester degree plans.

3. Failing a class

The degree to which CA can predict the effect of failing a class on the curricular options open to a student can be more directly probed by simulating failing a class. For the curricula which could be calculated, a related curriculum simulating the effect of failing Calculus 2 was created. Calculus 2 was selected because it was the central course for the majority of the 60 curricula studied. To simulate failing Calculus 2, a second Calculus 2 class was added to the curriculum. This new class was made prerequisite or corequisite to any course which had the original Calculus 2 as a prerequisite or corequisite. The new class had the original Calculus 2 as its prerequisite. Figure 6 plots the ratio of the number of degree plans available if Calculus 2 failed to the number available in the original curriculum. For all curricula, failing Calculus 2 greatly reduced the number of degree plans available. The lowest reduction was 82% (107 000 initial degree plans reduced to 19 000 degree plans). Three programs had no remaining degree plans if Calculus 2 was failed (initial degree plans 35 000, 170 000, and 224 000). The fraction of degree plans remaining decreases approximately linearly with increasing $\overline{C^{CM}}$ until $\overline{C^{CM}} = 10$. After this point, the fraction of degree plans remaining stops changing with $\overline{C^{CM}}$; the average remaining degree plans after this point is about 2% of the original plans with many programs having zero or nearly zero plans remaining. This suggests $\overline{C^{CM}} = 10$ as a threshold above which curricula are likely to be very unforgiving to students failing a core class. For the 19 programs with less than 2% remaining degree plans, the number of remaining plans had a broad range; 5 programs had less than 100 degree plans remaining, 13 had less than 20 000, while 3 had more than 1 000 000 remaining plans. The core complexity per course shows the same behavior with the plot leveling out at $\overline{C^{C}} = 10$. With the large number of available degree plans, even 2% of those plans may still be a substantial number of plans; however, for each student, the number of degree plans still available is greatly restricted by the courses they have already taken (and not taken). As such, it should be beneficial to maintain as much flexibility as possible to meet the individual needs of as many students as possible.

Supplemental Material [28] contains additional analysis of the relation of the number of degree plans with and without failing Calculus 2. A log-log plot of the number of degree plans if Calculus 2 is failed, DP_F, versus the number of degree plans if Calculus 2 is not failed, DP, is presented. The plot indicates a quasilinear relationship. A regression analysis suggests the relationship $DP_F = 0.0022 \cdot DF^{1.12}$, ($R^2 = 0.77$), further supporting the dramatic reduction of student flexibility if a central class is failed.

C. Modifications to curricular structures

In the previous two sections, which examined the curricula of many institutions, it was impossible to explore the effects of specific curricular changes. To examine the quantitative effect of common curricular modifications, we focus on a single middle tier institution called Middle Tier Public University (MTPU). The curriculum graph for the physics program at this institution is shown in Fig. 1 of Supplemental Material [28], and a common degree plan for this institution is presented in Fig. 1.

Seven curricular modifications were made to the MTPU curriculum, labeled Curriculum 1 to 7. The modifications maintained the number of classes. Each modification represents a change that had at one time been considered by the department. The modifications were as follows:

Curriculum 0 This is the current MTPU curriculum. Calculus 1 is a prerequisite for Physics 1. Introductory Physics 2 is used as the prerequisite for junior level classes such as Electromagnetism 1 (EM1) and Theoretical Mechanics 1 (TM1). Modern Physics is the prerequisite for more advanced classes such as Quantum Mechanics (QM) or Statistical Mechanics (SM).

Curriculum 1 This curriculum modified Curriculum 0 to make Calculus 1 a corequisite for Physics 1.

Curriculum 2 This curriculum modified Curriculum 0 to make Calculus 2 a prerequisite for Physics 1.

Curriculum 3 This curriculum modified Curriculum 0 to make Calculus 2 a prerequisite for Physics 2.

Curriculum 4 This curriculum modified Curriculum 0 to reduce the mathematics requirement for TM1 and EM1 from differential equations to multivariable calculus.

Curriculum 5 This curriculum modified Curriculum 0 by converting Electricity and Magnetism 2 into a linear algebra course with a Calculus 2 prerequisite. This course was made a prerequisite of TM1, EM1, and QM1.

Curriculum 6 This curriculum modified Curriculum 0 by adding an EM1 and TM1 prerequisite to QM and a TM1 prerequisite to SM.

Curriculum 7 This curriculum modified Curriculum 0 by making Calculus 2 a corequisite to Physics 1 and relaxing the math prerequisite of TM1 and EM1 to multivariable calculus.

The existing curriculum, Curriculum 0, had a number of features designed to allow students flexibility: using Physics 2 instead of modern physics as the prerequisite of TM1 and EM1 and having only differential equations and modern physics as the prerequisite to quantum mechanics. Table II presents curricular metrics for the modified curricula. Beyond metrics already introduced, the number of classes on modified longest paths of lengths less than 5, 6, 7, and 8 are presented as N_i^C. The overall number of degree plans, the number of degree plans requiring seven or fewer semester, $DP_{\leq 7}$, the degree plans requiring eight semesters, DP_8, and the ratio of the seven semester plans to the total number of plans are also presented. The range of the number of degree plans through the curricular modifications is broad. No modification created a modified longest path of length 8; only Curriculum 2 produced a longest path of length 7. For all curricula, the seven-semester degree plans were a small fraction of the total degree plans suggesting students not ready for Calculus 1 or students failing a class face a more restricted set of curricular choices if they wish to graduate in 4 years.

The modifications to the MTPU curriculum provide a more controlled way to examine how the number of degree plans changes with complexity. Figure 7 plots the number of degree plans against (a) the modified core complexity, (b) the modified core complexity per course, and (c) the core complexity on a logarithmic scale. All plots are fairly linear; a regression line is provided for each plot. The fit in

TABLE II. Curricular metrics for modifications to the MTPU curriculum.

Curriculum	C^C	$\overline{C^C}$	C^{CM}	$\overline{C^{CM}}$	$N_{\leq 5}^C$	N_6^C	N_7^C	N_8^C	DP	$DP_{\leq 7}$	DP_8	$DP_{\leq 7}/DP$
0	163	9.06	163	9.06	5	13	0	0	1 294 316	28 818	1 265 498	0.022
1	163	9.06	156	8.67	10	8	0	0	4 610 423	152 229	4 458 194	0.033
2	176	9.78	176	9.78	3	8	7	0	183 158	2090	181 068	0.011
3	165	9.17	165	9.17	5	13	0	0	1 161 755	27 408	1 134 347	0.024
4	152	8.44	152	8.44	12	6	0	0	6 782 182	197 411	6 584 771	0.029
5	160	8.89	160	8.89	7	11	0	0	2 965 225	82 082	2 883 143	0.028
6	169	9.39	169	9.39	3	15	0	0	470 466	9552	460 914	0.020
7	152	8.44	144	8.00	18	0	0	0	25 424 332	1 082 911	24 341 421	0.043

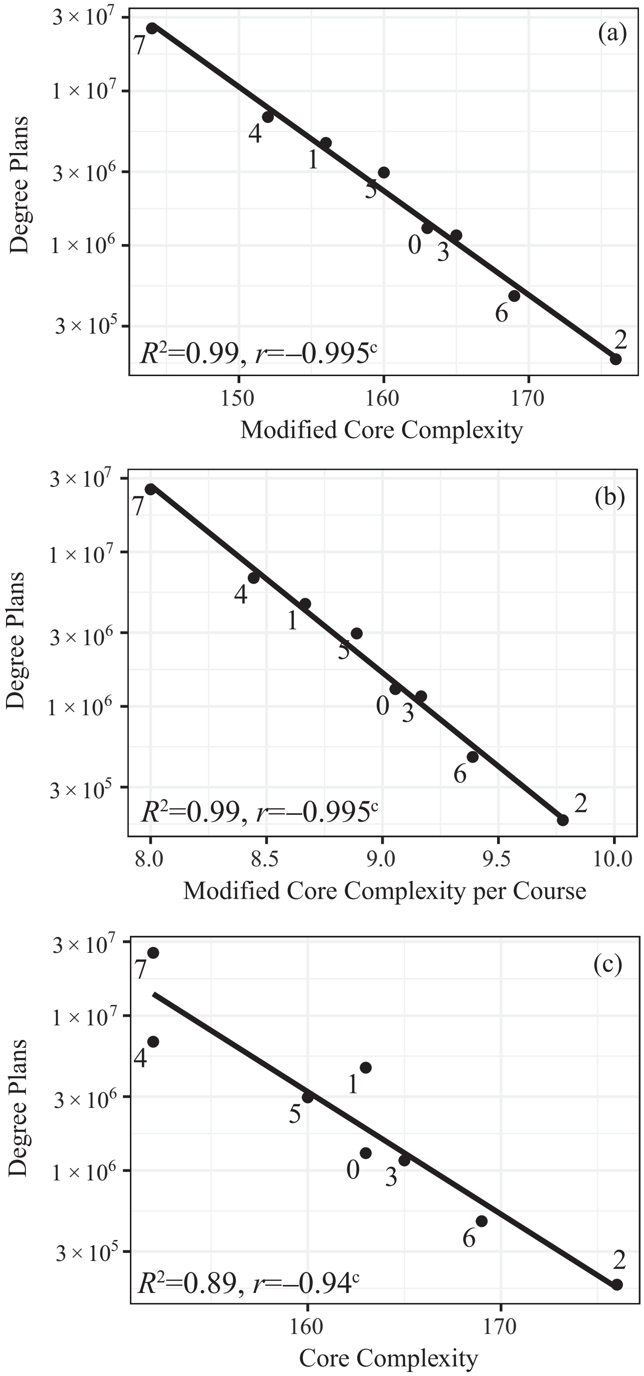

FIG. 7. Degree plans plotted against complexity for modifications of the MTPU curriculum. (a) Plots the number of degree plans against the modified core complexity, (b) the number of degree plans against the modified core complexity per course, and (c) the number of degree plans against the core complexity. All plots are on a logarithmic scale. The points are labeled with the curriculum number.

Fig. 7(a) has superior R^2 to that in Fig. 7(c) providing additional evidence that the modified core complexity is a better metric than the original definition of complexity.

A linear regression predicting $\ln(DP)$ using modified core complexity as the independent variable produces the equation $\ln(DP) = 39.4 - 0.16 \cdot C^{CM}$. Exponentiating yields, $DP = e^{39.4} \cdot e^{-0.16 \cdot C^{CM}}$; therefore, a 10 unit increase in modified core complexity multiplies the number of available degree plans by $e^{-0.16\times 10} = 0.20$. Curricula with a 10 point higher modified core complexity have only 20% of the degree plans of a lower complexity program. Likewise, predicting $\ln(DP)$ with $\overline{C^{CM}}$ yields $\ln(DP) = 39.4 - 2.8 \cdot \overline{C^{CM}}$; exponentiating yields, $DP = e^{39.4} \cdot e^{-2.8 \cdot \overline{C^{CM}}}$. A one unit increase in modified core complexity per course multiplies the number of available degree plans by $e^{-2.8} = 0.061$ leaving only 6% of the degree plans. As such, relatively small changes in curricular metrics generate large changes in the number of available pathways through the degree. These results were very similar to those found in Sec. III B 1 suggesting the quantitative relation between degree plans and curricular metrics may be fairly general.

The seven curricular changes explored produce modest changes to curricular metrics. Supplemental Material [28] provides a detailed example of a broader change in curricular structures which reduced complexity from 290 to 222.

Examination of the MTPU curriculum allows the exploration of two additional factors affecting the complexity of the curriculum experienced by a physics student: the role of math readiness and degree tracks. For all institutions studied, the first required mathematics course in the curriculum is Calculus 1. The analysis of math readiness and the analysis of degree tracks are presented in Supplemental Material [28]. As would be expected, the complexity increased approximately linearly as the number of required math courses increased (Fig. 3 in Supplemental Material [28]).

Many institutions offer degree tracks, sometimes called areas of emphasis or concentrations, to give students the opportunity to specialize in a subfield of interest. These degree tracks generally replace some physics electives with required courses and can require courses outside the physics department including mathematics, computer science, and engineering courses. Altering the general curriculum to accommodate these degree tracks influences the overall complexity. For math-ready students, there was a difference of 75 complexity points between the least and most complex degree tracks at MTPU. This is commensurate with or larger than the complexity difference between the 25% and 75% programs in each tier in Fig. 2. As such, the degree tracks at MTPU differ substantially in complexity compared to the overall variation in complexity of physics programs. Figure 2 in Supplemental Material shows the linear trend of increasing complexity per additional math course was observed for all degree tracks at MTPU.

IV. DISCUSSION

A. Research questions

This work explored three research questions which will be addressed in the order proposed.

RQ1: How do curricular metrics vary for U.S. physics programs? Is there a correlation between program ranking and curricular metrics? The replication of the Heileman *et al.* [20] study showed no significant difference in complexity between the three tiers. As such, the trends observed in electrical engineering and computer science were not observed in physics programs. This might lead to the conclusion that physics curricula do not vary across institutional rankings. Further analysis showed that this null result was a combination of multiple factors. The inclusion of general education electives obscured differences in the core components of the curricula: the physics, mathematics, and chemistry courses. Using only core classes, differences between the three tiers emerged. While differences in the core complexity remained insignificant, there was a significant difference in core complexity per course and the number of core courses between the tiers. The core complexity per course was significantly higher for upper tier institutions than lower and mid tier institutions. Upper tier institutions required on average fewer core courses than lower and mid tier institutions (with the only significant difference being between the upper and lower tier). Upper tier institutions also had more seven- and eight-semester modified longest paths than those in the other tiers. These differences were not significant when the full curricula was used.

It is possible that engineering programs are generally more complex because of a larger number of core courses required and the prerequisite structure these courses share, perhaps because of the requirements of ABET accreditation. This may have resulted in the qualitative differences between the results of the present study and those of Heileman *et al.* [20]. The core complexities for the electrical engineering and mechanical engineering degrees at MTPU are 347 and 343, respectively; the programs are much more complex than the physics program at MTPU. These core complexity scores for engineering are higher than the most complex physics institutions' complexity but are within the mid tier's 25% to 75% range in the study of Heileman *et al.*

Beyond focusing on the core curriculum and disaggregating the competing effects of the number of required core courses and the core complexity per course, a modified definition of the longest path was proposed. This modified definition counted corequisite relations as zero length in the longest path calculation; in the original definition, they are counted as length 1. The original definition yields equal complexity for both corequisite and prerequisite relations. This modified definition better predicted the available degree plans while making the longest path metric more useful (longest paths longer than eight semesters could not be completed in 4 years).

RQ2: How are curricular metrics related to other measures of curricular flexibility? The core complexity was not a good measure to quantify differences in curricular flexibility measured by the number of available degree plans. Figure 4(a) shows little relation between the core complexity and the number of available degree plans. The core complexity was separated into two competing components: the number of core courses and the core complexity per course. Using both the modified and original longest path definition, as the core complexity per course increased, the number of available degree plans exponentially decreased. For the curricula that could be enumerated, a one unit increase in core complexity per course reduced the number of available degree plans by 80%.

One possible benefit of a flexible curriculum is that students, who are not math ready or students who fail a class, may still complete the degree in eight semesters. Because failing a class often delays progression to degree by a semester, the number of degree plans that allow graduation in seven or fewer semesters should be a good measure of the degree to which the curriculum accommodates these challenges. The number of seven-semester degree plans also exponentially decreased with increasing modified core complexity (and original core complexity) with a commensurate decrease in the number of seven-semester plans with a one unit increase in modified core complexity per course.

The change in the number of available degree plans when failing a course was also examined by simulating the effective curriculum for students who fail Calculus 2. Figure 6 shows the ratio of degree plans remaining to the student to the original degree plans decreases linearly until the modified core complexity per course reaches 10; after this point, it becomes much more difficult for students to complete a degree in eight semesters after failing Calculus 2.

RQ3: How much do curricular changes affect curricular flexibility? In both the degree plans of the 47 programs which could be calculated and the controlled modifications to the MTPU curriculum, the number of degree plans decrease exponentially with increasing modified core complexity. For the 47 programs, a one unit increase in modified core complexity decreased the number of available degree plans by 89%; for the changes to a single program made to the MTPU curriculum, a one unit increase in modified core complexity per course reduced the number of degree plans by 94%, As such, curricular modifications that increase modified curricular complexity by even one unit represent changes that dramatically restrict student flexibility.

B. Qualitative comparison of physics programs

While this work focused on the evaluation of quantitative curricular metrics, physics programs may also benefit from a general description of the similarities and differences found in the 60 programs. The greatest difference between the complexity of any two programs in the analysis is 178. In the study by Heileman *et al.*, where electrical engineering programs were compared across tiers [20], the difference between the least and most complex curriculum analyzed was over 400 complexity points. These differences in physics curricular complexity can often be attributed to the number of elective courses required beyond the core requirements for a physics degree; however, some programs had extra intermediate required physics or mathematics classes.

Most of the 60 physics programs analyzed require Calculus 1 through differential equations, and 39 of them also required a linear algebra course. Several require additional upper level math electives beyond the typical calculus and differential equations sequence. Of the 60 programs in the analysis, 44 required a two-course introductory physics sequence while 15 required a three course introductory sequence and one institution required a four-course introductory sequence. 37 of the institutions required Calculus 1 as a prerequisite for the first introductory physics course Physics 1, and nine institutions allowed Calculus 1 to be taken as a corequisite with Physics 1. Eight institutions required Calculus 2 as a prerequisite to Physics 1, and six institutions allowed Calculus 2 to be taken as a corequisite with Physics 1. All but 3 of the institutions required a modern physics course, and 11 of the institutions required two modern physics courses. For five of the institutions that required a three- or four-course introductory sequence, modern physics was included as one of the introductory courses; otherwise, it was required in addition to the introductory sequence. Forty one of the institutions required a mathematical methods in physics course (some required 2 mathematical methods courses), and 28 programs required computational physics or computational methods courses, with 6 more programs offering such a course as an elective. Only 16 programs required a specific wave mechanics course, although several institutions mention wave mechanics as a topic in one of the introductory physics courses. Twenty institutions required an advanced optics course and 30 offered it as an elective, while 10 programs had no such course offered in their catalog. Each of the programs required some form of an advanced laboratory course, and many of the programs required two advanced laboratory courses. Each program required an advanced course (beyond the level of the introductory physics course sequence) in the topics of classical mechanics, electricity and magnetism, and quantum mechanics. Thirty programs required two courses in advanced electricity and magnetism, 24 offered a second course in electricity and magnetism as an elective, and 6 programs did not offer a second advanced electricity and magnetism course. Twenty seven programs required a second advanced quantum mechanics course, 15 offered one as an elective, and 18 programs did not offer a second quantum mechanics course. Thirty eight programs only offered one course in advanced classical mechanics, while 11 required a second classical mechanics course, and 11 offered a second course in classical mechanics as an elective.

V. USING CURRICULAR ANALYTICS

Curricular reform is a complex process that involves quantitative analysis, qualitative considerations of program goals, and holistic input from students and faculty. CA provides all academic units with quantitative metrics to characterize the structural complexity of a curriculum. The current study introduced new metrics related to structural complexity which were needed to understand curricular differences in physics programs. The breadth of curricular structures present in U.S. physics programs was also presented in Fig. 2 and the qualitative discussion in the previous section. Physics departments may use these results to help inform the revision of their programs as follows:

Quantitatively characterize the program: Identify the number of core courses in the program. Using the tools available on the CA website, compute the core complexity and the core complexity per course. Examine Fig. 2 to determine where the program lies in the spectrum of U.S. physics programs. If the program is above the 75th percentile in either the number of core courses and the core complexity per course, examine program requirements to determine if there are points where the program could be simplified without losing quality. The majority of physics programs have found it possible to provide quality instruction with lower levels of these metrics. If modified complexity is made available at the CA website, use this metric.

Examine longest paths: Use the CA website to explore the longest path through each course. For courses with longest paths of length 8 or higher, determine the actual number of semesters required to complete the sequence considering corequisite relationships. Consider modifying curricula to remove eight (and potentially seven) semester paths to lower the risk that failing a course will automatically lead to the extension of time to degree.

Consider the breadth of physics curricula: Consider the qualitative discussion of the kinds of curricular structures found in the 60 programs examined presented in the previous section. Perhaps some of these structures, all found in physics programs at the time of writing, could help simplify the physics curricula under consideration.

Examine curricular changes: Try some curricular changes. Use the exponential relation between core complexity per course and available degree plans to evaluate the magnitude of the effect of the change on curricular flexibility. Evaluate the learning outcomes of the courses and ask if the prerequisite structure makes sense. Examine the changes made to the MTPU curriculum to develop an understanding of the magnitude of change one might expect from curricular changes.

Avoid overly complex curricula: If the core complexity per class is above the threshold of 10 shown in Fig. 4, consider revision. Curricula above this threshold provide the students with few options if an introductory mathematics, chemistry, or physics course is failed.

VI. LIMITATIONS AND FUTURE WORK

The rankings of the physics programs in tiers were taken from the 2022 U.S. News rankings of the best physics graduate schools [26]. These rankings are the product of a survey conducted by U.S. News that asked department chairs and department directors of graduate studies to rank schools with physics Ph.D. programs from 1 (marginal) to 5 (outstanding). The response rate for this survey in physics was 27.9%. If a school received less than ten ratings, it was not included in the rankings. This is not a scientific determination of hierarchy among physics programs in the United States but rather is a ranking based upon popular opinion and public perception. We feel this is still a useful tool and that most would agree that the groups of randomly selected institutions in the upper, middle, and lower tiers are approximately in the same general order as would be accomplished by a more rigorous classification system.

A program's curricular complexity is only part of the structural features influencing student success; each student must fit the curricular requirements into an eight-semester degree plan. The semester in which a class is offered (spring or fall) and the frequency the class is offered (every semester, every year, every other year) can further impact time to degree. All the results presented in Sec. III B represent the best-case scenario, where it is assumed that every course could be taken in every semester. For many institutions, this is not the case. Many courses beyond the introductory courses tend to be offered once a year while elective courses are often offered every other year. We plan to investigate this important effect in a future study.

This study also focuses on structural complexity while ignoring instructional complexity. Instructional complexity may alter the relation of curricular complexity to student success. Instructional complexity encompasses all aspects of a course's delivery and environment. Development of robust metrics of instructional complexity will be explored in future work.

VII. CONCLUSIONS

This study applied CA to compare undergraduate physics programs at 60 academic institutions in the United States, separated into three tiers based on the U.S. News and World Report rankings. There was no significant relationship between program ranking and program complexity. Further analysis showed this resulted from two competing effects: lower ranked programs requiring more mathematics, chemistry, and physics classes and higher ranked programs having more restrictive prerequisite structures. Both differences were statistically significant and large effects. To capture these effects, which were not detected using the metrics introduced in the original CA description, additional metrics were introduced, most importantly the core complexity per course. Further, CA treats prerequisite and corequisite relations as equivalent. A revised definition of the longest path, which is used to calculate the delay factor, was introduced; this metric in general had superior performance to the original longest path definition.

Using the additional metrics, an exponentially decreasing relationship was found between the number of available degree plans and increasing core complexity per course. Modifications to the MTPU curriculum demonstrated similar results. These findings support the hypothesis that higher core complexity per course leads to reduced student flexibility and suggest that core complexity per course should be considered when developing or refining curricula.

With these additional metrics and modifications, CA could be a valuable tool to allow a physics department to examine the potential effects of curricular changes under consideration.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under Grants No. ECR-1561517, No. HRD-1834569, and No. DUE-1833694 and by a grant to West Virginia University from the Howard Hughes Medical Institute through the Science Education Program. We also thank Andrew Heckler for introducing us to CA. We also thank the referees whose suggestions greatly improved this work.

[1] J. L. Docktor and J. P. Mestre, Synthesis of discipline-based education research in physics, Phys. Rev. Phys. Educ. Res. **10**, 020119 (2014).
[2] A. L. Traxler, X. C. Cid, J. Blue, and R. Barthelemy, Enriching gender in physics education research: A binary past and a complex future, Phys. Rev. Phys. Educ. Res. **12**, 020114 (2016).
[3] I. Rodriguez, E. Brewe, V. Sawtelle, and L. H. Kramer, Impact of equity models and statistical measures on interpretations of educational reform, Phys. Rev. Phys. Educ. Res. **8**, 020103 (2012).
[4] J. M. Aiken, R. Henderson, and M. D. Caballero, Modeling student pathways in a physics bachelor's degree program, Phys. Rev. Phys. Educ. Res. **15**, 010128 (2019).
[5] J. Stewart, J. Hansen, and E. Burkholder, Visualizing and predicting the path to an undergraduate physics degree at two different institutions, Phys. Rev. Phys. Educ. Res. **18**, 020117 (2022).
[6] T. Hodapp, The back page: The economics of education: Closing undergraduate physics programs, APS news **20**, 8 (2011).
[7] G. L. Heileman, C. T. Abdallah, A. Slim, and M. Hickman, Curricular analytics: A framework for quantifying the impact of curricular reforms and pedagogical innovations, arXiv:1811.09676.
[8] S. Salehi, E. Burkholder, G. P. Lepage, S. Pollock, and C. Wieman, Demographic gaps or preparation gaps?: The large impact of incoming preparation on performance of students in introductory physics, Phys. Rev. Phys. Educ. Res. **15**, 020114 (2019).
[9] R. Henderson, J. Stewart, and A. Traxler, Partitioning the gender gap in physics conceptual inventories: Force Concept Inventory, Force and Motion Conceptual Evaluation, and Conceptual Survey of Electricity and Magnetism, Phys. Rev. Phys. Educ. Res. **15**, 010131 (2019).
[10] J. Stewart, G. L. Cochran, R. Henderson, C. Zabriskie, S. DeVore, P. Miller, G. Stewart, and L. Michaluk, Mediational effect of prior preparation on performance differences of students underrepresented in physics, Phys. Rev. Phys. Educ. Res. **17**, 010107 (2021).
[11] P. R. Aschbacher, E. Li, and E. J. Roth, Is science me? High school students' identities, participation and aspirations in science, engineering, and medicine, J. Res. Sci. Teach. **47**, 564 (2010).
[12] Z. Hazari, R. H. Tai, and P. M. Sadler, Gender differences in introductory university physics performance: The influence of high school physics preparation and affective factors, Sci. Educ. **91**, 847 (2007).
[13] L. Stiles-Clarke and K. MacLeod, Demystifying the scaffolding required for first-year physics student retention: Contextualizing content and nurturing physics identity, Can. J. Phys. **96**, 29 (2018).
[14] J. P. Zwolak, R. Dou, E. A. Williams, and E. Brewe, Students' network integration as a predictor of persistence in introductory physics courses, Phys. Rev. Phys. Educ. Res. **13**, 010113 (2017).
[15] J. Forsman, R. Moll, and C. Linder, Extending the theoretical framing for physics education research: An illustrative application of complexity science, Phys. Rev. ST Phys. Educ. Res. **10**, 020122 (2014).
[16] R. Molontay, N. Horváth, J. Bergmann, D. Szekrényes, and M. Szabó, Characterizing curriculum prerequisite networks by a student flow approach, IEEE Trans. Learn. Technol. **13**, 491 (2020).
[17] N. W. Klingbeil and A. Bourne, The Wright State model for engineering mathematics education: Longitudinal impact on initially underprepared students, in *Proceedings of the 2015 ASEE Annual Conference & Exposition, Seattle, WA* (American Society for Engineering Education, Washington, DC, 2015), pp. 26–1580.
[18] A. Slim, G. L. Heileman, C. T. Abdallah, A. Slim, and N. N. Sirhan, Restructuring curricular patterns using Bayesian networks, in *Proceedings of the 14th International Conference on Educational Data Mining, EDM 2023, Paris, France* (International Educational Data Mining Society, Boston, MA, 2021).
[19] A. Slim, H. Al Yusuf, N. Abbas, C. T. Abdallah, G. L. Heileman, and A. Slim, A Markov decision processes modeling for curricular analytics, in *Proceedings of the 2021 20th IEEE International Conference on Machine Learning and Applications, ICMLA, Pasadena, CA* (IEEE, New York, 2021), pp. 415–421.
[20] G. L. Heileman, W. G. Thompson-Arjona, O. Abar, and H. W Free, Does curricular complexity imply program quality?, in *Proceedings of the 2019 ASEE Annual Conference & Exposition, Tampa, FL* (American Society for Engineering Education, Washington, DC, 2019).
[21] G. L. Heileman, H. W. Free, J. Flynn, C. Mackowiak, J. W. Jaromczyk, and C. T. Abdallah, Curricular complexity versus quality of computer science programs, arXiv:2006.06761.
[22] D. Reeping, D. M. Grote, and D. B. Knight, Effects of large-scale programmatic change on electrical and computer engineering transfer student pathways, IEEE Trans. Ed. **64**, 117 (2020).
[23] G. L. Heileman, C. T. Abdallah, and A. K. Koch, The transfer student experience: It's a lot like buying a used car, arXiv:2203.00610.
[24] A. M. DeRocchis, L. E. Boucheron, M. Garcia, and S. J. Stochaj, Curricular complexity of student schedules compared to a canonical degree roadmap, in *Proceedings of the 2021 IEEE Frontiers in Education Conference (FIE), Lincoln, NE* (IEEE, New York, 2021), pp. 1–5.
[25] A. Slim, G. L. Heileman, M. Hickman, and C. T. Abdallah, A geometric distributed probabilistic model to predict graduation rates, in *Proceedings of the 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), San Francisco, CA* (IEEE, New York, 2017), pp. 1–8.
[26] U.S. News & World Report: Education, US News and World Report, Washington, DC, https://premium.usnews.com/best-colleges [accessed July 23, 2022].
[27] National Center for Education Statistics, https://nces.ed.gov/collegenavigator [accessed January 30, 2022].
[28] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.20.020143 for an example of how a physics curriculum could be

simplified, for an exploration of math readiness and degree tracks, and for additional explanation of the statistics used.

[29] Curricular analytics, https://curricularanalytics.org/. Accessed [accessed May 30, 2022].

[30] D. C. Hoaglin, F. Mosteller, and J. W. Tukey, *Understanding Robust and Exploratory Data Analysis* (John Wiley & Sons, New York, 1983).

[31] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (Academic Press, New York, NY, 1977).

[32] John Stewart, Degree plan simulator, https://github.com/johns496/DegreePlanSimulator (2024).

[33] R Core Team, *R: A Language and Environment for Statistical Computing*, (R Foundation for Statistical Computing, Vienna, Austria, 2017).