Ellinas, C., Allan, N. D., & Johansson, A. F. (2016). Towards Project Complexity Evaluation: A Structural Perspective. IEEE Systems Journal. DOI: 10.1109/JSYST.2016.2562358

Towards Project Complexity Evaluation: A Structural Perspective

Christos Ellinas, Neil Allan, Anders Johansson

***Abstract* - Complexity is often quoted as an independent variable that challenges the utility of traditional project management tools and techniques. A large body of work has been devoted in exposing its numerous aspects, yet means for quantitatively assessing it have been scarce. Part of the challenge lies in the absence of hard evidence supporting the hypothesis that projects can be considered as complex systems, where techniques for measuring such complexity are better established. In response, this work uses empirical activity networks to account for the technological aspect of five projects. By doing so, the contribution of this work is two-fold. First, a procedure for the quantitative assessment of an aspect of project complexity is presented; namely structural complexity. Second, results of the analysis are used to highlight qualitatively similar behavior with a well-known complex system, the Internet. As such, it suggests a transition from the current, metaphorical view of projects being complex systems to a literal one.**

From a practical point of view, this work uses readily-captured and widely-used data, enabling practitioners to evaluate the structural complexity of their projects to explore system pathologies and hence, improve the decision making process around project bidding, resource allocation and risk management.

***Index Terms*— complex networks, project management, risk analysis, complex systems engineering, project engineering**

I. Introduction

On March 17th, 2000, a lightning bolt struck a Royal Philips Electronics semiconductor plant, leading to a ten-minute fire [1]. A small, random and rather minor event was enough to shock the status quo of the global cellular telecommunication industry. Nokia and Ericsson, the two dominant companies in the area, both sourced microchips from that plant, though under different supply chain strategies. Nokia quickly shifted production requirements to other suppliers, while Ericsson was trapped due to its single-source strategy. As a consequence, Ericsson reported a 3% market share decline resulting in financial losses of 400 million USD in sales in the impacted quarter, while its stock value tumbled by 14% in just a few hours [1, 2].

Single-point failures, such as the aforementioned case, are surprisingly common and observed in numerous, seemingly different, domains (e.g. single tree falling leading to extensive power outages [3, 4]; single financial institution failure leading to a collapse of the financial system [5, 6], single factory failing threatens global manufacturers [7, 8]). As a result, these systems are considered to be “Robust-Yet-Fragile” (RYF) [9-12], where random disruptions cause minimum damage, unless a disproportionally important component is affected.

These central components emerge from the highly heterogeneous nature of these systems; first noted in the pioneering work of Barabási and Albert [13] and more recently becoming a recognized feature of complex systems in general [14-17]. Consequently, it does not matter how large a single contribution is, in an absolute way, but rather how it compares to the overall ensemble of entries. In other words, the variance in the underlying distribution of individual contributions (or observations) is a necessary condition for complex phenomena (such as the RYF behavior) to emerge [10].

The objective of this work is to contextualize this line of enquiry in the project management literature, with a focus on engineering projects in general, and construction projects in particular. In this context, it is often presumed that projects can be considered to be complex systems – yet no hard evidence have been proposed to support this hypothesis [18]. In response, this work proposes a procedure for assessing an aspect of project complexity – structural complexity; see Section II, B – in a quantitative, evidence-based manner. Subsequently, results of an analysis involving five real-world engineering projects highlights the qualitative similarity with the behavior of a widely-recognized complex system – the Internet. As such, this work provides evidence supporting the transition from the current metaphorical view of project being complex system to a literal one.

II. Theoretical Background

A. Complexity Science

In his classic paper, Weaver [19] proposed that system complexity arises from the inability to accurately predict the properties of a system, even under the state of complete knowledge of the properties of its composing parts. As a

---

Manuscript received: TBD. This work was supported by the Systems Centre and the EPSRC funded Industrial Doctorate Centre (IDC) in Systems (Grant EP/G037353/1) and Systemic Consult Ltd.

Christos Ellinas is with the IDC in Systems, University of Bristol, Queens Building, Bristol BS8 1TR, UK, and Systemic Consult Ltd, Carpenter House, Bath, BA1 1UD, UK. (e-mail: ce12183@bristol.ac.uk).

Neil Allan is with the University of Bath, Quarry Rd, Bath, BA2 7AY, UK, and Systemic Consult Ltd, Carpenter House, Bath, BA1 1UD, UK (e-mail: mnsnda@management.bath.ac.uk).

Anders Johansson is with the IDC in Systems, University of Bristol, Queens Building, Bristol BS8 1TR, UK (e-mail: A.Johansson@bristol.ac.uk).

result, non-trivial, system-wide properties, such as the RYF, emerge.

Network science provides the tools for understanding how the architecture of a system allows such non-trivial behavior to emerge [20] (see [21, 22] for technical reviews of the field). Under this view, RYF is readily explained as the result of extreme *variance* (in fact, theoretically infinite) within the system, in terms of the connectedness of each individual component [10, 13]. In other words, the complexity of the system, as seen through the emergence of non-trivial behavior (e.g. RYF), resides in the extreme variance exhibited by a given *indicator* of the network structure (e.g. node degree).

The importance of variance as a contributing factor to complexity is widely recognized within the natural sciences [23, 24] (also see [10, 25] for examples on how it is used); similar views on its importance have occasionally surfaced within the management literature [26-28].

B. Project Management

During the past years, there has been a growing body of work around complexity and its role in project management [27-33]. Complexity has been described as an independent variable which limits the applicability of best practice tools and methods, as the means for achieving improved performance of project delivery [18]. Yet, the discussion around its relevance and potential benefits has been hindered by the lack of evidence supporting the hypothesis that a project can be equated to a complex system. As noted by the review of Geraldi, et al. [18] "not one of the publications identified under the heading of "complexity in projects" provided any evidence or justification that a project is a complex system (equivalence). We concur with the view that projects can exhibit many of the characteristics of complex systems (analogy), and there are insights to be gained from viewing projects through the lenses provided by the various complexity theories. However, equivalence has not been established. We believe that the discussion of complexity would benefit from work to clarify whether such equivalence is indeed justified, and under what circumstances". Fuelled by this vagueness, a multitude of attributes and indicators attempting to describe the nature of project complexity have flooded the research space – for example, Geraldi, et al. [18] report a total of 34 different manifestations for measuring one single aspect of project complexity. For clarity, we note that the term complexity does not refer to its everyday use (which can be discounted as "merely complicated" [29, 34]) but rather on the narrower view of complexity science [35], focusing on the "emergence, dynamics, non-linearity and other behaviors present in systems of interrelated elements" [18].

Part of the challenge lies in the multiplicity of sources deemed responsible for fuelling this complexity. In response, both industry and academia have focused their efforts in constructing frameworks that capture those aspects – see [36, 37] and [18] for respective reviews. Throughout the literature, structural complexity emerges as a core aspect (e.g. [24, 27, 29, 32-34]; also see [18] for its frequency within the literature. Specifically, *structural complexity* refers to the (potentially) non-linear interactions between the activities of a project [38]. Interestingly, structural complexity is the earliest aspect deemed to contribute to project complexity (Fig. 1) and has remained relevant to both academics (e.g. [18]) and practitioners (e.g. [39]) ever since. Hence, understanding (and measuring) the role of structural complexity is central in understanding the wider concept of project complexity.

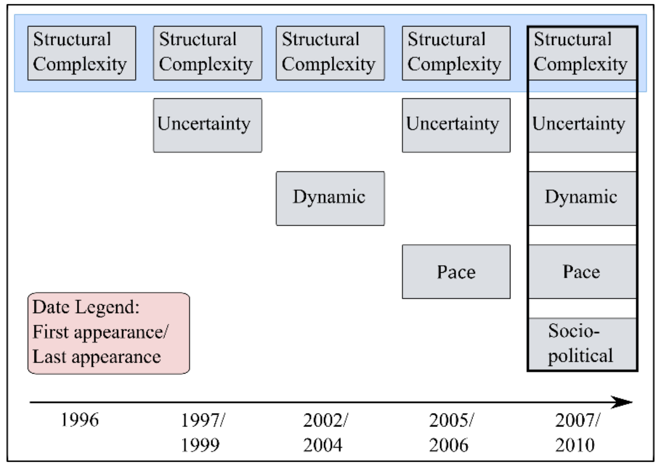

Fig. 1: Historical development of aspects contributing to project complexity, adapted from [18]. Note that since 2007, the nature of these aspects has converged.

Structural complexity, as with the of rest project complexity aspects, spans across several dimensions relevant to the project management process, such as organizational, environmental, *technological* etc. The technological dimension is of special interest, as it can provide the grounds for exploring structural complexity in a quantitative, evidence-based way. Baccarini [29] defines the technological dimensions as "the transformation process which converts inputs into outputs" and includes the act of task sequencing. As such, activity networks provide a suitable ground for capturing structural complexity of a project, across its technological dimension, in a quantitative, evidence-based way.

Activity networks are a key concept in a number of important project management tools. In practice, they are widely used to identify critical tasks (e.g. Critical Path Method, Program Evaluation and Review Technique) [42]; in academia they form the basis upon which various optimization problems are set [43-45]. The latter stream of work is of some relevance to this work, as it adopts a graph theoretic approach in an attempt to *measure* the perceived complexity of these networks [46]. Nonetheless, the applicability of this body of work is restricted within the domain of mathematical scheduling, where measuring the complexity of activity networks serves as a proxy on the algorithmic complexity of various solution procedures [47]. It does so by assuming that complexity is proportional to the hardness upon which a given algorithm can solve a given optimization problem (i.e. linear structure is desirable over a parallel one); an assumption debunked in [48].

III. Methodology

A. Overview

The premise of this work is grounded on the use of tools found within the study of complex systems throughout the field of natural sciences. Specifically, variability (quantified as variance [26]) is commonly used to classify a system as being complex (e.g. [13, 14, 23, 49]). As such, variance serves as the quantity of interest. By focusing on the technological dimension of projects (i.e. the transformation process of converting inputs to outputs) [29], activity networks are used to provide the underlying structure of a project. As such, activity networks form the subject of this quantitative analysis. In response to issues raised by Geraldi, et al. [18], the contribution of this work, to the domain of project management, is two-fold. First, evidence is provided in supporting the hypothesis that activity networks (and thus, to an extent, projects) can be considered as a complex system, where the latter is described as "a system composed of many interacting parts, such that the collective behavior of those parts together is more than the sum of their individual behaviors" [20]. Second, the proposed procedure provides a step forward in operationalizing the quantification of project complexity whilst using readily available data. By doing so, practitioners can identify, and proactively manage, complex engineering projects. Note that this work does not claim that other project aspects (e.g. supply chains, organizational learning requirement etc.) do not contribute towards their complex character. Rather, it focuses on quantifying the extent by which non-trivial technological dependencies (i.e. structural complexity) describes the transformation process that sustains an engineering project (i.e. across its technological dimension).

The adopted approach can be summarized as follows. Project schedules of five engineering projects are obtained and converted into activity networks, using the activity-on-node notation [50, 51]. Specifically, every project schedule corresponds to a directed graph $G = \{\{N\}\{E\}\}$, where every task i is abstracted as node $i, i \in N$ and a functional dependency between task i and j corresponds to a directed link $e_{i,j}$, where $e_{i,j} \in E$.

Subsequently, four indicators are defined, providing the means for measuring, in a direct or indirect way, the structural complexity of an activity network. Based on the distribution of each indicator, an artificial set of observations (using a suitably parameterized, truncated Normal distribution) is obtained - this set, corresponds to a homogeneous representation of the same indicator (which satisfies the research hypothesis – see Section III, C). The histogram of the empirical sample is subsequently compared to its empirical counterpart, using the Bhattacharyya measure. Based on this comparison, the variability of every indicator is directly assessed in order to decide whether it contributes to structural complexity - see Table 1 for an algorithmic description of the method (MATLAB implementation can be provided upon request). Note that results are to be plotted in a semi-log fashion, exposing the exponential character of the artificial set (sharp declining plot). Finally, similarities in behavior with other complex systems will be appropriately drawn.

TABLE I
ALGORITHMIC DESCRIPTION OF METHOD

Extract precedence data from raw Gantt chart
Compute adjacency matrix, **A**
Compute the indicator of interest – this is stored in vector I, with size $n \times 1$
Using I, parameterize the truncated Gaussian probability density function (pdf); see eq. 4
For x number of times
Sample an n number of (observations from the parameterized pdf; store it in vector R, with size $n \times 1$
Normalize R and I between intervals [0, 1], using eq. 1, where the minimum and maximum values are equal to $\min K$ and $\max K$ respectively, where $K = R \cup I$.
Compute the probability of obtaining each $P(R_i = \alpha)$, where $\alpha \in R$, using eq. 2.
Store in matrix $M(R_i, x)$, where each column is sorted in an ascending manner
End
Average matrix **M** across its second dimension; store resulting vector as $R2$
Use the Bhattacharyya measure to assess the similarity between the normalized histogram capturing the frequency of values in I and $R2$.

B. Data

Project management is a field dominated by professional associations [18]. Consequently, best practice, as advised by their respective Bodies of Knowledge, drives current data availability. This reality needs to be reflected by the data requirements of the procedure proposed herein otherwise there is a real risk of being theoretically valid, yet practically irrelevant. As such, project schedules, in the form Gantt charts [42], are used as the sole data input, as they adequately capture the technological aspect of a project [29] – see Fig. 2.

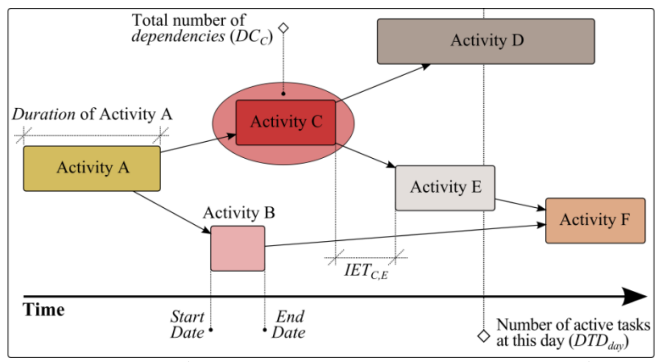

Fig. 2: An example of a project schedule, along with the typical information that it contains. Of relevance to this work are: a) the total number of links of a given task; b) the extend upon which a task keeps the network together; c) the free float between two tasks (referred to as the Inter-Event Time, IET) and d) the number of active tasks per day. For details see Section II, D (1).

Engineering projects with a focus on technical activities form the majority of modern organization activity [52], with construction projects being a typical example. As such, five real-world projects are considered, with their respective Gantt charts being produced at various stages of the project. Specifically, Project 1 corresponds to the delivery of an educational institution with an agreed cost of approximately 15,000,000 USD and an expected duration of 366 days. The Gantt chart used herein was produced 40 days after the project was launched. In terms of its activity network, it is composed of 5 connected components, where 935 nodes are connected through 1070 links. Project 2 corresponds to a commercial office complex, with an original contract sum of approximately 13,000,000 USD and an expected duration of 744 days. The Gantt chart used herein was produced 603 days after its launch. Its respective activity network is composed of 47 connected components, containing a total of 833 nodes and 806 links. Project 3 corresponds to an extension and renovation project, with an original contract sum of approximately 3,000,000 USD and expected duration of 596 days. The Gantt chart used herein was produced 8 days after the launch of the project. Its respective activity network is composed of 12 connected components, with 521 nodes and 563 links. Project 4 corresponds to an undisclosed project (undisclosed cost) with an expected duration of 418 days. The Gantt chart used herein was produced 45 days after the project launch. Its respective activity networks is composed of 79 connected components, with a total of 774 nodes and 822 links. Finally, Project 5 corresponds to the delivery of a commercial office complex, with an original contract sum of approximately 5,000,000 USD and an agreed duration of 549 days. The Gantt chart used herein was produced 12 days *before* the launch of the project. Its respective activity network is composed of 26 connected components, with a total of 326 nodes and 435 links.

C. Hypothesis

Variance is used as the means to evaluate the variability of an indicator, where the indicator captures a specific feature of the structure of the activity network[1]. This general interpretation will serve as the grounding principle for developing the research hypothesis (H1):

H1: Activity networks of engineering projects do not exhibit extreme variation in their topological structure.

In the case of a homogenous sample (H1 is true), the resultant distribution will resemble a *fast-decaying, probability density function* (FD_{pdf}) with a very thin tail (e.g. Normal, exponential etc.). On the other hand, if the sample is highly heterogeneous (*H1* is false), the resulting distribution will resemble a *slow-decay*, *probability density function* (SD_{pdf}) with a fat-tail (e.g. power law, log-normal distributions etc.) - see [53-55] for further discussion on the implications of identifying the specific functions. It is worth noting that from a risk management point of view, appreciating the difference between a FD_{pdf} and SD_{pdf} is key as it affects one's ability to confidently dismiss the existence of disproportionately important components within a system [10, 56] and thus, assess the likelihood of a system undertaking a systemic failure [57].

D. Method

1) Indicators

By focusing on the structure of the activity network, four indicators are considered; three are directly drawn from the domain of network science, with remaining being a new contribution – see Table II. Specifically, the first two indicators perform direct measurements on the topology of the network – they are referred to as the degree centrality (DC) and betweenness centrality (BC). The remaining two indicators emphasize on the indirect implications of the structure, from a temporal point of view[2]. These are the inter-event time (IET) and the daily task density (DTD).

In order to avoid potential scaling issues due to the different activity network sizes, each set of measurements is normalized accordingly:

$$X_i' = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

where X represents the vector containing the appropriate set of observations (i.e. empirical (I) or artificial $(R2)$), where each individual observation within X is denoted by the subscript i, where $i = 1,2,3,\dots,n$ and $i \in X$.

TABLE II

INDICATORS USED TO MEASURE STRUCTURAL COMPLEXITY

Measurement of Structural Complexity	Indicator
Direct	Degree Centrality (DC)
Direct	Betweenness Centrality (BC)
Indirect	Daily Task Density (DTD)
Indirect	Inter-Event Time (IET)

DC_i refers to the number of connections node i has. The extreme variance (and consequently, deviation from Normal distribution) noted across a wide range of real-world systems [53], first noted in [13], provided the first evidence of examining complex systems through a network lens [11, 12]. DC_i can be computed as:

$$DC_i = \sum_{i=1}^{n} A(i,j) \tag{2}$$

where $\mathbf{A}$ corresponds to the adjacency matrix of the network, defined as:

$$\mathbf{A}(i,j) = \begin{cases} 1 \text{ if there is a link between } i \text{ and } j \\ 0 \text{ otherwise} \end{cases}$$

Note that although activity networks are directed (reflected by an asymmetric $\mathbf{A}$ $\left(\text{i.e. } \mathbf{A}(i,j) \neq \mathbf{A}(j,i)\right)$, these are to be simplified as undirected, reducing the number of indicators needed. Such simplification would not have been appropriate if we were focusing on dynamical processes that are heavily dependent on the directionality of the task network e.g. failure propagation, where a task's capacity to trigger a failure cascade, or be influenced by one, is a function of its out-degree and in-degree respectively [59]. However, as we are focusing on the overall *structural* importance of each task, by considering the undirected case, we are able to highlight nodes that have both high in-degree *and* out-degree. This feature is an important perquisite of structural complexity (as it highlights non-trivial structural correlations due to assortative mixing [60]) emphasizing the existence of tasks that unlock an increased number of tasks *and* are increasingly dependent on the completion of other tasks – an aspect likely to increase the challenge of effectively delivering the project.

BC_i essentially reflects the number of times node i is found within the shortest path that connects any two other nodes.

[1] For example, if one was to focus on the RYF property (see Section I), *connectivity* would be a suitable indicator; its uniformity evaluated by examining the connectivity distribution. By doing so, the contribution of each individual node could be exposed i.e. if the connectivity distribution is increasingly uniform, the contribution of each node is comparable to every other node; if the connectivity distribution is not uniform, then the majority of nodes will be of little importance, with few nodes serving as central connection points (the so-called hubs).

[2] Such indirect implications arise from the fact that the Euclidean space of an activity network corresponds to project time, a feature unique to activity networks. For reference, another major class of networks that assign a meaningful aspect to the Euclidean space of a network are spatial networks [58], where distance between two connected nodes corresponds to the physical distance between them.

Nodes with high BC play a key role in holding the network together, and consequently, relate to several resilience-related concepts [15, 61, 62]. BC_i can be defined as:

$$BC_i = \sum_{jk} n_{jk}^i \tag{3}$$

where n_{jk}^i is 1 if node i is found within the shortest path that connects node j and k and 0 if it does not or if there is no such path [63].

$\mathrm{IET}_{i,j}$ captures the time interval between two consecutive tasks i and j, measured in days. This measure has a direct equivalent within the project management literature, coined as task free-float [42]. However, this more general term was chosen in order to reflect the generality of the concept. For example, it has been shown that the numerous complex systems follow a bursty behavior, captured by extreme variance in the IET distribution, compared to its normally-distributed counterpart [25, 64-66].

Finally, DTD corresponds to the number of active tasks per day, providing a proxy for the daily coordination effort required, noting that higher DTD will demand a greater amount of coordination effort.

It should be noted that neither the number nor the nature of the indictors proposed is exclusive. One can easily extend this procedure to include further aspects that are deemed to have a direct or indirect effect on the structural complexity of a project.

2) Artificial set of observations

Due to the focus of H1 (i.e. variability), the precise nature of the $\mathrm{FD_{pdf}}$ is of little importance, as long as two main conditions are satisfied: a) the presence of a fast decaying tail, and b) be defined within positive bounds. The first condition is central in the argument (i.e. caps the potential variability of individual contributions); the second condition is set to restrict the emergence of negative values for each indicator.

With respect to the first condition, the Normal (i.e. Gaussian) distribution is one of the most widely used pdfs [67]. Its exponential nature strictly limits the range of possible values that an observation can take. As a result, a meaningful average emerges, where the majority of observations fall within a narrow band of values (99.7% are found within 3 standard deviations). Consequently, the contribution of each component to the overall aspect is similar (and consequently, H1 is satisfied). Yet, the Normal distribution is defined with infinite bounds $[-\infty, +\infty]$ and thus, allows for both positive *and* negative values to emerge, conflicting with condition b). In response, a truncated variant of the Normal distribution can be used, where finite bounds [a, b] can be set [68]. Following the work of Mazet [69], a truncated Normal pdf is defined as:

$$p(x) = \frac{1}{Z}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{4}$$

where $Z = \sqrt{\frac{\pi}{2}}\sigma\left[\mathrm{erf}\left(\frac{b-\mu}{\sqrt{2\sigma}}\right) - \mathrm{erf}\left(\frac{\alpha-\mu}{\sqrt{2\sigma}}\right)\right]$, erf is the error function, defined as $\frac{2}{\sqrt{\pi}}\int_0^x e^{t^2}dt$, $\mu =$ mean and $\sigma^2 =$ variance.

To evaluate H1, eq. 4 can be used to generate a sample of observations that corresponds to a specific aspect of the topology of an activity network. This set of (artificial) observations will necessarily satisfy H1 (i.e. the range of values will be tightly bound) and thus, will serve as the reference point. First, eq.4 must be parameterized against the empirical set of observations, as stored in vector I. In order to do so, parameters a, b, μ and σ need to be accordingly set. The latter two parameters correspond to the mean and standard deviation of I; a and b correspond to the minimum and maximum values in I. Once the precise form of $p(x)$ has been obtained, 1,000 samples of n entries (in the form of 1,000 column vectors) are drawn. The probability of encountering a value of a given size X is computed for every column vector, and subsequently stored in matrix $\mathbf{M}$. Finally, matrix $\mathbf{M}$ is averaged across its second dimension in order to limit the emergence of outlier values. The resulting column vector $(R2)$ essentially corresponds to the artificial observations obtained by the null model, and can be used to assess the difference in variability of I.

3) Distance measure

To compare the variability of the empirical set of observations (I) against its artificial counterpart $(R2)$, and consequently evaluate *H1*, the histogram containing the frequency of each observation will be used. Specifically, let I_i^{freq} (and $R2_i^{\mathrm{freq}}$) contain the frequency-coded empirical (and artificial) observation in bin i. Note that each histogram is normalized, such that $\sum_i I_i^{\mathrm{freq}} = \sum_i R2_i^{\mathrm{freq}} = 1$).

As such, the Bhattacharyya measure (B) can be used [70], to assess the difference between the two histograms, defined as:

$$B = \sum_i \sqrt{I_i^{\mathrm{freq}}} \times \sqrt{R2_i^{\mathrm{freq}}} \tag{5}$$

with $B = 1$ indicating a perfect match between the two histograms. As such, $1 - B$ refers to the structural complexity of an activity network, as captured by the appropriate indicator. Note that the Bhattacharyya measure was chosen over more traditional similarity measures (such as the chi-squared statistic) as it has a number of benefits, including lifting the assumption that the content within each bin follows a Poisson distribution (which is especially useful as the underlying distribution of I is unknown) [71].

IV. Results

As the purpose of this work is to exemplify the applicability and utility of the procedure, detailed analysis will be limited to one direct (BC) and one indirect (IET) indicator – for completeness, results for the remaining indictors (DC and DTD) are included in the Appendix. For each indicator, the empirical set of observations is first computed (and normalized using eq.1). Similarly, the artificial set of observations is obtained by using the truncated Normal pdf

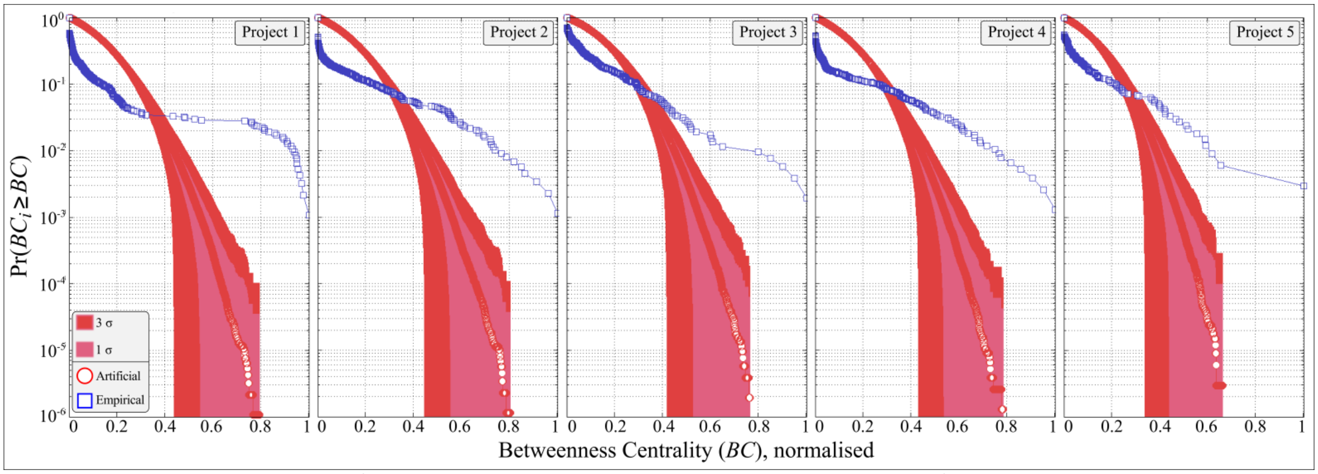

Fig. 3: Cumulative probability distribution of betweenness centrality (BC), capturing the probability (y-axis) of encountering an entry equal or greater than a given size (x-axis). Increased variance is exhibited by the empirical set of observations (blue, square marker) compared to the normally distributed, artificial set of observations (red, circle).

(eq.4). Subsequently, the cumulative probability distributions of BC and IET, for both empirical and artificial observations, are plotted – see Fig.3 and Fig. 5 respectively. By doing so, we can proceed to a visual inspection, providing for a qualitative assessment of the differences between the empirical and artificial set of observations for each respective indicator. For the quantitative assessment, the normalized histograms for both empirical and artificial set of observations are first obtained. The Bhattacharyya measure is subsequently computed in order to provide for a measure of the absolute similarity between the two histograms. By doing so, H1 can be evaluated for each individual indicator.

As such, the degree of structural complexity for each project is provided in the form of an average, along with the percentile contribution of each aspect – see Fig. 6. Where appropriate, parallels with the wider set of complex systems are drawn, either from original work (Fig. 4) or through relevant literature.

A. Direct Measurement – Betweenness Centrality

A number of general observations can be made across all five activity networks – see Fig. 3. Specifically, the cumulative probability distribution of all five artificial sets of observations decay exponentially, limiting the allowable size of an observation – this is to be expected due to the exponential tail of the (truncated) Normal pdf from which they are drawn. This is in stark contrast to the empirical set of observations, where the probability of obtaining relatively low values declines rather swiftly at the beginning, but subsequently slows down. As a result, it is significantly more probable to encounter large values (typically, this difference is several orders of magnitude) along with allowing for greater values to emerge.

Project 2, 3 and 4 follow similar behaviors, where the empirical set of observations can account for roughly 20% to 30% larger values. At the same time, the probability for encountering an observation in the empirical set, of equal size to the maximum value of the artificial is approximately 4 orders of magnitude higher.

Despite the fact that Project 1 follows the same general points, it shows an added peculiar feature. Specifically, a transitional stage is observed near the mid-point of the distribution. At this stage, the probability of obtaining a value of a given size remains roughly the same, yet its potential value more than triples (from $BC \cong 0.2$ to almost $BC \cong 0.8$). The mechanism responsible for the emergence of this effect may be worth further exploration; yet it is beyond the scope of this work and will not be explored further.

Moving to a more quantitative approach, the Bhattacharyya measure is computed using the histograms of all five sets of empirical observations, with respect to their artificial counterpart. As such, Project 1 is the most structurally complex (i.e. lowest $B(BC)$ value) with a Bhattacharyya value of 0.5463, followed by Project 5 (0.6220), Project 2 (0.7023), Project 4 (0.8398) and Project 3 (0.9703). Note that a *higher* Bhattacharyya value corresponds to increased similarity with the respective artificial set of observations, and hence to *reduced* structural complexity.

At this point, let us use the same indicator to examine the variance of a system that is widely accepted as being complex. Consider the Internet, as the operating level of autonomous systems (AS) – it is widely considered as one of the most complex systems currently in operation [63], with numerous researchers exploring its various properties through a networks view (e.g. [13, 14, 21]). Note that the Internet differs from the WWW, where the former corresponds to a physical network of autonomous systems, at an inter-domain level, and the latter corresponds to an information network composed of websites and hypertext links, and is of a significantly larger scale. By using publically-available data, the network representation of the system (composed of 22,963 nodes and 48,436 links) can be explored, and the BC indicator accordingly computed. Note that despite the difference in size between the Internet AS network and the activity networks used, the normalization process ensures that scaling conflicts are avoided. By visual inspection of Fig. 4, a qualitatively similar behavior is observed, yet to a greater extend. This increase can be explained by the increased scale of the system, where the difference in size implies an increased capacity for asymmetry to grow – this is reflected by the lower Bhattacharyya measure (0.4232; approximately 29.09% higher than the most structurally complex project). Importantly, this qualitatively similar behavior between the Internet AS network and the

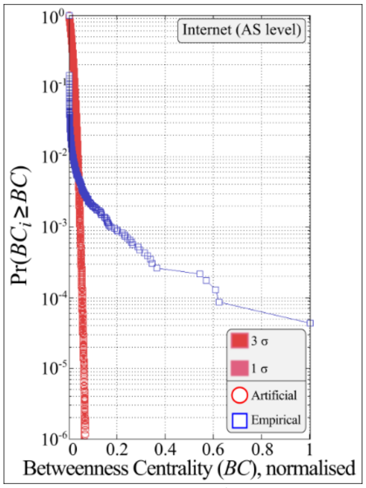

Fig. 4: Cumulative probability distribution of betweenness centrality across a widely-accepted complex system - the Internet. Notice the extreme difference between the observed and artificial sample and the qualitatively similar behavior to Fig. 3.

activity networks is important for two reasons: a) it illustrates that the method is consistent and able to show that the Internet, as the AS level, is complex and b) that projects, at least from a structural point of view, are also complex, yet to a lesser degree.

B. *Indirect measurement – Inter-Event Time*

IET is directly coupled to the structure of the activity network, and thus, provides an indication of the structural complexity of the project. Despite the fact that it builds on topological information (i.e. task precedence) it further requires the temporal signature of each task (i.e. start and end date). By doing so, it captures the structural complexity of project in an *indirect* way.

As with the case of BC, the entirety of projects follows a qualitatively similar behavior, with a number of general observations: the cumulative probability distribution of the empirical set of observations deteriorates faster than its artificial counterpart but as the reference value increases, the rate of decrease slows down (Fig. 5). As a result, the maximum value found within the empirical set of observations is significantly higher (ranging from approximately 20% to 60%). At the same time, the probability of encountering the maximum artificial observation in the empirical set is significantly higher, typically 3 orders of magnitude higher. With respect to the Bhattacharyya measure, Project 1 is the most structurally complex (i.e. lowest $B(IET)$ value), with a value of 0.6274, followed by Project 2 (0.6879), Project 5 (0.7207), Project 4 (0.8186) and Project 3 being the least structurally complex, with a value of 0.9681.

Although we are unable to compute the IET (nor the DTD) indicator for the Internet AS network (since nodes do not have any temporal signature), numerous systems have been noted to have heavily asymmetric IET distributions, similar to the ones captured in Fig. 5. Examples range from various forms of human communication (including the use of email, letters and phone) [25, 65, 72, 73], library activity [73] and Internet traffic [64] to earth-quake activity [74] and brain activity [72]. Such bursty behavior is considered to be another universal feature of several complex systems [72]. Consequently, the similarity between activity networks and complex systems is reinforced, further strengthening the view that activity networks (and to an extent, projects) are indeed complex, at least from a structural point of view.

C. *Overall Structural Complexity*

The structural complexity of each activity network, quantified using the Bhattacharyya measure for all four indicators, is given in Table III. In step with both qualitative and quantitative observations, all 5 projects are shown to significantly deviate from their artificial counterparts and hence, H1 can be confidently falsified. It is worth noting that since the scope of the work revolves around the falsification of a hypothesis (rather its validation), the finite size of the sample shouldn't affect the validity of this insight.

TABLE III
BHATTACHARYYA MEASURE ON THE FOUR INDICATORS OF STRUCTURAL COMPLEXITY. HIGHER VALUES SUGGEST LOWER COMPLEXITY

Project	**Measurement of Structural Complexity**			
	Direct		*Indirect*	
	$B(DC)$	$B(BC)$	$B(DTD)$	$B(IET)$
1	0.5750	*0.5463*	*0.5599*	*0.6274*
2	0.7001	*0.7023*	*0.8294*	*0.6879*
3	0.9189	*0.9703*	*0.9906*	*0.9681*
4	0.9348	*0.8398*	*0.8088*	*0.8186*
5	0.2420	*0.6220*	*0.8107*	*0.7207*

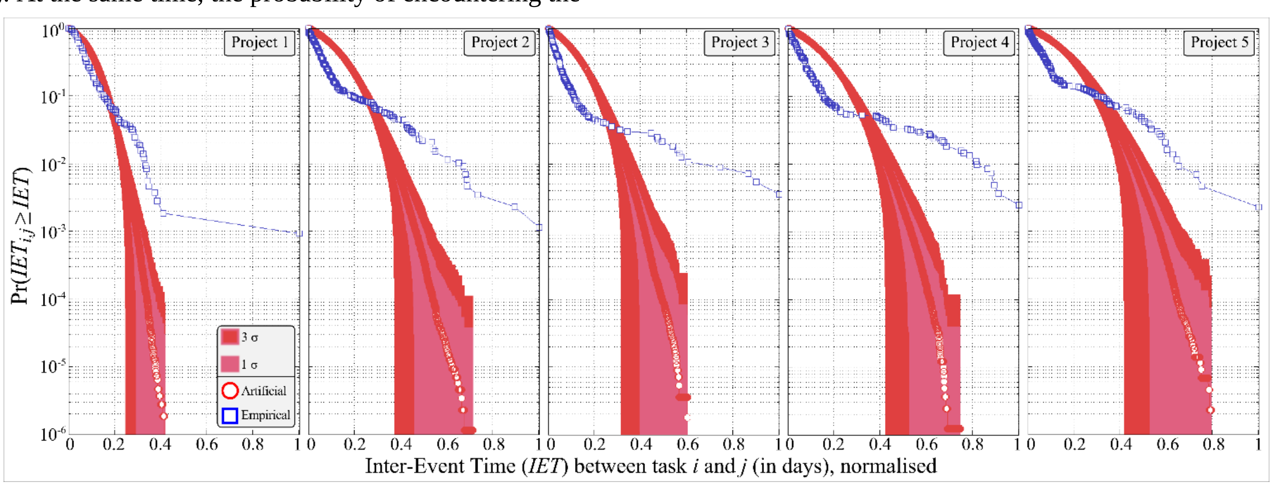

Fig. 5: Cumulative probability distribution of IET, capturing the probability (y-axis) of encountering an entry equal or greater than a given size (x-axis). As with Fig. 3, increased variance is noted in the empirical set of observations compared to the artificial set.

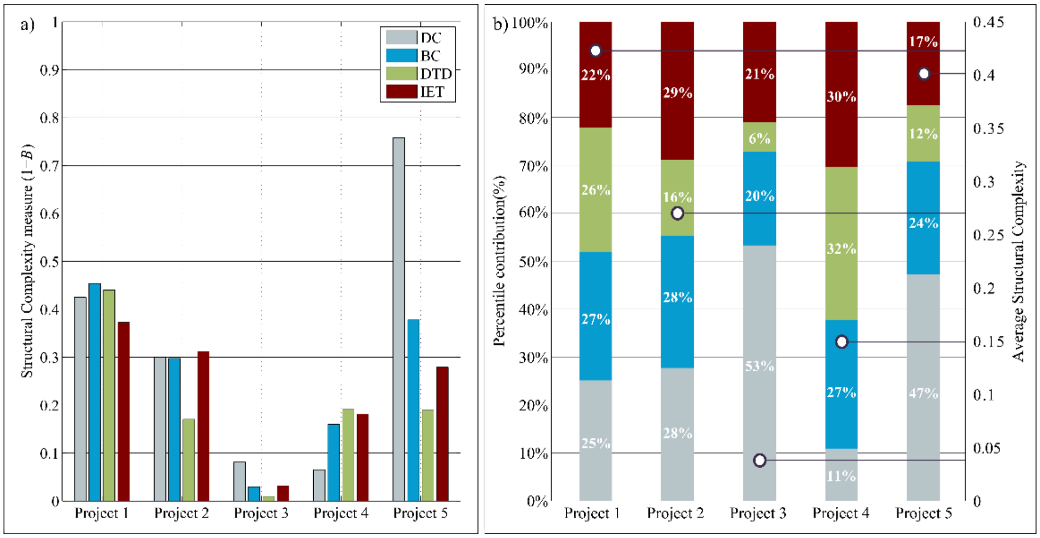

Fig. 6: (a) Structural complexity, quantified by the $1-B$ value (y-axis) for each project, where each color represents a specific indicator; (b) contribution of each indicator to the overall structural complexity of each project, in terms of a percentage (left y-axis; numerical value in each bar). The white, circle marker represents the average structural complexity of each project (right y-axis).

Following these results, the structural complexity (i.e. $1 - B$) of each activity network is visualized in Fig. 6a. Fig.6b emphasizes the distinct profiles of structural complexity within each project, based on the contribution of each distinct aspect. For example, consider Project 1 and 5, with an average structural complexity of 0.5772 and 0.5988 respectively (a relative difference of 3.74%). Despite their comparable average structural complexity, their composition is distinctly different – Project 5 has an asymmetric profile, with DC contributing almost half of its overall structural and DTD being significantly low. On the other hand, Project 1 shows a balanced profile with each aspect roughly contributing in equal parts – see Fig. 6b. Similarly, Project 3 appears to be rather similar to Project 5, despite their significant difference in terms of average structural complexity (a relative difference of 954.96%). Specifically, the larger contributor of structural complexity for both Project 3 and 5 is DC, with DTD having a significantly lower contribution.

V. Discussion

The aim of this paper is: (a) to develop a procedure which enables the measurement of an aspect of project complexity, and by doing so, (b) provide evidence supporting the equivalence between projects and other complex systems – both aspects have been recognized as key challenges within the field of project management [18]. Despite the broad range of aspects that define a project, and consequently project complexity, this work focuses on the structural complexity of a project across its technological dimension (i.e. the transformation process of converting inputs to outputs [29]).

Even within this narrow context, this task is inherently challenging. Part of the complexity in describing certain systems lies at the necessity for describing each aspect separately and how it relates with all the remaining aspects [40]. In the context of project complexity, this challenge translates directly to the composing aspects of complexity (Fig. 1), where structural complexity must be described both independently aspect but also with respect to the remaining aspects of complexity (e.g. structural complexity is implicitly coupled with uncertainty [41]). This work focuses on the first task (describing, and quantifying, structural complexity) – enabling future work to delve into the latter aspect of evaluating the relationship between structural complexity and other sources of project complexity in an evidence-based, quantitative way.

Even within the limited scope of this work we note that project complexity does *not* necessarily scale with project cost or size (in terms of number of activities). Consider Project 1 and 5, having very little difference in terms of average structural complexity (Fig. 6b), despite the fact that Project 1 has more than twice of Project's 5 cost, duration and scale (in terms of its compositing activities). Similar insight was put forward by Williams [33], where the case of the Kuwait oil field reconstruction and the Automated London Ambulance project are used as examples where a massive and costly project (Kuwait oil field) can have significantly lower complexity than what its size suggests, with the converse being also true (i.e. London ambulance project). The consequences of such underestimations can be serious – in the case of the Automated London Ambulance project, it translates to 2.5 million USD and 20 lives lost [30]. On the same grounds, one may consider Project 5 as a similar case, where its relatively small size can mask its structural complexity, potentially leading to ineffective/inefficient decisions being made.

A. *Measuring Structural Complexity*

Operationalizing the means for measuring structural complexity of projects must reflect the practical ethos of project management and consequently, dictate the data requirements of the proposed procedure. In response, this work endorses the use of readily-captured and widely-used data; specifically project schedules in the form of Gantt charts [42]. As such, it can be readily adopted by current practice.

The proposed procedure has been applied in an *a priori* fashion, allowing for the quantification of the structural complexity of five engineering projects. By doing so, numerous insights can be put forward. For example, projects with a relatively low $B(DC)$ value (e.g. Project 5) will be largely composed of tasks with a relatively low number of links. At the same time, a few tasks will have a surprisingly large number of links these will be central tasks, and thus should be constantly monitored as their failure is bound to affect a large portion of the remaining activities. As such, project manager faced with low $B(DC)$ projects should expect a heightened need for identifying these central tasks, and prioritize resource allocation accordingly. At the same time, shielding these tasks against possible perturbation should be part of any proactive risk management plan. This is of special importance in cases where the largest contribution of structural complexity is DC, as in the case of Project 5.

The case of $B(BC)$ is subtly different, where a low value suggest the emergence of nodes that are increasingly important in unlocking an increased number of activities. These nodes are tasks that have increased control over the remaining nodes and thus, a project manager dealing with low $B(BC)$ project should follow a similar approach and actively identify, and appropriately manage, them.

In the case of indirect measures, and specifically from a DTD point of view, a project with a low $B(DTD)$ would have highly asymmetric coordination effort requirements, where few task will be active per day, for the majority of the project's duration (low coordination requirement), but with few days having a surprisingly high number of active tasks (high coordination requirement). These few days are bound to stress the coordination capabilities set to manage the project, especially if the resources in place were tailored to deal with the remaining, low-requirement days. Furthermore, an increased capability to toy with the level of coordination effort dispensed must be in place, to cope with the highly asymmetric coordination requirements. With respect to the IET indicator, one can identify projects that exhibit an increasingly bursty behavior (i.e. the majority of tasks starts soon after its predecessor(s) are completed, yet some are scheduled surprisingly late). Such bursty behavior, and the implied long wait-time between consecutive tasks, can introduce a temporal buffer. As such, potential issues in the delivery of task i have an increased amount of time to be resolved, lower the probability of the consecutive task(s) to be affects. A similar effect is noted by Karsai, et al. [75] with regard to the general process of failure cascades, of which the previous project-specific example is a subset of.

This procedure can further be applied *in-situ*, where changes in structurally complexity can be monitored, as they arise from changes in a project's schedule. As such, an increased in the structural complexity of a project may serve as an early-warning mechanism, calling for immediate mitigation action to be taken.

Finally, this procedure can be used on a portfolio level, where it can be used to evaluate the capacity of an organization in successfully delivering structurally complex projects. Insight of this sort can be used to guide the bidding process for future projects, a process being widely considered as a major cause of engineering project failure [76, 77].

B. Projects/Complex Systems Equivalency

The Internet AS network will be used as a reference point in order to assess whether projects, even under the limited representation of their activity network, can be considered to be complex systems. There is little doubt that the Internet is considered to be a complex system, exhibiting numerous trademark characteristics – extreme variance in its topology being one of them [13, 14, 78]; Fig.4 is a representative example. As such, all five projects exhibit a qualitatively similar behavior, where the largest empirical observed value cannot be accounted by the artificial set of observations. Furthermore, significant differences are noted in terms of the probability of encountering a value of a given size, in times spanning several orders of magnitudes. Importantly, both of these qualitative observations are scale invariant, as they are present despite the normalization process that took place. Nonetheless, one would expect that the Internet AS network would somehow be "more complex" as its larger size allows for certain nodes to increase in importance (from a BC point of view), further skewing the BC distribution. Consequently, one would expect that quantitatively, the Internet would still rank higher in terms of complex – this is exactly what is observed when the Bhattacharyya measure of the two classes is compared (0.4232 for the Internet AS network; 0.5463 for the lowest ranking activity network).

Activity networks capture both topological and temporal aspects and thus, the implications of structural complexity can be measured directly or indirectly. In contrast, the Internet AS network is limited to topological information, restricting our ability to use it as a reference point for the indirect indicators. Nonetheless, the bursty behavior noted by the activity networks (a consequence of increased variance in its distribution) has been noted in several complex systems (ranging from the human brain to the earth's crust [25, 65, 66, 72-74]), and has been suggested as a universal features of a large class of complex systems [72]. As such, activity networks exhibit another qualitatively similar behavior, further reinforcing the equivalence between projects and complex systems.

VI. Limitations and Future Work

Establishing a link between project performance and project complexity (even in the limited view of structural complexity) is bound to be challenging, simply because project complexity extends across numerous aspects [18] and dimensions [29]. Despite the fact that some external validation for the results of the propose procedure can be achieved by examining the actual project performance data[3], taking this route can be ambiguous. This is because it implies that other aspects of project complexity (beyond the structural aspect examined herein) have been kept constant between all projects – a clearly faulty assumption. Nonetheless, the fact that empirical data does not contradict the results of Fig. 6 (or Table III) should add confidence to its use. As an alternative, one can isolate the impact of structural complexity by using computer experiments (numerical simulation) to examine its impact of various processes that can affect the delivery of a project.

[3] For example, both Project 1 and 2 have shown significant delays (5 months for Project 1; 1.5 months for Project 2) and cost overruns (5 million USD, or 33% of its overall budget, for Project 1; 0.7 million USD, or 4% of its overall budget, for Project 2), on par with the fact that both Projects have been shown to be structurally complex, with Project 1 being more complex than Project 2.

Cascading process is one such process [57], where failure to deliver task i can affect the delivery of a portion (or even all) of its successor tasks. Recent work has shown that structural complexity (in terms of the 4 indicators used) significantly affects the propagation of such failures [79] and hence, impacts the likelihood of successfully delivering a project. In other words, a non-structurally complex project is less sensitive to failure cascades across its activity network, compared to its structurally complex equivalent.

Building on this insight, the procedure presented herein forms a minimal attempt to quantify the structural complexity of an engineering project across its technological dimension. Nonetheless, the flexibility of the approach allows for further information to be introduced into the analysis, broadening its focus. For example, if task i is deemed to be of increased importance (e.g. lies on the critical part; its precise nature is uncertain etc.), then its increased contribution can be captured by using a weighted, rather than a binary, adjacency matrix, where *every* $\mathbf{A}(i,j)$ entry is suitably adjusted. Similarly, if the relationship between tasks i and j is deemed to be of increased importance (e.g. due to increased resource requirements; scarce resource availability; complicated supply chain etc.) the entry $\mathbf{A}(i,j)$ can, again, be suitable adjusted. Such aspects are increasingly important in practice, where milestone nodes and/or links exist and can be determinant for the life of a project i.e. funding being conditional to their accomplishment of the respective node and/or link.

VII. Conclusion

Complex project management is relatively young, with the term “project complexity” being first introduced in 1996. As the field strives for maturity, attempts to expose various aspects of project complexity have converged to five distinct aspects of project complexity – these are socio-political; pace; dynamics; uncertainty and structural complexity. As such, attempt to quantify these various aspects, such as the one contained within this work, are expected to grow. Part of the challenge lies in identifying suitable mediums for measuring these various aspects – with several being rather hard to conceptualize, let alone measured.

By focusing on the technological dimension of projects, activity networks capture the structure of a project, and hence, serve as a suitable medium for measuring structural complexity. The practical utility of this procedure revolves around the capability to use structural complexity as a proxy for project complexity. By doing so, several important decision making processes can be improved, including prioritizing resource allocation and project bidding. From a theoretical point of view, results herein highlight a qualitatively similar behavior between activity networks and a well-known complex system – the Internet. As such, it validates the analogy of projects being complex systems, and builds towards establishing an equivalency.

VIII. Acknowledgments

CE would like to thank Derek K. Hitchins and Hilary Sillitto for helpful discussions. Lois Builders Ltd is gratefully acknowledged for providing the project schedules; Mark Newman for making the Internet dataset available (see http://www-personal.umich.edu/~mejn/netdata/); Vincent Mazet and Jonathan Lansey for code implementations, available from (http://miv.u-strasbg.fr/mazet/rtnorm/) and http://www.jonathan.lansey.net/publications.html respectively. Finally, the authors would like to thank the three reviewers for their constructive comments.

IX. References

[1] A. Latour, "Trial by fire: A blaze in Albuquerque sets off major crisis for cell-phone giants," *Wall Street Journal,* vol. 1, p. 29, 2001.

[2] C. S. Tang, "Robust strategies for mitigating supply chain disruptions," *International Journal of Logistics: Research and Applications,* vol. 9, pp. 33-45, 2006.

[3] G. Andersson, P. Donalek, R. Farmer, N. Hatziargyriou, I. Kamwa, P. Kundur, *et al.*, "Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance," *Power Systems, IEEE Transactions on,* vol. 20, pp. 1922-1928, 2005.

[4] B. Walsh, "10 Years After the Great Blackout, the Grid Is Stronger — but Vulnerable to Extreme Weather," in *TIME*, 2013 ed, 2013.

[5] P. Gai and S. Kapadia, "Contagion in financial networks," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science,* p. rspa20090410, 2010.

[6] F. Schweitzer, G. Fagiolo, D. Sornette, F. Vega-Redondo, A. Vespignani, and D. R. White, "Economic networks: The new challenges," *Science,* vol. 325, p. 422, 2009.

[7] D. L. Olson, *Supply chain risk management: tools for analysis*: Business Expert Press, 2011.

[8] A. Punter, "Supply Chain Failures: A study of the nature, causes and complexity of supply chain disruptions," airmic, London2013.

[9] A. Vespignani, "Modelling dynamical processes in complex socio-technical systems," *Nature Physics,* vol. 8, pp. 32-39, 2012.

[10] A.-L. Barabási, "The architecture of complexity," *Control Systems, IEEE,* vol. 27, pp. 33-42, 2007.

[11] M. Mitchell, "Complex systems: Network thinking," *Artificial Intelligence,* vol. 170, pp. 1194-1212, 2006.

[12] S. H. Strogatz, "Exploring complex networks," *Nature,* vol. 410, pp. 268-276, 2001.

[13] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *science,* vol. 286, pp. 509-512, 1999.

[14] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature,* vol. 406, pp. 378-382, 2000.

[15] M. E. J. Newman, "The structure and function of complex networks," *SIAM review,* vol. 45, pp. 167-256, 2003.

[16] X. F. Wang and G. Chen, "Complex networks: small-world, scale-free and beyond," *Circuits and Systems Magazine, IEEE,* vol. 3, pp. 6-20, 2003.

[17] C. Guo, L. Wang, F. Zhou, L. Huang, and Z. Peng, "Analysis on the “robust yet fragile” nature of Internet: load, capacity and the cascading failure avalanche effect," in *Young Computer Scientists, 2008. ICYCS 2008. The 9th International Conference for*, 2008, pp. 2149-2153.

[18] J. Geraldi, H. Maylor, and T. Williams, "Now, let's make it really complex (complicated) A systematic review of the complexities of projects," *International Journal of Operations & Production Management,* vol. 31, pp. 966-990, 2011.

[19] W. Weaver, "Science and complexity," *American scientist,* vol. 36, pp. 536-544, 1948.

[20] M. E. J. Newman, "Complex Systems: A Survey," *Am. J. Phys.*, vol. 79, pp. 800-810, 2011.

[21] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics,* vol. 74, p. 47, 2002.

[22] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang, "Complex networks: Structure and dynamics," *Physics reports,* vol. 424, pp. 175-308, 2006.

[23] P. Bak, *How nature works*: Oxford university press Oxford, 1997.

[24] F. Reitsma, "A response to simplifying complexity," *Geoforum,* vol. 34, pp. 13-16, 2003.

[25] A.-L. Barabási, "The origin of bursts and heavy tails in human dynamics," *Nature,* vol. 435, pp. 207-211, 2005.

[26] S. E. Elmaghraby, "On the fallacy of averages in project risk management," *European Journal of Operational Research,* vol. 165, pp. 307-313, 2005.
[27] L.-A. Vidal and F. Marle, "Understanding project complexity: implications on project management," *Kybernetes,* vol. 37, pp. 1094-1110, 2008.
[28] C. Mele, J. Pels, and F. Polese, "A brief review of systems theories and their managerial applications," *Service Science,* vol. 2, pp. 126-135, 2010.
[29] D. Baccarini, "The concept of project complexity—a review," *International Journal of Project Management,* vol. 14, pp. 201-204, 1996.
[30] Y. Bar-Yam, "When systems engineering fails-toward complex systems engineering," in *Systems, Man and Cybernetics, 2003. IEEE International Conference on*, 2003, pp. 2021-2028.
[31] T. Cooke-Davies, S. Cicmil, L. Crawford, and K. Richardson, "We're Not in Kansas Anymore, Toto: Mapping the Strange Landscape of Complexity Theory, and Its Relationship to Project Mangement," *IEEE Engineering Management Review,* pp. 5-21, 2008.
[32] T. M. Williams, "The need for new paradigms for complex projects," *International Journal of Project Management,* vol. 17, pp. 269-273, 1999.
[33] T. M. Williams, *Modelling complex projects*: John Wiley & Sons, 2002.
[34] K. Remington and J. Pollack, *Tools for complex projects*: Gower Publishing, Ltd., 2007.
[35] P. Anderson, "Perspective: Complexity theory and organization science," *Organization science,* vol. 10, pp. 216-232, 1999.
[36] ICCPM, "Global Perspectives and the Strategic Agenda to 2025," International Centre for Complex Project Management, Australia, 2011.
[37] PMI, "Navigating Complexity," Project Management Institute, 2013.
[38] K. Remington, R. Zolin, and R. Turner, "A model of project complexity: distinguishing dimensions of complexity from severity" in *Proceedings of the 9th International Research Network of Project Management Conference*, 2009.
[39] J. Oehmen, C. Thuesen, P. P. Ruiz, and J. Geraldi, *Complexity Management for Projects, Programmes, and Portfolios: An Engineering Systems Perspective (White Paper)*. USA: Project Management Institute, 2015.
[40] Y. Bar-Yam, *Dynamics of complex systems* vol. 213: Addison-Wesley Reading, MA, 1997.
[41] A. De Meyer, C. H. Loch, and M. T. Pich, "Managing project uncertainty: from variation to chaos," *MIT Sloan Management Review,* vol. 43, p. 60, 2002.
[42] PMI, "A Guide to the Project Management Body of Knowledge: PMBOK® Guide," 5th ed: Project Management Institute, 2008.
[43] L. V. Tavares, "A review of the contribution of operational research to project management," *European Journal of Operational Research,* vol. 136, pp. 1-18, 2002.
[44] S. Hartmann and D. Briskorn, "A survey of variants and extensions of the resource-constrained project scheduling problem," *European Journal of Operational Research,* vol. 207, pp. 1-14, 2010.
[45] S. E. Elmaghraby, "Activity nets: A guided tour through some recent developments," *European Journal of Operational Research,* vol. 82, pp. 383-408, 1995.
[46] M. Vanhoucke, "An overview of recent research results and future research avenues using simulation studies in project management," *ISRN Computational Mathematics,* vol. 2013, 2013.
[47] E. Demeulemeester, M. Vanhoucke, and W. Herroelen, "RanGen: A random network generator for activity-on-the-node networks," *Journal of Scheduling,* vol. 6, pp. 17-38, 2003.
[48] S. E. Elmaghraby and W. S. Herroelen, "On the measurement of complexity in activity networks," *European Journal of Operational Research,* vol. 5, pp. 223-234, 1980.
[49] W. R. Ashby, "Requisite variety and its implications for the control of complex systems," in *Facets of Systems Science*, ed: Springer, 1991, pp. 405-417.
[50] R. C. Leachman and S. Kim, "A revised critical path method for networks including both overlap relationships and variable-duration activities," *European Journal of Operational Research,* vol. 64, pp. 229-248, 1993.
[51] V. Valls and P. Lino, "Criticality analysis in activity-on-node networks with minimal time lags," *Annals of operations research,* vol. 102, pp. 17-37, 2001.
[52] A. J. Shenhar, "One size does not fit all projects: Exploring classical contingency domains," *Management Science,* vol. 47, pp. 394-414, 2001.
[53] A. Clauset, C. R. Shalizi, and M. E. Newman, "Power-law distributions in empirical data," *SIAM review,* vol. 51, pp. 661-703, 2009.
[54] M. Mitzenmacher, "A brief history of generative models for power law and lognormal distributions," *Internet mathematics,* vol. 1, pp. 226-251, 2004.
[55] M. E. J. Newman, "Power laws, Pareto distributions and Zipf's law," *Contemporary physics,* vol. 46, pp. 323-351, 2005.
[56] N. N. Taleb, *The Black Swan:: The Impact of the Highly Improbable Fragility*: Random House Digital, Inc., 2010.
[57] D. Helbing, "Globally networked risks and how to respond," *Nature,* vol. 497, pp. 51-59, 2013.
[58] M. Barthélemy, "Spatial networks," *Physics Reports,* vol. 499, pp. 1-101, 2011.
[59] C. Ellinas, N. Allan, C. Durugbo, and A. Johansson, "How Robust Is Your Project? From Local Failures to Global Catastrophes: A Complex Networks Approach to Project Systemic Risk," *PLoS ONE,* vol. 10, p. e0142469, 2015.
[60] M. E. Newman, "Assortative mixing in networks," *Physical review letters,* vol. 89, p. 208701, 2002.
[61] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han, "Attack vulnerability of complex networks," *Physical Review E,* vol. 65, p. 056109, 2002.
[62] M. E. J. Newman, "Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality," *Physical Review E,* vol. 64, p. 016132, 2001.
[63] M. E. J. Newman, *Networks: an introduction*: Oxford University Press, 2009.
[64] V. Paxson and S. Floyd, "Wide area traffic: the failure of Poisson modeling," *IEEE/ACM Transactions on Networking (ToN),* vol. 3, pp. 226-244, 1995.
[65] J. G. Oliveira and A.-L. Barabási, "Human dynamics: Darwin and Einstein correspondence patterns," *Nature,* vol. 437, pp. 1251-1251, 2005.
[66] J. L. Iribarren and E. Moro, "Impact of human activity patterns on the dynamics of information diffusion," *Physical Review Letters,* vol. 103, p. 038702, 2009.
[67] G. Grimmett and D. Stirzaker, *Probability and random processes* vol. 2: Oxford Univ Press, 1992.
[68] N. Chopin, "Fast simulation of truncated Gaussian distributions," *Statistics and Computing,* vol. 21, pp. 275-288, 2011.
[69] V. Mazet, "Simulation d'une distribution gaussienne tronquée sur un intervalle fini," Technical report, Universite de Strasbourg/CNRS2012.
[70] A. Bhattachayya, "On a measure of divergence between two statistical population defined by their population distributions," *Bulletin Calcutta Mathematical Society,* vol. 35, pp. 99-109, 1943.
[71] F. J. Aherne, N. A. Thacker, and P. I. Rockett, "The Bhattacharyya metric as an absolute similarity measure for frequency coded data," *Kybernetika,* vol. 34, pp. [363]-368, 1998.
[72] M. Karsai, K. Kaski, A.-L. Barabási, and J. Kertész, "Universal features of correlated bursty behaviour," *Scientific reports,* vol. 2, 2012.
[73] K.-I. Goh and A.-L. Barabási, "Burstiness and memory in complex systems," *EPL (Europhysics Letters),* vol. 81, p. 48002, 2008.
[74] S. Lennartz, V. Livina, A. Bunde, and S. Havlin, "Long-term memory in earthquakes and the distribution of interoccurrence times," *EPL (Europhysics Letters),* vol. 81, p. 69001, 2008.
[75] M. Karsai, M. Kivelä, R. K. Pan, K. Kaski, J. Kertész, A.-L. Barabási, *et al.*, "Small but slow world: How network topology and burstiness slow down spreading," *Physical Review E,* vol. 83, p. 025102, 2011.
[76] J. D. Frame, *The new project management: tools for an age of rapid change, complexity, and other business realities*: John Wiley & Sons, 2002.
[77] A. Bagies and C. Fortune, "Bid/no bid decision modelling for construction projects," in *Proceedings of the 22nd Annual ARCOM Conference, 4-6 September*, 2006, pp. 511-21.

[78] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology," in *ACM SIGCOMM Computer Communication Review*, 1999, pp. 251-262.
[79] D. J. Watts, "A simple model of global cascades on random networks," *Proceedings of the National Academy of Sciences,* vol. 99, pp. 5766-5771, 2002.

X. APPENDIX

Results of the analysis, with respect to DC and DTD, are presented in Fig. 1SI and 2SI respectively

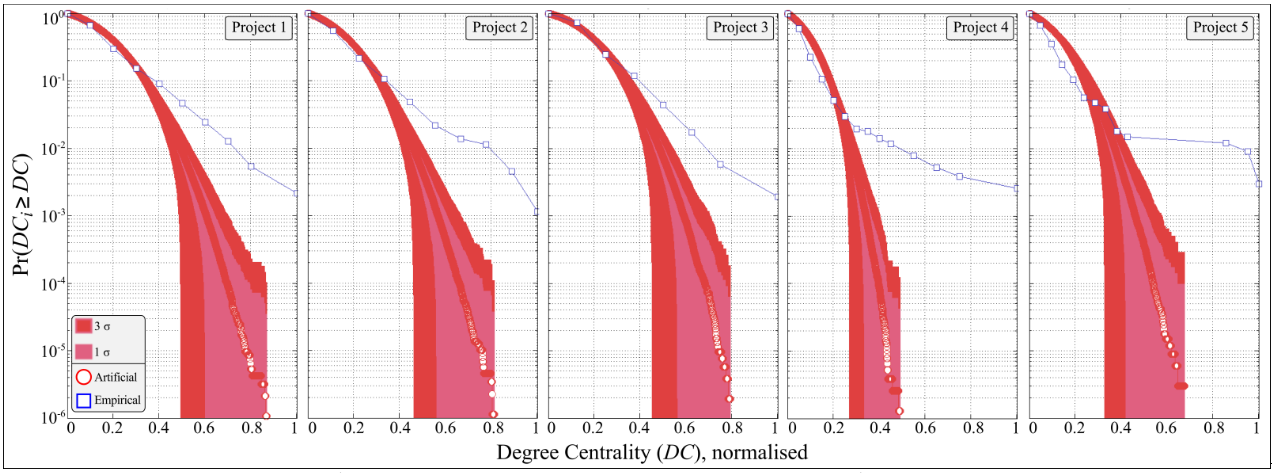

Fig. 1A: Cumulative probability distribution of degree centrality (DC) capturing the probability (y-axis) of encountering an entry equal or greater than a given size (x-axis). In this case, empirical observations (blue, square) exhibit limited variance and are reasonably mapped by the normally-distributed, null sample (red, circle). As values increase in size, variance dominates and difference of several orders of magnitude is observed.

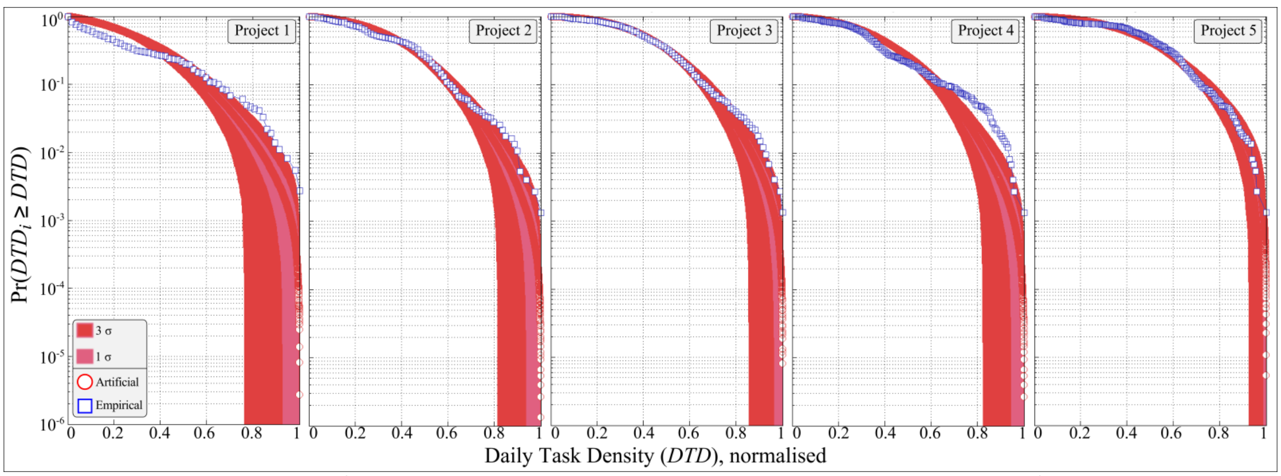

Fig. 2A: Cumulative probability distribution of the daily task density (DTD) distribution, capturing the probability (y-axis) of encountering an entry equal or greater than a given size (x-axis). In this case, empirical observations (blue, square) exhibit limited amounts of variance, with Project 2,3 and 5 being well mapped by the normally-distributed, null sample (red, circle); Project 1 and 3 are also reasonably mapped.