Contents

1 Executive Summary 1

2 Goals and Impact 2

3 Technical Approach 3

4 Management Plan 9

5 Capabilities 9

6 Statement of Work 9

References 9

1 Executive Summary

The proposed work will be carried out over a period of 24 months in the base period, followed by 12 months in the option period, at a total cost of XX USD.

2 Goals and Impact

Detailed simulation of social phenomena is currently used by the DoD to undertsand the evolution of social structures, and the emergence of relevant organizational hierarchies in theatres of conflict. In the modern reality of asymmetric warfare often against local insurgencies in conflict countries, informing military strategy with expected social implications is crucial for optimal long-term outcomes, and US foreign policy success. In this work, we aim to define, develop and demonstrate proof-of-concept principles for validating complex social simulations. Additionally, our goal is to extend social theory to 1) compare and constrast complex systems, 2) disambiguate real and simulated phenomena, 3) chart effective principles to narrow or erase such identifiable distinctions to ultimately realize more realistic models and predictions, and 4) develop computable strategies to disambiguate closed vs open systems, $i.e.$, identify the existence and the role of unobserved and unknown external influences.

Social phenomena unfold as complex interdependent system of systems operating at multiple scales of orgamization in space and time. Centuries of work in social theory has teased out a few of the important guiding principles around which such large scale systems organize; nevertheless first principle quantitative rules akin to the laws of physics, are generally missing. Thus, unlike physics, social scientists do not have a "standard model" — a neat set of equations believed to be not just a *good* model of the physical univese, but rather a representation of the exact ground truth. Under this convenient scenario, physicists can work out simulations that confidently reflect reality, and build gargantuan particle accelerators to test when-and-if rare events subtly deviate from simulated outcomes to search for *new physics*, and validate existing theory. The reality for social scientists is harsher — with no such universal set of equations (*or the hope of ever finding one*), the veracity of complex social simulations is forever suspect. How do we know we have built in the right amount of complexity? How do we know if the simulated systemes have the same emergent structures, and any conclusion or observation in such systems have even a tenuous connection to reality? How do we *quantify* the deviations from and uncertainties between real systems of interest and engineered simulations built to interrogate them, often at great expense?

These questions are of crucial importance to national security. Without quantified confidence on the large scale social simulations that are becoming increasingly important in military strategy and foreign policy, incorrect recommendations have the potential for catastrophic long-term consequences.

In this work we propose to address these issues by crafting rigorous computable measures that characterize diverse aspects of the emergent dynamics in social interactions. The challenge here is to craft measures that are application agnostic, and thus capable of evaluating objectively diverse real-life and simulated scenarios. Technically speaking, we are designing characterizations for complex spatio-temporal systems, with unknown or poorly understood rules, operating at multiple spatial and temporal scales, with variables that can be a mix of categorical, ordinal as well as discrete and continuous, and potentially subject to noisy and adversarially corrupted observations.

Our approach is predicated on our ability to effectively distill good predictive models of such systems from data, in a manner that is agnostic to the explicit details of the application. To that effect we leverage our recent work on a novel spatio-temporal stochastic modeling framework — the Granger nets — which are demonstrably superior in predictive ability and sample complexity to existing off-the-shelf deep learning architectures.

Broadly our scheme is as follows: Given observation logs from a system (simulated or real), we construct the Granger net, and then interrogate the inferred predictive structures through the lens of a range of carefully constructed measures (See Table 1 for overview) that illuminate various dynamical characteristics pertaining to complexity, stability, reslilience, and evidence of self-organized criticality. In our preliminary studies, these measures clearly disambiguate real and simulated systems — even ones constructed with considerable effort aiming to erase any such distinctions.

Thus, the primary over-arching goal of this work is to investigate the missing elements that leave such tell-tale signatures in high fidelity simulations, and ultimately move towards future design principles

TABLE 1: Dynamical Measures Proposed For Precise Chracterization of Complex Systems

	Measure	Property Measured	Description
1.	μ_c	Complexity	The level of complexity of multi-scale multi-variate spatially extended systems require new measures of compelxity that capture statistical compelxity, structure of connectedness, and cross-talk.
2.	μ_s	Stability	Stability in systems of interest might be finely poised between instability regimes, with the possibile manifestation of self-orgnized criticality.
3.	μ_ω	Frequency-like	Generalization of frequecny domain tools and measures to non-linear multi-scale spatio-temporal system of systems with categorical and ordinal variables.
4.	μ_e	External Influence	Estimate the pssibility having an open vs a closed system.

that better simulate reality.

3 Technical Approach

We lay out our apprach which proceeds by conceiving, refining and validating computable measures that interrioghate key dynamical chracteritcs of compelx systems from empirical observations. The primary objective here is to explore iof high fidelity social simulations are recognizable as simulations from theur real world counetrparts. And then, develop tools thatnarrow this gap. Additionally, we plan to investigate measures that estimate the probabiloity of external influences or unmodelled effects in social systems, both simulated and real.

Systems of Interest

Our real-world systems of interest are urban crime and global terrorism. Simualted worlds from current DARPA programs, Robo Soccer, Real-life Soccer provide aditional examples.

Measures of Stability: The May Constraint & Self-organized Criticality In Real-World Systems

For our purpose, a stable system is one that can carry on operating with low probability of generating event patterns or behaviors that threaten catastrophic failures, and cessation of operation; a stable system is in no danger of dying suddenly. Stability criteria are well understood for systems with known structures. However, for real-world multi-component multi-scale complex sytems revealed only through finite set of empirical observations, it is less clear how to establish or analyze stability properties. *We hypothesize that explicit measures of expected stability may be designed that potentially disambiguate real world complex systems of social interactions from simulated ones.* To see how such measures are designed, we briefly describe the seminal work of Robert May[1], [2] on stability of complex systems.

May's work on large random systems show that complex systems tend to be inherently unstable, implying that for ecosystems with more species, more interspecific interactions per species (connectance), or stronger interactions are not as likely to be as stable as systems with fewer of these attributes. However, these results are at odds with the empirical observations that large, highly complex ecosystems are MORE stable NOT LESS.[3]

In other words, May's work identifies computable constraints on system parameters, violations of which that are overwhelmingly likely to cause systemic instability in large random systems; nevertheless natural complex systems seem to be able to operate in those regimes with no sign of an unstable demise. More specifically, May's theorem deals with S variables (species), where the ith species interacts with the jth one with a strength drawn from a normal distribution with probability C and are 0 otherwise. For large S, he shows that the system is stable with probability 1 if:

$$\gamma\sqrt{SC} < 1 \tag{1}$$

and with probability 0 otherwise, where γ is the average strength of interaction. The elegance of May's theorem, coupled with the verified contradiction in large scale natural systems suggests that complex ecosystems are perhaps NOT randomly wired. Real world systems exist in small islands of stability t

Zipf's Law (Frequency $\propto$ rank^{-1}) for Inferred Strength of Predictive Influence (γ)

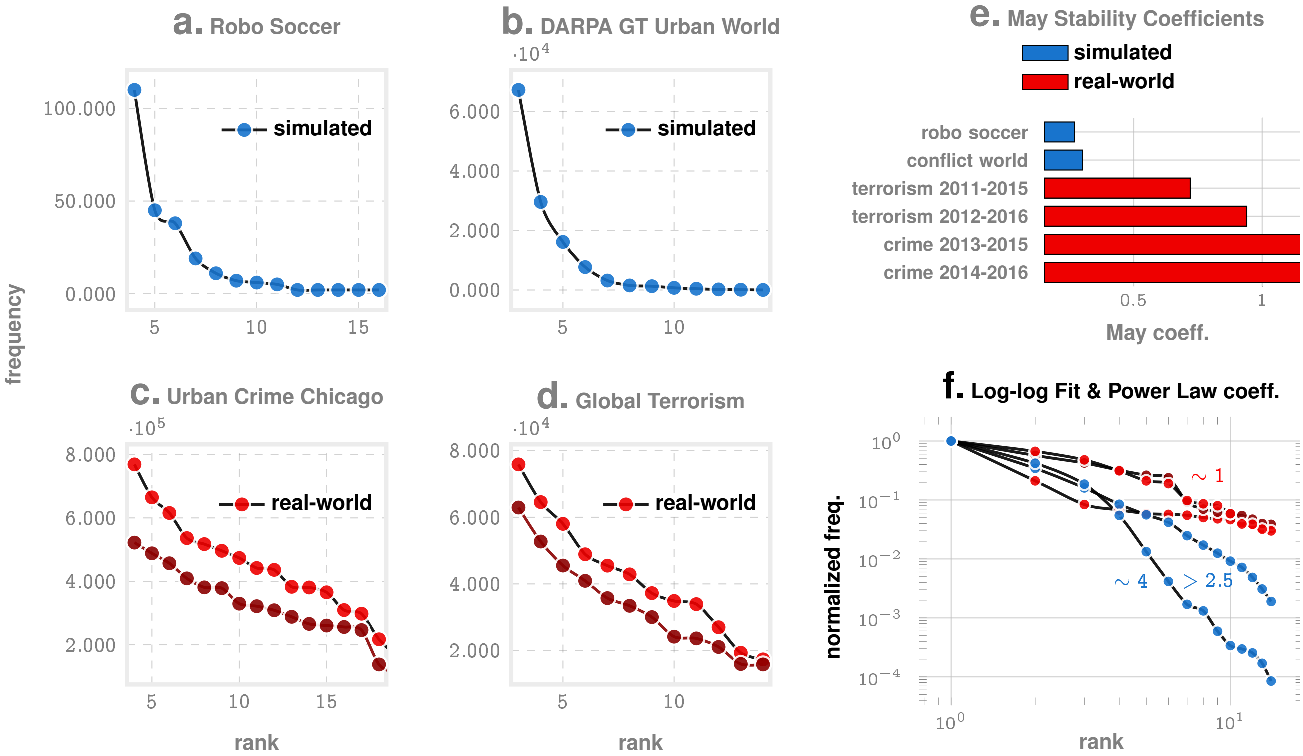

Fig. 1: Zipf's Law holds in real-world systems, but not in simulated ones, even the ones where significant effort and resources are expended to replicate reality. Thus, it appears complex system of systems in the real world often exhibit critical phenomena, whereas simulated systems often do not. This discrepancy is also reflected in the stability criterion of Robert May: it seems that real worls systems should be unstable but they are not: they exist in islands of stability surrounded by unstable regimes in high dimensional parameter spaces.

in complex high dimensional parameter spaces, where large perturbations would risk incursion into unstable regimes by May's theorem.

Thus, we define our first measure of stability as:

$$\mu_s = \gamma \sqrt{SC} \tag{2}$$

We hypothesize that simulated systems will typically have a significantly lower value of μ_s, while real world systems will often violate the apparent stability criterion ($\mu_s < 1$), while continuing to be actually stable in practice.

This notion of natural systems being poised at special miniscule "critical" regions in high-dimensional parameter spaces is strongly indicative of *self-organized criticality*.[4] Our preliminary studies — and the discussion below — suggests that real-world complex systems tend to be *poised at criticality*, operating at or near systemic phase transitions, while simulated systems often lack this behavior. We exploit these insights for our second class of stability-related characterization measures. First, we briefly dilineate this deep connection, before charting our application strategy.

For decades, physicists have hoped that the emergent, collective phenomena of living systems, both at the moelcular and the social scales, could be captured using ideas from statistical mechanics. States of such emergent systems are almost never at equilibrium, maintained by a flow of energy and material (in biology) or societal exchange of relevant entities (people, ideas, beliefs etc). At the same time, it seems impossible that such balance is attained via accidental choice of system parameters. We believe that recognizing the emergence of self-organized criticality across scales is key to distinguishing real-world systems from existing simulation efforts.

Classicially, critical plhenomena has to do with phase transitions, and discontinuities in relevant order parameters near such transitions. Criticality, however, is a much more general concept than its instantiation by phase transitions in equilibrium systems. Assume that we are able to model the statistically stationary states of a complex system with a many-parameter model. Clearly random choice of these parameters is not likely to redproduce meaningful function; and if the system we are studying has many components, then any reasonable probabilistic model will break the parameter space into regions corresponding to different phases. Thus, for complex systems it is likely that the parameter space supports a phase diagram, in which regimes of qualitatively distinct behavior are separated by critical surfaces. *We plan to investigate the intriguing possibility that most complex systems in the real-world are not deep in one phase or another, but rather poised near a critical surface in the natural parameter space.*

Directly verifying this hypothesis is problematic. However, we can indeed validate if the systems carry signatures of critical organization, $e.g.$, the emergence of Zipf's Law in the distribution of strength of inferred predictive links between observables, or the in/out degree distribution of the inferred influence network. The justification of looking for such characterization in the distrbution of links strengths or degrees arises from the idea that an enumeration of these properties may be seen as a specification of the state the complex system — and existence of the Zipf's law amounts to the existence of a critical temperature below which the system "freezes", and above which thermodynamic quantities diverge.[4]

As shown in Fig 1, our preliminary studies corroborate this hypothesis, at least with the systems under consideration.

Quantifying Complexity of Complex Systems: Are Real-world Systems More Complex?

The proactice of using various functions of system size as a measure of complexity is potentially misleading. In contrast, we plan to develop measures that capture the complexity of information processing within the system, $e.g.$, the intricacies of how the system is internally wired, and the number of intrinsic states that the system has for it to manifest the observed behavior. Furthermore, since for real-world systems the ground truth of internal structures may only be estimated, our complexity measures must always be calculated from structures inferred or estimated from observed data and event streams.

We plan to consider two broad compelxity measures: 1) *I-complexity* ($\mu_{\mathbf{c}}^{I}$): the number of non-trivial models we can infer given the data and an upper bound on temporal memeory, normalized by the maximum number of possible models, and 2) the *Z-Complexity*($\mu_{\mathbf{c}}^{Z}$): which is the average number of states in the set of inferred models. I-complexity is the connectance that appears in May's condition of stability of complex systems discussed before: what is the probability that two varaibles or entities interact. And Z-complexity captures the average complexity of the intervariable dependence.

Note that these measures are computable effectively only if we have an inference algorithm that models non-trivial dependence between observed data streams. Any such inference algorithm must return generative models capturing the complexity of signal tranduction across observables only if there is actual dependence, and indicate direction-specific independence otherwise. Thus, in the trivial case where we are considering N independent coin tosses (implying no temporal memory, since outcomes are iid), we have N trivial single state models. Thus, our complexity measures in this case:

$$\mu_{\mathbf{c}}^{Z} = N/(N^2) = 1/N \tag{3}$$

$$\mu_{\mathbf{c}}^{I} = 1 \tag{4}$$

We plan to investigate if real-world systems are inherently more complex compared to simulated counterparts. In contrast to the stability measures, the complexity measures are easier to replicate, since we can always define in more dependencies in our simulation rules. Nevertheless, we find that for the systems we considered for our preliminary studies, real-world systems are indeed more complex when such complexity is measured through the lens of the above metrics (See Table **??**).

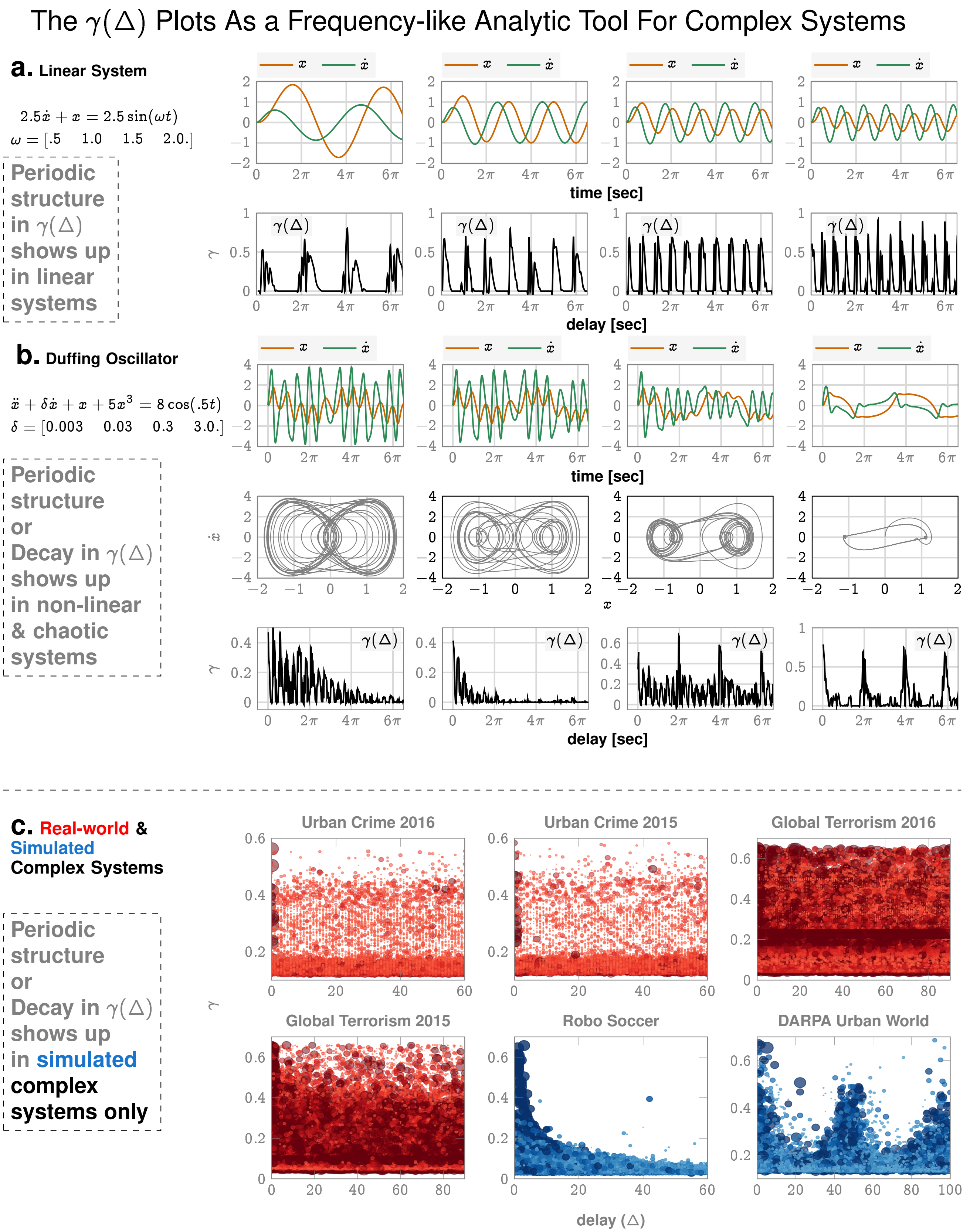

Fig. 2: Learning Processes Cross-talk in Complex Phenomena. Learning Processes Cross-talk in Complex Phenomena.Learning Processes Cross-talk in Complex Phenomena.Learning Processes Cross-talk in Complex Phenomena.Learning Processes Cross-talk in Complex Phenomena.

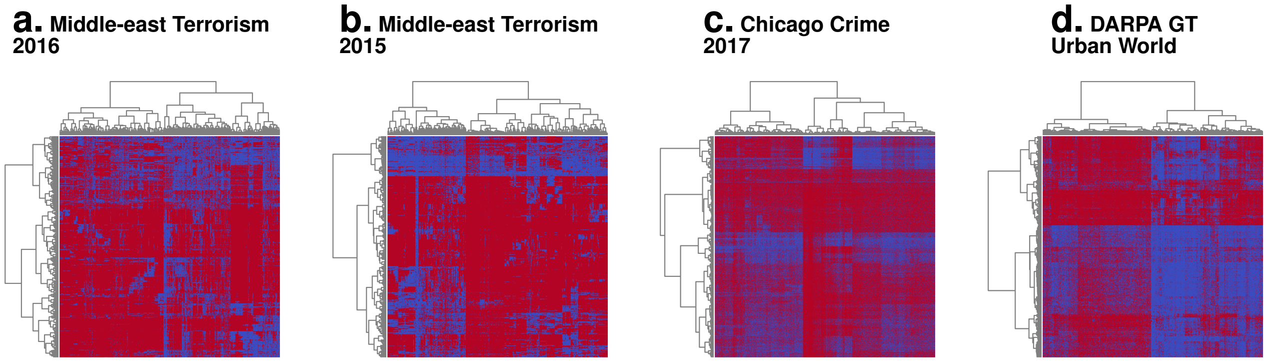

Fig. 3: Estimating Effect of External/Unmodeled Drivers: Biclustering on the difference of influnece matrix and its transpose. More red equates to more unmodeled effects. Terror event modeling restricted to the middle-east has the most unmodeled or external influences, while simulated DARPA urban world has the least, consistent with intuition.

Frequency-domain (like) Measures & Analytics: Exploring Temporally Non-localized Patterns

Frequency-domain analysis is a classical tool for linear time-invariant systems, that has enabled development of sophisticated engineering systems over the last century. Unfortunately, the concept of frequency does not easily generalize to non-linear, mixed variable, multi-scale complexities of social interactions. Leveraging our ability to quantify directional influence propagation between observables as a function of possible time delays, we aim to investigate one possible design of an analogous tool. We call this the $\gamma(\Delta)$ plot, where γ is an inferred measure of predictive influence, and Δ is the time delay over which the influence transpires.

For simple systems, the $\gamma(\Delta)$ plot may be interpreted as a auto-correlation measure, depicting the decay of influence as we move away from the source. For complex systems, the variation of γ with time delays may be used to illuminate key globally emergent properties such as scale invaraince, and process memory.

In our preliminary work, we evaluated the $\gamma(\Delta)$ analytic for a diversity of dynamical systems, ranging from simple linear, to non-linear to chaotic to simuated and real-world complex systems (See Fig. 2). We concluded that in simple systems, the $\gamma(\Delta)$ plot reveals clear macroscopic structure corresponding to periodic behaviors and/or rapid decay of influence with increasing delay. Similar behavior is observed in simulated complex systems. In contrast, in the real-world systems, all obvious macroscopic structure vanishes — suggesting the lack of any chracteristic scales in the dynamics. In line with our previousd observations, such scale-free behavior is associated with large to unbounded correlation lengths, and is consistent with the existence of critical phenomena and self-organization.

Thus, the proposed measure μ_ω is defined to be a measure of dependence of the $\gamma(\Delta)$ on Δ; if macro-structures are missing this measure will be small, implying that we expect this to be larger in simple systems, and smaller with real-world complexity.

These observations might be atypical, as not all complex systems will necessarily manifest similar scale-free behavior; Some systems may experience periodicity enforced by the nature of the dynamics being analyzed. Nevertheless, one of our goals here is to investigate the applicability of the $\gamma(\Delta)$ plot in analyzing non-temporally localized dynamical chracteristics — much in the same manner that the frequency domain tools are used for linear systems.

Exploiting Assymetry of Causality To Measure External Influence & Unmodeled Effects

Our second goal in this effort is to develop tools that uncover the possible existence of external drivers in complex social systems, $i.e.$, analyze if the system is open or closed.

This is particularly challenging as no social system is truely isolated: cities rise and fall, technology advances, beliefs and opinions evolve — sometimes such changes are brought about by internal processes, and even then it is unlikely that the effect of external infusion of knowledge, ideas and information can be ignored. Thus, what we can hope to meaningfully identify here is the probability of large external or unmodeled influences from empirical observations.

A particularly intriguing idea here is that we can exploit the intrinsic assymtery in the intuitive notion of causal influence to design a computable measure of the possibility of existence of external drivers. Note that if a variable x is causally influencing y, then an inference algorithm that recovers directional cross-dependencies between varaibles will register a larger influence from x to y, and a smaller influence in the reverse direction, $i.e.$, in other words the matrix of inferred influences will significantly deviate from symmetry. Similalrly if an unmodeled effect is driving both x and y, then it might appear that there are strong influences operating in both directions, $i.e.$, the influence matrix will tend towards being more symmetric.

More generally, we hypothesize that a closed system has *relatively long causal chains*, which conduct influence reltively unidirectionally, whereas in an open system we have an *over-abundance of short loops*. This idea is not categorically true for every situation, since one can easily fine-tune counter-examples in simple systems — but the question is if natural complex systems operate with such fine-tuned structures, or if the above scheme does indeed hold true broadly.

We plan to begin investigating this idea with two measures of isolation:

$$\mu^1_{\mathbf{e}} = \mathbf{E}_{i,j}\left(|\gamma^j_i - \gamma^i_j|\right) \tag{5}$$

$$\mu^2_{\mathbf{e}} = \text{ average path length in the graph of inferred directional influences} \tag{6}$$

where γ^i_j is the influence from node or variable i on variable j, and $\mathbf{E}(\cdot)$ is the expectation over all observed pairs in the system.

In our preliminary investigations, we apply this notion to terror event dynamics in the middle-east, to urban crime in Chicago, and to the simulated urban world from the DARPA Ground Truth program. Consistent with intuition we find that terrorism in middle-east seems to have the most external or unmodeled influences, while the simulated world has the least.

Importantly, this idea is not generative, in the sense that estimating the presence of unmodeled effects does not necessarily tell us how to correct our models or simulation to account for the possibly missing information. We will investigate the possibility of such corrections by evolving and perturbing the inferred models such that the above described measures iteratively improve.

Closing The Loop: Designing Better Simulations

Research Tasks

■ **Task 1. Definition of Complexity, Stability, Frequency-like Measures.** The key ideas described in the technical approach will be implemented and validated in a diversity of simulated and real-world systems. We will rank order simulated systems by their inferred realism, as indicated by our measures, and investigate commonalities that drive real vs simulated distinctions.

■ **Task 2. Development of Measures of External Influence.** Investigate the validity of the proposed notion of identifyiing the presence of external influence by expliting the assyemtry fo causal dependencies. In particular, we will carry out both theoretical investigations aiming to relate this idea to more mainstream notions unmodeled effects, and validate our approach extensively in real and simualetd systems.

■ **Task 3. Refining Simualtions & Engineering Better Systems.** We will leverage the work in tasks 1 and 2 to develop approaches that refine existting simulation environments, and eventually allow better designs that better mimic reality.

4 Management Plan

5 Capabilities

6 Statement of Work

Schedule & Milestones

References

[1] May, R. M. Will a large complex system be stable? *Nature* **238**, 413–414 (1972).
[2] May, R. M. *Stability and complexity in model ecosystems*, vol. 6 (Princeton university press, 2001).
[3] Jacquet, C. *et al.* No complexity–stability relationship in empirical ecosystems. *Nature communications* **7**, 12573 (2016).
[4] Mora, T. & Bialek, W. Are biological systems poised at criticality? *Journal of Statistical Physics* **144**, 268–302 (2011).