Naval Research Laboratory
Washington, DC 20375-5320

NRL/5580/MR—2022/4

Information Geometry

DR. IRA S. MOSKOWITZ

Information and Decision Sciences Branch
Information Technology Division

September 29, 2022

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
29-09-2022	NRL Memorandum Report	10/01/2019 – 09/30/2022

4. TITLE AND SUBTITLE

Information Geometry

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Dr. Ira S. Moskowitz

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

1P82

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Research Laboratory
4555 Overlook Avenue, SW
Washington, DC 20375-5320

8. PERFORMING ORGANIZATION REPORT NUMBER

NRL/5580/MR--2022/4

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Office of Naval Research
One Liberty Center
875 N. Randolph St, Suite 1425
Arlington, VA 22203-1995

10. SPONSOR / MONITOR'S ACRONYM(S)

ONR

11. SPONSOR / MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A: Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This is a close out memorandum report on the Information Geometry work unit. The goal of this work unit was to combine new fundamental research, which used geometric and topological analysis to develop new techniques that describe data and image complexity to make machine learning, and associated systems, "smarter" with respect to their predictive nature, and more resilient to erroneous classification. Furthermore, these techniques were evaluated for use in machine decision making in general, concentrating on multi-agents systems/Teams.

15. SUBJECT TERMS

Manifold, Wavelet, Wiener process, Brownian motion, Hurst coefficient, Machine learning, Topological data analysis, Differential geometry, Shannon information theory

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Dr. Ira S. Moskowitz
U	U	U	U	14	**19b. TELEPHONE NUMBER** *(include area code)* (202) 404-7930

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

This page intentionally left blank.

CONTENTS

This page intentionally left blank.

EXECUTIVE SUMMARY

The goal of this work unit (fiscal years 2020-2022) was to use geometric and topological information to make machine and hybrid human-machine Team decisions less prone to error and more efficient. Our Information Geometry research involved several different techniques from the field. We analyzed the value of information in various problems aimed at assisting decision making, while concentrating on the geometric aspects via Fisher information, statistical manifolds and Riemannian geometry. Wavelet transforms and geometric aspects of them were used as a decision tool. Research was performed to optimize a decision over a geometric decision space for Contingent Attention Management. We applied homological techniques from topological data analysis (TDA) to improve the classification accuracy of neural nets under an adversarial attack on certain data sets. We investigated, using time series analysis and the Hurst coefficient, fractional Brownian motion models. Furthermore, we applied these techniques to interdependence problems in agent/Team science (which models agents that are completing a task). From this research, we obtained a new metric for agent Team interdependence. We continued our Brownian motion analysis to study geometric diffusion properties, and related this to decision making processes. Furthermore, we used Riemannian geometric techniques to model multi-agent/Team system decisions. We concluded our research by using metric techniques, together with Shannon theory, to model agent decision systems.

This page intentionally left blank.

INFORMATION GEOMETRY

1 PROJECT OBJECTIVE

Combine new fundamental research, which uses geometric and topological analysis to develop new techniques that describe data and image complexity to make machine learning, and associated systems, "smarter" with respect to their predictive nature, and more resilient to erroneous classification. Furthermore, these techniques will also be evaluated for use in machine decision making in general, concentrating on multi-agents systems/Teams.

2 PROJECT BACKGROUND

As with any scientific or engineering field that is moving from its birth into a mature and successful field, the necessity of re-examining and adjusting the field in terms of advanced mathematical concepts becomes necessary. That is the situation we were faced with when beginning this project. How could we improve the decision making of machines? How could we lessen the probability of an adversarial attack? We chose to use techniques from differential geometry and algebraic topology to mature the area. In particular, we concentrated on Fisher information theory and Riemannian geometry and topological data analysis. However, we included geometric aspects of wavelet transforms, the value of information, and geometric aspects of Brownian motion. The mathematical techniques we used have shown their value in other applied areas of research, but their use in machine decision analysis was nascent. We successfully applied those techniques.

Riemannian geometry is a well-known mathematical area that lets us understand concepts in non-Euclidean spaces. In particular, Riemannian geometry changes the standard metric by which distance is measured. Instead of equal weighting to all directions, one can adjust, on an infinitesimal level, the distance in particular directions. Riemannian geometry showed itself to be “the mathematics” necessary for Einstein’s general theory of relativity. It successfully measured the curvature caused by gravity.

Fisher information develops a very particular type of Riemannian metric related to how information is transferred in various physical situations. Fisher information has its roots in statistics via the Cramer–Rao bound. Amari took this approach and successfully applied it to what he named “statistical manifolds.”

Topological data analysis, and, in particular, persistent homology theory show how the knowledge of gaps in a topological space can give one meaningful information about that space. It has proven its worth in analyzing data sets.

Manuscript approved September 29, 2022.

The geometric paths of Brownian motion, which are the infinitesimal version of random walks, have significant probabilistic information embedded in them. Many physical systems obey a diffusion equation similar to the base case of a Wiener process.

Shannon's information theoretic concept of channel capacity can tell one how multi-agents systems are transferring information based upon the underlying geometric constraints. Shannon theory has been the basis for modern telecommunications engineering.

3 TECHNICAL APPROACH & RESULTS

We first present our approaches and results. We then summarize, at the end of this section, by publication.

Our previous base program research in steganography and the existing literature showed that what we call "regions of high complexity" (these are regions of high power Fourier frequencies, residual high bit plane artifacts, etc.) are good places to both hide and look for information. By "complexity," we mean rapid changes in standard curvature features, and/or sudden topological changes in either shape of color/luminance, and large, relative (to other regions) volume with respect to the statistical manifold structure. We hypothesized that the complexity of an image makes the image more vulnerable to being manipulated to fool a machine learner (neural net) or biometric device. To test our hypotheses, we developed new metrics for scene complexity, based upon our analytic mathematical techniques (e.g., rapid changes of surface normals, high speed variation of curvatures, sudden topological changes, large relative Fisher-Rao volume), and used these metrics to assess if high complexity areas are ripe for steganographic hiding, and whether the incorrect output of a deep machine learner is correlated with changes in regions of "complex" image features. Our 6.1 research, aside from assisting the Navy, advanced the state of the art by allowing us to develop and apply new techniques and research results in machine learning and in statistical manifold theory. We used as a testbed the MNIST database with several well-known adversarial attacks on the classification power of the standard neural net model for this data. We then applied homological techniques from topological data analysis (TDA) to see if we could increase the classification accuracy when the neural net was under an adversarial attack (Figs. 1,2). We showed a modest improvement against the attacks, and are attempting to increase the neural net defense by using more than the first homology group approach that we took. Our approach tied two different fields together, and has promise for further exploitation.

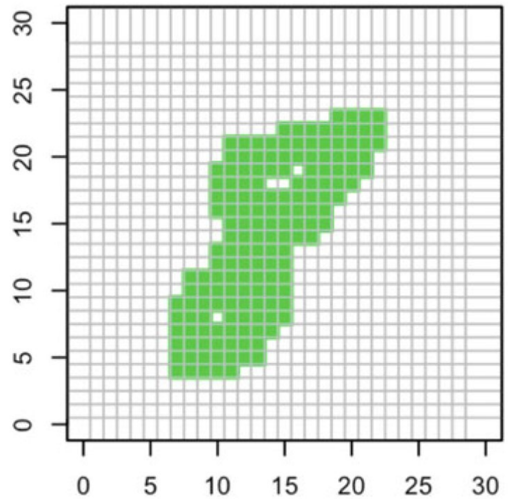

Fig. 1. MNIST DIGIT WITH THREE HOLES (RANK OF FIRST HOMOLOGY GROUP=3)

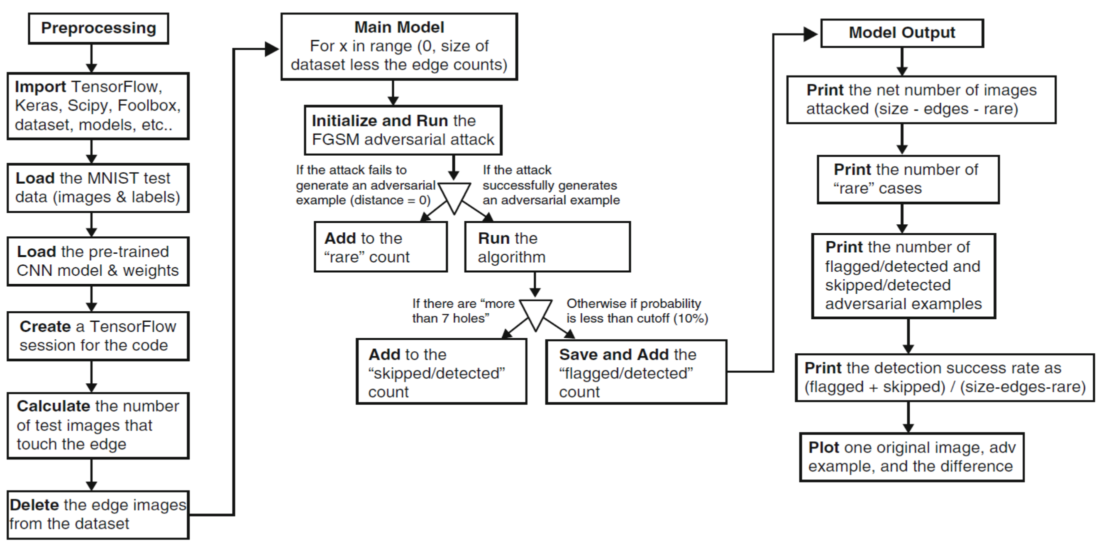

Fig. 9 Flowchart

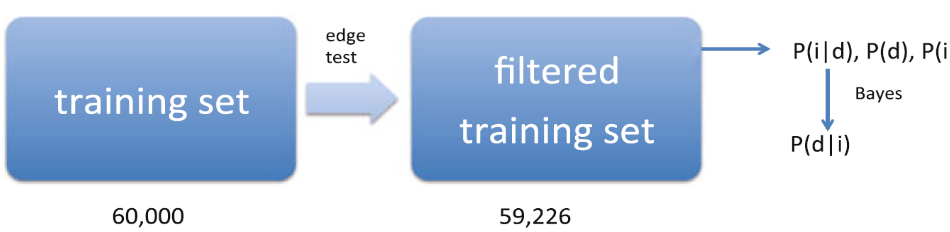

Fig. 2. ALGORITHM DETERMINING PROBABILITY BASED ON 7 HOLES

In addition to using standard Gaussian surface geometry and algebraic topological techniques, we used a natural and intuitive metric on probability distributions, the Fisher-Rao metric, to form a non-standard Riemannian manifold called the statistical manifold. Statistical manifolds have recently been used in many applications, running the gamut from belief propagation, manifold learning, and neural nets, to Grover's search algorithm in quantum computation. Using a statistical manifold is the preferred way to view parameterized distributions as metric spaces, so that the concepts of “close” and “far” are well defined, intuitive, and tractable.

Connections between differential geometry and algebraic topology are numerous and were utilized in our research. For example, there is the famous Gauss-Bonnet theorem, Morse theory, Ricci curvature, etc. These and similar techniques must be investigated to fully understand what characterizes an image, and how it can be manipulated. However, there has been very little work on combining geometry and topology with respect to analyzing machine learning. A major question we plan to explore is how much variance (wiggle room) is available to hide information, while keeping the curvature and topological signatures essentially unchanged. We started with the classical approach of using the standard Gaussian surface normal and principal curvatures to describe a surface, and then added the second geometric structure of the statistical manifold structure induced by the Fisher-Rao metric.

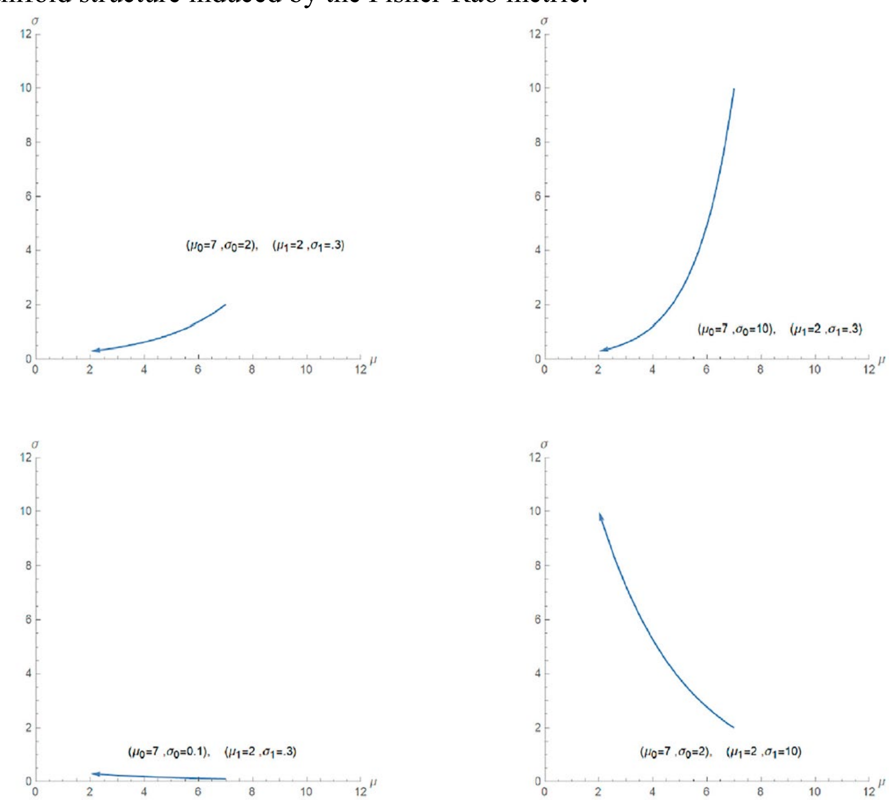

Fig. 3. VERY NON-EUCLIDEAN GEODESIC “STRAIGHT LINES” IN A MANIFOLD

A recent approach for image recognition of faces used probability distributions to represent images. Image identification is achieved by traveling along geodesics given by the Fisher-Rao metric. Our hypothesis is that regions of high volume (similar to regions of high frequency in discrete Fourier image transforms), with respect to the Fisher-Rao metric, are ripe for steganography and of course affect the distinguishability of information parameters. The Fisher-Rao metric allows one to obtain the volume element for the statistical manifold. This result is difficult to obtain in closed form, so we will develop numerical methods, when needed, to determine stego-rich and stego-poor regions of the image in terms of the Fisher-Rao volume element, and in terms of standard Gaussian and topological techniques. Additionally, a detailed analysis of the geodesics (Fig. 3) for the Fisher-Rao Riemannian metric is a necessity for calculating distance. Closed form solutions of these geodesics are difficult to obtain. However, in the situations that we analyzed, we solved for the geodesics and obtained a distance measure.

Howard's value of information research, which was originally applied to making economic decisions, was reexamined in terms of differential geometry, which led to new results incorporating the statistical manifold. Our approach was analyzed in terms of multi-agent decision making and power allotment. This theme of resource allocation was behind much of the research for this project.

We showed that geometric wavelets and discrete wavelet transformations can abstract data preprocessing, provide contextual conditioning, and otherwise affect often obfuscated underlying effects in models. Wavelets' properties of data reduction and feature retention, combined with discretizing the trend from local fluctuations, provided a mathematically sound technique for enhancing data-driven decision-making support. Our research discussed these benefits and presented a geometric method to answer the critical question of where an appropriate decomposition level for training data may exist. These benefits and the preferred geometric level selection technique were evaluated with a decision analysis application and simulation. Our research showed that the preferred geometric level selection technique can identify, in the case of voluminous data with many local geometric fluctuations, the Discrete Wavelet Transformation (DWT) level most appropriate for selecting decision-related values. Our analysis also showed that DWTs can make a significant difference in quantitative data-driven decision outcomes. The results of this research demonstrate that wavelets can improve quantitative algorithmic decision outcomes. In addition, our study illustrated how geometric DWT techniques reduce the volume of data while retaining its original features and characteristics, making the approach appropriate for time-series data and for many decision domains. Other decision-making scenarios may require the identification of exceptions and fluctuation-based features in the data, as opposed to selection for trending elements and sensitivities to other wavelets besides the Daubechies series. Additionally, our research laid the groundwork for using research results on multi-resolution multi-scale topological optimization to obtain optimal objectives.

Geometric results on Multiple Criteria Decision Analysis, derived from our prior base 6.2 Applied Network Science project, were modified and used to perform the geometric analysis needed in the associated base 6.1 project Contingent Attention Management. Analysis was required for how to deal with the problem of human performance degradation in highly multitasked environments. Research was performed to optimize a decision over a geometric decision space. The motivation for this research was to introduce a capability that monitors multiple, incoming streams of information for a user with expert decision-making responsibilities to assess the value of that information and to direct the user's attention to higher priority content. We performed the Multiple Criteria Geometric Decision Analysis needed to solve this problem by adapting Analytic Hierarchy Process techniques. The decisions to be made form a geometric decision space. Results from linear algebra were used to analyze the boundary conditions of this decision space. We then developed a weighting function to provide the functionality required. The theory was based upon exponential weighting. Analysis of the rankings for various scenarios was given and our method showed good and realistic performance.

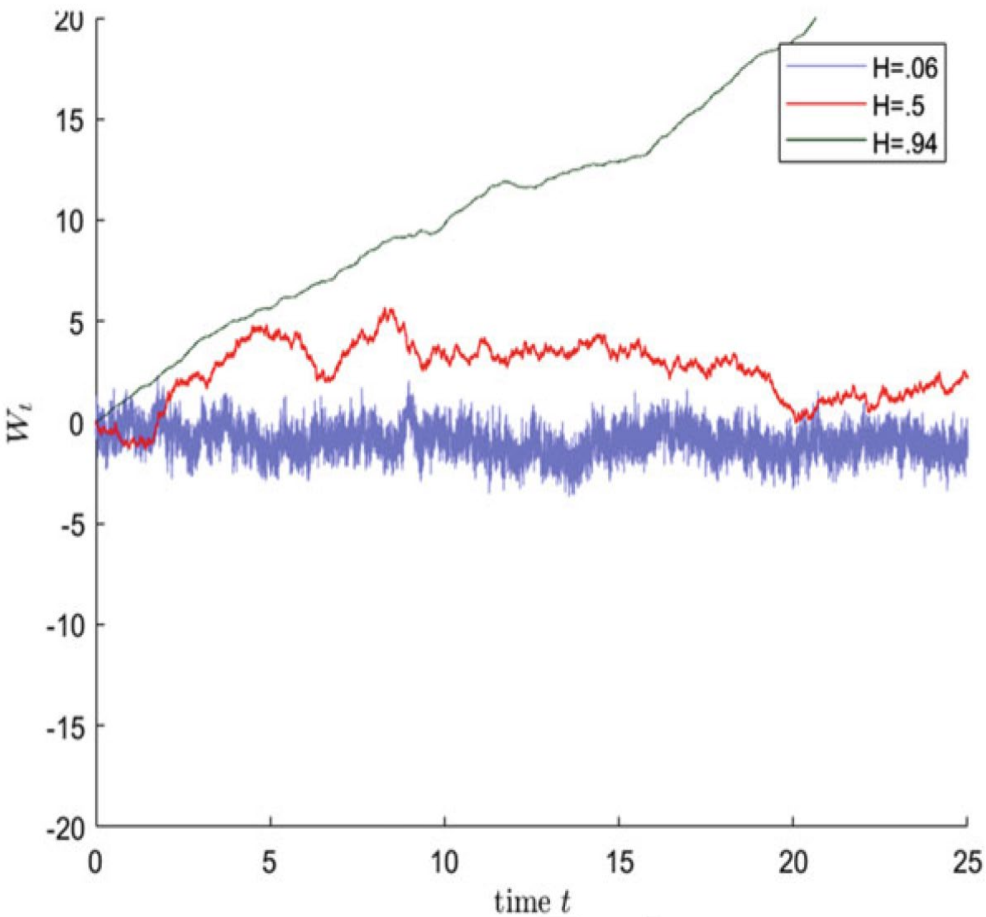

Fig. 4. DIFFERENT HURST COEFFICIENTS FOR THE WIENER PROCESS
(H=.94 largest slope, .H=.06 closest to 0, H=.5 in the middle (uncorrelated standard Wiener)

This project also extended the existing artificial intelligence modeling of human decision making by generalizing the Wiener process approach to the more general fractional Brownian motion. This called into question the absorbing time models that exist in the literature. We also investigated, using time series analysis, to determine the Hurst coefficient for fractional Brownian motion. Furthermore, we applied these techniques to interdependence problems in agent/Team science (which model agents that are completing a task). From this research, we obtained a new metric for agent Team interdependence. Additionally, various stopping time properties that hold for martingales do not hold for the Hurst-type processes that we are interested in (Fig. 4).

Advances in machine intelligence have led to an increase in human-agent teaming. In this context, one or more machines act as semi-autonomous or autonomous agents interacting with other machine teammates and/or their human proxies. This phenomenon has led to cooperative work models where the role of agent can be, interchangeably, a human, or machine, support system. Human counterparts that interact with automation become less like operators, supervisors, or monitors, and more like equal-authority peers.

Critical to the success of any Team is efficient and effective communication. Multi-agent systems are no different. Information sharing is a key element in building collective cognition, and it enables agents to cooperate and ultimately achieve shared goals successfully. Information sharing, or communication, provides the foundation for a Team's success. In complex multi-agent engagements, information is not always universally available to all agents. Such engagements are often characterized by distributed entities with limited communication channels among them, where no agent has a complete view of the solution space, and information relevant to Team goals only becomes available to Team members in spontaneous, unpredictable and even unanticipated ways. Moreover, there is always a resource cost to inter-agent communication. Finding highly efficient and effective communication patterns is a recurring problem in any multi-agent system, particularly if the system agents are distributed.

We researched how a Multi-agent System or Team, sends information between agents or teammates. By "how" we mean "how" in an information theoretic sense---in particular, we did not concentrate on the mechanics or physics of the transmission other than how it impacts information theory. We were concerned with what strategy an agent can use to maximize its information flow to another agent (Fig. 5).

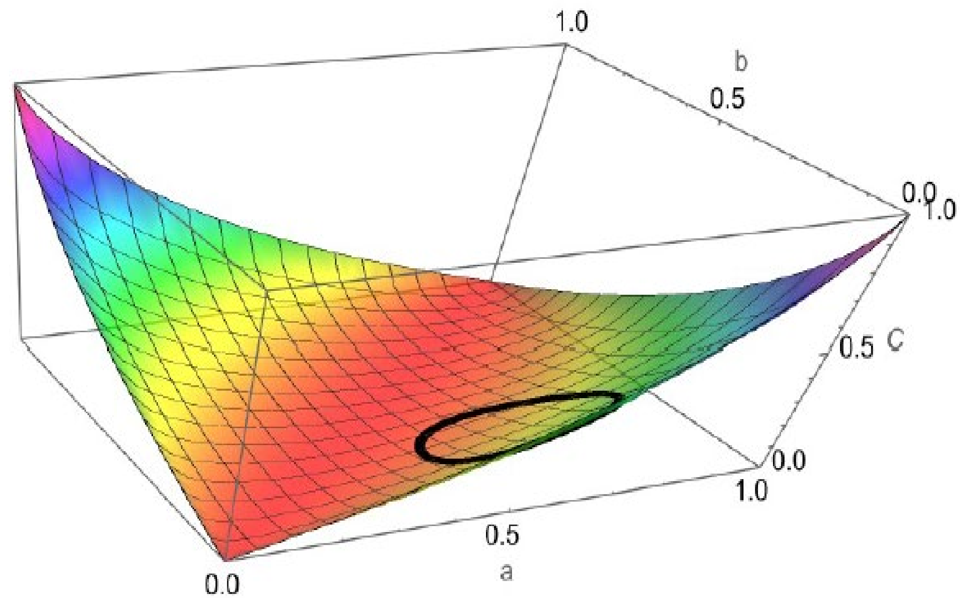

Fig. 5. LOCUS OF POINTS WITHIN A CERTAIN POWER CONSTRAINT IN A MULTI-AGENT SYSTEM

FY 2020 RESULTS

Publication [1] concerned itself with Howard's value of information and how we extended it via differential geometry to statistical manifolds and derived new decision making tools. The ability to value information may be easy for humans, but when it comes to a machine attempting to use the tools of artificial intelligence (AI), a valuation may not be easy. This is because even though we have a very good quantitative theory of information, we do not have a very good qualitative theory. In this paper we explored how hybrid machine-human Teams need to communicate in a way that can come to be accepted and trusted.

Publication [2] gave our wavelet-based results for decision making algorithms based upon the geometric patterns we obtained. We posited that to achieve general AI, techniques that abstract data preprocessing, provide contextual conditioning, and otherwise affect often obfuscated underlying models are necessary.

Publication [3] was involved in the geometric decision space via Multiple Criteria Decision Analysis. We used the Analytic Hierarchical Process in particular to develop our results.

FY 2021 RESULTS

Publication [4] is the result of our persistent homological topological data analysis of the MNIST database. We showed modest improvement over the state of the art image classifier by using homological side-information. This research holds promise for future improvements but needs to be analyzed in terms of method complexity, which is a topic of a FY2023 new start (Framework for Domain Focused Interpretable Machine Learning).

Publication [5] discusses various strengths and weaknesses of state-of-the-art AI systems. The paper discusses interdependence and entropy which we developed into information geometrical models in our other papers. Our research primarily addressed the application of interdependence theory to autonomous human-machine Teams.

Publication [6] discusses how to make machine decision making "smarter." We concentrate on various mathematical techniques but, in particular, focus on emergent behavior and its possibilities for geometric analysis.

In publication [7] we took a problem in human cognition and extrapolated it to a Teams science problem, which motivated some of our work for the next fiscal year. Geometric properties of fractional Brownian motion were considered and we developed a convolutional neural net to determine the Hurst coefficient.

FY 2022 RESULTS

Publication [8] was a continuation of our geometric aspects of Brownian motion research. In this paper we concentrated on Teams and multi-agent systems. We discussed the limiting behavior of probabilities of interest and also discussed how an information geometric approach can assist in analyzing various situations of interest.

In publication [9] we used a Riemannian metric as a cost metric when it comes to the optimal decisions that should be made in a multi-agent/Team scenario. The two parameters of interest to us are Team skill and Team interdependence, which are modeled as Wiener process drift and the inverse of Wiener process diffusion, respectively.

Publication [10] was an application of information theory to multi-agent systems with a rudimentary metric. We laid the groundwork for future research for more advanced Riemannian metrics. New results in information theory were obtained in the course of this research.

4 ASSOCIATIONS

BASE PROGRAMS

NRL base new start work unit T008-23 "Framework for Domain Focused Interpretable Machine Learning."
ONR funded 6.2 work unit 1M30 "Characterizing Decision Making in Complex Environments."
NRL base 6.1 work unit 1G29 "Contingent Attention Management."

EXTERNAL ORGANIZATIONS

We have collaborated with Dr. Paul Cotae at the University of the District of Columbia (Electrical and Computer Engineering), and Dr. Stephen Russell when he was at the Army Research Laboratory as the Division Chief for Information Science.

5 PUBLICATIONS

1--- **An Information Geometric Look at the Valuing of Information,** Ira S. Moskowitz, Stephen Russell and William F. Lawless, Chapter 9, Human-Machine Shared Contexts, Elsevier, pp. 177-204, June 2020.

2---**Context: Separating the Forest and The Trees -- Wavelet Contextual Conditioning For AI**, Chapter 4, Stephen Russell, Ira S. Moskowitz and Brian Jalain, Human-Machine Shared Contexts, Elsevier, pp. 67-91, June 2020.

3---**Integrating Expert Human Decision-Making in Artificial Intelligence Applications,** Chapter 13, Hesham Fouad, Ira S. Moskowitz, Derek Brock and Michael Scott, Human-Machine Shared Contexts, Elsevier, pp. 257-275, June 2020.

4--- **Homology as an Attack Indicator,** Ira S. Moskowitz, Nolan Bay, Brian Jalaian, and Arnold Tunick, Adversary-Aware Learning Techniques and Trends in Cybersecurity, pp. 167-196, Springer, January 2021.

5---**Cyber-(in)security, revisited: Proactive Cyber-Defenses, Interdependence and Autonomous Human-Machine Teams (A-HMTs),** W.F. Lawless, R. Mittu, I.S. Moskowitz, D.A. Sofge, and S. Russell., Adversary-Aware Learning Techniques and Trends in Cybersecurity, Springer, January 2021.

6---**Re-orienting towards the Science of the Artificial: Engineering AI Systems,** Stephen Russell, Brian Jalaian and Ira S. Moskowitz**,** book chapter for the book "Systems Engineering and Artificial Intelligence," published by Springer, editors Lawless et al., pp. 149-174, Fall, 2021.

7--- **A Fractional Brownian Motion Approach to Psychological and Team Diffusion Problems,** Ira S. Moskowitz, Noelle Brown and Zvi Goldstein, book chapter for the book "Systems Engineering and Artificial Intelligence," published by Springer, editors Lawless et al., pp.213-246, Fall, 2021.

8---**Agent Team Action, Brownian Motion and Gambler's Ruin,** Ira S. Moskowitz, book chapter in "Engineering Artificially Intelligent Systems," edited by W. Lawless, J. Llinas, D. Sofge & R. Mittu, Lecture Notes in Computer Science (LNCS 13000), pp. 90-108, Springer, Nov 2021.

9---**A Cost Metric for Team Efficiency**, Ira S. Moskowitz, Frontiers in Physics, Section: Interdisciplinary Physics, ed. W. Lawless, Frontiers, April 25, 2022.

10---**Mutual Information and Multi-Agent Systems**. Ira S. Moskowitz, Pi Rogers and Stephen Russell. Entropy, under review 2022.