REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
19-01-2020	Final Report	15-Jul-2016 - 14-Jul-2017

4. TITLE AND SUBTITLE	
Final Report: Information Geometry: Geometrization of Science of Information	5a. CONTRACT NUMBER W911NF-16-1-0383
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER 611102
6. AUTHORS	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES	8. PERFORMING ORGANIZATION REPORT NUMBER
University of Michigan - Ann Arbor 3003 South State Street Ann Arbor, MI 48109 -1274	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211	10. SPONSOR/MONITOR'S ACRONYM(S) ARO 11. SPONSOR/MONITOR'S REPORT NUMBER(S) 69546-CS-DRP.10

12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Jun Zhang
UU	UU	UU	UU		19b. TELEPHONE NUMBER 734-763-6161

RPPR Final Report
as of 23-Jan-2020

Agency Code:

Proposal Number: 69546CSDRP **Agreement Number: W911NF-16-1-0383**

INVESTIGATOR(S):

Name: Jun Zhang
Email: junz@umich.edu
Phone Number: 7347636161
Principal: Y

Organization: **University of Michigan - Ann Arbor**
Address: 3003 South State Street, Ann Arbor, MI 481091274
Country: USA
DUNS Number: 073133571 EIN: 386006309

Report Date: 14-Oct-2017 Date Received: 19-Jan-2020
Final Report for Period Beginning 15-Jul-2016 and Ending 14-Jul-2017
Title: Information Geometry: Geometrization of Science of Information
Begin Performance Period: 15-Jul-2016 **End Performance Period:** 14-Jul-2019
Report Term: 0-Other

Submitted By: Jun Zhang Email: junz@umich.edu
Phone: (734) 763-6161

Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees: **STEM Participants:**

Major Goals: Information geometry is an emerging mathematical framework for modeling the space of probability density functions as forming a (possibly infinite-dimensional) differentiable manifold equipped with a Riemannian metric and a pair of conjugate connections that need not be metrical. Information geometry has so far been applied to many disciplines such as asymptotic statistics, Bayesian inference, information theory and coding, machine learning, neural computation, econometrics, cognitive psychology, etc. It is seen by some as a potentially unifying framework for “geometrizing” information in the same way that physics has been geometrized.

This project investigated the geometric structures of statistical manifolds and examined its applications in aforementioned branches. Under the Geometry Thrust, the project investigated affine immersion, equiaffine, symplectic, Hermitian and almost Hermitian, a Kahler and almost Kahler structures linked to statistical manifolds. Equiaffine structure allows one to define volume forms and enables affine (including centro-affine) immersions, i.e., building hypersurface representation of a statistical manifold in a flat ambient space -- this is immensely useful for scientific visualization and reasoning. Symplectic structure allows one to characterize the dynamics of pairs of dual variables, and hence study the time evolution of information systems. Hermitian structure allows one to complexify (i.e., use complex numbers) the underlying representation of information entities, enriching computation power. Kahler structure allows one to switch between symplectic and complex frameworks and offer deeper understanding of information processing. Under the Information Thrust, the project applied geometric constructs to provide quantitative measures of model complexity and data compression, extend source/channel coding results from the classic Shannon entropy (which characterizes extensive structures with independent components) to the Tsallis entropy (which arises from non-extensive structures with interactive components), and to an even more general family of entropy/cross-entropy that is characterized by arbitrary embedding function. Our project linked mirror symmetry, which emerged in the study of Calabi-Yau manifold in String Theory, to reference-representational biduality in information geometry. So the project achieved the overall goal of deepening our understanding of the mathematical foundation of information geometry which provided powerful tools for developing novel algorithms in statistics, learning, inference, control, optimization, coding, etc.

Accomplishments: 0. Background.

The space of *all* probability models can be viewed as differentiable manifold such that algorithms for estimating, scoring, updating, learning those models based on dynamic data streams have natural geometric interpretations. Information geometry prescribes standard recipe to capture efficiency of estimators and analyze their biases through scoring functions, and to provide invariant priors for Bayesian inference. Geometric mechanics solves, moving beyond classic mechanics, optimal control via the interplay of Hamiltonian representation (on cotangent manifold) and Lagrangian representation (on tangent manifold) with, in general, non-holonomic constraints.

Various geometric structures on a differentiable manifold M are:
1. Metric g: symmetric, positive-definite bilinear form, geodesic, Riemannian submanifold, conformal;
2. Affine connection: identifies tangent spaces at different points, auto-parallel curves, projective;
3. Volume form r: top differentiable form, orientability of manifold;
4. Symplectic form W: skew-symmetric, non-degenerate bilinear form, canonically prescribed on cotangent bundle T*M and specifies Hamiltonian dynamics of conservation systems;
5. Almost complex structure J: on tangent bundle TM, allows (if J is integrable) complex coordinates;
6. Kähler structure: mutually compatible metric+symplectic+complex structures; a subclass is the "Calabi-Yau manifold" which models "super-strings" in mathematical physics;
7. Dirac structure: direct sum of TM and T*M, with its base coordinates (x,v,p) "decoupled" (i.e., with duality gap) between the Hamiltonian (x,p) and the Lagrangian (x,v) systems.

Information Geometry provided the following geometric tools for statistical estimation and inference
1. Contrast (a.k.a. divergence) function: capturing (asymmetric) difference of probability models (e.g., KL divergence); they are objective functions in machine learning/neural nets;
2. Curvature: related to second-order efficiency of estimators, statistical embedding;
3. Torsion: related to integrable condition for local estimating equations;
4. Volume form: invariant, non-informative Bayesian prior (generalizing Jeffrey's prior);
5. Symplectic volume: Hamiltonian Monte Carlo algorithms for sampling probability mode
6. Biorthogonal coordinates: dual relationship between natural vs expectation parameters in specifying probability models; intersecting submanifold when mixed specifications;
7. Pythogorian relation: projections between exponential family and mixture families of probability models; e/m alternating optimization algorithms;
8. Hessian geometry: generalized linear (deformed log) model, generalized entropy;

1. Geometric thrust

A. We established the complex structure for information geometry, by proving the remarkable fact that statistical manifold (where information lives) can be a Kahler/para-Kahler manifold (where superstrings live). [Tao and Zhang, 2016; Fei and Zhang, 2017; Zhang and Fei, 2018; Grigorian and Zhang, 2019]
Why this is important: Information Geometry, using Riemannian metric and affine connections of statistical models, provides geometric understanding of statistical inference: both likelihood-based point estimation (second-order efficiency, integrability of local estimating functions) and Bayesian statistics (non-informative prior of statistical model). We now show that the entire geometric framework may admit complex-variable representation. The meaning of imaginary part for statistical estimation and inference is under current investigation.

B. We investigated the role of torsion of an affine connection in information geometry. This provides an alternative framework for the manifold of parametric family of probability models, one based on partially flat geometry. [Zhang and Khan, 2018, Zhang and Khan 2019, Zhang and Khan, under review]
Why is this important: The standard formulation of information geometry is based on a pair of torsion-free affine connections that are conjugate with respect to the Fisher-Rao Riemannian metric. A special case is the dually-flat (Hessian) manifold showing the duality of natural parameter and expectation parameter of an exponential family, and has been widely used in statistics, machine learning, optimization, etc. We now have an alternative differential geometric framework, namely "partially flat" geometry in which only one of the connections is flat and the other is curvature-free but carries torsion. This partially-flat construction turns out to be universally valid for any parametric family of probability models (not just the exponential family). So this is a substantial generalization to our current conceptualization of statistical manifold.

2. Computational thrust

A. We found the most general expression of Entropy, Cross-Entropy, Divergence, while demonstrating the Maxent Principle for Generalized Linear Model [Naudts and Zhang, 2017]
Why this is important: Maximum Entropy/Minimum Divergence Principle has been one of the fundamental first principles for complex, biological systems, and for designing adaptive learning algorithms in those systems. We are now able to map these concepts to general functions, not just exponential/logarithmic functions. This broadens tremendously the applicability of such principles of information theory, thermodynamics and statistical physics.

B. We revealed a link of optimal transport to information geometry, by showing that regularity of OT problems, as mediated through an MTW tensor, is actually related to the sign of some curvature tensors in the complex geometry of tangent bundle of a Hessian manifold. [Khan and Zhang, 2020]
Why this is important:
Optimal transport is an active, interdisciplinary field of mathematics (linking analysis, probability theory, Riemannian geometry, PDE) that has found wide applications in natural sciences, engineering, economics, machine learning, etc. The geometry it generates is complimentary to the information geometry. Previously OT was shown to be linked to para-Kahler geometry on MxM. Now we show a Kahler geometry for TM. We are working towards a full understanding of MxM, TM, and T*M.

C. We were able to link information geometry with geometry mechanics. In geometric mechanics, Hamiltonian H(x, p) and Lagrangian L(x,v) dynamics represent one and the same dynamics – they are “coupled”; this is because H (x,p) and L(x,v) are Legendre transformation pair. We use the divergence function D(x,v,p) of information geometry to measure the discrepancy of the two dynamics without assuming that H(x,p) and L(x,v) are coupled. When they are coupled (no duality gap), then Hessian manifold results. In general, the dynamics of H and L, as “decoupled” systems, could be described in the Dirac manifold, with almost Kähler and almost para-Kähler structures. [Leok and Zhang, 2018].
Why this is important:
We view L as describing the Lagrangian dynamics of action minimization system (optimization of an objective function) whereas H as the Hamiltonian dynamics of a conservation system (or constraints). So they collectively describe the generic scheme of “constrained optimization.” The next step is to work out “information mechanics,” or canonical systems of equations that characterize dynamics of information systems viewed as generalization to mechanical systems, through the dualistic framework of information geometry.

Training Opportunities: A major training accomplishment is postdoc Gabriel Khan (for a total of 9.5months), who obtained Ph.D. in Mathematics from the Ohio State University. Dr. Khan, whose specialty is in geometry, obtained his expertise in information geometry through postdoctoral work with the PI.

Another senior personnel trained through the grant is Dr. Yinbin Lei (for a total of 1.9 months), who is a pure topologist. Dr. Lei received training in cognitive modeling, and became a co-author with the PI on several joint publications.

University of Michigan students were recruited to work on this project as a part of education/training mission. This included graduate student temps (60 hours) and undergraduate student temp (172 hours).

Effort was undertaken to reach members of communities of these research activities, and for the purpose of enhancing public understanding and awareness of science and technology. In 2016-2017, the PI is a co-organizer of two AFRL-Academic workshop on “Mathematical Foundation of Probability, Uncertainty, and Decision” with AFRL/RY researchers Jared Culberson, Roman Ilin, and Ben Robinson. In August of 2017 and July of 2018, the PI offered lecture series (8 lectures) on information geometry at the Yau Center for Mathematical Sciences at Tsinghua University.

The PI plays an instrumental role in negotiation with Springer in launching a new journal Information Geometry. It is the first journal dedicated to the field of information geometry. The PI is serving as Co-Editor of the Journal.

Results Dissemination: Publications arising out of research activities as supported by the grant

Publications in peer-reviewed journals

1. Tao, J. and Zhang, J. (2016). Transformations and coupling relations for affine connections. Journal of Differential Geometry and Applications. 49: 111-130.
2. Fei, T. and Zhang, J. (2017). Interaction of Codazzi couplings with (para)-Kahler geometry. Results in Mathematics. 72: 2037-2056
3. Leok, M. and Zhang, J. (2017). Connecting information geometry to geometric mechanics. Entropy, 19: 518; doi: 10.3390/e19100518
4. Greenfield M. and Zhang, J (2018) Resolution to the topological social choice paradox. Mathematical Social Sciences. 93: 47-51.
5. Naudts, J. and Zhang, J. (2018) Information geometry under monotone embedding. Information Geometry. 1: 79-115.
6. Zhang, J. (2019) Characterizing projective geometry of binocular visual space through Mobius transformation. Journal of Mathematical Psychology. 88: 15-26.
7. Grigorian, S. and Zhang, J. (2019). (Para)-holomorphic and conjugate on (para-)Hermitian and (para-)Kahler manifolds. Results in Mathematics, 74: 150.
8. Zhang, J. and Khan, G. (2019). From Hessian to Weitzenbock: Manifolds with torsion-carrying connections. Information Geometry, 2: 77-98.
9. Zhang, J. and Shi Z. (2019). Bayesian inference as probability transfer across sample spaces. Decision. (in print)
10. Khan, G. and Zhang, J. (2020) The Kahler geometry of certain optimal transport problems. Archive for Rational Mechanics and Analysis. (in print)

Publications under review

1. Zhang, J. and Khan, G. (under review) Statistical mirror symmetry. Submitted to Differential Geometry and Applications.
2. Sun, Y. and Zhang, J. (under review). Characterizing the structure of interval and semi-order. Submitted to Journal of Mathematical Psychology.
3. Lei, Y. and Zhang. (under review). Limit and convergence of a sequence. Submitted to Journal of Mathematical Psychology
4. Lin, R., Zhang, J. and Zhang, H. (under review). On Reproducing Kernel Banach Spaces: Generic definitions and unified framework of construction. Submitted to Journal of Functional Analysis.

Publications in peer-reviewed proceedings:

1. Zhang, J. and Naudts, J. (2017). Information geometry under monotone embedding.Part I: divergence functions. In Nielson, F. and Barbaresco, F. (Eds), Geometric Science of Information, 3rd International Conference GSI2017 Springer LNCS 10589 (pp. 205-214).
2. Naudts, J. and Zhang, J. (2017). Information geometry under monotone embedding.Part II: geometry. In Nielson, F. and Barbaresco, F. (Eds), Geometric Science of Information, 3rd International Conference GSI2017 Springer LNCS 10589 (pp. 215-222).
3. Grigorian, S. and Zhang, J. (2017). (Para)-holomorphic connections for information geometry. In Nielson, F. and Barbaresco, F. (Eds), Geometric Science of Information, GSI2017 (pp. 186-194).

4. Zhang, J. and Fei, T. (2018). (Para-)Kahler structure in information geometry. In Ay, N., Gibilisco, P., Matus, F. (Eds) Information Geometry and Its Applications, IGAIA4 Proceedings. Springer Proceedings in Mathematics & Statistics Vol 252 (pp 297-321).
5. Zhang, J. and Khan, G. (2019). New geometry of parametric statistical models. In Nielson, F. and Barbaresco, F. (Eds), Geometric Science of Information,4th International Conference GSI2019 Springer LNCS 11712 (pp. 288-296).
6. Khan, G. and Zhang, J. (2019). Hessian curvature and optimal transport. To appear Nielson, F. and Barbaresco, F. (Eds), Geometric Science of Information, 4th International Conference GSI2019 Springer LNCS 11712 (pp. 423-430).

Publications as invited book chapters

1. Zhang, J. and Sun, Y. (2016). Subset systems: Mathematical abstraction of object and context. In Houpt and Blaha (Eds.).Mathematical Models of Perception and Cognition, Volume I: A Festschrift for James T. Townsend. Psychological Press.
2. Zhang, H. and Zhang, J. (2017). Learning with reproducing kernel Banach spaces. New Trends in Analysis and Interdisciplinary Applications. (pp.417-423). Birkhauser, Cham.
3. Zhang, J. and Zhang, H. (2018). Categorization based on similarity and features: The reproducing kernel Banach space (RKBS) approach. In W. Batchelder, H. Colonius, E.N. Dzhafarov, J. Myung (Eds.) New Handbook of Mathematical Psychology, Volume 2, Springer.
4. Naudts, J. and Zhang, J. (2019). Rho-tau embedding of statistical models. Geometric Structures of Information, Springer, Cham (pp.1-13).
5. Lei, Y. and Zhang, J. (2019). Generalizing topological set operators. Electronic Notes in Computer Science, Vol 345: 63-76.

Honors and Awards: Nothing to Report

Protocol Activity Status:

Technology Transfer: Nothing to Report

CONFERENCE PAPERS:

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: Geometric Science of Information GSI2017
Date Received: Conference Date: 06-Nov-2017 Date Published: 06-Nov-2017
Conference Location: Paris, France
Paper Title: Information geometry under monotone embedding. Part I: divergence functions
Authors: J Zhang, J. Naudts
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: Geometric Science of Information GSI2017
Date Received: Conference Date: 06-Nov-2017 Date Published: 06-Nov-2017
Conference Location: Paris, France
Paper Title: Information geometry under monotone embedding.Part II: geometry
Authors: J Naudts, J Zhang
Acknowledged Federal Support: **Y**

Publication Type: Conference Paper or Presentation **Publication Status:** 1-Published
Conference Name: Geometric Science of Information GSI2017
Date Received: Conference Date: 06-Nov-2017 Date Published:
Conference Location: Paris, France
Paper Title: (Para)-holomorphic connections for information geometry
Authors: S Grigorian, J Zhang
Acknowledged Federal Support: **Y**

RPPR Final Report
as of 23-Jan-2020

ARO Final Report

Project title: Information Geometry: Geometerization of Science of Information

Contract/grant number: W911NF1610383

Author of the report: Jun Zhang (PI)

Period covered by report: 7/15/2016-7/14/2019

Performing organization and address: The University of Michigan, 530 Church Street, Ann Arbor, MI 48109

PAF: F043728

Abstract

Information geometry is an emerging mathematical framework for modeling the space of probability density functions as forming a (possibly infinite-dimensional) differentiable manifold equipped with a Riemannian metric and a pair of conjugate connections that need not be metrical. Information geometry has so far been applied to many disciplines such as asymptotic statistics, Bayesian inference, information theory and coding, machine learning, neural computation, econometrics, cognitive psychology, etc. It is seen by some as a potentially unifying framework for "geometrizing" information in the same way that physics has been geometrized.

This project investigated the geometric structures of statistical manifolds and examined its applications in aforementioned branches. Under the *Geometry Thrust*, the project investigated affine immersion, equiaffine, symplectic, Hermitian and almost Hermitian, a Kahler and almost Kahler structures linked to statistical manifolds. Equiaffine structure allows one to define volume forms and enables affine (including centro-affine) immersions, i.e., building hypersurface representation of a statistical manifold in a flat ambient space -- this is immensely useful for scientific visualization and reasoning. Symplectic structure allows one to characterize the dynamics of pairs of dual variables, and hence study the time evolution of information systems. Hermitian structure allows one to complexify (i.e., use complex numbers) the underlying representation of information entities, enriching computation power. Kahler structure allows one to switch between symplectic and complex frameworks and offer deeper understanding of information processing. Under the *Information Thrust*, the project applied geometric constructs to provide quantitative measures of model complexity and data compression, extend source/channel coding results from the classic Shannon entropy (which characterizes extensive structures with independent components) to the Tsallis entropy (which arises from non-extensive structures with interactive components), and to an even more general family of entropy/cross-entropy that is characterized by arbitrary embedding function. Our project linked mirror symmetry, which emerged in the study of

Calabi-Yau manifold in String Theory, to reference-representational biduality in information geometry. So the project achieved the overall goal of deepening our understanding of the mathematical foundation of information geometry which provided powerful tools for developing novel algorithms in statistics, learning, inference, control, optimization, coding, etc.

Accomplishment

0. Background.

The space of *all* probability models can be viewed as differentiable manifold such that algorithms for estimating, scoring, updating, learning those models based on dynamic data streams have natural geometric interpretations. Information geometry prescribes standard recipe to capture efficiency of estimators and analyze their biases through scoring functions, and to provide invariant priors for Bayesian inference. Geometric mechanics solves, moving beyond classic mechanics, optimal control via the interplay of Hamiltonian representation (on cotangent manifold) and Lagrangian representation (on tangent manifold) with, in general, non-holonomic constraints.

Various geometric structures on a differentiable manifold *M* are:

1. *Metric g:* symmetric, positive-definite bilinear form, geodesic, Riemannian submanifold, conformal;
2. *Affine connection*: identifies tangent spaces at different points, auto-parallel curves, projective;
3. *Volume form r:* top differentiable form, orientability of manifold;
4. *Symplectic form W:* skew-symmetric, non-degenerate bilinear form, canonically prescribed on cotangent bundle *T*M* and specifies Hamiltonian dynamics of conservation systems;
5. *Almost complex structure J:* on tangent bundle *TM*, allows (if *J* is integrable) complex coordinates;
6. *Kähler structure*: mutually compatible metric+symplectic+complex structures; a subclass is the "Calabi-Yau manifold" which models "super-strings" in mathematical physics;
7. *Dirac structure:* direct sum of *TM* and *T*M,* with its base coordinates *(x,v,p)* "decoupled" (i.e., with duality gap) between the Hamiltonian *(x,p)* and the Lagrangian *(x,v)* systems.

Information Geometry provided the following geometric tools for statistical estimation and inference

1. *Contrast (a.k.a. divergence) function*: capturing (asymmetric) difference of probability models (e.g., KL divergence); they are objective functions in machine learning/neural nets;
2. *Curvature*: related to second-order efficiency of estimators, statistical embedding;
3. *Torsion*: related to integrable condition for local estimating equations;

4. *Volume form*: invariant, non-informative Bayesian prior (generalizing Jeffrey's prior);
5. *Symplectic volume*: Hamiltonian Monte Carlo algorithms for sampling probability mode
6. *Biorthogonal coordinates*: dual relationship between natural vs expectation parameters in specifying probability models; intersecting submanifold when mixed specifications;
7. *Pythogorian relation*: projections between exponential family and mixture families of probability models; e/m alternating optimization algorithms;
8. *Hessian geometry*: generalized linear (deformed log) model, generalized entropy;

<u>*1. Geometric thrust*</u>

A. We established the complex structure for information geometry, by proving the remarkable fact that statistical manifold (where information lives) can be a Kahler/para-Kahler manifold (where superstrings live). [Tao and Zhang, 2016; Fei and Zhang, 2017; Zhang and Fei, 2018; Grigorian and Zhang, 2019]
<u>Why this is important:</u> Information Geometry, using Riemannian metric and affine connections of statistical models, provides geometric understanding of statistical inference: both likelihood-based point estimation (second-order efficiency, integrability of local estimating functions) and Bayesian statistics (non-informative prior of statistical model). We now show that the entire geometric framework may admit complex-variable representation. The meaning of imaginary part for statistical estimation and inference is under current investigation.

B. We investigated the role of torsion of an affine connection in information geometry. This provides an alternative framework for the manifold of parametric family of probability models, one based on partially flat geometry. [Zhang and Khan, 2018, Zhang and Khan 2019, Zhang and Khan, under review]
<u>Why is this important:</u> The standard formulation of information geometry is based on a pair of torsion-free affine connections that are conjugate with respect to the Fisher-Rao Riemannian metric. A special case is the dually-flat (Hessian) manifold showing the duality of natural parameter and expectation parameter of an exponential family, and has been widely used in statistics, machine learning, optimization, etc. We now have an alternative differential geometric framework, namely "partially flat" geometry in which only one of the connections is flat and the other is curvature-free but carries torsion. This partially-flat construction turns out to be universally valid for any parametric family of probability models (not just the exponential family). So this is a substantial generalization to our current conceptualization of statistical manifold.

<u>*2. Computational thrust*</u>

A. We found the most general expression of Entropy, Cross-Entropy, Divergence, while demonstrating the Maxent Principle for Generalized Linear Model [Naudts and Zhang, 2017]
Why this is important*:* Maximum Entropy/Minimum Divergence Principle has been one of the fundamental first principles for complex, biological systems, and for designing adaptive learning algorithms in those systems. We are now able to map these concepts to general functions, not just exponential/logarithmic functions. This broadens tremendously the applicability of such principles of information theory, thermodynamics and statistical physics.

B. We revealed a link of optimal transport to information geometry, by showing that regularity of OT problems, as mediated through an MTW tensor, is actually related to the sign of some curvature tensors in the complex geometry of tangent bundle of a Hessian manifold. [Khan and Zhang, 2020]
Why this is important*:*
Optimal transport is an active, interdisciplinary field of mathematics (linking analysis, probability theory, Riemannian geometry, PDE) that has found wide applications in natural sciences, engineering, economics, machine learning, etc. The geometry it generates is complimentary to the information geometry. Previously OT was shown to be linked to para-Kahler geometry on MxM. Now we show a Kahler geometry for TM. We are working towards a full understanding of MxM, TM, and T*M.

C. We were able to link information geometry with geometry mechanics. In geometric mechanics, Hamiltonian H(x,p) and Lagrangian L(x,v) dynamics represent one and the same dynamics – they are "coupled"; this is because H(x,p) and L(x,v) are Legendre transformation pair. We use the divergence function D(x,v,p) of information geometry to measure the discrepancy of the two dynamics without assuming that H(x,p) and L(x,v) are coupled. When they are coupled (no duality gap), then Hessian manifold results. In general, the dynamics of H and L, as "decoupled" systems, could be described in the Dirac manifold, with almost Kähler and almost para-Kähler structures. [Leok and Zhang, 2018].
Why this is important:
We view L as describing the Lagrangian dynamics of action minimization system (optimization of an objective function) whereas H as the Hamiltonian dynamics of a conservation system (or constraints). So they collectively describe the generic scheme of "constrained optimization." The next step is to work out "information mechanics," or canonical systems of equations that characterize dynamics of information systems viewed as generalization to mechanical systems, through the dualistic framework of information geometry.

Publications arising out of research activities as supported by the grant

Publications in peer-reviewed journals

1. Tao, J. and Zhang, J. (2016). Transformations and coupling relations for affine

connections. *Journal of Differential Geometry and Applications*. **49**: 111-130.
2. Fei, T. and Zhang, J. (2017). Interaction of Codazzi couplings with (para)-Kahler geometry. *Results in Mathematics*. **72:** 2037-2056
3. Leok, M. and Zhang, J. (2017). Connecting information geometry to geometric mechanics. *Entropy,* **19:** 518; doi:10.3390/e19100518
4. Greenfield M. and Zhang, J (2018) Resolution to the topological social choice paradox. *Mathematical Social Sciences.* **93***: 47-51.*
5. Naudts, J. and Zhang, J. (2018) Information geometry under monotone embedding. *Information Geometry*. 1: 79-115.
6. Zhang, J. (2019) Characterizing projective geometry of binocular visual space through Mobius transformation. *Journal of Mathematical Psychology*. **88**: 15-26.
7. Grigorian, S. and Zhang, J. (2019). (Para)-holomorphic and conjugate on (para-)Hermitian and (para-)Kahler manifolds. *Results in Mathematics,* **74**: 150.
8. Zhang, J. and Khan, G. (2019). From Hessian to Weitzenbock: Manifolds with torsion-carrying connections. *Information Geometry,* **2:** 77-98.
9. Zhang, J. and Shi Z. (2019). Bayesian inference as probability transfer across sample spaces. *Decision. (in print)*
10. Khan, G. and Zhang, J. (2020) The Kahler geometry of certain optimal transport problems. *Archive for Rational Mechanics and Analysis. (in print)*

Publications under review

1. Zhang, J. and Khan, G. (under review) Statistical mirror symmetry. Submitted to Differential Geometry and Applications.
2. Sun, Y. and Zhang, J. (under review). Characterizing the structure of interval and semi-order. Submitted to *Journal of Mathematical Psychology.*
3. Lei, Y. and Zhang. (under review). Limit and convergence of a sequence. Submitted to *Journal of Mathematical Psychology*
4. Lin, R., Zhang, J. and Zhang, H. (under review). On Reproducing Kernel Banach Spaces: Generic definitions and unified framework of construction. Submitted to *Journal of Functional Analysis.*

Publications in peer-reviewed proceedings:

1. Zhang, J. and Naudts, J. (2017). Information geometry under monotone embedding.Part I: divergence functions. In Nielson, F. and Barbaresco, F. (Eds), *Geometric Science of Information, 3rd International Conference GSI2017* Springer LNCS 10589 (pp. 205-214).
2. Naudts, J. and Zhang, J. (2017). Information geometry under monotone embedding.Part II: geometry. In Nielson, F. and Barbaresco, F. (Eds), *Geometric Science of Information, 3rd International Conference GSI2017* Springer LNCS 10589 (pp. 215-222).
3. Grigorian, S. and Zhang, J. (2017). (Para)-holomorphic connections for information geometry. In Nielson, F. and Barbaresco, F. (Eds), *Geometric Science of Information, GSI2017* (pp. 186-194).

4. Zhang, J. and Fei, T. (2018). (Para-)Kahler structure in information geometry. In Ay, N., Gibilisco, P., Matus, F. (Eds) *Information Geometry and Its Applications, IGAIA4 Proceedings. Springer Proceedings in Mathematics & Statistics Vol 252* (pp 297-321).
5. Zhang, J. and Khan, G. (2019). New geometry of parametric statistical models. In Nielson, F. and Barbaresco, F. (Eds), *Geometric Science of Information,4th International Conference GSI2019* Springer LNCS 11712 (pp. 288-296).
6. Khan, G. and Zhang, J. (2019). Hessian curvature and optimal transport. To appear Nielson, F. and Barbaresco, F. (Eds), *Geometric Science of Information, 4th International Conference GSI2019* Springer LNCS 11712 (pp. 423-430).

Publications as invited book chapters

1. Zhang, J. and Sun, Y. (2016). Subset systems: Mathematical abstraction of object and context. In Houpt and Blaha (Eds.).Mathematical Models of Perception and Cognition, Volume I: A Festschrift for James T. Townsend. Psychological Press.
2. Zhang, H. and Zhang, J. (2017). Learning with reproducing kernel Banach spaces*.* *New Trends in Analysis and Interdisciplinary Applications*. (pp.417-423). Birkhauser, Cham.
3. Zhang, J. and Zhang, H. (2018). Categorization based on similarity and features: The reproducing kernel Banach space (RKBS) approach. In W. Batchelder, H. Colonius, E.N. Dzhafarov, J. Myung (Eds.) New Handbook of Mathematical Psychology, Volume 2, Springer.
4. Naudts, J. and Zhang, J. (2019). Rho-tau embedding of statistical models. Geometric Structures of Information, Springer, Cham (pp.1-13).
5. Lei, Y. and Zhang, J. (2019). Generalizing topological set operators. Electronic Notes in Computer Science, Vol 345: 63-76.

Supported Personnel and Subcontracts

PI: Jun Zhang effort – total 10.4 months
Postdoc: *Gabriel Khan* - 9.5 month

Subcontractors

- *Sergey Grigorian*, Assistant Professor of Mathematics, University of Texas Grande Rio Valley. Expertise in differentiable manifold, Hermitian manifold, Calabi-Yau manifold, G2 space.
- *Teng Fei*, Assistant Professor of Mathematics, Columbia University. Expertise in Kahler and para-Kahler manifold, Calabi-Yau manifold, hyperkahler manifold, etc.
- *Shing-Tung Yau*, Professor of Mathematics and Professor of Physics, Harvard University, Director of Center for Math Science and Applications. Fields Medalist, Wolf Prize winner, Clifford Prize winner, National Medal of Science awardee.

Invited Visitors and Research Participants:

- *Nihat Ay* (Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany). Gave a talk to Complex System Program "Mathematical Aspects of Embodied Intelligence"
- *Gerald Miseolek* (University of Notre Dame, USA). Gave a talk to Math Department "Diffeomorphism Groups and Nonlinear PDE in Information Geometry."
- *Zhongmin Shen* (Indiana University, IUPUI, USA). Gave a talk to Math Department "Information Geometry from Finsler Geometry Perspective".
- *Jan Naudts* (University of Antwerpen, Belgium). Gave a talk to Complex Systems program and to Physics Department at the University of Michigan Winter'18.
- *Fangyang Zheng* (Ohio State University). Research collaborations.

Local Collaborators:

In Winter'2017, we stated a biweekly reading group on Hamiltonian Markov Chain "HMC Focus Group": Fred Feinberg (Stats and Business School), Xuanlong Nygen (Stats), Yves Atchade (Stats), Antony Block (Math).

Research Dissemination

The results have been disseminated through academic conferences, workshops, invited symposia, and invited colloquium/seminar presentations.

Invited Colloquium and Seminar Presentations:

- 4/16 Harvard University, Center for Mathematical Sciences and Applications, *Kahler and para-Kahler structures in information geometry.* (Host: S.T. Yau)
- 5/16 Chinese Academy of Science, Academy of Mathematics and Systems Science, *Kahler and para-Kahler structures in information geometry.* (Host: Feiming Huang)
- 10/16 Wayne State University, Department of Mathematics Colloquium, *Information geometry: Geometric science of information* (Host: Hengguang Li)
- 10/16 University of California Los Angeles, Department of Statistics Colloquium, *Information geometry: From divergence functions to geometric structures.* (Host: Arash Amini)
- 10/16 University of California Irvine, Institute of Mathematical Behavioral Sciences (IMBS) Colloquium, *Information geometry: From divergence functions to geometric structures.* (Host: Louis Narens).
- 7/17 University College London, UK *Information geometry: Geometrization of science of information* (Host: Karl Friston)
- 9/17 University of Notre Dame. *Information Geometry: Geometrization of statistical inference and information.* (Host: Gerald Miseolek)
- 9/17 University of Michigan, Department of Mathematics Colloquium Talk. *Information geometry: Geometrization of statistical inference and information.*

(Host: Lizhen Ji)

- 11/17 Centre Nationale Research Scientifique CNRS-L2S, France. *Differential geometry for statistical and entropy-based inference.* (Host: Ali Mohhamad-Djafari)
- 11/17 University of Montpelier, Institute of Alexander Grothendieck, France. *Information geometry: differentiable manifold for statistical and entropy-based Inference*. (Host: Michel Boyom).
- 11/17 Max Planck Institute of Mathematics in the Sciences, Leipzig, Germany. *Statistical manifold and entropy-based inference*. (Host: Nihat Ay).
- 3/18 Center for Mathematical Science and Applications, Harvard University *Statistical manifold with almost (para-)complex structures.* (Host: ST Yau)
- 4/18 Chern Institute of Mathematics, Tianjin, China. *Statistical manifold and entropy-based inference.* (Host: Huitao Feng)
- 4/18 School of Mathematics, Xiamen University, Xiamen, China. *Statistical manifold with almost (para-)complex structures.* (Host: Chunping Zhong)
- 10/18 Department of Mathematics, Ohio State University. *Statistical manifold and entropy-based inference.* (Host: Fangyang Zheng).
- 11/18 Department of Statistics, University of Michigan. *Information geometry and maximum entropy inference.*
- 1/19 Department of Mathematics, Brandeis University. (Host: An Huang). *Information geometry and entropy-based inference.*
- 2/19 University of Montpelier, Institute of Alexander Grothendieck, France. *Information geometry: Differentiable manifold for statistical and entropy-based Inference*. (Host: Michel Boyom).
- 2/19 Department of Mathematics, Memphis University *Information geometry of statistical inference.* (Host: Robert Kozma).
- 2/19 Department of Electronic and Computer Engineering, University of Michigan *Information geometry: Geometrizing statistical inference.*
- 3/19 Institute of Statistical Mathematics (ISM), Japan *Some recent progress in information geometry.* (Host: Shiro Ikeda).
- 3/19 University of Tokyo, Japan *Some recent progress in information geometry.* (Host: Komaki).
- 3/19 University of California, Irvine, Department of Mathematics Infor*mation geometry and entropy-based inference.* (Host: Hongkai Zhao).
- 4/19 University of California, Los Angeles, Department of Mathematics *Information geometry: Geometry for statistical inference* (Host: Kefeng Liu/Gang Liu).
- 4/19 University of California San Diego, Department of Applied Physics and Mathematics *Information geometry: Geometrization of information and statistical inference.* (Host: Melvin Leok).
- 5/19 Harvard University, Department of Statistics *Information geometry: Geometrization of statistical inference.* (Host: Xiaoli Meng).
- 5/19 Chongqin Science and Technology University, China *Information geometry: Geometrization of statistical inference.* (Host: Huitao Feng).
- 7/19 University of Toronto, Canada. *Information geometry: Geometrization of statistical inference.* (Host: T.L. Wong)

Conference, Workshop, Invited Panel Presentations:

- 5/16 Plenary talk to International Conference on Intelligent Science and Big Data Engineering (IScIDE 2016), Guangzhou, China. *Geometric structures induced from divergence functions.*
- 06/16 Present a talk (with T Fei) to International Conference of Information Geometry and Its Applications IV (IGAIA IV), Prague, Czech*Kahler and para-Kahler structures for information geometry*
- 12/16 Plenary talk to International Workshop on Mathematical Issues of Information Sciences (MIIS'2016). Shenzhen, China. *Information Geometry: Mathematical foundation of information science.*
- 11/17 Invited talk to Max Planck Institute of Mathematical Sciences. (Organizer: Nihat Ay). *Statistical manifold and entropy-based inference.*
- 11/17 Presented a talk (with Sergey Grigorian as first author) to Geometric Science of Information GSI'2017, Paris France. *(Para-)Hermitian connections for information geometry.*
- 11/17 Presented a talk (with Jan Naudts as second author) to Geometric Science of Information GSI'2017, Paris France *Information geometry under monotone embedding, Part I.*
- 11/17 Presented a talk (with Jan Naudts as first author) to Geometric Science of Information GSI'2017, Paris France. *Information geometry under monotone embedding. Part II.*
- 1/18. Invited talk to Deep Learning Workshop Hong Kong University of Science and Technology, Hong Kong (Organizer: Yang Wang and YuanYao). *Differential geometry for statistical and entropy-based inference.*
- 2/18 Invited talk to Topology and Algebraic Geometry in Statistics (TAGS) workshop. Max Planck Institute of Mathematical Sciences, Leipzig, Germany (Organizer: Bernd Sturmfels). *Statistical manifold and entropy-based inference.*
- 3/18 Invited talk to Geometry, Imaging, Computing (GIC) workshop, Harvard CMSA (Organizer: Yang Wang) *Geometry of maximal entropy inference.*
- 5/18 Presented a talk (with Jan Naudts as second author) to Entropy 2018, Barcelona, Spain. *Information geometry under monotone embedding.*
- 5/18 Presented a poster (with Melvin Leok as second author) to Entropy 2018, Barcelona, Spain. *Connecting information geometry with geometric mechanics.*
- 5/18 Invited talk Fifth Bayesian, Fiducial, and Frequentist Conference: Foundation of Data Science. BFF5, Ann Arbor, Michigan (Organizer: Peter Song). *Geometry of maximal entropy inference.*
- 11/18 Invited talk to International Conference on Mathematics of Data Science (ICMDS), Old Dominian University (Organizer: Yuesheng Xu*). Information geometry and maximum entropy inference.*
- 11/18 Invited talk to Symposium on Contextual Probability, Purdue Winer Memorial Lectures. (Organizer: Ehtibar Dzhafarov) *Probability normalization,*

marginalization, and maximum entropy inference: an information geometric approach.

- 1/19 Invited talk to Global Aspects of Projective and Kahler Geometry, Tsinghua-Sanya International Mathematics Forum, Hainan, China (Organizer: Akito Futaiki). *Interaction of statistical structure with almost (para-)complex structures.*
- 3/19 Invited talk to Information Geometry and Affine Differential Geometry IGADG III, Nagoya, Japan (Organizer: Hideyuki Ishi) *Some recent progress in information geometry.*
- 6/19 Presented an invited talk (with Gabriel Khan as second author) to ICCM International Congress of Chinese Mathematicians, Beijing, China. *Statistical mirror symmetry.*
- 7/19 Presented a talk to 52th Annual Conference of Society for Mathematical Psychology, Montreal Canada. *Information geometry: A mathematical foundation of statistical inference, learning, and information.*
- 8/19 Presented a talk (with Gabriel Khan as second author) to Geometric Science of Information GSI'2019. Toulouse, France. *New geometry of parametric statistical models.*
- 8/19 Presented a talk (with Gabriel Khan as first author) to Geometric Science of Information GSI'2019. Toulouse France. *Hessian curvature and optimal transport.*

Training, Outreach, and Synergistic Activities

A major training accomplishment is postdoc Gabriel Khan (for a total of 9.5months), who obtained Ph.D. in Mathematics from the Ohio State University. Dr. Khan, whose specialty is in geometry, obtained his expertise in information geometry through postdoctoral work with the PI.

Another senior personnel trained through the grant is Dr. Yinbin Lei (for a total of 1.9 months), who is a pure topologist. Dr. Lei received training in cognitive modeling, and became a co-author with the PI on several joint publications.

University of Michigan students were recruited to work on this project as a part of education/training mission. This included graduate student temps (60 hours) and undergraduate student temp (172 hours).

Effort was undertaken to reach members of communities of these research activities, and for the purpose of enhancing public understanding and awareness of science and technology. In 2016-2017, the PI is a co-organizer of two AFRL-Academic workshop on "Mathematical Foundation of Probability, Uncertainty, and Decision" with AFRL/RY researchers Jared Culberson, Roman Ilin, and Ben Robinson. In August of 2017 and July of 2018, the PI offered lecture series (8 lectures) on information geometry at the Yau Center for Mathematical Sciences at Tsinghua University.

The PI plays an instrumental role in negotiation with Springer in launching a new journal Information Geometry. It is the first journal dedicated to the field of information geometry. The PI is serving as Co-Editor of the Journal.