Mathematicians meet Machine Learning

TMS Meeting January 17, 2022

Kenichi Bannai (Keio University/RIKEN)

Self Introduction

Kenichi Bannai（坂内健一）

2000 Ph.D. The University of Tokyo
2001–2008 Nagoya University
2008— Keio University

Speciality: **Arithmetic Geometry** ***pure mathematics***

- Various realizations of Polylogarithms
- Special Values of Hasse-Weil L-functions
- Bloch-Beilinson-Kato Conjecture

2016— RIKEN AIP/ Team Leader

Today

- RIKEN AIP
- Research conducted by our team
- Some Thoughts

RIKEN AIP

RIKEN（理化学研究所）

Research Institute dedicated to fundamental research in the natural sciences, including Physics, Chemistry, Biology, Medicine, Engineering, Informatics & more recently, Mathematics

Head Quarters:

Wakoshi

和光市

Saitama Prefecture

埼玉県

RIKEN（理化学研究所）

Founded in 1917

Tradition of Autonomy for Researchers

“Paradise for Researchers”





…

Gave rise to many companies:

RICOH imagine. change.

RIKEN KEIKI

KYOWA

Riken

…

RIKEN & Mathematics

Dairoku Kikuchi

1917 2016

iTHEMS Interdisciplinary Theoretical and Mathematical Sciences Program (Extension of iTHES)

AIP Center for Advanced Intelligence Project

RIKEN AIP

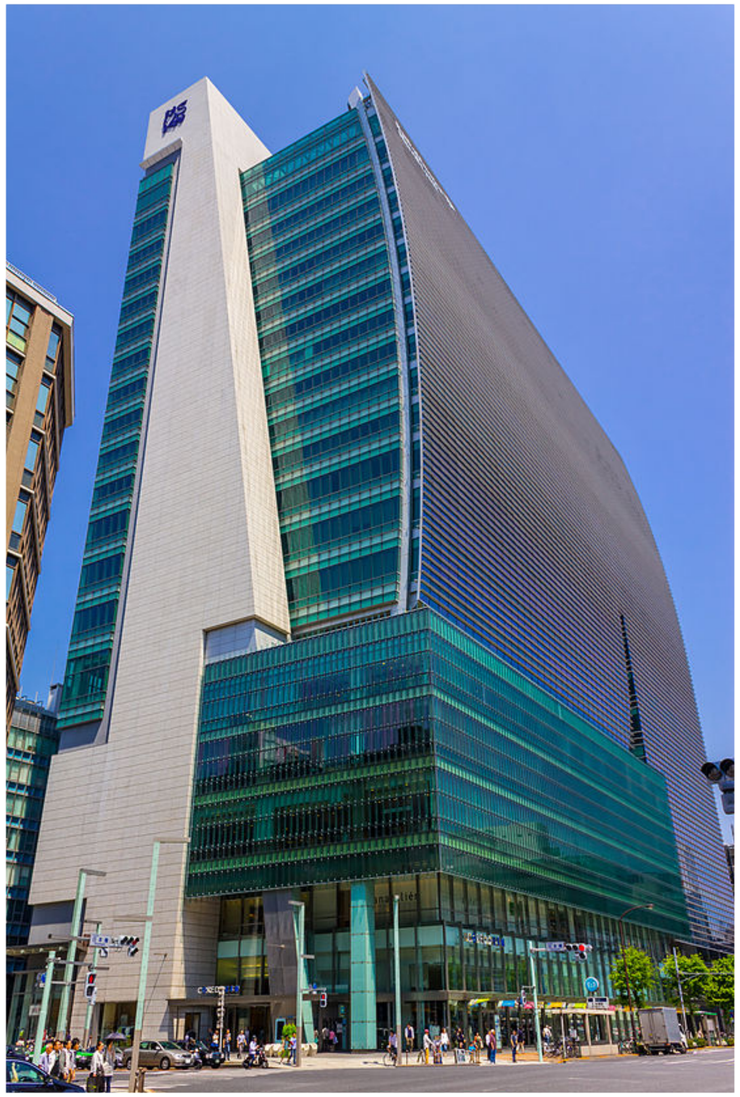
MEXT AIP Project

April 2016 — March 2026

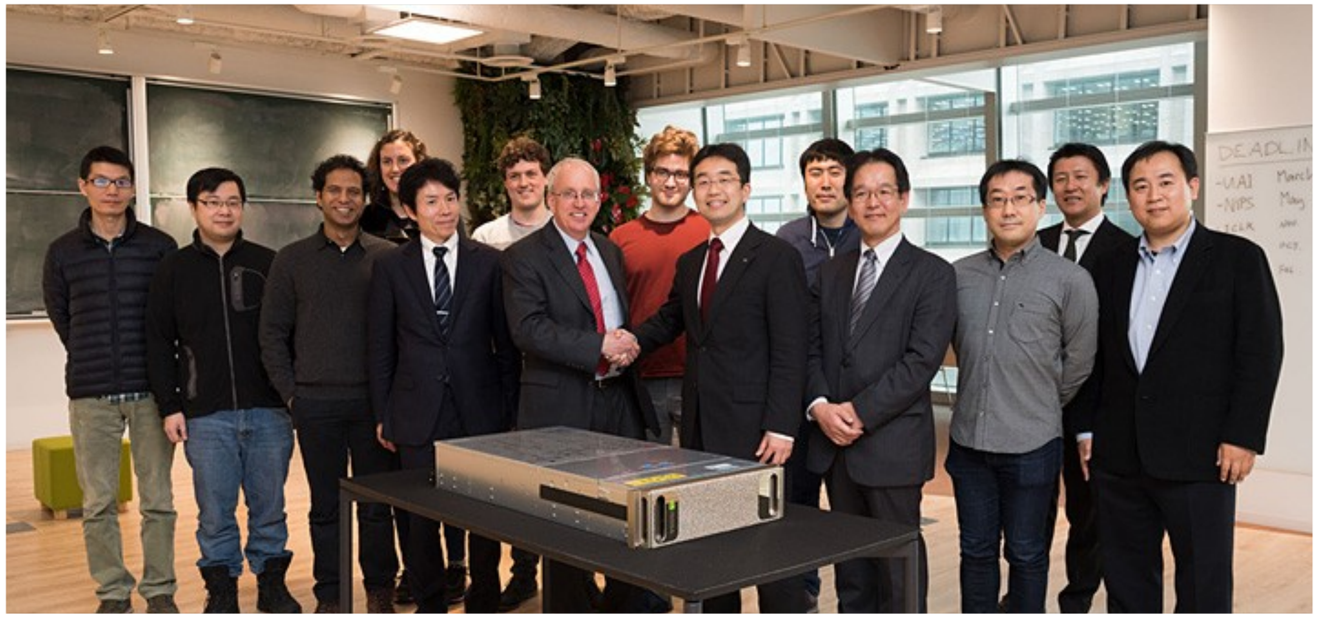
Mission

- Develop Next-Generation AI Technology
- Accelerate Scientific Research
- Solve Socially Critical Problems
- Consider Ethical, Legal and Social issues of AI
- Develop Next Generation of AI researchers

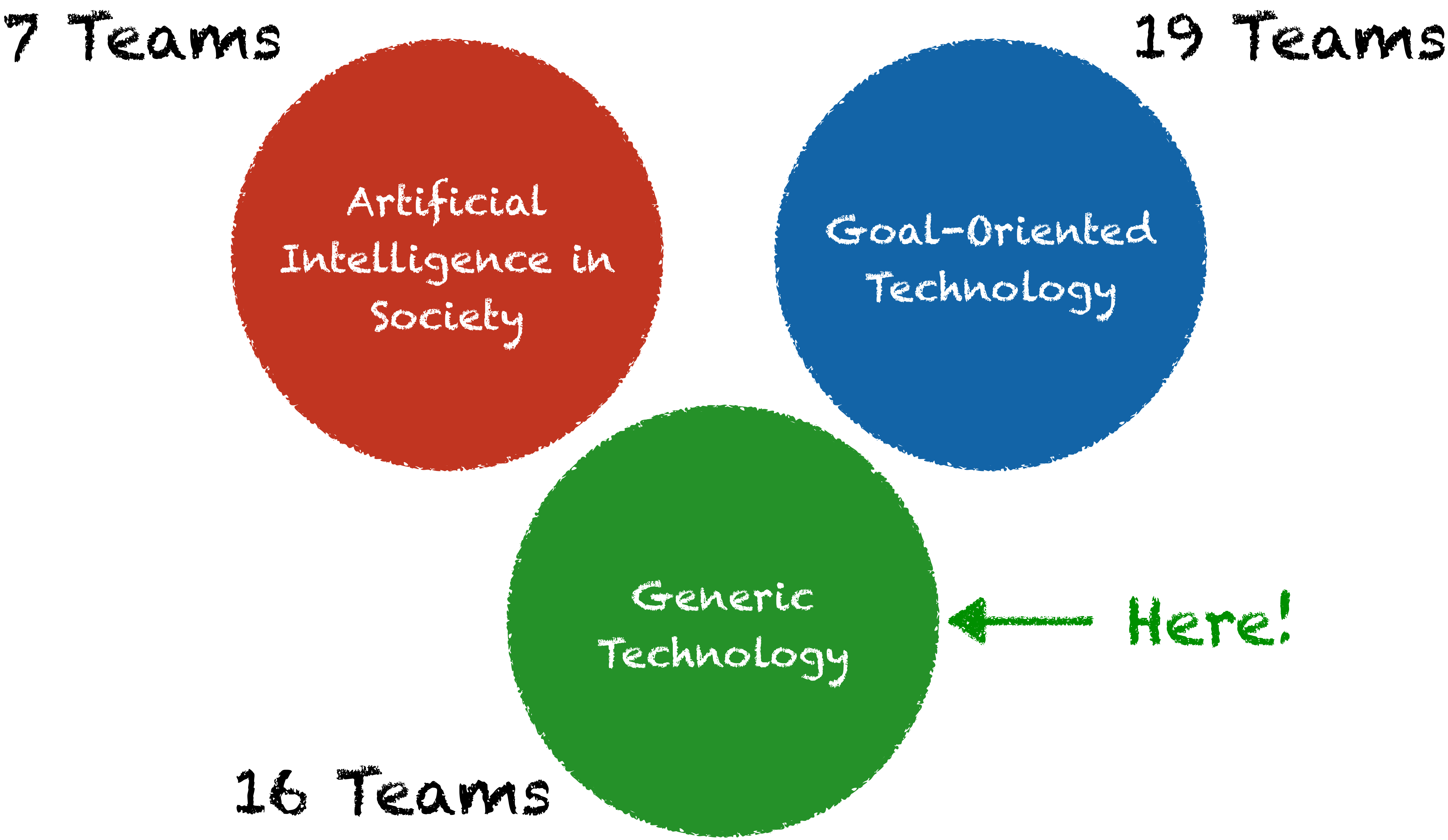
RIKEN AIP



Nihonbashi Office



RIKEN AIP


7 Teams
Artificial Intelligence in Society
19 Teams
Goal-Oriented Technology
Generic Technology
Here!
16 Teams

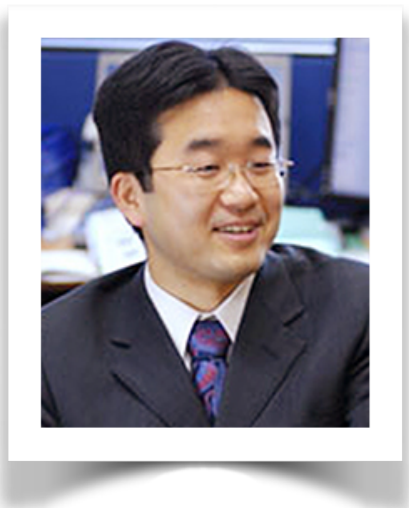
RIKEN AIP

2016 No Experience except pure mathematics

Masashi Sugiyama
Director AIP

Please organize a team of mathematicians to achieve a breakthrough in AI

RIKEN AIP



Behind each breakthrough,
a mathematical theory

Statistical Inference →

Optimization Theory →

Manifold Theory →

? → !

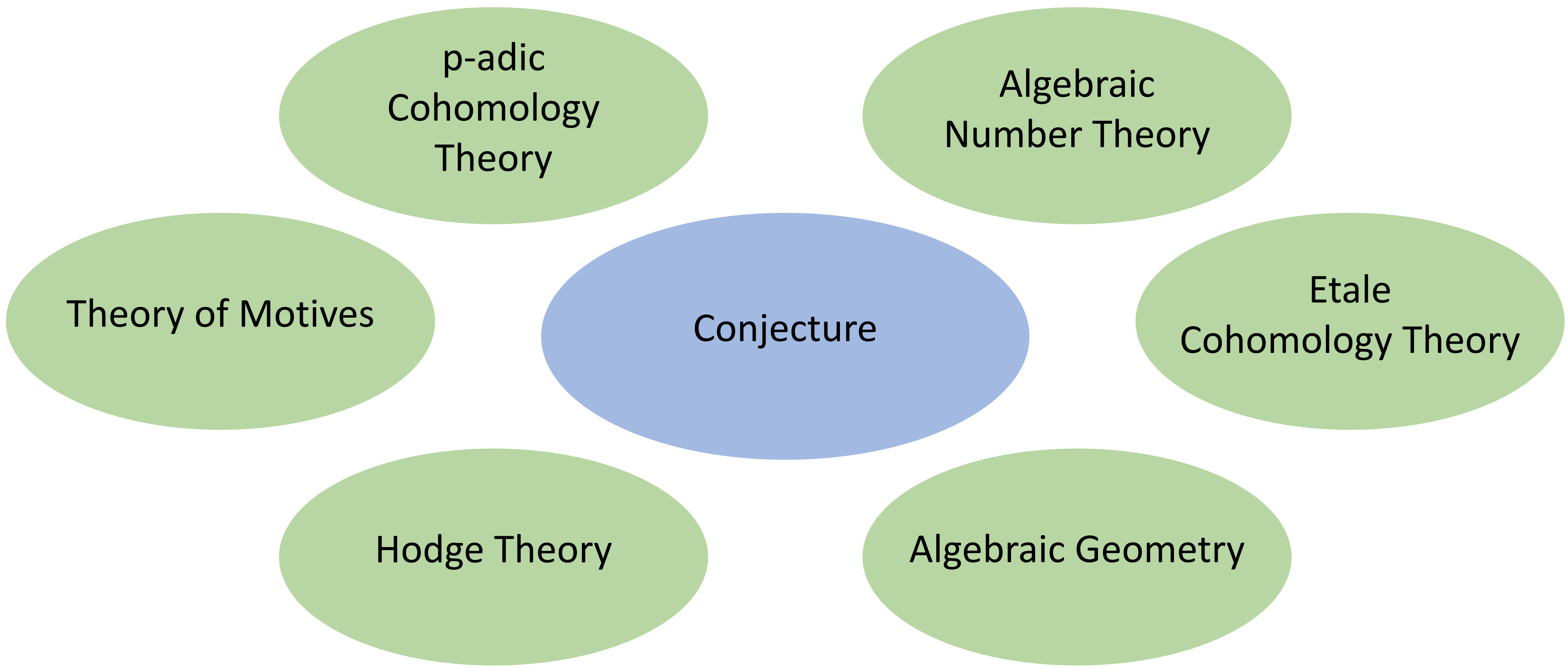
Machine Learning

Breakthrough with New Mathematical Theory ?!

RIKEN AIP

Bloch-Beilinson-Kato Conjecture

Pure Mathematics Special Values of L-functions of Algebraic Varieties

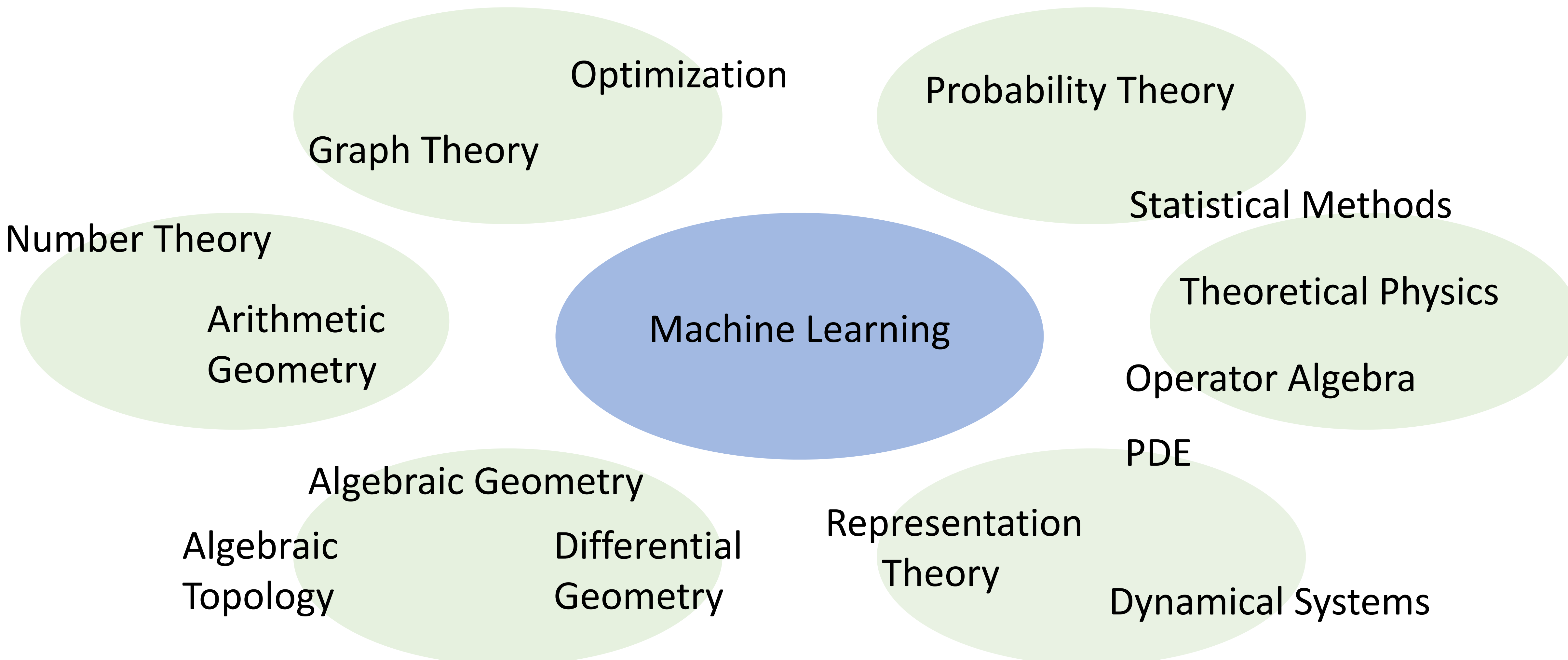


Collaborative Research with Experts of Different Fields

RIKEN AIP

Machine Learning

Machine Learning Fundamental Problems in Machine Learning



Collaborative Research with Experts of Different Fields

The Team

Research Scientist/Postdoctoral Researcher

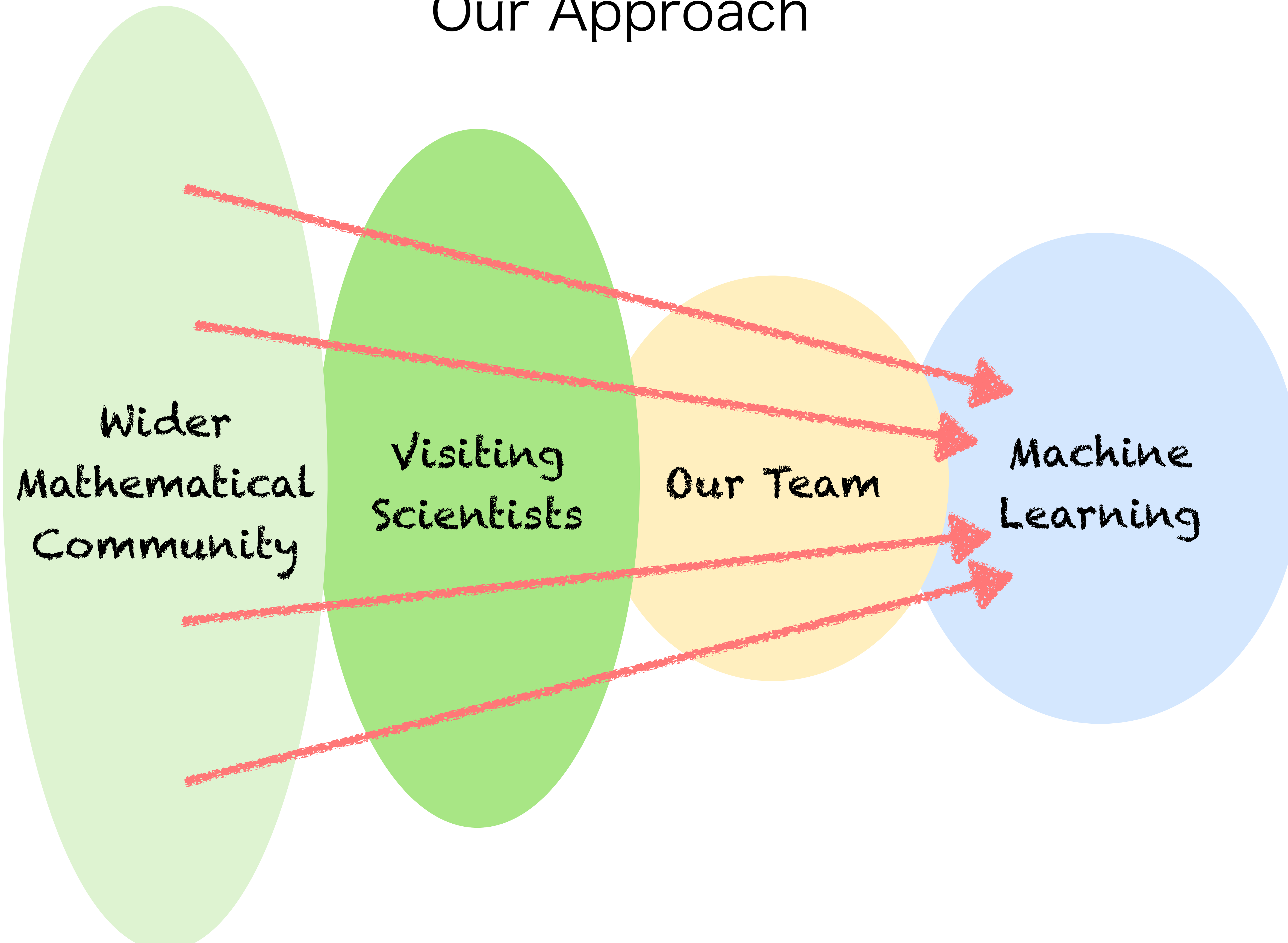
K. Hagihara (Arithmetic Geometry)
M. Ikeda (PDE)
T. Kuwahara (Mathematical Physics)
A. Sannai (Algebraic Geometry)
K. Tojo (Representation Theory)

RIKEN SPDR (Own Research)

E. Kiral (Analytic Number Theory)
D. Takeuchi (Arithmetic Geometry)
R. Sakamoto (Number Theory)

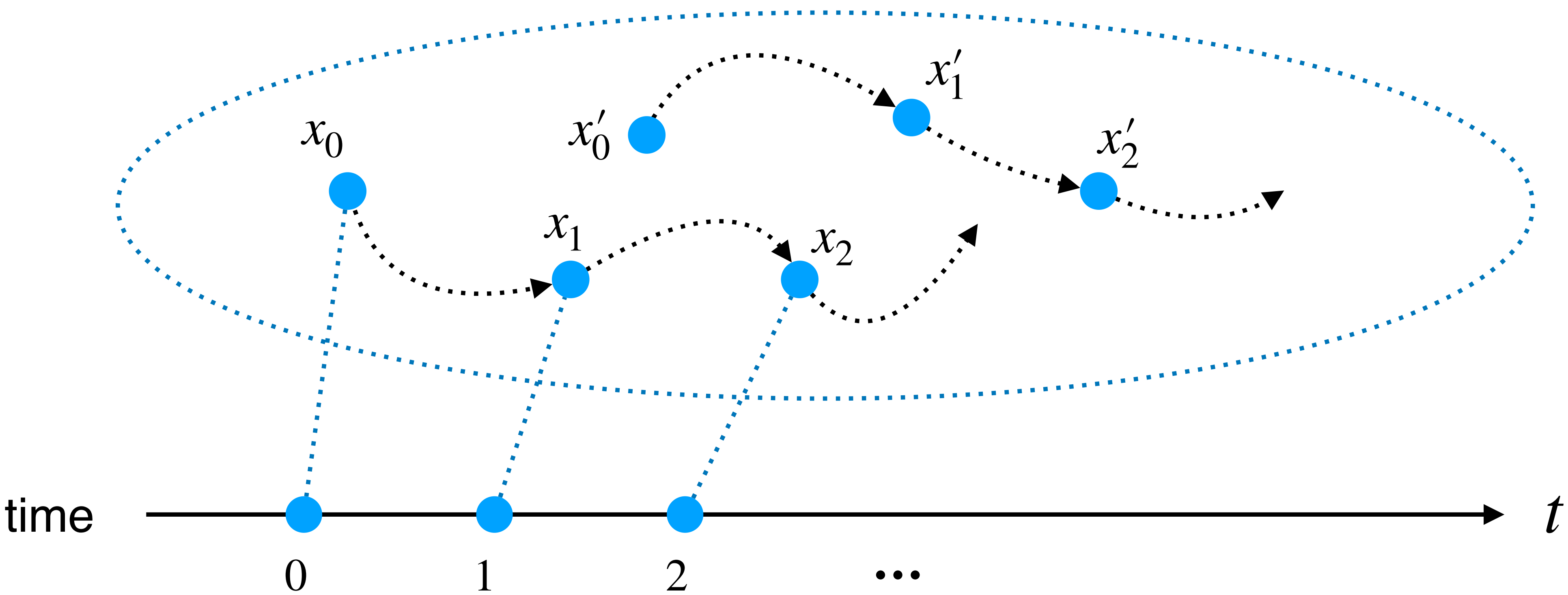
6 Ph.D. Students, Over **13 visiting scientists** in Arithmetic Geometry, Algebraic Topology, Differential Geometry, Graph Theory, Dynamical Systems, Real Analysis, Probability Theory, Optimization, Statistics

Our Approach



First Example

Times Series Data



$\{x_0, x_1, x_2, \ldots\}$ $\{x'_0, x'_1, x'_2, \ldots\}$

How different?

First Example

Times Series Data

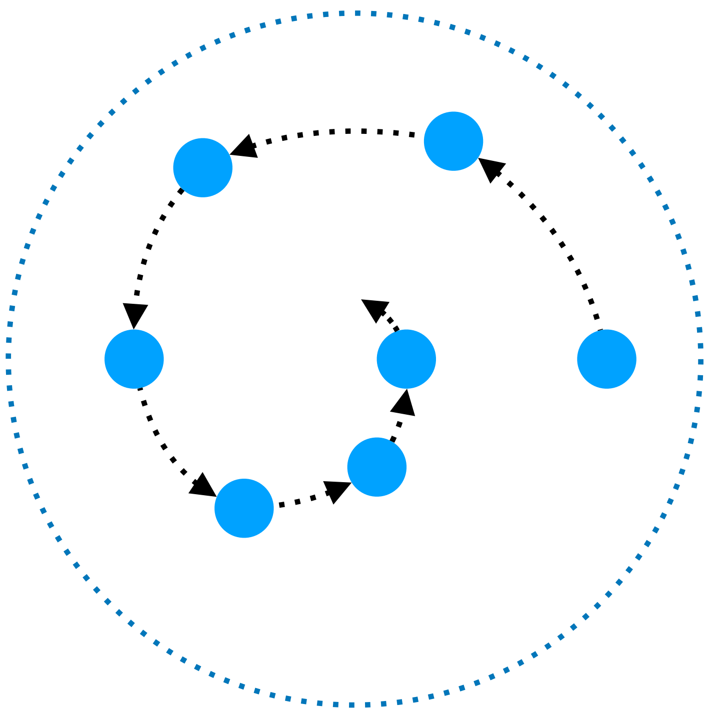
Linear and **Stable** case

$$x_{n+1} = Ax_n$$

Martin, *A Metric for ARMA Processes,*
IEEE Transactions on Signal Processing,
VOL. 48, NO. 4, APRIL 2000

K. De Cock and B. De Moor, Subspace angles between ARMA models, Systems & Control Letters, 46:4 (2002), pp. 265–270.

Example: Converging Rotation



Defined “distance” or “angle” measuring the difference of time-series data

non-Linear or **non-Stable** case?

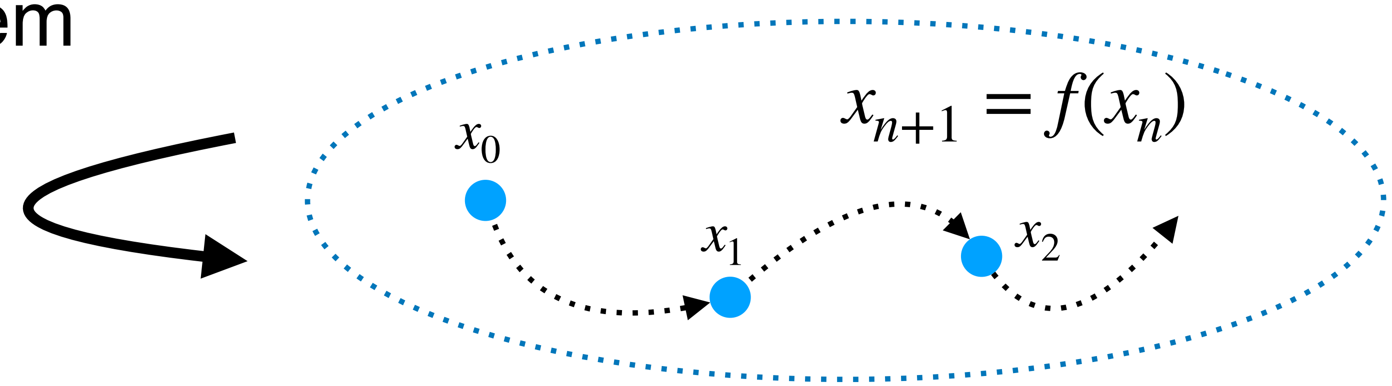
First Example

Times Series Data

non-Linear or **non-Stable** case?

Dynamical System

non-Linear f



AIP Kawahara TL

- **Reproducing Kernel Hilbert Space (RKHS)**
- **Koopman/Perron-Frobenius Operator**

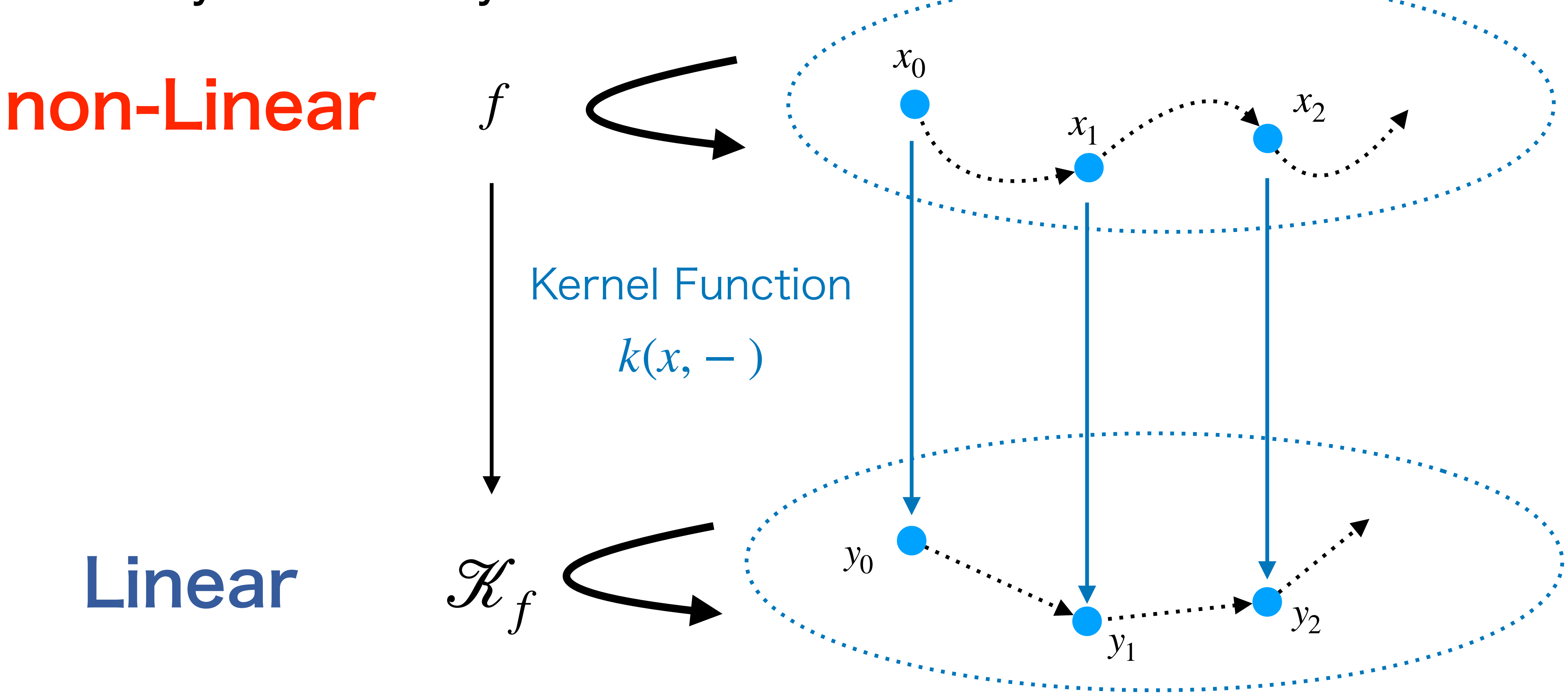
Y. Kawahara, Dynamic Mode Decomposition with Reproducing Kernels for Koopman Spectral Analysis, NeurIPS2016

First Example

Times Series Data

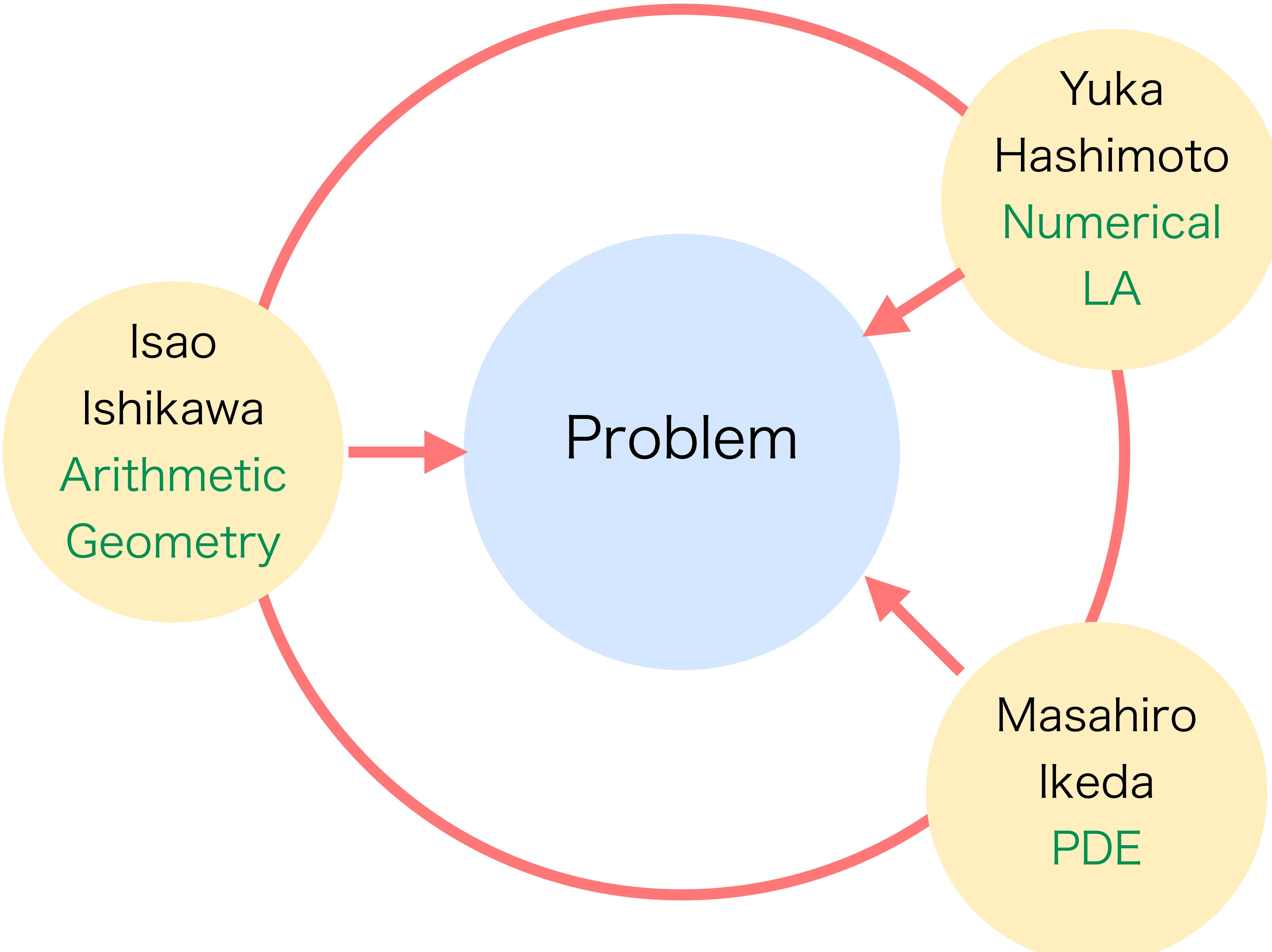
non-Linear or **non-Stable** case?

Dynamical System



Perron-Frobenius Operator

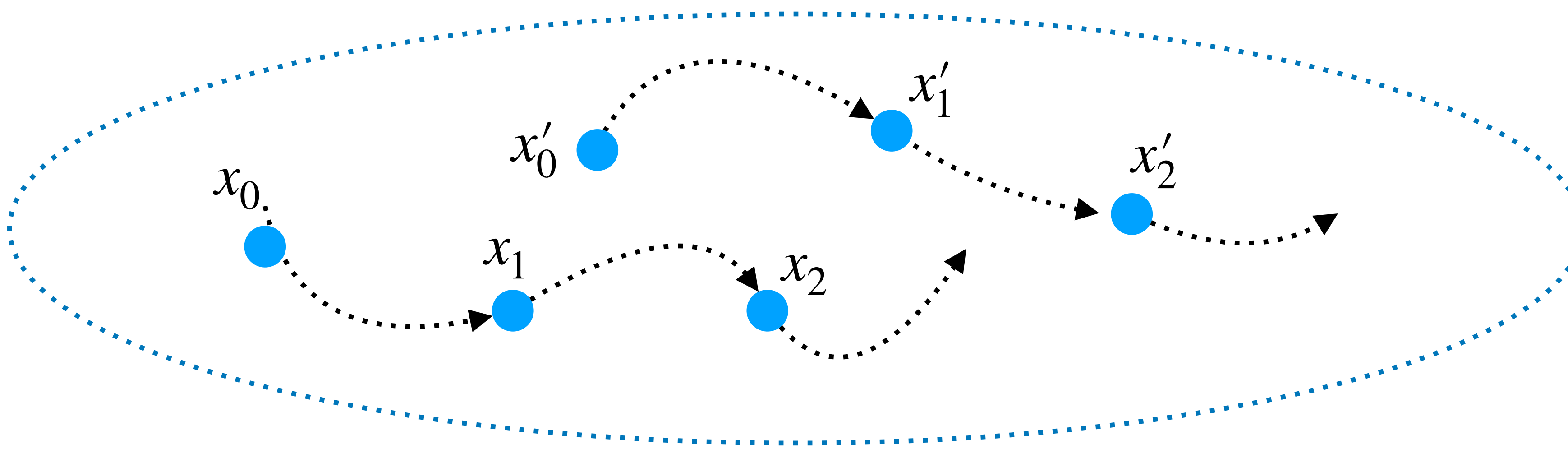
First Example



First Example

Times Series Data

non-Linear or **non-Stable** case?



$$\{x_0, x_1, x_2, \ldots\} \longleftrightarrow \{x'_0, x'_1, x'_2, \ldots\}$$

Succeeded in Defining "distance"

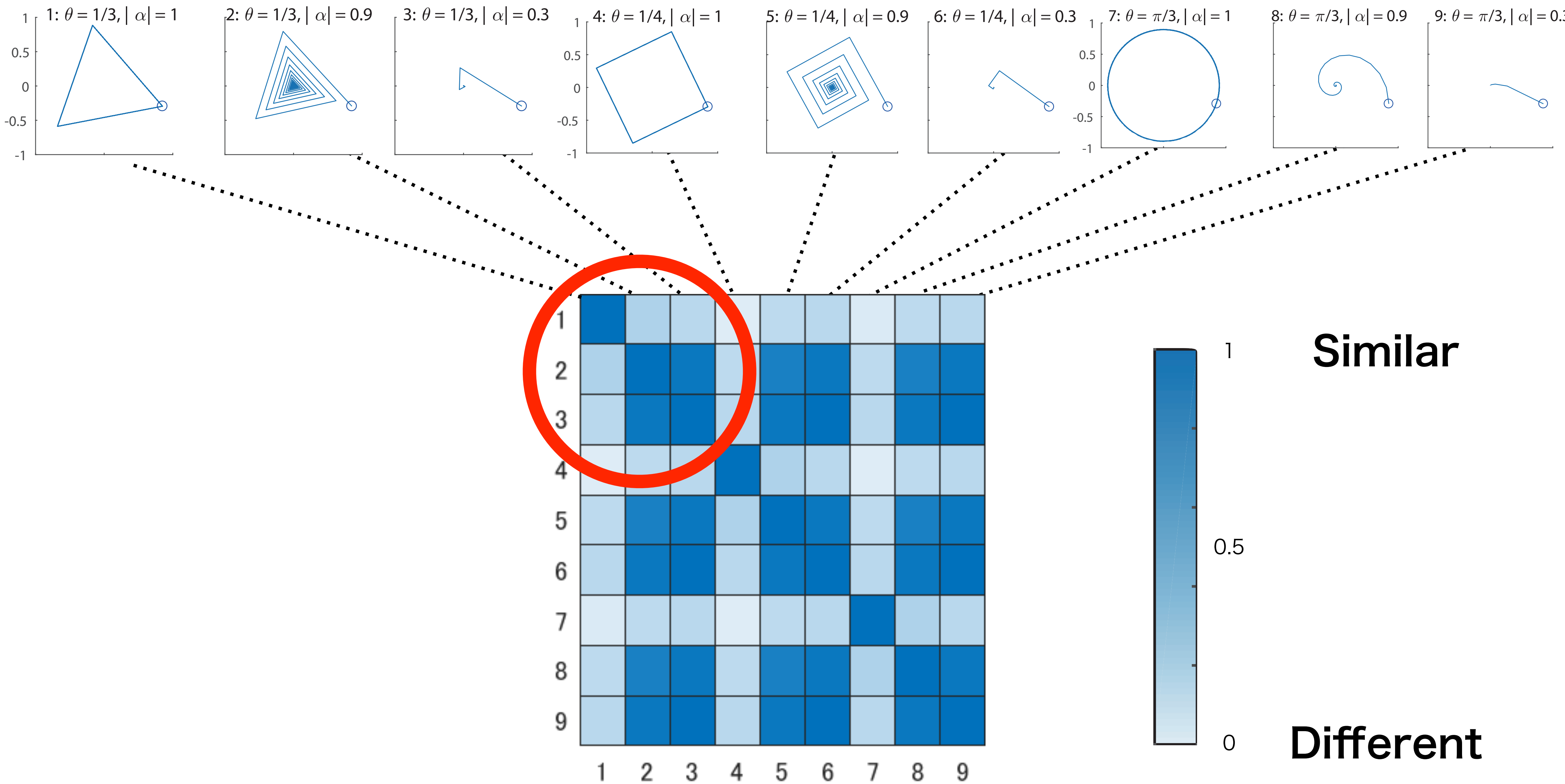
$$A_m(\boldsymbol{x}, \boldsymbol{x}') := \lim_{t \to \infty} A_m^t(\{x_0, \ldots, x_t\}, \{x'_0, \ldots, x'_t\})$$

I. Ishikawa, K. Fujii, M. Ikeda, Y. Hashimoto and Y. Kawahara,
Metric on Nonlinear Dynamical Systems with Perron-Frobenius Operators,
NeurIPS2018

I. Ishikawa

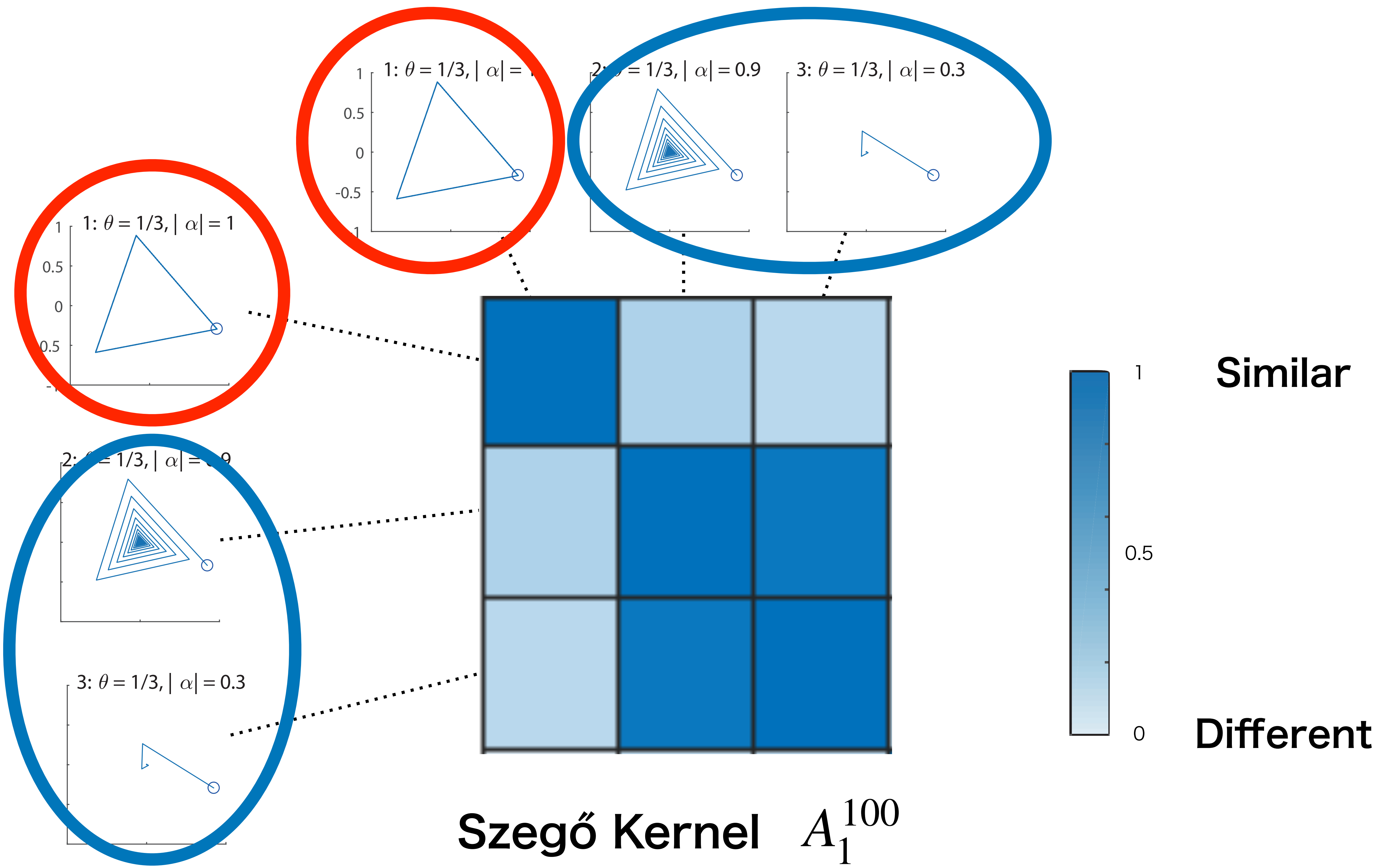
Awarded 2019 RIKEN Ohbu Prize（理研櫻舞賞）

Experimental Results



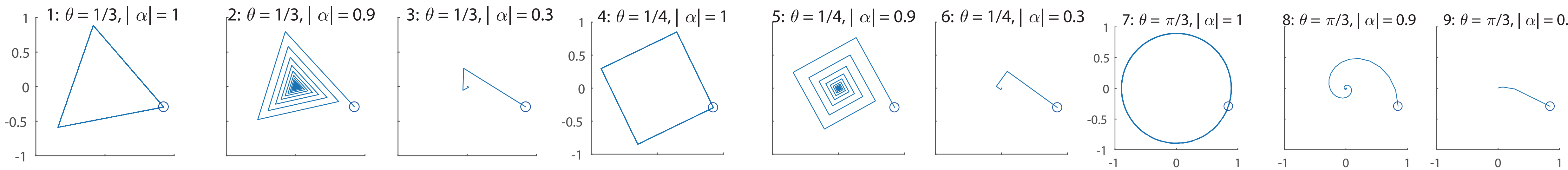
Szegő Kernel A_1^{100}

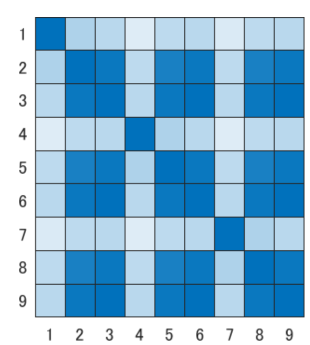
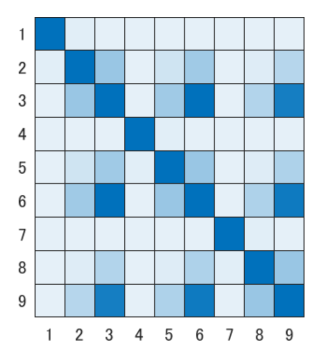
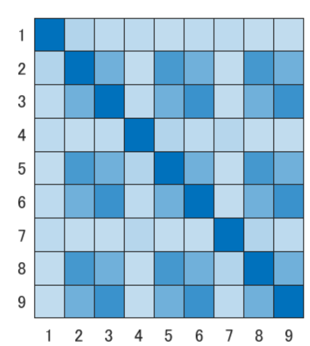
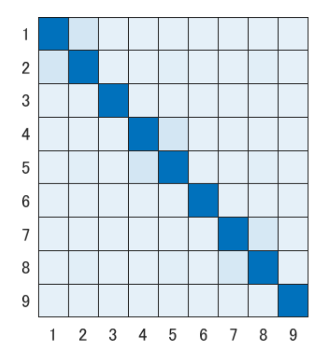
Experimental Results

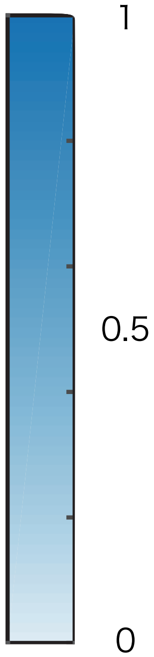


Szegő Kernel A_1^{100}

Experimental Results



	Szegö kernel		Gaussian kernel	
z_0	A_1^{100}	A_2^{100}	A_1^{100}	A_2^{100}
0.9	 (a)	 (b)	 (c)	 (d)



Similar

Different

Development

Times Series Data

Kawahara CREST

COMPUT DYN

CREST: Operator theoretic data analysis of complicated dynamics and its integrated utilization with mathematical models
FY2019—2024 (Oct. 2019—March 2025)

Relation to Higher Mathematics

- C*-algebras
- Boundedness of Operators
- Random Noise (Probability Theory)

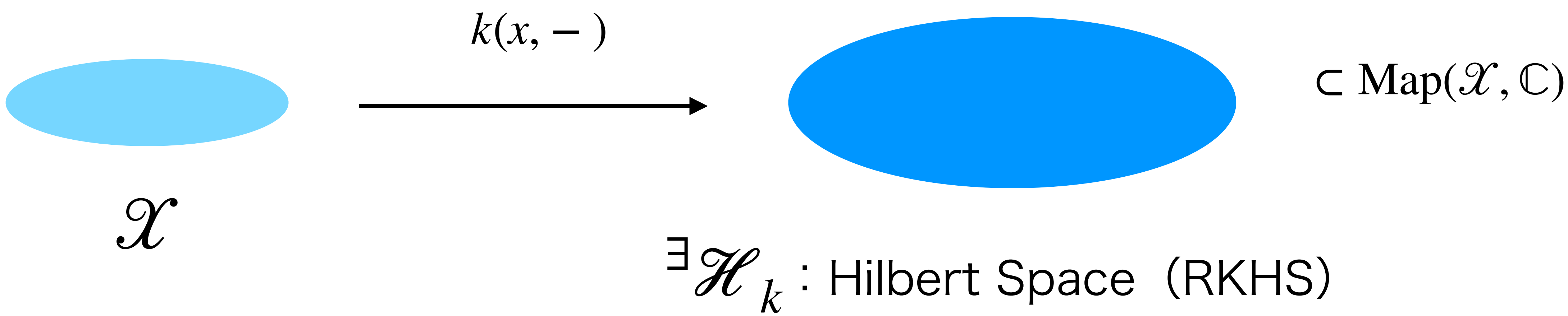
Recent

Reproducing Kernel Hilbert Space（RKHS）

Def. $\mathscr{X}$: set, $k : \mathscr{X} \times \mathscr{X} \to \mathbb{C}$: positive definite kernel

$\Leftrightarrow$ (i) $k(x, y) = \overline{k(y, x)}$

(ii) $\sum_{i,j=1}^{n} \overline{c_i} k(x_i, x_j) c_j \geq 0$ for any $x_1, \ldots, x_n \in \mathscr{X}, c_1, \ldots, c_n \in \mathbb{C}$



Recent

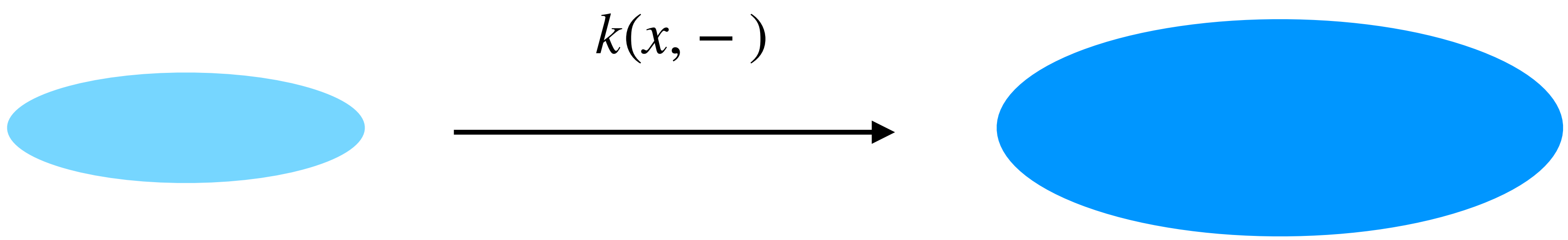
Reproducing Kernel Hilbert C* module (RKHM)

Def. $\mathscr{X}$: set, $k: \mathscr{X} \times \mathscr{X} \to \mathscr{A}$ $\mathscr{A}$: C* algebra

(i) $k(x, y) = k(y, x)^*$

(ii) $\sum_{i,j=1}^{n} c_i^* k(x_i, x_j) c_j \geq 0$ for any $x_1, \ldots, x_n \in \mathscr{X}, c_1, \ldots, c_n \in \mathbb{C}$

$k(x, -)$



$\subset \mathrm{Map}(\mathscr{X}, \mathscr{A})$

$\mathscr{X}$

$^\exists \mathscr{H}_k$: Hilbert C* module (RKHM)

+ Takeshi Katura, Fuyuta Komura

C^* algebra

Def （C^* algebra）

$\mathscr{A}$ ：C^* algebra ⇔ (i) $\mathscr{A}$ ：Banach Algebra over $\mathbb{C}$

(ii) ∃ involution $*: \mathscr{A} \to \mathscr{A}$

(iii) $\forall \lambda \in \mathbb{C} \quad a \in \mathscr{A}, \quad (\lambda a)^* = \bar{\lambda} a^*$

(iv) $\forall a \in \mathscr{A} \quad \|aa^*\|_{\mathscr{A}} = \|a\|_{\mathscr{A}} \|a^*\|_{\mathscr{A}}$

Example $\mathscr{A} = M_n(\mathbb{C}) \qquad A^* := \overline{{}^t A} \quad \forall A \in M_n(\mathbb{C})$

Bounded operators on Hilbert space

von Neumann algebras

First Idea: Replace $\mathbb{C}$ by $M_n(\mathbb{C})$

Hilbert C^* modules

$\mathscr{A}$: C^* algebra $\mathscr{M}$: right $\mathscr{A}$ module

Def ($\mathscr{A}$ inner product on $\mathscr{M}$)

$\langle -, - \rangle : \mathscr{M} \times \mathscr{M} \to \mathscr{A}$: inner product $\Leftrightarrow$ (i) $\mathscr{A}$ bilinear

(ii) $\langle u, v \rangle = \langle v, u \rangle^* \quad \forall u, v \in \mathscr{M}$

(iii) $\langle u, u \rangle \geq 0, \quad \langle u, v \rangle = 0 \Leftrightarrow u = 0$

Norm $\|u\| := \|\langle u, u \rangle\|_{\mathscr{A}}^{1/2}$ gives distance (topology) on $\mathscr{M}$

Def (Hilbert C^* module)

$\mathscr{A}$ module with $\mathscr{A}$ inner product, whose topology is complete

Representer Theorem

Thm (Representer Theorem)

$\mathscr{A}$: von Neuman algebra

$\mathscr{X}$: set, $k \colon \mathscr{X} \times \mathscr{X} \to \mathscr{A}$: positive definite kernel

$\mathscr{H}_k \subset \mathrm{Hom}(\mathscr{X}, \mathscr{A})$: Hilbert C*-module

$h \colon \mathscr{X} \times \mathscr{A}^2 \to \mathscr{A}_+$: loss function, $\mathscr{A}_+ := \{aa^* \mid a \in \mathscr{A}\}$

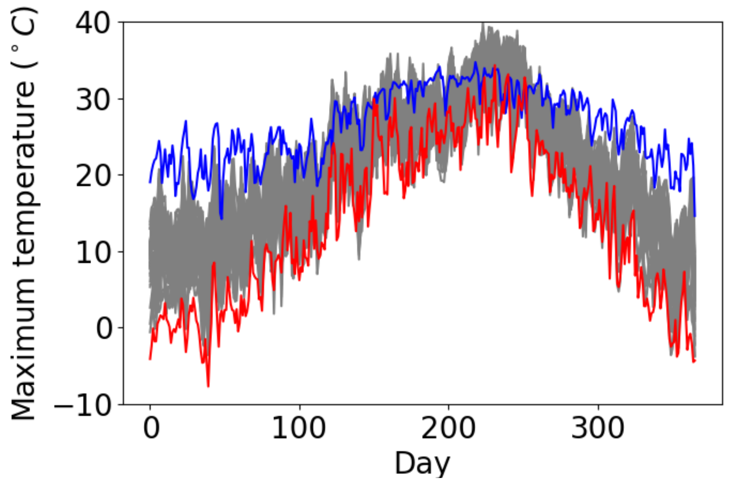
For any data $x_1, \ldots, x_n \in \mathscr{X}$ and $a_1, \ldots, a_n \in \mathscr{A}$,

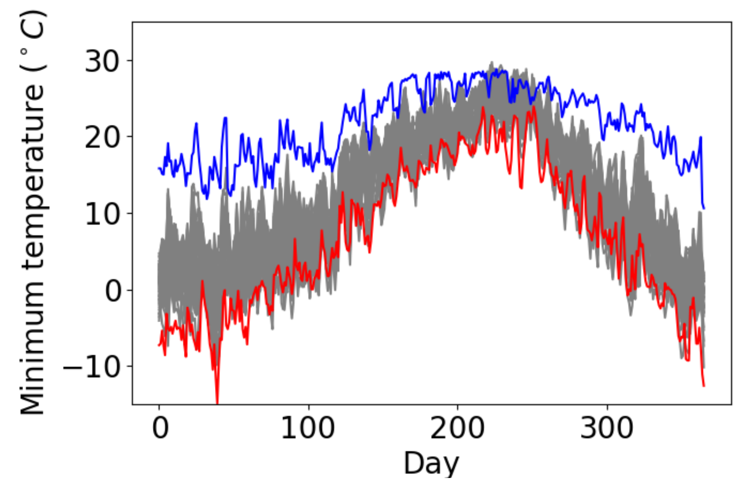
$u \in \mathscr{H}_k$ minimizing $\sum_{i=1}^{n} h(x_i, a_i, u(x_i))$ is of the form

$$u(-) = \sum_{i=1}^{n} c_i \langle x_i, - \rangle \text{ for some } c_1, \ldots, c_n \in \mathscr{A}$$

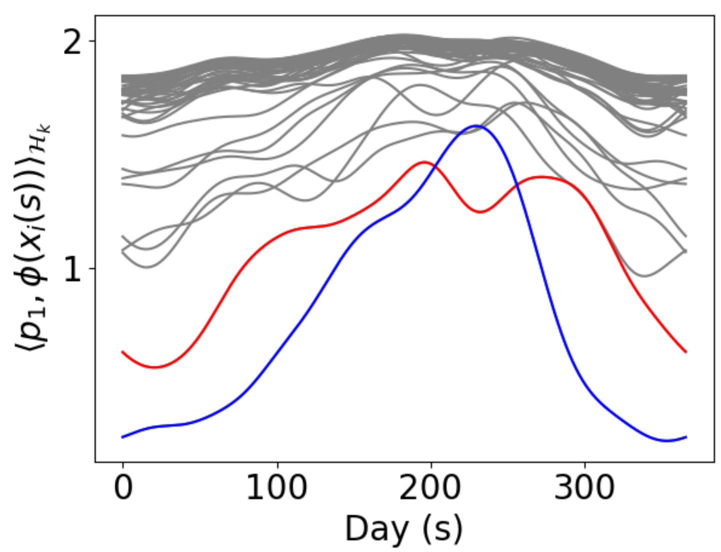
Experiments

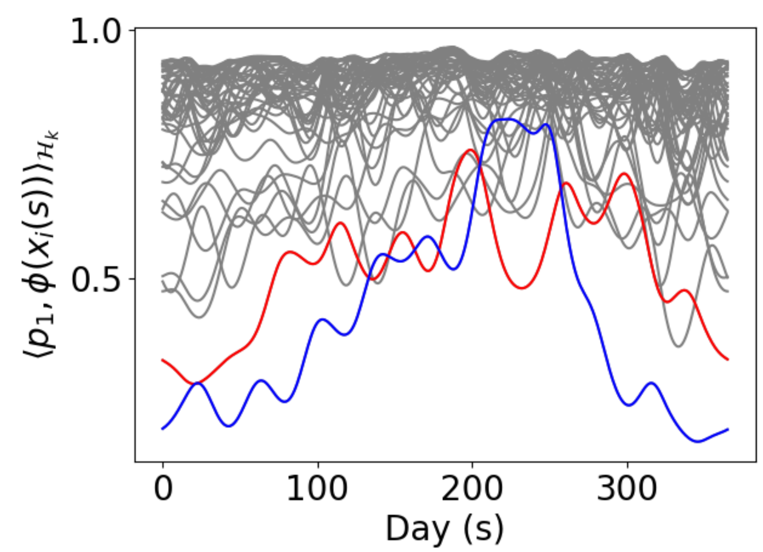
Experiments with climate data in Japan
(available at `https://www.data.jma.go.jp/gmd/risk/obsdl/`).





$\mathscr{X} = C([0,366], \mathbb{R}^2)$ $\quad k(x, y) = -\exp(-\|x - y\|^2)$ $\quad \mathscr{A} = L^\infty([0,366], \mathbb{R})$





Y. Hashimoto, I. Ishikawa, M. Ikeda, F. Komura, T. Katsura, Y. Kawahara,
Reproducing kernel hilbert C*-module and kernel mean embeddings,
In: Journal of Machine Learning Research. 2021 ; Vol. 22.

Other Collaborations

Sugiyama Team

Takeshi Teshima, Isao Ishikawa, Koichi Tojo, Kenta Oono, Masahiro Ikeda and Masashi Sugiyama, Coupling-based invertible neural networks are universal diffeomorphism approximator, Proc. of NeurIPS 2020.

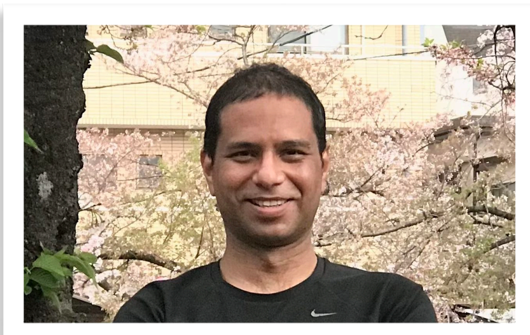
RAFCC (AIP-Fujitsu)

Kobayashi, K., Hamada, N., Sannai, A., Tanaka, A., Bannai, K., & Sugiyama, M. (2019). Bézier Simplex Fitting: Describing Pareto Fronts of Simplicial Problems with Small Samples in Multi-Objective Optimization. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*(01), 2304-2313.

Tanaka, A., Sannai, A., Kobayashi, K., & Hamada, N. (2020). Asymptotic Risk of Bézier Simplex Fitting. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(03), 2416-2424.

Many other collaborations with eg. Suzuki Team, Hatano Team, Takeda Team, etc.

The Bayes-Duality Project



Khan CREST

Emti Kahn
RIKEN AIP

CREST: Operator theoretic data analysis of complicated dynamics and its integrated utilization with mathematical models
FY2021—2026 (Oct. 2021—March 2027)

Other Research

See AIP Open Seminar Videos

18th AIP Open Seminar
https://aip.riken.jp/events/event_113730/

- **Koiichi Tojo:** A method to construct exponential families by representation theory
- **Tomotaka Kuwahara:** Information-theoretic structure of quantum Boltzmann distribution
- **Masahiro Ikeda:** Operator-theoretic approach for time-series data generated by nonlinear dynamical system
- **Akiyoshi Sannai:** Deep learning with symmetry

Thoughts

Mathematics vs Mathematical Science

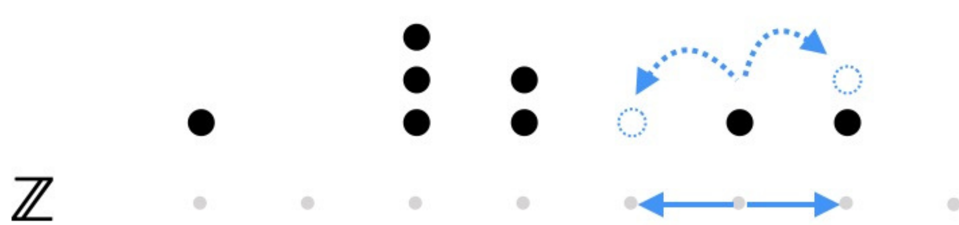
Pure Mathematics		Pure & Applied Mathematics
数学	→	数理科学
"Mathematic"	+Numerical Computation	
Axiomatic	+Algorithms	Reality
Rigor	+Applications	
	+Statistics	
Bourbaki	+Logic +Category Theory	

Feedback to Pure Mathematics

(with Makiko Sasada and Yukio Kametani)

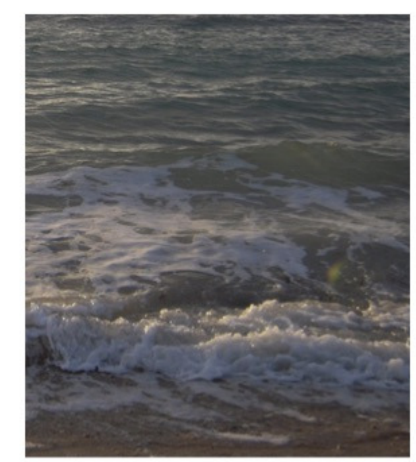
Hydrodynamic Limit

Derive deterministic macroscopic partial differential equations from stochastic microscopic dynamics



Microscopic

scaling limit



Macroscopic

Feedback to Pure Mathematics

(with Makiko Sasada and Yukio Kametani)

Results

- Proposed a general framework to describe the microscopic model (axiomatized the model)
- Divided the data of the model into the Geometric and Stochastic Data
- Interpreted the number of parameters of the deterministic PDE in terms of invariants of the Geometric Data

Group Cohomology
Cohomology of Graphs
Projective Systems and Systematic Use of Duality

Feedback to Pure Mathematics

- K. Bannai, Y. Kametani, and M. Sasada, Topological Structures of Large Scale Interacting Systems via Uniform Locality, arXiv:2009.04699 **[math.PR]**

- K. Bannai and M. Sasada, A Decomposition Theorem of Varadhan Type for Co-local Forms on Large Scale Interacting Systems, arXiv:2105.06043 **[math.PR]**

- K. Bannai and M. Sasada, A Decomposition Theorem of Varadhan Type for Co-local Forms on Large Scale Interacting Systems, arXiv:2111.08934 **[math.PR]**

Feedback to Pure Mathematics

Working with people of various discipline gives very unexpected results

Diversity, Equity and Inclusion is VERY IMPORTANT