

---

# Towards Outlier-Robust Statistical Inference on Kernel-Enriched Domains\*

---

Zoltán Szabó<sup>†</sup>

## Abstract

Mean embedding and maximum mean discrepancy (MMD) form the basis of modern statistical inference techniques on kernel-endowed domains with a wide variety of successful applications. In several examples the underlying kernel used is unbounded (examples include polynomial, exponential, string or graph kernels) in which case even a single outlier can completely ruin the available mean embedding and MMD estimators. In this talk, I will (i) focus on how to design systematically outlier-robust mean embedding and MMD estimators relying on the recently emerged principle of median-of-means, (ii) detail the consistency and excessive outlier-robustness of the constructed estimators, and (iii) illustrate their efficiency in discrimination of DNA subsequences.

Preprint: <https://arxiv.org/abs/1802.04784>

---

\*Workshop at RIKEN AIP, Tokyo, Japan, 15 April 2019; abstract.

<sup>†</sup>Joint work with Matthieu Lerasle<sup>1</sup>, Timothée Mathieu<sup>2</sup> and Guillaume Lecué<sup>3</sup>. <sup>1</sup>Laboratoire de Mathématiques d'Orsay, Univ. Paris-Sud; CNRS, Université Paris Saclay, France. <sup>2</sup>Laboratoire de Mathématiques d'Orsay, Univ. Paris-Sud, France. <sup>3</sup>CREST ENSAE ParisTech, France.