Automated Metric Discovery: Navigating Quantum Geometry with Symbolic Regression

Xianquan Yan[a]

[a] *Department of {Physics, Computer Science}, National University of Singapore, Singapore 117551* yanx@u.nus.edu

* Presenting author

1. Introduction and Motivation

Quantum geometry [1, 2], a branch of quantum mechanics exploring the distances and shapes formed by quantum states in the Hilbert space, plays a crucial role in quantum computing, quantum information, and condensed matter physics [3, 4].

Quantum geometric metrics allow us to quantify how "far apart" two quantum states are, enabling insights into quantum state distinguishability, evolution, and information processing. The Bures metric [5] for mixed states and the Fubini–Study metric [6, 7] for pure states are especially well documented, as they maximize the information content of the set of states. In computer science parlance, these metrics can be viewed as specialized kernels for quantum data analysis.

Despite the significance of quantum metrics, an automatic method to uncover new metrics directly from quantum states that optimizes an interested objective remains elusive.

To address this gap, we introduce a methodology leveraging symbolic regression [8, 9, 10, 11] to unsupervisedly discover such problem-specific quantum metrics. We show that, with a defined target objective/loss function, symbolic regression can identify analytic expressions of the optimal metrics, thereby opening a new way for exploring quantum geometry.

2. Methodology

Our approach to discovering metrics comprises the following steps:

1. **Quantum State Generation:** We generate datasets in various Hilbert space dimensions (starting from 2 upwards). For pure states, we create random complex state vectors. For mixed states, we create random Hermitian density matrices by normalizing random positive semi-definite matrices.

2. **Symbolic Regression with Objective:** We then apply symbolic regression to learn a function of a pair of quantum states that maximizes a chosen objective function.

3. **Trangular Inequality Validation:** We validate the learned metric by checking if it satisfies the triangle inequality and discard all expressions that fail this test.

4. **Scoring and Selection:** The symbolic expressions are scored based on both their loss and complexity, so that the best formula achieves the optimal trade-off between accuracy and simplicity.

This approach allows researchers to *rediscover* known quantum metrics or *discover* new ones customized for a specific application.

3. Results

In our experiments, symbolic regression accurately reproduces known metrics from simulated datasets.

We set the objective function to be the sum of mutual information over all pairs of states. For pure states, we aim to learn a function $d(\psi_1 - \phi_1, \psi_2 - \phi_2, \ldots, \psi_n - \phi_n)$ of a pair of state vector $|\psi\rangle$ and $|\phi\rangle$; for mixed states, we aim to learn a function $d(\rho_{11} - \sigma_{11}, \rho_{12} - \sigma_{12}, \ldots, \rho_{nn} - \sigma_{nn})$ of a pair of density matrices ρ and σ.

Table 1 illustrates the recovery of the Fubini–Study distance in pure states, while Table 2 shows the rediscovery of the Bures distance in mixed states.

Notice in each case the best symbolic expression matches the known analytic formula, thereby confirming the effectiveness of the methodology.

4. Conclusion and Outlook

We have demonstrated how symbolic regression can *automatically learn* quantum geometric metrics from numerical data, recovering known formulas (e.g. Fubini–Study and Bures) in a transparent, closed-form fashion. Beyond these canonical examples, the same approach opens the door to exploring *customized* quantum metrics that maximize a problem-specific objective.

Moving forward, one may: (1) Extend the methodology to more complex quantum states. (2) Develop loss functions tailored to practical quantum computing tasks. (3) Explore hybrid strategies that combine symbolic regression with other machine learning approaches.

We envision that combining symbolic regression with domain knowledge will spur new ways to interpret, design, and optimize quantum technologies.

Table 1: Symbolic regression results on pure states, maximizing mutual information. We show a snippet of the discovered candidate formulae. The best result recovers the known formula for the FS distance $d_{\mathrm{FS}} = \arccos(F)$ where $F = |\langle\psi|\phi\rangle|$ is the the overlap (fidelity).

Score	Equation	Complexity	Loss
0.000000	1.0130168	1	2.553290×10^{-1}
16.338898	$\boldsymbol{\arccos(F)}$	**2**	$\mathbf{2.047420 \times 10^{-8}}$
0.002554	$\arccos(F) - 2.2369306 \times 10^{-8}$	4	2.036989×10^{-8}
0.001832	$\sqrt{(\arccos(F) - 1.7406188 \times 10^{-9})^2}$	6	2.029538×10^{-8}
0.009961	$\arccos(F) - 1.8471347 \times 10^{-8} F^2$	7	2.009422×10^{-8}
0.003343	$\sqrt{(\arccos(F) + 5.4268083 \times 10^{-9})^2 + 7.167 \times 10^{-9}}$	8	2.002716×10^{-8}

Table 2: Symbolic regression results on mixed states, maximizing mutual information. We show a snippet of the discovered candidate formulae. The best result recovers the known formula for the Bures distance $d_B = \sqrt{2(1-F)}$ where $F(\rho, \sigma) = \mathrm{Tr}\big(\sqrt{\sqrt{\rho}\,\sigma\,\sqrt{\rho}}\big)$ is the Uhlmann fidelity.

Score	Equation	Complexity	Loss
0.000000	0.6523421	1	2.575764×10^{-2}
4.396659	$\arccos(F)$	2	3.172936×10^{-4}
0.989465	$0.97597754 \times \arccos(F)$	4	4.385534×10^{-5}
4.399256	$\sqrt{\frac{1.0000458-F}{0.5}}$	6	6.620290×10^{-9}
7.624512	$\boldsymbol{\sqrt{\frac{(\frac{F}{F}-F)}{0.5}}}$	**8**	$\mathbf{1.578737 \times 10^{-15}}$
0.002821	$\sqrt{\frac{(\frac{F}{F}-F)/0.27199987}{1.8382362}}$	10	1.569855×10^{-15}

References

[1] Ingemar Bengtsson and Karol Życzkowski. *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge university press, 2017.

[2] Dorje C Brody and Lane P Hughston. Geometric quantum mechanics. *Journal of geometry and physics*, 38(1):19–53, 2001.

[3] Michael A Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[4] Tianyu Liu, Xiao-Bin Qiang, Hai-Zhou Lu, and XC Xie. Quantum geometry in condensed matter. *National Science Review*, page nwae334, 2024.

[5] Donald Bures. An extension of kakutani's theorem on infinite product measures to the tensor product of semifinite w*-algebras. *Transactions of the American Mathematical Society*, 135:199–212, 1969.

[6] Guido Fubini. *Sulle metriche definite da una forma hermitiana: nota*. Office graf. C. Ferrari, 1904.

[7] Eduard Study. Kürzeste wege im komplexen gebiet. *Mathematische Annalen*, 60(3):321–378, 1905.

[8] Michael Schmidt and Hod Lipson. Distilling free-form natural laws from experimental data. *science*, 324(5923):81–85, 2009.

[9] Steven L Brunton, Joshua L Proctor, and J Nathan Kutz. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the national academy of sciences*, 113(15):3932–3937, 2016.

[10] Miles Cranmer, Alvaro Sanchez Gonzalez, Peter Battaglia, Rui Xu, Kyle Cranmer, David Spergel, and Shirley Ho. Discovering symbolic models from deep learning with inductive biases. *A ances in neural information processing systems*, 33:17429–17442, 2020.

[11] Miles Cranmer. Interpretable machine learning for science with pysr and symbolicregression.jl. *arXiv preprint arXiv:2305.01582*, 2023.