Quantum Information Elicitation

RAFAEL FRONGILLO, University of Colorado Boulder, USA

Information elicitation is the study of mechanisms which incentivize self-minded agents to reveal their private information. Perhaps the most fundamental scenario is the scoring rule setting, where a principal wishes to incentivize a single agent to reveal their private belief about the probability of a future outcome. Specifically, given a reported probability distribution $p \in \Delta_{\mathcal{Y}}$, and the realized outcome $y \in \mathcal{Y}$, the principal wishes to design a *proper* scoring rule $S(p, y)$, such that the agent will maximize their expected score by reporting their true belief. This basic setting is fundamental in statistics and machine learning [Gneiting 2011], and forms the basis of more complex mechanisms like prediction markets, wagering mechanisms, and peer prediction.

Since the turn of the century, however, another type of information has been increasingly relevant: quantum information. The elicitation of classical information is well-studied, but what of quantum information? That is, how could one design a scoring rule when the private information to be elicited is not a classical probability distribution, but a quantum mixed state? By analogy to the classical setting, this problem is fundamentally important to learning, inference, and economic mechanisms involving quantum information.

One interesting complication is that, from a physical standpoint, any sample "drawn" from a quantum mixed state must be mediated by a choice of measurement. Roughly speaking, the measurement procedure takes a particular projection of the density matrix onto a classical distribution over some set of outcomes, from which the actual outcome of the measurement is drawn. How then can the principal incentivize an agent to reveal their entire mixed state, when given only measurement access?

This paper introduces quantum information elicitation, presenting a model and several results to answer these and other questions. In the basic model, the principal chooses both a scoring function s and a measurement function μ, together called a *quantum score*, and the agent reports a density matrix ρ. The principal then measures the true state using $\mu(\rho)$, observes the outcome y, and awards the agent with score $s(\rho, y)$.

This work leverages the affine score framework of Frongillo and Kash [2021] to give several characterization results, all of which echo a familiar structure based on convex analysis. Many results concern particularly natural classes of measurements, such as a fixed measurements (that do not depend on ρ), or those that form an orthonormal basis, the most common type of measurement in the literature.

We find that fixed-measurement scores cover nearly all truthful quantum scores, but not all. In particular, fixed-measurement scores cannot capture the quantum analog of the log scoring rule, arguably the most natural quantum score. Another interesting class of quantum scores is spectral scores, wherein the measurement function is given by the spectral decomposition of the report ρ; we show that spectral scores have particularly elegant structure and connections to quantum information theory.

It is also natural to ask about eliciting *properties* (e.g., summary statistics) of a density matrix, such as its eigenvectors or eigenvalues. We find that one can elicit eigenvectors of a density matrix, but not its eigenvalues or entropy. In fact, the *elicitation complexity* of eigenvalues and entropy, a measure of how hard it is to elicit them, is as large as that of the full density matrix itself. We conclude with a discussion of other quantum information elicitation settings and connections to the literature.

A full version of this paper can be found at https://arxiv.org/abs/2203.07469.

ACM Reference Format:
Rafael Frongillo. 2023. Quantum Information Elicitation. In *Proceedings of the 24th ACM Conference on Economics and Computation (EC '23), July 9–12, 2023, London, United Kingdom.* ACM, New York, NY, USA, 1 page. https://doi.org/10.1145/3580507.3597703

EC '23, July 9–12, 2023, London, United Kingdom

ACM ISBN 979-8-4007-0104-7/23/07.
https://doi.org/10.1145/3580507.3597703