Iterative Algorithms for Quantum State Discrimination

Joint work with Peter Brown and Cambyse Rouzé

July 2, 2024 – QURIOSITY

Ali Almasi

Institut Polytechnique de Paris

ali.almasi@polytechnique.edu   ali-almasi.github.io

Introduction

Iterative Methods in Optimization

- Make a guess, and iterate on it until you converge!
- It is a quite old method:
  - Heron (60 AD) described an iterative method for finding the square root.
  - Iranian mathematician *Jamshid Kashi* (1380-1429) used an iterative method to compute $\sin 1^\circ$ to a high precision.
  - Newton used an iterative method for finding the root of a polynomial.
- Iterative methods are widely used today:
  - Gradient descent, hill climbing, Newton’s method, quasi-Newton methods, etc.
- They can be more efficient than the direct methods.

Quantum State Discrimination

> ***State Discrimination***
> Alice possesses an ensemble $\mathcal{E} = \{(\rho_j, p_j)\}_{j\in[N]}$. She picks a state $\rho_?$ according to the distribution $(p_j)_{j\in[N]}$ and sends it to Bob. Bob's task is to guess $i \in [N]$ such that $\rho_i = \rho_?$.

- If the states are not mutually orthogonal, then one can not *perfectly* discriminate.
- In the case of imperfect distinguishability, different strategies might be considered:
  - minimum-error, unambiguous, maximum-confidence, etc.
- The problem has applications in quantum information theory, quantum cryptography and quantum query complexity.

A Quick Recap on SDPs (Based on [SC23])

Semidefinite Programs

- *Semidefinite Programs* are generalizations of linear programs: constraint optimization problems
  - in Hermitian variables,
  - with a linear objective function $\mathrm{tr}(AX)$, for some Hermitian operator A,
  - and a number of linear equality and inequality constraints $\Phi_i(X) = B_i$ and $\Gamma_j(X) \leqslant C_j$, where Φ_i, Γ_j are linear hermiticity preserving maps, and B_i, C_j are Hermitian.

$$
\begin{array}{rll}
\multicolumn{3}{c}{\text{SDP}} \\
\hline
\text{maximize}: & \mathrm{tr}(AX) & \\
\text{subject to}: & \Phi_i(X) = B_i & i \in [m], \\
& \Gamma_j(X) \leqslant C_j & j \in [n].
\end{array}
$$

- SDP power: many problems can be cast as SDPs!

Duality

- We already introduced the *Primal* problem. The *Dual* is also an optimization problem providing an alternative formulation of the primal.
- Every feasible point of the Dual, provides an upper-bound on the optimal value of the primal.

Primal SDP	Dual SDP
maximize : $\mathrm{tr}(AX)$	minimize : $\sum_{i=1}^{m} \mathrm{tr}(Y_i B_i) + \sum_{j=1}^{n} \mathrm{tr}(Z_j C_j)$
subject to : $\Phi_i(X) = B_i \quad i \in [m],$	subject to : $A - \sum_{i=1}^{m} \Phi_i^*(Y_i) - \sum_{j=1}^{n} \Gamma_j^*(Z_j) = 0,$
$\Gamma_j(X) \leqslant C_j \quad j \in [n].$	$Z_j \geqslant 0 \quad j \in [n].$

- The optimal value of the dual is always an upper-bound of the optimal value of the primal (Weak Duality).

Duality

Primal SDP	
maximize :	$\mathrm{tr}(AX)$
subject to :	$\Phi_i(X) = B_i \qquad i \in [m],$
	$\Gamma_j(X) \leqslant C_j \qquad j \in [n].$

Dual SDP

$$\text{minimize}: \sum_{i=1}^{m} \mathrm{tr}(Y_i B_i) + \sum_{j=1}^{n} \mathrm{tr}(Z_j C_j)$$

$$\text{subject to}: A - \sum_{i=1}^{m} \Phi_i^*(Y_i) - \sum_{j=1}^{n} \Gamma_j^*(Z_j) = 0,$$

$$Z_j \geqslant 0 \quad j \in [n].$$

- Under some mild conditions the optimal value of the primal and the dual are equal (Strong Duality).
  - If the primal (dual) has finite optimal value and it is strictly feasible, then the strong duality holds.
- Strong duality holds for almost all the SDPs arising in QIT.
- When the strong duality holds, we have *Complementary Slackness*:

$$Z_j^* \left[C_j - \Gamma_j(X^*)\right] = 0 \quad \text{for } j = 1, \ldots, n$$

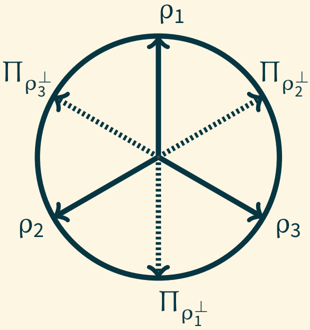
State Discrimination Revisited

State Antidiscrimination

State Antidiscrimination

Alice possesses an ensemble $\mathcal{E} = \{(\rho_j, p_j)\}_{j\in[N]}$. She picks a state $\rho_?$ according to the distribution $(p_j)_{j\in[N]}$ and sends it to Bob. Bob's task is to guess $i \in [N]$ such that $\rho_i \neq \rho_?$.

- Antidiscrimination is *weaker* than the discrimination.
- It turned out to be useful in proving ψ—ontology theorems [Lei14].



A More General Setting: Quantum Guessing Games

Quantum Guessing Game

Alice possesses an ensemble $\mathcal{E} = \{(\rho_j, p_j)\}_{j\in[N]}$, and picks a state $\rho_?$ according to the distribution $(p_j)_{j\in[N]}$ and sends it to Bob. If Bob guesses that $\rho_?$ is ρ_i and the true index of $\rho_?$ is j, then Bob receives the reward $f(i,j) \in \mathbb{R}$. Bob's task is to maximize his expected reward.

- The problem can be cast as a state discrimination [CHT22].
- Many problems can be seen as the special cases of the quantum guessing games [CHT22; MSU23]:
    - State Discrimination ($f(i,j) = \delta_{i,j}$), State Antidiscrimination ($f(i,j) = 1 - \delta_{i,j}$), Set Discrimination, etc.

Formulation of the Problem As a SDP

A quantum guessing game can be formulated as the following SDP:

Primal SDP

$$\text{maximize}: \ \mathcal{R}_M = \sum_{i=1}^{L} \sum_{j=1}^{N} f(i,j) \operatorname{tr}(M_i \widetilde{\rho}_j)$$

$$\text{subject to}: \ \sum_{k=1}^{N} M_k = \mathbb{1},$$

$$M_k \geqslant 0 \quad k \in [L],$$

where $\widetilde{\rho}_k \stackrel{\text{def}}{=\!=} p_k \rho_k$.

Formulation of the Problem As a SDP

In particular, the SDP formulation of the state discrimination is:

Primal SDP

$$\text{maximize}: \ \mathcal{P}_M = \sum_{i=1}^{N} \operatorname{tr}(M_i \widetilde{\rho}_i)$$

$$\text{subject to}: \ \sum_{k=1}^{N} M_k = \mathbb{1},$$

$$M_k \geqslant 0 \quad k \in [N].$$

Dual SDP

$$\text{minimize}: \ \operatorname{tr}(Y)$$

$$\text{subject to}: \ Y \geqslant \widetilde{\rho}_i \quad i \in [N]$$

- Primal and Dual are both strictly feasible (Strong Duality).

Strategies for Solving the State Discrimination Problem

There are typically two general strategies:

- Finding the optimal value:
  - Analytically: An analytical solution is only known for few cases:
    - 2-state ensemble: $\mathcal{P}_{\text{opt}} = \frac{1}{2} + \frac{1}{2}\|\widetilde{\rho}_1 - \widetilde{\rho}_2\|_1$
    - equiprobable qubits [DT10], geometrically uniform states [EMV04], mirror-symmetric states [And+02].
  - Numerically: Using the SDP solvers.
- Using sub-optimal measurements with an acceptable degree of quality:
  - Pretty good (bad, ugly) measurements:

$$M_i = \Sigma^{-1/2}\widetilde{\rho}_i\Sigma^{-1/2}, \qquad \text{where} \qquad \Sigma \stackrel{\text{def}}{=} \sum_{k=1}^{N}\widetilde{\rho}_k.$$

  - Belavkin Measurements: Having an array of weight matrices $\{w_k \in \text{Pos}(\mathbb{C}^{\text{rank}\,\rho_k}) \;:\; k \in [N]\}$ and writing $\rho_i = \psi_i\psi_i^\dagger$, define

$$M_i \stackrel{\text{def}}{=} \Sigma_w{}^{-\frac{1}{2}}\psi_i w_i \psi_i{}^\dagger \Sigma_w{}^{-\frac{1}{2}}, \qquad \text{where} \qquad \Sigma_w \stackrel{\text{def}}{=} \sum_{i=1}^{N}\psi_i w_i \psi_i{}^\dagger.$$

Optimality Conditions

Using the complementary slackness, we can obtain the following necessary and sufficient condition for an optimal measurement.

Theorem ([Hol73; YKL75])

A measurement $M = \{M_k\}_{k \in [N]}$ is optimal iff there exists an operator $G \in \mathrm{Pos}(\mathbb{C}^d)$ such that $GM_k = \widetilde{\rho}_k M_k$ and $G \geqslant \widetilde{\rho}_k$ for all $k \in [N]$.

There is also an optimality condition for the Belavkin measurement.

Theorem ([BM87])

A measurement $M = \{M_k\}_{k \in [N]}$ is optimal iff it is identical to a Belavkin measurement with weights $\{w_k\}_{k \in [N]}$ such that there exist a positive c satisfying $p_k Y_k w_k = c w_k$ and $p_k Y_k \leqslant c\mathbb{1}$, where $Y_k \overset{\text{def}}{=\!=} \psi_k^\dagger \Sigma_w^{-\frac{1}{2}} \psi_k$, for all $k \in [N]$.

Towards an Iterative Algorithm

An Algorithm by Ježek et al. [JŘF02]

Theorem

A measurement $M = \{M_k\}_{k\in[N]}$ is optimal iff there exists an operator $G \in \mathrm{Pos}(\mathbb{C}^d)$ such that $GM_k = \widetilde{\rho}_k M_k$ and $G \geqslant \widetilde{\rho}_k$ for all $k \in [N]$.

$$M_k = G^{-1}\widetilde{\rho}_k M_k \widetilde{\rho}_k G^{-1}$$

If we take G to be $\left(\sum_{i=1}^N \widetilde{\rho}_i M_i \widetilde{\rho}_i\right)^{1/2}$, we have

$$M_k = \left(\sum_{i=1}^N \widetilde{\rho}_i M_i \widetilde{\rho}_i\right)^{-1/2} \widetilde{\rho}_k M_k \widetilde{\rho}_k \left(\sum_{i=1}^N \widetilde{\rho}_i M_i \widetilde{\rho}_i\right)^{-1/2}.$$

JFR iteration

$$M_k^{(+)} \stackrel{\text{def}}{=\!=} \left(\sum_{i=1}^N \widetilde{\rho}_i M_i \widetilde{\rho}_i\right)^{-1/2} \widetilde{\rho}_k M_k \widetilde{\rho}_k \left(\sum_{i=1}^N \widetilde{\rho}_i M_i \widetilde{\rho}_i\right)^{-1/2}$$

An Algorithm by Ježek et al. [JŘF02]

JFR iteration

$$M_k^{(+)} \stackrel{\text{def}}{=\!=} \left(\sum_{i=1}^{N} \widetilde{\rho}_i M_i \widetilde{\rho}_i \right)^{-1/2} \widetilde{\rho}_k M_k \widetilde{\rho}_k \left(\sum_{i=1}^{N} \widetilde{\rho}_i M_i \widetilde{\rho}_i \right)^{-1/2}$$

They observe that:

> *"In the many tests we did a monotonic convergence to the true global maximum of the success rate always had been observed, though we have no analytic proof of this behavior in general."*

A Modification by Nakashira et al. [NKU15]

They proposed iterating on weights instead of measurements.

- Advantages: lower computational costs, accelerating the convergence speed

Theorem

A measurement $M = \{M_k\}_{k\in[N]}$ is optimal iff it is identical to a Belavkin measurement with weights $\{w_k\}_{k\in[N]}$ such that there exist a positive c satisfying $p_k Y_k w_k = c w_k$ and $p_k Y_k \leqslant c\mathbb{1}$, where $Y_k \overset{\text{def}}{=\!=} \psi_k^\dagger \Sigma_w^{-\frac{1}{2}} \psi_k$, for all $k \in [N]$.

NKU Iteration

$$w_k^{(+)} = p_k^2 Y_k w_k Y_k$$

A More General Framework: Directional Iterations [Tys10]

For an ensemble $\mathcal{E} = \{\widetilde{\rho}_k\}_{k\in[N]}$ define the semidefinite inner product $\langle\cdot,\cdot\rangle_{\mathcal{E}}$ on the space of $[\mathcal{L}(\mathbb{C}^d)]^N$ as

$$\langle E, F\rangle_{\mathcal{E}} \stackrel{\text{def}}{=\!=} \sum_{i=1}^{N} \operatorname{tr}(E_i^\dagger F_i \widetilde{\rho}_i),$$

for $E = \{E_k\}_{k\in[N]}$ and $F = \{F_k\}_{k\in[N]}$.

- For a measurement $M = \{M_k\}_{k\in[N]}$, where $M_k = E_k^\dagger E_k$,

$$\mathcal{P}_M = \|E\|_{\mathcal{E}}.$$

A More General Framework: Directional Iterations [Tys10]

$$\langle E, F \rangle_{\mathcal{E}} \overset{\text{def}}{=\!=} \sum_{i=1}^{N} \operatorname{tr}(E_i^\dagger F_i \widetilde{\rho}_i),$$

Maximal Seminorm Problem

Let V be a linear (real or complex) space equipped with a semidefinite inner product $\langle \cdot, \cdot \rangle : V \times V \to V$, and consider the seminorm induced by this semidefinite inner product on V. Let $S \subseteq V$. Find an element of S which is maximal with respect to this seminorm.

$$V_{\mathcal{E}} \overset{\text{def}}{=\!=} \left\{ E \in \mathcal{L}(\mathbb{C}^d)^L \mid \|E\|_{\mathcal{E}} < \infty \right\},$$

$$S_{\mathcal{E}} \overset{\text{def}}{=\!=} \{ E \in V_{\mathcal{E}} \mid \sum_{i=1}^{L} E_i^\dagger E_i = \mathbb{1}_d \}$$

$$\mathcal{P}_{\text{opt}} = \max_{s \in S_{\mathcal{E}}} \|s\|_{\mathcal{E}}^2$$

A More General Framework: Directional Iterations [Tys10]

Directional Iteration

A directional iterate of $v \in V$, is an element $v^{(+)} \in S$ such that

$$v^{(+)} = \arg\max_{s \in S} \operatorname{Re}\langle s, v \rangle.$$

We immediately conclude that

Theorem

$$\left\| v^{(+)} \right\|^2 \geqslant \|v\|^2 + \left\| v^{(+)} - v \right\|^2.$$

Tyson showed that the JFR iteration is a dirctional iteration.

The Convergence of NKU in the Case of Linearly Independent Pure States [NKU15]

- Note that from the previous theorem, it is implied that

$$S^{(r)} = \sum_{k=1}^{N} \operatorname{tr}\left[E_k^{(r+1)} - E_k^{(r)}\right]^{\dagger}\left[E_k^{(r+1)} - E_k^{(r)}\right]\widetilde{\rho}_k,$$

  converges to zero when $r \to \infty$.
- When we have pure states, weights are positive numbers, as well as Y_ks ($Y_k = \psi_k^{\dagger}\Sigma_w^{-\frac{1}{2}}\psi_k$). Thus, many things commute!
- One can use the linear independence to show that $p_k Y_k^{(r)}$ tends to 1 when $r \to \infty$.
- Because of this convergence, we have

$$\mathcal{P}_{M^{(r)}} \geqslant (1-\epsilon)^2\, \mathcal{P}_{\text{opt}}\,.$$

A Fixed-point Theorem

Let T be the set-valued directional iteration, and consider the sequence $(E^{(r)})_{r=1}^{\infty}$, where $E^{(r+1)} \in T(E^{(r)})$.

Proposition

Let $(E^{(r)})_{r=1}^{\infty}$ be a sequence obtained by consecutively applying the directional iteration to an arbitrary $E^{(0)} \in V_G$. Then, every limit-point of this sequence is a fixed point of T.

The proof is by using the continuity of the semidefinite inner-product and the induced seminorm.

Future Steps

Future Steps

- Our ultimate goal is to prove convergence in the most general case. A possible next step is to prove that a fixed-point is optimal using the previously mentioned optimality conditions.
- A more ambitious goal: Is there a fundamental unifying theory? There are other similar iterative algorithms whose convergence has not yet been shown.
- Comparing the time and space complexity of the iterative method with the complexity of solving the SDP.

Thank you all for your attention!