Computation of the Quantum Rate-Distortion Function

K. He[1] J. Saunderson[1] H. Fawzi[2]

[1]Department of Electrical and Computer System Engineering
Monash University

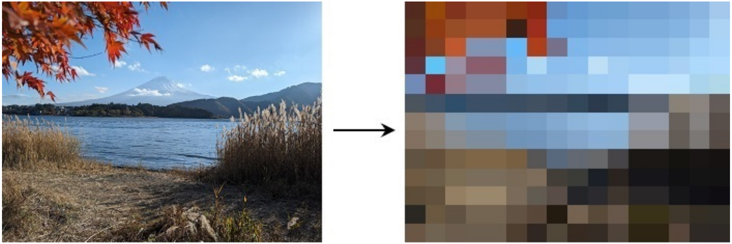
[2]Department of Applied Mathematics and Theoretical Physics
University of Cambridge

67th Annual Meeting of the Australian Mathematical Society,
December 2023

Introduction

Quantum rate-distortion problem: Consider the problem data

- Quantum signal $\rho \in \mathbb{H}^n_+$
- Distortion matrix $\Delta \in \mathbb{H}^{n^2}_+$
- Maximum allowable distortion $D \geq 0$



Input signal — Compressed signal

Introduction

Quantum rate-distortion problem: Consider the problem data

- Quantum signal $\rho \in \mathbb{H}^n_+$
- Distortion matrix $\Delta \in \mathbb{H}^{n^2}_+$
- Maximum allowable distortion $D \geq 0$

Maximum achievable compression of quantum signal is

$$\begin{aligned} \min_{\sigma \in \mathbb{H}^{n^2}_+} \quad & \operatorname{tr}[\sigma \log(\sigma)] - \operatorname{tr}[\mathcal{A}(\sigma) \log(\mathcal{A}(\sigma))] \\ \text{subj. to} \quad & \operatorname{tr}[\sigma \Delta] \leq D, \quad \mathcal{B}(\sigma) = \rho, \end{aligned}$$

for linear operators $\mathcal{A} : \mathbb{H}^{n^2} \to \mathbb{H}^n$ and $\mathcal{B} : \mathbb{H}^{n^2} \to \mathbb{H}^n$

Quantum entropy: If λ_i are the eigenvalues of σ, then

$$\operatorname{tr}[\sigma \log(\sigma)] = \sum_{i=1}^{n^2} \lambda_i \log(\lambda_i)$$

Introduction

Problem dimension can scale quickly

- Need to optimize over $\mathbb{H}^{n^2}$, i.e., optimization problem scales $O(n^4)$
- Is there structure in problem we can take advantage of?

Introduction

Problem dimension can scale quickly

- Need to optimize over $\mathbb{H}^{n^2}$, i.e., optimization problem scales $O(n^4)$
- Is there structure in problem we can take advantage of?
- Use **symmetry reduction**

Introduction

Problem dimension can scale quickly

- Need to optimize over $\mathbb{H}^{n^2}$, i.e., optimization problem scales $O(n^4)$
- Is there structure in problem we can take advantage of?
- Use **symmetry reduction**

Not many good algorithms available to compute quantum functions

- Approximate logarithms as linear matrix inequalities [Fawzi & Saunderson, 2023]
- Primal-dual interior point methods for general non-symmetric conic programs [Dahl & Andersen, 2022], [Coey et al., 2022], [Papp & Yildiz, 2022], [Karimi & Tuncel, 2020]

Introduction

Problem dimension can scale quickly

- Need to optimize over $\mathbb{H}^{n^2}$, i.e., optimization problem scales $O(n^4)$
- Is there structure in problem we can take advantage of?
- Use **symmetry reduction**

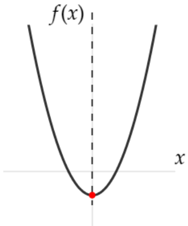
Not many good algorithms available to compute quantum functions

- Approximate logarithms as linear matrix inequalities [Fawzi & Saunderson, 2023]
- Primal-dual interior point methods for general non-symmetric conic programs [Dahl & Andersen, 2022], [Coey et al., 2022], [Papp & Yildiz, 2022], [Karimi & Tuncel, 2020]
- Use **mirror descent**

Table of Contents

Preliminaries

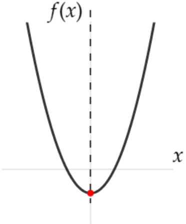
Many optimization problems possess symmetries



If problem is convex, symmetries inform us about problem solutions

Preliminaries

Many optimization problems possess symmetries



If problem is convex, symmetries inform us about problem solutions

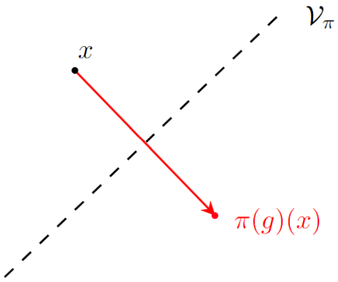
A **representation** of a group $\mathcal{G}$ is a pair $(\mathbb{V}, \pi)$, where

- $\mathbb{V}$ is a vector space
- $\pi : \mathcal{G} \rightarrow GL(\mathbb{V})$ is a group homomorphism

e.g., for group $\mathcal{G}$ consisting of matrices, let $(\mathbb{H}^n, \pi)$ be a congruence trans.

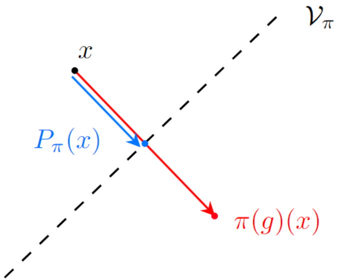
$$\pi(g)(X) = gXg^\dagger, \quad \forall g \in \mathcal{G}$$

Preliminaries



Fixed-point subspace $\mathcal{V}_\pi$: Set of all pnts. fixed under $\pi(g)$ for all $g \in \mathcal{G}$.

Preliminaries



Fixed-point subspace $\mathcal{V}_\pi$: Set of all pnts. fixed under $\pi(g)$ for all $g \in \mathcal{G}$.

Projection operator P_π: Linear proj. onto the fixed-point subspace $\mathcal{V}_\pi$. Given by the group average

$$P_\pi = \frac{1}{|\mathcal{G}|} \sum_{g \in \mathcal{G}} \pi(g).$$

Lemma 1

Consider a representation $(\mathbb{V}, \pi)$ of a group $\mathcal{G}$. If a convex optimization problem

$$\min_{x} f(x), \quad \text{subj. to } x \in \mathcal{X},$$

is invariant under π, meaning

$$\begin{aligned} & f(\pi(g)(x)) = f(x) && \forall g \in \mathcal{G},\ \forall x \in \mathcal{X} \\ \text{and} \quad & \pi(g)(x) \in \mathcal{X} && \forall g \in \mathcal{G},\ \forall x \in \mathcal{X}, \end{aligned}$$

then there is an optimal point for the optimization problem in $\mathcal{V}_\pi$.

Preliminaries

Lemma 1

Consider a representation $(\mathbb{V}, \pi)$ of a group $\mathcal{G}$. If a convex optimization problem

$$\min_{x} f(x), \quad \text{subj. to } x \in \mathcal{X},$$

is invariant under π, meaning

$$\begin{aligned} & f(\pi(g)(x)) = f(x) && \forall g \in \mathcal{G},\ \forall x \in \mathcal{X} \\ \text{and} \quad & \pi(g)(x) \in \mathcal{X} && \forall g \in \mathcal{G},\ \forall x \in \mathcal{X}, \end{aligned}$$

then there is an optimal point for the optimization problem in $\mathcal{V}_\pi$.

Proof:

$$f^* \leq f\left(\frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}} \pi(g)(x^*)\right) \leq \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}} f(\pi(g)(x^*)) = \frac{1}{|\mathcal{G}|}\sum_{g\in\mathcal{G}} f^* = f^*.$$

Quantum rate-distortion

Typically use the **entanglement fidelity** distortion matrix

$$\Delta = \sum_{ij}^{n} \sqrt{\lambda_i \lambda_j}\, v_i v_j^\dagger \otimes v_i v_j^\dagger, \quad \text{where} \quad \rho = \sum_{i=1}^{n} \lambda_i v_i v_i^\dagger.$$

Quantum rate-distortion

Typically use the **entanglement fidelity** distortion matrix

$$\Delta = \sum_{ij}^{n} \sqrt{\lambda_i \lambda_j}\, v_i v_j^\dagger \otimes v_i v_j^\dagger, \quad \text{where} \quad \rho = \sum_{i=1}^{n} \lambda_i v_i v_i^\dagger.$$

Theorem 2 (HSF, 2023)

Consider the group

$$\mathcal{G}_{ea} = \left\{ \sum_{i=1}^{n} z_i v_i v_i^\dagger : z \in \{\pm 1, \pm\sqrt{-1}\}^n \right\}$$

and corresponding representation $(\mathbb{H}^{n^2}, \pi_{cc})$ where

$$\pi_{cc}(g)(X) = (g \otimes \bar{g}) X (g \otimes \bar{g})^\dagger.$$

The quantum rate-distortion problem is invariant under this representation.

Quantum rate-distortion

Typically use the **entanglement fidelity** distortion matrix

$$\Delta = \sum_{ij}^{n} \sqrt{\lambda_i \lambda_j}\, v_i v_j^\dagger \otimes v_i v_j^\dagger, \quad \text{where} \quad \rho = \sum_{i=1}^{n} \lambda_i v_i v_i^\dagger.$$

Corollary 2 (HSF, 2023)

A solution to the quantum rate-distortion problem is in

$$\mathcal{V}_{ea} = \left\{ \sum_{i \neq j}^{n} \alpha_{ij} v_i v_i^\dagger \otimes v_j v_j^\dagger + \sum_{ij}^{n} \beta_{ij} v_i v_j^\dagger \otimes v_i v_j^\dagger : \alpha_{ij} \in \mathbb{R}\ \forall i \neq j,\ \beta \in \mathbb{H}^n \right\}.$$

Quantum rate-distortion

Typically use the **entanglement fidelity** distortion matrix

$$\Delta = \sum_{ij}^{n} \sqrt{\lambda_i \lambda_j}\, v_i v_j^\dagger \otimes v_i v_j^\dagger, \quad \text{where} \quad \rho = \sum_{i=1}^{n} \lambda_i v_i v_i^\dagger .$$

Corollary 2 (HSF, 2023)

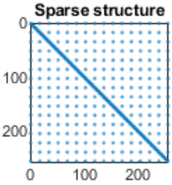
A solution to the quantum rate-distortion problem is in

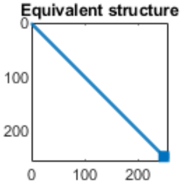
$$\mathcal{V}_{ea} = \left\{ \sum_{i \neq j}^{n} \alpha_{ij} v_i v_i^\dagger \otimes v_j v_j^\dagger + \sum_{ij}^{n} \beta_{ij} v_i v_j^\dagger \otimes v_i v_j^\dagger : \alpha_{ij} \in \mathbb{R}\ \forall i \neq j,\ \beta \in \mathbb{H}^n \right\}.$$

This subspace $\mathcal{V}_{ea} \subset \mathbb{H}^{n^2}$ has a real dimension of $2n^2 - n$

Quantum rate-distortion

Visualizing sparsity structure when v_i is the standard basis and $n = 16$:





Isomorphic to

- $n^2 - n$ blocks of size 1×1,
- ones block of size $n \times n$.

Easy to take eigendecomposition, quantum entropies, etc.

Table of Contents

Mirror descent

Consider constrained convex optimization problem

$$\min_{x \in \mathcal{X}} \quad f(x).$$

Projected gradient descent can be represented as

$$x^{k+1} = \arg\min_{x \in \mathcal{X}} \langle \nabla f(x^k), x \rangle + \frac{1}{2t_k} \|x - x^k\|_2^2$$

Mirror descent

Consider constrained convex optimization problem

$$\min_{x \in \mathcal{X}} \quad f(x).$$

Projected gradient descent can be represented as

$$x^{k+1} = \arg\min_{x \in \mathcal{X}} \langle \nabla f(x^k), x \rangle + \frac{1}{2t_k} \| x - x^k \|_2^2$$

Mirror descent replaces Euclidean norm with Bregman divergence

$$x^{k+1} = \arg\min_{x \in \mathcal{X}} \langle \nabla f(x^k), x \rangle + \frac{1}{t_k} D_\varphi(x \| y)$$

where

$$D_\varphi(x \| y) \coloneqq \varphi(x) - (\varphi(y) + \langle \nabla \varphi(y), x - y \rangle).$$

Mirror descent – convergence

A function f is L-smooth relative to φ if for $L > 0$

$$L\varphi - f \quad \text{convex}$$

Mirror descent w/ $t_k = 1/L$ converges sublinearly if f is L-smooth rel. to φ

Mirror descent – convergence

A function f is L-smooth relative to φ if for $L > 0$

$$L\varphi - f \quad \text{convex}$$

Mirror descent w/ $t_k = 1/L$ converges sublinearly if f is L-smooth rel. to φ

Theorem 3 (HSF, 2023)

The objective function of the quantum rate-distortion problem is 1-smooth relative to $\varphi(x) = \text{tr}[x \log(x)]$.

Therefore, mirror descent applied to QRD problem with unit step size and $\varphi(x) = \text{tr}[x \log(x)]$ will converge sublinearly to global optimum.

Mirror descent – convergence

A function f is L-smooth relative to φ if for $L > 0$

$$L\varphi - f \quad \text{convex}$$

Mirror descent w/ $t_k = 1/L$ converges sublinearly if f is L-smooth rel. to φ

Theorem 3 (HSF, 2023)

The objective function of the quantum rate-distortion problem is 1-smooth relative to $\varphi(x) = \operatorname{tr}[x \log(x)]$.

Therefore, mirror descent applied to QRD problem with unit step size and $\varphi(x) = \operatorname{tr}[x \log(x)]$ will converge sublinearly to global optimum.

Caveat:

- Each iteration requires solving a convex subproblem
- Can do efficiently by solving the dual problem inexactly (while retaining convergence guarantees!)

Numerical experiments

(a) Without symmetry reduction

			Ours		CvxQuad	
n	#variables	D	Time (s)	Gap	Time (s)	Gap
8	4×10^3	0.8	4.88	$2e{-}8$	Out of memory	
		0.5	1.12	$1e{-}8$	Out of memory	
32	1×10^6	0.8	> 3600.00	$2e{-}7$	Out of memory	
		0.6	1321.85	$1e{-}7$	Out of memory	
512	7×10^{10}	0.9	Out of memory		Out of memory	
		0.6	Out of memory		Out of memory	

(b) With symmetry reduction

			Ours		CvxQuad	
n	#variables	D	Time (s)	Gap	Time (s)	Gap
8	1×10^2	0.8	.17	$7e{-}9$	454.98	$2e{-}8$
		0.5	.07	$4e{-}9$	686.11	$6e{-}8$
32	2×10^3	0.8	1.52	$6e{-}8$	Out of memory	
		0.6	.32	$5e{-}9$	Out of memory	
512	5×10^5	0.9	2174.38	$7e{-}8$	Out of memory	
		0.6	1216.96	$5e{-}9$	Out of memory	

Conclusion

Summary:

- Rate-distortion problems possess symmetries that can be exploited to significantly reduce dimensionality of the optimization problem.
- Mirror descent algorithm can efficiently solve the problem

Conclusion

Summary:

- Rate-distortion problems possess symmetries that can be exploited to significantly reduce dimensionality of the optimization problem.
- Mirror descent algorithm can efficiently solve the problem

Outlook:

- Other problems in quantum inf. theory w/ symmetries we can exploit?
- Study how to solve mirror descent subproblems in more detail

Conclusion

Summary:

- Rate-distortion problems possess symmetries that can be exploited to significantly reduce dimensionality of the optimization problem.
- Mirror descent algorithm can efficiently solve the problem

Outlook:

- Other problems in quantum inf. theory w/ symmetries we can exploit?
- Study how to solve mirror descent subproblems in more detail

Paper: `https://arxiv.org/abs/2309.15919`
Code: `https://github.com/kerry-he/efficient-qrd`