Robust Interior Point Method for Quantum Key Distribution Rate Computation

Hao Hu, Jiyoung Im, Jie Lin, Norbert Lütkenhaus, Henry Wolkowicz

Friday Seminar

Li Chen

August 20, 2021

Outline

Problem Description

Regularization by Facial Reduction

Interior Point Method using Gauss-Newton Direction

Numerical Results

Quantum Key Distribution (QKD) Problem

We focus on the convex nonlinear SDP problem:

$$\begin{array}{cl} \min\limits_{\rho \in \mathbb{C}^{n\times n}} & D(\mathcal{G}(\rho) || \mathcal{Z}(\mathcal{G}(\rho))) \\ \text{s.t.} & \Gamma(\rho) = \gamma \\ & \rho \succeq 0 \end{array} \tag{1}$$

- $\mathbb{C}^{n\times n}$ is a real Hilbert space with $\langle Y, X \rangle = \Re\left(\mathrm{Tr}\left(Y^* X\right)\right)$
- $\mathbb{H}^n$ is the Hermitian matrices: n^2-dim subspace of $2n^2$-dim $\mathbb{C}^{n\times n}$
- $\rho \succeq 0$ means $\rho \in \mathbb{H}^n_+$

Quantum Key Distribution (QKD) Problem

We focus on the convex nonlinear SDP problem:

$$\begin{array}{ll} \min\limits_{\rho \in \mathbb{C}^{n\times n}} & D(\mathcal{G}(\rho)||\mathcal{Z}(\mathcal{G}(\rho))) \\ \text{s.t.} & \Gamma(\rho) = \gamma \\ & \rho \succeq 0 \end{array} \tag{1}$$

- $\mathbb{C}^{n\times n}$ is a real Hilbert space with $\langle Y, X\rangle = \Re\left(\mathrm{Tr}\left(Y^* X\right)\right)$
- $\mathbb{H}^n$ is the Hermitian matrices: n^2-dim subspace of $2n^2$-dim $\mathbb{C}^{n\times n}$
- $\rho \succeq 0$ means $\rho \in \mathbb{H}^n_+$
- $D(\delta||\sigma) = \mathrm{Tr}\left(\delta(\log \delta - \log \sigma)\right)$ is the von-Neumann relative entropy
- The maps $\mathcal{G}$ and $\mathcal{Z}$ are linear, completely positive
- $\Gamma : \mathbb{H}^n \to \mathbb{R}^m$ is a linear map: $\Gamma(\rho)_i := \langle \Gamma_i, \rho\rangle$ where $\Gamma_i \in \mathbb{H}^n$

Objective Function

- von-Neumann relative entropy:

$$D(\delta||\sigma) = \begin{cases} \mathrm{Tr}\left(\delta(\log\delta - \log\sigma)\right) & \text{if range}\,(\delta) \subseteq \text{range}\,(\sigma) \\ +\infty & \text{otherwise} \end{cases}$$

  with convention $0 \cdot \log 0 = 0$.
  - Jointly convex in δ and σ
  - Nonegative and equals 0 only when $\delta = \sigma$
- Linear map $\mathcal{G} : \mathbb{H}^n \to \mathbb{H}^k$ with $k > n$ is

$$\mathcal{G}(\rho) := \sum_{j=1}^{\ell} K_j \rho K_j^*$$

  where $K_j \in \mathbb{C}^{k \times n}$ and $\sum_{j=1}^{\ell} K_j^* K_j \preceq I$.
  - $\mathcal{G}(\rho)$ is singular for any $\rho \succ 0$
- Linear map $\mathcal{Z} : \mathbb{H}^k \to \mathbb{H}^k$ is

$$\mathcal{Z}(\delta) := \sum_{j=1}^{N} Z_j \delta Z_j^*$$

  where $Z_j = Z_j^2 = Z_j^* \in \mathbb{H}_+^k$ and $\sum_{j=1}^{\ell} Z_j = I$.

Objective Function

Proposition 1

Let $X \succeq 0$, *then* $\operatorname{range}(X) \subseteq \operatorname{range}(\mathcal{Z}(X))$.

Implication:

$$\rho \succeq 0 \Longrightarrow \mathcal{G}(\rho) \succeq 0 \Longrightarrow \mathcal{Z}(\mathcal{G}(\rho)) \succeq 0$$

and

$$\operatorname{range}(\mathcal{G}(\rho)) \subseteq \operatorname{range}(\mathcal{Z}(\mathcal{G}(\rho)))$$

so $D(\mathcal{G}(\rho)||\mathcal{Z}(\mathcal{G}(\rho)))$ is well-defined.

Proposition 2

The linear map $\mathcal{Z}$ *is an orthogonal projection on* $\mathbb{H}_+^k$. *Moreover,*

$$\operatorname{Tr}(\delta) \leq 1, \delta \succ 0 \Longrightarrow \operatorname{Tr}(\delta \log \mathcal{Z}(\delta)) = \operatorname{Tr}(\mathcal{Z}(\delta) \log \mathcal{Z}(\delta))$$

It is useful to simplify computation later.

Constraints

$\Gamma(\rho) = \gamma$ has two parts:

- Observational constraints:

$$S_O := \left\{\rho \succeq 0 : \langle P_s^A \otimes P_t^B, \rho \rangle = p_{st}, \forall s, t\right\}$$

  where $P_s^A \in \mathbb{H}^{n_A}$ and $P_t^B \in \mathbb{H}^{n_B}$, and $\otimes$ is Kronecker product. Note $n = n_A n_B$ is the order of ρ.

- Reduced density operator constraints:

$$\begin{aligned} S_R := & \quad \{\rho \succeq 0 : \mathrm{Tr}_B(\rho) = \rho_A\} \\ = & \quad \{\rho \succeq 0 : \langle \Theta_j \otimes I_{n_B}, \rho \rangle = \langle \Theta_j, \rho_A \rangle, \forall j = 1, ..., m_R\} \end{aligned}$$

  where $\{\Theta_j\}$ forms a orthonormal basis of real vector space of Hermitian matrices on system A. Note $\mathrm{Tr}_B^*(Y) = Y \otimes I_B$.

- Add density constraint:

$$\mathrm{Tr}\,(\rho) = 1$$

  into S_R

Idea of The Paper

- Goal: obtain a reliable lower bound of

$$\begin{array}{ll} \min_{\rho,\delta,\sigma} & D(\delta||\sigma) \\ \text{s.t.} & \Gamma(\rho) = \gamma \\ & \delta = \mathcal{G}(\rho) \\ & \sigma = \mathcal{Z}(\delta) \\ & \rho \succeq 0 \end{array} \tag{2}$$

- Challenge: degeneracy of the constraints
  - Degeneracy from linear constraints
  - Degeneracy from von-Neumann entropy
- Solution:
  - regularize the problem by Facial Reduction (FR) first
  - solve the regularized problem by interior point method

Outline

Facial Reduction Basics

Definition 1 (Faces of convex cones)

A convex cone F is a face of a convex cone K, denoted as $F \trianglelefteq K$, if

$$x, y \in K, x + y \in F \Longrightarrow x, y \in F$$

Faces of PSD cones are characterized by the range or nullspace of any element in the relative interior of its faces.

Lemma 1 (Characterization of faces of PSD cone)

Let F be a convex subset of $\mathbb{H}^n_+$ with $X \in \mathrm{ri}(F)$. Let $X = [P\ Q]\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}[P\ Q]^*$ be the orthogonal spectral decomposition with $D \in \mathbb{H}^r_{++}$. Then the following are equivalent:

1. $F \trianglelefteq \mathbb{H}^n_+$
2. $F = \{Y \succeq 0 : \mathrm{range}\,(Y) \subseteq \mathrm{range}\,(X)\} = \{Y \succeq 0 : \mathrm{null}\,(X) \subseteq \mathrm{null}\,(Y)\}$
3. $F = P\mathbb{H}^r_+ P^*$
4. $F = \mathbb{H}^n_+ \cap (QQ^*)^\perp$ *(QQ^* is called an exposing vector for face F)*

Facial Reduction Basics (cont.)

Definition 2 (Minimal Faces)

Let K be a closed convex cone and $X \in K$, then $\mathrm{face}\,(X) \trianglelefteq K$ is the minimal face, the intersection of all faces of K containing X.

Facial reduction is the procedure of finding minimal face containing the feasible region, which can be done via the following lemma.

Lemma 2 (Theorem of Alternative)

Suppose $\{\rho \succeq 0 : \Gamma(\rho) = \gamma\}$ is feasible, then exactly one of the following is true:

1. *there exists $\rho \succ 0$ such that $\Gamma(\rho) = \gamma$*
2. *there exists y such that $0 \neq \Gamma^*(y) \succeq 0$, $\langle \gamma, y \rangle = 0$.*

Note if 2. holds, then $\Gamma^*(y)$ is the exposing vector of the face, i.e. $\{\rho \succeq 0 : \Gamma(\rho) = \gamma\} \subseteq \Gamma^*(y)^\perp$.

Degeneracy from Linear Map

If ρ_A is singular, then there is no strictly feasible solution of

$$\begin{aligned} S_R = & \quad \{\rho \succeq 0 : \mathrm{Tr}_B(\rho) = \rho_A\} \\ = & \quad \{\rho \succeq 0 : \langle \Theta_j \otimes I_{n_B}, \rho \rangle = \langle \Theta_j, \rho_A \rangle, \forall j = 1, ..., m_R\} \end{aligned}$$

Theorem 1

Let $\mathrm{range}\,(P) = \mathrm{range}\,(\rho_A) \subset \mathbb{H}^{n_A}$, $P^*P = I_r$, *let* $V = P \otimes I_B$. *Then*

$$\rho \in S_R \Longrightarrow \rho = VRV^*, \text{ for some } R \in \mathbb{H}_+^{r \cdot n_B}.$$

Proof:
Let $[P\ Q]$ be a unitary matrix such that $\mathrm{range}\,(P) = \mathrm{range}\,(\rho_A)$ and $\mathrm{range}\,(Q) = \mathrm{null}\,(\rho_A)$ and let $W = QQ^* \succeq 0$, then for any $\rho \in S_R$,

$$\langle W \otimes I_{n_B}, \rho \rangle = \langle \mathrm{Tr}_B^*(W), \rho \rangle = \langle W, \mathrm{Tr}_B(\rho) \rangle = \langle W, \rho_A \rangle = 0$$

So $\mathrm{Tr}_B^*(W) \succeq 0$, $\langle W, \rho_A \rangle = 0$. By Lemma 2 S_R is not strictly feasible.
Note $S_R \subseteq \mathrm{Tr}_B^*(W)^\perp$, by Lemma 1, $S_R \subseteq V \mathbb{H}_+^{r \cdot n_B} V^*$.

Degeneracy from von-Neumann Entropy

Recall the objective function is

$$f(\rho) := D(\mathcal{G}(\rho)||\mathcal{Z}(\mathcal{G}(\rho))) = \mathrm{Tr}\left(\mathcal{G}(\rho)\log \mathcal{G}(\rho)\right) - \mathrm{Tr}\left(\mathcal{G}(\rho)\log \mathcal{Z}(\mathcal{G}(\rho))\right)$$

Although $f(\rho)$ is well-defined for any $\rho \succ 0$, its gradient is not.

Proposition 3

Let $\mathcal{A} : \mathbb{H}_+^n \to \mathbb{H}_+^k$ be a linear map preserving positive semidefiniteness. Assume $\mathcal{A}(\rho) \succ 0$, then the gradient of $g(\rho) = \mathrm{Tr}\left(\mathcal{A}(\rho)\log \mathcal{A}(\rho)\right)$ is

$$\nabla g(\rho) = \mathcal{A}^*(I) + \mathcal{A}^*(\log(\mathcal{A}(\rho)))$$

and its Hessian at ρ acting on $\Delta\rho$ is

$$\nabla^2 g(\rho)[\Delta\rho] = \mathcal{A}^*(\log' \mathcal{A}(\rho)[\mathcal{A}(\Delta\rho)])$$

where $\log'$ is the Fréchet derivative.

However, $\mathcal{G}(\rho)$ is always rank-deficient.

FR for Linear Maps $\mathcal{G}$, $\mathcal{Z}$

Lemma 3

Let $\mathcal{C} \subseteq \mathbb{H}^n_+$ be a given closed convex set with nonempty interior. Let $Q_i \in \mathbb{H}^{k\times n}$, $i = 1, ..., t$, be given matrices. Define the linear map $\mathcal{A}$: $\mathbb{H}^n \to \mathbb{H}^k$ and V by

$$\mathcal{A}(X) = \sum_{i=1}^{t} Q_i X Q_i^*, \ \mathrm{range}\,(V) = \mathrm{range}\left(\sum_{i=1}^{t} Q_i Q_i^*\right),$$

then the minimal face, $\mathrm{face}\,(\mathcal{A}(\mathcal{C})) = V\mathbb{H}^r_+ V^*$.

FR for Linear Maps $\mathcal{G}$, $\mathcal{Z}$

Lemma 3

Let $\mathcal{C} \subseteq \mathbb{H}^n_+$ be a given closed convex set with nonempty interior. Let $Q_i \in \mathbb{H}^{k\times n}$, $i = 1, \ldots, t$, be given matrices. Define the linear map $\mathcal{A}$: $\mathbb{H}^n \to \mathbb{H}^k$ and V by

$$\mathcal{A}(X) = \sum_{i=1}^{t} Q_i X Q_i^*, \ \mathrm{range}\,(V) = \mathrm{range}\left(\sum_{i=1}^{t} Q_i Q_i^*\right),$$

then the minimal face, $\mathrm{face}\,(\mathcal{A}(\mathcal{C})) = V\mathbb{H}^r_+ V^*$.

Proof.

$$\begin{aligned}
0 \neq W \succeq 0, \langle W, \mathcal{A}(\mathcal{C})\rangle = 0 \quad &\iff 0 \neq W \succeq 0, \langle W, Y\rangle = 0, \forall Y \in \mathcal{A}(\mathcal{C}) \\
&\iff 0 \neq W \succeq 0, \langle \mathcal{A}^*(W), X\rangle = 0, \forall X \in \mathcal{C} \\
&\iff 0 \neq W \succeq 0, W \in \mathrm{null}\,(\mathcal{A}^*) \\
&\iff 0 \neq W \succeq 0, Q_i^* W Q_i = 0, \forall i \\
&\iff 0 \neq \mathrm{range}\,(W) \subseteq \mathrm{null}\left(\textstyle\sum_{i=1}^t Q_i Q_i^*\right)
\end{aligned}$$

So $\mathrm{face}\,(\mathcal{A}(\mathcal{C})) = \{Y \succeq 0 : \mathrm{range}\,(Y) \subseteq \mathrm{range}\left(\sum_{i=1}^t Q_i Q_i^*\right)$. Then use lemma 1.

Summary of FR Regularization

1. Apply FR to $\{\rho \succeq 0 : \Gamma(\rho) = \gamma\}$ to find its minimal face in $\mathbb{H}^n_+$ represented by

   $$\rho = V_\rho R_\rho V_\rho^* \in \mathbb{H}^n_+, \ R_\rho \in \mathbb{H}^{n_\rho}_+$$

   Then $\mathcal{R}_\rho := \{R_\rho \in \mathbb{H}^{n_\rho}_+ : \Gamma_V(R_\rho) = \gamma\}$ is strictly feasible.
   Let $\Gamma_V(R_\rho) := \Gamma(V_\rho R_\rho V_\rho^*)$, $\mathcal{G}_V(R_\rho) := \mathcal{G}(V_\rho R_\rho V_\rho^*)$.
2. Apply FR to $\{\mathcal{G}_V(R_\rho) \in \mathbb{H}^k_+ : R_\rho \in \mathcal{R}_\rho\}$ by choosing V_δ such that

   $$\operatorname{range}(V_\delta) = \operatorname{range}(\mathcal{G}_V(I))$$

   Then $\mathcal{R}_\delta := \{R_\delta \in \mathbb{H}^{k_\delta}_+ : \mathcal{G}_V(R_\rho) = V_\delta R_\delta V_\delta^*\}$ is strictly feasible.
   Let $\mathcal{Z}_V(R_\delta) := \mathcal{Z}(V_\delta R_\delta V_\delta^*)$.
3. Apply FR to $\{\mathcal{Z}_V(R_\delta) \in \mathbb{H}^k_+ : R_\delta \in \mathcal{R}_\delta\}$ by choosing V_σ such that

   $$\operatorname{range}(V_\sigma) = \operatorname{range}(\mathcal{Z}_V(I))$$

   Then $\mathcal{R}_\sigma := \{R_\sigma \in \mathbb{H}^{k_\sigma}_+ : \mathcal{Z}_V(R_\delta) = V_\sigma R_\sigma V_\sigma^*\}$ is strictly feasible.

W.l.o.g, we assume $V_M^* V_M = I$ for $M = \rho, \delta, \sigma$.

Regularized Problem

Let $\mathcal{V}_\delta(R_\delta) := V_\delta R_\delta V_\delta^*$ and $\mathcal{V}_\sigma(R_\sigma) := V_\sigma R_\sigma V_\sigma^*$, the constraints of QKD transforms

$$
\begin{array}{l}
\Gamma(\rho) = \gamma \\
\delta = \mathcal{G}(\rho) \\
\sigma = \mathcal{Z}(\delta) \\
\rho \succeq 0, \delta \succeq 0, \sigma \succeq 0
\end{array}
\quad \iff \quad
\begin{array}{l}
\Gamma_V(R_\rho) = \gamma \\
\mathcal{V}_\delta(R_\delta) = \mathcal{G}_V(R_\rho) \\
\mathcal{V}_\sigma(R_\sigma) = \mathcal{Z}_V(R_\delta) \\
R_\rho \succeq 0, R_\delta \succeq 0, R_\sigma \succeq 0
\end{array}
$$

Regularized Problem

Let $\mathcal{V}_\delta(R_\delta) := V_\delta R_\delta V_\delta^*$ and $\mathcal{V}_\sigma(R_\sigma) := V_\sigma R_\sigma V_\sigma^*$, the constraints of QKD transforms

$$\begin{array}{l} \Gamma(\rho) = \gamma \\ \delta = \mathcal{G}(\rho) \\ \sigma = \mathcal{Z}(\delta) \\ \rho \succeq 0, \delta \succeq 0, \sigma \succeq 0 \end{array} \quad \iff \quad \begin{array}{l} \Gamma_V(R_\rho) = \gamma \\ \mathcal{V}_\delta(R_\delta) = \mathcal{G}_V(R_\rho) \\ \mathcal{V}_\sigma(R_\sigma) = \mathcal{Z}_V(R_\delta) \\ R_\rho \succeq 0, R_\delta \succeq 0, R_\sigma \succeq 0 \end{array}$$

Remove the redundant constraints by defining $\mathcal{G}_{UV}(\cdot) := V_\delta^* \mathcal{G}_V(\cdot) V_\delta$ and $\mathcal{Z}_{UV}(\cdot) := V_\sigma^* \mathcal{Z}_V(\cdot) V_\sigma$:

$$\begin{array}{l} \Gamma_V(R_\rho) = \gamma \\ \mathcal{V}_\delta(R_\delta) = \mathcal{G}_V(R_\rho) \\ \mathcal{V}_\sigma(R_\sigma) = \mathcal{Z}_V(R_\delta) \\ R_\rho \succeq 0, R_\delta \succeq 0, R_\sigma \succeq 0 \end{array} \quad \iff \quad \begin{array}{l} \Gamma_V(R_\rho) = \gamma \\ R_\delta = \mathcal{G}_{UV}(R_\rho) \\ R_\sigma = \mathcal{Z}_{UV}(R_\delta) \\ R_\rho \succeq 0, R_\delta \succeq 0, R_\sigma \succeq 0 \end{array}$$

Regularized Problem

The objective transforms as :

$$\begin{aligned} D(\delta||\sigma) \quad &= \mathrm{Tr}\left(\delta \log \delta - \delta \log \sigma\right) \\ \text{(Proposition 2)} \quad &= \mathrm{Tr}\left(\delta \log \delta - \sigma \log \sigma\right) \\ &= \mathrm{Tr}\left(R_\delta \log R_\delta - R_\sigma \log R_\sigma\right) \end{aligned}$$

Lemma 4

Let $Y = VRV^* \succeq 0$, $R \succ 0$ *be the compact spectral decomposition of* Y *with* $V^*V = I$. *Then* $\mathrm{Tr}\left(Y \log Y\right) = \mathrm{Tr}\left(R \log R\right)$.

Regularized Problem

Overall, the regularized QKD problem is

$$\begin{array}{cl} \min\limits_{R_\rho, R_\delta, R_\sigma} & \operatorname{Tr}\left(R_\delta \log R_\delta\right) - \operatorname{Tr}\left(R_\sigma \log R_\sigma\right) \\ \text{s.t.} & \Gamma_V(R_\rho) = \gamma \\ & R_\delta = \mathcal{G}_{UV}(R_\rho) \\ & R_\sigma = \mathcal{Z}_{UV}(R_\delta) \\ & R_\rho \succeq 0, R_\delta \succeq 0, R_\sigma \succeq 0 \end{array} \tag{3}$$

For simplicity in the following, we redefine notation

$$\rho \longleftarrow R_\rho,\ \delta \longleftarrow R_\delta,\ \sigma \longleftarrow R_\sigma$$

and mapping

$$\hat{\mathcal{G}} := \mathcal{G}_{UV},\ \hat{\mathcal{Z}} := \mathcal{Z}_{UV} \circ \mathcal{G}_{UV},$$

the final model is

$$\begin{array}{rcl} p^* = & \min\limits_{\rho} & f(\rho) = \operatorname{Tr}\left(\hat{\mathcal{G}}(\rho) \log \hat{\mathcal{G}}(\rho)\right) - \operatorname{Tr}\left(\hat{\mathcal{Z}}(\rho) \log \hat{\mathcal{Z}}(\rho)\right) \\ & \text{s.t.} & \Gamma_V(\rho) = \gamma_V \\ & & \rho \in \mathbb{H}_+^{n_\rho} \end{array} \tag{4}$$

Outline

Regularized QKD Problem

$$
\begin{aligned}
p^* = \quad \min_{\rho} \quad & f(\rho) = \operatorname{Tr}\left(\hat{\mathcal{G}}(\rho)\log\hat{\mathcal{G}}(\rho)\right) - \operatorname{Tr}\left(\hat{\mathcal{Z}}(\rho)\log\hat{\mathcal{Z}}(\rho)\right) \\
\text{s.t.} \quad & \Gamma_V(\rho) = \gamma_V \\
& \rho \in \mathbb{H}_+^{n_\rho}
\end{aligned} \tag{5}
$$

- Strong duality holds
- Note that

$$\rho \succ 0 \Longrightarrow \hat{\mathcal{G}}(\rho) \succ 0 \Longrightarrow \hat{\mathcal{Z}}(\rho) \succ 0$$

- The gradient of the objective is

$$\nabla f(\rho) = \hat{\mathcal{G}}^*(I + \log\hat{\mathcal{G}}(\rho)) - \hat{\mathcal{Z}}^*(I + \log\hat{\mathcal{Z}}(\rho))$$

and the Hessian is

$$\nabla^2 f(\rho)[\Delta\rho] = \hat{\mathcal{G}}^*(\log'\hat{\mathcal{G}}(\rho)[\hat{\mathcal{G}}(\Delta\rho)]) - \hat{\mathcal{Z}}^*(\log'\hat{\mathcal{Z}}(\rho)[\hat{\mathcal{Z}}(\Delta\rho)])$$

Perturbed KKT

Interior point method iteratively finds $\rho \succ 0$, $Z \succ 0$ and y to solve

$$
\begin{array}{ll}
F_\mu^d := \nabla f(\rho) + \Gamma_V^*(y) - Z = 0 & \text{(Dual feasibility)} \\
F_\mu^p := \Gamma_V(\rho) - \gamma_V = 0 & \text{(Primal feasibility)} \\
F_\mu^c := Z\rho - \mu I = 0 & \text{(Perturbed complementary slackness)}
\end{array}
\tag{6}
$$

while decreasing the perturbation parameter $\mu > 0$ to 0.

Note $F_\mu = \begin{pmatrix} F_\mu^d \\ F_\mu^p \\ F_\mu^c \end{pmatrix} : \mathbb{H}^{n_\rho} \times \mathbb{R}^{m_V} \times \mathbb{H}^{n_\rho} \longrightarrow \mathbb{H}^{n_\rho} \times \mathbb{R}^{m_V} \times \mathbb{C}^{n_\rho \times n_\rho}$ is overdetermined [1]. So they instead solve

$$\min_{\rho, y, Z} \frac{1}{2} \| F_\mu(\rho, y, Z) \|^2$$

by Gauss-Newton method:

$$F_\mu'^*(F_\mu' d + F_\mu) = 0$$

[1] **Question: why not do symmetrization?**

Nullspace Representation of Primal Feasibility

- To reduce the size of linear system, let $\hat{\rho} \in \mathbb{H}^{n_\rho}$ be feasible to $\Gamma_V(\hat{\rho}) = \gamma_V$ and define $\mathcal{N}^* : \mathbb{R}^{n_\rho^2 - m_V} \to \mathbb{H}^{n_\rho}$ such that

$$\Gamma_V(\rho) = \gamma_V \Longleftrightarrow \mathcal{N}^*(v) + \hat{\rho} = \rho, \quad \text{for some } v$$

  Redefine the primal residual as

$$F_\mu^p := \mathcal{N}^*(v) + \hat{\rho} - \rho$$

  and the perturbed KKT becomes

$$F_\mu(\rho, v, y, Z) = \begin{bmatrix} F_\mu^d \\ F_\mu^p \\ F_\mu^c \end{bmatrix} = \begin{bmatrix} \nabla f(\rho) + \Gamma_V^*(y) - Z \\ \mathcal{N}^*(v) + \hat{\rho} - \rho \\ Z\rho - \mu I \end{bmatrix}$$

Projected Gauss-Newton Search Direction

Search direction d is the least squares solution of

$$F'_\mu d + F_\mu = 0$$

Note that

$$F'_\mu d + F_\mu = \begin{bmatrix} \nabla^2 f(\rho)\Delta\rho + \Gamma_V^*(\Delta y) - \Delta Z \\ \mathcal{N}^*(\Delta v) - \Delta\rho \\ \Delta Z\rho + Z\Delta\rho \end{bmatrix} + \begin{bmatrix} F_\mu^d \\ F_\mu^p \\ F_\mu^c \end{bmatrix}$$

Use variable elimination

$$\begin{aligned} \Delta\rho &= \mathcal{N}^*(\Delta v) + F_\mu^p \\ \Delta Z &= \nabla^2 f(\rho)\Delta\rho + \Gamma_V^*\Delta y + F_\mu^d \end{aligned} \tag{7}$$

and solve least squares solution $(\Delta v, \Delta y)$ of

$$Z\mathcal{N}^*(\Delta v) + \nabla^2 f(\rho)\mathcal{N}^*(\Delta v)\rho + \Gamma_V^*(\Delta y)\rho = -F_\mu^c - ZF_\mu^p - (F_\mu^d + \nabla^2 f(\rho)F_\mu^\rho)\rho \tag{8}$$

Properties of Search Direction

Theorem 2

Let α be a step size and consider the update
$\rho_+ \leftarrow \rho + \alpha\Delta\rho = \rho + \alpha F_\mu^p + \alpha\mathcal{N}^*(\Delta v)$,

1. *If $\alpha = 1$, then exact primal feasibility holds, $\mathcal{N}^*(v_+) + \hat{\rho} - \rho_+ = 0$*
2. *Suppose exact primal feasibility is achieved, then it is maintained regardless of subsequential step sizes.*

Proof. Calculate the new primal residual

$$\begin{aligned} F_\mu^p \quad &= \mathcal{N}^*(v + \Delta v) + \hat{\rho} - \rho - \Delta\rho \\ &= F_\mu^p + \mathcal{N}^*(\Delta v) - \Delta\rho = 0 \end{aligned}$$

Suppose $\Gamma_V(\rho) = \gamma_V$,

$$\Gamma_V(\rho_+) = \Gamma_V(\rho + \alpha\Delta\rho) = \gamma_V + \alpha\Gamma_V(\mathcal{N}^*(\Delta v) + F_\mu^p) = \gamma_V$$

So they take step size of one as quickly as possible.

Algorithm

Algorithm 1: Projected Gauss-Newton Primal-Dual Interior Point

---

Input: $\hat{\rho} \succ 0$, $\mu > 0$, $\eta \in (0, 1)$

while *stopping criteria not met* **do**

  Solve least squres solution $(\Delta v, \Delta y)$ of equation (8);

  Set $(\Delta \rho, \Delta Z)$ by equation (7);

  Choose step size α by backtracking to ensure positive definiteness;

  Set $(\rho, y, Z) \leftarrow (\rho, y, Z) + \alpha(\Delta \rho, \Delta y, \Delta Z)$;

  Set $\mu \leftarrow \eta \cdot \langle \rho, Z \rangle / n_\rho$;

end

- Sparse null space representation
- Diagonal preconditioning for least squares
- Step size: start from $\alpha^* = -\langle \nabla f(\rho), \Delta \rho \rangle / \langle \Delta \rho, \nabla^2 f(\rho) \Delta \rho \rangle$
- Stopping criteria: let RHS be RHS of equation (8),

$$\text{relstopgap} = \frac{\max\{bestub - bestlb, \|RHS\|}{1 + \frac{1}{2}\min\{\|\rho\| + \|Z\|, |bestub| + |bestlb|\}} < \epsilon$$

Optimal Value Bounds

- Upper bound: Given primal iterate $\bar{\rho} \succ 0$, compute

$$\rho = \bar{\rho} - \Gamma_V^{\dagger}(\Gamma_V(\rho) - \gamma_V) = \arg\min_{\rho} \left\{ \frac{1}{2} \|\rho - \bar{\rho}\|^2 : \Gamma_V(\rho) = \gamma_V \right\}$$

  If $\rho \succeq 0$, then $f(\rho) \geq p^*$.

- Lower bound: Given primal-dual iterates $(\bar{\rho}, \bar{y})$ with $\bar{\rho} \succ 0$, let $Z = \nabla f(\bar{\rho}) + \Gamma^*(\bar{y})$. If $Z \succeq 0$, then

$$p^* \geq f(\bar{\rho}) + \langle \bar{y}, \Gamma_V(\bar{\rho}) - \gamma_V \rangle - \langle \bar{\rho}, Z \rangle$$

  Proof. Weak duality.

Outline

Comparison with Other Approaches

- Frank-Wolfe[2]
- Semidefinite approximation of matrix logarithm[3]

Problem Data			Gauss-Newton		Frank-Wolfe with FR		Frank-Wolfe w/o FR		cvxquad with FR	
protocol	parameter	size	gap	time	gap	time	gap	time	gap	time
ebBB84	(0.50,0.05)	(4,16)	5.98e-13	0.63	1.01e-04	84.39	1.17e-04	94.71	5.46e-01	216.37
ebBB84	(0.90,0.07)	(4,16)	2.33e-13	0.25	2.32e-04	85.09	2.54e-04	113.20	7.39e-01	647.60
pmBB84	(0.50,0.05)	(8,32)	5.51e-13	0.24	3.13e-05	1.85	6.47e-04	1.47	5.26e-01	170.12
pmBB84	(0.90,0.07)	(8,32)	1.01e-12	0.17	7.31e-05	1.04	6.25e-04	31.77	6.84e-01	235.89
mdiBB84	(0.50,0.05)	(48,96)	7.86e-13	1.08	9.62e-05	1.54	5.39e-04	134.79	1.82e-01	588.71
mdiBB84	(0.90,0.07)	(48,96)	2.96e-13	1.12	1.51e-04	101.84	3.48e-03	408.26	4.57e-01	574.31
TFQKD	(0.80,100,0.70)	(12,24)	7.67e-13	1.20	1.98e-04	96.08	1.55e-03	179.57	3.98e-03	990.92
TFQKD	(0.90,200,0.70)	(12,24)	3.42e-12	0.96	1.92e-05	2.07	1.65e-04	2.15	2.26e-04	875.44
DMCV	(10,60,0.05,0.35)	(44,176)	2.74e-09	510.66	2.44e-06	1015.14	3.36e-06	1709.65	⋆⋆	0.86
DMCV	(11,120,0.05,0.35)	(48,192)	3.23e-09	720.61	2.60e-06	348.81	1.98e-06	628.25	⋆⋆	1.24
dprBB84	(1,0.08,30)	(12,48)	4.92e-13	0.93	3.79e-06	77.86	9.38e-05	108.50	⋆⋆	119.20
dprBB84	(2,0.14,30)	(24,96)	1.04e-12	10.07	6.19e-06	15.61	3.62e-06	27.79	⋆⋆	105.40
dprBB84	(3,0.10,30)	(36,144)	4.96e-13	61.32	6.48e-04	7.89	2.08e-02	28.46	⋆⋆	614.71
dprBB84	(4,0.12,30)	(48,192)	1.13e-12	272.09	4.41e-05	15.28	9.79e-04	184.42	⋆⋆	3397.34

Table 5.1: Numerical Report from Three Algorithms

"gap" is the relative gap $\frac{bestub - bestlb}{1+\frac{|bestub|+|bestlb|}{2}}$; $\epsilon = 10^{-9}$ or 10^{-12}.

- Gauss-Newton outperforms
- FR can help

[2] Adam Winick, Norbert Lütkenhaus, and Patrick J Coles. "Reliable numerical key rates for quantum key distribution". In: *Quantum* 2 (2018), p. 77.

[3] Hamza Fawzi, James Saunderson, and Pablo A Parrilo. "Semidefinite approximations of the matrix logarithm". In: *Foundations of Computational Mathematics* 19.2 (2019), pp. 259–296.

Comparison with Analytical Solution

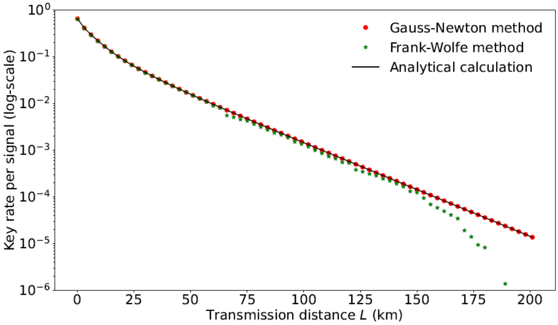

Figure 5.2: Comparison of key rate for discrete-modulated continuous-variable **QKD** (Appendix C.5) among our Gauss-Newton method, the Frank-Wolfe method and analytical key rate for the noise $\xi = 0$ case.

Conclusion

- Solve a nonlinear SDP from QKD accurately and efficiently
  - Regularize degenerate problem by FR (avoid perturbation)
  - Stable primal-dual interior point method
- Lack of theoretical analysis on convergence
- Many other nonlinear SDP problems from quantum computing

Conclusion

- Solve a nonlinear SDP from QKD accurately and efficiently
  - Regularize degenerate problem by FR (avoid perturbation)
  - Stable primal-dual interior point method
- Lack of theoretical analysis on convergence
- Many other nonlinear SDP problems from quantum computing

Thank You! Questions?