Team 15
Quantum State Tomography.
Project Report

Alexandr Talitsky, Andrey Vlasov, Anton Bozhedarov,
Nikolay Shvetsov, Polina Pilyugina

December 2018

Abstract

In current work we consider and implement several algorithms for quantum state tomography, such that pseudo-inverse matrix, direct gradient algorithm, hedged likelihood method and semi-definite programming methods. As testing model we develop procedure of numerical experiment of measurement qubits in pure state.

1 Introduction

Quantum tomography (QT) is applied on a source of systems, to determine what the quantum state is of the output of that source. Unlike a measurement on a single system, which determines the system's current state after the measurement (in general, the act of making a measurement alters the quantum state), quantum tomography works to determine the state prior to the measurements.

It is used in quantum computing to examine the result of quantum algorithm (i.e. state of a qubit) or in quantum communications.

Quantum state may be represented as a density operator $\rho \in \mathcal{L}(\mathcal{H}^d)$. And has following properties:

- $\rho = \rho^*$
- $\rho \geq 0$
- $Tr[\rho] = 1$

where * denotes hermitian conjugation.

Density matrix is fully described by $d^2 - 1$ parameters ($d = 2$ for qubit). To find that parameters one have to make $d^2 - 1$ measurements and one addition measurement to normalize the experimental frequencies into probabilities.

Measurements in quantum mechanics can be described using a positive-operator valued measure (POVM) formalism. Operator E_j from a POVM set

($\sum_j E_j = I$) corresponds to "j" measurement outcome. The probability of getting j outcome after a single measurement is

$$p_j = Tr[E_j \rho]. \tag{1}$$

2 Methodology

2.1 Linear Inversion

A state ρ of a qubit can be parameterized in the following way:

$$\rho = \frac{I + \sum_k (a_k \cdot \sigma_k)}{2} \qquad k = 1, 2, 3, \tag{2}$$

where $\begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix}^T$ lies inside a unit ball in $\mathbb{R}^3$ and σ_k are standard Puali matrices.

One can use vector notation:

$$\rho = \begin{pmatrix} I & \sigma_1 & \sigma_2 & \sigma_3 \end{pmatrix} \begin{pmatrix} 1 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \mathbf{\Sigma}^T \mathbf{a}.$$

Measurement operator can also be parameterized in the same notation:

$$E_j = s_0 I + \sum_k s_{k,j} \sigma_k = \mathbf{\Sigma}^T \mathbf{s_j}, \tag{3}$$

And $s_{k,j} \in \mathbb{R}$.

We can rewrite probability of j-th outcome in i-th measurement as follow

$$p_j^{(i)} = (\mathbf{s_j^{(i)}})^{\mathbf{T}} \mathbf{a}$$

And for the set all measurements, i.e. the experiment statistics:

$$\mathbf{p} = \mathbf{S} \mathbf{a} \tag{4}$$

Solution of that system of linear equations is estimated density matrix

$$\mathbf{a} = (\mathbf{S}^* \mathbf{S})^{\mathbf{-1}} \mathbf{S}^* \mathbf{p} \tag{5}$$

2.2 Direct Gradient Method

Another common approach to the solution of the problem is called Maximum Likelihood Estimation. From experimental data one can construct Likelihood functional (LH)

$$\mathcal{L}(\{n_j\}|\rho) = \prod_j tr[\rho E_j]^{n_j} = \prod_j p_j^{n_j}, \tag{6}$$

where n_j is amount of outcomes j from N total measurements. To recover parameters of ρ one need to solve optimization task and maximize the likelihood function, i.e.

$$\rho_{est} = \arg\max \mathcal{L}(\{n_j\}|\rho)$$

It is more convenient to maximize $\log(\mathcal{L})$ (LLH). Since log is convex function, it has the same maximum as LH. A possible approach to maximize LLH is Gradient Descent Method. As was shown in [6] step of iterative procedure is following:

$$\rho_{k+1} = \frac{[1+\frac{\epsilon}{2}(R_k-1)]\rho_k[1+\frac{\epsilon}{2}(R_k-1)]}{tr\{[1+\frac{\epsilon}{2}(R_k-1)]\rho_k[1+\frac{\epsilon}{2}(R_k-1)]\}} \tag{7}$$

where $R_k = \sum_j \frac{f_j}{p_j^{(k)}} E_j$.

The optimal choice of ϵ in (7) may be done as line search procedure. One use two trial value of ϵ and compute LLH for them. Then, these two points and current value of LLH (its corresponding $\epsilon is 0$) are used to construct quadratic polynomial that interpolates LLH(ϵ). $\hat{\epsilon_k}$ that maximize polynomial is using to compute ρ_{k+1}. In current work we use another approach. One chooses several trial ϵ and compute LLH. One of these points (which maximize LLH) is used as a parameter of the next step.

2.3 Hedged Quantum State Estimation

Quantum state tomography procedure uses finite number of measurements and as a result problem of zero probabilities is arising. As was suggested by Robin Blume-Kohout, the Hedged Maximum Likelihood (HML) method results in full-rank estimator, so that the problem of mistakenly obtained zero eigenvalues is reduced. Also, HML estimator allows for examination of not only the peak (as in likelihood), but also the victimity of maximum.

In our work we will focus on maximization of convex HML functional, defined as follows:

$$\mathcal{L}_H(\{n_j\};\rho) = (\det\rho)^\beta\, \mathcal{L}(\{n_j\};\rho) \tag{8}$$

In order to search for HML maximum, the following iteration method is used:

1. Start from an arbitrary state, usually $\rho_0 = 1/D$, where D is identity matrix multiplied by dimension of Hilbert space.

2. Follow the direction of the steepest gradient.

3. Look for the extremal state extimator $\hat{\rho}_{HML}$ with the following iterative equations:

$$\rho_{k+1} = \frac{[\,1+\Delta_k\,]\rho_k[\,1+\Delta_k\,]}{\mathrm{tr}\{[\,1+\Delta_k\,]\rho_k[\,1+\Delta_k\,]\}}, \tag{9}$$

where

$$\Delta_k = \frac{\epsilon}{2}[\,\beta(\rho_k^{-1}-D) + N(R_k-1) \tag{10}$$

As was suggested in [6] the best choices for constants are: $\beta = \frac{1}{2}, \epsilon = \frac{1}{N}$.

After solving the optimization problem using methods from sections 2.2, 2.3 or solving system of linear equations from 2.1 one can recover a quantum state $\hat{\rho}$. But there is a problem which may happens after solution of this particular problem: state $\hat{\rho}$ is not physical (i.e. $Tr[\rho] > 1$ or $\rho < 0$). And hence, it is necessary to find the nearest state $\tilde{\rho}$ which is physical by these conditions. The following method was suggested to deal with this problem.

2.4 Semi-definite programming

Optimization problem under given constraints can be solved with help of semi-definite programming. Task is:

$$\max_x f_0(x)$$

$$\text{cond } f_i(x) > 0 \text{ and } Ax = b$$

Lets define F:

$$F = \{x | Ax = b, f_i(x) \leq 0\} \tag{11}$$

In this case F is convex set.

General Idea

1. x_k is initial point. $x_k \in F$
2. Let's find c is direction.
3. Calculate $x_{k+1} = x_k + c\alpha$, such that $x_{k+1} \in F$ and $f_0(x_{k+1}) < f_0(x_k)$

The question is how to choose the right direction, one of solutions is Newton Barrier Function.

$$B(x, \mu) = f(x) - \lambda \sum_i ln(f_i(x)) \text{ is Barrier Function.}$$

$$p_B = x + \tfrac{1}{\lambda} X^2 (A^* \mu - b) \text{ is Newton Direction.}$$

But for numerical algorithms more complicated methods are used, some of them are presented in **SVXPY** package for python and for semidefinite programming have been used **SVXOPT** solver

There are 2 estimators, which we used in this case:

The matrix Dantzig selector

$$\hat{\rho}_{DS} = \arg\min_X \|X\|_{tr}, \text{ s.t. } \|A^*(A(X) - y)\| < \lambda$$

The Lasso matrix:

$$\hat{\rho}_{Lasso} = \arg\min_X \frac{1}{2} \|A(X) - y\|_2^2 + \mu \|X\|_{tr}$$

There are Theorems, which connect parameters μ and λ with Error Bounds and prove that it is correct estimated value.

2.5 Evaluation of performance

In order to evaluate the performance of each method, we needed to measure an error in some norm. For this purpose we used the norms, specified below:

- Hilbert-Schmidt

$$\Delta^{HS}(\rho',\,\rho) = \frac{1}{\sqrt{2}}\mathrm{Tr}[\,(\rho'-\rho)^2\,]^{1/2}$$

- Trace-distance

$$\Delta^{T}(\rho',\,\rho) = \frac{1}{2}\mathrm{Tr}[\,|\rho'-\rho|\,]$$

- Infidelity

$$\Delta^{IF}(\rho',\,\rho) = 1 - \mathrm{Tr}\Big[\,\sqrt{\sqrt{\rho'}\rho\sqrt{\rho'}}\,\Big]^2$$

3 Results

Current work requires data of measurements of quantum system. We decided to simulate experiment and develop for this purpose next procedure:

1. choose set of measurement operators E_j (minimal POVM set). The set of operators was suggested in [5];
2. generate state of a qubit as a random point on a unitary sphere;
3. calculate probabilities of outcome for given set of POVM
4. generate N outcomes, where each outcome was chosen with appropriate probability (we consider $N = 1000$);
5. calculate frequency probability from from obtained list of outcomes.

To simulate noise in measurements we use model of depolarizing channel with level of noice p:

$$\Phi_p(\rho) = \frac{I}{2}\rho + (1-p)\rho. \tag{12}$$

i.e. every outcome was chosen as equiprobable outcome with probability p and remain the same as in point 4 with probability $1-p$. Then, we use described above strategies to reconstruct density matrix ρ_{est}.

In the appendix one can see numerical comparison of tomography algorithms. As was expected the algorithms give wrong result with noise increasing fig.1, fig.2b, fig.3b, fig.4. Convergence of the direct gradient method is broken with the addition of noise (fig. 2a).

One can note that performance of different algorithms almost the same (fig.**??**). This result may be a consequence of testing model. Here we consider tomography of pure state qubit, but differences in methods' performance may be significant for mixed states and many-qubits system.

So, the optimal method in as a trade-off between implementation complexity and convergence speed for single qubit is hedged likelihood estimator. To expand this conclusion more complex analysis is required.

4 Problems

In this project our team faced with following problems:

- **Scallability**. In order to increase complexity of the solving problem we developed algorithm for generating pure state of many-qubits systems. In this case we increase the dimensionality of the corresponding Hilbert space that but in physical meaning we do not increase complexity of the task (we still can perform one qubit tomography many times). In other words we have to simulate entanglement between particles, but this task does not require special methods of NLA and, moreover, does not change considered algorithms of tomoghraphy. So we decided to test all methods on a single qubit system.

- **Quantum state restrictions**. Quantum state should require properties described in 1. In order to satisfy the specified constraints, we set zero eigenvalues of the density matrix to zero and use trace normalization. But this is probably not the best choice. This problem requires more careful study but it is not correlated to the course topics.

5 Contribution

Alexandr Talitsky: Implementation of method of semi-definite programming.

Andrey Vlasov: Implementation of method of direct gradient algorithm.

Anton Bozhedarov: Development of algorithm of the experimental modeling qubit tomography and pseudo-inverse matrix method implementation.

Nikolay Shvetsov: Implementation of the hedged likelihood method.

Polina Pilyugina: Comparison of different distance metrics, implementation of the hedged likelihood method.

6 Related works

General overview of the quantum tomography and resent achievements in this area was described in [3]. Choice of minimal POVM set for optimal measurement one can find in [5]. Description of used algorithms one can find in [6]. In [4] were described several metrics to compare two quantum states.

References

[1] `https://github.com/Anton-31-96/NLA_2018_team_15_tomography`

[2] Martin Ringbauer *Exploring Quantum Foundations with Single Photons*

[3] Alessandro Bisio, Giulio Chiribella, Giacomo Mauro, D'Ariano, Stefano Facchini, Paolo Perinotti *Optimal quantum tomography*
https://arxiv.org/abs/1702.08751

[4] Takanori Sugiyama, Peter S. Turner, Mio Murao *Precision-guaranteed quantum tomography*
https://arxiv.org/abs/1306.4191

[5] Jaroslav Rehacek, Berthold-Georg Englert, Dagomir Kaszlikowski *Minimal qubit tomography*
https://arxiv.org/abs/quant-ph/0405084

[6] Teo Yong Siah *Numerical Estimation Schemes for Quantum Tomography*

A Figures

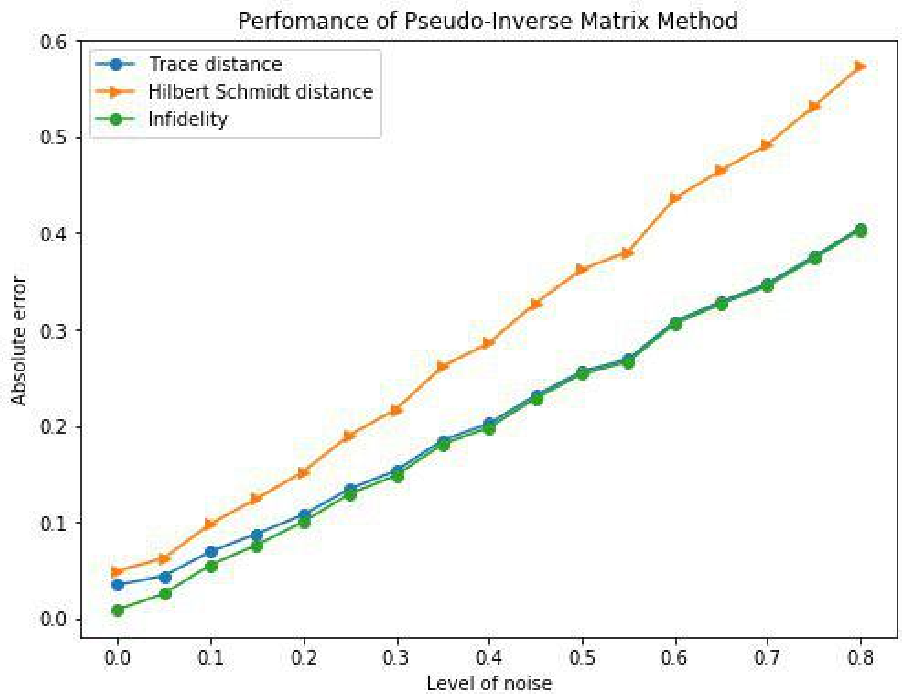
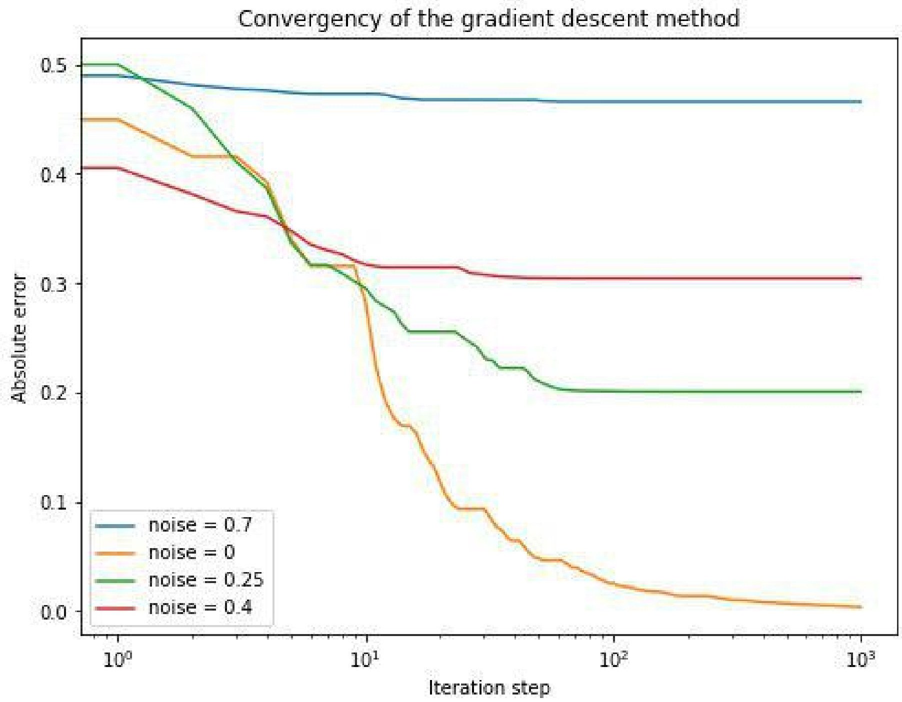
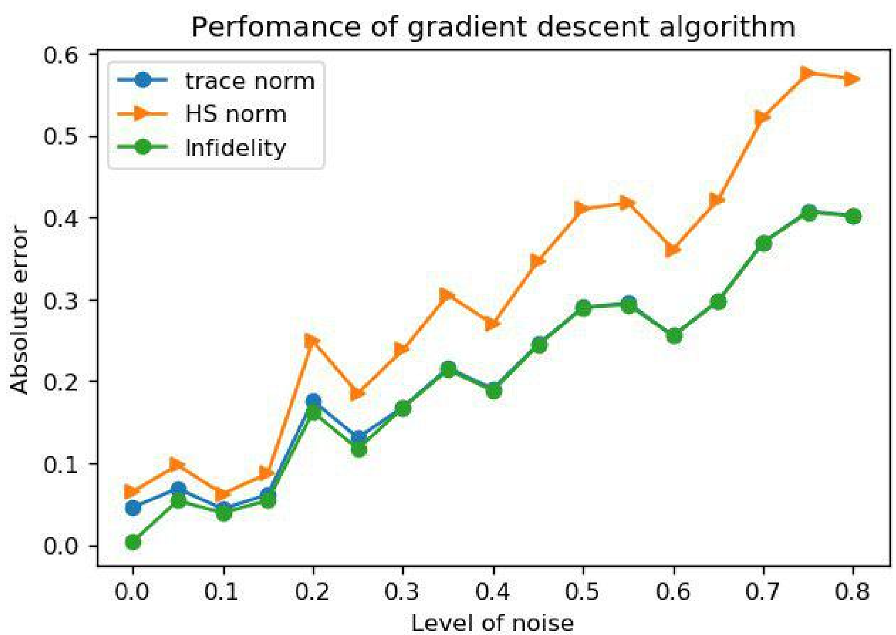


Figure 1: Performance of pseudo-inverse matrix method for different levels of noise.

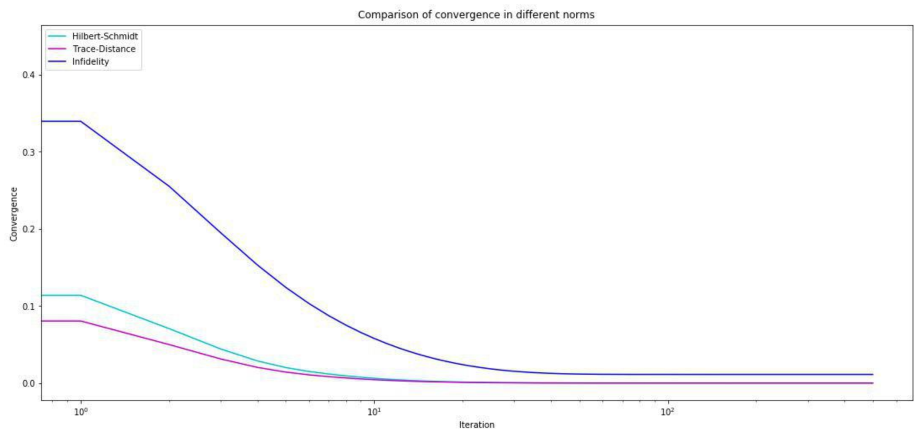


(a) Convergence of direct gradient algorithm for different level of noise.

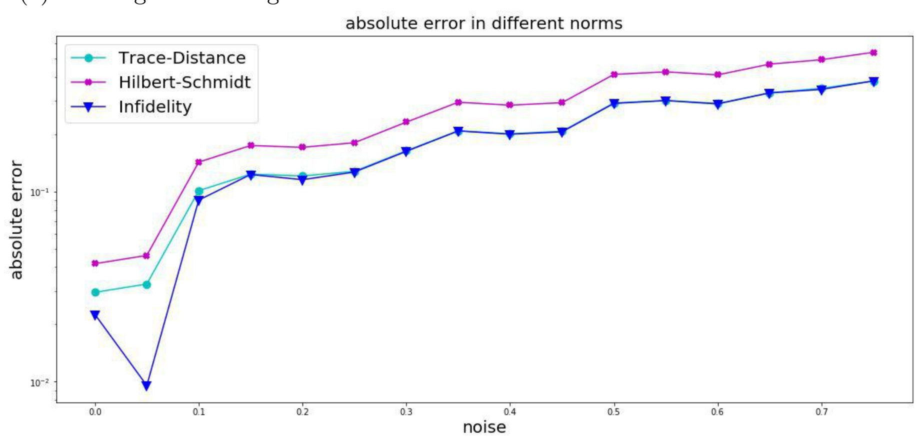


(b) Error of direct gradient algorithm result for different levels of noise

Figure 2: Performance of direct gradient algorithm

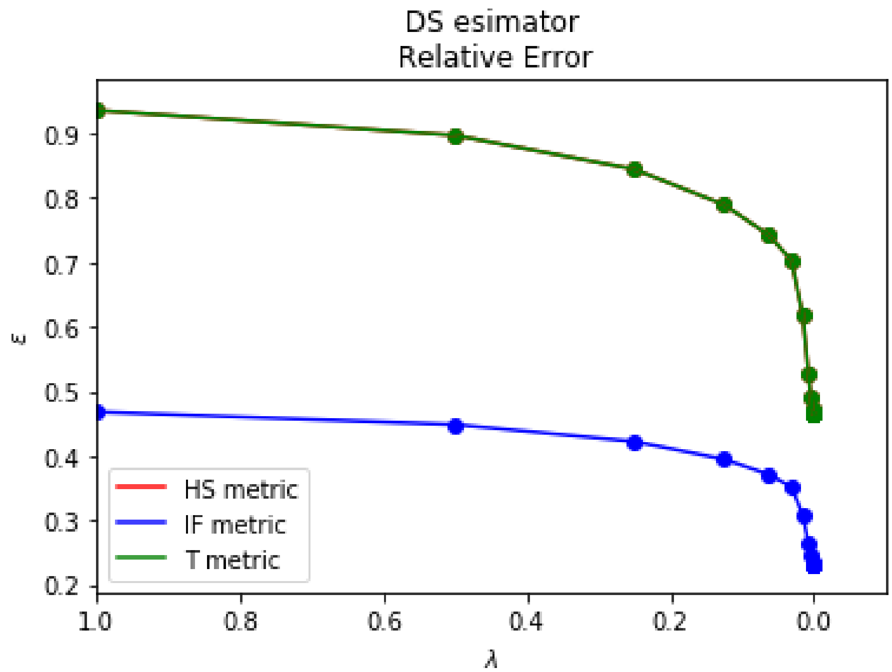


(a) Convergence of hedged likelihood estimator for different metrics without noise.

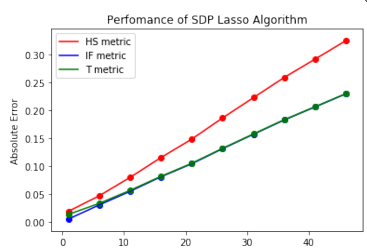


(b) Error of hedged likelihood estimator result

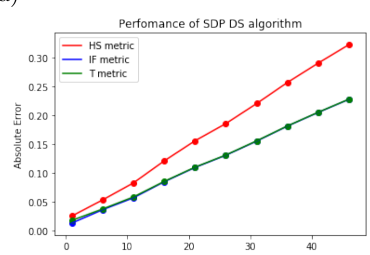
Figure 3: Performance of hedged likelihood estimator



(a)

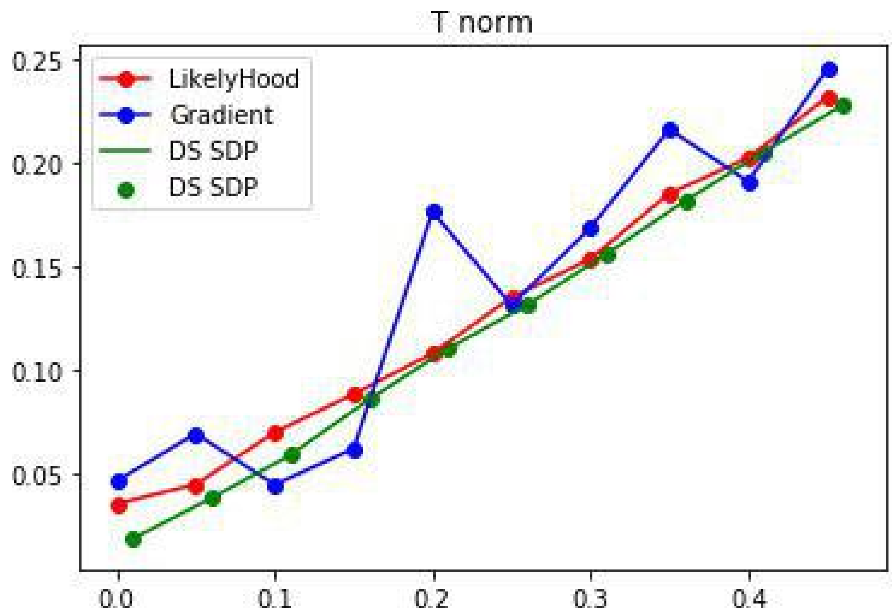


(b) Error of Lasso algorithm result

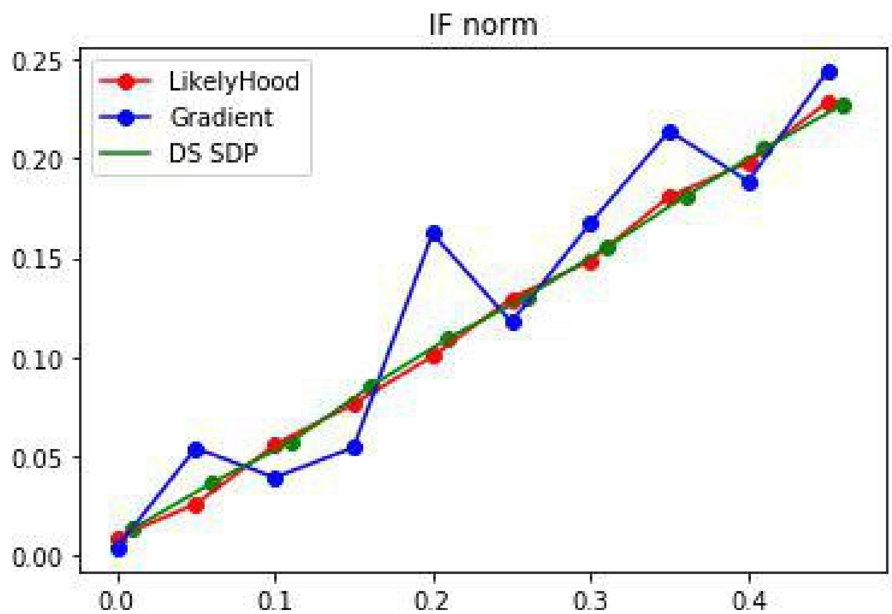


(c) Error of DS algorithm result

Figure 4: Performance of SDP algorithms



(a) Comparison of algorithms with respect to level of noise in trace-distance metrics



(b) Comparison of algorithms with respect to level of noise for infidelity metrics

Figure 5: Comparison of performance of algorithms in case of single qubit