

# Quantum Tomography Stats Model

May 2023

## 1 Introduction

Given the measurement data we wish to determine the best possible estimate for the quantum state. We can use Bayes' theorem to define the probability of the state given the data:

$$\Pr(\rho|\text{Data}) = \frac{\Pr(\text{Data}|\rho)\Pr(\rho)}{\Pr(\text{Data})} .$$

We define these functions as the following

- Posterior:  $\Pr(\rho|\text{Data})$
- Likelihood:  $\mathcal{L}(\rho) = \Pr(\text{Data}|\rho)$
- Prior:  $\Pr(\rho)$
- Evidence:  $\Pr(\text{Data})$

## 2 Distribution of Counts

The counts of the measurements can be modeled by a Poisson Distribution. With high enough counts we can approximate this with a Normal Distribution using the Central Limit Theorem. A important fact of the poisson distribution is the variance is equal to the mean.

## 3 1det/qubit

For each measurement we have 1 count number. From the Photonic State Tomography paper, the distribution of the counts follows:

$$n_i \sim \text{Poiss}(\mu_i) \xrightarrow{\text{CLT}} n_i \sim \text{Norm}(\mu_i, \sigma_i) ,$$
$$\mu_i = I_0 I_i \text{Tr} (M'_i \rho) + a_i .$$

### 3.1 Variables:

- $n_i$ : Number of counts on measurement i
- $\mu_i$ : Expected number of counts on measurement i given  $\rho$
- $i \in [1, m]$  |  $m$  = Number of Measurements
- $\sigma_i^2$ : Variance for the number of counts on measurement i given  $\rho$   
 –  $\sigma_i^2 = \mu_i$  from assuming a poisson distribution
- $I_0$ : Is the overall intensity
- $I_i$ : Is the relative intensity of measurement i given as an input. Default is 1
- $M'_i$ : Is measurement i with cross talk correction
- $a_i$ : Is the predicted accidental counts for measurement i.

### 3.2 Log-Likelihood

$$\begin{aligned}\mathcal{L}(\rho) &= \Pr(\text{Data}|\rho) , \\ \mathcal{L}(\rho) &= \prod_{i=1}^m \Pr(n_i|\mu_i) \\ \mathcal{L}(\rho) &= \prod_{i=1}^m \frac{1}{\sqrt{\sigma_i^2 2\pi}} \exp\left(-\frac{(n_i - \mu_i)^2}{2\sigma_i^2}\right)\end{aligned}$$

The  $\frac{1}{\sqrt{\sigma_i^2 2\pi}}$  is a normalization factor. Since we are maximizing this function we can ignore this term

$$\mathcal{L}(\rho) \propto \prod_{i=1}^m \exp\left(-\frac{(n_i - \mu_i)^2}{2\sigma_i^2}\right)$$

We can plug in  $\sigma_i^2 = \mu_i$  and take the log of this function to get our log-Likelihood function.

$$\log(\mathcal{L}(\rho)) = \sum_{i=1}^m -\frac{(n_i - \mu_i)^2}{2\mu_i^2}$$

## 4 2det/qubit

For each measurement a complete number of counts on all the possible outcomes. We define the following variables:

From the Photonic State Tomography paper:

$$\begin{aligned}n_{ij} &\sim \text{Pois}(\mu_{ij}) \xrightarrow{\text{CLT}} n_{ij} \sim \text{Norm}(\mu_{ij}, \sigma_{ij}) , \\ \mu_{ij} &= I_0 I_i E_j \text{Tr}(M'_{ij} \rho) + a_{ij} .\end{aligned}$$

## 4.1 Variables

- $n_{ij}$ : Number of counts on measurement i for detector j
- $\mu_{ij}$ : Expected number of counts on measurement i, detector j given  $\rho$
- $i \in [1, m]$  |  $m$  = Number of Measurements
- $j \in [1, k]$  |  $k$  = Number of Detector Pairs
- $\sigma_{ij}^2$ : Variance for the number of counts on measurement i, detector j given  $\rho$ 
  - $\sigma_{ij}^2 = \mu_{ij}$  from assuming a poisson distribution
- $I_0$ : Is the overall intensity
- $I_i$ : Is the relative intensity of measurement i given as an input. Default is 1
- $E_j$  Is the relative efficiency on the jth basis
- $M'_{ij}$ : Is the jth basis of measurement i with cross talk correction

## 4.2 Log-Likelyhood

$$\mathcal{L}(\rho) = \Pr(\text{Data}|\rho)$$

$$\mathcal{L}(\rho) = \prod_{i=1}^m \prod_{j=1}^k \Pr(n_{ij}|\mu_{ij})$$

$$\mathcal{L}(\rho) = \prod_{i=1}^m \prod_{j=1}^k \frac{1}{\sqrt{\sigma_{ij}^2 2\pi}} \exp\left(-\frac{(n_{ij} - \mu_{ij})^2}{2\sigma_{ij}^2}\right)$$

The  $\frac{1}{\sqrt{\sigma_{ij}^2 2\pi}}$  is a normalization factor. Since we are maximizing this function we can ignore this term

$$\mathcal{L}(\rho) \propto \prod_{i=1}^m \prod_{j=1}^k \exp\left(-\frac{(n_{ij} - \mu_{ij})^2}{2\sigma_{ij}^2}\right)$$

We can plug in  $\sigma_{ij}^2 = \mu_{ij}$  and take the log of this function to get our log-Likelyhood function.

$$\log(\mathcal{L}(\rho)) = \sum_{i=1}^m \sum_{j=1}^k -\frac{(n_{ij} - \mu_{ij})^2}{2\mu_{ij}}$$

## 5 Error Correction

### 5.1 Accidental Correction

Accidental Correction is used for 2 detectors per qubit.

$$a_{ij} = \frac{W_j}{T_i * 10^9} \prod_{k=1}^2 S_{ijk}$$

- $T_i$  Is the time in nanoseconds of measurement i given as an input. Default is 1
- $S_{ijk}$  Is the kth single count on measurement i given as an input. Default is 0
- $W_j$  Is the coincidence window duration for the jth basis as an input. Default is 0

### 5.2 Crosstalk

$$M'_{i,j} = \sum_{j'=1}^k C_{j,j'} M_{i,j}$$

- $M'_{ij}$ : Is the jth basis of measurement i with cross talk correction
- $M_{ij}$ : Is the jth basis of measurement i. The state the target quantum state is projected on. These are pure states here represented as density matrices
- $C_{ij}$ : The index at the jth row and j'th column of the crosstalk matrix
- $i \in [1, m] \mid m = \text{Number of Measurements}$
- $j, j' \in [1, k] \mid k = \text{Number of Detector Pairs, which is equal to the number of qubits squared.}$

## 6 Lab setup example: 2 qubits with 2 det/qubit vs 1 det/qubit

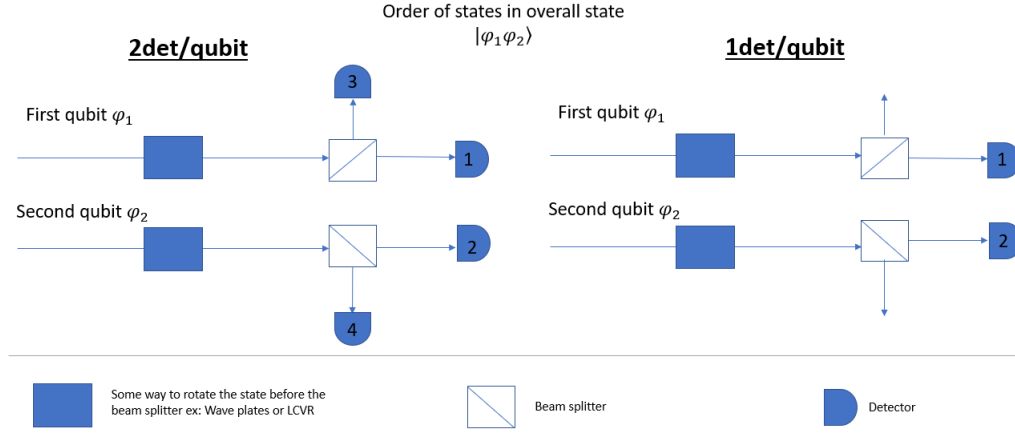
We need to specify 2 states per measurement. The first is the state that the first qubit is projected onto when it ends up at detector 1. The second is the state that the second qubit is projected onto when it ends up at detector 2.

#### 2det/qubit

1. Det-pair 1 : 1-2
2. Det-pair 2 : 1-4
3. Det-pair 3 : 3-2
4. Det-pair 4 : 3-4

#### 1det/qubit

1. Det-pair 1 : 1-2



## 7 Estimators

### 7.1 Maximum Likelihood Estimator(MLE)

The MLE estimator is found by maximizing the log of the likelihood function:

$$\rho_{\text{MLE}} = \underset{\rho}{\operatorname{argmin}} -\log(\mathcal{L}(\rho)) .$$

### 7.2 Hedged Maximum Likelihood Estimator(HMLE)

Hedged maximum likelihood is a simple modification of the maximum likelihood approach. Instead of maximizing the likelihood, the estimate is the one that maximizes the product of the likelihood with the following hedge function:

$$h(\rho) = \det(\rho)^\beta \mid \beta \in \left[0, \frac{1}{2}\right] ,$$

$$\rho_{\text{HMLE}} = \underset{\rho}{\operatorname{argmin}} -\log(\mathcal{L}(\rho)h(\rho)) .$$

The value of Beta can be defined in the conf settings; default is 1/2. Robin Blume-Kohout covers the hedged likelihood function in more detail in the paper titled [Hedged Maximum Likelihood Estimation](<https://arxiv.org/pdf/1001.2029.pdf>).

### 7.3 Bayesian Estimator(BME)

The Bayes estimator is the expected value of the posterior

$$\hat{\rho}_{\text{BME}} = \mathbb{E}_{\text{posterior}}[\rho|\text{Data}]$$

$$\hat{\rho}_{\text{BME}} = \int \rho \text{Pr}(\rho | \text{Data}) d\rho$$

$$\hat{\rho}_{\text{BME}} = \int \rho \frac{\text{Pr}(\text{Data} | \rho) \text{Pr}(\rho)}{\text{Pr}(\text{Data})} d\rho$$

We only care about an estimator that is proportional to the density matrix since we can normalize the matrix at the end.

$$\hat{\rho}_{\text{BME}} \propto \int \rho \text{Pr}(\text{Data} | \rho) \text{Pr}(\rho) d\rho$$

$$\hat{\rho}_{\text{BME}} \propto \mathbb{E}_{\text{prior}}[\rho \text{Pr}(\text{Data} | \rho)]$$

Monte Carlo Approximation of the Bayesian Estimator

$$\hat{\rho}_{\text{BME}} = \mathbb{E}_{\text{posterior}}[\rho | \text{Data}] \propto \mathbb{E}_{\text{prior}}[\rho \text{Pr}(\text{Data} | \rho)]$$

$\rho_i \sim \text{Pr}(\rho)$  is a random sample from the prior

$$\hat{\rho}_{\text{BME}} \propto \sum_i \rho_i \text{Pr}(\text{Data} | \rho_i)$$

$\rho_i \sim \text{Pr}(\rho | \text{Data})$  is a random sample from posterior

$$\hat{\rho}_{\text{BME}} \propto \sum_i \rho_i$$

## 8 Linear Inversion

The Linear Inversion method is a simpler approach to State Tomography. It is used to get a starting state for MLE and is typically not recommended on its own as it can give non-valid density matrices.

### 8.1 1 Qubit

We can represent any density matrix in terms of the stokes parameters and the pauli matrices.

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_2 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\rho = \frac{1}{2} \sum_{i=0}^3 S_i \sigma_i \quad \Bigg| \quad S_i = \text{Tr}(\rho \sigma_i)$$

$$\rho = \frac{1}{2} \vec{S}_\rho \cdot \vec{\sigma} \quad \Bigg| \quad \vec{\sigma} = \begin{bmatrix} \sigma_0 \\ \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{bmatrix}, \quad \vec{S}_\rho = \begin{bmatrix} \text{Tr}(\rho \sigma_0) \\ \text{Tr}(\rho \sigma_1) \\ \text{Tr}(\rho \sigma_2) \\ \text{Tr}(\rho \sigma_3) \end{bmatrix}$$

For a general set of projectors  $\{M_i\}$ , the probability that  $\rho$  will collapse to  $M_i$  is  $P_i = \text{Tr}(M_i\rho)$ .

$$\vec{P} = \begin{bmatrix} \text{Tr}(M_0\rho) \\ \text{Tr}(M_1\rho) \\ \vdots \\ \text{Tr}(M_{m-1}\rho) \end{bmatrix} = \begin{bmatrix} \text{Tr}\left(M_0\left(\frac{1}{2}\sum_{j=0}^3 S_j\sigma_j\right)\right) \\ \text{Tr}\left(M_1\left(\frac{1}{2}\sum_{j=0}^3 S_j\sigma_j\right)\right) \\ \vdots \\ \text{Tr}\left(M_{m-1}\left(\frac{1}{2}\sum_{j=0}^3 S_j\sigma_j\right)\right) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\sum_{j=0}^3 \text{Tr}(M_0\sigma_j) S_j \\ \frac{1}{2}\sum_{j=0}^3 \text{Tr}(M_1\sigma_j) S_j \\ \vdots \\ \frac{1}{2}\sum_{j=0}^3 \text{Tr}(M_{m-1}\sigma_j) S_j \end{bmatrix}$$

Define the matrix  $\mathbf{B} \mid B_{ij} = \text{Tr}(M_i\sigma_j)$

$$\vec{P} = \mathbf{B}\vec{S}_\rho$$

Finally we use the psuedo-inverse to get  $\vec{S}_\rho$ :

$$\begin{aligned} \vec{P} &= \mathbf{B}\vec{S}_\rho \\ \mathbf{B}^T \vec{P} &= \mathbf{B}^T \mathbf{B} \vec{S}_\rho \\ (\mathbf{B}^T \mathbf{B})^{-1} \mathbf{B}^T \vec{P} &= \vec{S}_\rho \end{aligned}$$

Our probability vector can be approximated by our counts

## 8.2 N Qubits

This can easily be scaled to larger dimensions. We simply scale the pauli basis by doing:

$$\vec{\sigma} = \bigotimes_{n=1}^N \begin{bmatrix} \sigma_0 \\ \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{bmatrix}$$

$$\rho = \frac{1}{2^N} \sum_{i=0}^{4^N-1} S_i \sigma_i \quad \Bigg| \quad S_i = \text{Tr}(\rho \sigma_i)$$

For example, for 2 qubits our pauli basis would be

$$\vec{\sigma} = \begin{bmatrix} \sigma_0 \otimes \sigma_0 \\ \sigma_0 \otimes \sigma_1 \\ \sigma_0 \otimes \sigma_2 \\ \sigma_0 \otimes \sigma_3 \\ \sigma_1 \otimes \sigma_0 \\ \sigma_1 \otimes \sigma_1 \\ \sigma_1 \otimes \sigma_2 \\ \sigma_1 \otimes \sigma_3 \\ \sigma_2 \otimes \sigma_0 \\ \sigma_2 \otimes \sigma_1 \\ \sigma_2 \otimes \sigma_2 \\ \sigma_2 \otimes \sigma_3 \\ \sigma_3 \otimes \sigma_0 \\ \sigma_3 \otimes \sigma_1 \\ \sigma_3 \otimes \sigma_2 \\ \sigma_3 \otimes \sigma_3 \end{bmatrix}$$

## 9 Parameterization

During the optimization we parameterize the density matrix as follows in order to ensure we are only considering valid density matrices:

$$\rho = TT^\dagger, \quad T = \begin{bmatrix} t_1 & 0 & \dots & 0 \\ t_{2^n+1} + it_{2^n+2} & t_2 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ t_{4^n-1} + it_{4^n} & t_{4^n-3} + it_{4^n-2} & t_{4^n-5} + it_{4^n-4} & t_{2n} \end{bmatrix}.$$