

Supplemental Information:

Readout error mitigated quantum state tomography tested on superconducting qubits

Adrian Skasberg Aasen,^{1,2,*} Andras Di Giovanni,^{3,*} Hannes Rotzinger,^{3,4} Alexey V. Ustinov,^{3,4} and Martin Gärttner^{2,†}

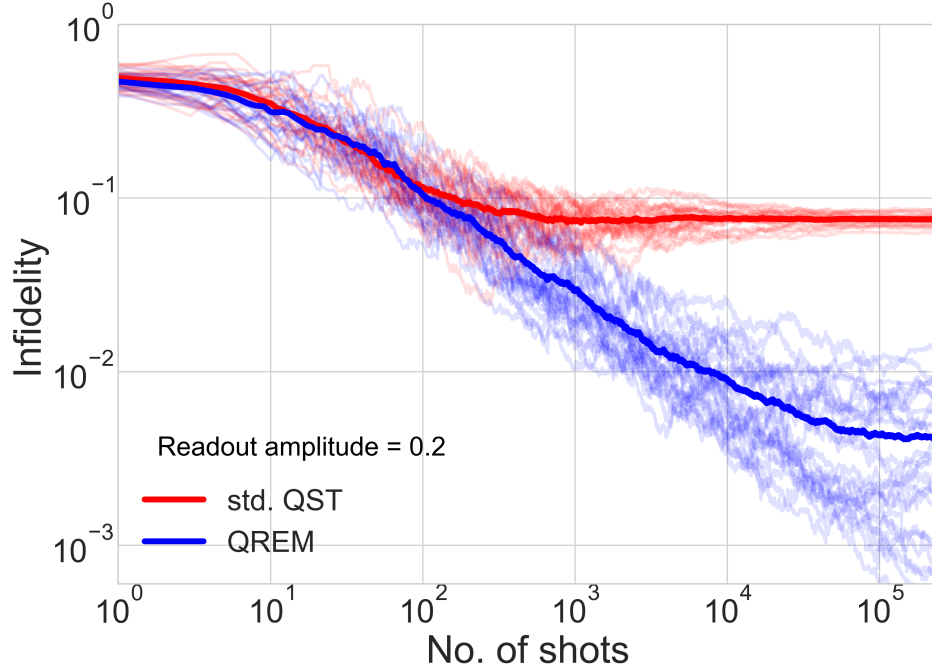
¹*Kirchhoff-Institut für Physik, Universität Heidelberg,
Im Neuenheimer Feld 227, 69120 Heidelberg, Germany*

²*Institut für Festkörpertheorie und -optik, Friedrich-Schiller-Universität Jena, Max-Wien-Platz 1, 07743 Jena, Germany*

³*Physikalisches Institut, Karlsruher Institut für Technologie, Kaiserstraße 12, 76131 Karlsruhe, Germany*

⁴*Institut für QuantenMaterialien und Technologien,
Karlsruher Institut für Technologie, Kaiserstraße 12, 76131 Karlsruhe, Germany*

(Dated: August 16, 2024)



Supplementary Figure 1. Replotted Fig. 10 (a) from the main text with all averaged curves plotted to show spread. The blue curves are readout error mitigated (QREM) and the red curves are standard quantum state tomography (std. QST). Each of the 25 Haar-random state reconstruction infidelities are plotted as translucent curves of the respective colors, while the solid colored curves are the average.

SUPPLEMENTARY NOTE 1: CONVERGENCE OF MODIFIED ESTIMATOR

Here we show that the protocol estimator, eq. (12) in main text, converges to the noiseless state with the modifications presented. By convergence of the estimator, we mean that the expected state converges to the true state for all quantum states, in the limit of infinite statistics. To make this explicit, we show that an estimator with the standard likelihood function, eq. (8) in the main text, converges to any given true state ρ_T . The generalization to noisy measurements follows immediately.

For simplicity, we will assume that the measurement device used implements an IC and minimal POVM. By extension, any quantum state is uniquely defined by the probabilities $p_i = \text{Tr}(\rho M_i) = \langle M_i \rangle$. The natural estimator for

* These authors contributed equally to this work.

† martin.gaerttner@uni-jena.de

the probabilities is the frequency of each outcome,

$$\hat{p}_i = \frac{n_i}{N}. \quad (\text{S1})$$

All we need to show is that using this estimator for the probabilities maximizes the likelihood function, eq. (8) in the main text. To be explicit, we need to show that $\hat{p}_i = \frac{n_i}{N}$ converges to the probability $p_i = \text{Tr}(\rho M_i)$. This can be done by maximizing the log-likelihood function $\log(\mathcal{L}_M(\rho))$ with Lagrange multipliers. This yields that $p_i = \hat{p}_i$ maximizes the likelihood function. Therefore, if the experiment implements the POVM $\{M_i\}$, which means it samples from the probabilities $p_i = \text{Tr}(\rho_T M_i)$, it is clear that we have a reconstruction $\hat{\rho}$ that converges to ρ_T .

Consider the case where the measurement device is imperfect and implements the noisy POVM $\{\tilde{M}_i\}$, and we have the reconstructed POVM from the QDT $\{\tilde{M}_i^{\text{estm}}\}$. Analogous to the noiseless case, $\hat{p}_i = \frac{\tilde{n}_i}{N}$ converges to $\tilde{p}_i = \text{Tr}(\rho_T \tilde{M}_i)$, as long as $\langle \tilde{M}_i^{\text{estm}} \rangle$ converges to $\tilde{M}_i$. This is guaranteed by using the MLE reconstruction outlined in Ref. [1].

SUPPLEMENTARY NOTE 2: STATE PREPARATION ERRORS AS EFFECTIVE STATE-DEPENDENT READOUT ERROR CHANNELS

We present a straightforward model demonstrating how unitary gate errors can create an effective state-dependent error channel during state preparation, and how this influences the readout error mitigation protocol. By state-dependent, we mean that the effective error channel introduced during state preparation depends on specific properties of the prepared state. Our examination will be restricted to the case of a single qubit.

Consider a channel that describes the preparation of basis state ρ_S from the ground state $|0\rangle$.

$$\mathcal{E}_S(|0\rangle\langle 0|) = \rho_S. \quad (\text{S2})$$

In the operator-sum representation a general channel can be represented as [2]

$$\mathcal{E}(\rho) = \sum_i K_i \rho K_i^\dagger, \quad (\text{S3})$$

where $\sum_i K_i^\dagger K_i = \mathbb{1}$. When the state preparation is noiseless, the operator-sum representation is particularly simple

$$\mathcal{E}_{S,\text{ideal}}(\rho) = U_S \rho U_S^\dagger, \quad (\text{S4})$$

where U_S unitary that implements the transformation $|0\rangle\langle 0| \rightarrow \rho_S$. We can decompose any single-qubit unitary [2]

$$U_S = R_Z(\theta_2) R_Y(\theta_1) R_Z(\theta_0), \quad (\text{S5})$$

where $R_\alpha(\theta) = e^{-i\theta\sigma_\alpha/2}$ are the single qubit rotation operators about the Cartesian axis α of the Bloch sphere. We will now consider the channel $\mathcal{E}_S$ where with probability p_i implements a unitary $\tilde{U}_{S,i}$ which deviates from the ideal U_S . We want to show that $\mathcal{E}_S = \mathcal{E}_{S,\text{error}} \circ \mathcal{E}_{S,\text{ideal}}$.

Proposition 1. *The state preparation channel can be separated into $\mathcal{E}_S = \mathcal{E}_{S,\text{error}} \circ \mathcal{E}_{S,\text{ideal}}$, where $\mathcal{E}_{S,\text{error}}$ generally depends on the state S being prepared.*

Proof. Our strategy is to show that each $\tilde{U}_{S,i} = \tilde{U}_{S,i,\text{error}} U_S$, where $\tilde{U}_{S,i,\text{error}}$ is some unitary, that may or may not depend on S . From such a decomposition of each i it immediately follows that the channels also decompose. We start by expressing $\tilde{U}_{S,i}$ in the decomposition in eq. (S5) (where we drop the index i for convenience)

$$\tilde{U}_S = R_Z(\tilde{\theta}_2) R_Y(\tilde{\theta}_1) R_Z(\tilde{\theta}_0). \quad (\text{S6})$$

This channel will always be applied to the ground state $|0\rangle$, so we can ignore $R_Z(\tilde{\theta}_0)$ as it does not affect the ground state. The remaining angles can be expressed in terms of the ideal ones as

$$\tilde{\theta}_k = \theta_k + \Delta_k. \quad (\text{S7})$$

Since the rotation gate commutes with itself, we can split them

$$R_\alpha(\tilde{\theta}_k) = R_\alpha(\theta_k) R_\alpha(\Delta_k) = R_\alpha(\Delta_k) R_\alpha(\theta_k). \quad (\text{S8})$$

We can then write the the noisy unitary as

$$\begin{aligned}
\tilde{U}_S &= R_Z(\tilde{\theta}_2) R_Y(\tilde{\theta}_1) \\
&= R_Z(\Delta_2) R_Z(\theta_2) R_Y(\Delta_1) R_Y(\theta_1) \\
&= \underbrace{R_Z(\Delta_2) R_Z(\theta_2) R_Y(\Delta_1) R_Z^\dagger(\theta_2)}_{U_{S,\text{error}}} \underbrace{R_Z(\theta_2) R_Y(\theta_1)}_{U_S} \\
&= U_{S,\text{error}} U_S.
\end{aligned} \tag{S9}$$

The error unitary generally depends on the rotation that was performed through θ_2 .

Consider now the channel $\mathcal{E}_S$ where we with some probability p_i implement the unitary $\tilde{U}_{S,i}$,

$$\mathcal{E}_S(\rho) = \sum_i p_i \tilde{U}_{S,i} \rho \tilde{U}_{S,i}^\dagger \tag{S10}$$

Inserting the noisy unitary from eq. (S9) into the channel we have

$$\begin{aligned}
\mathcal{E}_S(\rho) &= \sum_i p_i U_{S,\text{error}} \overbrace{U_S \rho U_S^\dagger}^{\mathcal{E}_{S,\text{ideal}}(\rho)} U_{S,\text{error}}^\dagger \\
&= \sum_i p_i U_{S,\text{error}} \mathcal{E}_{S,\text{ideal}}(\rho) U_{S,\text{error}}^\dagger \\
&= \mathcal{E}_{S,\text{error}} \circ \mathcal{E}_{S,\text{ideal}}(\rho)
\end{aligned} \tag{S11}$$

and we have showed that the noise channels separate. $\square$

This separation of the state preparation channel means that $\mathcal{E}_{S,\text{error}}$ will be interpreted as part of the readout errors in the passive picture. Importantly, we see the explicit dependence of the error channel on the prepared state and the distribution p_i . With a state-dependent error channel present, one cannot reconstruct a consistent set of POVM elements based on the $I \times S$ constraints, since each set of constraints per calibration state (S) comes from a different channel. This means, that the POVM reconstructed in the detector tomography stage, which is based on a finite set of calibration states, will in general not faithfully represent the error incurred for the calibration states themselves or any other prepared states.

With this simple and experimentally motivated model we can provide an explanation of why the protocol still works for some cases, while in others it does not. Note that the operator-sum representation is not unique, so it could still be possible to find representations where the explicit dependence vanishes, see e.g. increased temperature.

Qubit detuning

Qubit detuning has non-trivial contributions from both Δ_1 and Δ_2 . We therefore expect there to be state-dependent errors present in state preparation, and subsequently our readout error mitigation protocol to under-perform.

Increased temperature

The primary source of error at elevated temperatures stems from the excitation of the qubit ground state. A secondary contribution arises from unitary gate errors, with finite temperatures interpreted as an error symmetrically distributed over fluctuating Δ_1 and Δ_2 . The symmetric convex combination of these unitary rotations results in a depolarizing error channel that does not depend on the prepared state, thereby enabling us to transfer these state-preparation errors to the readout stage as well.

Shorter T_1 and T_2 times

Increased manipulation times cannot easily be viewed as effective unitary gate errors. However, it is plausible that also non-unitary errors can lead to a state-dependent effective error channel. The protocol's under-performance during extended manipulation times might suggest that state-dependent errors are indeed induced.

SUPPLEMENTARY NOTE 3: INSUFFICIENCY OF CLASSICAL ERROR MITIGATION

To demonstrate that classical errors are not sufficient to capture all error present in the measurement device, we investigate coherent errors in the POVM elements. Classical errors assume that the reconstructed POVM elements are purely diagonal in their measurement basis, i.e. only contains a statistical redistribution of the diagonal elements. Coherent errors manifest in the off-diagonal elements and are beyond classical. For a superconducting qubit the Pauli-6 POVM is performed by measuring individually the three bases of the Pauli-operators. We can then investigate how off-diagonal each of these bases are with respect to their ideal eigenbasis.

In Supplementary Figure 2 we present the reconstructed POVMs for some of the noisy runs in their ideal measurement basis (i.e. we present the POVM elements $R_{X \rightarrow Z} M_X R_{X \rightarrow Z}^\dagger$, $R_{Y \rightarrow Z} M_Y R_{Y \rightarrow Z}^\dagger$ and M_Z). For a single qubit it is sufficient to consider only the element corresponding to the outcome 0.

We found that for almost all cases, the z-basis measurement (the computational basis measurement) does not have coherence errors beyond statistical fluctuations, while for many instances the x- and y-basis measurements suffer from large coherence errors in the induced noise scenarios.

SUPPLEMENTARY NOTE 4: DIGITIZING EXPERIMENTAL DATA

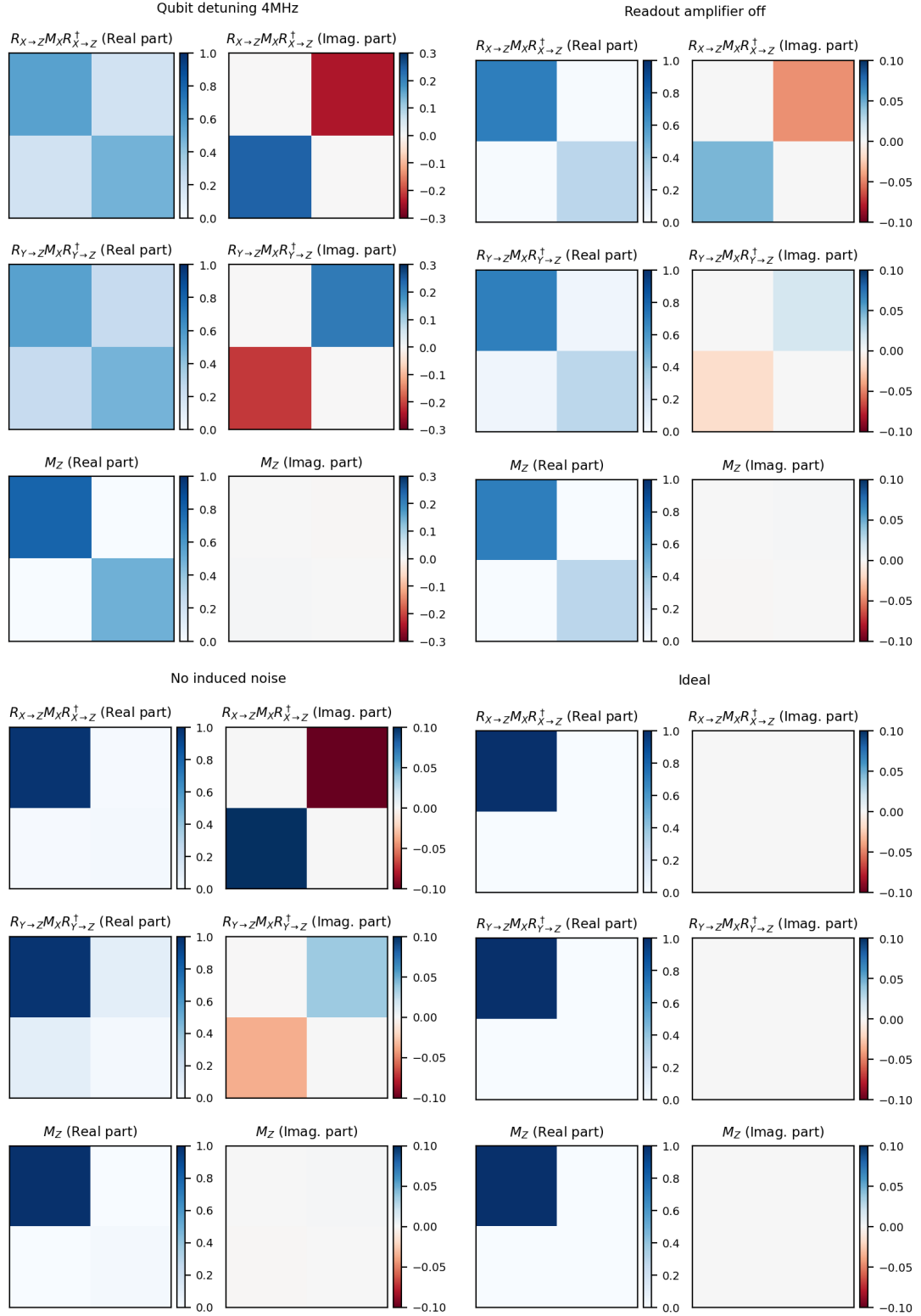
After measuring the length of a π -pulse, we use a supervised classification algorithm on the data to digitize the measured analog voltages (I and Q, see Supplementary Figure 3) into a single bit: 0 or 1. Often, two 2D Gaussian distributions can be fitted on the measured data points, but using machine-learning-based classification turned out to work more efficiently than Gaussian fitting, particularly for strong noises. We used the Python package sklearn, which runs in an intermediate level between the low-level experimental drivers and the high-level Bayesian and MLE algorithms.

SUPPLEMENTARY NOTE 5: OPEN-ACCESS DATASET

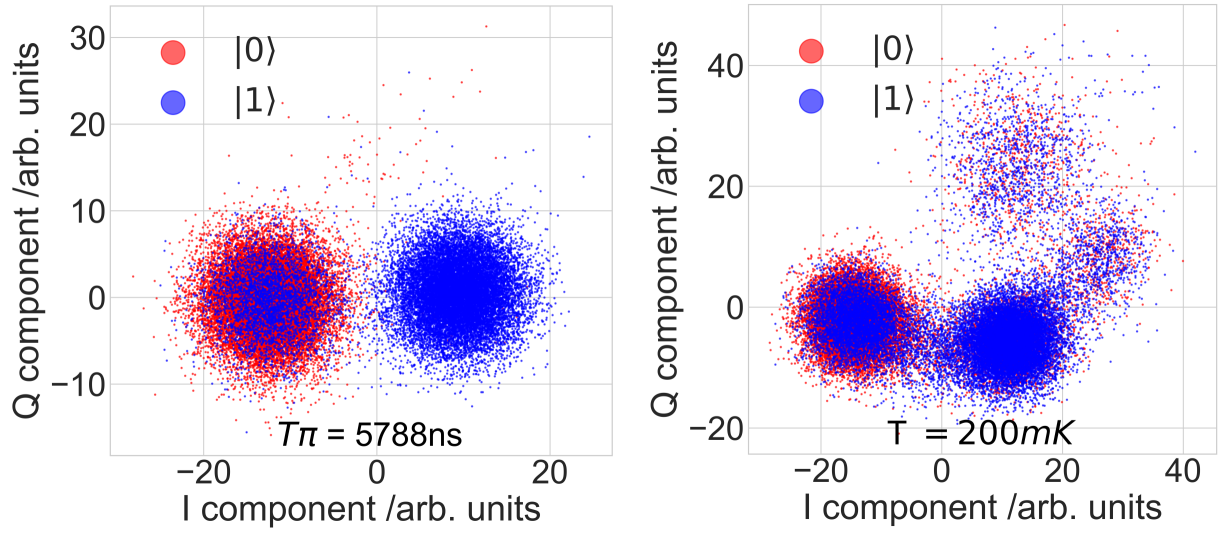
The datasets used for the experimental plots in the main text can be found on the KIT research data repository: <https://radar.kit.edu/radar/en/dataset/RNbuograoVUFQNB>. The dataset contains following folders:

Figure	Folder name
Figure 2	4noises
Figure 3	combined noises
Figure 4	dt shots
Figure 6	theory depol
Figure 7	temperature
Figure 10	two different readout amplitudes
Supplementary Figure 1	0.2 ro amp spreads

Note that Figure 6 contains purely simulated data.



Supplementary Figure 2. The positive operator-valued measure (POVM) elements in their ideal eigenbasis. Non-trivial off-diagonal elements signal coherent errors that are not captured by classical error mitigation. The plots shows, starting from the top left: Qubit detuning at 4 MHz, turning off the parametric amplifier, optimal operation of the experiment, and theoretical ideal POVM. The POVMs were estimated using $N = 240k$ single-shot measurements for each computational basis, which means statistical fluctuations are on the order of $\propto 3 \times 10^{-2}$, well below the coherence errors observed.



Supplementary Figure 3. Single-shot qubit measurements in the IQ plane, color-coded to the two qubit states. Long manipulation pulses allow the qubit to decay more, populating the ground state more often (left). Higher temperatures shift the readout to higher states, populating out-of-qubit states, too (right).

Supplementary Algorithm 1 Pseudo-code for readout error mitigation protocol

Input: $\{\rho_B\}$: List of calibration states for QDT $\{\rho_S\}$: (Haar-random) list of unknown states to measure. $\{\{M_i\}\}$: A set of POVMs forming an IC POVM, e.g basis measurement in $\sigma_x, \sigma_y, \sigma_z$ **Output:** $\{\rho_{\text{REM}}\}$: List of readout error mitigated reconstructed states $\{\{\tilde{M}_i\}\}$: A set of reconstructed noisy POVMs $\{I\}$: List of reconstruction infidelities between readout error mitigated QST and standard QST**Experiment:**

```

1: for each measurement basis  $\rho_B$  do
2:   for each POVM  $\{M_i\}$  do
3:     QDTdata  $\leftarrow$  Measure( $\rho_B, \{M_i\}$ )
4:   end for
5: end for
6: for each Haar-random state  $\rho_S$  do
7:   for each POVM  $\{M_i\}$  do
8:     QSTdata  $\leftarrow$  Measure( $\rho_S, \{M_i\}$ )
9:   end for
10: end for

```

Post-processing:

```

1: for each POVM in IC POVM do
2:    $\{\tilde{M}_i\} \leftarrow$  QDT(QDTdata,  $\rho_B$ ) [1]
3: end for
4: for each unknown state  $\rho_S$  do
5:   Construct modified likelihood function  $\mathcal{L}_{\tilde{M}}(\rho)$ , eq. (12) in main text
6:   Construct standard likelihood function  $\mathcal{L}_M(\rho)$ , eq. (8) in main text
7:   if BME then
8:     Integrate  $\mathcal{L}_{\tilde{M}}$  and find mean state  $\rho_{\text{BME}}$  [3]
9:     Integrate  $\mathcal{L}_M$  and find mean state  $\tilde{\rho}_{\text{BME}}$ 
10:   end if
11:   if MLE then
12:     Find  $\rho_{\text{MLE}}$  that maximizes  $\mathcal{L}_{\tilde{M}}$  [4]
13:     Find  $\tilde{\rho}_{\text{MLE}}$  that maximizes  $\mathcal{L}_M$ 
14:   end if
15:    $\{\rho_{\text{REM}}\} \leftarrow \rho_{\text{MLE/BME}}$ 
16:    $\{I\} \leftarrow \text{Inf}(\rho_{\text{MLE/BME}}, \tilde{\rho}_{\text{MLE/BME}})$ 
17: end for

```

SUPPLEMENTARY REFERENCES

- [1] J. Fiurášek, “Maximum-likelihood estimation of quantum measurement,” *Physical Review A*, vol. 64, July 2001.
- [2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, June 2012.
- [3] G. I. Struchalin, I. A. Pogorelov, S. S. Straupe, K. S. Kravtsov, I. V. Radchenko, and S. P. Kulik, “Experimental adaptive quantum tomography of two-qubit states,” *Physical Review A*, vol. 93, Jan. 2016.
- [4] A. I. Lvovsky, “Iterative maximum-likelihood reconstruction in quantum homodyne tomography,” *Journal of Optics B: Quantum and Semiclassical Optics*, vol. 6, pp. S556–S559, May 2004.