Quantum Bit Error Estimation Based on the Syndrome of a Linear Code

Christoph Pacher[1,†], **Gottfried Lechner**[2]

† Christoph.Pacher@AIT.ac.at

1 Safety & Security Department, AIT Austrian Institute of Technology GmbH, 1220 Vienna, Austria

2 University of South Australia, Institute for Telecommunications Research (ITR), Mawson Lakes, SA 5095, Australia

AUSTRIAN INSTITUTE OF TECHNOLOGY

University of South Australia

Institute for Telecommunications Research

SUMMARY

We derive and analyze a Maximum Likelihood (ML) estimator for the quantum bit error rate (QBER). The estimator is based on Low-Density Parity-Check (LDPC) codes. Bob takes as input only his raw key and the syndrome he has received from Alice. We focus our analysis [1] on check-regular LDPC codes where every row of the parity-check matrix has constant weight but briefly address the check-irregular case with non-constant weights as well. We obtain a quite accurate estimator that can be used for two tasks in QKD: as an improvement over the sampling estimator (which compares sets of individual bits), and to improve the efficiency in interactive reconciliation protocols.

1 RECONCILIATION WITH LOW-DENSITY PARITY-CHECK (LDPC) CODES

- One way to define a binary linear error correcting code is by means of its parity-check matrix H: The null-space of the parity-check matrix defines the set of all codewords: $\mathcal{C} = \left\{\boldsymbol{x} \in \{0,1\}^n : \boldsymbol{x}\mathrm{H}^T = \boldsymbol{0}\right\}$.
- If H is sparse the code is called *Low-Density Parity-Check (LDPC) code* [2].
- Codes with constant weight d (called check degree) in each row are *check-regular*.
- An important application is reconciliation of data in quantum crypto: Assume Alice and Bob have obtained correlated vectors, $\boldsymbol{x}_A$ and $\boldsymbol{x}_B = \boldsymbol{x}_A \oplus \boldsymbol{e}$, resp., where $\boldsymbol{e}$ is the errorword (of low weight).
  Then Alice calculates the syndrome $\boldsymbol{S}_A := \boldsymbol{x}_A\mathrm{H}^T$ of her vector $\boldsymbol{x}_A$ and an LDPC code with parity-check matrix H and sends $\boldsymbol{S}_A$ on an error-free channel to Bob. If the quantum bit error rate has not been too large, Bob can reconstruct $\boldsymbol{x}_A$ from $\boldsymbol{x}_B$ and $\boldsymbol{S}_A$.
- But, can we reuse the syndrome for further purposes?

2 ERROR ESTIMATION WITH LDPC CODES

Yes, Bob can :) estimate the quantum bit error rate prior to decoding!
We model the errors from the quantum channel, i.e. the individual bits of $\boldsymbol{e}$ as iid: $\Pr\{e_i = 1\} = \rho$, where ρ denotes the quantum bit error rate (QBER).

- Bob performs the calculation

$$\boldsymbol{S} := \boldsymbol{e}\mathrm{H}^T = (\boldsymbol{x}_A \oplus \boldsymbol{x}_B)\,\mathrm{H}^T = \boldsymbol{S}_A \oplus \boldsymbol{x}_B\mathrm{H}^T.$$

The individual bits of the syndrome $\boldsymbol{S}$ can be well approximated to also be i.i.d. The approximation consists in neglecting the (weak) correlation between syndrome bits that sum over a common data bit, x_i.

- With this approximation the probability q that a syndrome bit is one is [2]

$$q = f_d(\rho) := \sum_{\substack{1 \le i \le d \\ i \text{ odd}}} \binom{d}{i} \rho^i (1-\rho)^{d-i} = \frac{1-(1-2\rho)^d}{2}. \tag{1}$$

3 DERIVATION OF THE MAXIMUM LIKELIHOOD ESTIMATOR FOR THE QBER

Let m denote the length of $\boldsymbol{S}$, and $W = \mathrm{wt}\{\boldsymbol{S}\}$ denote the Hamming weight of $\boldsymbol{S}$. The syndrome weight W is a binomially distributed random variable, i.e.,

$$\Pr\{W = w\} = f_{binom}(w; m, q) := \binom{m}{w} q^w (1-q)^{m-w}, \tag{2}$$

and the maximum likelihood (ML) estimate for ρ given a syndrome weight w is

$$\hat{\rho}(w) = \arg\max_{\rho'} \left\{ f_{binom}(w; m, f_d(\rho')) \right\}, \tag{3}$$

which can be solved analytically. Equivalently, one can take the ML estimator for q

$$\hat{q}(w) = \frac{w}{m}, \tag{4}$$

and use it with (1) to obtain the estimate $\hat{\rho}$. Both approaches give the same result:

- The final estimator in closed form is

$$\hat{\rho}(w) = \begin{cases} \frac{1-\left(1-2\frac{w}{m}\right)^{\frac{1}{d}}}{2} & ; \frac{w}{m} \le 1/2 \\ \frac{1}{2} & ; \frac{w}{m} > 1/2 \end{cases}. \tag{5}$$

This approach can be generalized to check-irregular LDPC codes with different check degrees by replacing the binomial distribution in (2) with a multinomial distribution.

4 PROPERTIES OF THE MAXIMUM LIKELIHOOD ESTIMATOR

- Mean

$$\mu(d,\rho,m) = \mathbb{E}_W[\hat{\rho}(W)] = \frac{1}{2} - \frac{1}{2}\sum_{w=0}^{\lfloor m/2 \rfloor} f_{binom}(w; m, f_d(\rho)) \left(1 - 2\frac{w}{m}\right)^{\frac{1}{d}}, \tag{6}$$

- Bias

$$B(d,\rho,m) = \mu(d,\rho,m) - \rho. \tag{7}$$

- Mean squared error (MSE)

$$\begin{aligned} \mathrm{MSE}(d,\rho,m) = & \mathbb{E}_W\left[(\hat{\rho}(W)-\rho)^2\right] = \frac{1}{4} - 2\rho\mu(d,\rho,m) + \rho^2 \\ & + \frac{1}{4}\sum_{w=0}^{\lfloor m/2 \rfloor} f_{binom}(w; m, f_d(\rho)) \left(\left(1-2\frac{w}{m}\right)^{\frac{2}{d}} - 2\left(1-2\frac{w}{m}\right)^{\frac{1}{d}} \right). \end{aligned} \tag{8}$$

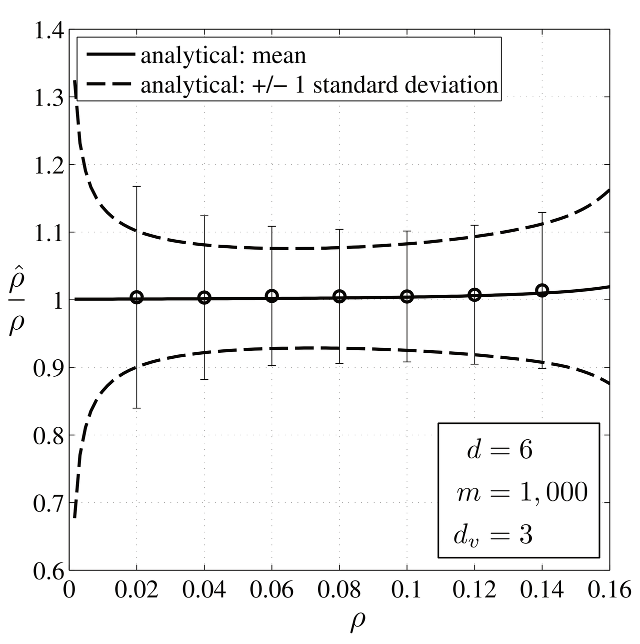
- Cramér-Rao Lower Bound
  The mean squared error of any biased estimator is lower bounded by

$$\mathrm{MSE}(d,\rho,m) \ge \frac{\left(\frac{\partial}{\partial\rho}\mu(d,\rho,m)\right)^2}{\mathcal{I}(\rho)} + B^2(d,\rho,m), \tag{9}$$

  where $\mathcal{I}(\rho)$ is the Fisher information that the syndrome $\boldsymbol{S}$ carries about ρ:

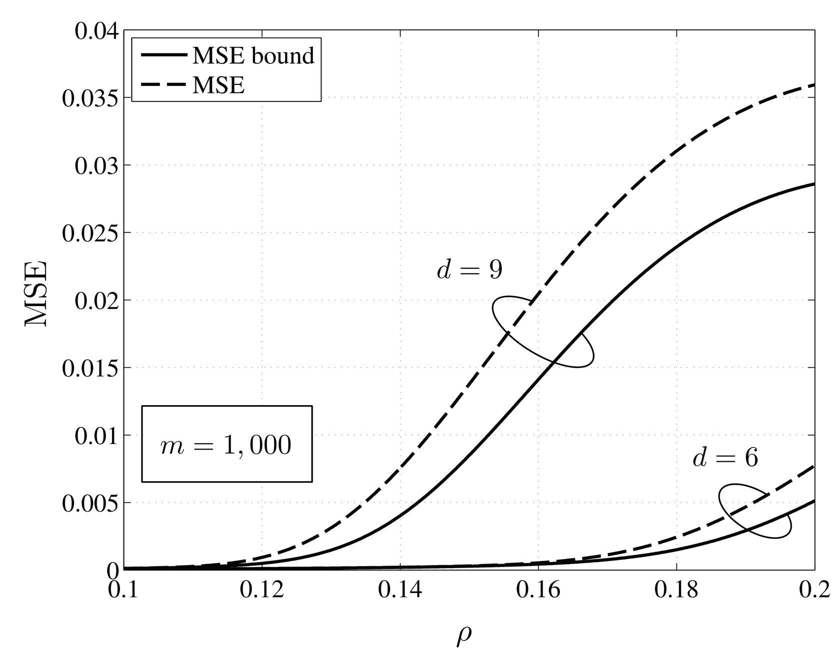
$$\mathcal{I}(\rho) = -\mathbb{E}_{\boldsymbol{S}}\left[\frac{\partial^2}{\partial\rho^2}\log\Pr\{\boldsymbol{S};\rho\}\right] = \frac{4md^2(1-2\rho)^{2d-2}}{1-(1-2\rho)^{2d}}. \tag{10}$$

5 NORMALIZED MEAN AND STANDARD DEVIATION



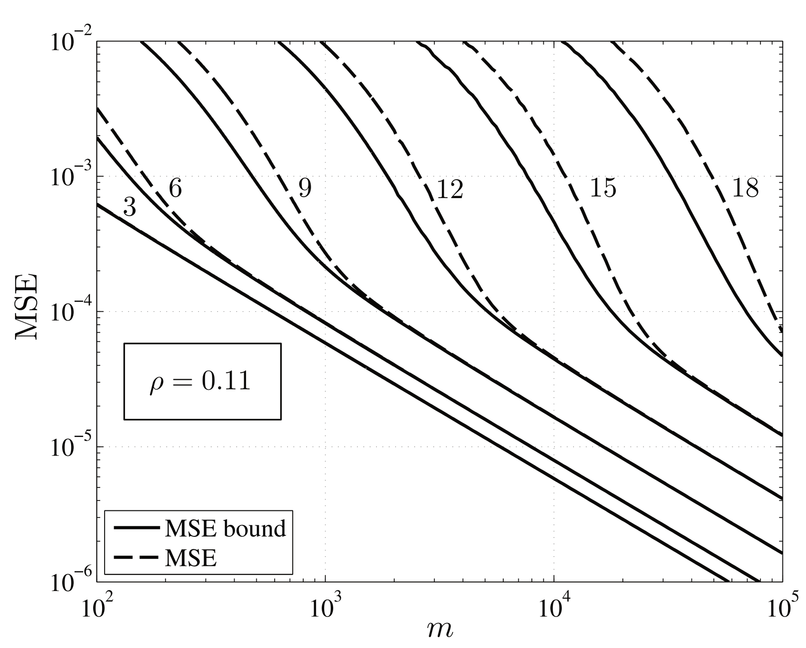
- The analytical mean (6) of the estimator is close to the true parameter ρ.
- The mean (shown as markers) of a simulation of a regular LDPC code matches the analytical result.
- The simulated normalised standard deviation (shown as error bars) is (slightly) larger than the analytical result due to the violation of the independence assumption of the syndrome bits.

6 MEAN SQUARED ERROR COMPARED TO CRAMÉR-RAO LOWER BOUND AS FUNCTIONS OF ERROR RATE AND CHECK DEGREE



- Due to the relatively small number of check nodes there is a relatively large gap between the MSE (8) of the estimator and the Cramér-Rao bound (9).
- A higher check node degree leads to a significant increase of the MSE.

7 MEAN SQUARED ERROR AS FUNCTION OF NUMBER OF CHECK NODES AND CHECK DEGREE



- For small check node degrees d already a relatively small number of check nodes m leads to a small MSE.
- For a large number of check nodes, the curves approach the inverse of the Fisher information.

REFERENCES

[1] G Lechner, C Pacher, *Estimating Channel Parameters from the Syndrome of a Linear Code*, accepted with minor revisions at IEEE Communications Letters (2013).

[2] R G Gallager, *Low-density parity-check codes*, IEEE Transactions on Information Theory **8**, 21–28 (1962).

[3] V Toto-Zarasoa, A Roumye, and C Guillemot, *Maximum Likelihood BSC Parameter Estimation for the Slepian-Wolf Problem*, IEEE Communications Letters **15**, 232–234 (2011).

ACKNOWLEDGEMENTS

This work has been supported by the Vienna Science and Technology Fund (WWTF), project ICT10-067 (HiPANQ).

WWTF
VIENNA SCIENCE AND TECHNOLOGY FUND