Efficient Codes for Quantum Key Distribution

MS Project

Jayanth Shreekumar

UCLA

December 1, 2023

Introduction - LDPC Codes[2]

- Sparse parity check matrix H.
- Elements of H are taken from a Galois Field GF(q)
- Message passing algorithm used for decoding.
- NB-LDPC codes drawback – decoding complexity [1].

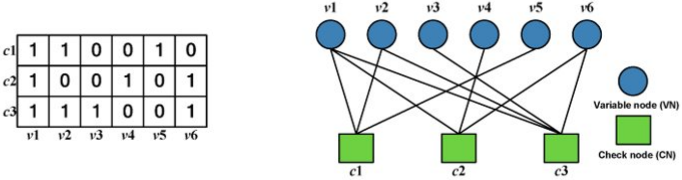


Figure: An LDPC code - Parity Check Matrix and its corresponding Tanner Graph. Taken from source

Introduction - Quantum Key Distribution [4]

- Secure communication protocol that involves several features from quantum mechanics.
- Time entanglement QKD [3] - Frames, bins, and binwidth.
- We observe a sequence of pairs of symbols (**X, Y**) with each symbol in the same Galois field size $GF(2^q)$.
- GOAL: Share symbols between two users through a public channel with little informatiom leak

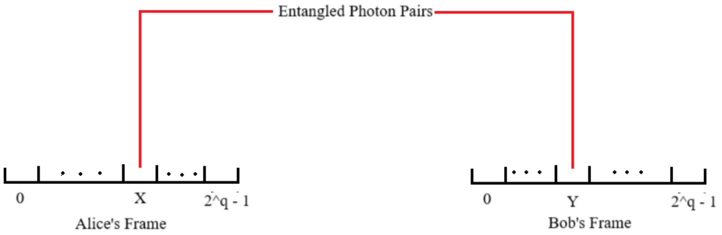


Figure: Key Generation using Time binning

Preliminaries - Channel Coding for Information Reconciliation

- Alice encodes her information by performing $\mathbf{R} = H\mathbf{X}$, and transmits this encoded message $\mathbf{R}$ over the public channel to Bob.
- Bob utilizes $\mathbf{R}$ as well as his own message $\mathbf{Y}$ to perform decoding and obtain $\hat{\mathbf{X}}$, his estimate of $\mathbf{X}$ which is Alice's message.
- Technique of utilizing side information $\mathbf{Y}$ to perform LDPC decoding of the message $\mathbf{R}$ is called the Slepian-Wolf scheme [5].
- To measure the performance of these techniques, the information reconciliation rate (IR rate) r is used:

$$r = q(1 - E)\frac{N - M}{N}$$

Preliminaries - Multi Layer Coding (MLC) Scheme [6]

- Using NB-LDPC codes for decoding at higher $GF(2^q)$ not scalable.
- Use MLC Scheme: map each symbol $X \in \mathbf{X}$, denoted by X^i to a sequence of k bits $[X^i_1, X^i_2, ..., X^i_k]$.
- Perform encoding $\mathbf{R}_j = H\mathbf{X}_j$ where j denotes layer and $\mathbf{X}_j = [X^1_j, X^2_j, ..., X^N_j]$ and transmit to Bob.
- Bob receives $\mathbf{R} = [R^1, R^2, ..., R^M]$ where each R^i is made up of bits given by $R^i = [R^i_1, R^i_2, ..., R^i_k]$, and performs binary LDPC Slepian-Wolf decoding on these bits layer-wise to recover $\hat{\mathbf{X}}_j = [\hat{X}^1_j, \hat{X}^2_j, ..., \hat{X}^N_j]$ for all j and thus also recover $\hat{\mathbf{X}}$.

Preliminaries - Multi Layer Coding (MLC) Scheme

- Generalizing this coding scheme, instead of using a bit for every layer, we can use l bits per layer.
- Convert every symbol $X \in \mathbf{X}$ into a sequence of symbols $[X_1, X_2, ..., X_b, X_{b+1}]$ where the first b symbols belong to the Galois field $\mathbb{F}(2^a)$ and the final symbol, if there is a reminder present, belongs to the Galois field $\mathbb{F}(2^r)$.
- To correspond to this, we also use parity check matrices whose elements belong to the same Galois field, i.e., $H_i \in \mathbb{F}(2^a)^{m_i \times N}, 1 \leq i \leq b$ and $H_{b+1} \in \mathbb{F}(2^r)^{m_{b+1} \times N}$.
- The modified total IR rate is given by:

$$r = \sum_{i=1}^{b} a(1 - E_i)\frac{N - m_i}{N} + r(1 - E_{b+1})\frac{N - m_{b+1}}{N}$$

where E_i are the frame error rates for layer i. Clearly, the total IR rate depends on the rates $\frac{N-m_i}{N}$ used for every layer.

Preliminaries - Multi Layer Coding (MLC) Scheme

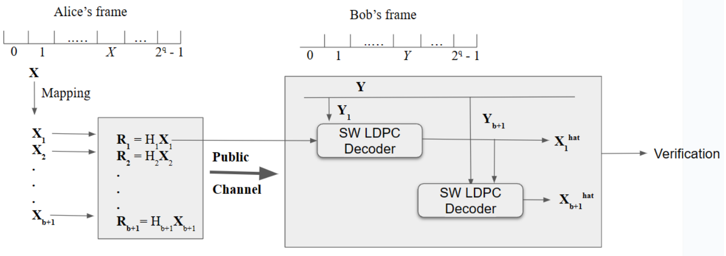


Figure: Non Binary MLC Protocol with Slepian Wolf Coding Scheme

Preliminaries - QKD Public Channel Model [7]

- Standard AWGN and BSC channels do not model the QKD channel effectively.
- Using a better channel transition probability will help LDPC codes in performing effective decoding.
- I model the QKD channel using a generative modelling approach and show that the channel can be modelled as a mixture of 2 Gaussians and a Uniform distribution given by the equation:

$$P_{Y|X}(y|x) = c\left(e^{\frac{(y-x-\mu_1)^2}{2\sigma_1^2}} + \alpha e^{(\frac{y-x-\mu_2)^2}{2\sigma_2^2}}\right) + \beta$$

Contributions - Progressive Edge Growth (PEG) Construction [8]

- NB-LDPC codes rely on sparsity of H to perform low complexity decoding by avoiding cycles of length 4 and 6.
- This strategy does not work well for shorter block lengths which are more prone to smaller girth.
- The PEG algorithm is a deterministic algorithm that works to construct large girth graphs by placing edges iteratively as best as possible by maximizing the local girth of the sub-graph at that stage.
- I incorporated the PEG algorithm, taken from here, into the pipeline.

Contributions - Mappings

- For MLC scheme, when mapping a symbol into bits, the simplest way is to use the binary mapping $GF(2^q) \to GF(2)^q$.
- However, is there a better arbitrary mapping that can be used to convert a symbol into a binary string to obtain higher key rates?
- For a Galois field of size 2^q, there are $2^q!$ number of different permutations which is infeasible.
- Simulated annealing [9] is a search algorithm that makes local changes to the state of the system to escape the local optima and approximate the global optima in discrete optimization problems.
- We use a modified version of simulated annealing for our problem. The different possible permutations of mappings in this problem correspond to the states of the system and local search is performed by swapping the binary representation of any two symbols in a mapping to find a neighbour.

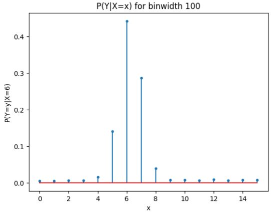
Contributions - Mappings

Algorithm Simulated Annealing Algorithm

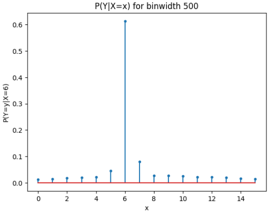
procedure SIMULATED ANNEALING($S_{init}, q, f_{th}, f_{de}, iter$)
  $S \leftarrow S_{init}$ ▷ Initial mapping (typically binaryz)
  $K \leftarrow$ GET IR RATE(S)
  $S_{best} \leftarrow S$
  for $T \leftarrow iter$ to 0 **do** ▷ loop over epoch
    $S_{new} \leftarrow$ SWAP(S) ▷ Find new mapping permutation
    $K_{new} \leftarrow$ GET IR RATE(S_{new})
    if $K_{new} - K \geq f_{th}$ **then** ▷ Better mapping found
      $S \leftarrow S_{new}$
      $K \leftarrow K_{new}$
      if $K_{new} > K_{best}$ **then** ▷ Best mapping at this point found
        $K_{best} = K_{new}$
        $S_{best} = S_{new}$
      end if
    else if $e^{\frac{\Delta K - f_{th}}{T \times f_{de}}} > rand(0,1)$ **then** ▷ Exploration
      $S \leftarrow S_{new}$
      $K \leftarrow K_{new}$
    end if
  end for
  return S_{best}, K_{best} ▷ Return the best mapping and its IR rate
end procedure

Contributions - QKD Channel Modelling

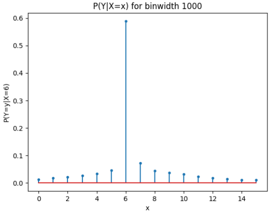
- QKD experiments relied on working with standard channels such as the AWGN channel or the BSC channel for the public channel model - not good approximations.
- Having a good general model that can be used to extrapolate to different binwidths as required, will be very useful.



(a) $P(Y|X=6)$ for Binwidth 100 ps



(b) $P(Y|X=6)$ for Binwidth 500 ps



(c) $P(Y|X=6)$ for Binwidth 1000 ps

Figure: Sample PMFs $P(Y|X=6)$ for binwidth 100 ps, 500 ps, and 1000 ps

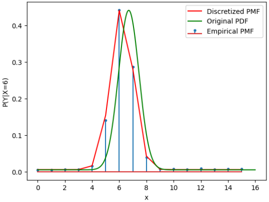
Contributions - QKD Channel Modelling

- I used a generative modelling approach, where I created a continuous distribution whose inputs were the parameters of the distribution and then discretized it to create a corresponding discrete probability distribution that resembled the empirical PMF as closely as possible.
- I used the mean absolute distance to calculate the error between the empirical PMF $P(Y|X=x)$ obtained from the dataset and the generated PMF $P(\tilde{Y}|X=x)$. Since $Y \in \{0, 1, 2, ..., 2^q - 1\}$, the average of the mean absolute differences $\tilde{M}$ over all values of Y is calculated as:

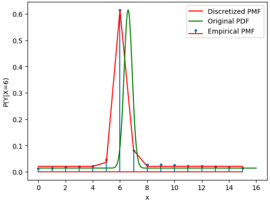
$$\tilde{M} = \frac{1}{2^q} \sum_{y=0}^{2^q - 1} |P(\hat{Y} = y|X = x, A) - P(Y = y|X = x)|$$

Contributions - QKD Channel Modelling

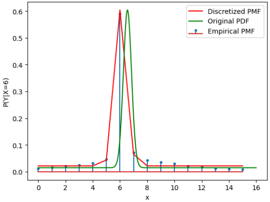
- A reasonable starting point to modelling is the basic assumption that the distribution is a combination of a uniform distribution, that models system noise, and another distribution (or a combination of distributions) that models photon jitter.
- A single Gaussian + Uniform model (with or without skew did not perform well).



(a) $P(Y|X=6)$ for Binwidth 100 ps



(b) $P(Y|X=6)$ for Binwidth 500 ps



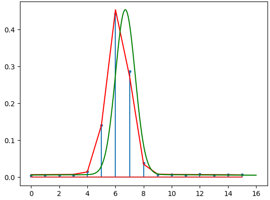
(c) $P(Y|X=6)$ for Binwidth 1000 ps

Figure: Sample PMFs $P(Y|X=6)$ for binwidth 100 ps, 500 ps, and 1000 ps along with a discretized model fit to the empirical PMF (in red) and the underlying continuous distribution (in green).

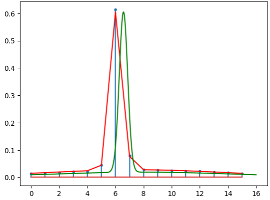
Contributions - QKD Channel Modelling

- We observe from figure 4 that there is a distinct second Gaussian distribution that can be used to model higher binwidths.

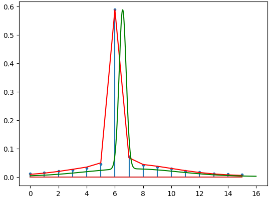
$$P_{Y|X}(y|x) = c\left(e^{\frac{(y-x-\mu_1)^2}{2\sigma_1^2}} + \alpha e^{\left(\frac{y-x-\mu_2)^2}{2\sigma_2^2}}\right) + \beta$$



(a) $P(Y|X=6)$ for Binwidth 100 ps



(b) $P(Y|X=6)$ for Binwidth 500 ps

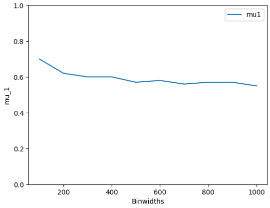


(c) $P(Y|X=6)$ for Binwidth 1000 ps

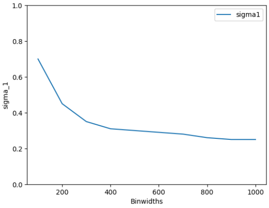
Figure: Sample PMFs $P(Y|X=6)$ for binwidth 100 ps, 500 ps, and 1000 ps along with a discretized model fit to the empirical PMF (in red) and the underlying continuous distribution (in green).

Contributions - QKD Channel Modelling

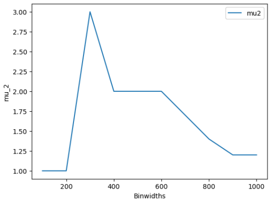
Finding a function across binwidths for each parameter:



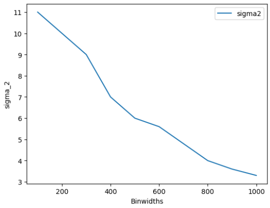
(a) Trend in μ_1



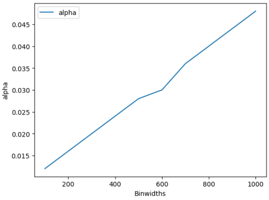
(b) Trend in σ_1



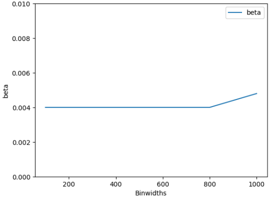
(c) Trend in μ_2



(a) Trend in σ_2



(b) Trend in α



(c) Trend in β

Contributions - QKD Channel Modelling

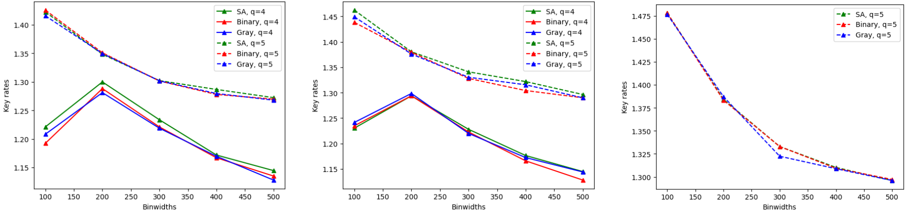
To obtain a general model for the channel that follows these trends, I performed leave one out cross validation (LOOCV) for each parameter.

Binwidth	μ_1	σ_1	μ_2	σ_2	α	β
100	0.616667	0.586812	2.230556	10.105556	0.011778	0.003767
200	0.638065	0.478297	2.025806	9.373387	0.015790	0.003877
300	0.626324	0.355109	1.466912	8.567647	0.019794	0.003962
400	0.611250	0.303912	1.658333	7.868750	0.023792	0.004033
500	0.601149	0.281018	1.629054	6.972973	0.027784	0.004100
600	0.586757	0.270975	1.590541	6.031757	0.032000	0.004168
700	0.576250	0.267032	1.583333	5.145833	0.035750	0.004242
800	0.560662	0.268927	1.598529	4.262500	0.039721	0.004329
900	0.543952	0.270475	1.637903	3.241935	0.043677	0.004310
1000	0.532500	0.270145	1.650000	1.994444	0.047611	0.004178

Table: Results of Leave One Out Cross Validation to Obtain the Parameters for each Binwidth when using 2 Gaussians + Uniform

Results - Simulated Annealing

- All simulations were performed on a Galois Field of size 2^4 and 2^5 on binwidths from 100 to 500.
- We see that the mapping obtained from SA performs slightly better for $q = 4, a = 2$, but about the same for the others.

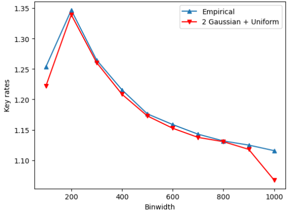


(a) a = 2 (b) a = 3 (c) a = 4

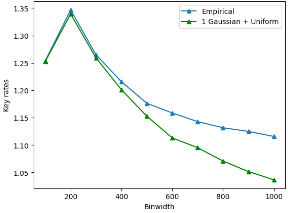
Figure: Simulation results of best mapping decoding results obtained using SA search

Results - QKD Channel Modelling

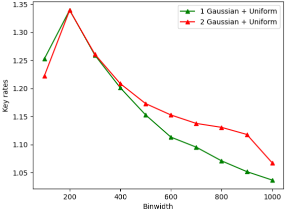
- All simulations were performed on a Galois Field of size 2^4.
- We see that while the 1 Gaussian + Uniform channel performs well for lower binwidths, but it performs significantly worse on higher binwidths.



(a) Empirical vs 2 Gaussian + Uniform



(b) Empirical vs 1 Gaussian + Uniform



(c) 2 Gaussian vs 1 Gaussian

Figure: Simulation results (key rates) of QKD Channel Modelling across binwidths

References

M. Davey and D. MacKay, "Low density parity check codes over gf(q)," in *1998 Information Theory Workshop (Cat. No.98EX131)*, pp. 70–71, 1998.

R. Gallager, "Low-density parity-check codes," *IRE Transactions on Information Theory*, vol. 8, no. 1, pp. 21–28, 1962.

T. Zhong, H. Zhou, R. D. Horansky, C. Lee, V. B. Verma, A. E. Lita, A. Restelli, J. C. Bienfang, R. P. Mirin, T. Gerrits, *et al.*, "Photon-efficient quantum key distribution using time–energy entanglement with high-dimensional encoding," *New Journal of Physics*, vol. 17, no. 2, p. 022002, 2015.

C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," *Theoretical Computer Science*, vol. 560, p. 7–11, Dec. 2014.

D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, 1973.

H. Zhou, L. Wang, and G. Wornell, "Layered schemes for large-alphabet secret key distribution," in *2013 Information Theory and Applications Workshop (ITA)*, pp. 1–10, 2013.

S. Yang, *Application-Driven Coding Techniques: From Cloud Storage to Quantum Communications*. University of California, Los Angeles, 2021.

X.-Y. Hu, E. Eleftheriou, and D. Arnold, "Regular and irregular progressive edge-growth tanner graphs," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 386–398, 2005.

F. W. Glover and G. A. Kochenberger, *Handbook of metaheuristics*, vol. 57. Springer Science & Business Media, 2006.