Combinatorial Designs and the Analysis of their Application to Channel Estimation

Philipp Grabenweger[1], Christoph Pacher[1] and Dimitris E. Simos[2]
[1] AIT Austrian Institute of Technology
[2] SBA Research

Algebraic Combinatorics and Applications
ALCOMA 2015, Kloster Banz, Germany
March 16, 2015

Outline of the Talk

Channel Estimation

- LDPC Codes
- Motivation/Previous Work
- Exact Mean and Variance of Syndrome Weight

secure sba-research.org

Outline of the Talk

Channel Estimation
- LDPC Codes
- Motivation/Previous Work
- Exact Mean and Variance of Syndrome Weight

Connections with Combinatorial Designs
- Regularity Conditions for Tanner Graphs
- Resolvable Graph Designs
- Comparison of Theory with Simulation

Outline of the Talk

Channel Estimation
- LDPC Codes
- Motivation/Previous Work
- Exact Mean and Variance of Syndrome Weight

Connections with Combinatorial Designs
- Regularity Conditions for Tanner Graphs
- Resolvable Graph Designs
- Comparison of Theory with Simulation

Research Problems

secure
sba-research.org

LDPC Codes

Binary Linear Error-Correcting Code

- Can be defined by means of its parity-check matrix H
- The null-space of the $m \times n$ binary parity-check matrix H defines the set of all codewords: $\mathcal{C} = \left\{ \boldsymbol{x} \in \{0,1\}^n : \boldsymbol{x}\mathrm{H}^T = \mathbf{0} \right\}$

LDPC Codes

- If H is sparse the code is called Low-Density Parity-Check (LDPC) code
- Codes with constant row weight d (called check node degree) are called check-regular
- Codes with constant row weight d and constant column weight d_v (called variable node degree) are called regular

Channel Estimation

Channel State Information (CSI)

- CSI at the receiver, i.e. knowledge of parameters like the crossover probability ρ or the signal-to-noise ratio, is often assumed when discussing forward error corection
- Of interest for prediction of successful decoding attempts

Estimating Channel Parameters

- **The problem:** Analyze estimation of CSI based on the syndrome of a linear code
- **Syndrome computation:** The receiver performs a hard decision on the receiver signal, thereby converting the channel to a binary symmetric channel (BSC)
- **CSI of original channel:** Derived from the estimated crossover probability of this BSC

Channel Estimation in Other Domains

Codes for Information Reconciliation in Quantum Cryptography

1. 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)
2. 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
3. If the quantum bit error rate ρ has not been too large (otherwise the decoder fails), Bob can reconstruct $\boldsymbol{x}_A$ from $\boldsymbol{x}_B$ and $\boldsymbol{s}_A$

Can Bob use the Syndrome for Further Purposes?

- **Yes**, Bob can calculate the syndrome $\boldsymbol{s}$ of the error word as $\boldsymbol{s} := \boldsymbol{e}\mathrm{H}^T = (\boldsymbol{x}_A \oplus \boldsymbol{x}_B)\,\mathrm{H}^T = \boldsymbol{s}_A \oplus \boldsymbol{x}_B\mathrm{H}^T$
- and estimate the bit error rate before he starts decoding!

Related Work on Channel Estimation

Recent Approaches for Channel Estimation

- CSI Estimation from LDPC Codes (Lechner and Pacher, 2013)
- Estimation of the bit error probability of received packets (Chen et al., 2012)
- Estimation for Quantum Error Correcting Codes (Fujiwara et al., 2013)

Plan of this Talk

1. Exploit regularity conditions of LDPC Codes used for channel parameters estimation via structured classes of combinatorial designs
2. Optimize the parameters of combinatorial designs in terms of bit error estimators (analytical results vs. numerical simulations)

Mean and Variance of Syndrome Weight

Computation of Mean of Syndrome Weight (Distribution)

Let I_i be the set of the indices of those variable nodes adjacent to check node i (e.g. $s_i = \bigoplus_{j \in I_i} e_j$), and let $w = \mathrm{wt}(\boldsymbol{s})$ be the syndrome weight:

- $w = \sum_{i=1}^{m} s_i$ (weight), $\mathbb{E}[w] = \sum_{i=1}^{m} \mathbb{E}[s_i]$ (expectation value of syndrome weight = the expectation value of the sum of syndrome bits)
- $\mathbb{E}[s_i] = P(s_i = 1) = \sum_{\substack{k \in \mathbb{N}_{\mathrm{odd}} \\ 0 \le k \le |I_i|}} \binom{|I_i|}{k} \rho^k (1-\rho)^{|I_i|-k} = \frac{1-(1-2\rho)^{|I_i|}}{2} =: f_{|I_i|}(\rho)$ (sum over all error-patterns with an odd number of 1s)

Computing the Variance of Syndrome Weight (Distribution)

Same assumptions as before:

- $\mathbb{V}[w] = \mathbb{E}\left[w^2\right] - (\mathbb{E}[w])^2 = \mathbb{E}\left[w^2\right] - \left(\sum_{i=1}^{m} \mathbb{E}[s_i]\right)^2$
- $\mathbb{E}\left[w^2\right] = \mathbb{E}\left[\left(\sum_{i=1}^{m} s_i\right) \cdot \left(\sum_{j=1}^{m} s_j\right)\right] = \sum_{i,j=1}^{m} \mathbb{E}[s_i \cdot s_j]$
- $\mathbb{E}[s_i \cdot s_j] = P(s_i = 1 \wedge s_j = 1) =$
  $f_{|I_i \setminus I_j|}(\rho) f_{|I_j \setminus I_i|}(\rho) + f_{|I_i \cap I_j|}(\rho) \left[1 - f_{|I_i \setminus I_j|}(\rho) - f_{|I_j \setminus I_i|}(\rho)\right]$

secure sba-research.org

Exact Mean and Variance of Syndrome Weight

The case of LDPC Codes

For a check-regular LDPC code with check degree (row weight) d:

- $|I_i| = |I_j| = d$, $|I_i \backslash I_j| = |I_j \backslash I_i| = d - c_{ij}$, $c_{ij} := |I_i \cap I_j| = c_{ji}$ (number of overlaps of two rows in the parity-check matrix H of the LDPC code)
- $\mathbb{E}\left[s_i \cdot s_j\right] = f_{d-c_{ij}}^2(\rho) + f_{c_{ij}}(\rho)\left(1 - 2f_{d-c_{ij}}(\rho)\right) = f_d(\rho) - \frac{1}{2} f_{2(d-c_{ij})}(\rho)$
- $\mathbb{V}\left[w\right] = \sum_{i,j=1}^{m} \mathbb{E}\left[s_i \cdot s_j\right] - (m\, f_d(\rho))^2 = \frac{1}{2}\left(m^2 f_{2d}(\rho) - \sum_{i,j=1}^{m} f_{2(d-c_{ij})}(\rho)\right)$

Exact Form for the Variance of Syndrome Weight

- Let $g_k := \left|\{(i,j) \in \{1, \ldots, m\}^2 | c_{ij} = k\}\right|$ be the number of all ordered pairs of check nodes which share exactly k variable nodes
- Assuming that all rows of H are distinct ($\implies g_d = m$):

$$\boxed{\mathbb{V}\left[w\right] = \underbrace{m\, f_d(\rho)\left(1 - f_d(\rho)\right)}_{\substack{:=\mathbb{V}[w]_{i.i.d} \\ \text{variance for i.i.d. syndrome bits}}} + \underbrace{\frac{1}{2}\sum_{k=1}^{d-1} g_k \left(f_{2d}(\rho) - f_{2(d-k)}(\rho)\right)}_{\text{correction term for dependent syndrome bits}}}$$

Regularity Conditions for LDPC Codes

Minimum of the Variance of Syndrome Weight

The $\mathbb{V}[w]$ attains its minimum if $g_k = 0$ for $2 \leq k \leq d-1$

Tanner Graph

1. The Tanner graph of an $m \times n$ parity check matrix H is a bipartite graph consisting of n variable nodes (vertices) (each corresponding to one column of H) and m check nodes (each corresponding to one row $=$ check equation of H)
2. An edge connects check node i with variable node j iff the variable node is checked by the corresponding parity-check equation, i.e. if $\mathrm{H}_{ij} = 1$

Cycles for Tanner Graphs of LPDC Codes

- If there exist (at least) two check nodes that share two variable nodes we have $g_2 \geq 1$
- In the Tanner graph of the LDPC code these four nodes will form a 4-cycle

LDPC Codes from Combinatorial Designs

Parity-Check Matrices as Incidence Matrices

The parity-check matrix of a regular LDPC code can also be regarded as a sparse incidence matrix of an incomplete block design (IBD)

Incomplete Block Designs

- An IBD of size (v, k, r) is an arrangement of v points set out in blocks of size k $(< v)$ such that each point occurs in exactly r blocks
- The number of blocks will be b, where $bk = vr$

Correspondence of IBDs and LDPC Codes

- The incidence matrix $\mathcal{D}$ of an IBD(v, k, r) has size $v \times b$ and constant row and column weights equal to r and k, respectively
- In that case, the blocks of the design form the columns of the $v \times b$ parity-check matrix H of a regular LDPC code with $d = r$ and $d_v = k$
- Blocks $\equiv$ variable nodes and points $\equiv$ check nodes

Example of an IBD with 4-cycles

An IBD($v = 4, k = 2, r = 4$)

- Has $\frac{4 \times 4}{2} = 8$ blocks
- $\mathcal{B} = \{\{0,3\},\{2,1\},\{0,2\},\{1,3\},\{0,1\},\{2,3\},\{0,1\},\{2,3\}\}$

The 4×8 Incidence Matrix $\mathcal{D}$ of an IBD(4, 2, 4)

$$\mathcal{D} = \left(\begin{array}{c|cccccccc} & v_0 & v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 \\ \hline c_0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ c_1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ c_2 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ c_3 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \end{array}\right)$$

Cycles in the World of Matrices

A 4-cycle in a Tanner graph is equivalent to a 2×2 all-one submatrix in the incidence matrix of the design (or the parity-check matrix of the corresponding LDPC code)

Regular Graph Designs

Concurrence Matrix of a Design

- The concurrence λ_{ij} of points i and j is r if $i = j$ and otherwise is the number of blocks in which i and j both occur
- The matrix $\Lambda = \mathcal{D}\mathcal{D}^T$ is called the concurrence matrix of the design
- **Remark:** $\lambda_{ij} = c_{ij}$ (number of overlaps of rows i and j in the corresponding parity-check matrix H)

Regular Graph Design (RGD)

An RGD is an IBD(v, k, r) where any two points belong to either λ or $\lambda + 1$ common blocks, for some constant λ and is denoted as RGD$_\lambda(v, k, r)$

Example of an IBD with 4-cycles (Cont.)

Cycles and Concurrences

Having 4-cycles in the Tanner graph of an LDPC code imply that the corresponding IBD has a concurrence equal to 2

The Concurrence Matrix Λ of an IBD$(4, 2, 4)$

$$\Lambda = \mathcal{D}\mathcal{D}^T = \left(\begin{array}{c|cccc} \text{points} & 0 & 1 & 2 & 3 \\ \hline 0 & 4 & 2 & 1 & 1 \\ 1 & 2 & 4 & 1 & 1 \\ 2 & 1 & 1 & 4 & 2 \\ 3 & 1 & 1 & 2 & 4 \end{array} \right)$$

secure
sba-research.org

Minimal Variance obtained from RGDs

RGDs and Tanner Graphs

An RGD with $\lambda = 0$ has the property that any two points occur in at most one block, which implies that the corresponding Tanner graph of the code is thus without 4-cycles

Minimal Variance of Syndrome Weight obtained from RGDs

Any code with minimal variance $\mathbb{V}[w]$ (that has $g_k = 0$ for $2 \leq k \leq d-1$) must be free of 4-cycles and is equivalent to an RGD with $\lambda = 0$!

secure
sba-research.org

Comparison of Theory with Simulation

- $m = 500$, $n = 1000$, $d = 6$, $d_v = 3$
- green: simulated bin heights, red: bin heights for $\mathcal{N}\left(\mathbb{E}\left[w\right], \mathbb{V}\left[w\right]_{i.i.d.}\right)$
- blue: bin heights for $\mathcal{N}\left(\mathbb{E}\left[w\right], \mathbb{V}\left[w\right]\right)$ (where $\mathcal{N}(\mu, \sigma^2)$ is the normal distribution)

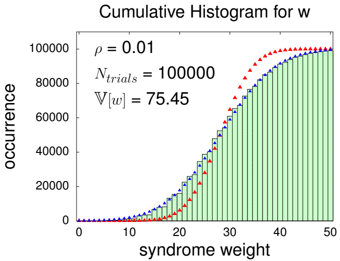

Figure : Simulation for $\text{RGD}_0(500, 3, 6)$ = regular LDPC code w/o 4-cycles (and w/o 6-cycles) (left) vs Simulation for $\text{IBD}(500, 3, 6)$ = regular LDPC code with 1000 4-cycles (right)

Comparison of Theory with Simulation

- $m = 500$, $n = 1000$, $d = 6$, $d_v = 3$
- green: simulated bin heights, red: bin heights for $\mathcal{N}\left(\mathbb{E}\left[w\right], \mathbb{V}\left[w\right]_{i.i.d.}\right)$
- blue: bin heights for $\mathcal{N}\left(\mathbb{E}\left[w\right], \mathbb{V}\left[w\right]\right)$ (where $\mathcal{N}(\mu, \sigma^2)$ is the normal distribution)

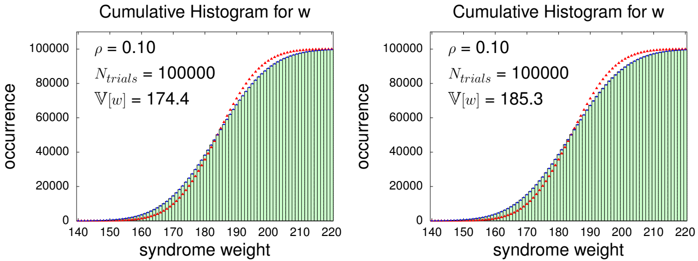

Figure : Simulation for $\mathrm{RGD}_0(500, 3, 6)$ = regular LDPC code w/o 4-cycles (and w/o 6-cycles) (left) vs Simulation for $\mathrm{IBD}(500, 3, 6)$ = regular LDPC code with 1000 4-cycles (right)

Research Problems

Related to Designs

Are there any other classes of designs suitable for CSI estimation?

Related to Codes

Relaxing regularity conditions for LDPC codes can lead to better estimators (obtained from designs)?

Summary

Highlights

1. We showed how combinatorial designs can be used for channel estimation
2. We demonstrated that regular graph designs correspond to codes with minimal variance for estimating the syndrome weight

Summary

Highlights

1. We showed how combinatorial designs can be used for channel estimation
2. We demonstrated that regular graph designs correspond to codes with minimal variance for estimating the syndrome weight

Future Work

Solve (some) of the research problems..!

secure
sba-research.org

References

G. Lechner, C. Pacher, Estimating Channel Parameters from the Syndrome of a Linear Code, IEEE Communications Letters 17, 2148-2151 (2013).

W. D. Wallis, Regular graph designs, Journal of Statistical Planning Inference 51, 273–281 (1996).

Y. Fujiwara, A. Gruner, P. Vandendriessche, High-rate quantum low-density parity-check codes assisted by reliable qubits, arXiv preprint arXiv:1309.5587 (2013).

secure
sba-research.org

Questions - Comments

Thanks for your Attention!

dsimos@sba-research.org