June 22, 2015

Non-Binary LDPC Erasure Codes with Separated Low-Degree Variable Nodes

Giuliano Garrammone

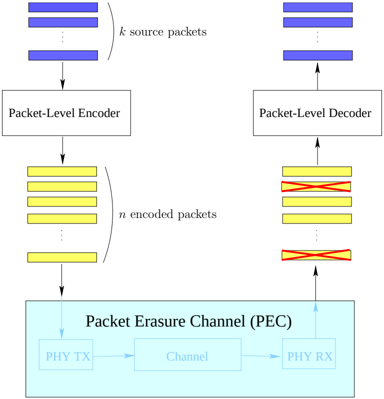
German Aerospace Center (DLR)
Institute of Communications and Navigation

Motivation (I)

- Error correcting codes are nowadays a fundamental component of modern communication networks.
- Coding at the upper-layers of the communication protocol is a simple technique to cope with packet losses.
- Applications of packet-level coding (erasure coding) within SATCOM:
  - Multicasting/broadcasting in land mobile satellite services: cope with long fading events (DVB-SH, DVB-RCS2).
  - Telemetry services in deep-space communication: reduce the average delay.
  - Free-space optical communication: compensate turbulence.

Principle of Packet-Level Coding

- k source packets (L bits), n encoded packets (L bits).
- (n, k) code on $\mathbb{F}_q$.
- CRC and error correcting code at physical layer.
- Erasures: packets whose CRC has failed after physical layer decoding.
- PEC: a packet is either correctly received or lost (erased).



Motivation (II)

- Typical erasure codes: binary low-density parity-check (LDPC) codes, Reed-Solomon codes, fountain codes (rate-less).
- Binary LDPC codes:
  - Poor performances for short codeword lengths.
  - Low decoding complexity $\mathcal{O}(n)$.
- RS codes:
  - Good performances for short codeword lengths.
  - Decoding complexity higher than $\mathcal{O}(n)$.
- Non-binary LDPC codes: good performances for short codeword lengths (AWGN).
- Non-binary LDPC erasure codes can be a flexible solution to bridge:
  - Good performances for short codeword lengths.
  - Low decoding complexity.

Outline

Outline

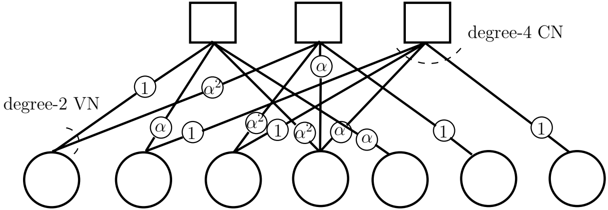
1. Introduction
2. Ensemble with Separated Variable Nodes
3. Weight Distribution and Its Growth Rate
4. Code Design for the q-ary Erasure Channel
5. Conclusion

Non-Binary Low-Density Parity-Check Codes

- Parity-check matrix:

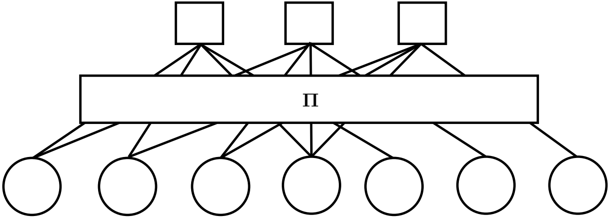
$$\mathbf{H} = \begin{bmatrix} 1 & \alpha & 0 & \alpha^2 & \alpha & 0 & 0 \\ \alpha^2 & 0 & \alpha^2 & \alpha & 0 & 1 & 0 \\ 0 & 1 & 1 & \alpha & 0 & 0 & 1 \end{bmatrix}$$

- Tanner graph:



- Degree distribution pair:
$\lambda(x) = \sum_{i=1}^{\mathrm{d}_{v,\max}} \lambda_i x^{i-1}$, $\rho(x) = \sum_{i=1}^{\mathrm{d}_c} \rho_i x^{i-1}$
λ_i, ρ_i: fractions of edges connected to degree-i VNs, CNs.

Non-Binary Unstructured LDPC Code Ensembles



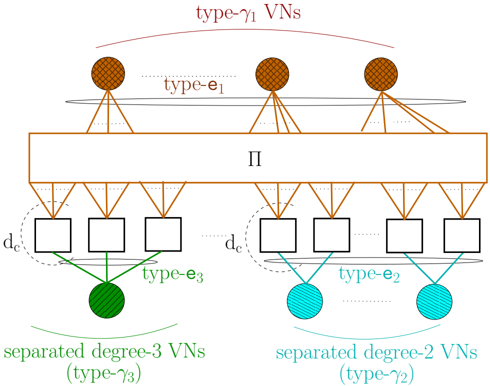
- Usually, we consider sets, or *ensembles*, of LDPC codes, fulfilling $(\lambda(x), \rho(x))$.
- The design rate of the ensemble is $R = 1 - \frac{\int_0^1 \rho(x)dx}{\int_0^1 \lambda(x)dx}$.
- Non-binary *unstructured* ensemble: all possible edge labelings from $\mathbb{F}_q \setminus \{0\}$ (uniform probability) and all possible edge permutations **Π**.

Structured LDPC Code Ensembles

- If not all edge permutations are allowed: *structured* ensemble.
- We focus on a structured LDPC code ensemble.
- A similar ensemble was heuristically introduced by MacKay (binary) [1].
- An ensemble similar to the one of MacKay was analyzed by C. Di (binary) [2].
- We extend the ensemble of MacKay, we provide an analytical analysis of the extended ensemble on non-binary Galois fields.
- This is the ensemble from which the progressive edge-growth (PEG) algorithm picks the codes.

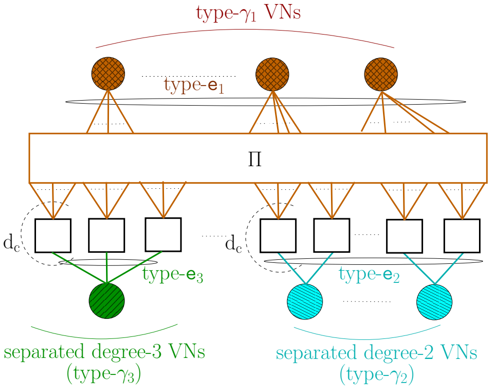
Outline

Ensemble with Separated Variable Nodes: Definition



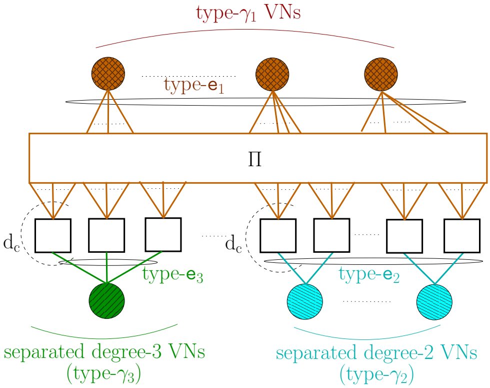
- The degree-2 VNs are all separated (type-e_2 edges).

Ensemble with Separated Variable Nodes: Definition



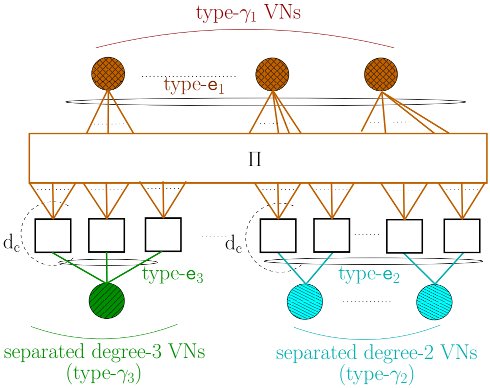
- The degree-2 VNs are all separated (type-e_2 edges).

Ensemble with Separated Variable Nodes: Definition



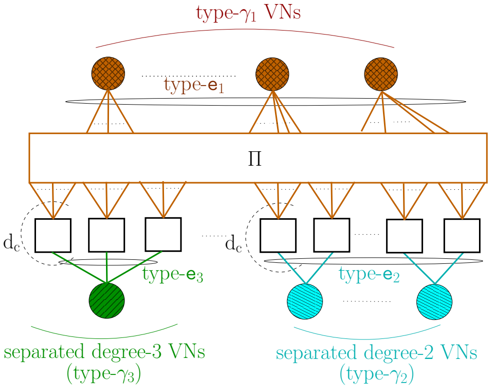
- Some of the degree-3 VNs are separated (type-e_3 edges).

Ensemble with Separated Variable Nodes: Definition



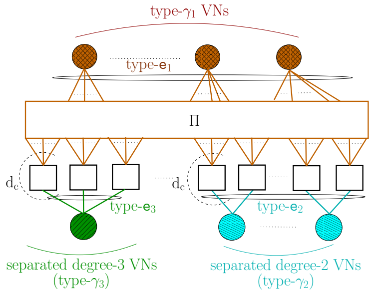
- Constant CN degree d_c.

Ensemble with Separated Variable Nodes: Definition



- All possible type-e_1 (brown) edge permutations Π and all possible edge labelings from $\mathbb{F}_q \setminus \{0\}$ (uniform probability).

Ensemble with Separated Variable Nodes: Notation



- n: number of VNs (codeword length in symbols from $\mathbb{F}_q$).
- m: number of CNs.
- V_2: number of degree-2 VNs (type γ_2).
- V_3^S: number of separated degree-3 VNs (type γ_3).
- $\tilde{V}_j$: number of degree-j VNs of type γ_1 (brown).

Outline

Codeword Weight Distribution

- The weight of a codeword is the number of its non-zero symbols.

Theorem 1 - $\mathbb{E}[A(\mathcal{C}, l)]$

The expected number of codewords of weight l for a code $\mathcal{C}$ picked from an ensemble with separated VNs (SVN ensemble) and distribution pair $(\lambda(x), \rho(x) = x^{\mathrm{d_c}-1})$ is

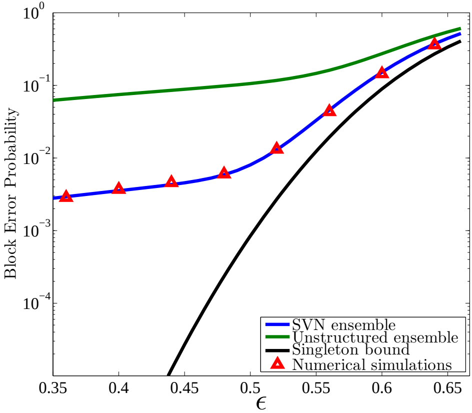
$$\mathbb{E}[A(\mathcal{C}, l)] = \sum_{\boldsymbol{l}: l_{\gamma_2}+l_{\gamma_3}+\sum_j \tilde{l}_j = l} \binom{V_2}{l_{\gamma_2}} \binom{V_3^S}{l_{\gamma_3}} \prod_j \binom{\tilde{V}_j}{\tilde{l}_j}$$

$$\times \frac{\mathrm{Coeff}\left(\left(N^-(z)\right)^{2l_{\gamma_2}+3l_{\gamma_3}} \left(N^+(z)\right)^{m-2l_{\gamma_2}-3l_{\gamma_3}}, z^{\sum_j \tilde{l}_j j} \right)}{(q-1)^{-(l_{\gamma_2}+l_{\gamma_3}+\sum_j \tilde{l}_j)} \binom{m(\mathrm{d_c}-1)}{\sum_j \tilde{l}_j j} (q-1)^{\sum_j \tilde{l}_j j + 2l_{\gamma_2}+3l_{\gamma_3}}}$$

with $\boldsymbol{l} = (\tilde{l}_3, \dots, \tilde{l}_{\mathrm{d}_{v,\max}}, l_{\gamma_2}, l_{\gamma_3})$ and $0 \le l_{\gamma_2} \le V_2, 0 \le l_{\gamma_3} \le V_3^S, 0 \le \tilde{l}_j \le \tilde{V}_j$. Further, $N^+(z)$ and $N^-(z)$ are univariate polynomials.

Expected Block Error Probability of a q-ary LDPC Code [3]

- E.g.: $(81, 27)$ structured vs. unstructured 4-ary LDPC codes, 4-ary EC.



Growth Rate of the Weight Distribution ($n \to \infty$)

- Normalized codeword weight: $\omega = l/n$. Thus, $0 \leq \omega \leq 1$.
- Growth rate: $G(\omega) = \lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}[A(\mathcal{C}, \lfloor \omega n \rfloor)]$

Theorem 2 - $G(\omega)$

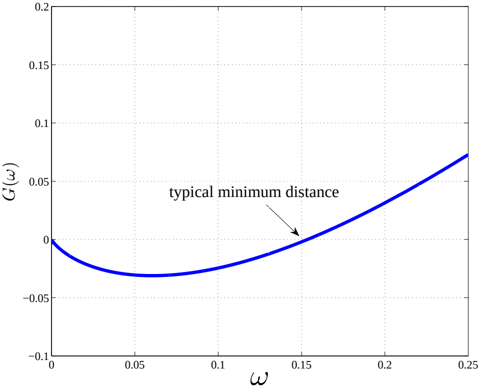
For an SVN ensemble with distribution pair $(\lambda(x), \rho(x) = x^{\mathrm{d_c}-1})$ the growth rate is

$$G(\omega) = \sum_{j=3}^{\mathrm{d}_{v,\max}} \tilde{\delta}_j \ln(B^{(j)}(x_0, y_{0,1})) + \sum_{i=2}^{3} \delta_i \ln(B^{(i)}(x_0, y_{0,\mathrm{s}})) - \omega \ln(x_0) + (1-R)\ln(N^+(z_0)) + \frac{\ln(1-\beta_1 t)}{t}$$

with $\tilde{\delta}_j = \tilde{V}_j/n$, $\delta_2 = V_2/n$, $\delta_3 = V_3^S/n$, $B^{(j)}(x,y) = 1 + (q-1)xy^j$ and $t = \frac{1}{(1-R)(\mathrm{d}_c - 1)}$. Further, $x_0, y_{0,1}, y_{0,\mathrm{s}}, z_0, \beta_1$ are the unique solutions to a 5×5 system of polynomial equations.

Example of Growth Rate Curve on $\mathbb{F}_4$

- $\lambda(x) = \frac{1}{5}x + \frac{4}{5}x^3$,
  $\rho(x) = x^4$.
- Typical minimum distance:
  $$\omega^* = \inf\{\omega > 0 : G(\omega) \geq 0\}.$$
- *Good growth rate behaviour*:
  - Large typical minimum distance.
  - Negative $G(\omega)$ for small ω.
- As $n \to \infty$,
  $$\mathbb{E}(A(\mathcal{C}, \lfloor \omega n \rfloor)) \to \exp\{nG(\omega)\}.$$



Growth Rate for Small (Normalized) Weight ω

Theorem 3 - $G(\omega)$ as $\omega \to 0$

The weight spectral shape of an SVN ensemble with distribution pair $(\lambda(x), \rho(x) = x^{\mathrm{d_c}-1})$ fulfills

$$G(\omega) = -\frac{3\xi_1\omega}{2} - \omega \ln\left(\frac{2(1-\xi_1)(\mathrm{d}_c-2)}{\nu_2(\mathrm{d}_c-1)\,(5\xi_1-2)}\right) + o(\omega)$$

with $\xi_1 = \frac{2}{5} + o(1)$ and $0 < \nu_2 \leq 1$.

- For small values of ω, $G(\omega)$ is always negative.
- The SVN ensemble is always characterized by a strictly positive typical minimum distance.

Typical Minimum Distances

- SVN ensemble vs. its unstructured counterpart:

$$\rho_2(x) = x^7$$

$$\lambda_2(x) = 0.1250x + 0.4951x^2 + 0.0254x^{12} + 0.2489x^{16} + 0.1056x^{17}.$$

- Typical minimum distances of the two ensembles:

Ensembles	$\mathbb{F}_2$	$\mathbb{F}_4$	$\mathbb{F}_{16}$	$\mathbb{F}_{64}$	$\mathbb{F}_{128}$	$\mathbb{F}_{256}$
SVN	0.0082	0.0178	0.0346	0.0408	0.0400	0.0373
Unstructured	0.0009	0.0017	0.0019	0.0009	0.0005	0.0003

- The SVN ensemble has much higher typical minimum distances.

Outline

DLR

Maximum A-Posteriori Decoding

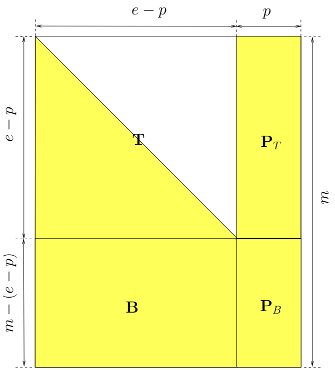
- Codeword $\mathbf{v}$ is transmitted, e erasures are introduced by the q-EC.
- MAP decoding: solve a linear system of m equations in e unknowns

$$\mathbf{H}_{\overline{\mathcal{I}}}\,\mathbf{v}_{\overline{\mathcal{I}}}^{T} = \mathbf{H}_{\mathcal{I}}\,\mathbf{v}_{\mathcal{I}}^{T}$$

- $\mathbf{v}_{\overline{\mathcal{I}}}$ and $\mathbf{v}_{\mathcal{I}}$: vector of e erased and $(n-e)$ received codeword symbols.
- $\mathbf{H}_{\overline{\mathcal{I}}}$ and $\mathbf{H}_{\mathcal{I}}$: sub-matrix composed of the corresponding columns of $\mathbf{H}$.
- The system can be solved with Gaussian elimination, complexity $\mathcal{O}(n^3)$.
- The sparseness of the parity-check matrix can be exploited in order to solve the system with reduced complexity [4].

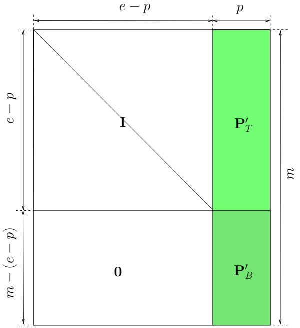
Efficient MAP Decoding for LDPC Codes (I)

- The matrix $\mathbf{H}_{\overline{\mathcal{I}}}$ is re-organized in an approximate lower triangular form.
- The codeword symbols associated with the right-most p columns: *pivots*.



Efficient MAP Decoding for LDPC Codes (II)

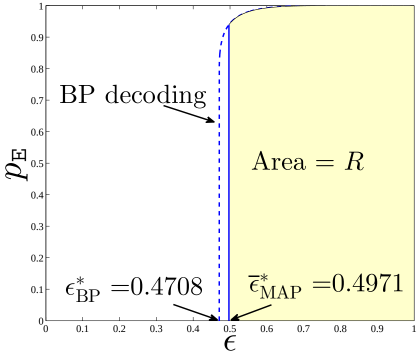
- Zeroing-out algorithm is applied (complexity $\mathcal{O}(n^2)$):



- Gaussian elimination to recover the p pivots (complexity $\mathcal{O}(p^3)$).
- BP decoding to recover the remaining unknowns (complexity $\mathcal{O}(n)$).
- The number of pivots can be controlled with a careful code design.

BP and MAP Decoding Thresholds [5] ϵ^* ($n \to \infty$)

- $p_E(\epsilon) \to 0, \forall \epsilon \leq \epsilon^*$.
- $p_E(\epsilon)$: average extrinsic symbol erasure probability at the output of a decoder.



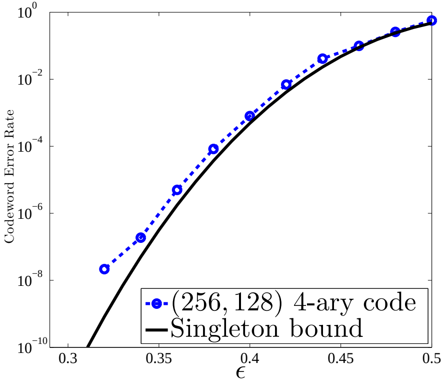
Design Guidelines, under MAP Decoding

- Design a code from an ensemble with separated variable nodes (SVNs).
- The code design in two phases (asymptotic and finite-length):
  1. Ensemble search with asymptotic tools.
  2. Construct finite-length parity-check matrix with girth optimization techniques.

- In practice:
  1. Search for SVN ensembles (degree distribution) with:
     - ★ MAP thresholds approaching the Shannon limit.
     - ★ BP thresholds close to the MAP threshold [6].
  2. Construct the finite-length parity-check matrix with PEG algorithm.

Code Performance, 4-EC (MAP decoding)

- Rate-$1/2$ SVN ensemble with $\bar{\epsilon}^*_{\text{MAP}} = 0.4971$, $\epsilon^*_{\text{BP}} = 0.4708$.
- Short 4-ary $(256, 128)$ LDPC code. $n = 256$ symbols of $\mathbb{F}_4$.



Decoding Complexity (MAP decoding)

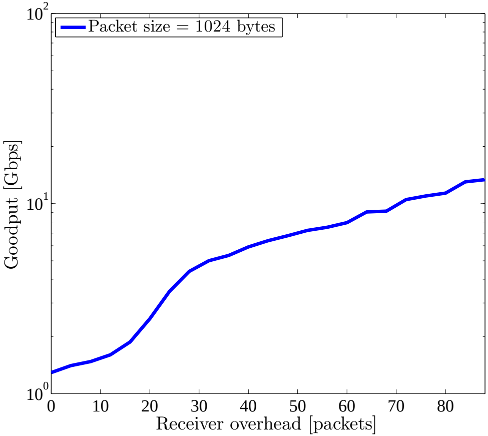
- Average number of pivots, $(256, 128)$ code, $n = 256$:

Ensembles	$\epsilon = 0.48$	$\epsilon = 0.46$	$\epsilon = 0.44$	$\epsilon = 0.42$	$\epsilon = 0.40$
SVN	6.96	5.01	3.35	1.74	0.75
Regular	22.95	18.99	15.59	11.62	7.62

- A $(256, 128)$ regular ($\mathrm{d}_v = 4$, $\mathrm{d}_c = 8$) code has been designed.
- The asymptotic thresholds of the regular ensemble are:
  - $\overline{\epsilon}^*_{\mathrm{MAP}} = 0.4977$.
  - $\epsilon^*_{\mathrm{BP}} = 0.3834$.
- The code from (irregular) SVN ensemble has much less pivots than the one from regular ensemble: less complexity.

Decoding Speed on the Packet Erasure Channel

- $(256, 128)$ code on $\mathbb{F}_4$ over the PEC. $n = 256$ packets of 1024 bytes.



Outline

Conclusion

- The design of non-binary LDPC erasure codes has been investigated.
- A promising ensemble of LDPC codes has been identified and analyzed in terms of:
  - Asymptotic thresholds.
  - Weight distribution.
  - Growth rate of the weight distribution.
- Codes from the ensemble designed and analyzed in terms of:
  - Performance (codeword error rate).
  - Decoding complexity.
- Codes from the ensemble provide excellent trade-offs between:
  - Waterfall performance, error-floor and decoding complexity.
- Thanks to their flexibility they can be used in many practical applications.

G. Garrammone, E. Paolini, B. Matuz, G. Liva, "*Non-Binary LDPC Erasure Codes with Separated Low-Degree Variable Nodes*", IEEE Transactions on Communications, submitted.

Thank you for your attention!
Questions?

G. Garrammone, E. Paolini, B. Matuz, G. Liva, "*Non-Binary LDPC Erasure Codes with Separated Low-Degree Variable Nodes*", IEEE Transactions on Communications, submitted.

Thank you for your attention!

Questions?

G. Garrammone, E. Paolini, B. Matuz, G. Liva, "*Non-Binary LDPC Erasure Codes with Separated Low-Degree Variable Nodes*", IEEE Transactions on Communications, submitted.

Thank you for your attention!
Questions?

References

1. MacKay et al., *Near Shannon limit performance of low density parity check codes*, Electronics Letters, vol. 32, no. 18, pp. 1645-1646, 1996.
2. C. Di, *Asymptotic and Finite-Length Analysis of Low-Density Parity-Check Codes*. PhD Thesis, E.P.F.L. Press, 2004.
3. Liva et al., *Bounds on the error probability of block codes over the q-ary erasure channel*, IEEE Trans. Commun., vol. 61, no. 6, pp. 2156-2165, Jun. 2013.
4. Burshtein et al., *An efficient maximum likelihood decoding of LDPC codes over the binary erasure channel*, IEEE Trans. Inf. Theory, vol. 50, no. 11, pp. 2837-2844, Nov. 2004.
5. Ashikhmin et al., *Extrinsic information transfer functions: Model and erasure channel properties*, IEEE Trans. Inform. Theory, vol. 50, no. 11, pp. 2657-2673, Nov. 2004.
6. Measson et al., *Maxwell construction: The hidden bridge between iterative and maximum a posteriori decoding*, IEEE Trans. Inform. Theory, vol. 54, no. 12, pp. 5277-5307, Dec. 2008.