Absorbing Set Analysis of Codes from Affine Planes

5th International Castle Meeting on Coding Theory and Applications
Vihula, Estonia

Katie Haymaker

August 29, 2017

VILLANOVA UNIVERSITY

Outline

1. Background and notation for absorbing sets
2. Define the classes of finite geometry codes that we consider in this talk
3. Results for codes based on finite Euclidean geometries
4. Further directions

Background: FG-LDPC codes

- Codes based on low-density parity-check (LDPC) matrices have been in the forefront of research in coding theory due to their low-complexity efficient decoders and near capacity performance at long block lengths.

Background: FG-LDPC codes

- Codes based on low-density parity-check (LDPC) matrices have been in the forefront of research in coding theory due to their low-complexity efficient decoders and near capacity performance at long block lengths.
- Constructing LDPC codes from finite geometries has a rich history (from Weldon 1957 to Kou, Lin, Fossorier 2000, and beyond).

Background: FG-LDPC codes

- Codes based on low-density parity-check (LDPC) matrices have been in the forefront of research in coding theory due to their low-complexity efficient decoders and near capacity performance at long block lengths.
- Constructing LDPC codes from finite geometries has a rich history (from Weldon 1957 to Kou, Lin, Fossorier 2000, and beyond).
- The resulting finite geometry LDPC codes are based on the incidence structure (often of points and lines) of finite Euclidean and projective geometries, the structure of which can be used to prove parameters of the code.

Motivation for absorbing sets

- Graphical substructures can contribute to the persistence of the error floor in the bit error rate (BER) curves for some families of LDPC codes under message-passing iterative decoders.

Motivation for absorbing sets

- Graphical substructures can contribute to the persistence of the error floor in the bit error rate (BER) curves for some families of LDPC codes under message-passing iterative decoders.
- Structural properties of the Tanner graph of the code are tied to pseudocodewords, absorbing sets, trapping sets, and stopping sets.

Motivation for absorbing sets

- Graphical substructures can contribute to the persistence of the error floor in the bit error rate (BER) curves for some families of LDPC codes under message-passing iterative decoders.
- Structural properties of the Tanner graph of the code are tied to pseudocodewords, absorbing sets, trapping sets, and stopping sets.
- While trapping sets depend on the decoder, absorbing sets are independent of the channel and are eventually stable under bit-flipping decoding.

Motivation for absorbing sets

- Graphical substructures can contribute to the persistence of the error floor in the bit error rate (BER) curves for some families of LDPC codes under message-passing iterative decoders.
- Structural properties of the Tanner graph of the code are tied to pseudocodewords, absorbing sets, trapping sets, and stopping sets.
- While trapping sets depend on the decoder, absorbing sets are independent of the channel and are eventually stable under bit-flipping decoding.
- In this talk we analyze the absorbing sets of the $\text{EG}(2, q)$ classes of finite geometry LDPC codes.

Finite geometries

Definition

- Let $\mathbb{F}_q$ denote the field with q elements.
- Let $V(m, q)$ be a vector space of rank m over $\mathbb{F}_q$.
- EG(m, q) denotes a geometry formed from the vectors in $V(m, q)$. The points, lines, planes, etc., of the geometry are cosets of the subspaces of $V(m, q)$ of rank $0, 1, 2, \ldots, m-1$.

Examples

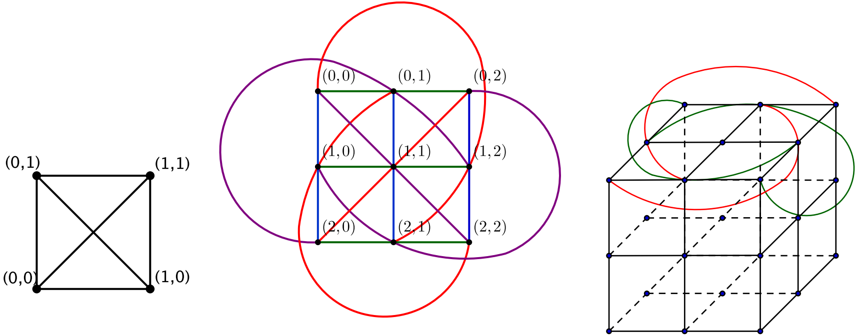


Figure: The finite Euclidean geometries EG(2, 2), EG(2, 3), and EG(3, 3).

Euclidean geometries

The m-dimensional finite Euclidean geometry $\mathrm{EG}(m, q)$ has the following parameters:

- There are q^m points and the number of lines is

$$\frac{q^{m-1}(q^m - 1)}{q - 1}.$$

$q(q+1)$ when $m = 2$

Euclidean geometries

The m-dimensional finite Euclidean geometry EG(m, q) has the following parameters:

- There are q^m points and the number of lines is
$$\frac{q^{m-1}(q^m - 1)}{q - 1}.$$
$q(q+1)$ when $m = 2$
- Each line contains q points and each point is on $\frac{q^m - 1}{q - 1}$ lines.
$q + 1$ lines when $m = 2$

Euclidean geometries

The m-dimensional finite Euclidean geometry EG(m, q) has the following parameters:

- There are q^m points and the number of lines is

$$\frac{q^{m-1}(q^m - 1)}{q - 1}.$$

$$q(q+1) \text{ when } m = 2$$

- Each line contains q points and each point is on $\dfrac{q^m - 1}{q - 1}$ lines.

  $q + 1$ lines when $m = 2$

- Any two points have exactly one line in common and any two lines either have one point in common or are parallel (i.e., have no points in common).

Euclidean geometries

The m-dimensional finite Euclidean geometry EG(m, q) has the following parameters:

- There are q^m points and the number of lines is
$$\frac{q^{m-1}(q^m - 1)}{q - 1}.$$
$q(q+1)$ when $m = 2$
- Each line contains q points and each point is on $\dfrac{q^m - 1}{q - 1}$ lines.
$q + 1$ lines when $m = 2$
- Any two points have exactly one line in common and any two lines either have one point in common or are parallel (i.e., have no points in common).
- A μ-dimensional subspace of a finite geometry is called a μ-*flat*.

Codes from EGs

- An LDPC code can be formed from an m-dimensional finite geometry by taking the incidence matrix of μ_1-flats and μ_2-flats, where $0 \leq \mu_1 < \mu_2 \leq m$.

Codes from EGs

- An LDPC code can be formed from an m-dimensional finite geometry by taking the incidence matrix of μ_1-flats and μ_2-flats, where $0 \leq \mu_1 < \mu_2 \leq m$.
- We use the notation $\mathcal{H}_{\mathrm{EG}}(m, p^s)$ to denote a parity-check matrix formed as the incidence matrix of points and lines in $\mathrm{EG}(m, p^s)$.

Codes from EGs

- An LDPC code can be formed from an m-dimensional finite geometry by taking the incidence matrix of μ_1-flats and μ_2-flats, where $0 \leq \mu_1 < \mu_2 \leq m$.
- We use the notation $\mathcal{H}_{\mathrm{EG}}(m, p^s)$ to denote a parity-check matrix formed as the incidence matrix of points and lines in $\mathrm{EG}(m, p^s)$.
- Points correspond to columns in $\mathcal{H}_{\mathrm{EG}}(m, p^s)$, and lines in the geometry correspond to rows in the parity-check matrix.

Codes from EGs

- An LDPC code can be formed from an m-dimensional finite geometry by taking the incidence matrix of μ_1-flats and μ_2-flats, where $0 \leq \mu_1 < \mu_2 \leq m$.
- We use the notation $\mathcal{H}_{\text{EG}}(m, p^s)$ to denote a parity-check matrix formed as the incidence matrix of points and lines in $\text{EG}(m, p^s)$.
- Points correspond to columns in $\mathcal{H}_{\text{EG}}(m, p^s)$, and lines in the geometry correspond to rows in the parity-check matrix.
- We focus on codes of the form $\mathcal{C}_{\text{EG}}(2, q)$. The minimum distance of such codes is given by: $d \geq q + 2$.

Notation: Tanner graphs

- Let $G = (V, W; E)$ be a bipartite Tanner graph corresponding to an LDPC code, where V and W denote the sets of variable nodes and check nodes, respectively, and E is the set of edges.

Notation: Tanner graphs

- Let $G = (V, W; E)$ be a bipartite Tanner graph corresponding to an LDPC code, where V and W denote the sets of variable nodes and check nodes, respectively, and E is the set of edges.
- For a subset S of V, let $G_S = (S, W_S; E_S)$ denote the subgraph induced by S in G.

Notation: Tanner graphs

- Let $G = (V, W; E)$ be a bipartite Tanner graph corresponding to an LDPC code, where V and W denote the sets of variable nodes and check nodes, respectively, and E is the set of edges.
- For a subset S of V, let $G_S = (S, W_S; E_S)$ denote the subgraph induced by S in G.
- In the induced subgraph, W_S is the set of constraint neighbors of S, and E_S is the set of edges between S and W_S.

Absorbing sets of Tanner graphs

Definition (Dolecek et al.)

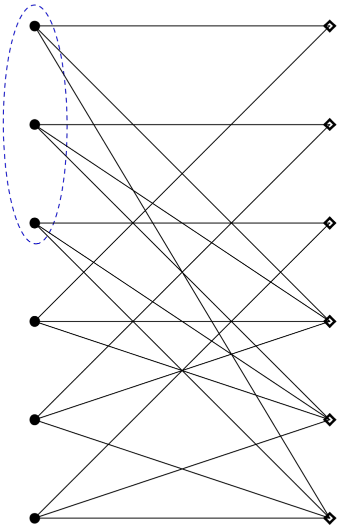
- An (a, b) **absorbing set** is a subset S of V where $|S| = a$, and there are b odd degree vertices in W_S, with the property that every vertex $v \in S$ has more even-degree than odd-degree neighbors in G_S. Let $\mathcal{O}(S)$ (resp., $\mathcal{E}(S)$) denote the vertices in W_S with odd degree (even degree) in G_S.
- If in addition, all variable nodes in $V \backslash S$ have strictly more neighbors in $W \backslash \mathcal{O}(S)$ than in $\mathcal{O}(S)$, then S is a **fully absorbing set**.
- An **elementary absorbing set** is an absorbing set in which all vertices in W_S have degree one or two in G_S.

Absorbing sets were introduced to qualitatively describe the convergent non-codeword state of the message passing algorithms, when transmitting across an AWGN channel.[1]

[1]Dolecek, L., Zhang, Z., Anantharam, V., Wainwright, M., and Nikolic, B.: Analysis of absorbing sets for array-based LDPC codes. In: IEEE Int. Conf. on Comm., 6261-6268, (2007)

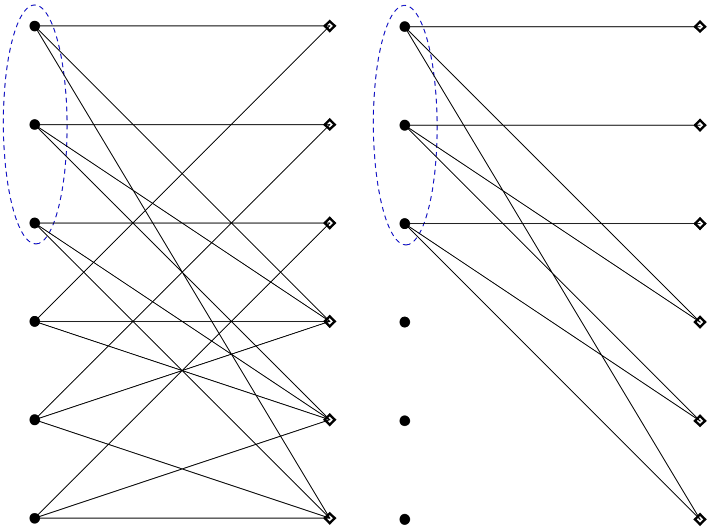
Example

The highlighted subset of V is a $(3, 3)$ (fully and elementary) absorbing set.



Example

The highlighted subset of V is a $(3,3)$ (fully and elementary) absorbing set.



Useful terminology

- Constraint nodes in the Tanner graph of $\mathcal{H}_{\mathsf{EG}}(2, q)$ correspond to lines in the geometry, so we say a line is *odd* if its corresponding constraint node is in $\mathcal{O}(S)$ with respect to some fixed set S.

Useful terminology

- Constraint nodes in the Tanner graph of $\mathcal{H}_{\mathrm{EG}}(2, q)$ correspond to lines in the geometry, so we say a line is *odd* if its corresponding constraint node is in $\mathcal{O}(S)$ with respect to some fixed set S.
- An *even line* has a constraint node in $\mathcal{E}(S)$.

Useful terminology

- Constraint nodes in the Tanner graph of $\mathcal{H}_{\mathsf{EG}}(2, q)$ correspond to lines in the geometry, so we say a line is *odd* if its corresponding constraint node is in $\mathcal{O}(S)$ with respect to some fixed set S.
- An *even line* has a constraint node in $\mathcal{E}(S)$.
- The *smallest* (a, b) absorbing set of the Tanner graph of a code is the size of the smallest a, and the corresponding smallest b for that given a, for which an absorbing set exists.

Useful terminology

- Constraint nodes in the Tanner graph of $\mathcal{H}_{\mathsf{EG}}(2, q)$ correspond to lines in the geometry, so we say a line is *odd* if its corresponding constraint node is in $\mathcal{O}(S)$ with respect to some fixed set S.
- An *even line* has a constraint node in $\mathcal{E}(S)$.
- The *smallest* (a, b) absorbing set of the Tanner graph of a code is the size of the smallest a, and the corresponding smallest b for that given a, for which an absorbing set exists.
- Small absorbing sets provide information that can be used to improve error-floor performance.

Lower bound on smallest absorbing sets

The following rephrases Lemma 1 of Dolecek, 2010 in terms of codes from finite Euclidean planes, $\mathcal{C}_{\text{EG}}(2,q)$:

Lemma (Dolecek, 2010)

The parameters (a^*, b^*) of the smallest absorbing sets for $\mathcal{H}_{\text{EG}}(2,q)$ satisfy $a^* \geq 2 + \left\lfloor \frac{q}{2} \right\rfloor$, and $b^* \geq a^* \cdot \left\lfloor \frac{q}{2} \right\rfloor$.

Small absorbing sets

Proposition

The only nontrivial EG-LDPC code that has a $(3,3)$ absorbing set is the code with parity-check matrix $\mathcal{H}_{\text{EG}}(2,2)$. There are four such sets in $\mathcal{H}_{\text{EG}}(2,2)$.

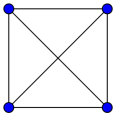


Figure: The Euclidean geometry EG$(2,2)$. Any subset of three points forms a $(3,3)$ absorbing set of $\mathcal{H}_{\text{EG}}(2,2)$.

Smallest AS in $\mathrm{EG}(2,4)$

The smallest absorbing sets of $\mathcal{H}_{\mathrm{EG}}(2,4)$ are $(4,8)$ absorbing sets, which are the minimal parameters according to the bound in Lemma 1.

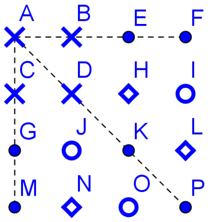


Figure: The set $\{A, B, C, D\}$ forms a $(4,8)$ absorbing set in $\mathrm{EG}(2,4)$. The only 'lines' represented are those containing the point A—three are shown by dotted lines. The fourth line containing A is denoted by open circles and the fifth by diamonds.

Smallest AS in EG$(2, 8)$

Smallest AS in $EG(2,8)$

Theorem

The smallest absorbing set of $\mathcal{H}_{\mathsf{EG}}(2,8)$ *has* $a = 8$, $b = 32$. *There are* $9 \cdot \binom{9}{2} \cdot \binom{8}{4}^2 = 1,587,600$ *distinct smallest absorbing sets in* $\mathcal{H}_{\mathsf{EG}}(2,8)$.

Smallest AS in $EG(2,8)$

Theorem

The smallest absorbing set of $\mathcal{H}_{\mathsf{EG}}(2,8)$ *has* $a = 8$, $b = 32$. *There are* $9 \cdot \binom{9}{2} \cdot \binom{8}{4}^2 = 1{,}587{,}600$ *distinct smallest absorbing sets in* $\mathcal{H}_{\mathsf{EG}}(2,8)$.

Corollary

There exist at least $(2^s+1) \cdot \binom{2^s+1}{2} \cdot \binom{2^s}{2^{s-1}}^2$ *absorbing sets with parameters* $(2^s, 2^{2s+1})$ *in* $\mathcal{H}_{\mathsf{EG}}(2,2^s)$.

A more general case

Theorem

Let a be an even integer, $a = 2\alpha$.
If $a \equiv 0 \pmod 4$ and $\left\lceil \frac{q}{2} \right\rceil \leq \alpha \leq q$, then there are at least
$(q+1) \cdot \binom{q+1}{2} \cdot \binom{q}{\alpha}^2$ *type (a, b) absorbing sets in $\mathcal{H}_{\mathrm{EG}}(2, q)$.*
If a is even but $a \not\equiv 0 \pmod 4$, and $\left\lceil \frac{q}{2} \right\rceil + 1 \leq \alpha \leq q$, then there are at least
$(q+1) \cdot \binom{q+1}{2} \cdot \binom{q}{\alpha}^2$ *type (a, b) absorbing sets in $\mathcal{H}_{\mathrm{EG}}(2, q)$.*

A more general case

Theorem

Let a be an even integer, $a = 2\alpha$.
If $a \equiv 0 \pmod 4$ *and* $\left\lceil \frac{q}{2} \right\rceil \le \alpha \le q$, *then there are at least*
$(q+1) \cdot \binom{q+1}{2} \cdot \binom{q}{\alpha}^2$ *type* (a, b) *absorbing sets in* $\mathcal{H}_{\mathsf{EG}}(2, q)$.
If a is even but $a \not\equiv 0 \pmod 4$, *and* $\left\lceil \frac{q}{2} \right\rceil + 1 \le \alpha \le q$, *then there are at least*
$(q+1) \cdot \binom{q+1}{2} \cdot \binom{q}{\alpha}^2$ *type* (a, b) *absorbing sets in* $\mathcal{H}_{\mathsf{EG}}(2, q)$.

Proof.

- Parallel bundles of lines are the basis for this construction of (a, b) absorbing sets.
- Choose two lines, L_1 and L_2, from a parallel bundle.
- Designate α points on each line, and call this set of points $\mathcal{A}$.
  Then $\mathcal{A}$ forms an absorbing set with a vertices.

□

Conjecture

- The smallest absorbing set in $\mathcal{H}_{\mathsf{EG}}(2, 2^s)$ has $a > 2^{s-1} + 2$, for $s \geq 3$.
- (In other words, apart from $\mathcal{H}_{\mathsf{EG}}(2, 2)$ and $\mathcal{H}_{\mathsf{EG}}(2, 4)$, the smallest a for an (a, b) absorbing set in $\mathcal{H}_{\mathsf{EG}}(2, 2^s)$ always exceeds the bound in Lemma 1.)
- We computationally verified that a^* is greater than the bound in Lemma 1 for $\mathcal{H}_{\mathsf{EG}}(2, 16)$.

Fully and elementary AS

Recall, if S is an absorbing set and all variable nodes in $V \backslash S$ have strictly more neighbors in $W \backslash \mathcal{O}(S)$ than in $\mathcal{O}(S)$, then S is a **fully absorbing set**.
The $(4, 8)$ absorbing sets in $\mathcal{H}_{\mathrm{EG}}(2, 4)$ are fully absorbing sets.

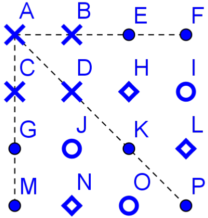


Figure: The set $\{A, B, C, D\}$ forms a $(4, 8)$ absorbing set in EG$(2, 4)$. The only 'lines' represented are those containing the point A—three are shown by dotted lines. The fourth line containing A is denoted by open circles and the fifth by diamonds.

Elementary AS

An **elementary absorbing set** is an absorbing set in which all vertices in W_S have degree one or two in G_S.

Elementary AS

An **elementary absorbing set** is an absorbing set in which all vertices in W_S have degree one or two in G_S.
The $(4, 8)$ absorbing sets in $\mathcal{H}_{\mathsf{EG}}(2, 4)$ are also all elementary absorbing sets, since each line of the geometry contains either $0, 1$, or 2 points in $\mathcal{A}$.

Elementary AS

An **elementary absorbing set** is an absorbing set in which all vertices in W_S have degree one or two in G_S.

The $(4, 8)$ absorbing sets in $\mathcal{H}_{\mathsf{EG}}(2, 4)$ are also all elementary absorbing sets, since each line of the geometry contains either $0, 1$, or 2 points in $\mathcal{A}$.

Conclusions and extensions

- We constructed and enumerated certain absorbing sets for codes from finite affine planes. We enumerated the smallest absorbing sets in $\mathcal{H}_{\mathsf{EG}}(2,8)$ and described their structure, which demonstrates that a sharper lower bound than the Dolecek Lemma on the size of the smallest absorbing sets in $\mathcal{H}_{\mathsf{EG}}(2,q)$ may be possible for this class of codes.
- Ongoing work includes extending these ideas to codes from geometries with dimension larger than two and obtaining a broader classification of fully and elementary absorbing sets in $\mathcal{H}_{\mathsf{EG}}(2,q)$.

Select References

Smarandache, R., Vontobel, P.: Pseudo-codeword analysis of Tanner graphs from projective and Euclidean planes. IEEE Transactions on Information Theory, vol. 53, no. 7, July. (2007)

Kou, Y., Lin, S., Fossorier, M.: Low density parity-check codes based on finite geometries: A rediscovery and new results. IEEE Transactions on Information Theory, vol. 47, no. 7, 2711-2736, Nov. (2001)

Kelley, C.A., Sridhara, D. and Rosenthal, J.: Tree-based construction of LDPC codes having good pseudocodeword weights. IEEE Trans. on Info. Thy, **53** (4), 1460-1478 (2007)

H. Liu, Y. Li, L. Ma, J. Chen.: On the Smallest Absorbing Sets of LDPC Codes From Finite Planes. IEEE Trans. on Infor. Thy, **58**, no. 6, (2012)

Dolecek, L., Zhang, Z., Anantharam, V., Wainwright, M., and Nikolic, B.: Analysis of absorbing sets for array-based LDPC codes. In: IEEE Int. Conf. on Comm., 6261-6268, (2007)

Schwartz, M., Vardy, A.: On the stopping distance and the stopping redundancy of codes. IEEE Trans. on Info. Thy, **52** (3), 922-932 (2006)

Batten, L. M.: Combinatorics of finite geometries, 2nd ed. Cambridge U. Press (1997)

Tanner, R.M.: Explicit concentrators from generalized N-gons. SIAM Journal on Algebraic Discrete Methods, **5** (3), 287-293 (1984)

Hirschfeld, J.W., Storme, L.: The packing problem in statistics, coding theory and finite projective spaces: Update 2001. In: Finite geometries, 201-246, Springer US (2001)

Dolecek, L.: On absorbing sets of structured sparse graph codes. In: Info. Thy and Applications Workshop (ITA), IEEE, 1-5 (2010)