

Quasi-uniform codes and their applications

Terence Chan Alex Grant

University of South Australia

6 August 2009

1 Linear Codes & Matroids (yet again)

- Matroids from Codes & Greene's Theorem
- The Fundamental Question
- Another Perspective - Entropy Functions

2 Quasi-uniform Codes

- Definition & Combinatorial Interpretation
- Construction from Groups

3 Main results

- Distance Invariance
- Puncturing & Shortening
- Generalized Greene's Theorem

4 Applications

- Robust Data Transmission
- Secret Sharing

5 Summary

1 Linear Codes & Matroids (yet again)

- Matroids from Codes & Greene's Theorem
- The Fundamental Question
- Another Perspective - Entropy Functions

2 Quasi-uniform Codes

- Definition & Combinatorial Interpretation
- Construction from Groups

3 Main results

- Distance Invariance
- Puncturing & Shortening
- Generalized Greene's Theorem

4 Applications

- Robust Data Transmission
- Secret Sharing

5 Summary

Definition

A matroid (\mathcal{X}, ρ) is a finite set \mathcal{X} and a “rank” function $\rho : 2^{\mathcal{X}} \mapsto \mathbb{R}^+$ satisfying for all $\mathcal{A}, \mathcal{B} \subseteq \mathcal{X}$,

- Cardinality bound: $\rho(\mathcal{A}) \leq |\mathcal{A}|$
- Integrality constraint: $\rho(\mathcal{A})$ is an integer
- Polymatroid constraints:

$$\rho(\emptyset) = 0 \tag{R1}$$

$$\mathcal{A} \subseteq \mathcal{B} \implies \rho(\mathcal{A}) \leq \rho(\mathcal{B}) \tag{R2}$$

$$\rho(\mathcal{A} \cup \mathcal{B}) + \rho(\mathcal{A} \cap \mathcal{B}) \leq \rho(\mathcal{A}) + \rho(\mathcal{B}). \tag{R3}$$

Linear Codes

- Dimension k subspace $C \subseteq \mathbb{F}_q^n$
- q^k codewords of length n
- Row space of a $k \times n$ generator matrix G
- Induces a vector matroid $M[G] = (\mathcal{N}, \rho)$

$$\rho(\mathcal{A}) = \text{rank } G_{\mathcal{A}}$$

where for all $\mathcal{A} \subseteq \mathcal{N} \triangleq \{1, 2, \dots, N\}$, $G_{\mathcal{A}}$ is the submatrix of G obtained by deleting columns not indexed by \mathcal{A} .

Definition (Weight enumerator polynomial)

Let C be a linear code of length n . Its weight enumerator polynomial is

$$W_C(x, y) \triangleq \sum_{z_{\mathcal{N}} \in C} x^{n-D(z_{\mathcal{N}})} y^{D(z_{\mathcal{N}})}$$

where $D(z_{\mathcal{N}})$ is the Hamming weight of $z_{\mathcal{N}}$.

Definition (Tutte polynomial)

Let ρ be the rank function of a matroid (\mathcal{N}, ρ) . Its Tutte polynomial is

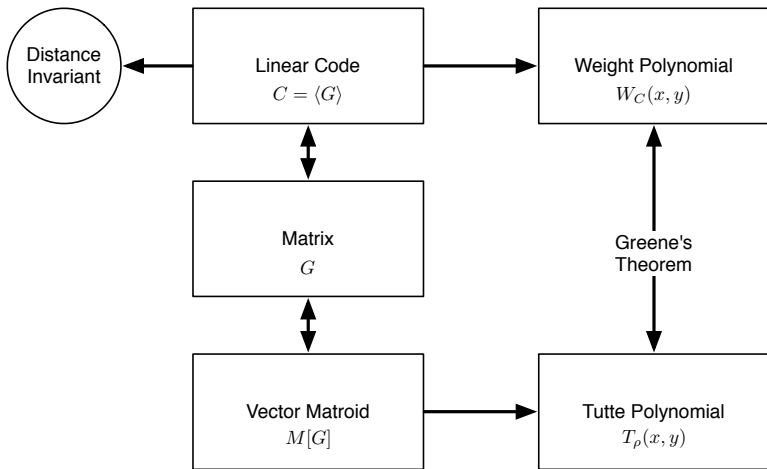
$$T_{\rho}(x, y) \triangleq \sum_{\mathcal{A} \subseteq \mathcal{N}} (x-1)^{\rho(\mathcal{N})-\rho(\mathcal{A})} (y-1)^{|\mathcal{A}|-\rho(\mathcal{A})}.$$

The Link: Greene's Theorem

Theorem (Greene)

Let C be a linear code of length n over a finite field \mathbb{F}_q and ρ be its associated matroid. Then

$$W_C(x, y) = y^{n-\rho(\mathcal{N})} (x-y)^{\rho(\mathcal{N})} T_\rho \left(\frac{x + (q-1)y}{x-y}, \frac{x}{y} \right). \quad (1)$$



The fundamental question

Besides linear codes, are there any other codes that exhibit similar properties?

From Codes to Random Variables

- Consider a linear code $C = \langle G \rangle$.
- Let $U = (U_1, U_2, \dots, U_k)$ be uniformly drawn from \mathbb{F}_q^k
- *Code symbol random variables*

$$(Z_1, \dots, Z_n) \triangleq U^T G$$

- Induces a probability distribution

$$\Pr(Z_{\mathcal{N}} = z_{\mathcal{N}}) = \begin{cases} 1/|C| & \text{if } z_{\mathcal{N}} \in C \\ 0 & \text{otherwise.} \end{cases}$$

- *A code and its induced random variables are in one-to-one correspondence.*

Another Perspective: Entropy Functions

- For any subset \mathcal{A} of \mathcal{N} ,

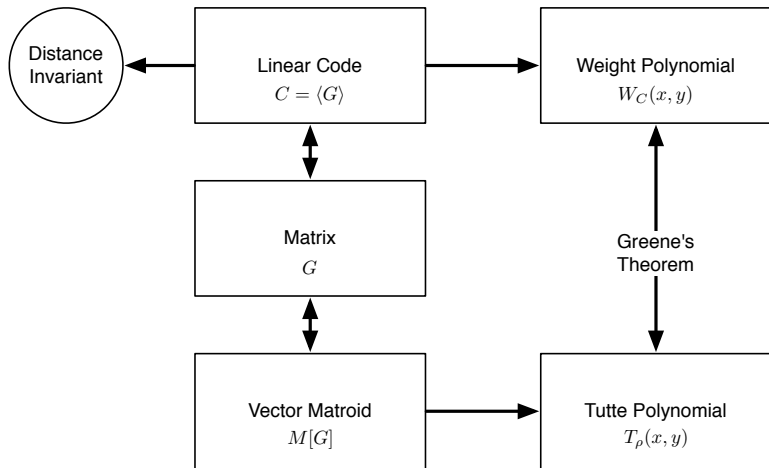
$$\Pr(Z_{\mathcal{A}} = z_{\mathcal{A}}) = \begin{cases} q^{-|G_{\mathcal{A}}|} & \text{if } \exists u \in \mathbb{F}_q^k, z_{\mathcal{A}} = u^T G_{\mathcal{A}} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

- Random variable $Z_{\mathcal{A}}$ is *uniformly distributed over its support*.
- $H(Z_{\mathcal{A}}) = |G_{\mathcal{A}}| = \rho(\mathcal{A})$ (w.r.t. \log_q)

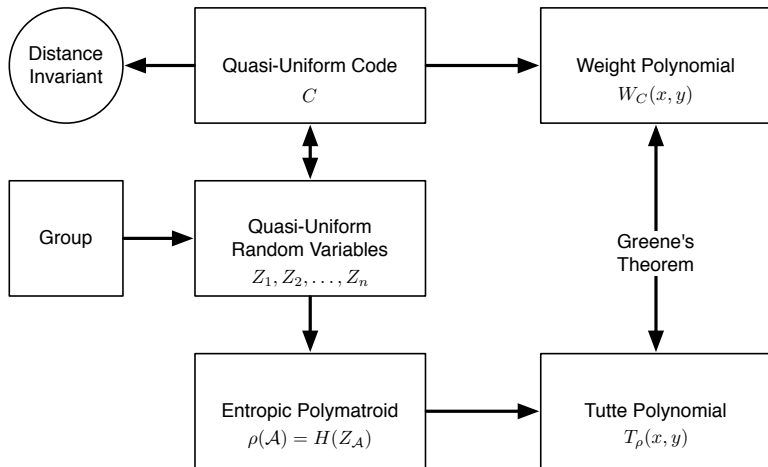
In other words ...

Greene's Theorem links the weight enumerator polynomial and the Tutte polynomial induced by the entropy function of the code symbol random variables.

Preview ...



Preview ...



1 Linear Codes & Matroids (yet again)

- Matroids from Codes & Greene's Theorem
- The Fundamental Question
- Another Perspective - Entropy Functions

2 Quasi-uniform Codes

- Definition & Combinatorial Interpretation
- Construction from Groups

3 Main results

- Distance Invariance
- Puncturing & Shortening
- Generalized Greene's Theorem

4 Applications

- Robust Data Transmission
- Secret Sharing

5 Summary

Definition (Quasi-Uniform Random Variables)

A set of random variables (Z_1, \dots, Z_n) is *quasi-uniform* if for any $\mathcal{A} \subseteq \mathcal{N}$, $Z_{\mathcal{A}} \triangleq (Z_i : i \in \mathcal{A})$ is uniformly distributed over its support $\lambda(Z_{\mathcal{A}})$,

$$\Pr(Z_{\mathcal{A}} = z_{\mathcal{A}}) = \begin{cases} 1/|\lambda(Z_{\mathcal{A}})| & \text{if } z_{\mathcal{A}} \in \lambda(Z_{\mathcal{A}}) \\ 0 & \text{otherwise.} \end{cases}$$

- A code C is called *quasi-uniform* if its induced code symbol random variables are quasi-uniform.
- Linear codes are quasi-uniform.

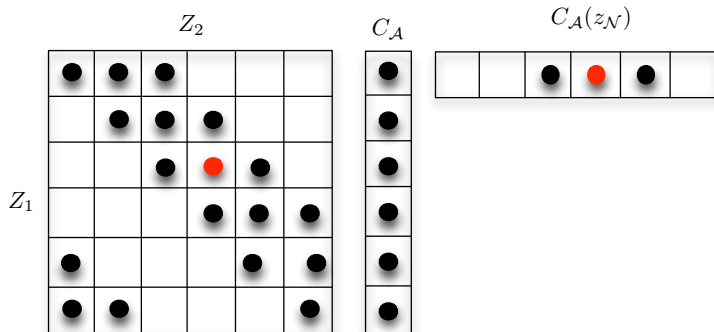
Combinatorial interpretation

Definition (Projection & Cross-Section)

Let $C \subseteq \prod_{i=1}^n \mathcal{Z}_i$. For any $\mathcal{A} \subseteq \mathcal{N}$ and $z_{\mathcal{N}} \in C$, we define

$$C_{\mathcal{A}} \triangleq \{z_{\mathcal{A}} \in \mathcal{Z}_{\mathcal{A}} : \exists z_{\mathcal{A}^c}, (z_{\mathcal{A}}, z_{\mathcal{A}^c}) \in C\} \quad \text{Projection}$$

$$C_{\mathcal{A}}(z_{\mathcal{N}}) \triangleq \{y_{\mathcal{A}^c} \in \mathcal{Z}_{\mathcal{A}^c} : (z_{\mathcal{A}}, y_{\mathcal{A}^c}) \in C\} \quad \text{Cross-section}$$



Proposition

A code C is quasi-uniform \Leftrightarrow for any $\mathcal{A} \subseteq \mathcal{N}$ and $z_{\mathcal{N}} \in C$,

- $|C_{\mathcal{A}}(z_{\mathcal{N}})|$ is constant for all $z_{\mathcal{N}} \in C$
- $|C| = |\mathcal{C}_{\mathcal{A}}| |C_{\mathcal{A}}(z_{\mathcal{N}})|$

How General are Quasi-Uniform Codes?

- A code is quasi-uniform if and only if every codeword looks essentially “the same”.
(quasi-uniform codes are not overly general)
- Can be nonlinear
- Can be constructed from any finite group
- Polymatroids, not necessarily representable matroids.
- Quasi-uniform random variables suffice to characterize $\bar{\Gamma}^*$

Quasi-uniform from Groups & Subgroups

- Let G be a finite group with subgroups G_1, \dots, G_n . Let U be uniformly distributed over G
- Each G_i induces a random variable U_i : the left coset of G_i in G that contains U .
- Then (U_1, \dots, U_n) is quasi-uniform.

Theorem (Linear codes are quasi-uniform)

Let $C = (Z_1, \dots, Z_n)$ be a linear code generated by M . Let

$$G = \mathbb{F}_q^k$$
$$G_i \triangleq \{u \in \mathbb{F}_q^k : u^T M_i = 0\}$$

Then Z_i is induced by the subgroup G_i .

Almost Affine Codes

Definition (Almost Affine Code)

Let $C \subseteq \prod_{i=1}^n \mathcal{Z}_i$ be a length n code where $|\mathcal{Z}_i| = q > 1$ for all $i \in \mathcal{N}$. The code C is called *almost affine* if for any $\mathcal{A} \subseteq \mathcal{N}$, $\log_q |C_{\mathcal{A}}|$ is a nonnegative integer.

Theorem (Almost affine codes are quasi-uniform)

A code C is almost affine if and only if it is quasi-uniform and its induced entropy function is a matroid.



J. Simonis and A. Ashikhmin, Almost affine codes, Designs, Codes and Cryptography, vol. 14, no. 2, pp. 179 - 197, 1998

1 Linear Codes & Matroids (yet again)

- Matroids from Codes & Greene's Theorem
- The Fundamental Question
- Another Perspective - Entropy Functions

2 Quasi-uniform Codes

- Definition & Combinatorial Interpretation
- Construction from Groups

3 Main results

- Distance Invariance
- Puncturing & Shortening
- Generalized Greene's Theorem

4 Applications

- Robust Data Transmission
- Secret Sharing

5 Summary

Distance invariance

- Consider any code C and a codeword $z_{\mathcal{N}} \in C$.
- *Distance profile* of C centered at $z_{\mathcal{N}}$

$$A(z_{\mathcal{N}}, r) = |\{y_{\mathcal{N}} : d_H(y_{\mathcal{N}}, z_{\mathcal{N}}) \leq r\}|$$

- A code is *distance-invariant* if $A(z_{\mathcal{N}}, r)$ independent of $z_{\mathcal{N}}$.
- Linear codes are distance-invariant.

Theorem

Quasi-uniform codes are distance-invariant.

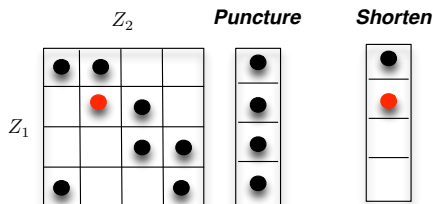
Puncturing and Shortening

- Puncturing (with respect to position i)

$$C_{\mathcal{N}\setminus i} = (Z_j : j \neq i)$$

- Shortening (with respect to position i)

$$C_i(\mathbf{0}) = \{y_{\mathcal{N}\setminus i} : (y_{\mathcal{N}\setminus i}, 0) \in C\}$$



Theorem

If C is quasi-uniform, then its punctured and shortened codes are also quasi-uniform.

Generalized Greene's Theorem

Theorem

Let $C = (Z_1, \dots, Z_n)$ be quasi-uniform code with entropy function $\rho(\mathcal{A}) \triangleq H_q(Z_i : i \in \mathcal{A})$. Define its weight enumerator polynomial and Tutte polynomial as follows:

$$W_C(x, y) \triangleq \sum_{z_{\mathcal{N}} \in C} x^{n-D(z_{\mathcal{N}})} y^{D(z_{\mathcal{N}})}$$

$$T_\rho(x, y) \triangleq \sum_{\mathcal{A} \subseteq \mathcal{N}} (x-1)^{\rho(\mathcal{N})-\rho(\mathcal{A})} (y-1)^{|\mathcal{A}|-\rho(\mathcal{A})}.$$

Then

$$W_C(x, y) = y^{n-\rho\mathcal{N}} (x-y)^{\rho\mathcal{N}} T_\rho \left(\frac{x+(q-1)y}{x-y}, \frac{x}{y} \right).$$

Sketch of Proof

Proof by induction.

- Let C be a quasi-uniform code of length n
- Let ρ be its induced entropy function.
- Fix $i \in \mathcal{N}$. Let $C_1 = C_{\mathcal{N} \setminus i}$ and $C_2 = C_i(\mathbf{0})$ be obtained respectively by puncturing and shortening w.r.t. Z_i .
- Induction uses two “splitting” lemmas:

$$W_C(x, y) = yq^{\rho(\mathcal{N}) - \rho(\mathcal{N} \setminus i)} W_{C_1}(x, y) + (x - y) W_{C_2}(x, y)$$

$$T_\rho(x, y) = (x - 1)^{\rho(\mathcal{N}) - \rho(\mathcal{N} \setminus i)} T_{\rho_1}(x, y) + (y - 1)^{1 - \rho(i)} T_{\rho_2}(x, y)$$

1 Linear Codes & Matroids (yet again)

- Matroids from Codes & Greene's Theorem
- The Fundamental Question
- Another Perspective - Entropy Functions

2 Quasi-uniform Codes

- Definition & Combinatorial Interpretation
- Construction from Groups

3 Main results

- Distance Invariance
- Puncturing & Shortening
- Generalized Greene's Theorem

4 Applications

- Robust Data Transmission
- Secret Sharing

5 Summary

Robust data transmission

- Transmitter and receiver connected via n parallel links
- Capacity of each link is different.
- An adversary aims to obstruct data transmission by replacing messages.
- Adversary can attack t links.
- Errors can be corrected if the code has a minimum Hamming distance at least $2t + 1$.
- Using linear or almost affine codes, all code symbols are drawn from the same set. The rate of the code will suffer if one link has small capacity forcing all the code symbols to take values in a smaller set.

Definition

A secret sharing scheme for an access structure Ω is a set of random variables (Z_0, \dots, Z_n) such that

- 1 Z_0 is the secret uniformly distributed over \mathcal{Z}_0 ;
- 2 $H(Z_0|Z_i, i \in \mathcal{A}) = 0$ if $\mathcal{A} \in \Omega$;
- 3 $I(Z_0; Z_{\mathcal{A}}) = 0$ if $\mathcal{A} \notin \Omega$.

- Quasi-uniform codes may be of interest for metric

$$\min_{i \in \mathcal{N}} \frac{H(Z_0)}{c_i H(Z_i)}$$

where almost affine codes may be inefficient.

- Non-ideal access structures?

1 Linear Codes & Matroids (yet again)

- Matroids from Codes & Greene's Theorem
- The Fundamental Question
- Another Perspective - Entropy Functions

2 Quasi-uniform Codes

- Definition & Combinatorial Interpretation
- Construction from Groups

3 Main results

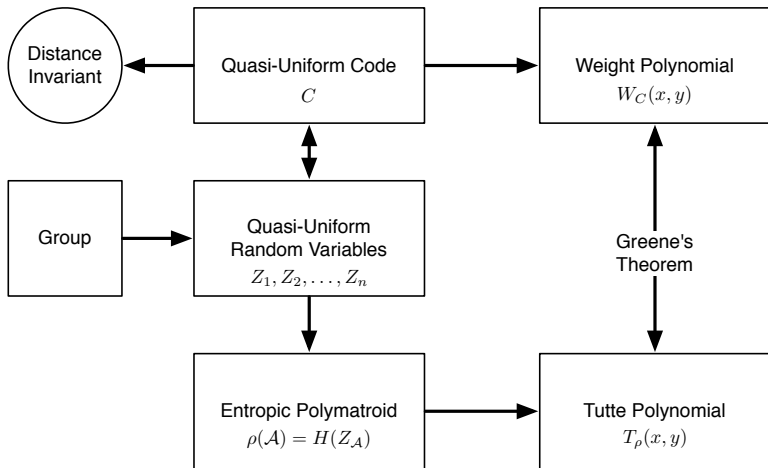
- Distance Invariance
- Puncturing & Shortening
- Generalized Greene's Theorem

4 Applications







- Robust Data Transmission
- Secret Sharing

5 Summary

Summary



If you enjoyed this talk, you might also like...

-  T. H. Chan, “A combinatorial approach to information inequalities,” *Communications in Information and Systems*, vol. 1, pp. 1–14, 2001.
-  T. H. Chan and R. W. Yeung, “On a relation between information inequalities and group theory,” *IEEE Trans. Inform. Theory*, vol. 48, pp. 1992–1995, 2002.
-  L. Guillé, T. Chan and A. Grant, “The minimal set of Ingleton inequalities,” *IEEE Int. Symp. Inform. Theory*, (Toronto, Canada), July 2008.
-  T. Chan and A. Grant, “Dualities between entropy functions and network codes,” *IEEE Trans. Inform. Theory*, vol. 54, pp. 4470 – 4487, Oct. 2008.
-  T. Chan and A. Grant, “Non-linear information inequalities,” *Entropy*, vol. 10, pp. 765–775, Dec. 2008.
-  T. Chan, A. Grant and D. Kern, “Existence of new inequalities for representable polymatroids,” [arXiv:0907.5030v1](https://arxiv.org/abs/0907.5030v1).