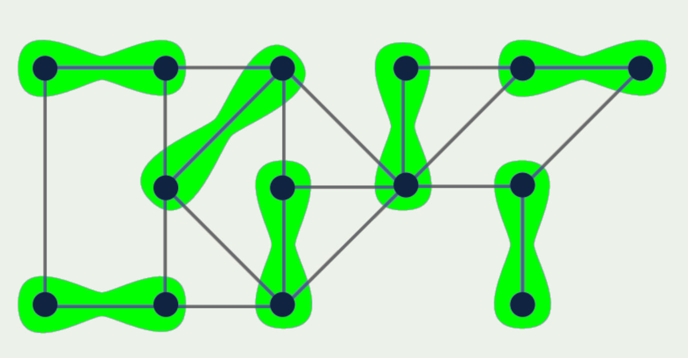
Basis Collapse in Holographic Algorithms Over All Domain Sizes

Sitan Chen
Harvard College

March 29, 2016

Holographic algorithms reduce counting problems into the problem of *counting perfect matchings* in a graph $G = (V, E)$.

- Perfect matching: $M \subset E$ for which every $v \in V$ belongs to exactly one edge $e \in M$
- [Valiant '79]: Counting perfect matchings in arbitrary graphs is #P-complete.
- [Fisher-Temperley 1961, Kasteleyn 1961]: Counting perfect matchings in planar graphs is in P.



More generally, if every edge e of G has some weight $w(e)$, define

$$\mathrm{PerfMatch}(G) = \sum_{\text{perfect matchings } M} \left(\prod_{e \in M} w(e) \right).$$

Theorem (FKT algorithm)

If G is a planar weighted graph, $\mathrm{PerfMatch}(G)$ *can be computed in polynomial time.*

Idea.

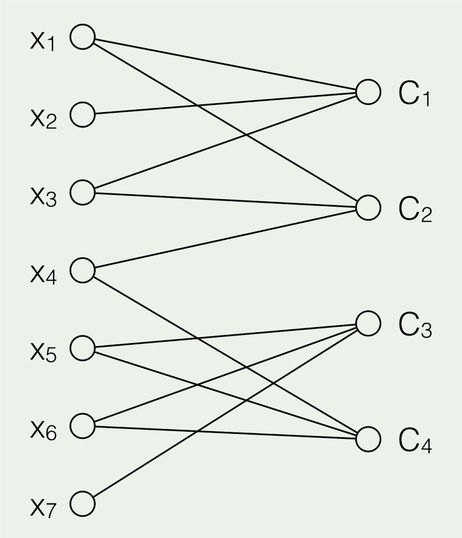
For an arbitrary graph G with adjacency matrix A, the *Pfaffian*

$$\mathrm{Pf}(A) = \sum_{\text{perfect matchings } M} \mathrm{sgn}(M) \left(\prod_{e \in M} w(e) \right)$$

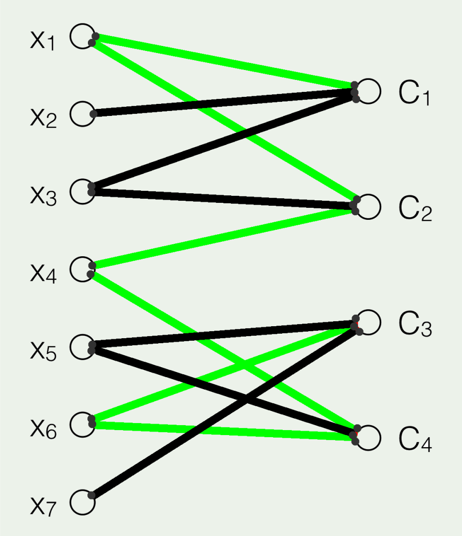
satisfies $\mathrm{Pf}(A)^2 = \det(A)$. For planar graphs, can flip the signs of some entries of A to make Pf and PerfMatch agree. □

$$(x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee x_3 \vee x_4) \wedge (x_5 \vee x_6 \vee x_7) \wedge (x_4 \vee x_5 \vee x_6)$$

$$(x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee x_3 \vee x_4) \wedge (x_5 \vee x_6 \vee x_7) \wedge (x_4 \vee x_5 \vee x_6)$$



Imagine: each vertex v on the left propagates signals along its outgoing edges indicating whether v is assigned 1 (green) or 0 (black).



Each satisfying assignment corresponds to a collection of signals satisfying two constraints:

Consistency: If x_i is a vertex on the left, the two signals x_i generates must be the same.

00	1
01	0
10	0
11	1

Satisfaction: If C_j is a vertex on the right, at least one of the three signals it receives must be 1.

000	0
001	1
010	1
011	1
100	1
101	1
110	1
111	1

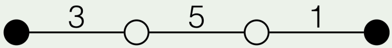
Goal: encode these bit vectors using the matching properties of graphs

Definition

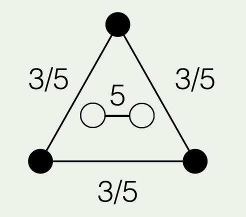
A *matchgate* is a weighted graph G with designated subsets of its vertices called *external nodes* X. We say that it is of *arity* $|X|$.

Definition

The *standard signature* $\underline{G}$ of matchgate G of arity n is a vector of dimension 2^n with entries indexed by bitstrings of length n. For $Z \subset X$ corresponding to bitstring α, $\underline{\Gamma}^\alpha = \mathrm{PerfMatch}(\Gamma \backslash Z)$.



00	3
01	0
10	0
11	5



000	0
001	3
010	3
011	0
100	3
101	0
110	0
111	5

We want *planar* matchgates G and R whose standard signatures respectively match the vectors encoding the consistency and satisfaction constraints:

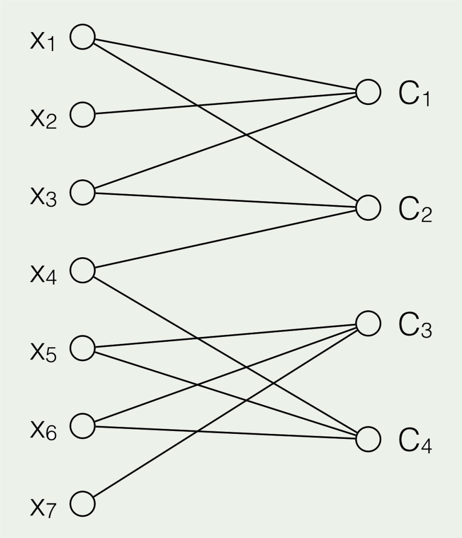
Consistency: If x_i is a vertex on the left, the two signals x_i generates must be the same.

00	1
01	0
10	0
11	1

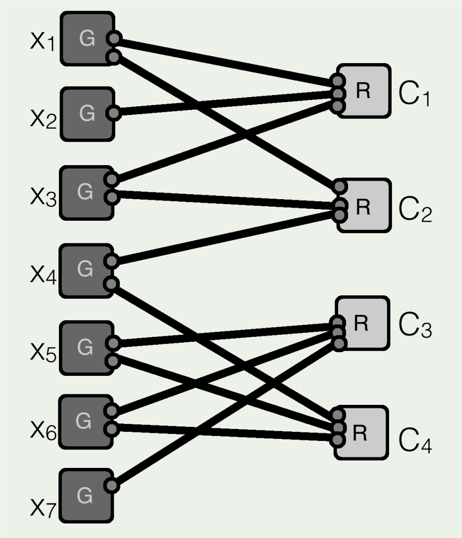
Satisfaction: If C_j is a vertex on the right, at least one of the three signals it receives must be 1.

000	0
001	1
010	1
011	1
100	1
101	1
110	1
111	1

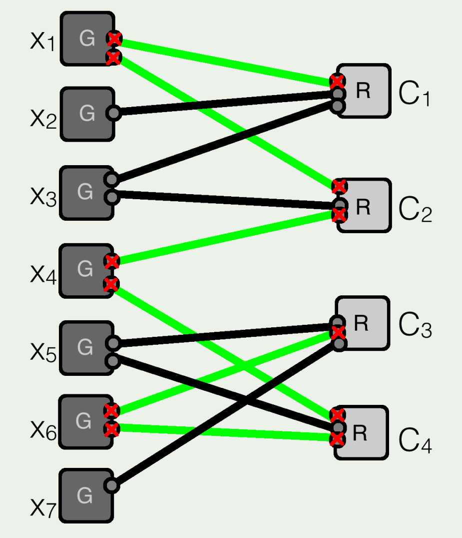
$$(x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee x_3 \vee x_4) \wedge (x_5 \vee x_6 \vee x_7) \wedge (x_4 \vee x_5 \vee x_6)$$



$$(x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee x_3 \vee x_4) \wedge (x_5 \vee x_6 \vee x_7) \wedge (x_4 \vee x_5 \vee x_6)$$



$$(x_1 \vee x_2 \vee x_3) \wedge (x_1 \vee x_3 \vee x_4) \wedge (x_5 \vee x_6 \vee x_7) \wedge (x_4 \vee x_5 \vee x_6)$$



Unfortunately, no recognizer has standard signature $(0, 1, 1, 1, 1, 1, 1, 1)$:

Observation (Parity Condition)

Because a graph with an odd number of vertices has no perfect matchings, given any matchgate G, the indices of the nonzero entries in its standard signature must have the same parity.

- The saving grace: rewrite number of perfect matchings of matchgrid Ω as an inner product and apply a change of basis.
- Suppose there are w wires in Ω, generators $G_1, \ldots G_g$, and $R_1, \ldots, R_r$ recognizers, then

$$\mathrm{PerfMatch}(\Omega) = \sum_{\substack{z \in \{0,1\}^w, \\ z = x_1 \circ \cdots \circ x_r \circ \\ y_1 \circ \cdots \circ y_g}} \left(\prod_{i=1}^{g} \underline{G_i}^{y_i} \prod_{j=1}^{r} \underline{R_j}^{x_j} \right) = \langle \mathbf{G}, \mathbf{R} \rangle,$$

  where $\mathbf{G} = \otimes_i \underline{G_i}$ and $\mathbf{R} = \otimes_i \underline{R_i}$ with the order of tensoring specified by the wires.
- Regard $\mathbf{G}$ as an element in $X = \mathbb{C}^{2^w}$ and $\mathbf{R}$ as an element in X^*: $\mathrm{PerfMatch}(\Omega)$ is the result of applying dual vector $\mathbf{R}$ to $\mathbf{G}$, which is *independent of the choice of basis for* X.

Definition

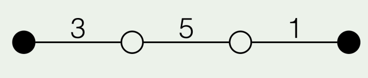
Given a 2×2 *basis matrix* M, the *signature with respect to* M *of a generator* G of arity n is the vector G satisfying

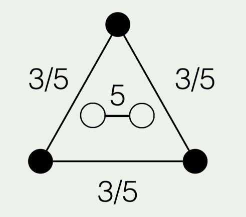
$$\underline{G} = M^{\otimes n} G.$$

The *signature with respect to* M *of a recognizer* R of arity n is the vector R satisfying

$$R = \underline{R} M^{\otimes n}.$$

- Suffices to find a basis M of matchgates G and R whose signatures with respect to M match the vectors encoding the consistency and satisfaction constraints.
- Over $\mathbb{C}$ and $\mathbb{F}_2$, this still cannot be done.
- [Valiant '06, Cai-Lu '07]: Over $\mathbb{F}_7$, take $M = \begin{pmatrix} 1 & 3 \\ 6 & 5 \end{pmatrix}$, $\underline{G} = (3, 0, 0, 5)$, and $\underline{R} = (0, 3, 3, 0, 3, 0, 0, 5)$.





00	3
01	0
10	0
11	5

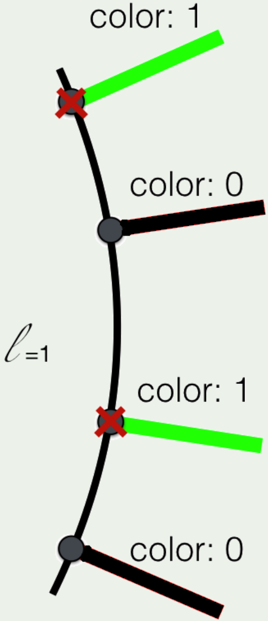
000	0
001	3
010	3
011	0
100	3
101	0
110	0
111	5

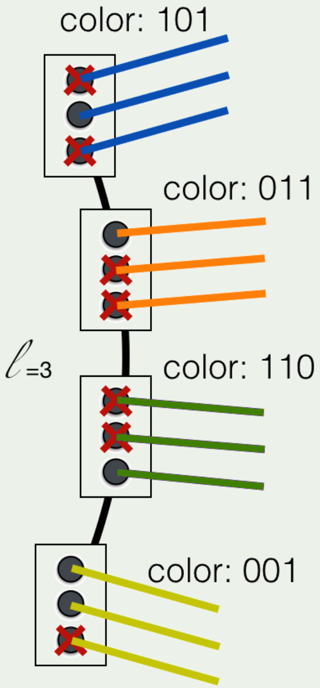
- The number of different values that objects in a counting problem can take on is called the domain size.
- Domain size 2:
  - Boolean satisfying assignments
  - Vertex covers
  - Perfect matchings
  - Ice problems
- Domain size k
  - k-colorings

Over domain size k:

- Arity-n signatures are now vectors of dimension k^n.
- M now has width k because

$$\underline{G} = M^{\otimes n} G \qquad R = \underline{R} M^{\otimes n}.$$


color: 1
color: 0
ℓ=1
color: 1
color: 0



color: 101
color: 011
$\ell_{=3}$
color: 110
color: 001

- Domain size 2: encode TRUE/FALSE by presence/absence of one external node
- Domain size k: encode colors $\{1, ..., k\}$ by removal of some subset of a group of ℓ external nodes
  - Arities are now multiples of ℓ
  - External nodes grouped into blocks of ℓ, with wires connecting matchgates blockwise.
  - If Γ has n blocks, $\underline{\Gamma}$ has $2^{\ell n}$ entries.
  - M has height 2^{ℓ} because

$$\underline{G} = M^{\otimes n} G \qquad R = \underline{R} M^{\otimes n}.$$

  - We call ℓ the basis size.

We will regard standard signatures as matrices:

Definition

For standard signature $\underline{G}$ of generator G, the *t-th matrix form* $\underline{G}(t)$ $(1 \le t \le n)$ is the $2^\ell \times 2^{(n-1)\ell}$ matrix of entries of $\underline{G}$ where the rows are indexed by $\alpha_t \in \{0,1\}^\ell$ and the columns are indexed by $\alpha_1 \cdots \alpha_{t-1}\alpha_{t+1} \cdots \alpha_n \in \{0,1\}^{(n-1)\ell}$.

We will also regard signatures as matrices:

Definition

For signature G of generator G, the *t-th matrix form* $G(t)$ $(1 \leq t \leq n)$ is the $k \times k^{n-1}$ matrix of entries of G where the rows are indexed by $\alpha_t \in [k]$ and the columns are indexed by $\alpha_1 \cdots \alpha_{t-1}\alpha_{t+1} \cdots \alpha_n \in [k]^{n-1}$.

Note: we will denote row indices by superscripts and column indices by subscripts.

Definition

A generator G is *full rank* if there exists t for which $\mathrm{rank}(G(t)) = k$.

It turns out we may assume that $\mathrm{rank}(M) = k$. But we know

$$\underline{G}(t) = MG(t)(M^T)^{\otimes(n-1)}.$$

So if G is of full rank,

$$\mathrm{rank}(\underline{G}(t)) = k.$$

Key to understanding the ultimate capabilities of holographic algorithms for solving counting problems over a given domain size:

Question

Given k, what is the smallest ℓ for which any holographic algorithm over domain size k with a full-rank matchgate can be simulated by one with basis size ℓ?

	domain size	basis size
Cai-Lu '08	2	1

Key to understanding the ultimate capabilities of holographic algorithms for solving counting problems over a given domain size:

Question

Given k, what is the smallest ℓ for which any holographic algorithm over domain size k with a full-rank matchgate can be simulated by one with basis size ℓ?

	domain size	basis size
Cai-Lu '08	2	1
Cai-Fu '14	3	1
	4	2

Key to understanding the ultimate capabilities of holographic algorithms for solving counting problems over a given domain size:

Question

Given k, what is the smallest ℓ for which any holographic algorithm over domain size k with a full-rank matchgate can be simulated by one with basis size ℓ?

	domain size	basis size
Cai-Lu '08	2	1
Cai-Fu '14	3	1
	4	2
C '15, Xia '15	k	$\lfloor \log_2 k \rfloor$

Definition

$Z \subset \{0,1\}^n$ is a *cluster* if there exists $s \in \{0,1\}^n$ and positions $p_1, ..., p_m \in [n]$ such that each member of Z is of the form $s \oplus \left(\bigoplus_{j \in J} e_{p_j}\right)$ for some $J \subset \{p_1, ..., p_m\}$, where e_{p_j} is the bitstring consisting of zeroes everywhere except position p_j.

We write Z as $s + \{e_{p_1}, ..., e_{p_m}\}$ (s only unique up to the bits outside of positions $p_1, ..., p_m$).

e.g. $\{000, 001, 100, 101\}$ is a cluster denoted $000 + \{e_1, e_3\}$.

For now, assume $k = 2^K$. Steps of proof:

1. **Cluster existence**: Any standard signature of rank at least 2^K contains a cluster of 2^K linearly independent rows
2. **Group property**: Inverses of standard signatures are also standard signatures
3. **Simulation**: Use 1 and 2 to simulate with a basis of size K

For now, assume $k = 2^K$. Steps of proof:

1. **Cluster existence**: Any standard signature of rank at least 2^K contains a cluster of 2^K linearly independent rows (*hardest part of the proof*)
2. **Group property**: Inverses of standard signatures are also standard signatures
3. **Simulation**: Use 1 and 2 to simulate with a basis of size K

For now, assume $k = 2^K$. Steps of proof:

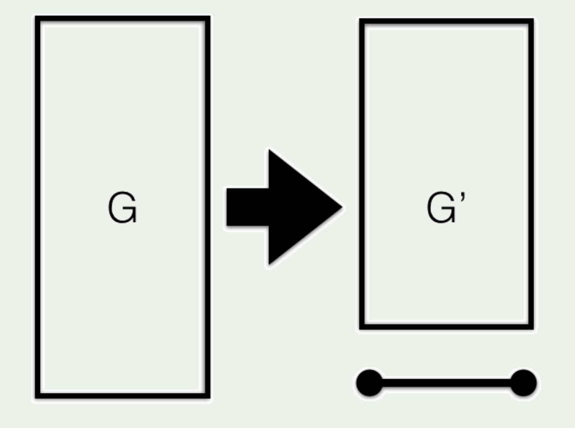
1. **Cluster existence**: Any standard signature of rank at least 2^K contains a cluster of 2^K linearly independent rows (*hardest part of the proof*)
2. **Group property**: Inverses of standard signatures are also standard signatures (*adapted from Li/Xia result in matchgate character theory*)
3. **Simulation**: Use 1 and 2 to simulate with a basis of size K

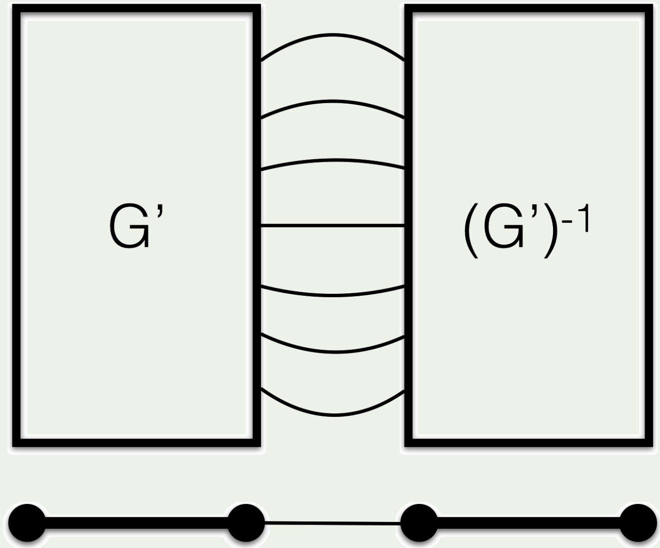
For now, assume $k = 2^K$. Steps of proof:

1. **Cluster existence**: Any standard signature of rank at least 2^K contains a cluster of 2^K linearly independent rows (*hardest part of the proof*)
2. **Group property**: Inverses of standard signatures are also standard signatures (*adapted from Li/Xia result in matchgate character theory*)
3. **Simulation**: Use 1 and 2 to simulate with a basis of size K (*technique due to Cai/Fu*)

Lemma (Group Property)

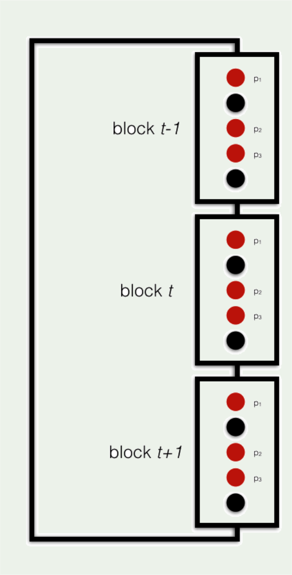
Full-rank $2^K \times 2^{(n-1)K}$ *standard signatures* $\underline{G}(t)$ *have right inverses (under matrix multiplication) that are also standard signatures.*

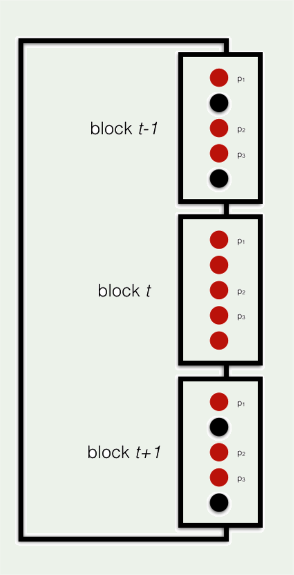



G'
(G')-1

We use the approach introduced by Cai-Fu for simulation given cluster existence and group property have been proven. Take any holographic algorithm over domain size $k = 2^K$.

1. By cluster existence, can pick out a generator G with full-rank signature and find a cluster $Z = s + \{e_{p_1}, ..., e_{p_K}\}$ of 2^K linearly independent rows. Suppose WLOG $s = 0^\ell$.
2. Let M^Z denote the submatrix of M with rows indexed by Z. This will be the basis of size $\log k = K$ we use for the simulation.





$$\underline{G}^{*\leftarrow Z} = (M^Z)^{\otimes n} G$$

$$\underline{G}^{t^c \leftarrow Z} = (M^Z)^{\otimes (t-1)} \otimes M \otimes (M^Z)^{\otimes (n-t)} G$$

- Modifying generators is easy:

$$\underline{G}_i^{*\leftarrow Z} = (M^Z)^{\otimes n_i} G_i$$

has signature G_i with respect to new basis M^Z

- Modifying recognizers is more subtle. Can write

$$R_j = \left(\underline{R_j}(M/M^Z)^{\otimes m_i}\right)(M^Z)^{\otimes m_j}.$$

Is $\underline{R_j}(M/M^Z)^{\otimes m_i}$ a valid recognizer standard signature?

3. Define $T = M(M^Z)^{-1}$.
4. By construction,

$$\underline{G}^{t^c \leftarrow Z}(t) = T\underline{G}^{* \leftarrow Z}(t).$$

5. By group property, $\underline{G}^{* \leftarrow Z}(t)$ has a right-inverse, so right-multiply by this on both sides to conclude that T is a standard signature.

Over the new basis M^Z:

6. Replace each recognizer $\underline{R}_i$ with $\underline{R}_i T^{\otimes m_i}$
7. Replace each generator $\underline{G}_j$ with $\underline{G}_j^{*\leftarrow Z}$.
8. These new matchgates have the same signatures as the originals, but over a basis of size K, so we're done.

The key ingredient:

Theorem (Rank Rigidity)

The rank of any standard signature Γ *(in matrix form) is always a power of two.*

Our methods are primarily algebraic and rely on the characterization of the set of all standard signatures as the variety cut out by a certain collection of quadratic relations:

Theorem (Matchgate Identities)

A $2^\ell \times 2^{(n-1)\ell}$ matrix Γ is the t-th matrix form of the standard signature of some generator matchgate iff for all $\zeta, \eta \in \{0,1\}^{(n-1)\ell}$ and $\sigma, \tau \in \{0,1\}^\ell$, the following matchgate identity (MGI) holds. Let $\zeta \oplus \eta = e_{q_1} \oplus \cdots \oplus e_{q_{d'}}$ and $\sigma \oplus \tau = e_{p_1} \oplus \cdots \oplus e_{p_d}$, where $q_1 < \cdots < q_{d'}$ and $p_1 < \cdots < p_d$. Then if d is even,

$$\sum_{i=1}^{d}(-1)^{i+1}\Gamma_{\zeta}^{(\sigma\oplus e_{p_1}\oplus e_{p_i})}\Gamma_{\eta}^{(\tau\oplus e_{p_1}\oplus e_{p_i})} = \pm\sum_{j=1}^{d'}(-1)^{j+1}\Gamma_{(\zeta\oplus e_{q_j})}^{(\sigma\oplus e_{p_1})}\Gamma_{(\eta\oplus e_{q_j})}^{(\tau\oplus e_{p_1})}.$$

$$\sum_{i=1}^{d}(-1)^{i+1}\Gamma_{\zeta}^{(\sigma\oplus e_{p_1}\oplus e_{p_i})}\Gamma_{\eta}^{(\tau\oplus e_{p_1}\oplus e_{p_i})} = \pm\sum_{j=1}^{d'}(-1)^{j+1}\Gamma_{(\zeta\oplus e_{q_j})}^{(\sigma\oplus e_{p_1})}\Gamma_{(\eta\oplus e_{q_j})}^{(\tau\oplus e_{p_1})}.$$

Definition

A $2^\ell \times 2^m$ matrix M is a *pseudo-signature* if for all σ, τ for which $\mathrm{wt}(\sigma \oplus \tau)$ is even, its entries satisfy the corresponding MGI up to a factor of ± 1 on the right-hand side.

E.g. (matrix-form) standard signatures, clusters of rows, and their transposes are all pseudo-signatures.

The matchgate identities allow us to deduce key linear algebraic relationships between the rows of any pseudo-signature Γ.

Example

Suppose that $d = 2$, $\sigma = 0000$, $\tau = 0011$, $\zeta = 1100$, $\eta = 1111$. Then the MGIs become

$$\Gamma^{0000}_{1100}\Gamma^{0011}_{1111} - \Gamma^{0011}_{1100}\Gamma^{0000}_{1111} = \pm\left(\Gamma^{0001}_{1101}\Gamma^{0010}_{1110} - \Gamma^{0001}_{1110}\Gamma^{0010}_{1101}\right).$$

	1100	1101	1110	1111
0000		0	0	
0001	0			0
0010	0			0
0011		0	0	

Example (Cont'd)

- Rows Γ^{1100} and Γ^{1111} are linearly dependent if Γ^{1101} and Γ^{1110} are linearly dependent.
- Similarly, rows Γ^{0000} and Γ^{1111} are linearly dependent if
  - Γ^{0001} and Γ^{1110}
  - Γ^{0010} and Γ^{1101}
  - Γ^{0100} and Γ^{1011}
  - Γ^{1000} and Γ^{0111}

  are linearly dependent

Lemma

Let $\sigma, \tau \in \{0,1\}^\ell$ *be such that* $\sigma \oplus \tau = \bigoplus_{j=1}^{2d} e_{p_i}$. *If row* $\Gamma^{(\sigma \oplus e_{p_i})}$ *is linearly dependent with row* $\Gamma^{(\tau \oplus e_{p_i})}$ *for all* $1 \leq i \leq 2d$, *then row* Γ^σ *is linearly dependent with row* Γ^τ.

Coordinate-free interpretation: linear relations among wedges of rows of even parity yield linear relations among wedges of rows of odd parity.

Definition

For V a vector space with basis $\{e_j\}$, the *second exterior power of* V, denoted $\Lambda^2 V$, is the vector space given by quotienting $V \otimes V$ by the relation $v \otimes w \sim -w \otimes v$ for all $v, w \in V$. We denote the image of $v \otimes w$ under this quotient map by $v \wedge w$. $\Lambda^2 V$ has basis $\{e_i \wedge e_j\}_{i<j}$.

Explicitly, if $v = \sum v_i e_i$ and $w = \sum w_i e_i$, then

$$v \wedge w = \sum_{i<j} (v_i w_j - v_j w_i) e_i \wedge e_j = \sum_{i<j} \begin{vmatrix} v_i & v_j \\ w_i & w_j \end{vmatrix} e_i \wedge e_j.$$

In particular, v and w are linearly dependent iff $v \wedge w = 0$.

By the MGIs, linear relations among wedges of rows of even parity yield linear relations among wedges of rows of odd parity.

Example

$$\Gamma^{0000} \wedge \Gamma^{1111} = 0$$

implies that

$$\Gamma^{0001} \wedge \Gamma^{1110} - \Gamma^{0010} \wedge \Gamma^{1101} + \Gamma^{0100} \wedge \Gamma^{1011} - \Gamma^{1000} \wedge \Gamma^{0111} = 0$$

By the MGIs, linear relations among wedges of rows of even parity yield linear relations among wedges of rows of odd parity.

Example

$$\mu \cdot (\Gamma^{0000} \wedge \Gamma^{1111}) + \nu \cdot (\Gamma^{0011} \wedge \Gamma^{1100}) = 0$$

implies that

$$\mu \cdot (\Gamma^{0001} \wedge \Gamma^{1110} - \Gamma^{0010} \wedge \Gamma^{1101} + \Gamma^{0100} \wedge \Gamma^{1011} - \Gamma^{1000} \wedge \Gamma^{0111}) \pm$$

$$\nu \cdot (\Gamma^{0010} \wedge \Gamma^{1101} - \Gamma^{0001} \wedge \Gamma^{1110} + \Gamma^{0111} \wedge \Gamma^{1000} - \Gamma^{1011} \wedge \Gamma^{0100}) = 0$$

Claim

Rank rigidity implies cluster existence.

Proof.

Suppose Γ is a $2^\ell \times 2^m$ pseudo-signature of rank 2^K.

Can assume Γ has no proper clusters of rows with the same rank as Γ. Otherwise, if there were such a cluster $Z = \sigma + \{e_{q_1}, ..., e_{q_{\ell'}}\}$, replace Γ by Γ^Z and ℓ by ℓ', and ignore bits in positions outside of $q_1, ..., q_{\ell'}$.

$$
\begin{array}{c|c}
\overbrace{0000}^{\ell-K+1} & \overbrace{0000}^{K-1} \\
0000 & 0001 \\
\vdots & \vdots \\
0000 & 1111
\end{array}
$$

$$
\begin{array}{c|c}
\overbrace{0}^{\ell-K+1=1} & \overbrace{0000}^{K-1} \\
0 & 0001 \\
\vdots & \vdots \\
0 & 1111
\end{array}
$$

$$
\begin{array}{c|c}
\overbrace{0000}^{\ell-K+1} & \overbrace{0000}^{K-1} \\
0000 & 0001 \\
\vdots & \vdots \\
0000 & 1111
\end{array}
$$

$$\begin{array}{c|c}
\overbrace{0000}^{\ell-K+1} & \overbrace{0000}^{K-1} \\
0000 & 0001 \\
\vdots & \vdots \\
0000 & 1111 \\
1110 & 0101
\end{array}$$

$$
\begin{array}{c|c}
\overbrace{000{\color{red}0}}^{\ell-K+1} & \overbrace{0000}^{K-1} \\
000{\color{red}0} & 0001 \\
\vdots & \vdots \\
000{\color{red}0} & 1111 \\
111{\color{red}0} & 0101
\end{array}
$$

$$
\begin{array}{c|c}
\overbrace{0000}^{\ell-K+1} & \overbrace{0000}^{K-1} \\
0000 & 0001 \\
\vdots & \vdots \\
0000 & 1111 \\
1111 & 0000 \\
1111 & 0001 \\
\vdots & \vdots \\
1111 & 1111
\end{array}
$$

$$
\begin{array}{c|c}
\overbrace{0000}^{\ell-K+1} & \overbrace{0000}^{K-1} \\
0000 & 0001 \\
\vdots & \vdots \\
0000 & 1111 \\
1111 & 0000 \\
1111 & 0001 \\
\vdots & \vdots \\
1111 & 1111 \\
1110 & 1010
\end{array}
$$

$$
\begin{array}{c|c}
\overbrace{000\textcolor{red}{0}}^{\ell-K+1} & \overbrace{0000}^{K-1} \\
000\textcolor{red}{0} & 0001 \\
\vdots & \vdots \\
000\textcolor{red}{0} & 1111 \\
1111 & 0000 \\
1111 & 0001 \\
\vdots & \vdots \\
1111 & 1111 \\
111\textcolor{red}{0} & 1010
\end{array}
$$

$$
\begin{array}{c|c}
\overbrace{0000}^{\ell-K+1} & \overbrace{0000}^{K-1} \\
0000 & 0001 \\
\vdots & \vdots \\
0000 & 1111 \\
1111 & 0000 \\
1111 & 0001 \\
\vdots & \vdots \\
1111 & 1111 \\
1110 & 1010
\end{array}
$$

$$
\begin{array}{c|c}
\overbrace{0000}^{\ell-K+1} & \overbrace{0000}^{K-1} \\
0000 & 0001 \\
\vdots & \vdots \\
0000 & 1111 \\
1111 & 0000 \\
1111 & 0001 \\
\vdots & \vdots \\
1111 & 1111
\end{array}
$$

All other rows must be zero. By the MGIs,

$$\Gamma^{0^z \circ 0^{K-1}} \wedge \Gamma^{1^z \circ 0^{K-1}} = 0,$$

a contradiction.

Theorem

If Γ *is a* $2^{K+1} \times 2^m$ *pseudo-signature with rank at least* $2^K + 1$, *then* $\mathrm{rank}(\Gamma) = 2^{K+1}$.

Sketch.

Inductively, we know that Γ contains a cluster Z of 2^K linearly independent rows, say $0^{K+1} \oplus \{e_2, ..., e_{K+1}\}$. Because $\mathrm{rank}(\Gamma) \geq 2^K + 1$, there exists a row outside the linear span of Z.

Even columns:	Odd columns:
0000	0001
0011	0010
0101	0100
0110	0111
1001	1000
1010	1011
1100	1101
1111	1110

Even columns:	Odd columns:
0000	0001
0011	0010
0101	0100
0110	0111
1001	1000
1010	1011
1100	1101
1111	1110

Suppose Γ^{1001} lay in the span of the red rows, so

$$\Gamma^{1001} \wedge \Gamma^{0000} = \sum_{\sigma \text{ red, even}} a_\sigma \cdot \left(\Gamma^\sigma \wedge \Gamma^{0000}\right).$$

LHS: $\Gamma^{1000} \wedge \Gamma^{0001}$

RHS: $\Gamma^{0***} \wedge \Gamma^{0***}$

Even columns:	Odd columns:
0000	0001
0011	0010
0101	0100
0110	0111
1001	1000
$\boxed{1010}$	1011
1100	1101
1111	1110

Suppose Γ^{1010} lay in the span of the red rows, so

$$\Gamma^{1010} \wedge \Gamma^{0000} = \sum_{\sigma \text{ red, even}} a_\sigma \cdot \left(\Gamma^\sigma \wedge \Gamma^{0000}\right).$$

LHS: $\Gamma^{1000} \wedge \Gamma^{0010}$

RHS: $\Gamma^{0***} \wedge \Gamma^{0***}$, $\Gamma^{1000} \wedge \Gamma^{0001}$

Even columns:	Odd columns:
0000	0001
0011	0010
0101	0100
0110	0111
1001	1000
1010	1011
$\boxed{1100}$	1101
1111	1110

Suppose Γ^{1100} lay in the span of the red rows, so

$$\Gamma^{1100} \wedge \Gamma^{0000} = \sum_{\sigma \text{ red, even}} a_\sigma \cdot \left(\Gamma^\sigma \wedge \Gamma^{0000}\right).$$

LHS: $\Gamma^{1000} \wedge \Gamma^{0100}$

RHS: $\Gamma^{0***} \wedge \Gamma^{0***}$, $\Gamma^{1000} \wedge \Gamma^{0001}$, $\Gamma^{1000} \wedge \Gamma^{0010}$

Even columns:	Odd columns:
0000	0001
0011	0010
0101	0100
0110	0111
1001	1000
1010	1011
1100	1101
1111	1110

Suppose Γ^{1011} lay in the span of the red rows, so

$$\Gamma^{1011} \wedge \Gamma^{0001} = \sum_{\sigma \text{ red, odd}} a_\sigma \cdot \left(\Gamma^\sigma \wedge \Gamma^{0001}\right).$$

LHS: $\Gamma^{1001} \wedge \Gamma^{0011}$
RHS: $\Gamma^{0***} \wedge \Gamma^{0***}$, $\Gamma^{0000} \wedge \Gamma^{1001}$

Even columns:	Odd columns:
0000	0001
0011	0010
0101	0100
0110	0111
1001	1000
1010	1011
1100	$\boxed{1101}$
1111	1110

Suppose Γ^{1101} lay in the span of the red rows, so

$$\Gamma^{1011} \wedge \Gamma^{0001} = \sum_{\sigma \text{ red, odd}} a_\sigma \cdot \left(\Gamma^\sigma \wedge \Gamma^{0001}\right).$$

LHS: $\Gamma^{1001} \wedge \Gamma^{0101}$

RHS: $\Gamma^{0***} \wedge \Gamma^{0***}$, $\Gamma^{0000} \wedge \Gamma^{1001}$, $\Gamma^{1001} \wedge \Gamma^{0011}$

Even columns:	Odd columns:
0000	0001
0011	0010
0101	0100
0110	0111
1001	1000
1010	1011
1100	1101
1111	$\boxed{1110}$

Suppose Γ^{1110} lay in the span of the red rows, so

$$\Gamma^{1110} \wedge \Gamma^{0001} = \sum_{\sigma \text{ red, odd}} a_\sigma \cdot \left(\Gamma^\sigma \wedge \Gamma^{0001}\right).$$

LHS: $\Gamma^{1111} \wedge \Gamma^{0000}$, $\Gamma^{1100} \wedge \Gamma^{0011}$, $\Gamma^{1010} \wedge \Gamma^{0101}$, $\Gamma^{0110} \wedge \Gamma^{1001}$
RHS: $\Gamma^{0***} \wedge \Gamma^{0***}$, $\Gamma^{0000} \wedge \Gamma^{1001}$, $\Gamma^{1001} \wedge \Gamma^{0011}$, $\Gamma^{1001} \wedge \Gamma^{0101}$

Even columns:	Odd columns:
0000	0001
0011	0010
0101	0100
0110	0111
1001	1000
1010	1011
1100	1101
$\boxed{1111}$	1110

Suppose Γ^{1111} lay in the span of the red rows, so

$$\Gamma^{1111} \wedge \Gamma^{0000} = \sum_{\sigma \text{ red, even}} a_\sigma \cdot \left(\Gamma^\sigma \wedge \Gamma^{0000}\right).$$

LHS: $\Gamma^{1110} \wedge \Gamma^{0001}$, $\Gamma^{1101} \wedge \Gamma^{0010}$, $\Gamma^{1011} \wedge \Gamma^{0100}$, $\Gamma^{0111} \wedge \Gamma^{1000}$
RHS: $\Gamma^{0***} \wedge \Gamma^{0***}$, $\Gamma^{1000} \wedge \Gamma^{0001}$, $\Gamma^{1000} \wedge \Gamma^{0010}$, $\Gamma^{1000} \wedge \Gamma^{0100}$

Even columns:	Odd columns:
0000	0001
0011	0010
0101	0100
0110	0111
1001	1000
1010	1011
1100	1101
1111	1110

Theorem

Suppose $\ell > K+1$. *If* Γ *is a* $2^\ell \times 2^m$ *pseudo-signature of rank* $\geq 2^K+1$, *then there exists a cluster* $Z \subsetneq \{0,1\}^\ell$ *for which* Γ^Z *is also of rank* $\geq 2^K+1$.

Suppose to the contrary. Inductively we know Γ has a cluster Z of 2^K linearly independent rows, say $0^\ell + \{e_{p_1}, ..., e_{p_K}\}$.

$$
\begin{array}{c|c}
\overbrace{0000}^{\ell-K} & \overbrace{0000}^{K} \\
0000 & 0001 \\
\vdots & \vdots \\
0000 & 1111
\end{array}
$$

$$
\begin{array}{c|c}
\overbrace{0000}^{\ell-K} & \overbrace{0000}^{K} \\
0000 & 0001 \\
\vdots & \vdots \\
0000 & 1111 \\
1110 & 0101
\end{array}
$$

$$\begin{array}{c|c} \overbrace{00000}^{\ell-K} & \overbrace{0000}^{K} \\ 00000 & 0001 \\ \vdots & \vdots \\ 00000 & 1111 \\ 11100 & 0101 \end{array}$$

$$
\begin{array}{c|c}
\overbrace{0000}^{\ell-K} & \overbrace{0000}^{K} \\
0000 & 0001 \\
\vdots & \vdots \\
0000 & 1111 \\
1111 & 0000
\end{array}
$$

$$
\begin{array}{c|c}
\overbrace{\color{red}{0000}}^{\ell-K} & \overbrace{\color{red}{0000}}^{K} \\
0000 & 0001 \\
\vdots & \vdots \\
0000 & 1111 \\
\color{red}{1111} & \color{red}{0000}
\end{array}
$$

To show Γ^{0^ℓ} and $\Gamma^{1^{\ell-K}\circ 0^\ell}$ are linearly dependent, by MGIs it's enough to show:

Lemma

$\Gamma^{0^\ell \oplus e_j} = 0$ *for all* $j \neq p_1, ..., p_K$.

Proof.

For $i \in \{p_1, ..., p_K\}$ and $j \notin \{p_1, ..., p_K\}$, define:

- T_i: all rows u for which $u_i = 0$
- T_i^j: all rows u for which $u_i = u_j = 0$
- Z_i: $Z \cap T_i$

Note that

$$Z_i \subset T_i^j \subset T_i.$$

So inductively, proper cluster T_i^j has rank a power of two, either 2^{K-1} or 2^K. □

Proof (Cont'd).

- If $\text{rank}(T_i^j) = 2^{K-1}$, then $\text{span}(T_i^j) = \text{span}(Z_i)$. This is true for all $i \in \{p_1, ..., p_K\}$, so

$$\Gamma^{0^\ell \oplus e_j} \in \cap_{i=1}^{K} \text{span}(Z_i) = \text{span}(\{\Gamma^{0^\ell}\}).$$

- If $\text{rank}(T_i^j) = 2^K$, then $\text{span}(T_i) = \text{span}(T_i^j) \subset \text{span}(Z)$, a contradiction.

□

Theorem (Fu/Yang '13)

Suppose a basis collapse theorem holds on domain size 2*. Then if a holographic algorithm uses a* $2^{\ell} \times k$ *basis of rank* 2*, then the same collapse theorem holds for this holographic algorithm.*

Theorem

Suppose a basis collapse theorem holds on domain size r. Then if a holographic algorithm uses a $2^\ell \times k$ basis of rank r, then the same collapse theorem holds for this holographic algorithm.

By rank rigidity, $\underline{G}(t)$ must have rank a power of two. If G is a full-rank signature, by

$$\underline{G}(t) = MG(t)(M^T)^{\otimes(n-1)},$$

we know M must have rank a power of two. So if $k \neq 2^K$, we're done inductively by the collapse theorem for domain size 2^K, where $K = \lfloor \log_2 k \rfloor$.

- Work out the case where no full-rank matchgate exists
- Use the collapse theorem to initiate a study of holographic algorithms over higher domains

Thank you to:

- Professor Leslie Valiant (Harvard University)
- Professor Jin-Yi Cai (University of Wisconsin)
- Harvard Herchel-Smith Research Fellowship
- Simons Institute for Computing