Computer Vision, Convolutions, Complexity and Algebraic Geometry

D. V. Chudnovsky, G.V. Chudnovsky
IMAS
Polytechnic Institute of NYU
6 MetroTech Center
Brooklyn, NY 11201

December 6, 2012

Fast Multiplication: glimpse

Gauss 1801-1809
Not dwarf planet Ceres, but an asteroid Pallas

$$(a_0 + a_1 i) \cdot (b_0 + b_1 i) = a_0 \cdot b_0 - a_1 \cdot b_1 + (a_0 \cdot b_1 + a_1 \cdot b_0)i$$

instead

$$a_0 \cdot b_0$$
$$a_1 \cdot b_1$$
$$(a_0 + a_1) \cdot (b_0 + b1)$$
$$(a_0 + a_1) \cdot (b_0 + b1) - (a_0 \cdot b_0 + a_1 \cdot b_1)$$

3 multiplies and 5 add/sub instead of 4 multiply and 2 add/sub.

Fast Multiplication: the beginning

Kolmogorov (Fall, 1960)
Karatsuba method

$$(A_0 + A_1 x) \cdot (B_0 + B_1 x) =$$
$$(A_0 B_0) + ((A_0 + A_1)(B_0 + B_1) - (A_0 B_0 + A_1 B_1))x + (A_1 B_1)x^2$$

Iterate to $O(n^{\log_2(3)})$ complexity of $n \times n$ convolution and thus to n-bit integer multiplication.

Polynomials and Interpolation

Karatsuba method as evaluation/interpolation at $x = 0, 1, \infty$.
Toom(-Cook) method

Multiply $P(x) = a_0 + \cdots + a_{n-1}x^{n-1}$ and
$Q(x) = b_0 + \cdots + b_{m-1}x^{m-1}$ using interpolation of the polynomial
$P(x) \cdot Q(x)$ at $n + m - 1$ points $x = 0, 1, \ldots, n + m - 2$.

Leads to $n + m - 1$ multiplication of linear forms in a_i, b_j but these linear forms have rational coefficients of the height $O(n^n)$.

Early papers on fast convolution

A. Karatsuba & Yu. Ofman, *Multiplication of many-digital numbers by automatic computers*, Doklady Akad. Nauk SSSR 145 (1962) 293–294.

A. L. Toom, *The complexity of a scheme of functional elements realizing the multiplication of integers*, Soviet Mathematics 3 (1963), 714–716.

S. A. Cook, *On the Minimum Computation Time of Functions*, PhD thesis, Harvard University, 1966

Bilinear Forms and Multiplicative Complexity

Strassen (1969) 7 for 8

Bilinear forms in variables $x_0, \ldots, x_{n-1}$ and $y_0, \ldots, y_{m-1}$

$$z_k = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} T_{i,j,k} \cdot x_i \cdot y_j : k = 1, \ldots, s$$

over some algebra $\mathbf{A}$ (typically a field $\mathbf{F}$).

The (multiplicative) complexity of computation of these forms over $\mathbf{A}$, $\mu_{\mathbf{A}}(T)$ is the rank of the tensor T over $\mathbf{A}$.

Bilinear Forms and Multiplicative Complexity (cont.)

Here rank 1 tensor is just like rank 1 matrix:

$$T_{i,j,k} = a_i \cdot b_j \cdot c_k$$

for scalars a_i, b_j, c_k from $\mathbf{A}$.

The tensor T has a rank μ if it is the sum of μ (and not less) of rank 1 tensors. This defines the multiplicative complexity of T over $\mathbf{A}$:

$$\mu = \mu_{\mathbf{A}}(T)$$

In the matrix form, we have A, B, C matrices (from $M_{\mu,n}(\mathbf{A}), M_{\mu,m}(\mathbf{A}), M_{s,\mu}(\mathbf{A})$) such that:

$$\vec{z} = C \cdot (A\vec{x} \otimes B\vec{y})$$

Complexity of one-dimensional convolution over infinite fields

S. Winograd's Theorems (1976-1977)

Theorem

Every minimal multiplicative complexity $(n + m - 1)$ *algorithm of computation of multiplication of* $P(x) = a_0 + \cdots + a_{n-1}x^{n-1}$ *and* $Q(x) = b_0 + \cdots + b_{m-1}x^{m-1}$ *over the field* $\mathbf{F}$ *is reduced to an interpolation of the polynomial* $P(x) \cdot Q(x)$ *at* $n + m - 1$ *distinct points in* $\mathbf{PF}^1$ *("interpolation algorithm")*

Similar result and complexity holds for a polynomial multiplication of $P(x)$ and $Q(x)$ of degrees $n - 1$ modulo an irreducible polynomial $R(x)$ of degree n, i.e.

$$\mu_{\mathbf{F}}(\mathbf{K}) = 2[\mathbf{K} : \mathbf{F}] - 1$$

for a finite extension $\mathbf{K}$ of an infinite field $\mathbf{F}$.

Towers of fields – Schönhage and Strassen

Over finite fields $\mathbf{F}_q$ and $n > q$ (or $\mathbf{Z}$) the minimal complexity algorithm does not work at all.
The idea then is to extend the fields $\mathbf{F}_q$ (or $\mathbf{Z}[\mathbf{x}]$) so we get enough interpolation points (especially if these interpolation points are roots of unity in corresponding extensions).
The standard tower of extensions is generated by "FFT-tower":

$$x^{2^{n+1}} + 1 \text{ over } x^{2^n} + 1$$

This method, initially used by Schönhage and Strassen (1971), and Nussbaumer (1976) gives, "log-linear" complexity.
When applied to integer multiplication it gives the famous Schönhage-Strassen bound for the n-bit multiplication

$$O(n \log n \log \log n)$$

improved by Fürer (2007) to $O(n \log n 2^{O(\log^* n)})$.

Papers on complexity of bilinear forms

V. Strassen, *Gaussian elimination is not optimal*, Numer. Math. **13** (1969) 354–356. 5–685.

S. Winograd, *Some bilinear forms whose multiplicative complexity depends on the field of constants*, Math. Systems Theory **10** (1977) 169–180.

C. Fiduccia & Y. Zalcstein, *Algebras having linear multiplicative complexities*, J. Assoc. Comput. Mach. **24** (1977) 311–331.

V. Strassen, *Rank and optimal computation of generic tensors*, Linear Algebra Appl. **52/53** (1983) 64

H. F. de Groote, *Lectures on the complexity of bilinear problems*, Lecture Notes in Comp. Science **245**, Springer-Verlag, 1987.

Linear Codes from Multiplication Algorithms

When the tensor

$$T = T_{i,j,k} \; i,j,k = 1,\ldots n$$

defines the multiplication in the (n-dimensional) algebra $\mathbf{A}$ over the finite field $\mathbf{F}_q$, any of its representation as a sum of rank 1 tensors gives rise to q-ary linear codes with special distance properties.

For example, if the algebra $\mathbf{A}$ has no divisors of 0, then all three $\mu \times n$ matrices A, B, C^T are generators of q-ary linear codes of length $\mu = \mu_{\mathbf{F}_q}(\mathbf{A})$ and the dimension n with the minimal weight (distance) of n.

Moreover, three codes, generated by A, B, C^T are *intersecting*, i.e. any code vectors from any two of these codes have non-empty common support.

Linear Codes and Better Lower Bounds

Applying Elias bound to binary codes, we get the lower bound for multiplication complexity:

$$\mu_{\mathbf{F}_2}(\mathbf{F}_{2^n}) \geq 3.5275.. \cdot n$$

In particular, a weak corollary:

$$\mu_{\mathbf{F}_2}(n \times n) \geq 3.5275.. \cdot n$$

In general, multiplication of polynomials of degree n modulo any polynomial of degree $\geq n$ over $\mathbf{F}_q$ has a multiplicative complexity

$$\mu_{\mathbf{F}_q} \geq (2 + \frac{1}{q-1}) \cdot n - o(n)$$

Hankel matrices..

Interpolate on Algebraic Curves of genus $g > 0$

In 1986 we needed better algorithms of integer multiplications, i.e. faster convolutions over $\mathbf{Z}$, or $\mathbf{Z}[\frac{1}{2}]$, $\mathbf{Z}[\frac{1}{3}]$, etc. – start with finite fields.

Do all minimal complexity convolution algorithms come from interpolations on algebraic surfaces?

Needed curves with many points over finite fields.
Ihara (1981) proved first results for modular curves.

$$N_q(g) = \max\{|X(\mathbf{F}_q)| : X \text{ is a curve of genus } g \text{ over } \mathbf{F}_q\}$$

$$\begin{aligned} A^-(q) &= \liminf_{g\to\infty} \frac{N_q(g)}{g} \\ A^+(q) &= \limsup_{g\to\infty} \frac{N_q(g)}{g} \end{aligned}$$

Algebraic Curves with many points on finite fields

$$\begin{aligned} A^+(q) &= \sqrt{q} - 1 \text{ for a square } q \\ A^-(q) &> c \cdot \log q \text{ for every } q \end{aligned}$$

The towers of algebraic curves providing the tight $A^+(q)$ bounds are typically modular (of elliptic, Shimura, or Drinfeld types). Garcia-Stichtenoth tower over $\mathbf{F}_{q^2}$ is defined by the sequence $(F_1, F_2, \dots)$ where

$$\begin{aligned} F_{k+1} &= F_k(z_{k+1}) \\ z_{k+1}^q + z_{k+1} &= \xi^{q+1} \\ \xi_k &= \frac{z_k}{\xi_{k-1}} \text{ in } F_k (k \geq 1) \end{aligned}$$

Our papers and a general review

D. V. & G. V. Chudnovsky, *Algebraic complexities and algebraic curves over finite fields*, Proc. Nat. Acad. Sci. USA **84** (1987) 1739–1743.

D. V. & G. V. Chudnovsky, *Algebraic complexities and algebraic curves over finite fields*, J. Complexity **4** (1988) 285–316.

P. Bürgisser, M. Clausen & A. Shokrollahi, *Algebraic complexity theory*, Grundlehren der Math. Wissenschaften **315**, Springer-Verlag, 1997.

Algorithms over Finite Fields

Our general interpolation results from 1986-1987 used mostly degree 1 divisors on curves $X(\mathbf{F}_q)$ – similar to distinct point interpolation on $\mathbf{PF}^1$.

The basic upper bound we obtained this way:

$$\mu_{\mathbf{F}_q}(\mathbf{F}_{q^n}) \ \leq \ 2(1+\frac{1}{\sqrt{q}-3})n+o(n)$$

for a square $q \geq 16$ and $n \to \infty$.

Recent interest in this problem prompted refined analysis of high degree divisors (cf. with the C.R.T) and high order interpolations.

Improved upper bounds

The recent series of results show (after corrections..) a slightly better bound:

$$\mu_{\mathbf{F}_q}(\mathbf{F}_{q^n}) \;\leq\; 2(1+\frac{1}{\sqrt{q}-2})n + o(n)$$

for a square $q \geq 9$ and sufficiently large n.

The original 1987 bound as well as these improved ones give a general linear bound for a one-dimensional convolution:

$$\mu_{\mathbf{F}_q}(n \times n) \;\leq\; C_q \cdot n$$

The upper bound on C_q (for a square $q \geq 9$) is $4(1+\frac{1}{\sqrt{q}-2})$, and a low bound is $3+\epsilon(q)$, for any q. E.g. for $q=3$ the low bound is 3.005.

Recent papers on convolutions and algebraic geometry

Hugues Randriambololona, *Bilinear complexity of algebras and the Chudnovsky–Chudnovsky interpolation method* Journal of Complexity 28 (2012) 489–517

S. Ballet, *On the tensor rank of the multiplication in the finite fields*, J. Number Theory **128** (2008) 1795–1806.

S. Ballet, D. Le Brigand & R. Rolland, "On an application of the definition field descent of a tower of function fields", in: F. Rodier & S. Vladut (eds.), *Proceedings of the Conference "Arithmetic, Geometry and Coding Theory" (AGCT 2005)*, Séminaires et Congrès **21**, Société Mathématique de France, 2010, pp. 187–203.

S. Ballet & J. Pieltant, *On the tensor rank of multiplication in any extension of* $\mathbf{F}_2$, J. Complexity **27** (2011) 230–245.

Recent papers on convolutions and algebraic geometry (cont.)

S. Ballet, C. Ritzenthaler & R. Rolland, *On the existence of dimension zero divisors in algebraic function fields defined over* $\mathbf{F}_q$, Acta Arith. **143** (2010) 377–392.

M. Cenk & F. Özbudak, *On multiplication in finite fields*, J. Complexity **26** (2010) 172–186.

J. Chaumine, "Multiplication in small finite fields using elliptic curves", in: J. Chaumine, J. Hirschfeld & R. Rolland (eds.), *Algebraic geometry and its applications, Proceedings of the first SAGA conference (Papeete, France, 7-11 May 2007)*, Ser. Number Theory Appl. **5**, World Sci. Publ., 2008, pp. 343–350.

Bounds for One-Dimensional Convolutions over $\mathbf{F}_3$ – part 1

Convolution	Complexity	Lower Bound method	Upper Bound Method
2×2	$\mu_{\mathbf{F}_3}(2 \times 2) = 3$	Dimensions	Standard (Karatsuba)
3×2	$\mu_{\mathbf{F}_3}(3 \times 2) = 4$	Dimensions	Standard (Toom-Cook)
3×3	$\mu_{\mathbf{F}_3}(3 \times 3) = 6$	Codes	C.R.T
4×2	$\mu_{\mathbf{F}_3}(4 \times 2) = 6$	Codes	C.R.T
4×3	$\mu_{\mathbf{F}_3}(4 \times 3) = 7$	Exhaustive	C.R.T
4×4	$\mu_{\mathbf{F}_3}(4 \times 4) = 9$	Codes	C.R.T
5×2	$\mu_{\mathbf{F}_3}(5 \times 2) = 7$	Exhaustive	C.R.T
5×3	$\mu_{\mathbf{F}_3}(5 \times 3) = 9$	Hankel matrices	C.R.T
5×4	$\mu_{\mathbf{F}_3}(5 \times 4) = 10$	Hankel matrices	C.R.T
5×5	$\mu_{\mathbf{F}_3}(5 \times 5) = 12$	Hankel matrices	C.R.T
6×2	$\mu_{\mathbf{F}_3}(6 \times 2) = 8$	Codes	C.R.T
6×3	$\mu_{\mathbf{F}_3}(6 \times 3) = 10$	Hankel matrices	C.R.T
6×4	$\mu_{\mathbf{F}_3}(6 \times 4) = 12$	Hankel matrices (New)	C.R.T
6×5	$\mu_{\mathbf{F}_3}(6 \times 5) = 13$	Hankel matrices (New)	C.R.T
6×6	$\mu_{\mathbf{F}_3}(6 \times 6) = 15$	Hankel matrices (New)	Field Extension
7×2	$\mu_{\mathbf{F}_3}(7 \times 2) = 10$	Codes	C.R.T
7×3	$\mu_{\mathbf{F}_3}(7 \times 3) = 12$	Hankel matrices (New)	C.R.T
7×4	$\mu_{\mathbf{F}_3}(7 \times 4) = 13$	Hankel matrices (New)	C.R.T
7×7	$18 \leq \mu_{\mathbf{F}_3}(7 \times 7) \leq 19$	Hankel matrices (New)	C.R.T
8×2	$\mu_{\mathbf{F}_3}(8 \times 2) = 11$	Codes	C.R.T
8×3	$\mu_{\mathbf{F}_3}(8 \times 3) = 13$	Hankel matrices (New)	C.R.T
8×8	$19 \leq \mu_{\mathbf{F}_3}(8 \times 8) \leq 23$	Hankel matrices (New)	Field Extension

Bounds for One-Dimensional Convolutions – part 2

Some new (mostly low bound) special interesting cases; a single new type of upper bound algorithm (6 non-equivalent solutions)

Convolution	Complexity	Lower Bound	Upper/State
6×5	$15 \leq \mu_{\mathbf{F}_2}(6 \times 5) \leq 16$	1D Hankel	Upper Bound C.R.T.
6×6	$\mu_{\mathbf{F}_2}(6 \times 6) = 17$	1D Hankel	Open since 1987
7×7	$19 \leq \mu_{\mathbf{F}_2}(7 \times 7) \leq 22$	1D Hankel	Upper Bound C.R.T.
$\mathbf{F}_{2^7}$	$18 \leq \mu_{\mathbf{F}_2}(\mathbf{F}_{2^7}) \leq 22$	1D Hankel	Upper Bound C.R.T.
mod $(Q_2(x)^2, 3)$	$\mu_{\mathbf{F}_3}(Q_2(x)^2) = 9$	Codes	New: Open since 1987

In the last case $A = B$ arise from ternary code generator matrix defining the algebra with 0 divisors up to the isotopy

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 & 1 & 2 & 0 \end{pmatrix}$$

Why to convolve

In 60s to 80s mostly for DSP – serial communications, audio, big-num multiply, building multi-dimensional FFT and filtering from 1D.

We finally approach the need for live visualization using two-dimensional filtering and convolutions.

Filtering algorithms from convolution: just flip the indices in the tensor $T_{i,j,k}$ to $T_{k,j,i}$.

...

Multidimensional Convolutions

Multiplication of polynomials in d variables $\vec{t} = (t_1, \ldots, t_d)$
$P(\vec{t}) = \sum_{i_1,\ldots,i_d=0}^{N-1} a_{i_1,\ldots,i_d} \cdot t_1^{i_1} \cdots t_d^{i_d}$ and
$Q(\vec{t}) = \sum_{j_1,\ldots,j_d=0}^{M-1} b_{j_1,\ldots,j_d} \cdot t_1^{j_1} \cdots t_d^{j_d}$
The coefficient list $c_{\vec{k}}$ of the polynomial product $P(\vec{t}) \cdot Q(\vec{t})$ is called an (acyclic) convolution of coefficient arrays $a_{\vec{i}}$ and $b_{\vec{j}}$:

$$c_{\vec{k}} = \sum_{\vec{i}+\vec{j}=\vec{k}} a_{\vec{i}} \cdot b_{\vec{j}}$$

If $M = N$ and this is a polynomial multiplication modulo $(t_1^N - 1) \cdots (t_d^N - 1)$ the resulting array $c_{\vec{k}}$ is called a cyclic convolution.

Complexity of Multidimensional Convolutions – Simple Bounds

Clearly, simple upper and low bounds of the d-dimensional convolution complexity are:

$$N^d \cdot M^d \geq \mu_{\mathbf{A}}(\underbrace{N \times \cdots \times N}_{d}, \underbrace{M \times \cdots \times M}_{d}) \geq (N + M - 1)^d$$

(the number of independent bilinear forms $c_{\vec{k}}$).

Theorem (Simple Bound)

Every minimal multiplicative complexity $(N + M - 1)^d$ *algorithm of computation of multiplication of* $P(\vec{t})$ *and* $Q(\vec{t})$ *over the field* $\mathbf{F}$ *with at least* $N + M - 2$ *elements is reduced to an* d*-dimensional interpolation of the polynomial* $P(\vec{t}) \cdot Q(\vec{t})$ *at* $(N + M - 1)^d$ *distinct points from* $\mathbf{PF}^d$.

Unlike $d = 1$, not all $(N + M - 1)^d$ distinct points from $\mathbf{PF}^d$ give rise to this algorithm!

2D Hankel matrices

The complexity bounds in the 2-dimensional case can be obtained using the 2D Hankel matrices ("Hankel-block-Hankel") having the form

$$\mathbf{H2} = (h_{\vec{i}+\vec{j}})_{\vec{i},\vec{j}=\vec{0}}^{(N-1)^d,(M-1)^d}$$

over finite fields $\mathbf{F}_q$.

Not much is known – study of two-dimensional recurrences over finite fields.

What is the number of 2D Hankel matrices of rank r? Is it a function of q only, when r is less than $\min((N-1)^d, (M-1)^d)$?

Better Bounds – can curves (surfaces) help us

Only a few examples with useful interpolation on surfaces.
Still, one can get linear upper bounds (with respect to the minimal complexity = number of terms).

For this reduce d-dimensional polynomial multiplication to a one-dimensional one using a "Kronecker trick" – a substitution of a d-dimensional monomial by a one-dimensional "sparse" monomial.

$$t_1^{i_1} \cdots t_d^{i_d} \rightarrow t^{i_1+(2N-1)i_2+\cdots+(2N-1)^{d-1}i_d}$$

This allows us to get a far from optimal, but still a linear bound, on the complexity of the d-dimensional convolution:

Better Bounds – curves can help us

$$\mu_{\mathbf{A}}(\underbrace{N \times \cdots \times N}_{d}, \underbrace{M \times \cdots \times M}_{d}) \leq C_q \cdot (2N-1)^d$$

where C_q is our linear factor for the 1D convolution bound; somewhere between $3 + \epsilon(q)$ and $4 + \delta(q)$, for $\epsilon, \delta > 0$ ($\to 0$ as $q \to \infty$).

This bound is far from tight – sparse polynomial multiplication is used here.

What about peculiar Weierstrass gaps – short intervals repeated in arithmetic progressions (like $3, 4, 8, 9$ in the 3×3 by 3×3 filters).

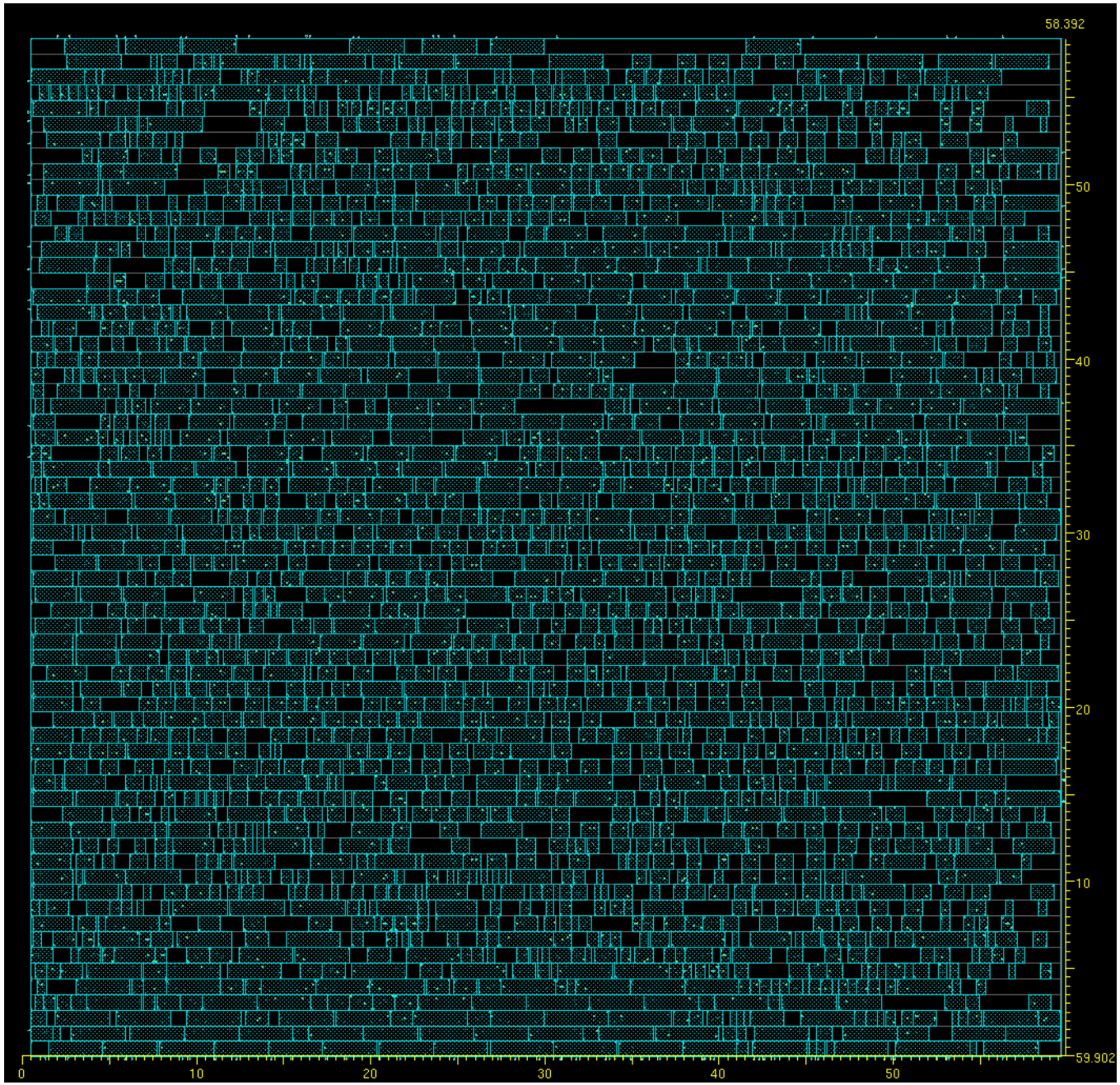
Practical cases of 2D Convolutions

New results; especially for lower bounds.

2D Convolution	Complexity	Lower Bound method
2×2 by 2×2	$\mu_{\mathbf{F}_2}(2 \times 2, 2 \times 2) = 9$	Dimensions
3×2 by 2×2	$\mu_{\mathbf{F}_2}(3 \times 2, 2 \times 2) = 15$	2D Hankel
3×2 by 3×2	$\mu_{\mathbf{F}_2}(3 \times 2, 3 \times 2) = 18$	2D Hankel
3×2 by 2×3	$19 \leq \mu_{\mathbf{F}_2}(3 \times 2, 2 \times 3) \leq 22$	2D Hankel
2×2 by 3×3	$19 \leq \mu_{\mathbf{F}_2}(2 \times 2, 3 \times 3) \leq 22$	2D Hankel
4×2 by 2×2	$\mu_{\mathbf{F}_2}(4 \times 2, 2 \times 2) = 18$	2D Hankel
3×3 by 3×3	$29 \leq \mu_{\mathbf{F}_2}(3 \times 3, 3 \times 3) \leq 33$	2D Hankel
4×2 by 3×2	$21 \leq \mu_{\mathbf{F}_2}(4 \times 2, 3 \times 2) \leq 24$	2D Hankel
4×2 by 2×3	$23 \leq \mu_{\mathbf{F}_2}(4 \times 2, 2 \times 3) \leq 27$	2D Hankel

Not much chance for a tight 8×8 by 8×8 bound.

But what is the size and power...



Wires...