Complexity of Matrix Multiplication and Bilinear Problems

François Le Gall

Graduate School of Informatics
Kyoto University

ADFOCS17 - Lecture 3
24 August 2017

Overview of the Lectures

✓ Fundamental techniques for fast matrix multiplication (1969~1987)

- Basics of bilinear complexity theory: exponent of matrix multiplication, Strassen's algorithm, bilinear algorithms
- First technique: tensor rank and recursion
- Second technique: border rank
- Third technique: the asymptotic sum inequality
- Fourth technique: the laser method

Lecture 1

✓ Recent progress on matrix multiplication (1987~)

Lecture 2

- Laser method on powers of tensors
- Other approaches
- Lower bounds
- Rectangular matrix multiplication

Lecture 3

✓ Applications of matrix multiplications, open problems

Overview of the Lectures

✓ Fundamental techniques for fast matrix multiplication (1969~1987)

- Basics of bilinear complexity theory: exponent of matrix multiplication, Strassen's algorithm, bilinear algorithms
- First technique: tensor rank and recursion
- Second technique: border rank
- Third technique: the asymptotic sum inequality

Lecture 1

- Fourth technique: the laser method

✓ Recent progress on matrix multiplication (1987~)

Lecture 2

- Laser method on powers of tensors
- Other approaches
- Lower bounds
- Rectangular matrix multiplication

Lecture 3

✓ Applications of matrix multiplications, open problems

Lower Bounds for Matrix Multiplication

$R(\langle 2,2,2\rangle) \geq 7$ [Hopcroft and Kerr 1971] [Winograd 1971]

$\underline{R}(\langle 2,2,2\rangle) \geq 7$ [Landsberg 2005]

$\underline{R}(\langle n,n,n\rangle) \geq \frac{3}{2}n^2$ [Strassen 1983]

$\underline{R}(\langle n,n,n\rangle) \geq \frac{3}{2}n^2 + \frac{1}{2}n - 1$ [Lickteig 1984]

$\underline{R}(\langle n,n,n\rangle) \geq \frac{3}{2}n^2 - 2$ [Bürgisser, Ikenmeyer 2011] ← using geometric complexity theory

$\underline{R}(\langle n,n,n\rangle) \geq 2n^2 - n$ [Landsberg, Ottaviani 2011]

$R(\langle n,n,n\rangle) \geq \frac{5}{2}n^2 - 3n$ [Bläser 1999]

$R(\langle n,n,n\rangle) \geq 3n^2 - 4n^{2/3} - n$ [Landsberg 2012]

$R(\langle n,n,n\rangle) \geq 3n^2 - 2\sqrt{2}n^{2/3} - 3n$ [Massaranti, Raviolo 2012]

Lower Bounds for Matrix Multiplication

Theorem ([Raz 2002])

> Any arithmetic circuit that computes the product of two n x n real matrices has size $\Omega(n^2 \log n)$, as long as the circuit does not use products with field elements of absolute value larger than 1.

Space-Time tradeoff (see, e.g., [Abrahamson 1991])

> For any algebraic algorithm computing the product of two n x n matrices using S space and T time we have $ST = \Omega(n^3)$.

Any subcubic-time algorithm for matrix multiplication has superlogarithmic space complexity.

Trivial algorithm: $T = O(n^3)$ $S = O(\log n)$

Strassen algorithm: $T = O(n^{2.81})$ $S = O(n^2)$

same quadratic space complexity for all the other fast algorithms we studied

Overview of the Lectures

✓ Fundamental techniques for fast matrix multiplication (1969~1987)

- Basics of bilinear complexity theory: exponent of matrix multiplication, Strassen's algorithm, bilinear algorithms
- First technique: tensor rank and recursion
- Second technique: border rank
- Third technique: the asymptotic sum inequality
- Fourth technique: the laser method

Lecture 1

✓ Recent progress on matrix multiplication (1987~)

Lecture 2

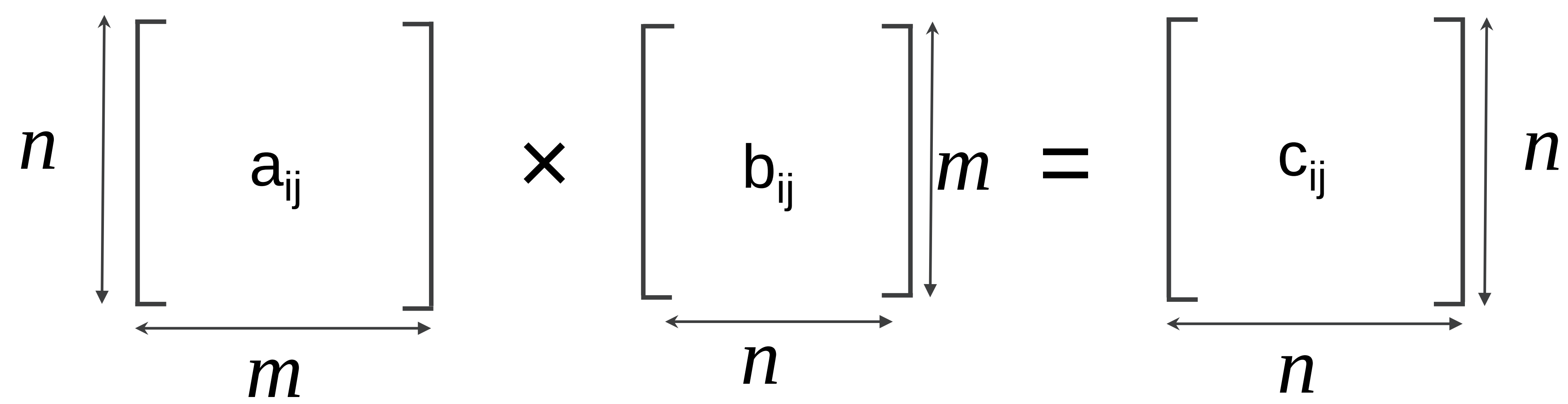
- Laser method on powers of tensors
- Other approaches
- Lower bounds
- Rectangular matrix multiplication

Lecture 3

✓ Applications of matrix multiplications, open problems

Rectangular Matrix Multiplication

Compute the product of an n x m matrix A and an m x n matrix B



m multiplications and $(m$-1) additions

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}$$ for all $1 \le i \le n$ and $1 \le j \le n$

Trivial algorithm: $n^2(2m\text{-}1)=O(mn^2)$ arithmetic operations

Rectangular Matrix Multiplication

Compute the product of an n x m matrix A and an m x n matrix B

Compute the product of an m x n matrix A and an n x n matrix B

Compute the product of an n x n matrix A and an n x m matrix B

same algebraic complexity (as seen yesterday)

✓ The problem (with $m \neq n$) appears as the bottleneck in many applications:

- linear algebra problems
- all-pairs shortest path problems
- dynamic computation of the transitive closure of a graph
- detecting directed cycles in a graph
- computational geometry (colored intersection searching)
- computational complexity (circuit lower bounds)

Exponent of Rectangular Matrix Multiplication

Compute the product of an n x n^k matrix A and an n^k x n matrix B

for any fixed k≥0

Exponent of rectangular matrix multiplication

ω(1,1,k) $= \inf \{ \tau \mid$ this product can be computed using $O(n^\tau)$ arithmetic operations $\}$

Exponent of rectangular matrix multiplication

$$\omega(1,1,k) = \inf \{ \tau \mid \underline{R}(\langle n, n, nk \rangle) = O(n^\tau) \}$$

trivial algorithm: $O(n^{2+k})$ arithmetic operations

→ ω(1,1,k) ≤ 2 + k

square matrices: ω(1,1,1) = ω ≤ 2.38

trivial lower bounds: ω(1,1,k) ≥ 2

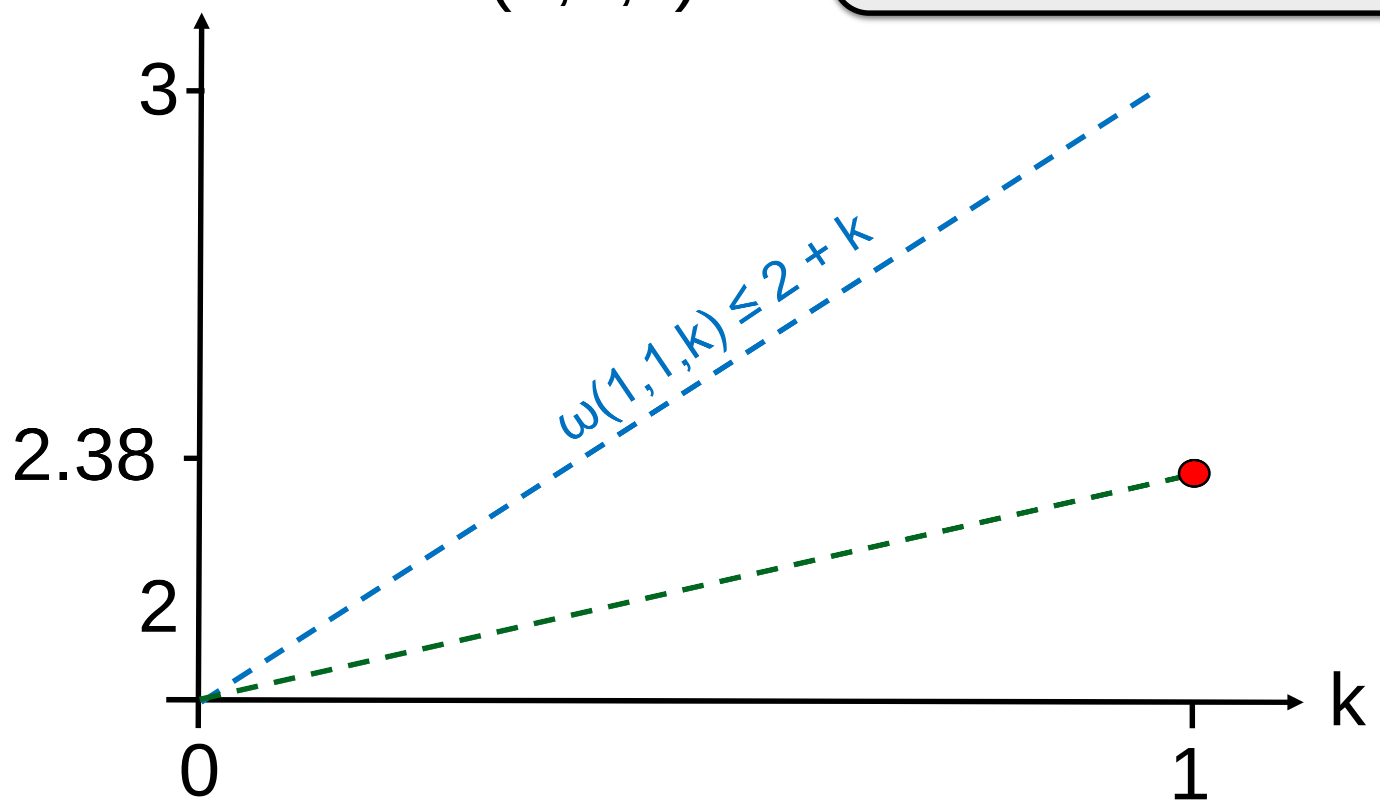
ω(1,1,k) ≥ 1+ k

Exponent of Rectangular Matrix Multiplication

Property [Lotti 83]

$\omega(1,1,k)$ is a convex function

upper bounds on $\omega(1,1,k)$



trivial algorithm: $O(n^{2+k})$ arithmetic operations

$\Rightarrow$ $\omega(1,1,k) \leq 2 + k$

square matrices: $\omega(1,1,1) = \omega \leq 2.38$

trivial lower bounds: $\omega(1,1,k) \geq 2$

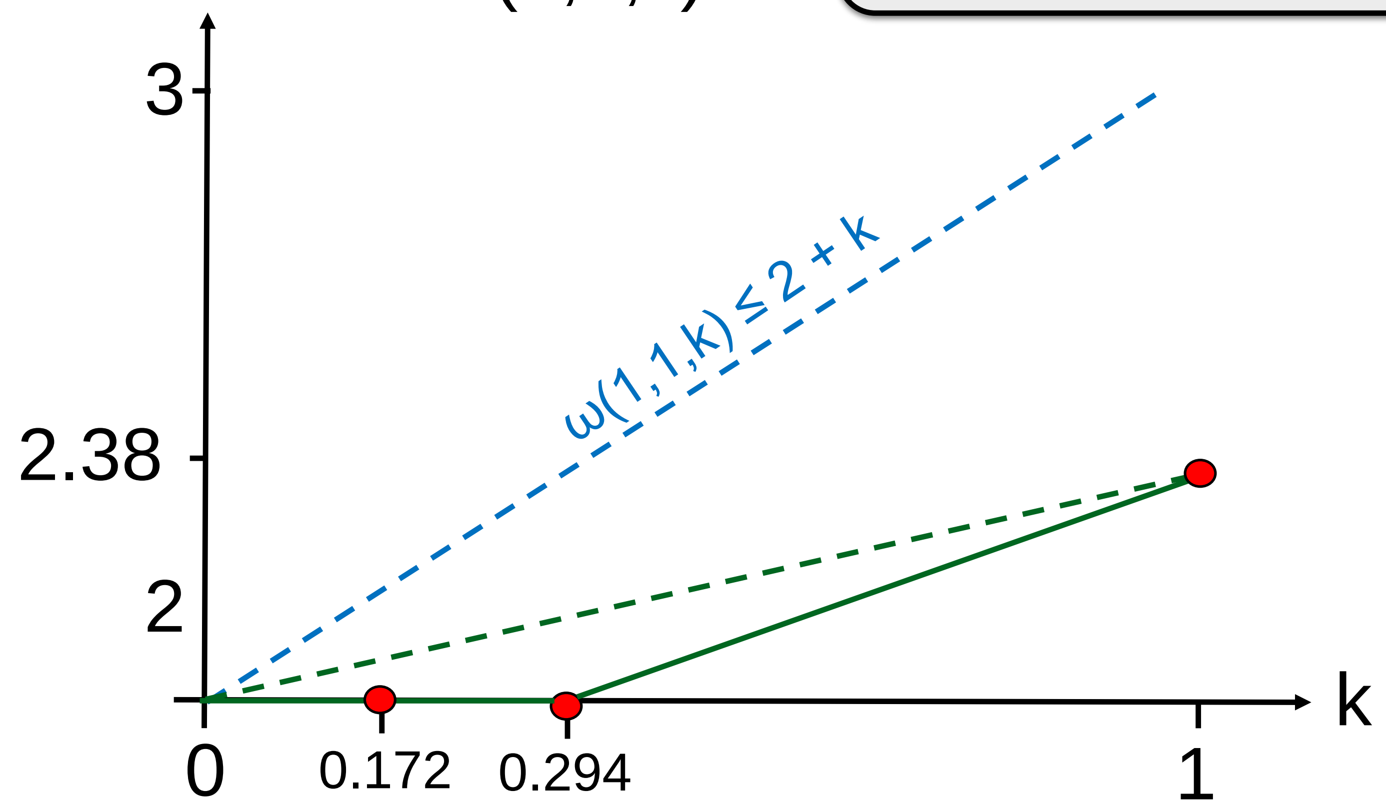
$\omega(1,1,k) \geq 1 + k$

Exponent of Rectangular Matrix Multiplication

Property [Lotti 83]

$\omega(1,1,k)$ is a convex function

upper bounds on $\omega(1,1,k)$



[Coppersmith 1982]: $\omega(1,1,0.172) = 2$

The product of an n x $n^{0.172}$ matrix by an $n^{0.172}$ x n matrix can be computed using $O(n^{2+\varepsilon})$ arithmetic operations for any $\varepsilon>0$

[Coppersmith 1997]: $\omega(1,1,0.294) = 2$

Exponent of Rectangular Matrix Multiplication

[Coppersmith 1982]: ω(1,1,0.172) = 2

The product of an n x $n^{0.172}$ matrix by an $n^{0.172}$ x n matrix can be computed using $O(n^{2+\varepsilon})$ arithmetic operations for any ε>0

Proof outline

Remember tuesday's exercise

Consider the computation of the product of two matrices A and B of the following form:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & 0 \\ b_{31} & 0 \end{pmatrix}.$$

(i) Write the tensor corresponding to this computational task.

(ii) Show that the border rank of this tensor is at most 5.

Taking power N of this tensor and combining it with power N of the same tensor with permuted variables we can show:

$$\underline{R}(\langle M, 4^{N/5}, M \rangle) \leq 5^{2N} \text{ with } M \approx 5^N$$

$$4^{N/5} = M^{\alpha} \text{ with } \alpha = \frac{1}{5}\log_5 4 = 0.1722\ldots$$

Exponent of Rectangular Matrix Multiplication

[Coppersmith 1997]: ω(1,1,0.294) = 2

Idea: Analyze the first power of the CW tensor in an asymmetric way

Tool: Rectangular version of the asymptotic sum inequality

Asymptotic sum inequality (square case)

$$\underline{R}\left(\bigoplus_{i=1}^{\ell}\langle m_i, m_i, m_i\rangle\right) \leq t \Longrightarrow \sum_{i=1}^{\ell} m_i^{\omega} \leq t$$

Asymptotic sum inequality (rectangular case)

$$\underline{R}\left(\bigoplus_{i=1}^{\ell}\langle m_i, m_i, m_i^k\rangle\right) \leq t \Longrightarrow \sum_{i=1}^{\ell} m_i^{\omega(1,1,k)} \leq t$$

As in the square case, it is enough to create a direct sum of matrix products of the desired format

Exponent of Rectangular Matrix Multiplication

[Coppersmith 1997]: $\omega(1,1,0.294) = 2$

Idea: Analyze the first power of the CW tensor in an asymmetric way

$$T_{CW} = T_{CW}^{011} + T_{CW}^{101} + T_{CW}^{110} + T_{CW}^{002} + T_{CW}^{020} + T_{CW}^{200}$$

$$T_{CW}^{011} \cong \langle 1,1,q \rangle \quad T_{CW}^{101} \cong \langle q,1,1 \rangle \quad T_{CW}^{110} \cong \langle 1,q,1 \rangle$$

and

$$T_{CW}^{002} \cong \langle 1,1,1 \rangle \quad T_{CW}^{020} \cong \langle 1,1,1 \rangle \quad T_{CW}^{200} \cong \langle 1,1,1 \rangle$$

square case →

Analysis of the second construction

Theorem 5

For any $0 \le \alpha \le 1/3$ (more variables) and for N large enough, the tensor $T_{CW}^{\otimes N}$ can be converted into a direct sum of

$$2^{\left(H\left(\frac{2}{3}-\alpha, 2\alpha, \frac{1}{3}-\alpha\right)-o(1)\right)N}$$ (recalculate)

terms, each isomorphic to

rectangular case → (α'N, αN, αN; adjust, adjust, adjust)

$$[T_{CW}^{011}]^{\otimes \alpha N} \otimes [T_{CW}^{101}]^{\otimes \alpha N} \otimes [T_{CW}^{110}]^{\otimes \alpha N} \otimes [T_{CW}^{002}]^{\otimes (\frac{1}{3}-\alpha)N} \otimes [T_{CW}^{020}]^{\otimes (\frac{1}{3}-\alpha)N} \otimes [T_{CW}^{200}]^{\otimes (\frac{1}{3}-\alpha)N}.$$

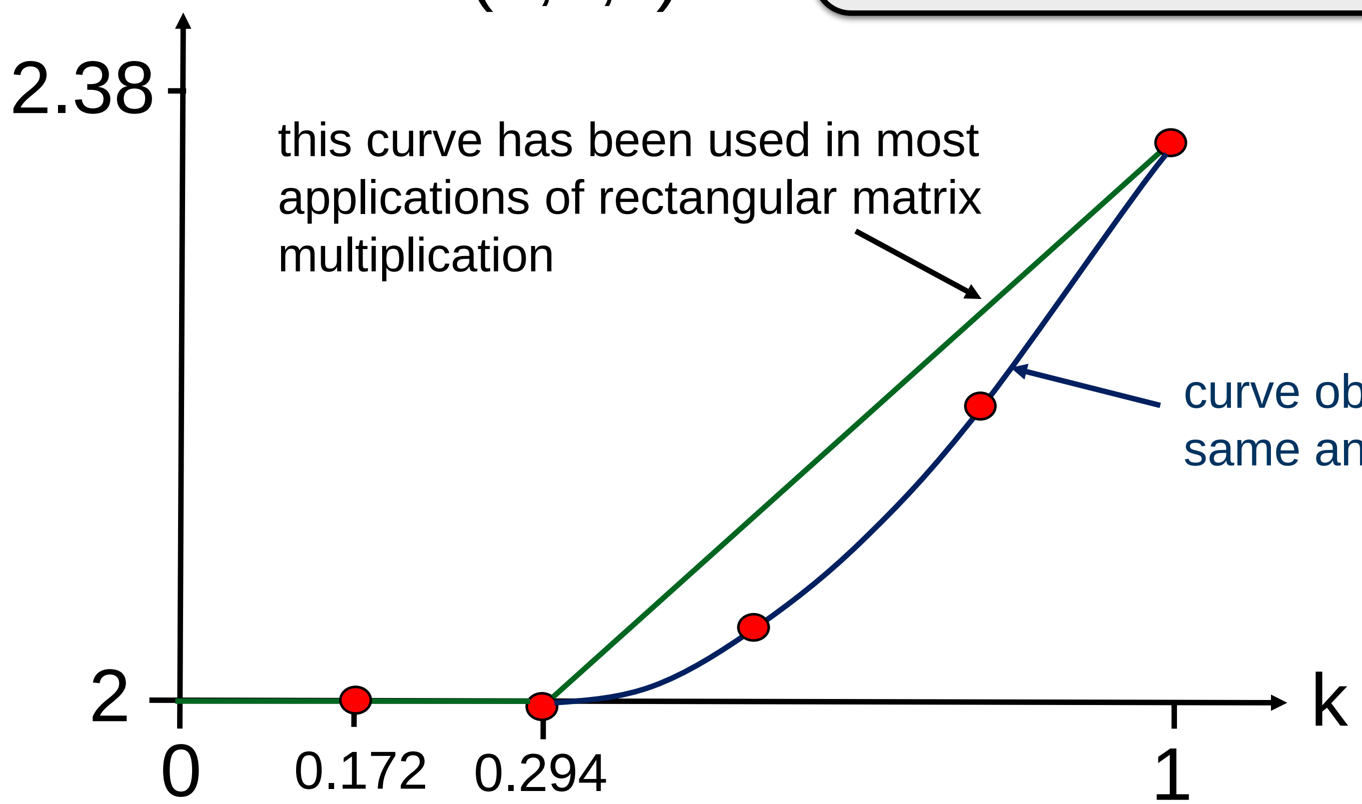
$$\cong \langle q^{\alpha N}, q^{\alpha N}, q^{\alpha N} \rangle \quad \langle q^{\alpha N}, q^{\alpha N}, q^{\alpha' N} \rangle$$

Exponent of Rectangular Matrix Multiplication

Property [Lotti 83]

ω(1,1,k) is a convex function

upper bounds on ω(1,1,k)



[Coppersmith 1982]: ω(1,1,0.172) = 2

[Coppersmith 1997]: ω(1,1,0.294) = 2

[Ke, Zeng, Han, Pan 2008]: ω(1,1,0.5356) < 2.0712

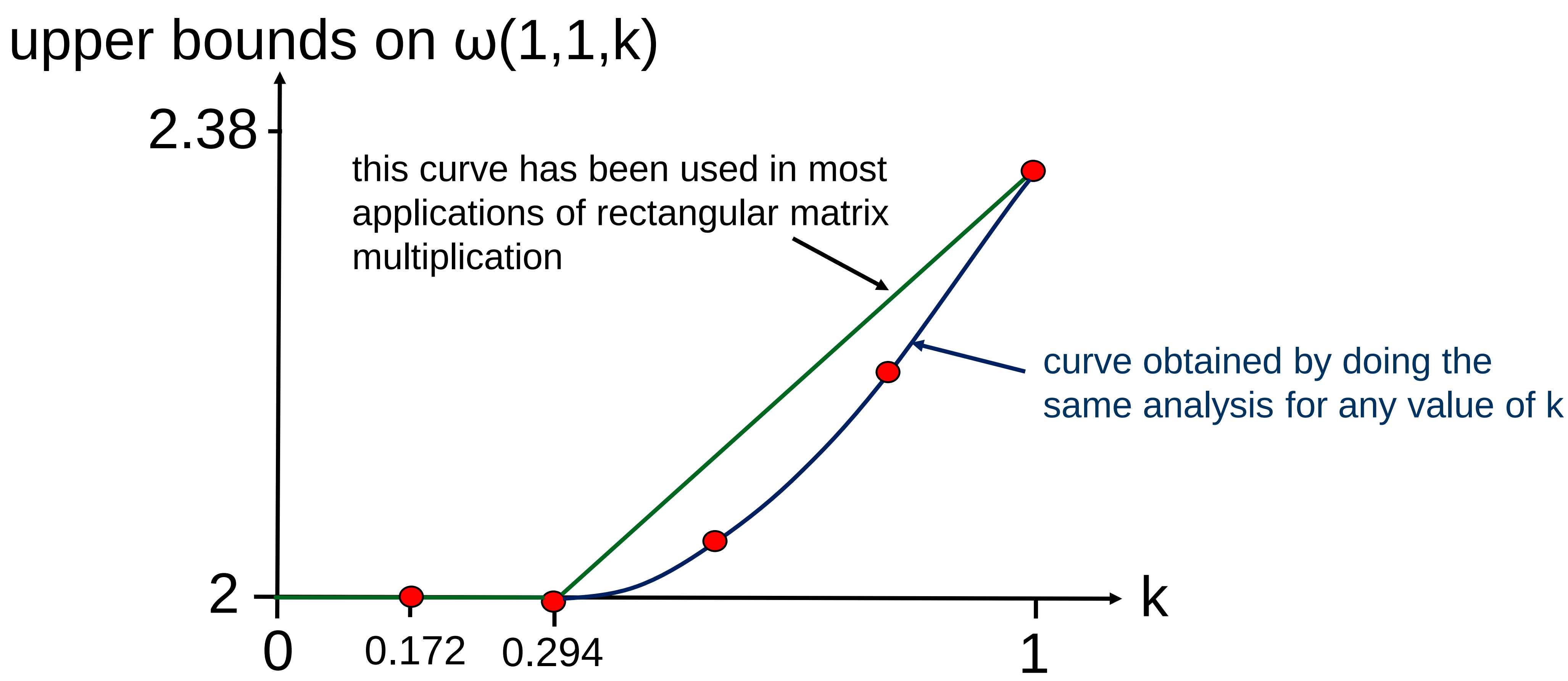
ω(1,1,0.8) < 2.2356

ω(1,1,2) < 3.2699

(slightly improving a bound from [Huang, pan 1998])

obtained by a similar asymmetric analysis of the first power of the CW tensor

Exponent of Rectangular Matrix Multiplication



[Coppersmith 1997]: $\omega(1,1,0.294) = 2$ ⟹ $\alpha > 0.294$

(obtained from the analysis of the first power of the CW tensor)

Dual exponent of matrix multiplication

$$\alpha = \sup\{\, k \mid \omega(1,1,k) = 2 \,\}$$

proving that $\alpha=1$ is equivalent to proving that $\omega=2$

[LG 2012]: $\omega(1,1,0.302) = 2$ ⟹ $\alpha > 0.302$

from the analysis of the second power of the CW tensor

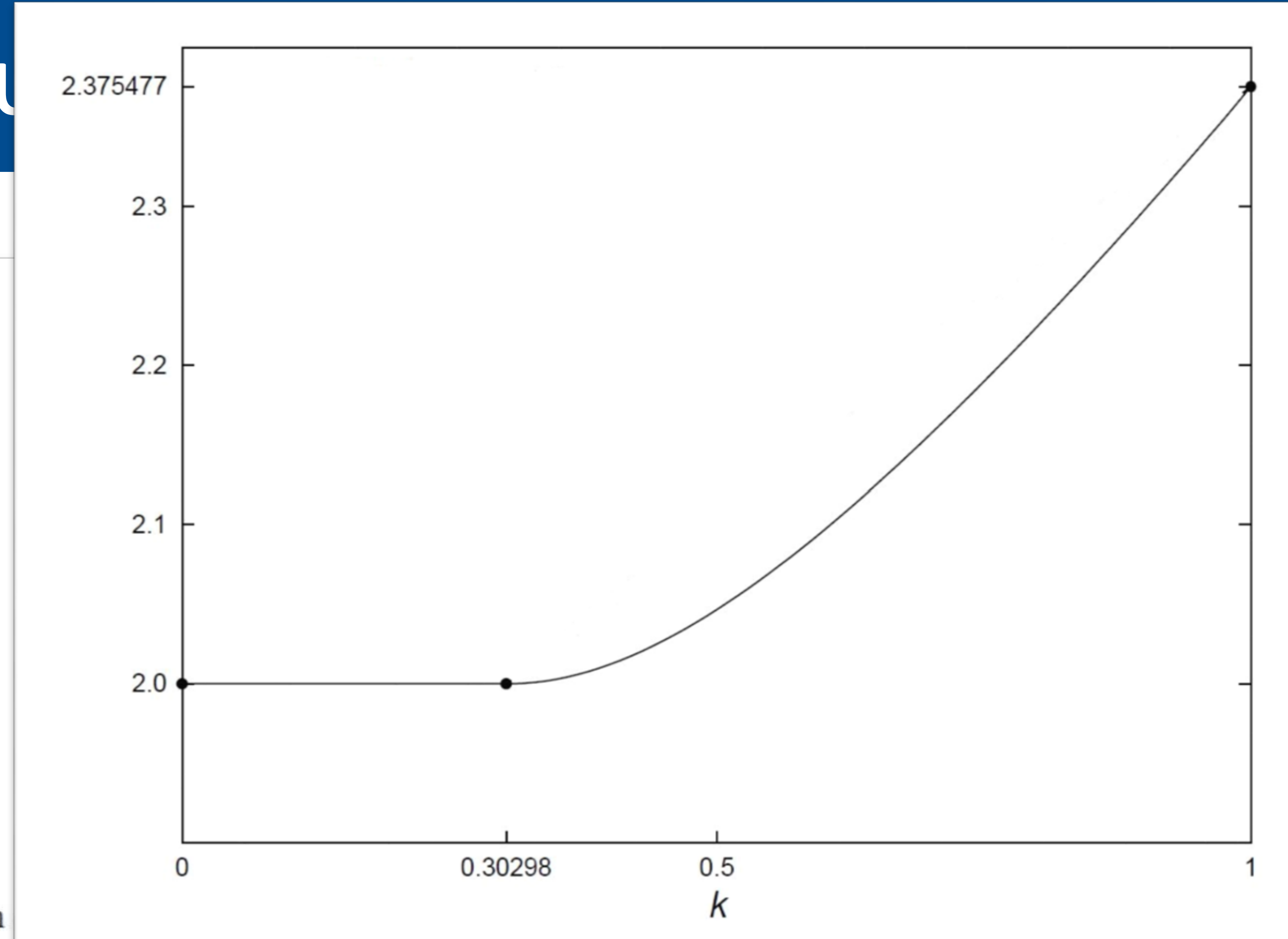
Exponent of Rectangu

from [LG 2012]

k	upper bound on $\omega(1,1,k)$
0.30298	2
0.31	2.000063
0.32	2.000371
0.33	2.000939
0.34	2.001771
0.35	2.002870
0.40	2.012175
0.45	2.027102
0.50	2.046681
0.5302	2.060396
0.55	2.070063

k	upper bound on $\omega(1,1,k)$
0.60	2.096571
0.65	2.125676
0.70	2.156959
0.75	2.190087
0.80	2.224790
0.85	2.260830
0.90	2.298048
0.95	2.336306
1.00	2.375477
1.10	2.456151
1.20	2.539392

Table 1: Our upper bounds on the exponent of the multiplication
matrix.



exactly the same bound as the one obtained by Coppersmith and Winograd for square matrix multiplication

curve of the same shape, but slightly below the previous curve

better than all previous bounds for $k \neq 1$

[LG 2012]: $\omega(1,1,0.302) = 2$ ➡ $\alpha > 0.302$

from the analysis of the second power of the CW tensor

Exponent of Rectangular Matrix Multiplication

Dual exponent of matrix multiplication

$\alpha = \sup \{ k \mid \omega(1,1,k) = 2 \}$

first power of the CW tensor:

ω < 2.3872 [CW 1987]

α > 0.294 [Coppersmith 1997]

second power of the CW tensor:

ω < 2.3755 [CW 1987]

α > 0.302 [LG 2012]

the gap increases

fourth power of the CW tensor:

ω < 2.3730 [Stothers 2010]

α > 0.313 [LG, Urrutia 2017]

eight power of the CW tensor:

ω < 2.3729 [Williams 2012]

???

What will happen with higher powers?

Overview of the Lectures

✓ Fundamental techniques for fast matrix multiplication (1969~1987)

- Basics of bilinear complexity theory: exponent of matrix multiplication, Strassen's algorithm, bilinear algorithms
- First technique: tensor rank and recursion
- Second technique: border rank
- Third technique: the asymptotic sum inequality
- Fourth technique: the laser method

Lecture 1

✓ Recent progress on matrix multiplication (1987~)

Lecture 2

- Laser method on powers of tensors
- Other approaches
- Lower bounds
- Rectangular matrix multiplication

Lecture 3

✓ Applications of matrix multiplications, open problems