

DATA FUSION

TENSOR FACTORIZATIONS BY COMPLEX OPTIMIZATION

Laurent Sorber

Supervisors:

prof. Lieven De Lathauwer

prof. Marc Van Barel

iwT
KU LEUVEN



**SELF-DRIVING CARS
ANALYSE 1GB/S**

**EQUIVALENT TO WATCHING
400 BLU-RAYS SIMULTANOUSLY**



**A BOEING 787 GENERATES
500GB/FLIGHT**

**1000 TREES MADE INTO 1M BOOKS'
WORTH OF DATA PER FLIGHT**



**THE LARGE HADRON COLLIDER
PRODUCES 30PB/YEAR**

**1PB IS THE ESTIMATED STORAGE
CAPACITY OF THE HUMAN BRAIN**

A large, jagged iceberg floats in the middle of a dark blue ocean. The sky above is a vibrant blue with wispy white clouds. The iceberg's surface is textured and shows some shadows, suggesting its massive size. The text is overlaid in the center of the image.

**THIS DOES NOT EVEN INCLUDE
THE MUCH LARGER DEEP WEB**

BIG DATA

1. Volume
2. Velocity
3. Variety

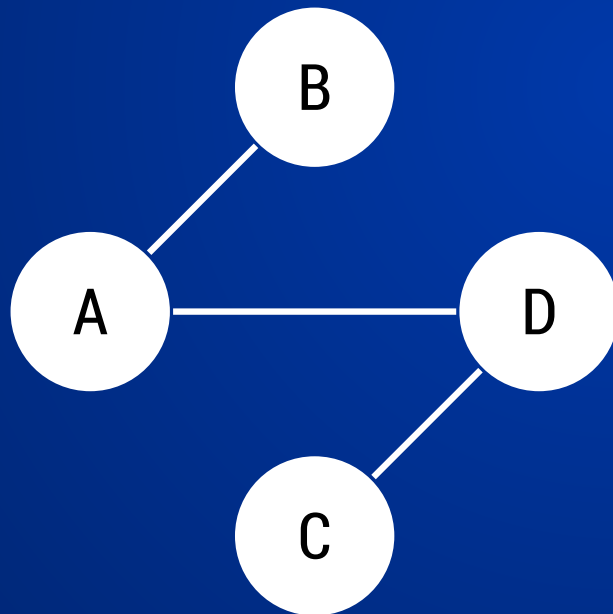
HOW IS DATA STORED?



GRAPHS

Web graph, Social graph, Electrical circuit,

...



	A	B	C	D
A	0	1	0	1
B	1	0	0	0
C	0	0	0	1
D	1	0	1	0

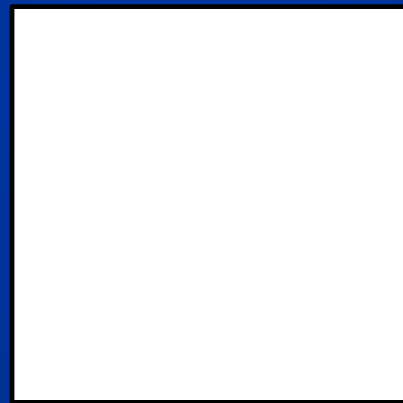
RELATIONAL DATA

user-rates-movie, student-performs-task,

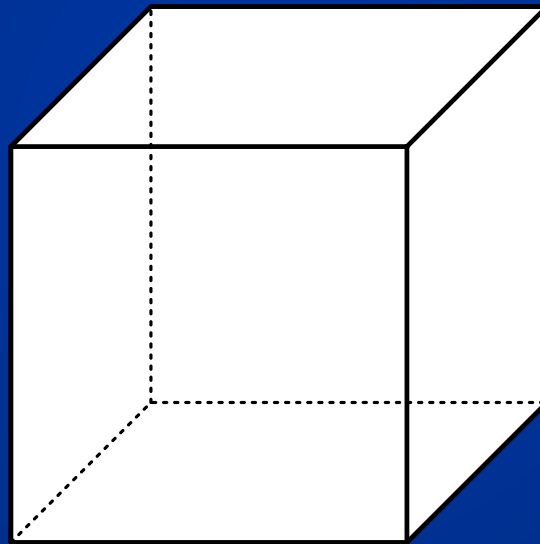
...

	Dries	Bart	Vicky	Clara
Pretty Woman	3.5	1.5	?	4.5
Titanic	?	2.25	3	4
Star Wars	4.25	?	2	3
The Terminator	4.5	4.5	1.5	?

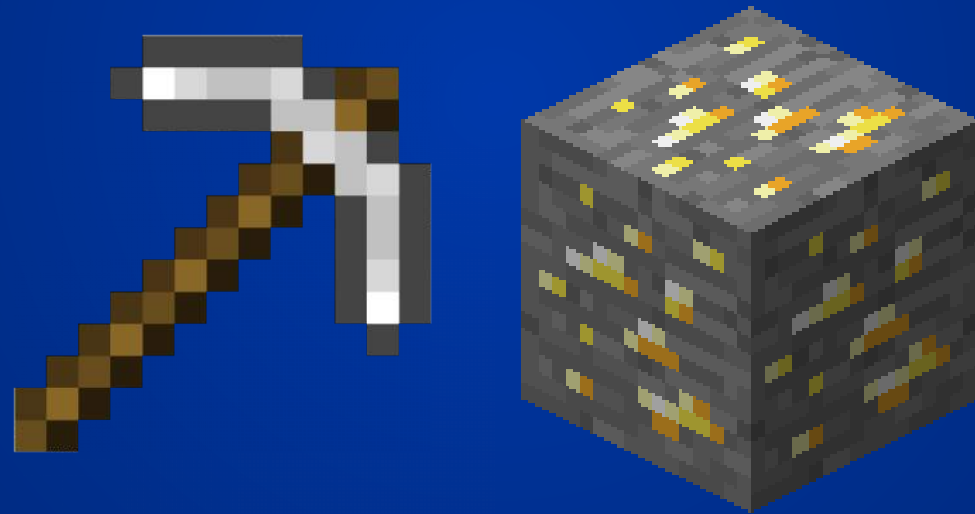
MATRIX



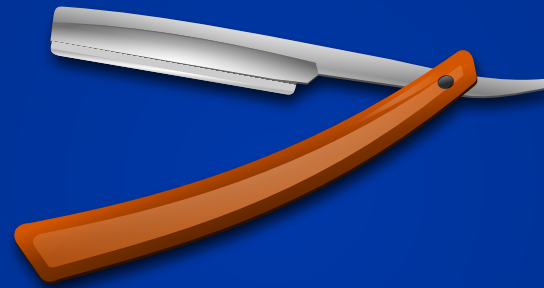
TENSOR



HOW CAN WE MINE THIS NEW NATURAL RESOURCE?



USE OCCAM'S RAZOR!



"Shave the model until you cut into its explanatory power."

PREDICTING MOVIE RATINGS

	Dries	Bart	Vicky	Clara
Pretty Woman	3.5	1.5	?	4.5
Titanic	?	2.25	3	4
Star Wars	4.25	?	2	3
The Terminator	4.5	4.5	1.5	?

	Action content (<i>a</i>)		Action preference (<i>b</i>)
Pretty Woman	0	Dries	1
Titanic	0.5	Bart	1
Star Wars	1.5	Vicky	-0.5
The Terminator	2	Clara	0

$$r_{m,u} = \mu + a_m \cdot b_u$$

$$4.5 = 2.5 + 2 \cdot 1$$

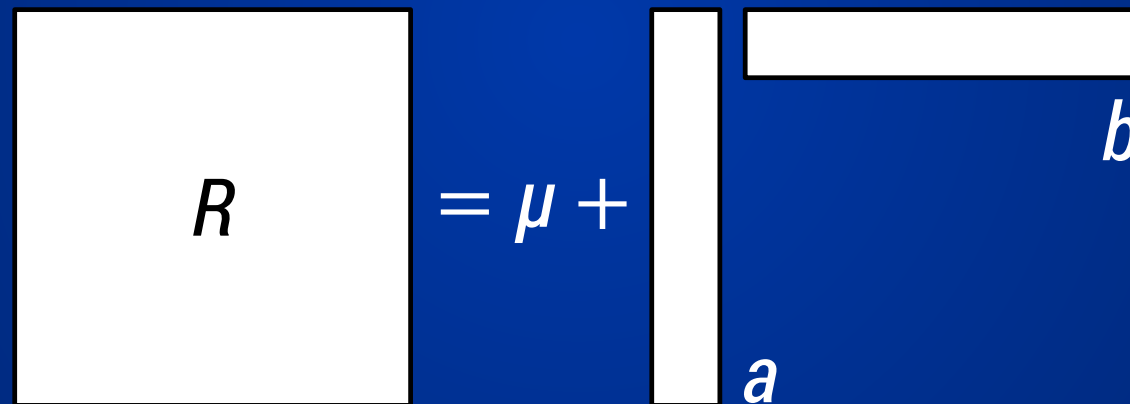
Dries likes The Terminator

$$1.5 = 2.5 + 2 \cdot (-0.5)$$

Vicky does not like The Terminator

WHAT DOES THE MODEL LOOK LIKE?

$$r_{m,u} = \mu + a_m \cdot b_u$$



IS THE MODEL ANY GOOD?

	Dries	Bart	Vicky	Clara
Pretty Woman	3.5−2.5	1.5−2.5	?−2.5	4.5−2.5
Titanic	?−3	2.25−3	3−2.25	4−2.5
Star Wars	4.25−4	?−4	2−1.75	3−2.5
The Terminator	4.5−4.5	4.5−4.5	1.5−1.5	?−2.5

$$f(a, b) := \sum_{m,u} [r_{m,u} - (\mu + a_m \cdot b_u)]^2$$

$$[3.5 - 2.5]^2 + [4.25 - 4]^2 + [4.5 - 4.5]^2 + [1.5 - 2.5]^2 + \dots$$

MINIMIZE f TO FIND THE BEST SOLUTION

$$\underset{a, b}{\text{minimize}} \quad f(a, b) := \left\| R - \left(\mu + \begin{matrix} \text{---} \\ | \\ \text{---} \\ a \end{matrix} \begin{matrix} \text{---} \\ | \\ \text{---} \\ b \end{matrix} \right) \right\|^2$$

However, even the best solution for a and b can not predict all the ratings accurately because the model is too simple!

	Action content (<i>a</i>)	Romance content (<i>c</i>)		Action preference (<i>b</i>)	Romance preference (<i>d</i>)
Pretty Woman	0	2	Dries	1	0.5
Titanic	0.5	1.5	Bart	1	-0.5
Star Wars	1.5	0.5	Vicky	-0.5	0.5
The Terminator	2	0	Clara	0	1

$$r_{m,u} = \mu + a_m \cdot b_u + c_m \cdot d_u$$

$$2.25 = 2.5 + 0.5 \cdot 1 + 1.5 \cdot (-0.5)$$

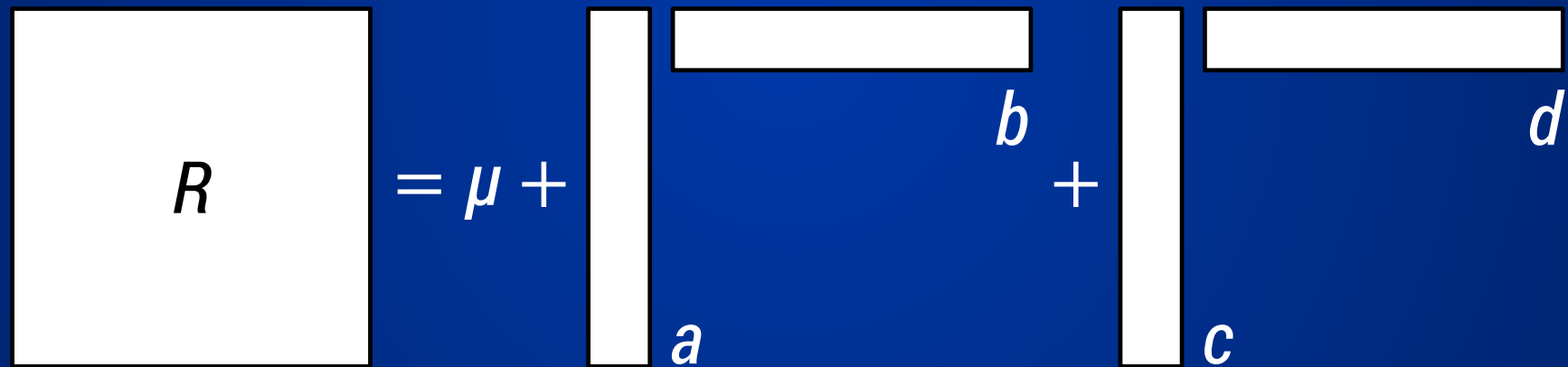
Bart does not like Titanic very much

$$4 = 2.5 + 0.5 \cdot 0 + 1.5 \cdot 1$$

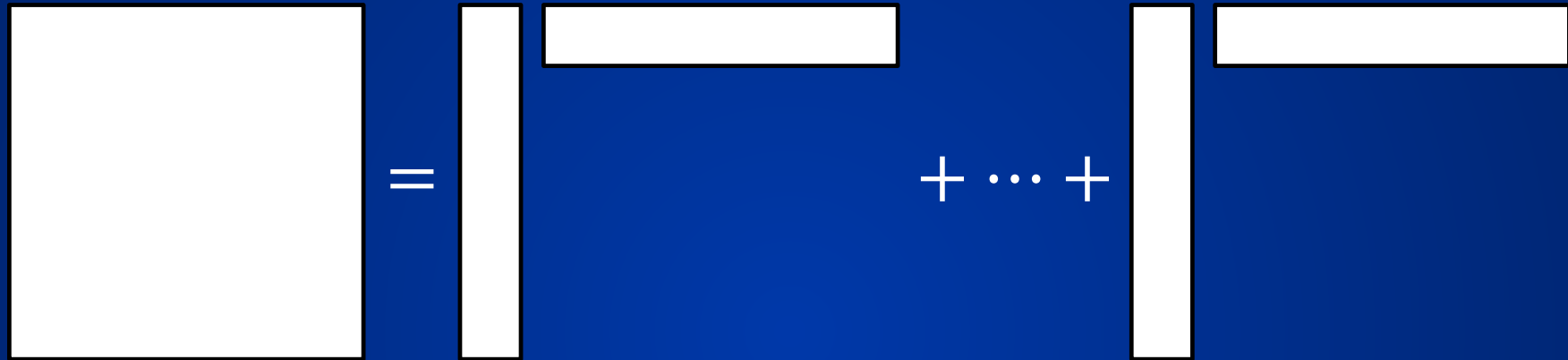
Clara likes Titanic

WHAT DOES THE NEW MODEL LOOK LIKE?

$$r_{m,u} = \mu + a_m \cdot b_u + c_m \cdot d_u$$

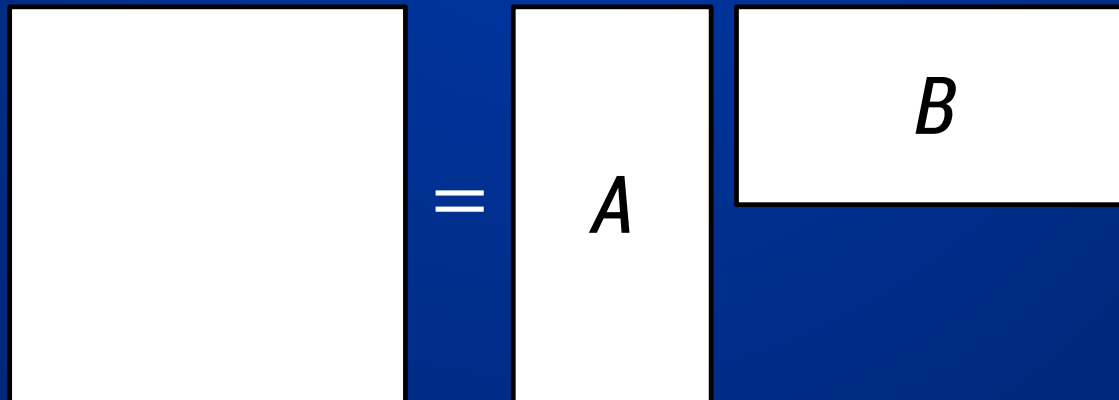


MATRIX FACTORIZATION



A diagram illustrating matrix factorization as a sum of rank-1 matrices. On the left is a large white square representing a matrix. To its right is an equals sign. This is followed by a vertical white bar (representing a column vector) and a horizontal white bar (representing a row vector). To the right of this pair is a plus sign, followed by an ellipsis, followed by another plus sign. This is followed by another vertical white bar and horizontal white bar pair, representing a second rank-1 matrix.

Or equivalently,

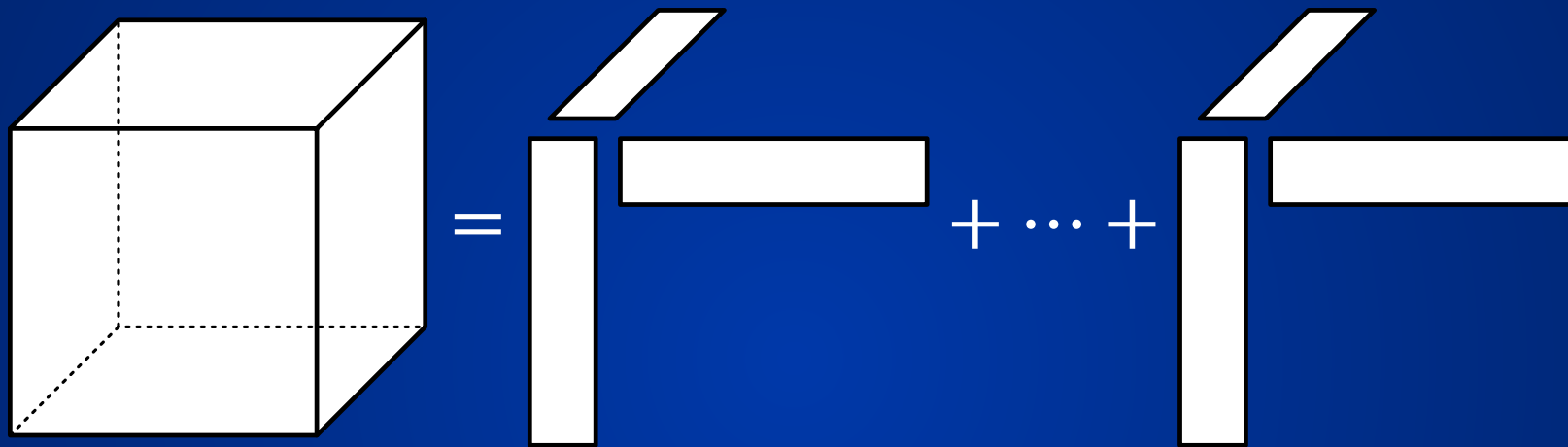


A diagram illustrating matrix factorization as a product of two matrices. On the left is a large white square representing a matrix. To its right is an equals sign. This is followed by a vertical white bar labeled with the letter *A* (representing a column vector) and a horizontal white bar labeled with the letter *B* (representing a row vector).

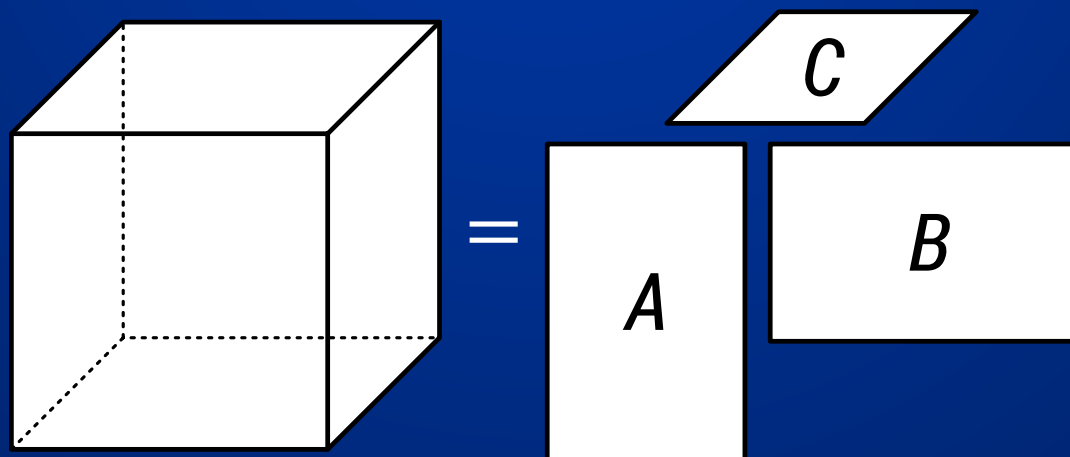
**THEN WHAT ARE
TENSOR FACTORIZATIONS?**

TENSOR FACTORIZATION

Canonical polyadic decomposition (CPD)

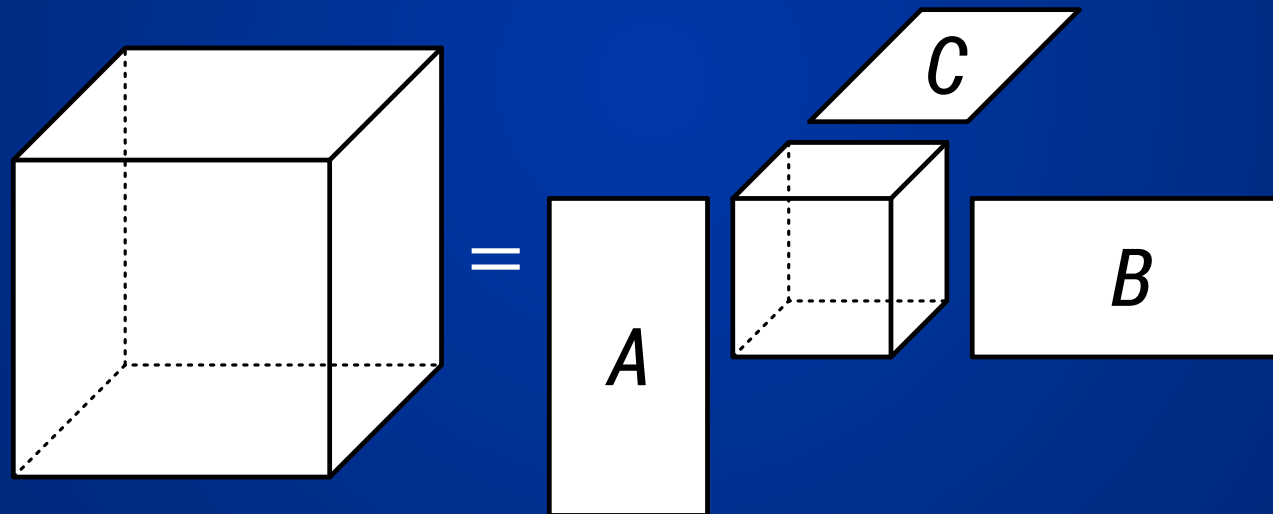


Or equivalently,

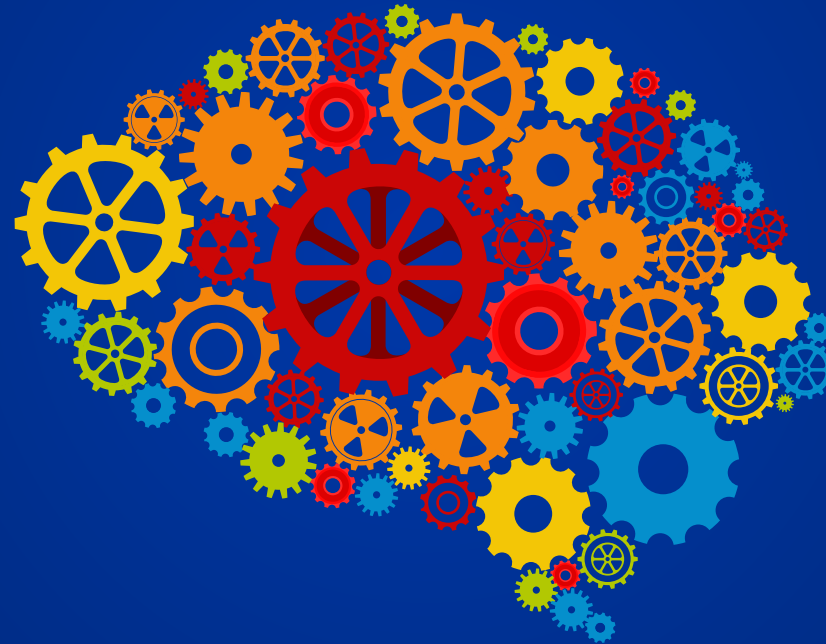


TENSOR FACTORIZATION

Low multilinear rank approximation (LMLRA)

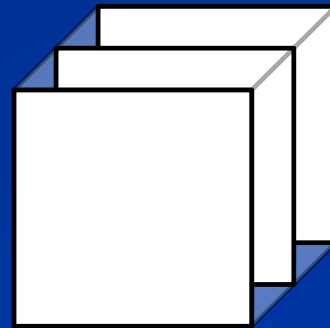


SO WHAT IS DATA FUSION?



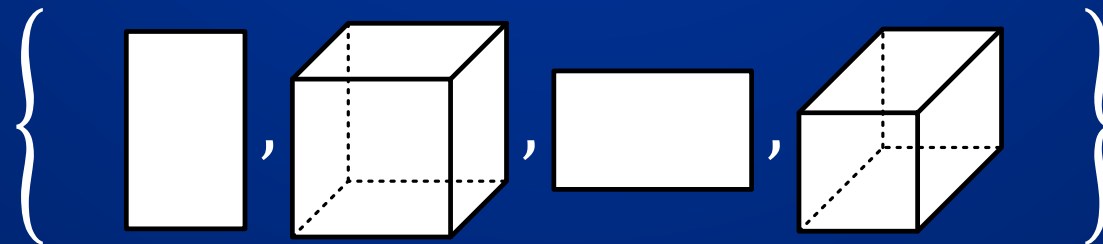
TENSOR FACTORIZATION

Fusion by jointly factorizing a set of matrices that are homogeneous in size



DATA FUSION

Fusion by jointly factorizing a set of tensors that are heterogeneous in size

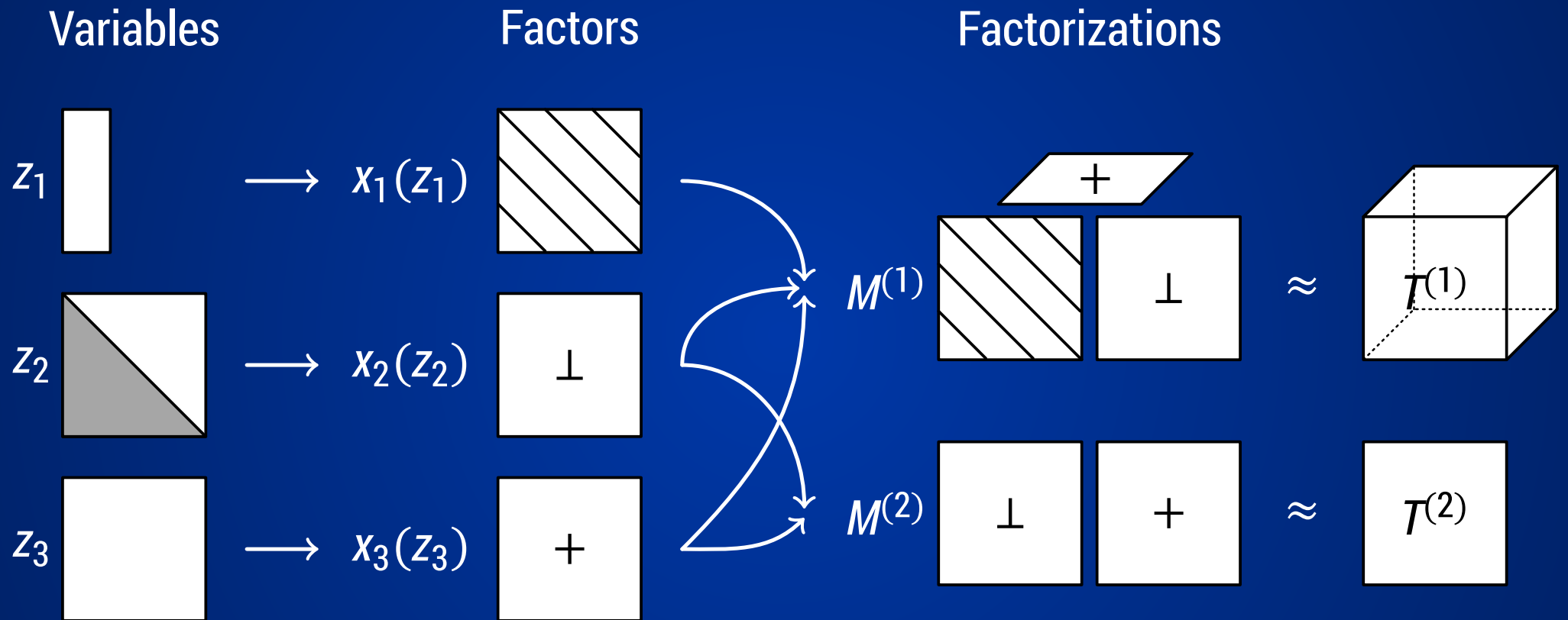


GOAL

For each data set (stored as a tensor):

1. Pick a type of tensor decomposition
CPD, LMLRA, BTD, TT, ...
2. Impose structure on their factors
Nonnegative, orthogonal, inverse, ...
3. Choose which factors are shared
1st and 2nd tensor share factor matrix A , ...

STRUCTURED DATA FUSION

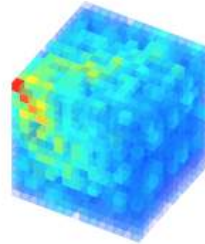


$$\underset{z}{\text{minimize}} \quad \sum_d \omega_d \left\| M^{(d)}(X(z)) - \mathcal{T}^{(d)} \right\|^2$$

WWW.TENSORLAB.NET

Tensorlab

A MATLAB toolbox for tensor computations



About

Tensorlab is a MATLAB toolbox that offers algorithms for

- **structured data fusion:** define your own (coupled) matrix and tensor factorizations with structured factors and support for dense, sparse and incomplete data sets,
- **tensor decompositions:** canonical polyadic decomposition (CPD), multilinear singular value decomposition (MLSVD), block term decompositions (BTD) and low multilinear rank approximation (LMLRA),
- **complex optimization:** quasi-Newton and nonlinear-least squares optimization with complex variables including numerical complex differentiation,
- **global minimization of bivariate polynomials and rational functions:** both real and complex exact line search (LS) and real exact plane search (PS) for tensor optimization,
- **and much more:** cumulants, tensor visualization, estimating a tensor's rank or multilinear rank, ...

Download the [Tensorlab user guide](#) (preview on the right) to get started with Tensorlab. Alternatively, see Tensorlab's Contents.m for an overview of the toolbox's functionality. For questions, bug reports or other inquiries, please contact tensorlab@esat.kuleuven.be.

Download 2014-05-07

To download Tensorlab, please fill out the form below. Your email address will not be used for marketing purposes, sold or shared with third parties.

First name

userguide.pdf

Tensorlab
User Guide 2014-05-07

Laurent Sorber^{†††} Marc Van Barel^{*} Lieven De Lathauwer^{††§}

Contents

1	Getting started	1
2	Data sets: dense, incomplete and sparse tensors	5
2.1	Representation	5
2.2	Tensor operations	6
3	Canonical polyadic decomposition	10
3.1	Problem and tensor generation	10
3.2	Computing the CPD	11
3.3	Choosing the number of rank-one terms R	12
4	Low multilinear rank approximation	13
4.1	Problem and tensor generation	14
4.2	Computing a LMLRA	14
4.3	Choosing the size of the core tensor	16
5	Block term decomposition	17
5.1	Problem and tensor generation	17
5.2	Computing a BTD	18
6	Structured data fusion	19
6.1	Domain specific language for SDF	20
6.2	Implementing a new factor structure	28
7	Complex optimization	29
7.1	Complex derivatives	30
7.2	Nonlinear least squares	36
7.3	Unconstrained nonlinear optimization	40
8	Global minimization of bivariate functions	42
8.1	Stationary points of polynomials and rational functions	43
8.2	Isolated solutions of a system of two bivariate polynomials	45

^{*}NALAG, Department of Computer Science, KU Leuven, Celestijnenlaan 200A, BE-3001 Leuven, Belgium (Laurent.Sorber@cs.kuleuven.be, Marc.VanBarel@cs.kuleuven.be)
[†]Group Science, Engineering and Technology, KU Leuven, Kulak, E. Sabbelaan 53, BE-3000 Kortrijk, Belgium (Lieven.DeLathauwer@kuleuven-kulak.be)
[§]STADIUS, Department of Electrical Engineering (ESAT), KU Leuven, Kasteelpark Arenberg 10, BE-3001 Leuven, Belgium (Lieven.DeLathauwer@esat.kuleuven.be)
^{||}Minds Future Health Department, Kasteelpark Arenberg 10, BE-3001 Leuven, Belgium.

Page: 1 of 49 Automatic

ZEISS



BOEING®

D:wave

The Quantum Computing Company™

SAMSUNG



HITACHI
Inspire the Next

EIGENVALUE DECOMPOSITION

is a CPD with structured factors

$$A = \begin{bmatrix} 0 & 0.5 & & & & \\ 1 & 0 & 0.5 & & & \\ & 0.5 & 0 & 0.5 & & \\ & & 0.5 & 0 & 0.5 & \\ & & & \ddots & \ddots & \ddots \end{bmatrix} = \begin{array}{c} \lambda \\ \hline \begin{array}{|c|c|} \hline V & V^{-1} \\ \hline \end{array} \end{array}$$

$$\lambda_n = \cos\left(\frac{\pi(2n-1)}{2N}\right) \text{ for } n = 1, \dots, N$$

MATLAB

```
[V,L] = eig(A);
```

TENSORLAB

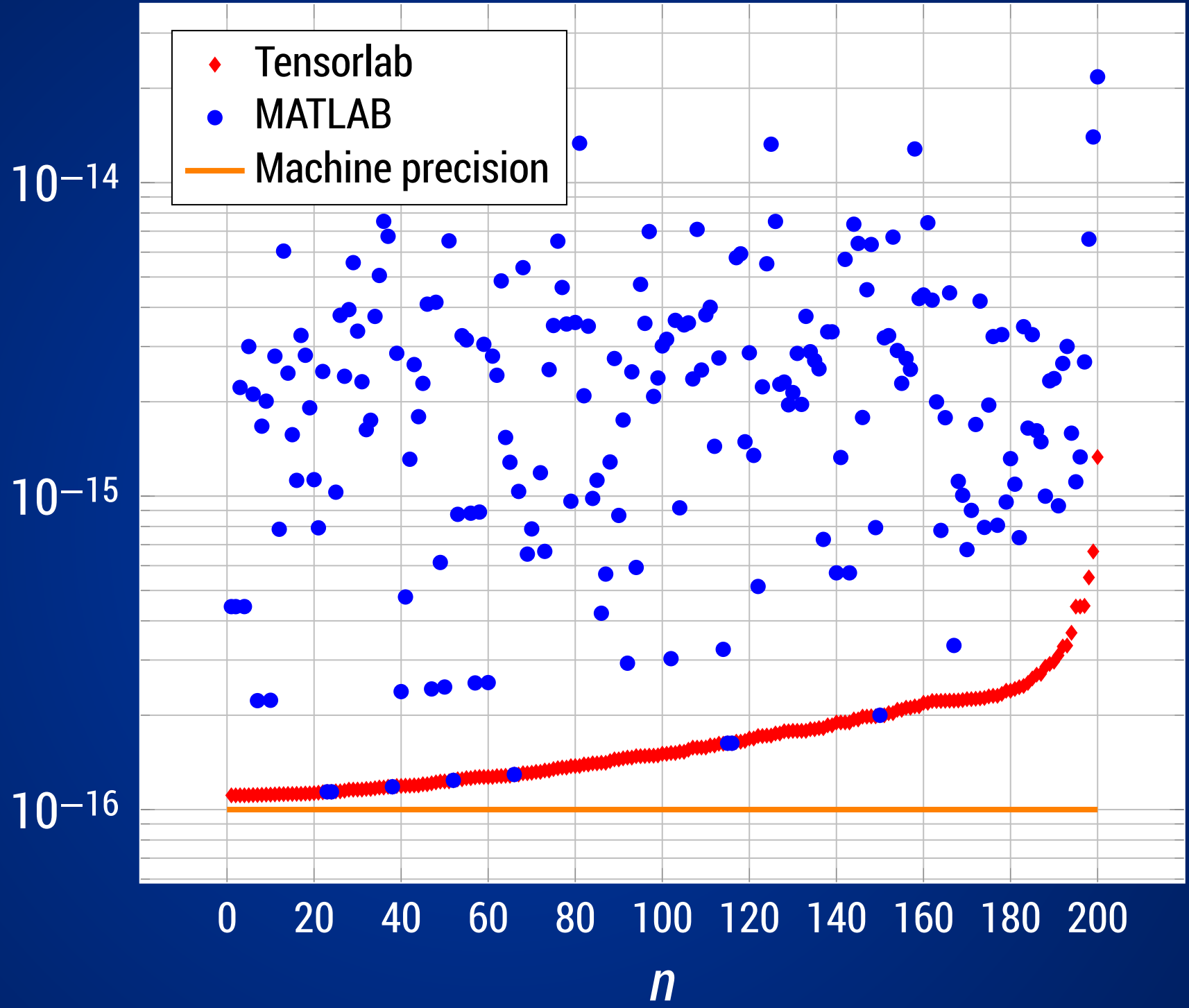
```
model.variables.v = randn(N,N); % N-by-N matrix  
model.variables.l = randn(1,N); % 1-by-N vector
```

```
model.factors.V      = 'v';  
model.factors.Vinv   = {'v',@struct_invtransp}; % One of 32 s  
model.factors.L      = 'l';
```

```
model.factorizations.evd.data = A;  
model.factorizations.evd.cpd  = {'V', 'Vinv', 'L'};
```

```
sol = sdf_nls(model); % sol.factors, sol.variables
```


$$\left| \frac{\lambda_n - \hat{\lambda}_n}{\lambda_n} \right|$$



ALGORITHMS

$$\underset{z}{\text{minimize}} \quad \sum_d \omega_d \left\| M^{(d)}(X(z)) - T^{(d)} \right\|^2$$

- User's choice of underlying solver
 - Quasi-Newton, nonlinear least squares, ...
- Solver exploits the structure in the factors
 - Nonnegative, orthogonal, inverse, ...
- Solver exploits the structure of the decomposition
 - CPD, LMLRA, BTD, TT, ...
- Based on complex optimization
 - Solve complex-valued problems with the same code

LINE AND PLANE SEARCH

$$\underset{z}{\text{minimize}} \quad f(z)$$

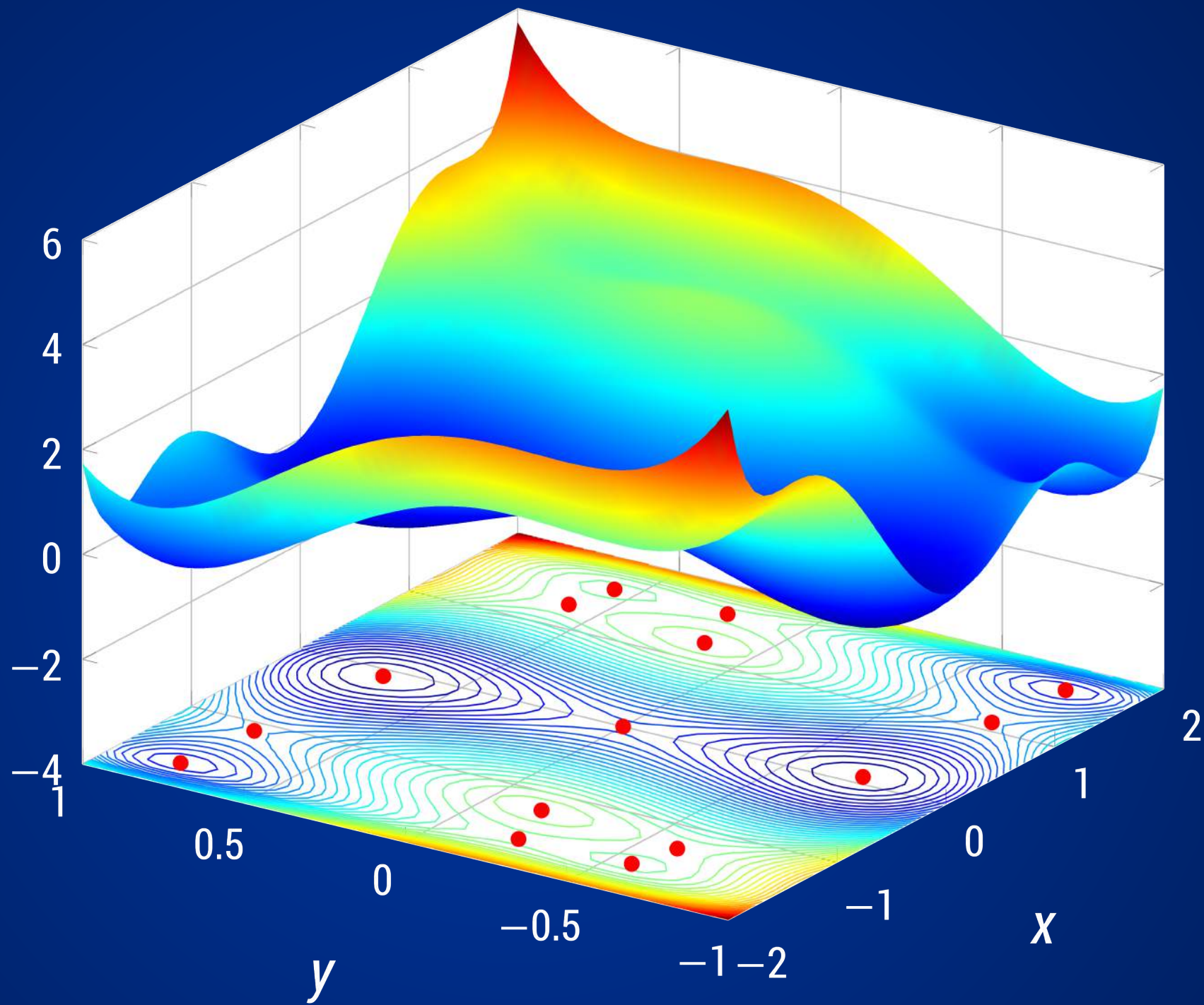
Line search subproblem

$$\underset{x \in \mathbb{R}}{\text{minimize}} \quad f(z_k + x \cdot p)$$

Plane search subproblem

$$\underset{x, y \in \mathbb{R}}{\text{minimize}} \quad f(z_k + x \cdot p + y \cdot q)$$

$f(x, y)$

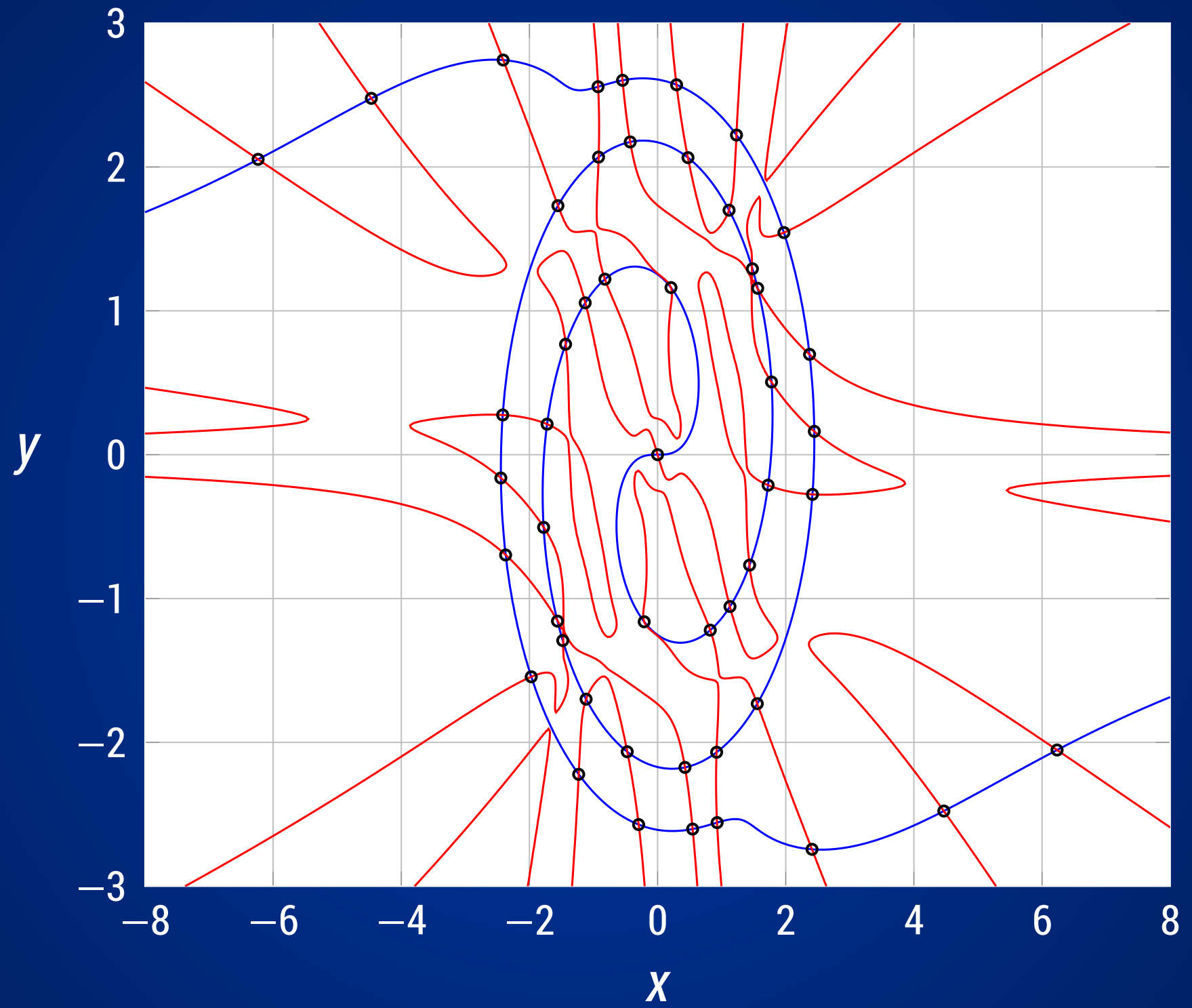


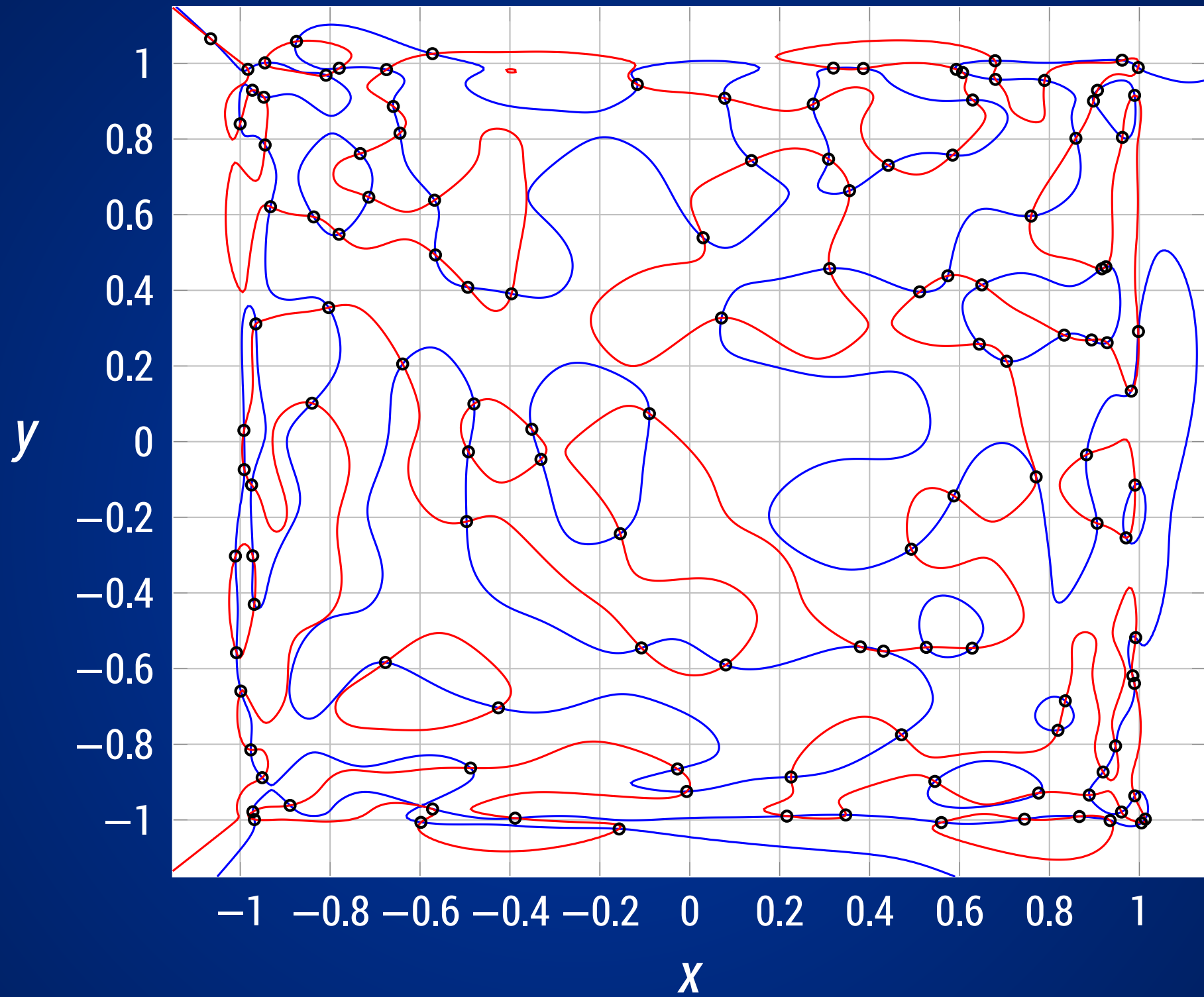
The global minimizer of $f(x,y)$ is a solution of a bivariate polynomial system, e.g.,

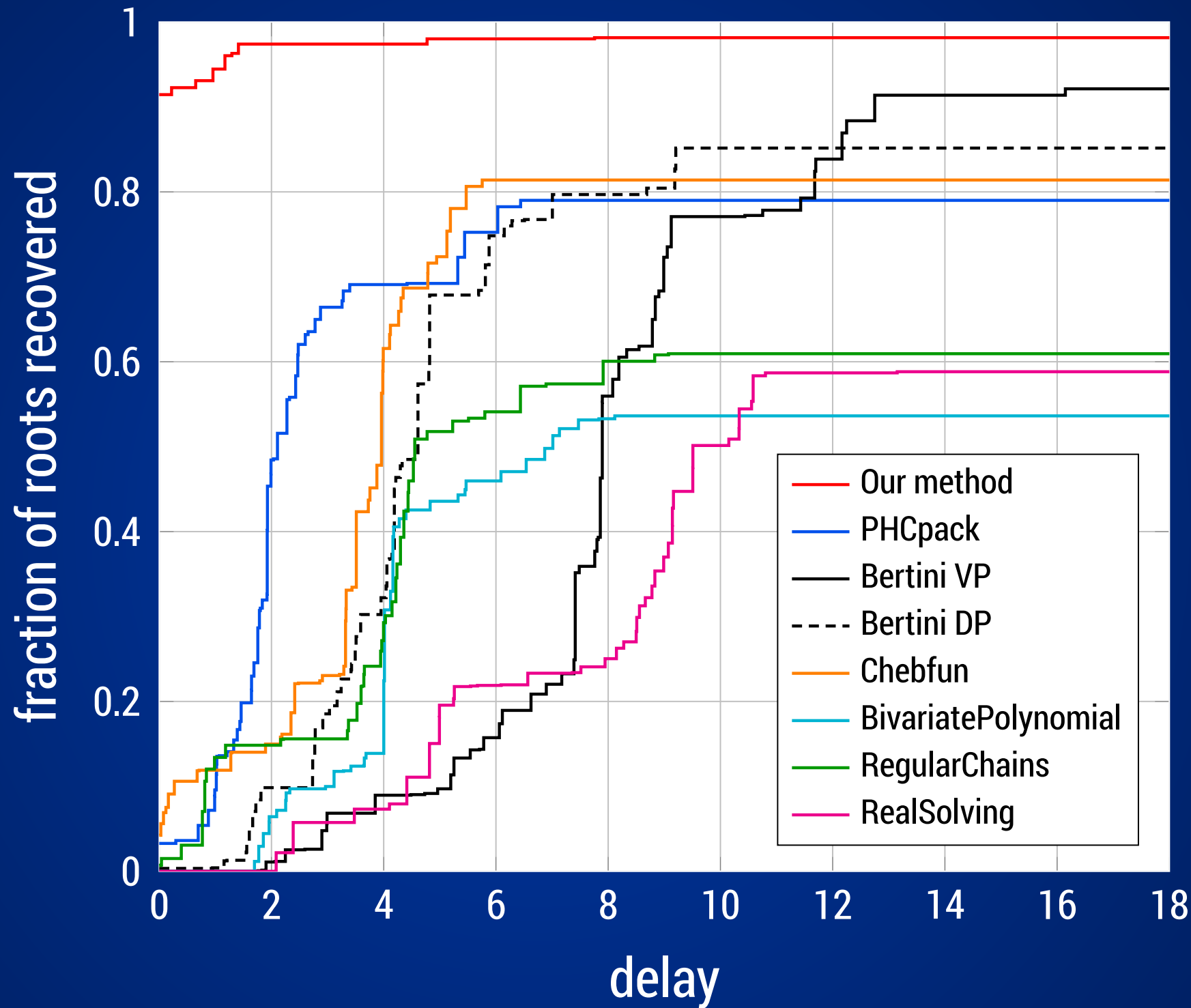
$$\begin{cases} 2x^2 - xy + y^4 - 8 = 0 \\ 5y^3 + 7xy^2 + 3 = 0 \end{cases}$$

Developed an algorithm that computes all solutions simultaneously by solving an eigenvalue decomposition

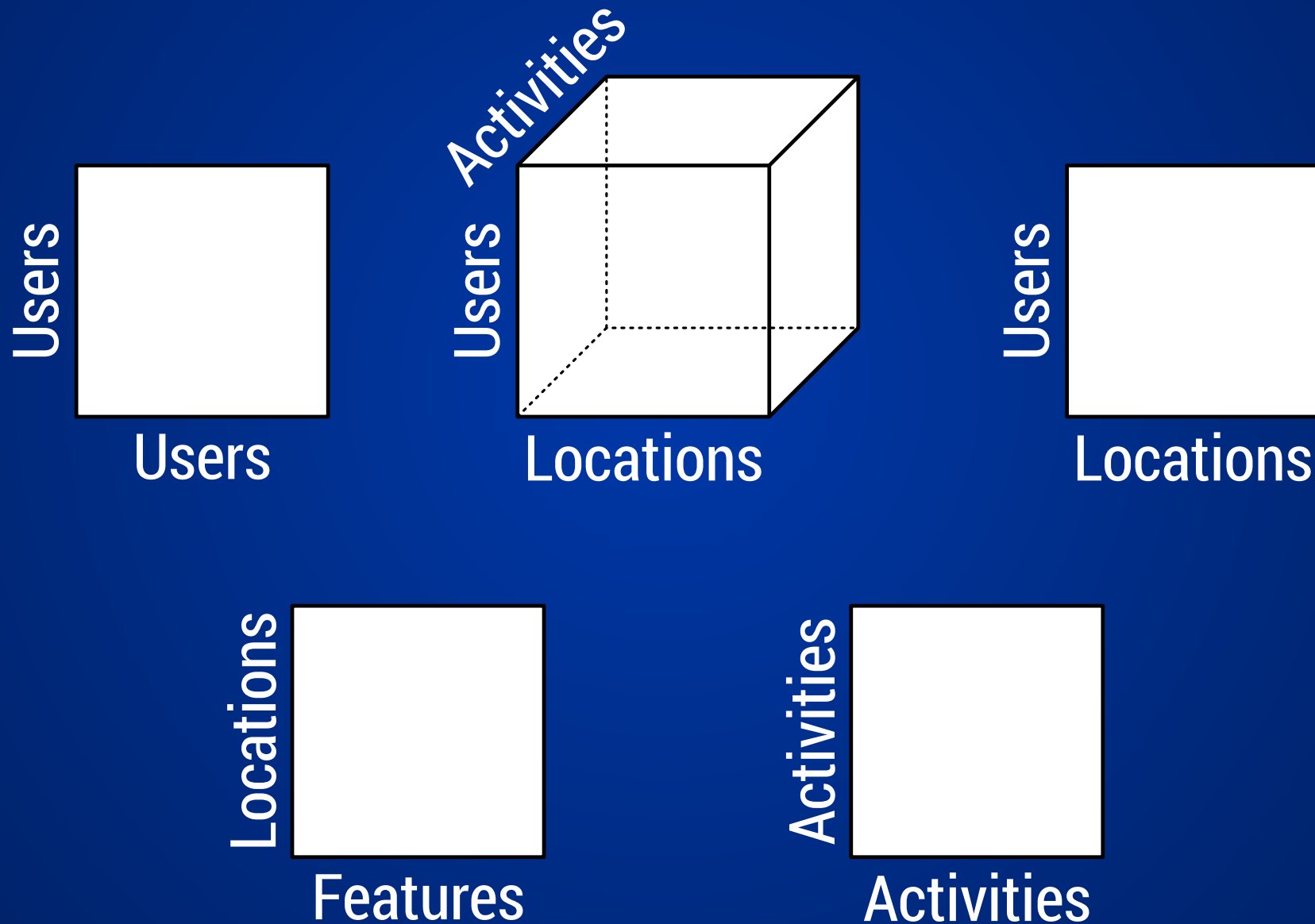
$$A \cdot x = \lambda \cdot B \cdot x$$

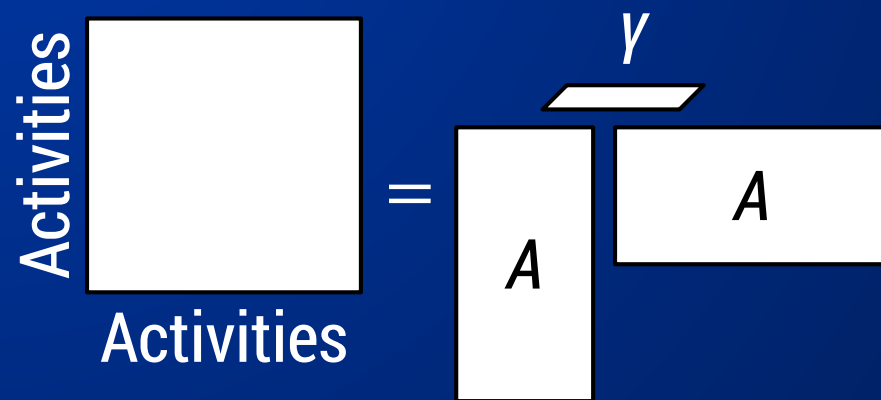
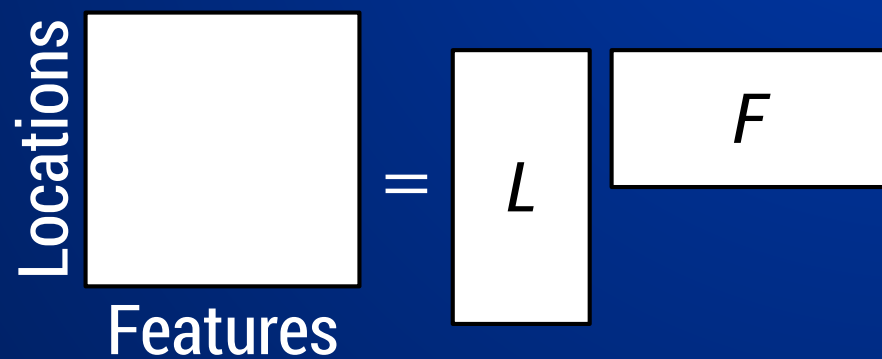
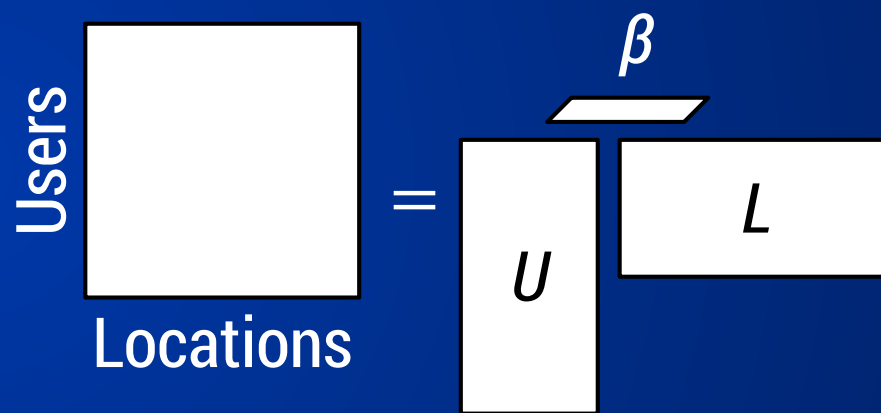
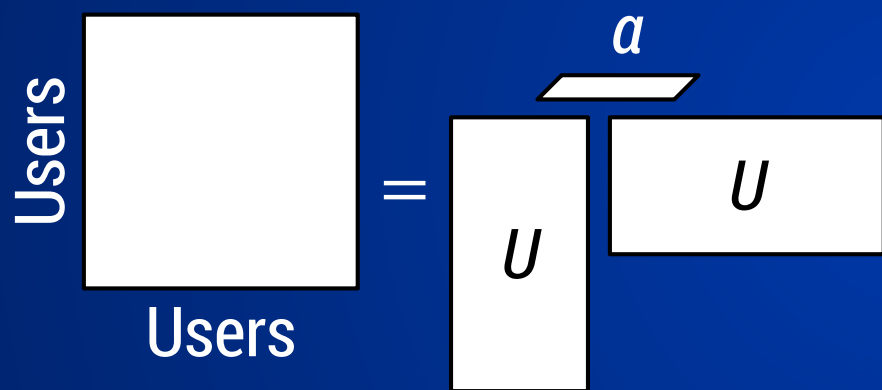
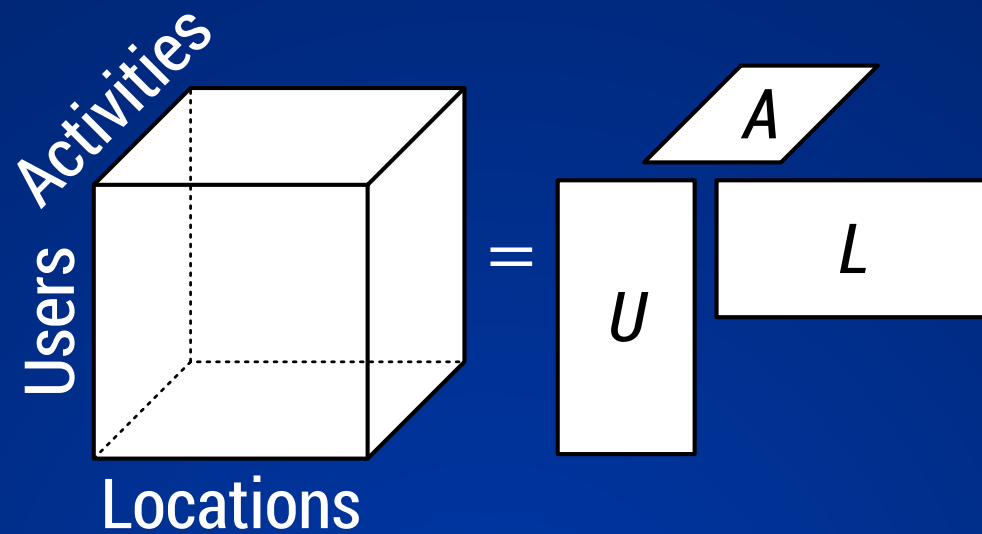






Microsoft Research tracked 164 users' GPS coordinates for 2.5 years
Let's apply data fusion to predict if users took part in activities!





TENSORLAB

```
model.factorizations.UserLocAct.data = UserLocAct;
model.factorizations.UserLocAct.cpd = {'U', 'L', 'A'};

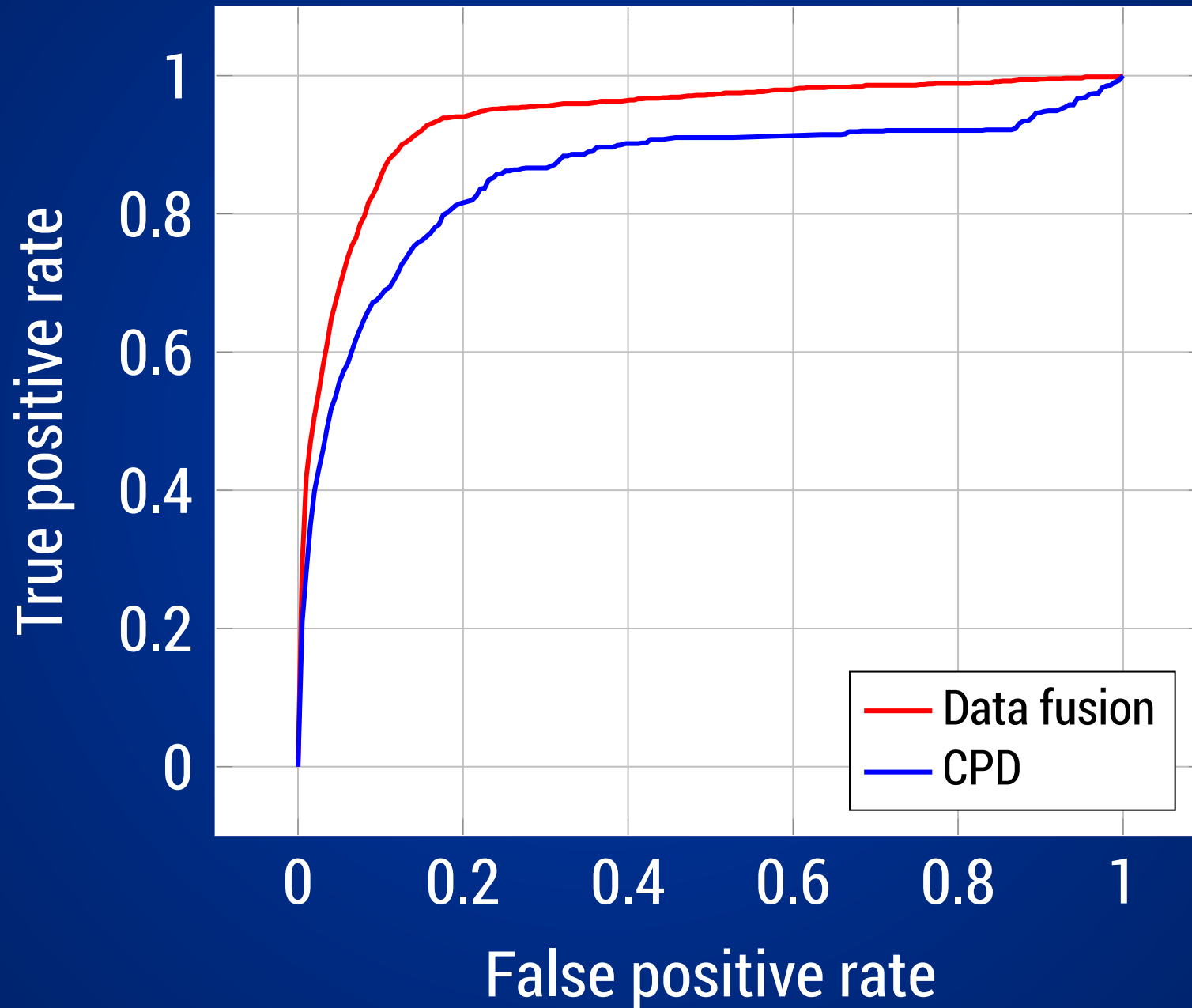
model.factorizations.UserUser.data = UserUser;
model.factorizations.UserUser.cpd = {'U', 'U', 'alpha'};

model.factorizations.LocFea.data = LocFea;
model.factorizations.LocFea.cpd = {'L', 'F'};

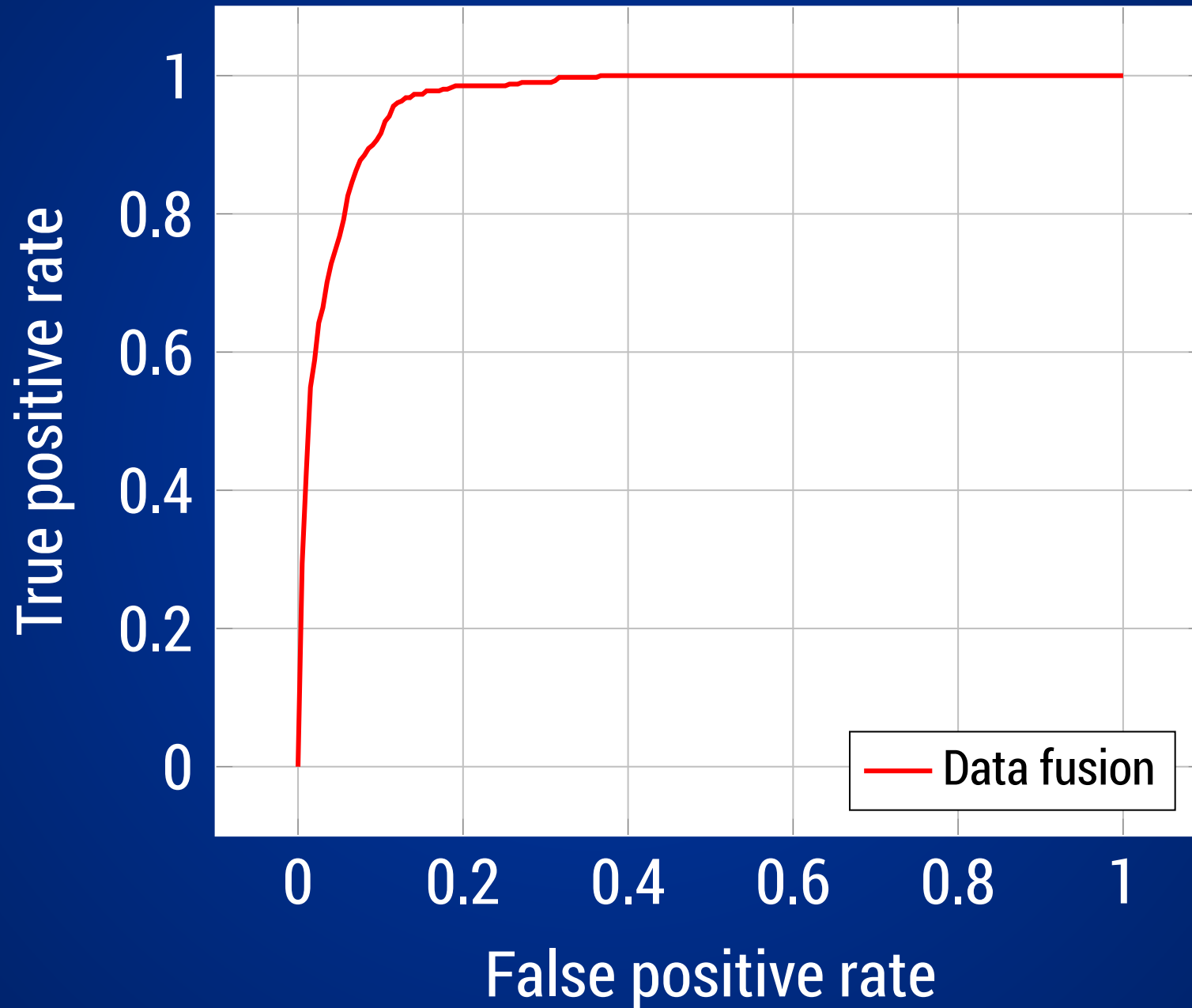
model.factorizations.ActAct.data = ActAct;
model.factorizations.ActAct.cpd = {'A', 'A', 'beta'};

model.factorizations.UserLoc.data = UserLoc;
model.factorizations.UserLoc.cpd = {'U', 'L', 'gamma'};
```

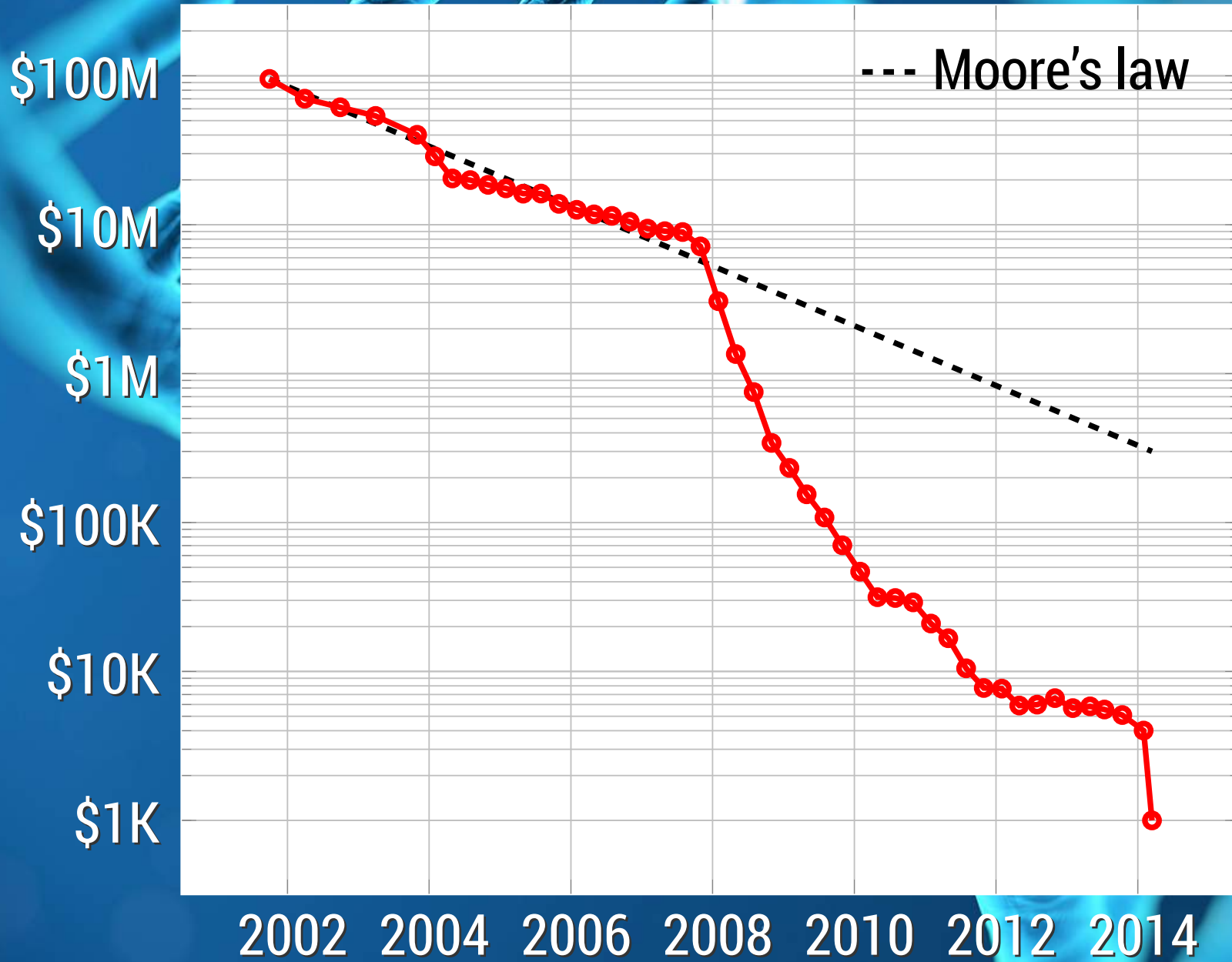
80% OF USER-LOCATION-ACTIVITY MISSING



50 USERS MISSING



COST OF SEQUENCING A HUMAN GENOME



THE END?

- Big data is upon us!
- Structured data fusion is a paradigm shift in our capabilities to extract knowledge from this exciting new resource



THANK YOU!

REFERENCES

- Sorber, L., Van Barel, M., De Lathauwer, L. Structured data fusion. ESAT-STADIUS IR 13-177, KU Leuven. Submitted.
- Sorber, L., Van Barel, M., De Lathauwer, L. Numerical solution of bivariate and polyanalytic polynomial systems. *SIAM Journal on Numerical Analysis*. Accepted.
- Sorber, L., Domanov, I., Van Barel, M., De Lathauwer, L. Exact line and plane search for tensor optimization. ESAT-STADIUS IR 13-02, KU Leuven. Submitted.
- Sorber, L., Van Barel, M., De Lathauwer, L. Optimization-based algorithms for tensor decompositions. *SIAM Journal on Optimization* 23, 2 (2013), 695–720.
- Sorber, L., Van Barel, M., De Lathauwer, L. Unconstrained optimization of real functions in complex variables. *SIAM Journal on Optimization* 22, 3 (2012), 879–898.