GRETINA

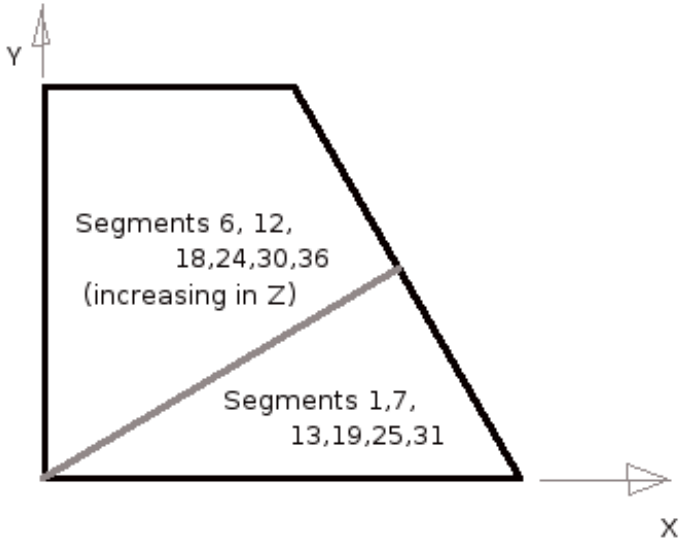
Signal Decomposition

Isidoros Doxas, Chet Nieter
Tech-X Corporation

David Radford
ORNL

signal decomposition is inverse problem

- Electromagnetic code can calculate signals for interaction at a given site, but not inverse.
- Using signals to calculate interaction site is a search operation.
- Current crystal has 4800-8800 sites per segment.



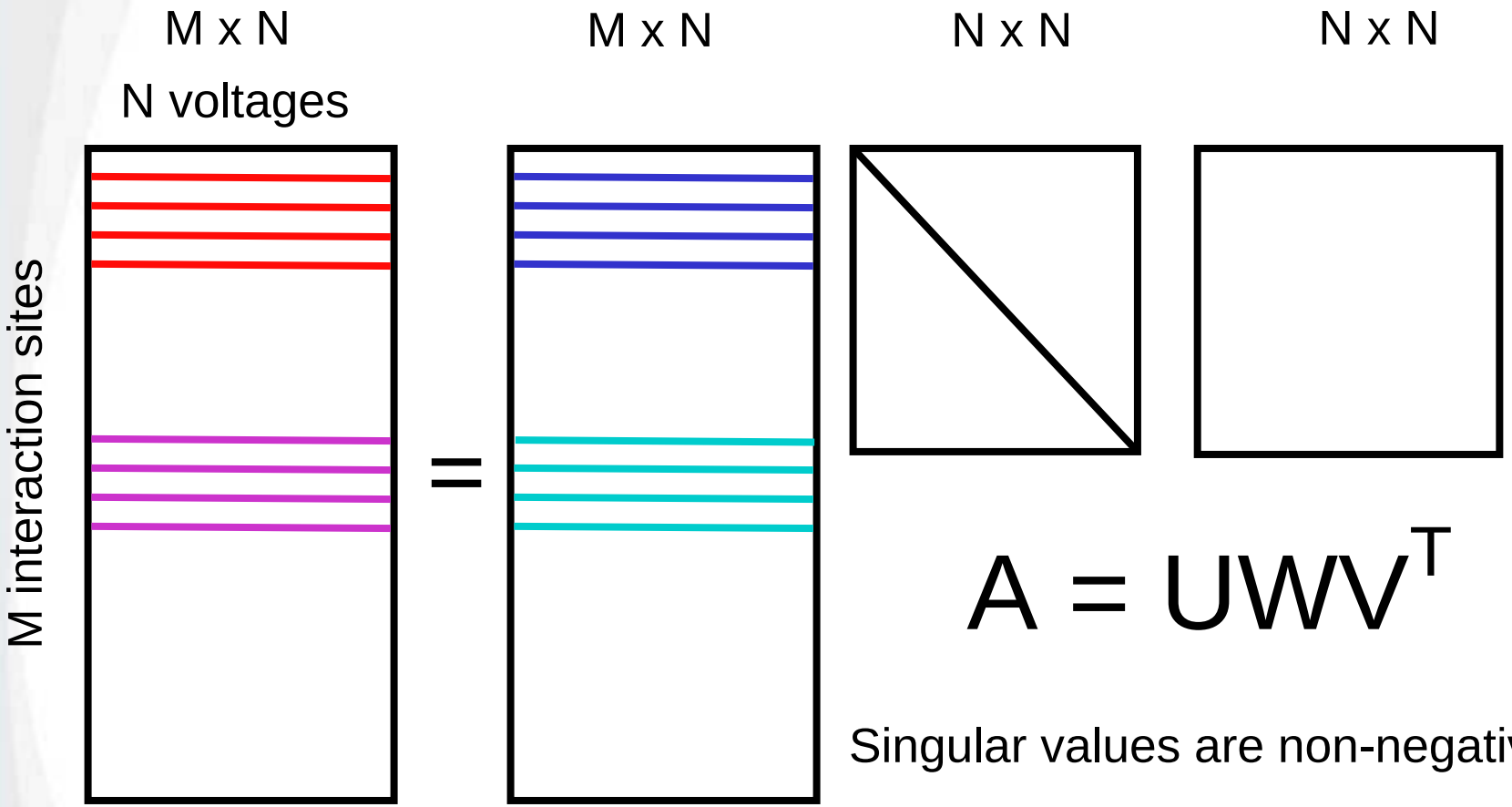
Standard search is $O(N^m)$

- For m interactions on N crystal sites, all $O(N^m)$ combinations of signals are calculated, and a χ^2 test identifies the best fit.
- For an average 6800-site segment, localizing a 2-interaction event requires a search over 23×10^6 vectors of length 1850.
- Calculation needs to complete in ~4ms to prepare for next shot, which translates to computational requirements of $O(10^{13})$FLOPS per CPU (for a ~100 CPU cluster).
- Three-interaction events will need an extra three orders of magnitude.
- Larger segments are even more expensive.

SVD search is O(Nm)

- Singular Value Decomposition is a standard method for searching large databases for similarity.
- Dimensionality of vectors is drastically reduced (~50 instead of 1850).
- The candidate signal is compared to each possible component independently, and a similarity measure is calculated.

SVD possible for any MxN matrix with M>N



$$A = UWV^T$$

Singular values are non-negative

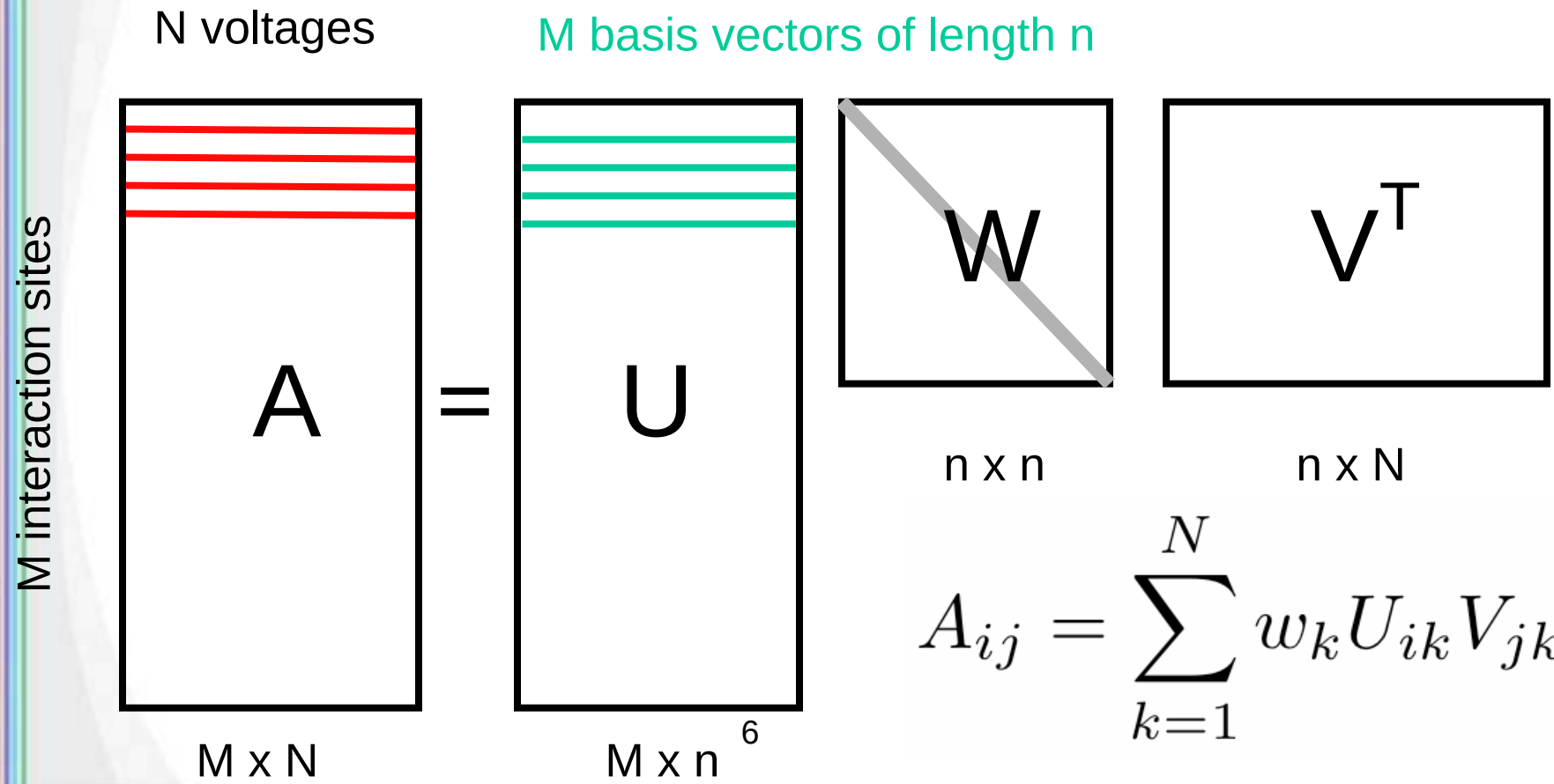
Decomposition is unique up to the interchange of entire raws or columns

U, V are column orthonormal: $U^TU=V^TV=I$

Dimension reduction: keep only a few singular values

If the singular value spectrum decays quickly, A can be approximated well by only a few singular values

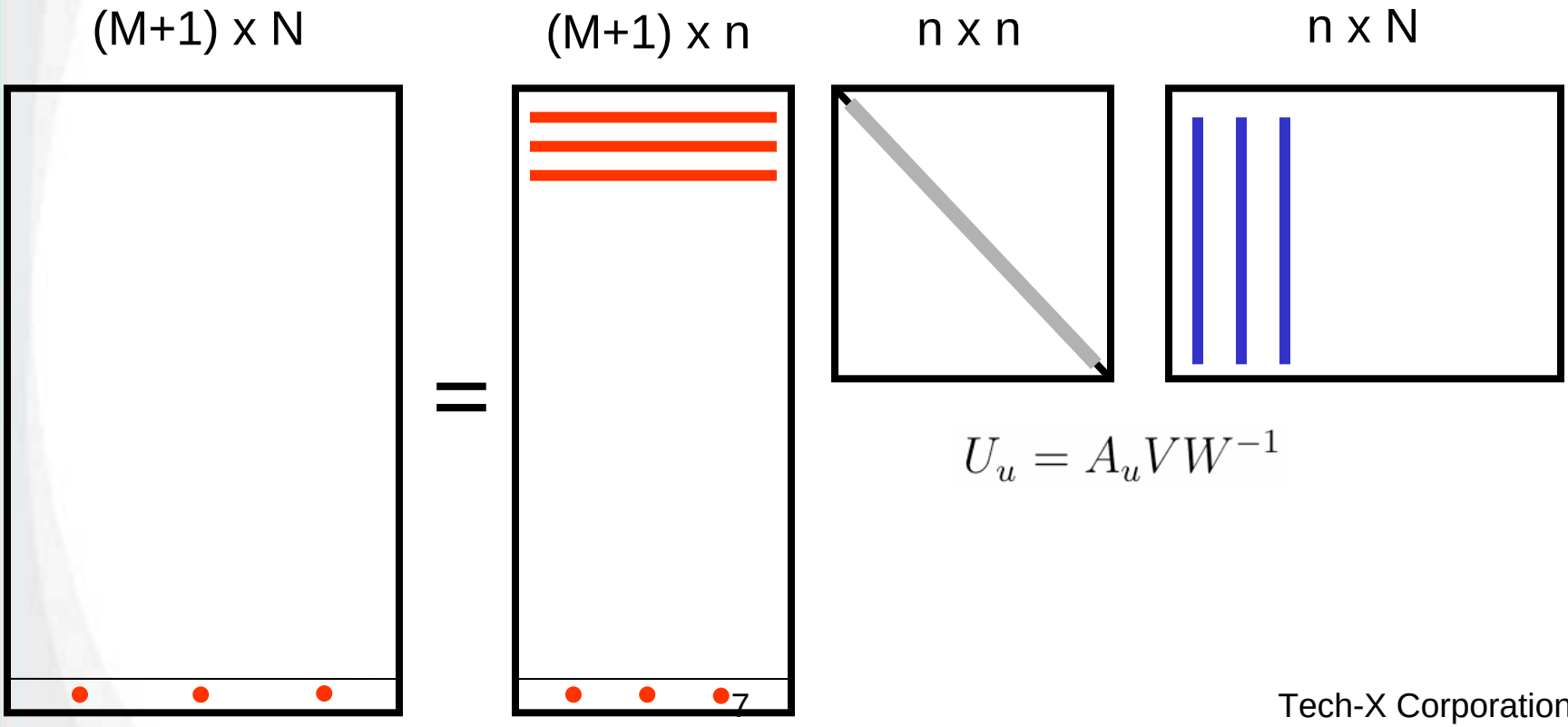
eg. in language applications $10^5 < M < 10^6$, $10^5 < N < 10^6$, n~300



$$A_{ij} = \sum_{k=1}^{N} w_k U_{ik} V_{jk}$$

Arbitrary vector can be projected onto basis

$$
\begin{aligned}
[A : A_u] &= [U : U_u]\, W V^T \\
[A : A_u]\, V &= [U : U_u]\, W V^T V \\
[A : A_u]\, V W^{-1} &= [U : U_u]
\end{aligned}
$$



$$U_u = A_u V W^{-1}$$

Different similarity measures emphasize different aspects

- Dot Product $$x.y = \sum_{i=1}^{N} x_i y_i$$
- Cosine $$\cos(\theta_{xy}) = \frac{x.y}{|x||y|}$$
- Euclidean $$euclid(x, y) = \sqrt{\sum_{i=1}^{N} (x_i - y_i)^2}$$

We have used cosine distances so far.
It de-emphasizes vector length in similarity measure.

Grid Designed for 1% Noise

- Current grid is laid on cylindrical coordinates.
- Maximum signal is 1.00V
- Experimental noise, typically +/- 0.01V.
- Criterion for nearest neighbour is that RMS difference be 0.01

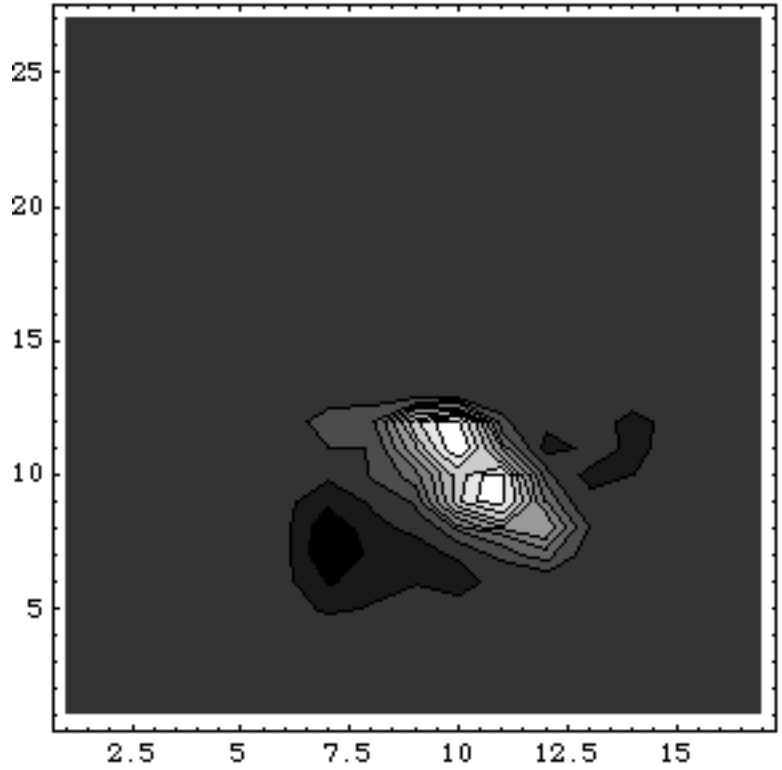
Time Shift Increases Uncertainty

- The measured timeseries does not necessarily start at the same time as the simulation timeseries.
- Difference can be 0-1 timestep.
- For standard χ^2 test this introduces a small error at every bin.

SVD search very efficient for GRETINA

With +- 1% noise
time: 5.620000 ms
crit= 2
scale= 169.000000
nev=200
nevlo=50
319 **6.380000**
118, 117, 110

Without noise
time: 5.620000 ms
crit= 2
scale= 169.000000
nev=200
nevlo=50
313 **6.260000**
114, 112, 112



SVD search of high precision when using many s-values

No noise

time: 13.540000 ms	time: 6.816000 ms	time: 3.628000 ms
crit= 1	crit= 1	crit= 1
scale= 121.000000	scale= 121.000000	scale= 121.000000
nevlo=400	nevlo=200	nevlo=100
4 0.080000	161 3.220000	404 8.080000

Interesting to see if there is a number of singular values for which the linear search can give 100% hits.

With noise, SVD search most efficient with fewer s-values

+- 1% Noise

time: 13.052000 ms
crit= 2
scale= 144.000000
nevlo=400
4714 **94.280000**

time: 10.228000 ms
crit= 2
scale= 144.000000
nevlo=300
2110 **42.200000**

time: 6.600000 ms
crit= 2
scale= 144.000000
nevlo=200
194 **3.880000**

time: 3.488000 ms
crit= 2
scale= 144.000000
nevlo=100
322 **6.440000**

time: 1.964000 ms
crit= 2
scale= 144.000000
nevlo=50
409 **8.180000**

Including Time Shift can "Improve" Performance

+/- 1% Noise, No Time Shift

time: 6.600000 ms
crit= 2
scale= 144.000000
nevlo=200
194 **3.880000**

time: 3.488000 ms
crit= 2
scale= 144.000000
nevlo=100
322 **6.440000**

time: 1.964000 ms
crit= 2
scale= 144.000000
nevlo=50
409 **8.180000**

+/- 1% Noise, Random Time Shift (basis at .33, .66)

time: 9.608000 ms
crit= 2
scale= 144.000000
nevlo=200
459 **9.180000**

time: 5.144000 ms
crit= 2
scale= 144.000000
nevlo=100
229 **4.580000**

time: 2.912000 ms
crit= 2
scale= 144.000000
nevlo=50
339 **6.780000**

Three-Interaction events also possible

time: 6.592000 ms
crit= 2
scale= 100.000000
nevlo=120
888 **11.840000**

Random Multiplicity Events also Possible

time: 4.140000 ms
crit= 2
mcrit1=0.890000, mcrit2=0.910000
scale= 100.000000
nevlo=80
629, 636, **12.650000**