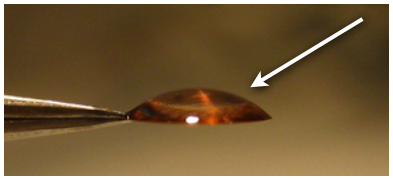
Fast Algorithms for Inverse Problems in Acoustic Scattering and Cryo-Electron Microscopy

Leslie Greengard
Simons Foundation & Courant Institute, NYU

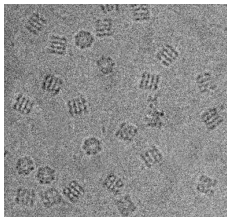
with A. Barnett, C. Borges, A. Gillman, A. Pataki, M. Spivak

2016 New York Scientific Data Summit

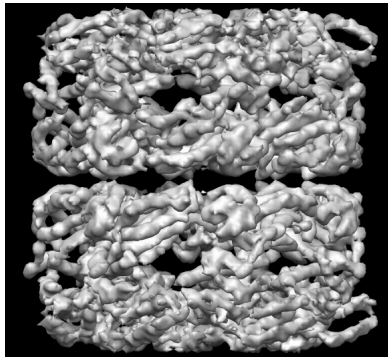
1) Freeze sample



2) EM image

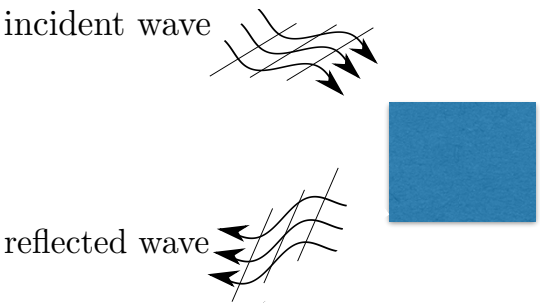


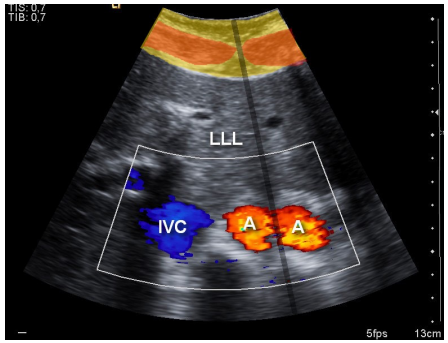
3) Process



courtesy Gabriel Lander

Cryo – electron microscopy





wikipedia.org/wiki/Medical_ultrasound#doppler

© Nevit Dilmen

Ultrasound imaging

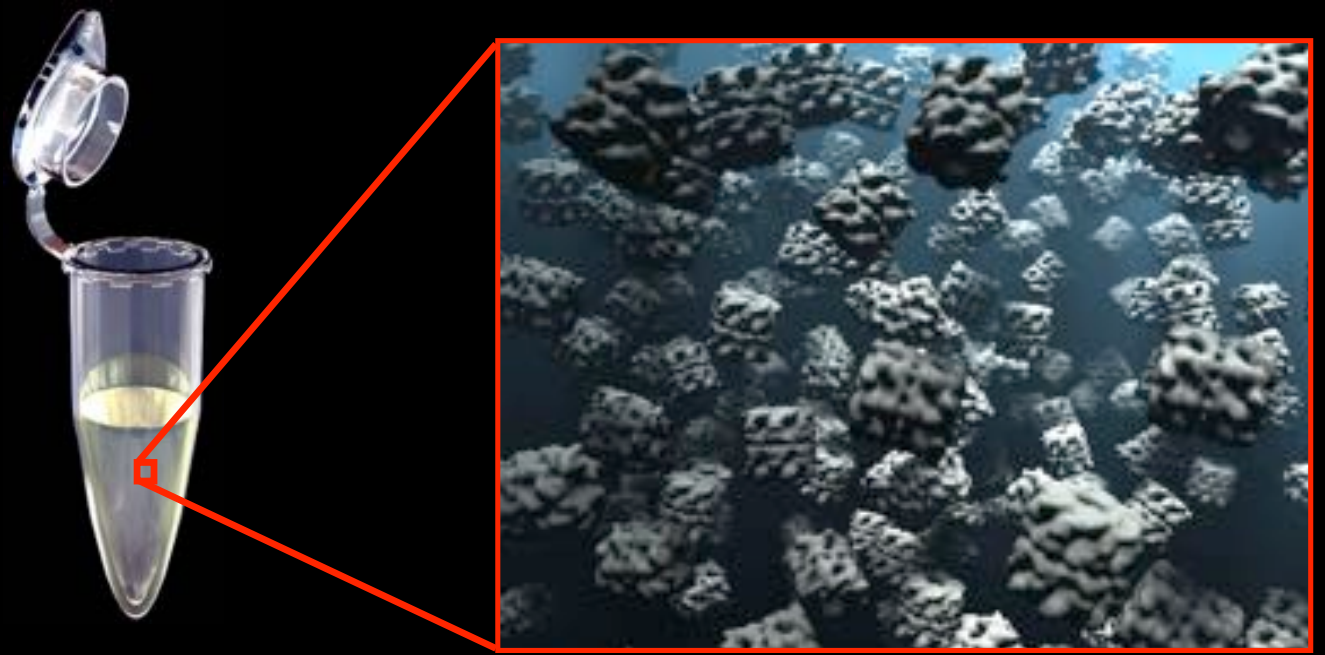
Outline of talk

- Introduction to cryo-electron microscopy
- Reconstruction by frequency marching
- Inverse acoustic scattering
- Reconstruction by frequency marching/ recursive linearization (Y. Chen)
- Examples
- Future work

Transmission Cryo-Electron Microscopy

- Allows near atomic resolution structure from proteins without X-ray crystallography
- No need for crystal growth
- Closer to native environment
- Leads to large, noisy data sets and complex processing task
- Thanks to Gabriel Lander (Scripps Research Institute), Bridget Carragher, Clint Potter, NY Structural Biology Center
- Joint work with Alex Barnett, Marina Spivak, Andras Pataki

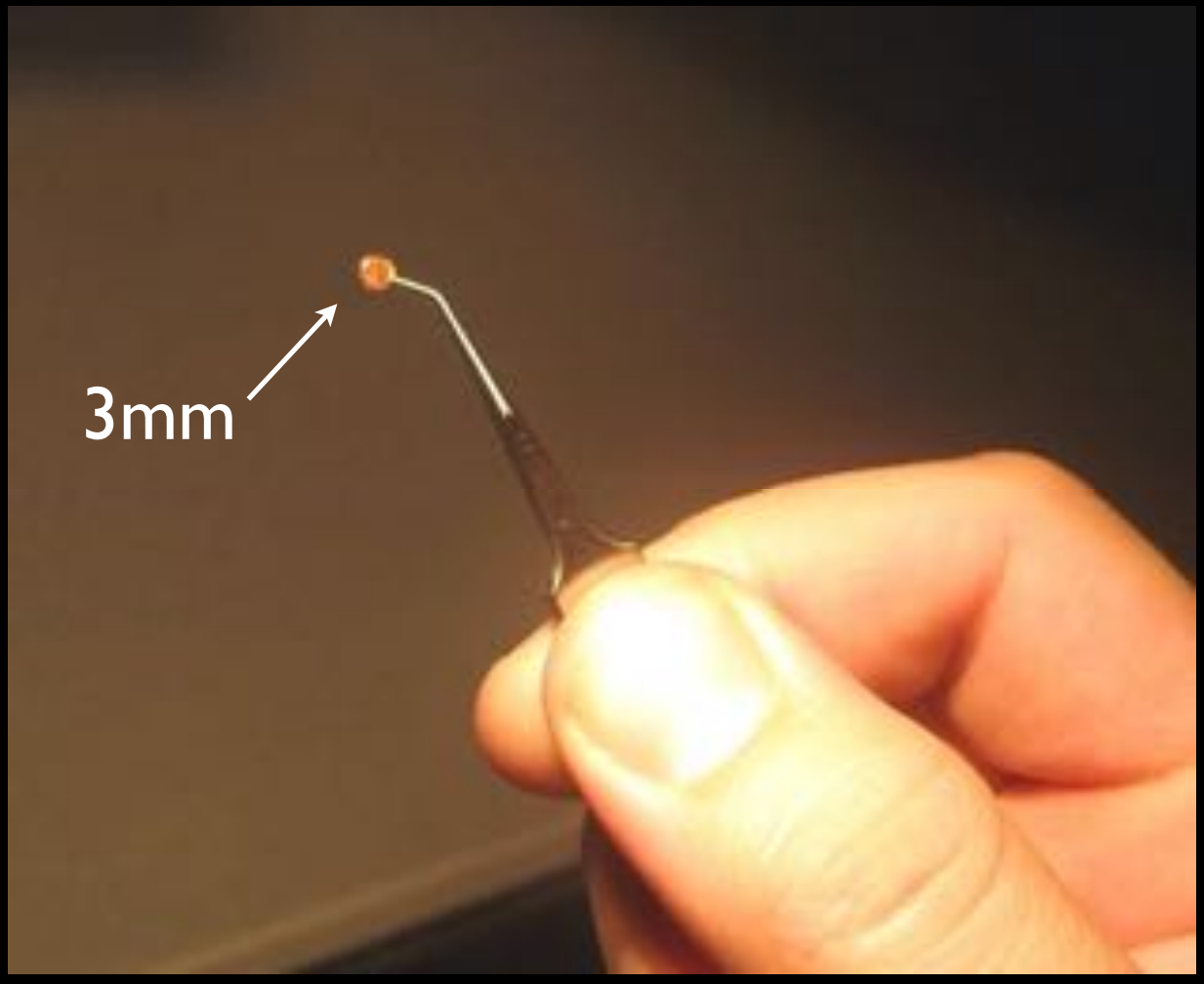
This plastic tube contains our molecular sample of interest



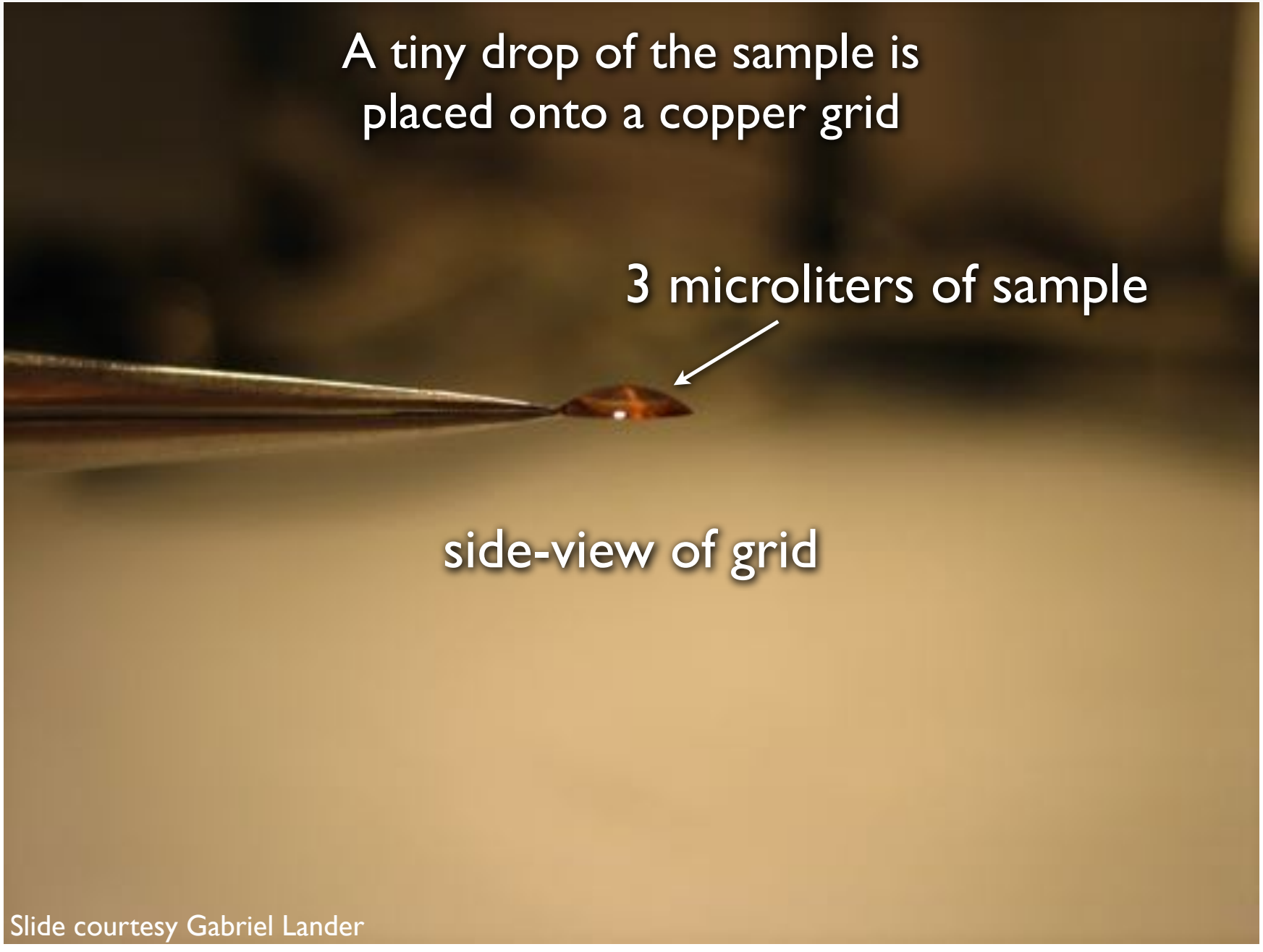
Inside the tube, the nanoparticles are freely mobile. The sample must be frozen in order to image it

Slide courtesy Gabriel Lander

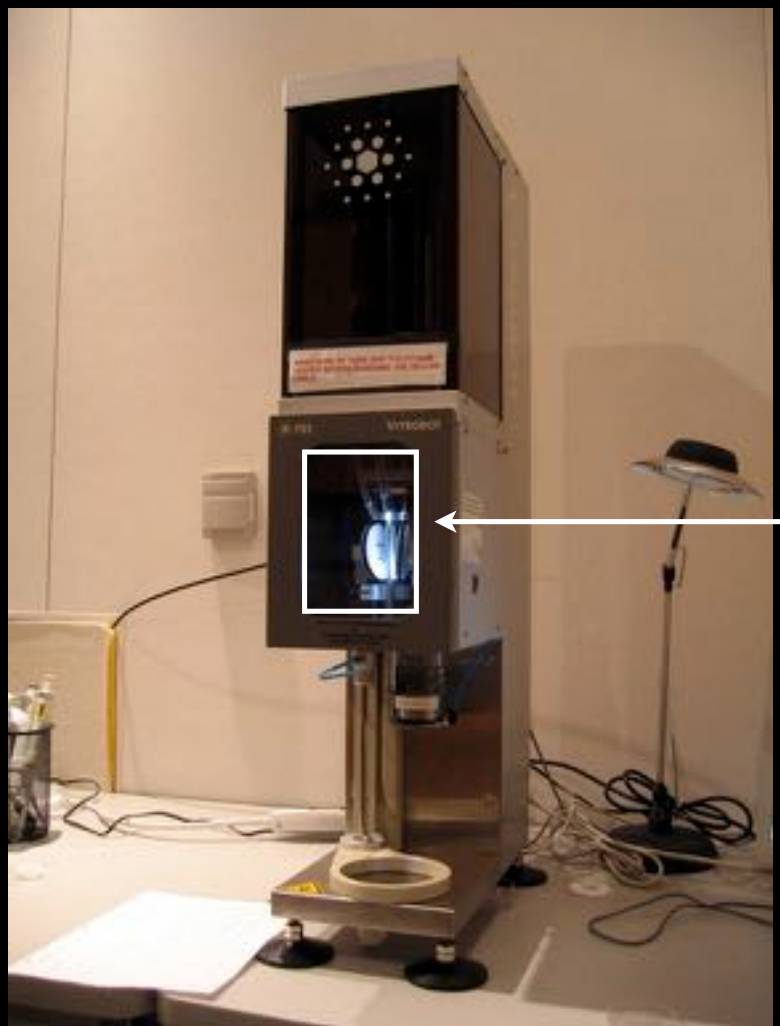
A tiny drop of the sample is placed onto a copper grid



Slide courtesy Gabriel Lander


A tiny drop of the sample is
placed onto a copper grid
3 microliters of sample
side-view of grid
Slide courtesy Gabriel Lander

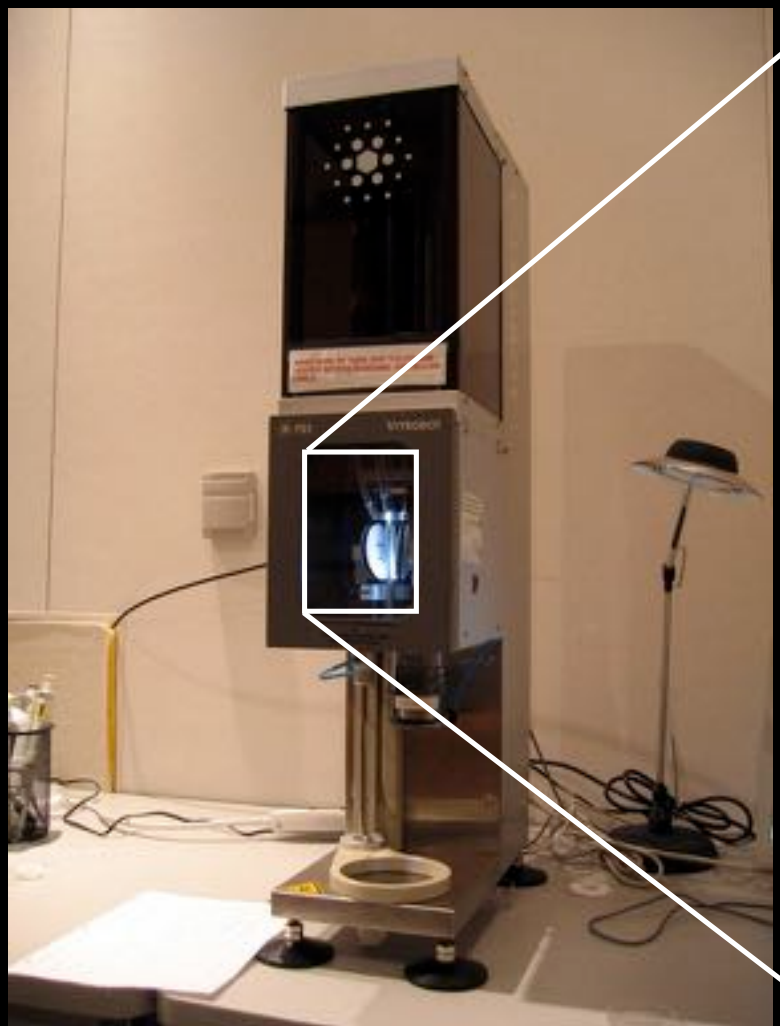
The sample is loaded into a machine called a Vitrobot



this chamber maintains 100% humidity and a temperature of 4°C

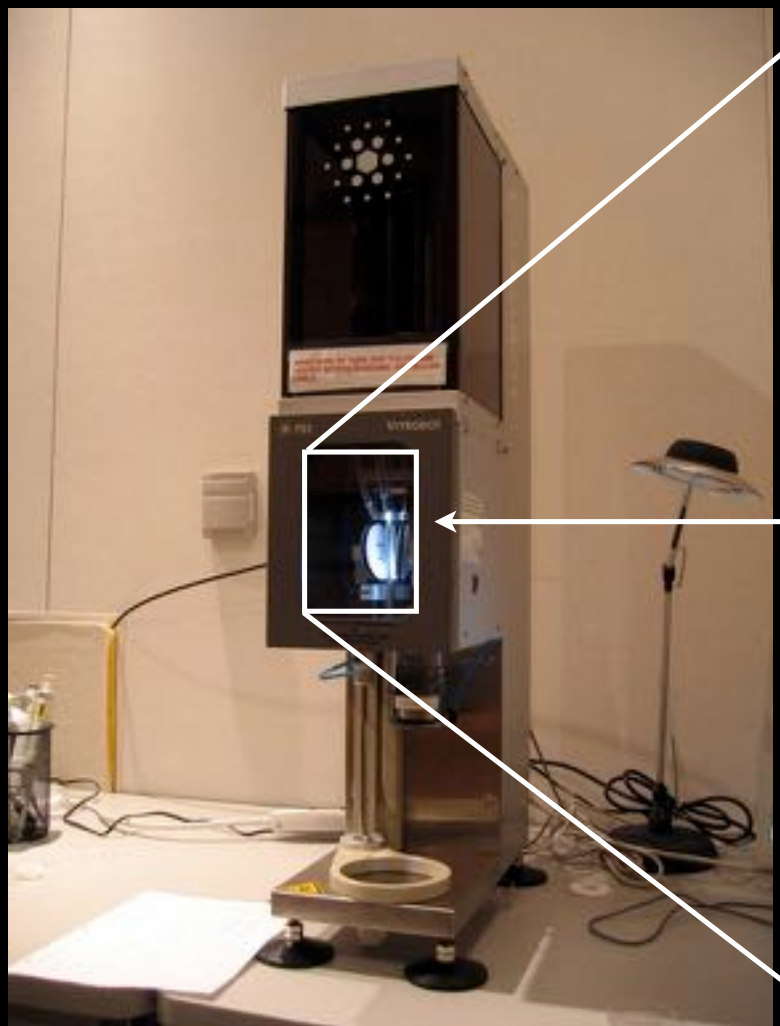
Slide courtesy Gabriel Lander

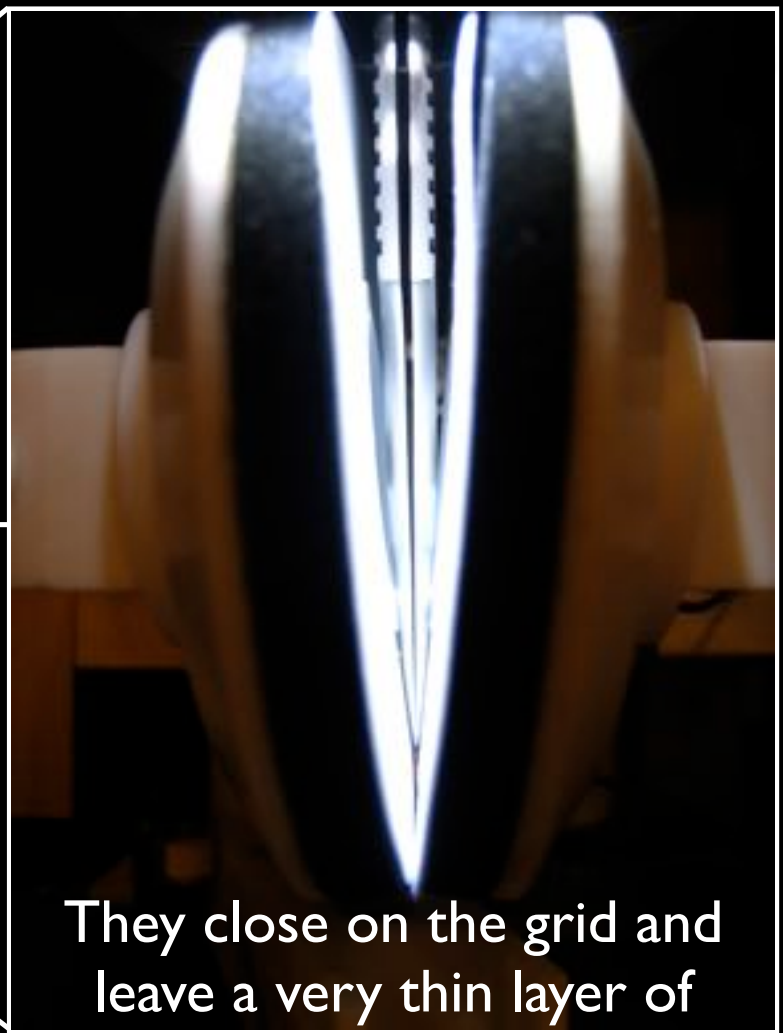
The sample is loaded into a machine called a Vitrobot



Slide courtesy Gabriel Lander

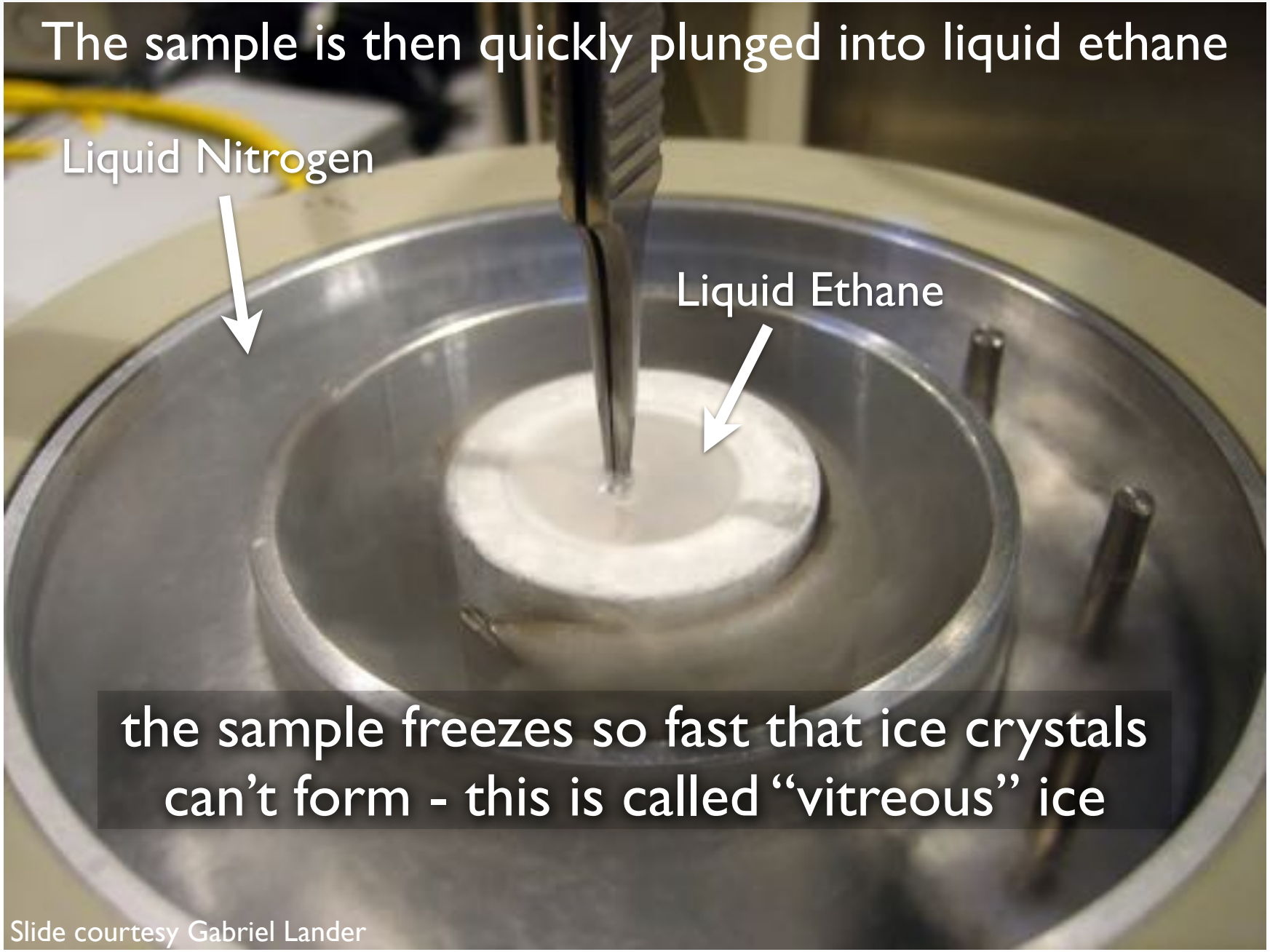
The sample is loaded into a machine called a Vitrobot



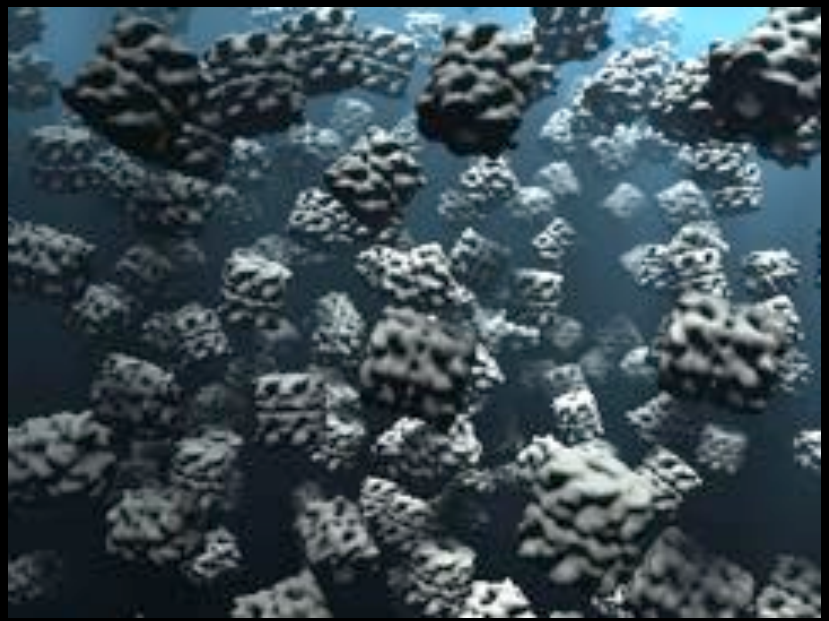


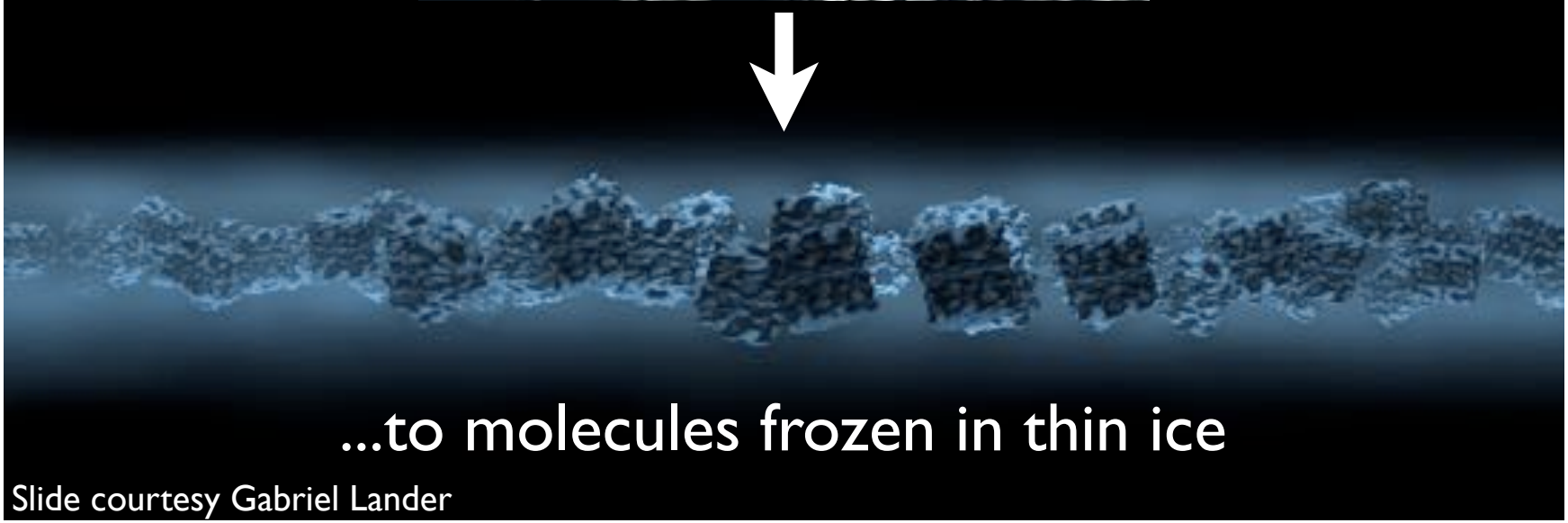
They close on the grid and leave a very thin layer of molecules in solution

Slide courtesy Gabriel Lander


The sample is then quickly plunged into liquid ethane
Liquid Nitrogen
Liquid Ethane
the sample freezes so fast that ice crystals can't form - this is called "vitreous" ice
Slide courtesy Gabriel Lander

That is how we go from sample in solution...





...to molecules frozen in thin ice

Slide courtesy Gabriel Lander

The grid containing the frozen molecules
is loaded into an electron microscope

Now we can reconstruct a 3D structure
from the frozen molecules

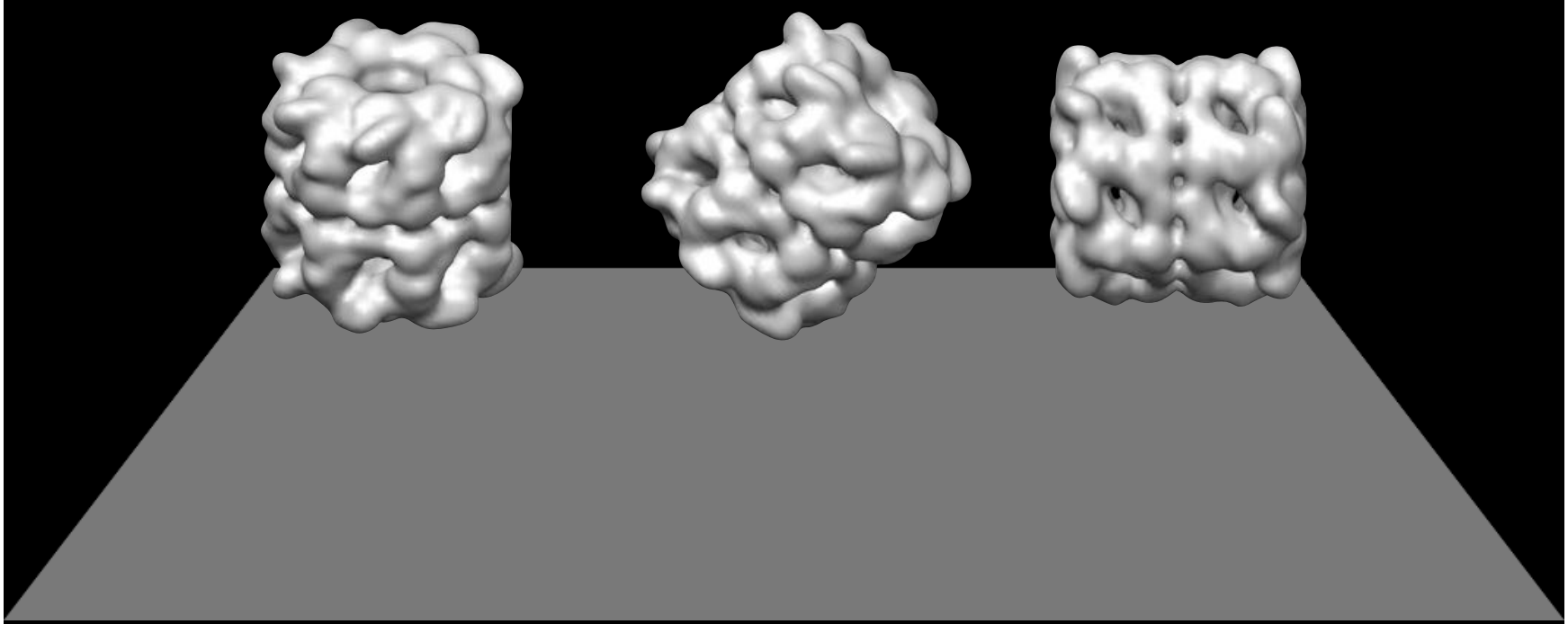
Slide courtesy Gabriel Lander

This is an example of a transmission electron microscope (TEM)

FEI Tecnai F20

Slide courtesy Gabriel Lander

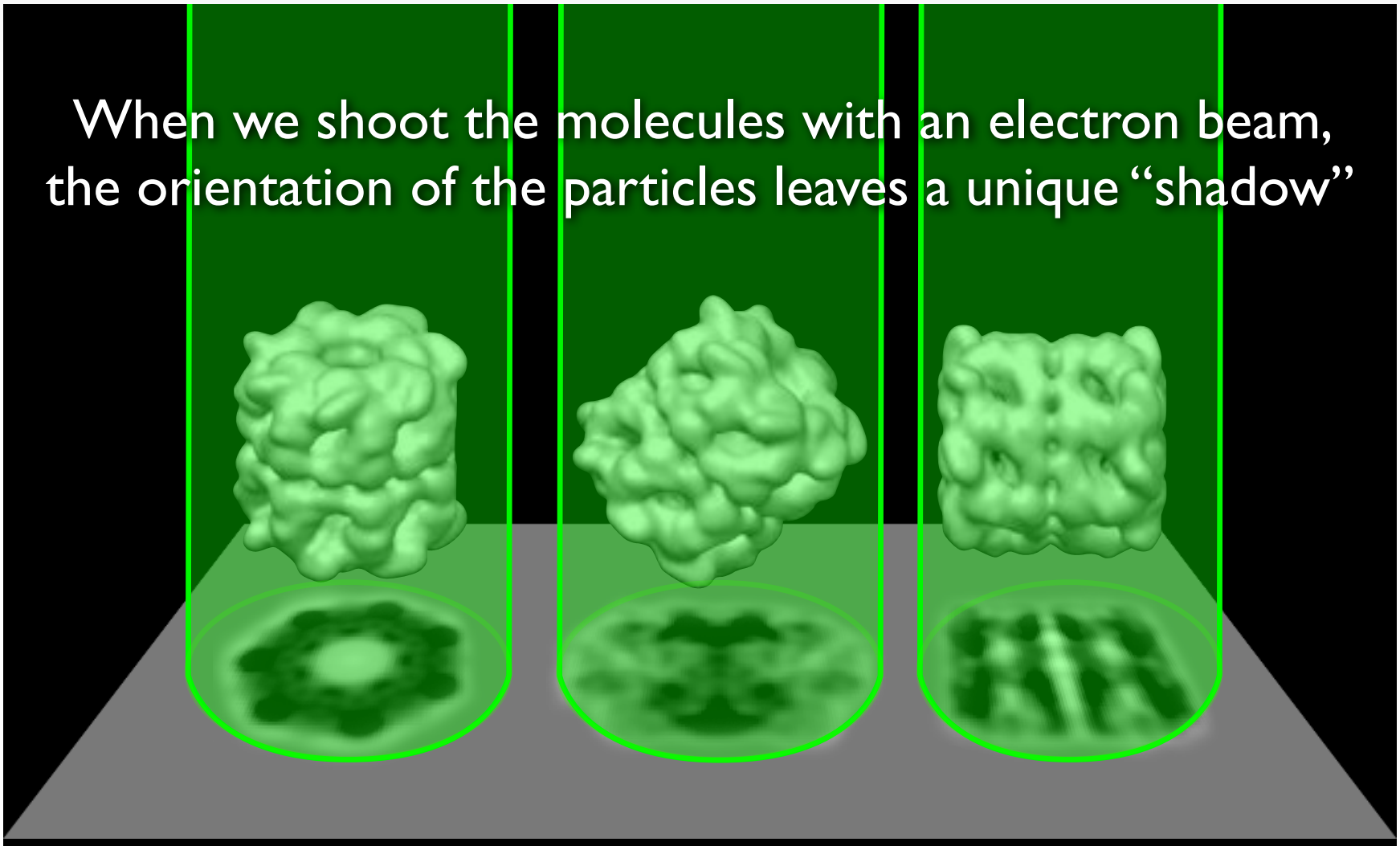
Here are three molecules trapped
in different orientations



Underneath is a detector (such as photographic film)

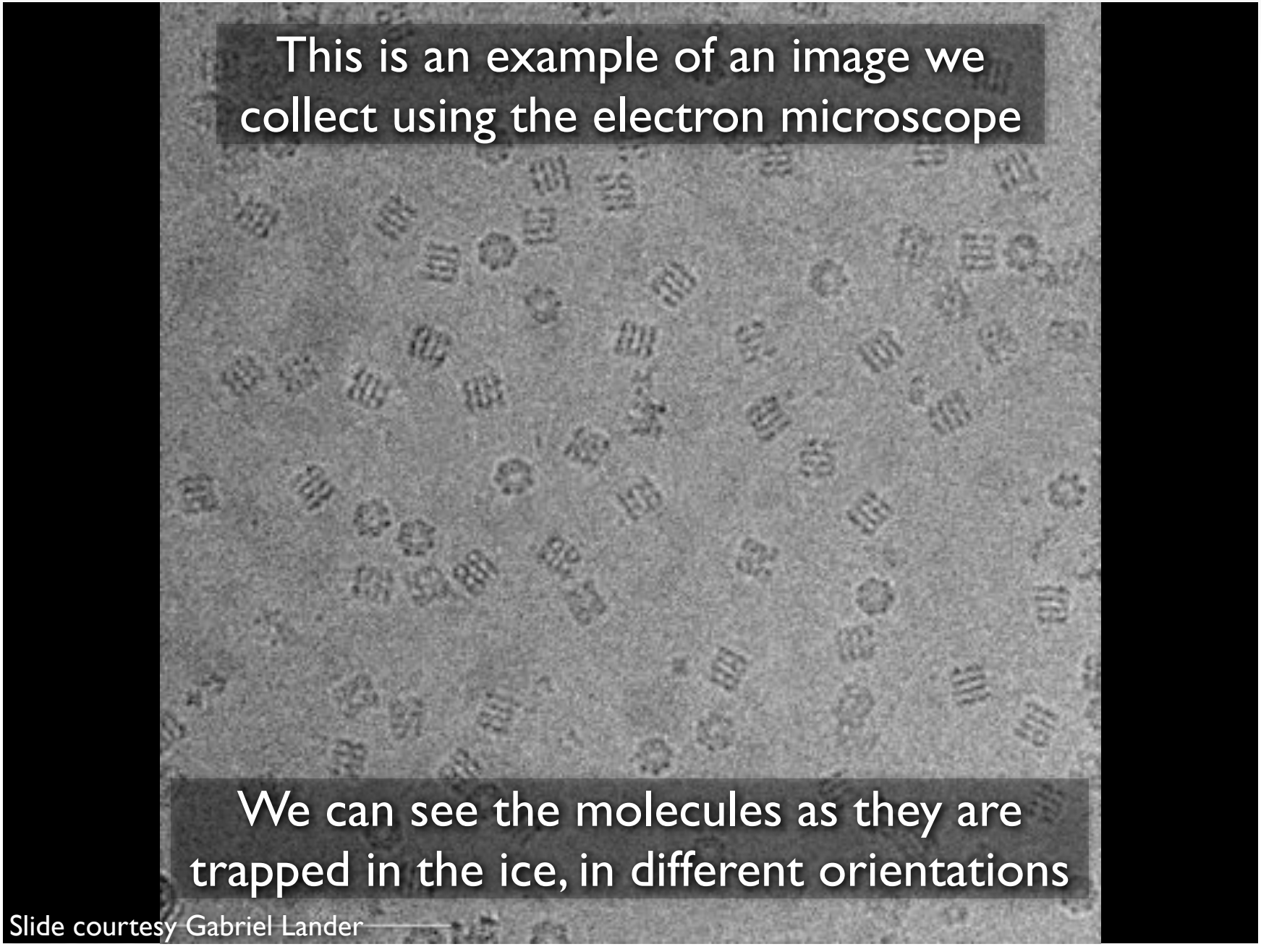
Slide courtesy Gabriel Lander

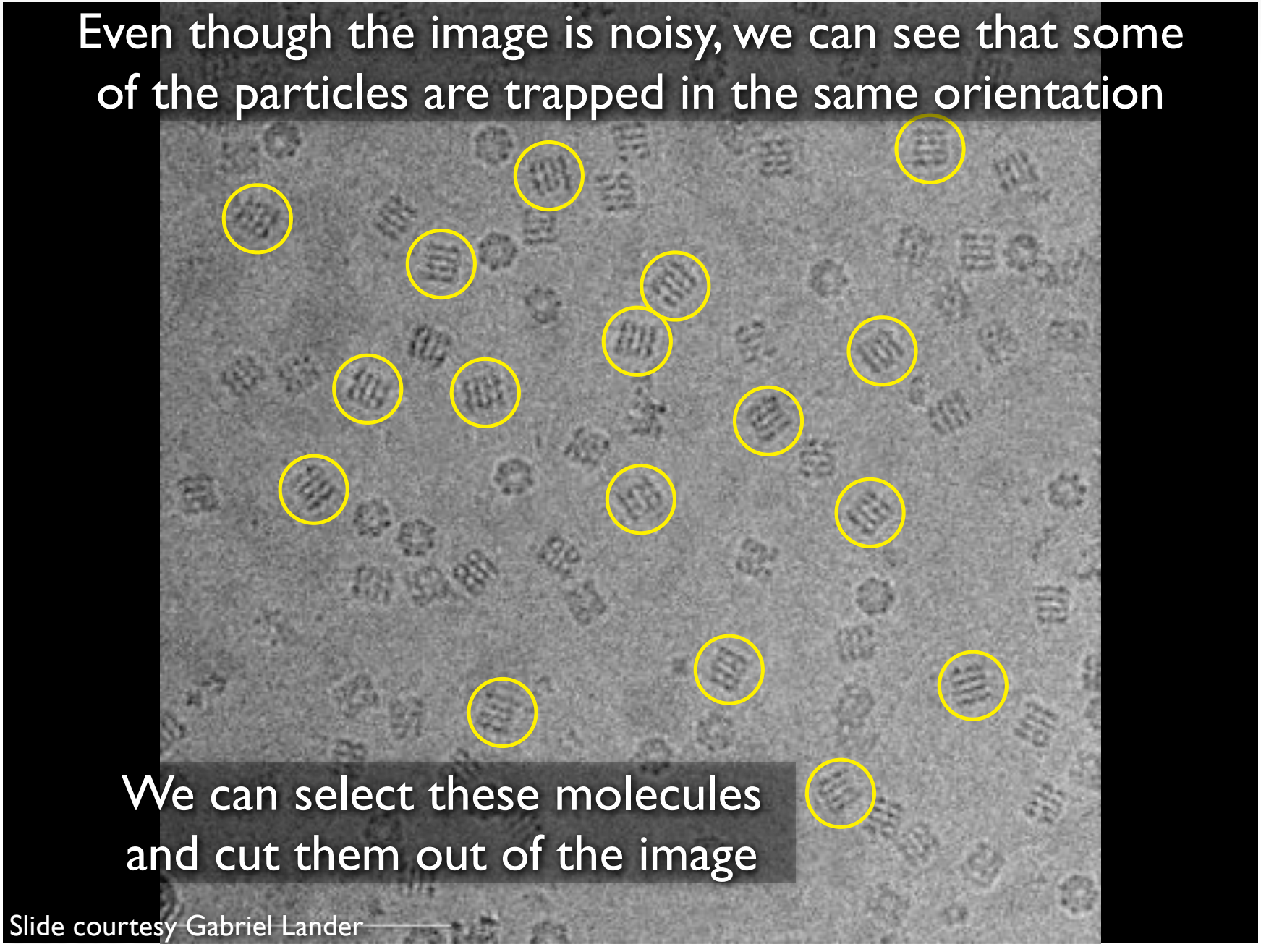
When we shoot the molecules with an electron beam, the orientation of the particles leaves a unique "shadow"



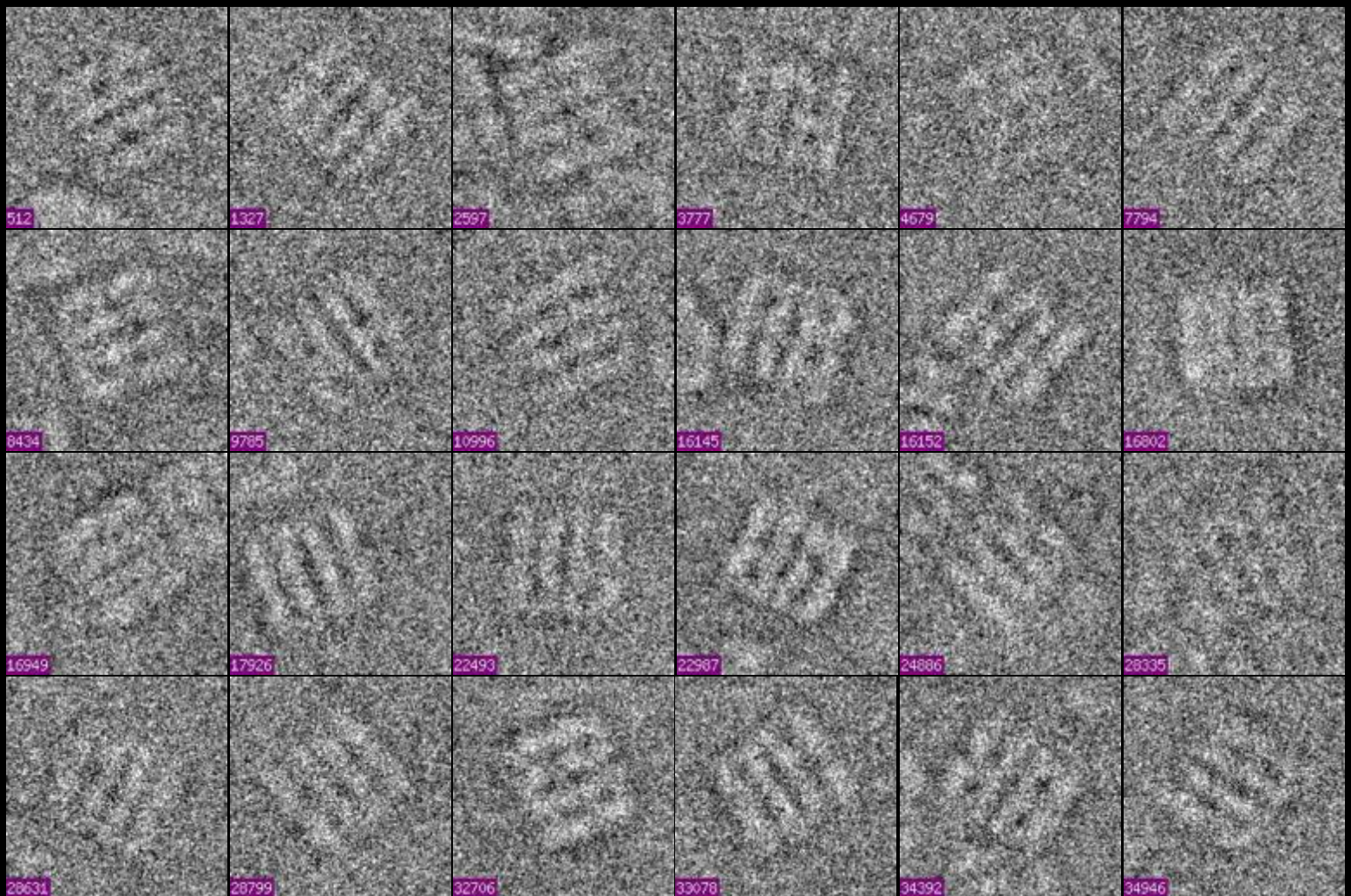
These "shadows" contain all the 3-dimensional information of the molecule, compressed into a 2D image

Slide courtesy Gabriel Lander


This is an example of an image we collect using the electron microscope
We can see the molecules as they are trapped in the ice, in different orientations
Slide courtesy Gabriel Lander

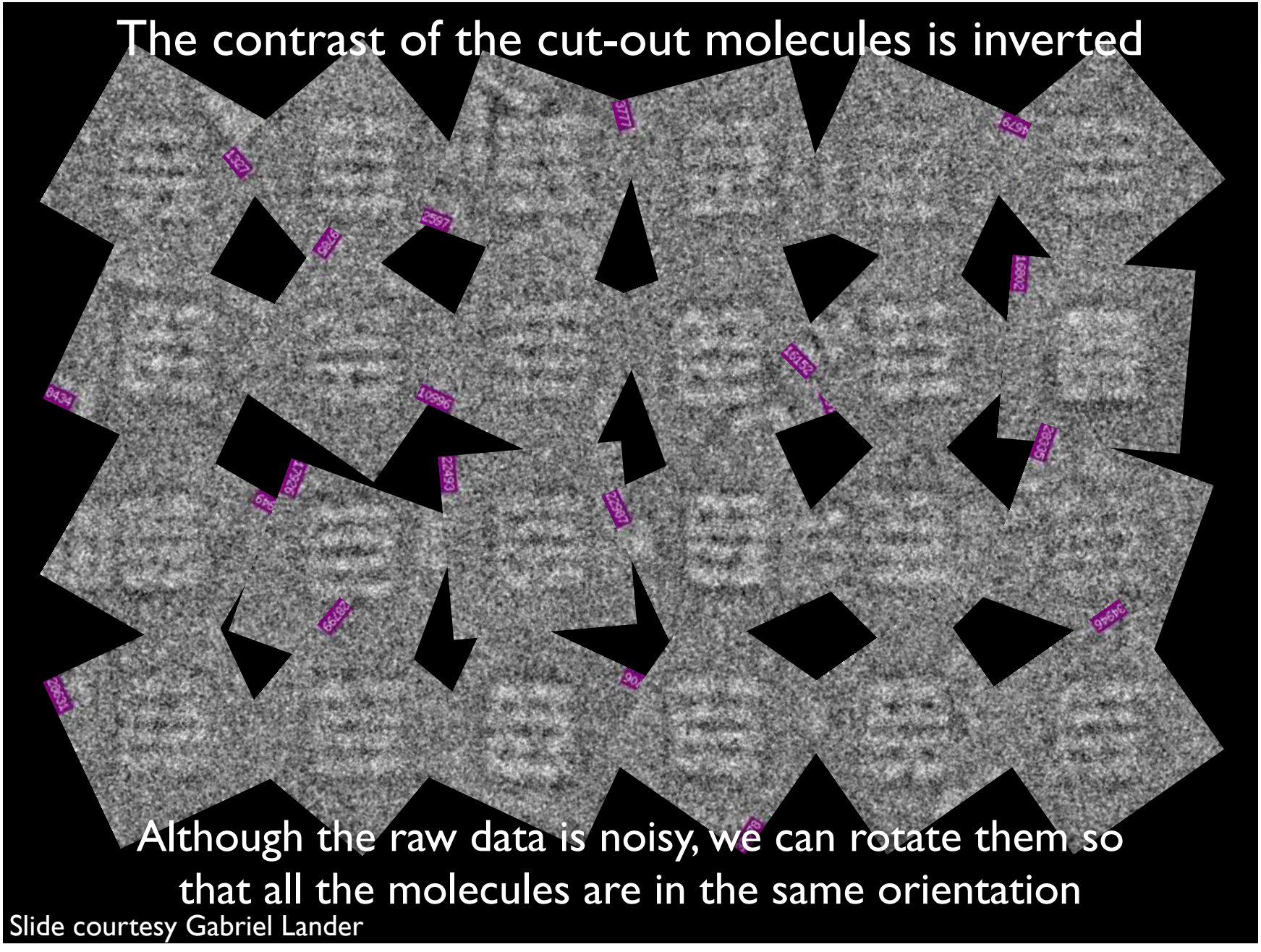

Even though the image is noisy, we can see that some of the particles are trapped in the same orientation
We can select these molecules and cut them out of the image
Slide courtesy Gabriel Lander

The contrast of the cut-out molecules is inverted

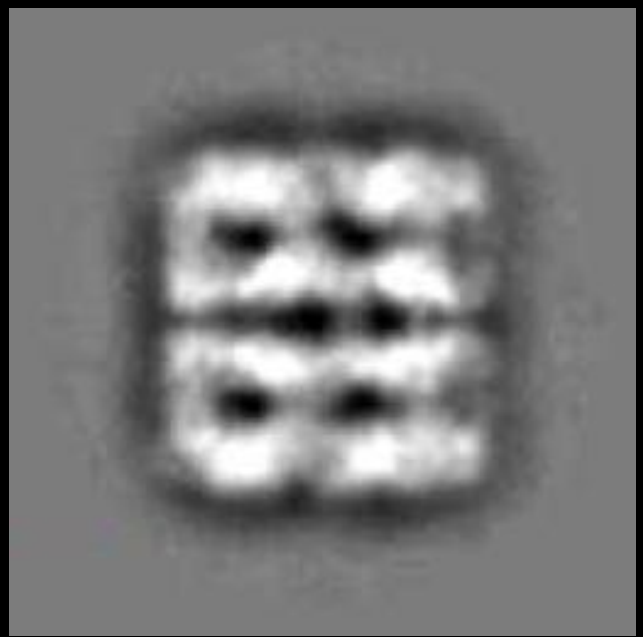


Although the raw data is noisy, we can rotate them so that all the molecules are in the same orientation

Slide courtesy Gabriel Lander

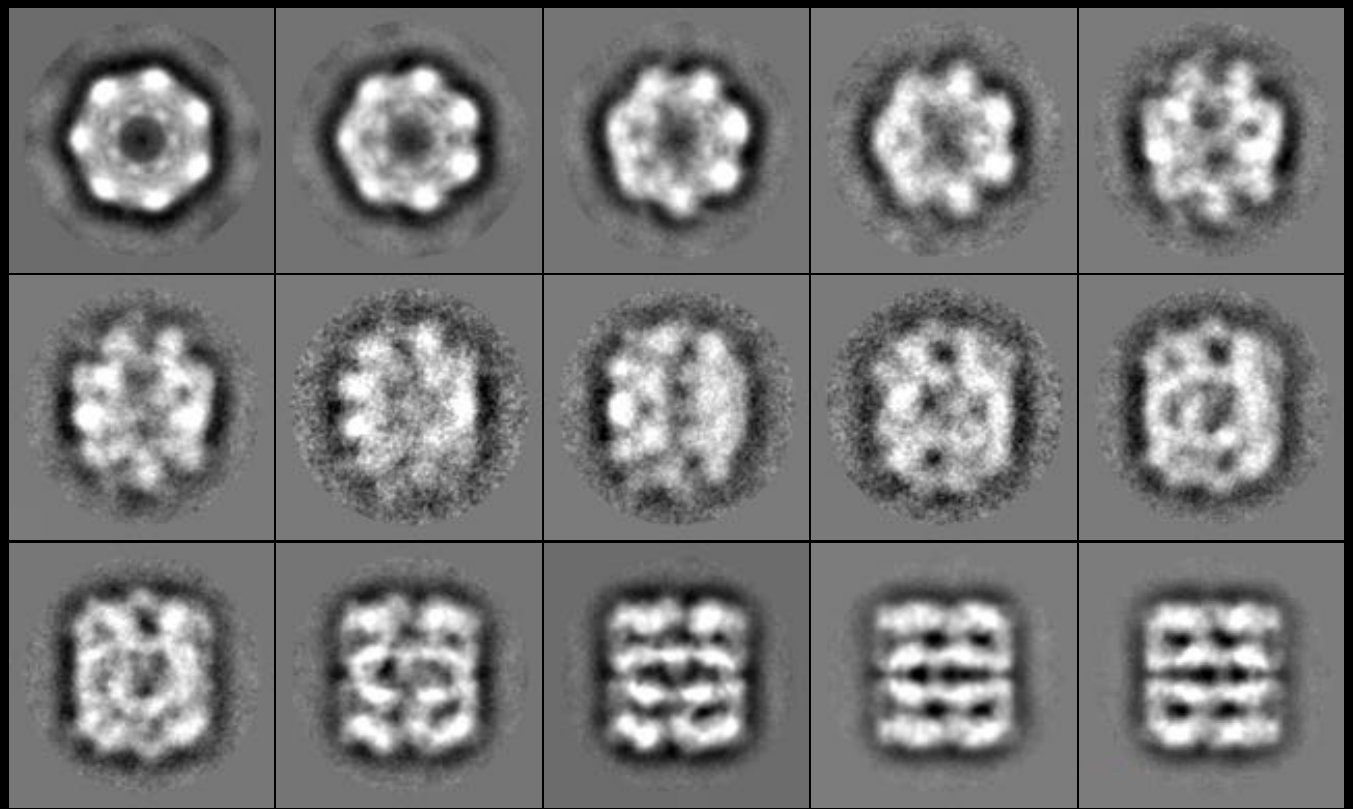

The contrast of the cut-out molecules is inverted
Although the raw data is noisy, we can rotate them so that all the molecules are in the same orientation
Slide courtesy Gabriel Lander

These “aligned” molecules are then added together, and this summed image provides a more detailed view of the molecule in this orientation



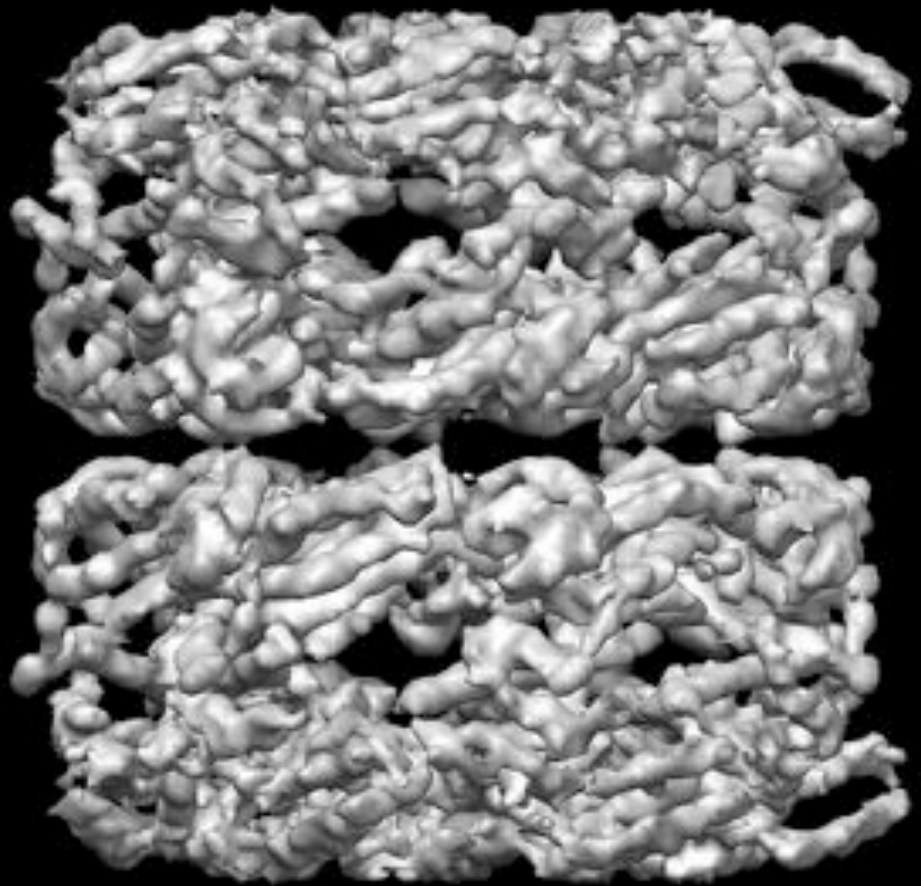
Slide courtesy Gabriel Lander

We perform this same process for the different orientations of the molecule



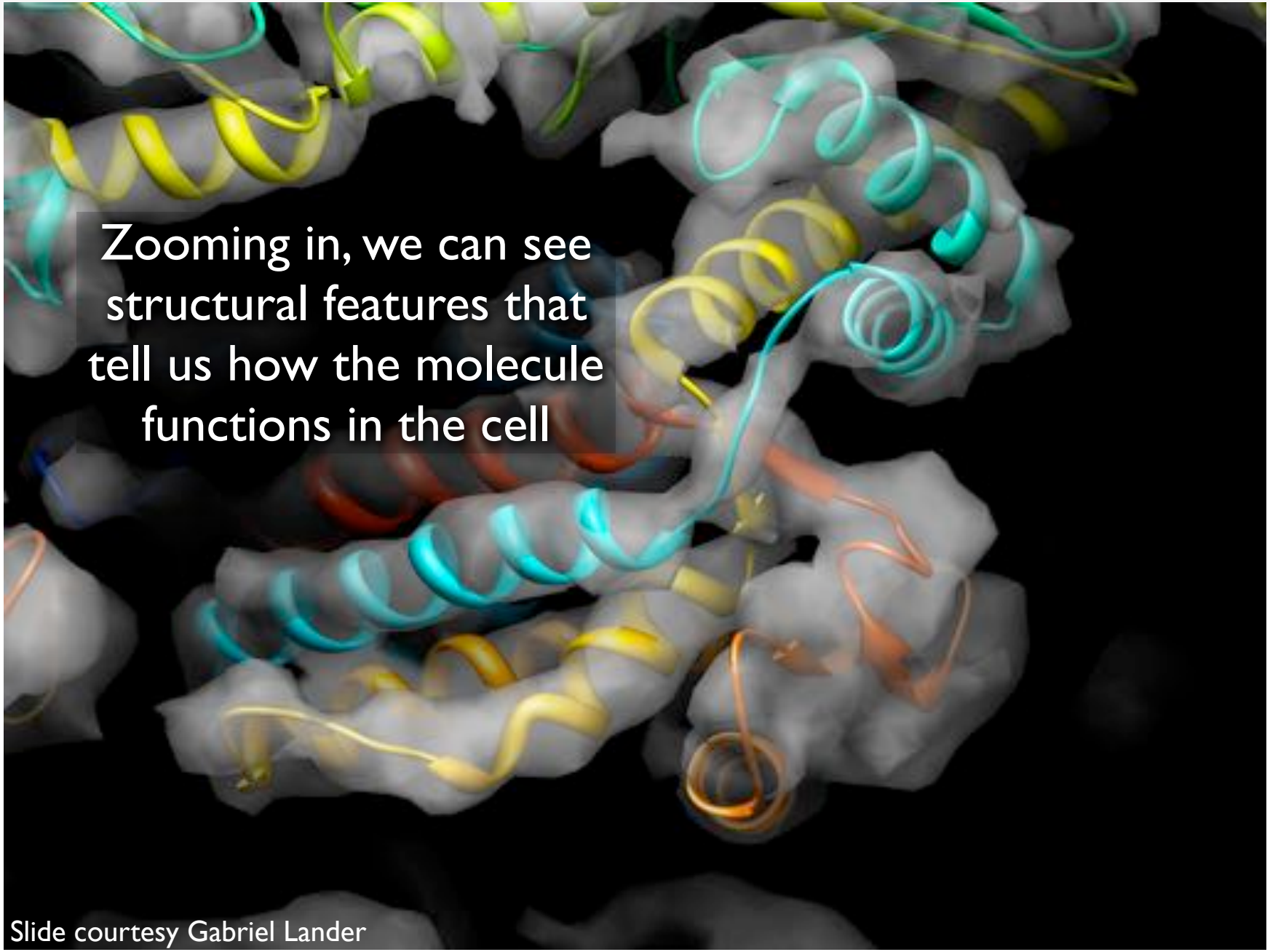
Slide courtesy Gabriel Lander

Then we gather all these views and combine them computationally

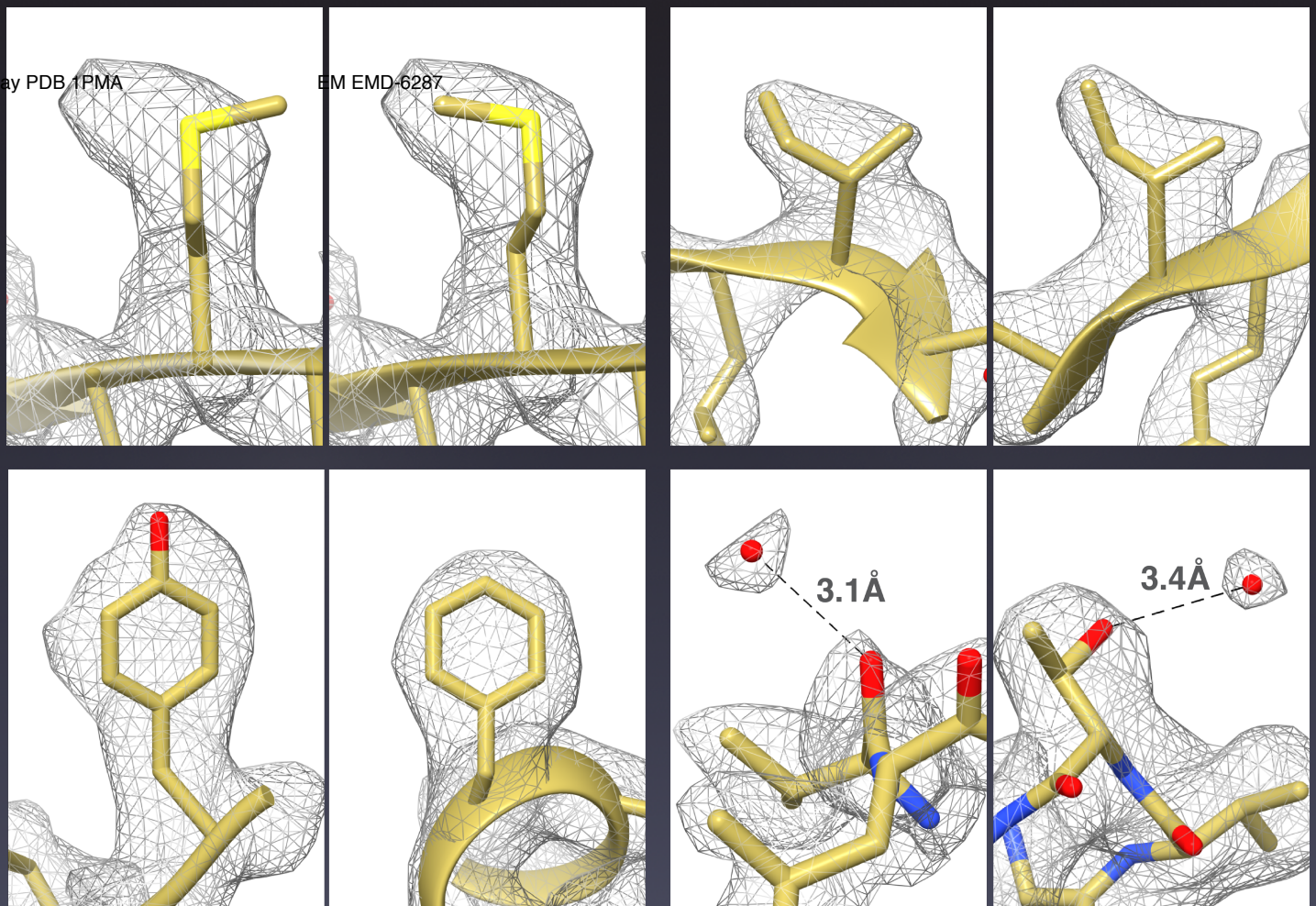


This provides us with a 3D reconstruction of our molecule

Slide courtesy Gabriel Lander


Zooming in, we can see structural features that tell us how the molecule functions in the cell
Slide courtesy Gabriel Lander

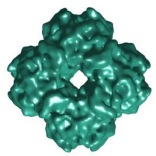
2.8 Å resolution reconstruction of 20 S proteasome



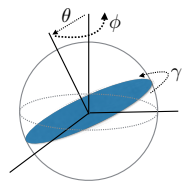
- Melody Campbell, David Veesler, Anchi Cheng, Bridget Carragher, and Clinton S. Potter (2015). 2.8 Å resolution reconstruction of the Thermoplasma acidophilum 20 S proteasome using cryo-electron microscopy. eLife.

Idealized CryoEM Setting

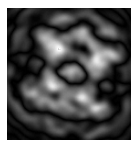
- $f(x, y, z)$ describes the structure of the unknown object



- (θ, ϕ) specify the direction of projection
- γ is in-plane rotation of projection



- $P_\alpha(f)$ is an operator that takes projection of $f(x, y, z)$ with Euler angles $\alpha = (\theta, \phi, \gamma)$

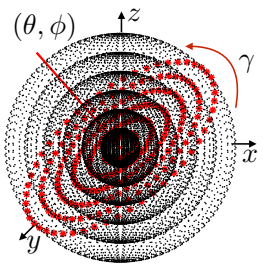


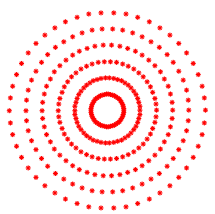
Problem

- Given a set of projections of the object from unknown angles α, reconstruct the function $f(x, y, z)$
- Challenges:
  - Highly non-convex problem with noisy data
  - Reconstruction is slow
  - Requires initialization and refinement
  - Validation is not robust

Reconstruct in Fourier Space

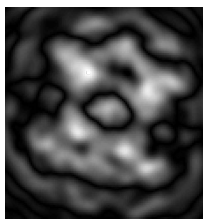
- Function $f(x, y, z)$ describes unknown object
- Fourier transform in spherical coordinates is $\hat{f}(k, \theta, \phi)$
- Let $S_\alpha(\hat{f})$ be an operator that takes a slice through $\hat{f}$ with normal (θ, ϕ) and rotates the slice by angle γ
- Slice $s_\alpha = S_\alpha(\hat{f})$ is Fourier transform of projection $P_\alpha(f)$ of function $f(x, y, z)$
- $s_\alpha(k, \xi)$ is sampled on a polar grid (k, ξ)



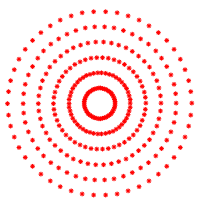


Data

- Data is a set of reference images $R = \{r_1, ..., r_{N_r}\}$



- $\hat{R} = \{\hat{r}_1, ..., \hat{r}_{N_r}\}$ is (2D) Fourier transform of reference images, sampled on a polar grid (k, ξ)



Objective Function

If the angles $\mathcal{A} = \{\alpha_1, .., \alpha_{N_r}\}$ of the reference images were given, $\hat{f}$ could be recovered by a least squares problem $L(\hat{R}, \mathcal{A})$

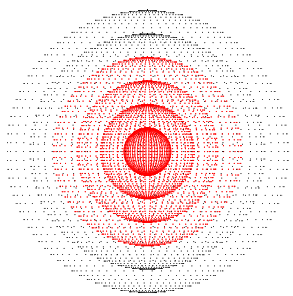
$$\hat{f}^* = L(\hat{R}, \mathcal{A}) = \arg\min_{\hat{f}} \sum_{j=1}^{N_r} \|S_{\alpha_j}(\hat{f}) - \hat{r}_j\|_2^2$$

Since the angles $\mathcal{A}$ are unknown, we need to solve the nonlinear problem:

$$\{\hat{f}^*, \mathcal{A}^*\} = \arg\min_{\{\hat{f}, \mathcal{A}\}} \sum_{j=1}^{N_r} \|S_{\alpha_j}(\hat{f}) - \hat{r}_j\|_2^2$$

Objective Function Restricted to Frequency k

- Let $\hat{R}([0,k])$ denote reference images restricted to ball up to frequency k
- Let $\hat{f}([0,k])$ denote the reconstruction up to frequency k using only the data $\hat{R}([0,k])$
- Let $\mathcal{A} = \{\alpha_1, .., \alpha_{N_r}\}$ denote the angles of the reference images



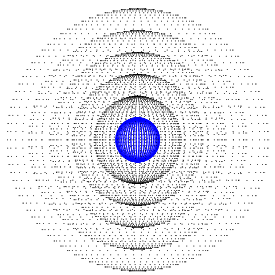
$$\{\hat{f}^*([0,k]), \mathcal{A}^*\} = \underset{\{\hat{f}([0,k]), \mathcal{A}\}}{\arg\min} \sum_{j=1}^{N_r} \|S_{\alpha_j}(\hat{f}([0,k])) - \hat{r}_j([0,k])\|_2^2$$

Marching In Frequency

Given $\hat{f}([0,k])$ restricted to frequency k

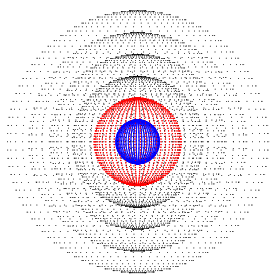
- Estimate the angles of reference images

$$\alpha_j = \arg\min_{\mathcal{A}} \sum_{j=0}^{N_r} \|S_{\alpha_j}\hat{f}([0,k]) - \hat{r}_j([0,k])\|^2, j = \{1..N_r\}$$



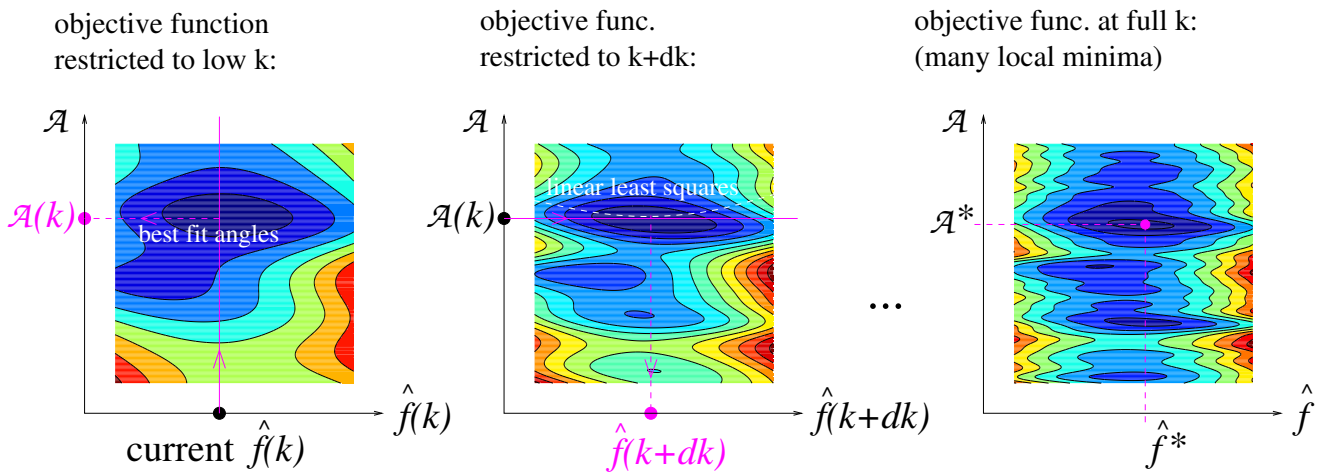
- Reconstruct $\hat{f}$ restricted to frequency $k + \delta k$

$$\hat{f}([0,k+\delta k]) = L(\hat{R}([0,k+\delta k]), \mathcal{A})$$



- Repeat

Intuition for Frequency Marching



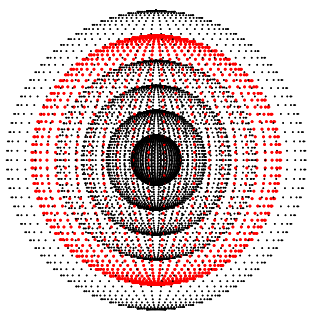
- At low frequency, the problem is ill-posed but nearly convex $\Rightarrow$ recover only a few parameters
- "Likely" that global minimum at $k+\delta k$ can be reached by coordinate descent starting from global minimum at k
- Reasonable to assume that, as k increases continuously, there is a smooth path from global minimum at k_{min} to global minimum at k_{max}

Implementation: Model Representation

- Discretize $\hat{f}(k, \theta, \phi)$ using K nodes in each dimension
- On each spherical shell at radius k, represent $\hat{f}(k, \theta, \phi)$ in the form

$$\hat{f}(k, \theta, \phi) = \sum_{n=0}^{p(k)} \sum_{m=-n}^{n} f_{nm}(k) Y_n^m(\theta, \phi) \,,$$

$$Y_n^m(\theta, \phi) = P_n^{|m|}(\cos\theta) e^{im\phi}$$



  - $P_n^{|m|}$ are associated Legendre functions of degree n and order m
  - $p(k) = k + 2$ is the degree of the spherical harmonic expansion on the sphere of radius k

Implementation : Fast algorithms

Operation counts for K spheres and N_r reference images:

- Find angles $\mathcal{A}$ by matching reference images $\hat{R}$ to projections of the current model
  - naive: $O(N_r K^5)$ optimized: $O(N_r K^3)$
- Reconstruct new model at current frequency $k+\delta k$ by solving the least squares problem $\hat{f}([0, k+\delta k]) = L(\hat{R}([0, k+\delta k]), \mathcal{A})$.
  - naive: $O(N_r K^6)$ optimized: $O(N_r K^2 + K^3)$.

Least Squares

For each sphere at radius k, we have $M = N_r K$ point values, denoted by b_l and wish to solve

$$\{f^*_{nm}(k)\} = \arg\min_{\{f_{nm}\}} \sum_{l=1}^{M} \|(\sum_{m,n} f_{nm}(k) Y_n^m(\theta_l, \phi_l)) - b_l\|_2^2.$$

This is a linear least squares problem of the form $T\mathbf{f} = \mathbf{b}$, where $\mathbf{f} = \{f_{nm}\}$, $\mathbf{b} = \{b_l\}$ and $T \in \mathbf{C}^{M \times K^2}$.

We will solve this system using conjugate gradient on the normal equations.

Note that a naive matrix-vector product requires $O(MK^4)$ work.

Fast Application of T

1. $T\mathbf{f}$ can be evaluated on a regular spherical grid with K^2 points using separation of variables in $O(K^3)$ operations.

2. Given values on a regular grid, $T\mathbf{f}$ can be evaluated at any target point with qth order accuracy using q^2 interpolation nodes.

3. The cost of applying T (or its adjoint) to a vector is reduced to $O(Mq^2 + K^3)$.

4. With sufficient sampling on the sphere, the least squares problem is well-conditioned.

Time

- M - number of points
- K - order of spherical harmonic expansion

M	K^2	Time	rel err	# CG iters
$100,000$	900	$2s.$	$3\ 10^{-6}$	8
$10,000,000$	900	$160s.$	$3\ 10^{-7}$	4
$10,000,000$	10000	$162s.$	$7\ 10^{-8}$	5

Interpolation order set to 7 and the CG tolerance set to 10^{-7}.

Experiments with Simulated Data

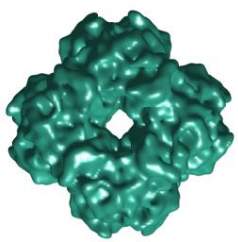
PDB contains coordinates of atoms and types of atoms from X-ray crystallography structures

1. Represent atoms as Gaussians with variance = (atomic radius)/2
2. Blur to 2 Angstrom resolution by convolving potential with Gaussian
3. Generate 50,000 projections with random Euler angles
4. Convolve with CTF
5. Add uniform noise

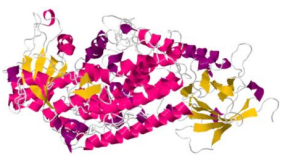
All experiments were run on 14 cores

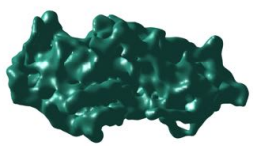
Molecules

Rubisco

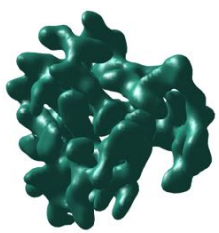


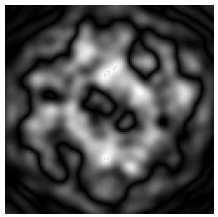
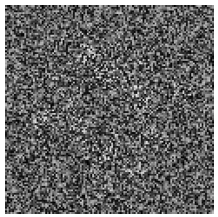
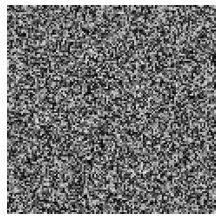
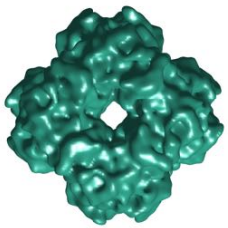
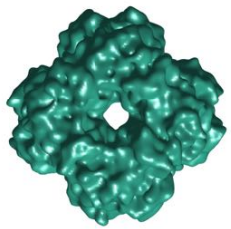
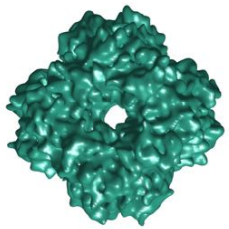
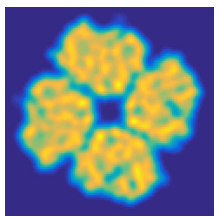
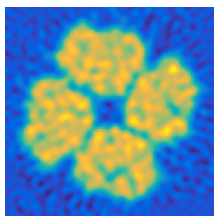
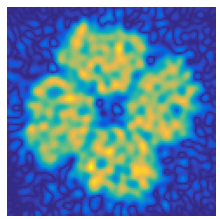
Lipoxygenase-1

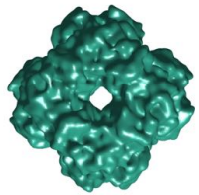
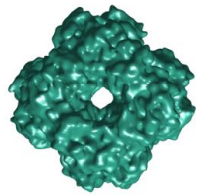
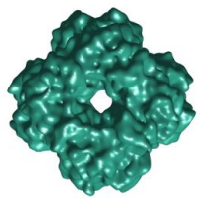
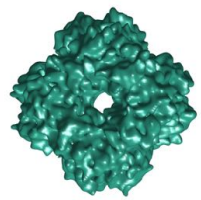


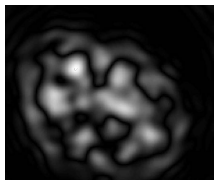
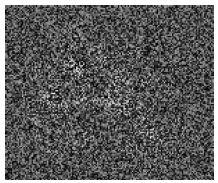
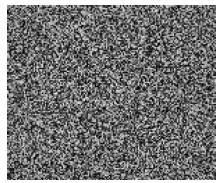
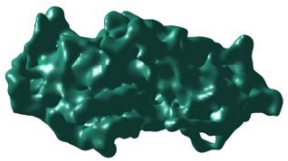
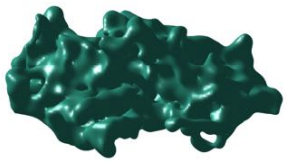
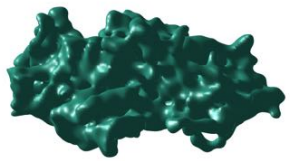
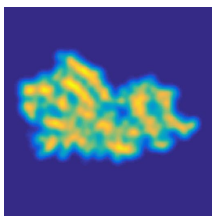
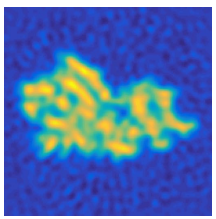
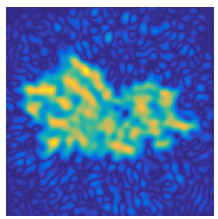


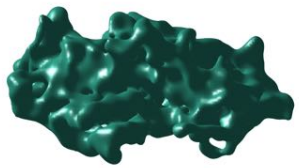
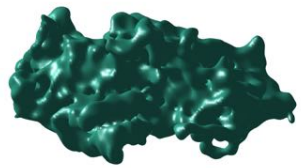
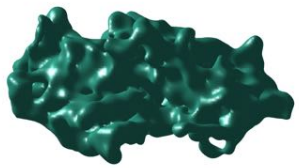
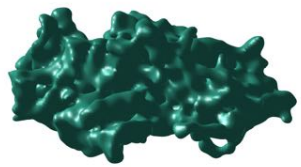
Neurotoxin

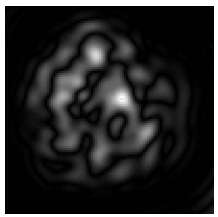
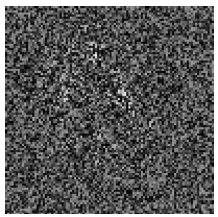
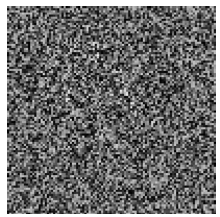
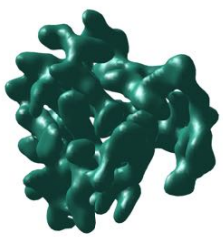
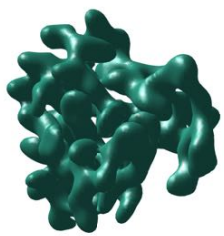
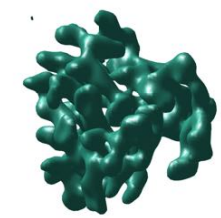
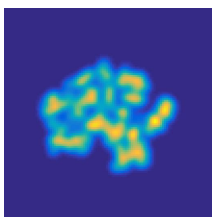
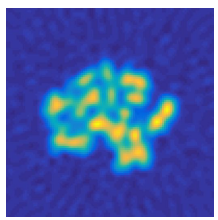
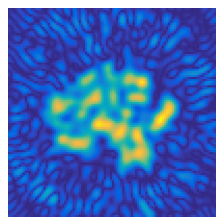


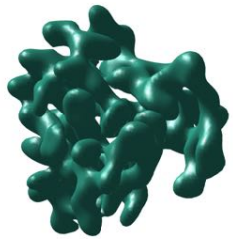
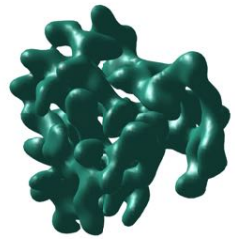
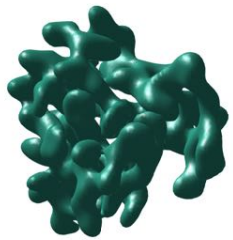
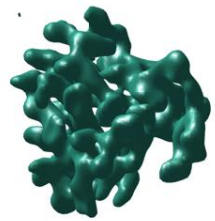
	Original	SNR 0.3	SNR 0.1
Reference images			
Recon-struction			
Slice			
Time		1.65 hours	1.7 hours

	SNR 0.3	SNR 0.1
Least squares with known angles		
$L2$ norm of difference from original	0.09	0.17
Marching in frequency		
$L2$ norm of difference from original	0.12	0.31

	Original	SNR 0.3	SNR 0.1
Reference images			
Recon-structions			
Slice			
Time		1.56 hours	1.76 hours

	SNR 0.3	SNR 0.1
Least squares with known angles		
$L2$ norm of difference from original	0.10	0.26
Marching in frequency		
$L2$ norm of difference from original	0.11	0.38

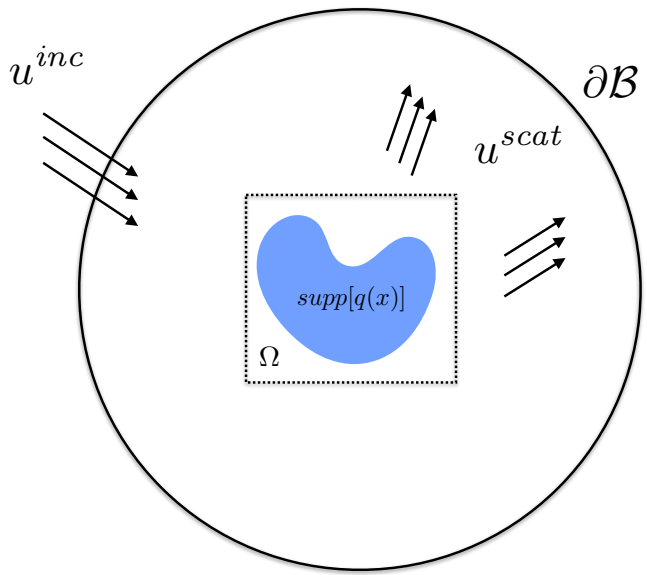
	original	SNR 0.3	SNR 0.1
Reference images			
Recon-structions			
Slice			
Time		1.1 hours	1.25 hours

	SNR 0.3	SNR 0.1
Least squares with known angles		
$L2$ norm of difference from original	0.08	0.14
Marching in Frequency		
$L2$ norm of difference from original	0.09	0.41

Part II: *Inverse Scattering*

Inverse scattering

Reconstruct the constrast function in a penetrable medium from measurements of the scattered field induced by time harmonic plane waves with varying angles of incidence and frequencies.



$$\Delta u^{scat}(x) + k^2(1 - q(x))u^{scat}(x) = k^2 q(x) u^{inc}(x)$$

Joint work with Carlos Borges, Adrianna Gillman

Inverse scattering

- The problem as stated is nonlinear and ill-posed.
- To deal with the nonlinearity, we apply Newton's method, linearizing the problem about the current guess for the unknown material properties. Iteration involves solving a large number of *forward* scattering problems. Thus, fast and accurate Helmholtz solvers are essential.
- To deal with the ill-posedness of this problem, we use the data from multiple directions and seek a successively more resolved solution as we increase the frequency of the interrogating waves, using Chen's method of recursive linearization.

The single frequency problem

- Suppose that, for a fixed frequency k, a series of experiments is carried out, with M distinct plane waves impinging on a domain Ω which contains the support of an unknown contrast function $q(\boldsymbol{x})$.

- Let the incident directions be denoted by $\{\boldsymbol{d}_m, m = 1, \ldots, M\}$.

- The *single frequency inverse scattering problem* consists of determining $q(\boldsymbol{x})$ from $\{u^{far}_{k,\boldsymbol{d}_m}(\theta), m = 1, \ldots, M\}$.

The single frequency problem

- In the far field, no more than $O(k)$ independent measurements can reasonably be made on $\partial\mathcal{B}$ (Heisenberg).
- In physical terms, Fourier modes on $\partial\mathcal{B}$ whose frequency exceeds k correspond to rapidly decaying fields. Acquiring such data would impose exponential accuracy requirements.
- For similar reasons, only $O(k)$ directions of incidence are useful, leading to a total of $O(k^2)$ independent measurements.
- Thus, in two dimensions, the single frequency inverse problem is at the limits of feasibility in seeking to reconstruct a model for $q(\boldsymbol{x})$ with $O(k^2)$ unknowns.

The multi-frequency problem

- Suppose now that we probe the unknown function $q(\boldsymbol{x})$ at a set of frequencies $\{k_j,\, j = 1, \ldots, Q\}$, with incident directions at each frequency k_j denoted by $\{\boldsymbol{d}_{j,m},\, m = 1, \ldots, M_j\}$.

- The *multi-frequency inverse scattering problem* consists of determining $q(\boldsymbol{x})$ from $\{u^{far}_{k_j,\boldsymbol{d}_{j,m}}(\theta);\, j = 1, \ldots, Q,\, m = 1, \ldots, M_j\}$.

Recursive linearization

- The recursive linearization method (Yu Chen, 1995) overcomes the non-convexity of the problem by "continuation" in frequency.

- One solves a sequence of *single-frequency* inverse problems for higher and higher values of k, using the approximation of $q(\boldsymbol{x})$ obtained at the preceding frequency as an initial guess.

- For sufficiently small steps in k, one can use only one Newton iteration - hence, the name.

Linearization

The solution of the direct scattering problem with a single incident plane wave $u_j^{inc}(x) = \exp(ikx \cdot d_j)$ defines an operator $F_j : q \to u^{far}$, where u^{far} is measured on the boundary ∂B. Given a set of M incident directions, the forward model is

$$\mathbf{F}(q) = \mathbf{u^s}(x)|_{\partial B}$$

where $\mathbf{u^s}(x)|_{\partial B}$ denotes the vector of scattered data from all incident angles.
Newton's method for the inverse problem is based on writing $q = q_0 + \delta q$ and

$$\mathbf{F}(q_0) + \mathbf{J_q}(\delta q) = \mathbf{u^s}(x)|_{\partial B}$$

where q_0 is known, $\mathbf{J_{q_0}}$ is the Jacobian (Fréchet derivative of $\mathbf{F}$) at q_0 and δq is to be determined.

Linearization

The Newton iteration proceeds as follows:

- Given an approximation $q^{(i)}$ of the domain, solve the equation
$$\mathbf{J}_{q^{(i)}} \delta q = \mathbf{u^s}|_{\partial B} - \mathbf{F}(q^{(i)}).$$
- Compute the update $q^{(i+1)} = q^{(i)} + \delta q$.
- Stop when the residual tolerance has been achieved.
- The system (**??**) is overdetermined and solved by LSQR or CGN:
$$\mathbf{J}^*_{q^{(i)}} \mathbf{J}_{q^{(i)}} \delta q = \mathbf{J}^*_{q^{(i)}} \left(\mathbf{u^s}|_{\partial B} - \mathbf{F}(q^{(i)}) \right),$$
where $\mathbf{J}^*_{\mathbf{q^{(i)}}}$ is the adjoint of $\mathbf{J}_{\mathbf{q^{(i)}}}$.

Applying $\mathbf{J}_q$ and $\mathbf{J}_q^*$

Both $\mathbf{J}_q$ and $\mathbf{J}_q^*$ can be applied by solving the forward scattering problem M times.
Modern fast solvers require $O(N^{3/2})$ work to factor the system matrix, where N is the number of grid points. The solution time for each subsequent right-hand side is of the order $O(N)$. [1]
Thus, the total cost for a single frequency is

$$O(N_{newton} N^{3/2}) + O(N_{newton}\,(2N_{iter} + 1)\,M\,N),$$

Since $N = O(k^2)$ and $M = O(k)$, assuming N_{newton} and N_{iter} are $O(1)$, the cost scales like

$$O(k^3).$$

[1] A. Gillman, A. Barnett, and P. Martinsson, A spectrally accurate direct solution technique for frequency- domain scattering problems with variable media, BIT Numerical Mathematics, 2014

Fast, direct solvers

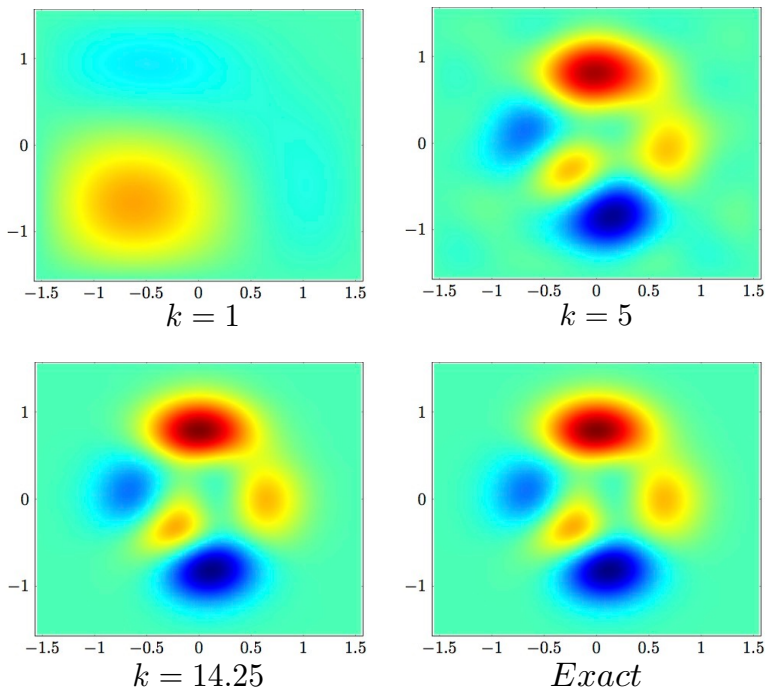
- A number of methods have been developed requiring $\mathcal{O}(N^{3/2})$ work in 2D: Chen, Chew, Michielssen, Boag, Gillman, Young, Martinsson, Rokhlin, G–, Ho, Engquist, Ying, Zepeda-Núñez, Demanet, Xia, Chandrasekaran, Gu, Pals, Hackbusch, Börm, Sauter, Bebendorf, Kapur, Long, Goreinov, Tyrtyshnikov, Zamarashkin, Gope, Jandhyala, ... apologies to those omitted
- More recently, $\mathcal{O}(N \log N)$ variants have been developed: Corona, Martinsson, Zorin, Ambikasaran and Darve, Ho and Ying, etc.
- Still a very active area with many open problems, including extension to 3D, reduction of constants implicit in $O(N \log N)$ notation, memory requirements, etc.

Recursive Linearization

Given measurements $\mathbf{u}^{far}_{k_j, \boldsymbol{d}_{j,m}}$, for $j = 1, \ldots, Q$, $m = 1, \ldots, M_j$:

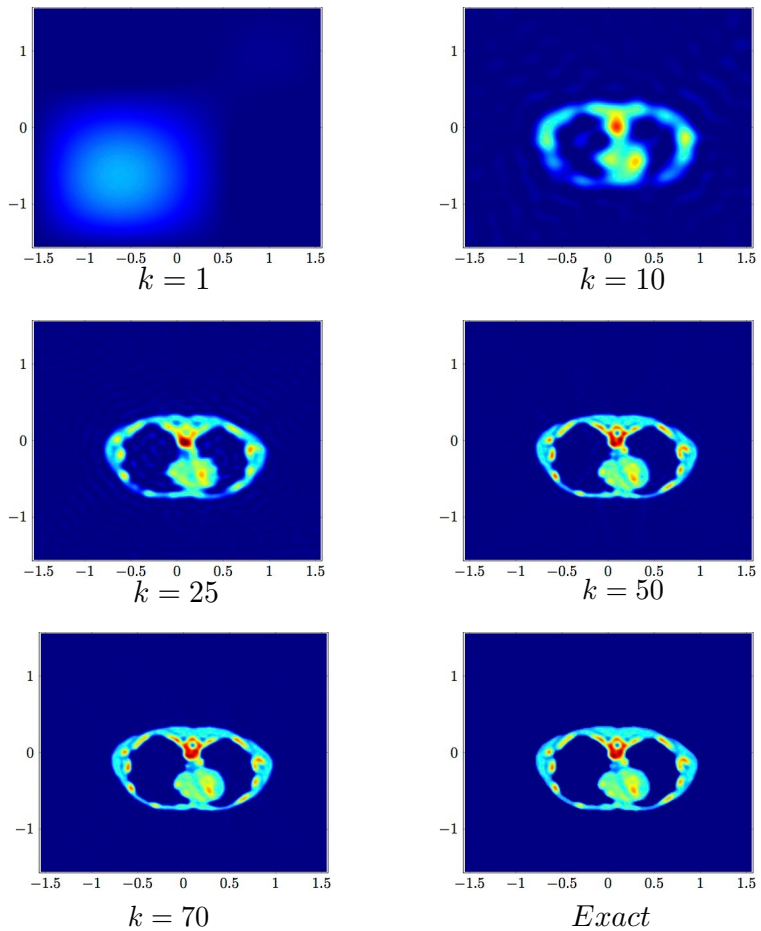
1. For $j = 1, \ldots, Q$:
   - 1.1 Use Newton's method with initial guess $q^{(0)}_{k_j}$ and far field $\mathbf{u}^{far}_{k_j}$. Let the result obtained after i iterations by $q^{(i)}_{k_j}$.
   - 1.2 Set $q^{(0)}_{k_{j+1}} = q^{(i)}_{k_j}$.
   - 1.3 Increase the number of degrees of freedom (Fourier modes) in the representation for $q_{k_{j+1}}$.

Simple cross-section



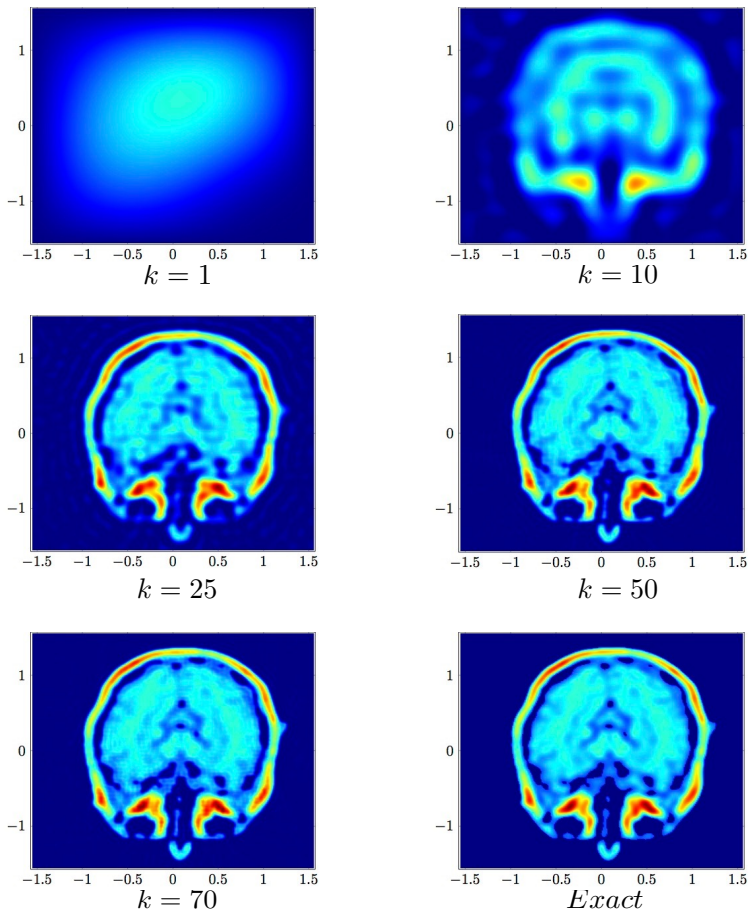
(a) Simple

Thorax-like cross-section



(b) Thorax

Head-like cross-section



(c) Head

Conclusions

- There are tremendous opportunities for discovery in the biological and biomedical sciences which require significant advances in scientific computing
- A realization of the last few years: *our ability to cope with the data influx is falling behind*
- The creation of tools that allow us to process vast data sets reliably is a central challenge.
- Recent progress in optimization and fast algorithms can have a significant impact on modern imaging technology.