Data-Compatible T-Matrix Completion (DCTMC)

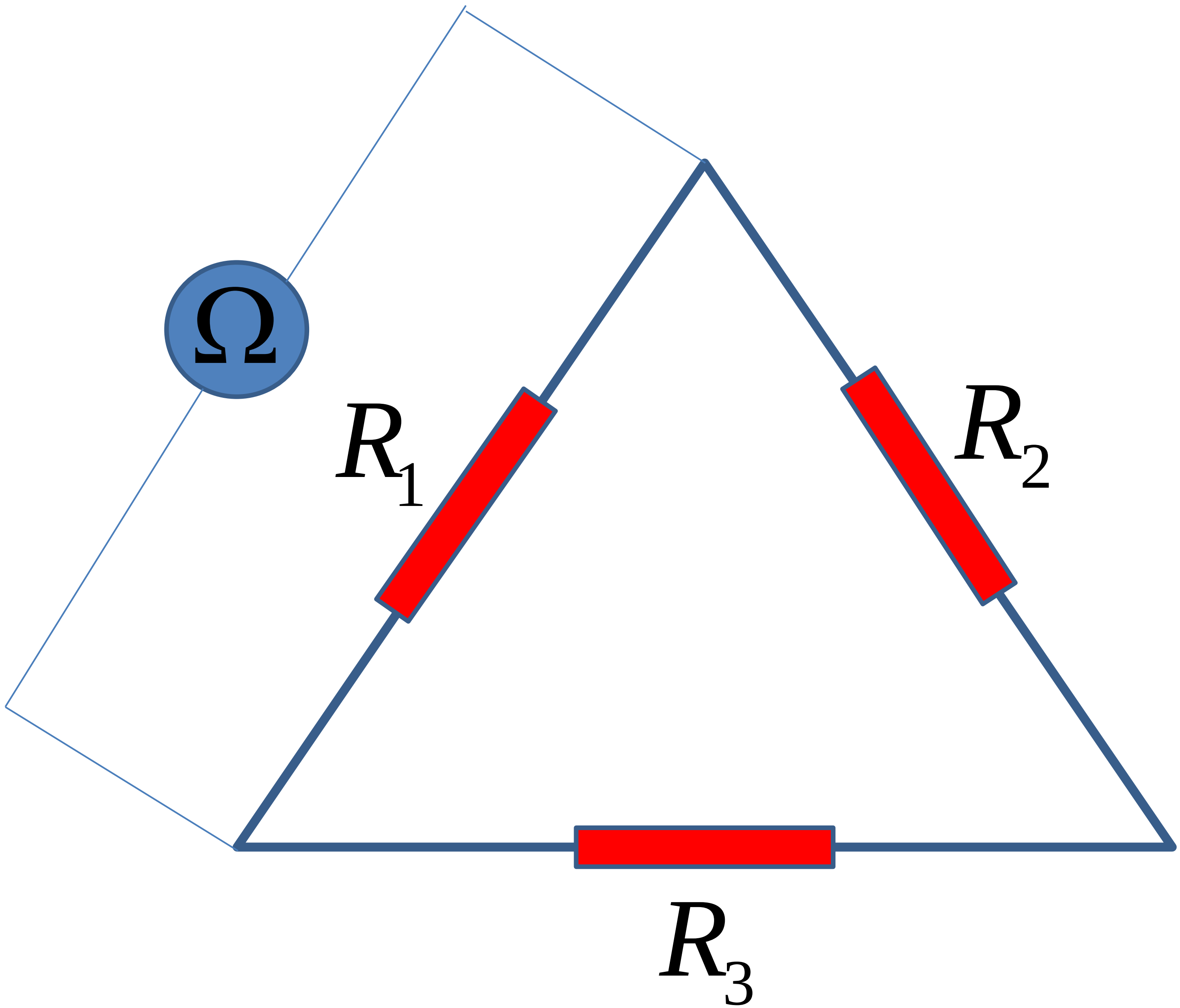
and a few examples of exactly-solvable nonlinear inverse toy problems

Vadim A Markel (Institut Fresnel, on leave from UPenn)

joint work with

Howard W. Levinson (formerly at UPenn, currently at UMich Math)

Three resistor problem



The nonlinear system of equations to determine the unknown resistors R_1, R_2, R_3 from the measurements of resistance between all pairs of vertices, $Z_1,\ Z_2,\ Z_3$

$$\frac{1}{Z_1} = \frac{1}{R_1} + \frac{1}{R_2 + R_3}$$

$$\frac{1}{Z_2} = \frac{1}{R_2} + \frac{1}{R_1 + R_3}$$

$$\frac{1}{Z_3} = \frac{1}{R_3} + \frac{1}{R_1 + R_2}$$

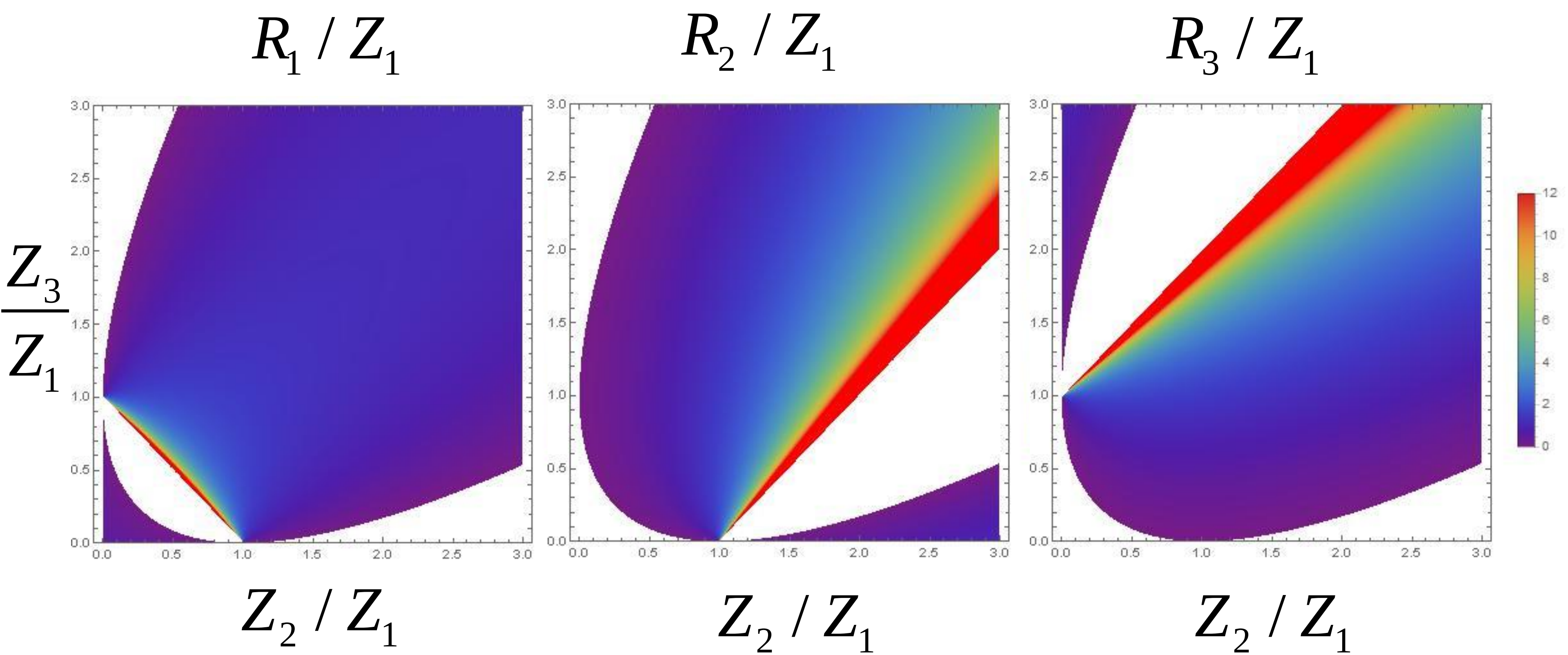
Inverse solution:

$$R_1 = Z_1 - \frac{1}{2}\frac{Z_1^2 - \left(Z_2 - Z_3\right)^2}{Z_1 - (Z_2 + Z_3)}$$

$$R_2 = Z_2 - \frac{1}{2}\frac{Z_2^2 - \left(Z_1 - Z_3\right)^2}{Z_2 - (Z_1 + Z_3)}$$

$$R_3 = Z_3 - \frac{1}{2}\frac{Z_3^2 - \left(Z_1 - Z_2\right)^2}{Z_3 - (Z_1 + Z_2)}$$

Density plots of solutions as functions of Z_2 / Z_1 and Z_3 / Z_1



Linearization:

Let $R_n = R(1+\xi_n)$ where $|\xi_n| << 1$ and R is known

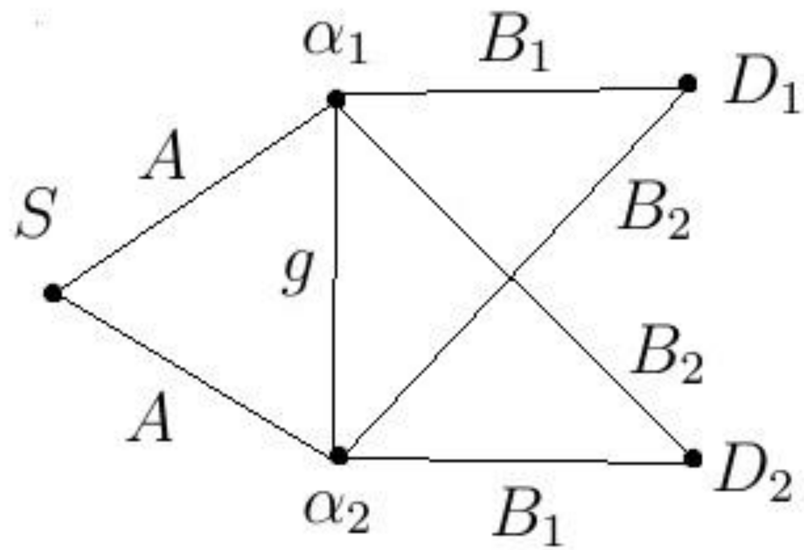
Let $\phi_n = \dfrac{R}{Z_n} - \dfrac{3}{2}$ $\leftarrow$ new definition of the datapoint

Then

$$\begin{pmatrix} 1 & 1/4 & 1/4 \\ 1/4 & 1 & 1/4 \\ 1/4 & 1/4 & 1 \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix}$$

This is a well-posed linear problem (matrix is invertible).

Simplest inverse scattering problem: Two scatterers



$$\begin{cases} d_1 = \alpha_1(A + g d_2) \\ d_2 = \alpha_2(A + g d_1) \end{cases}$$

$$E(D_1) = B_1 d_1 + B_2 d_2$$
$$E(D_2) = B_2 d_1 + B_1 d_2$$

Let $\phi_1 = \dfrac{E(D_1)}{AB_1}$, $\phi_2 = \dfrac{E(D_2)}{AB_1}$ and $\dfrac{B_2}{B_1} = \beta \neq 1$

Then the nonlinear equations are:

$$\begin{cases} \phi_1 = \dfrac{\alpha_1(1+g\alpha_2)+\beta\alpha_2(1+g\alpha_1)}{1-g^2\alpha_1\alpha_2} \\ \phi_2 = \dfrac{\alpha_2(1+g\alpha_1)+\beta\alpha_1(1+g\alpha_2)}{1-g^2\alpha_1\alpha_2} \end{cases}$$

If we multiply both sides by $1-g^2\alpha_1\alpha_2$, the system will have a spurrious solution $\alpha_1 = \alpha_2 = -1/g$

Nonlinear inverse solutions:

$$\begin{cases} \alpha_1^{\text{inv}} = \dfrac{\beta\phi_2 - \phi_1}{\beta^2 - 1 + g(\beta\phi_1 - \phi_2)} \\[2ex] \alpha_2^{\text{inv}} = \dfrac{\beta\phi_1 - \phi_2}{\beta^2 - 1 + g(\beta\phi_2 - \phi_1)} \end{cases}$$

What if β=1 ?

We will obtain the spurrious solution $\alpha_1^{\text{inv}} = \alpha_2^{\text{inv}} = -\dfrac{1}{g}$

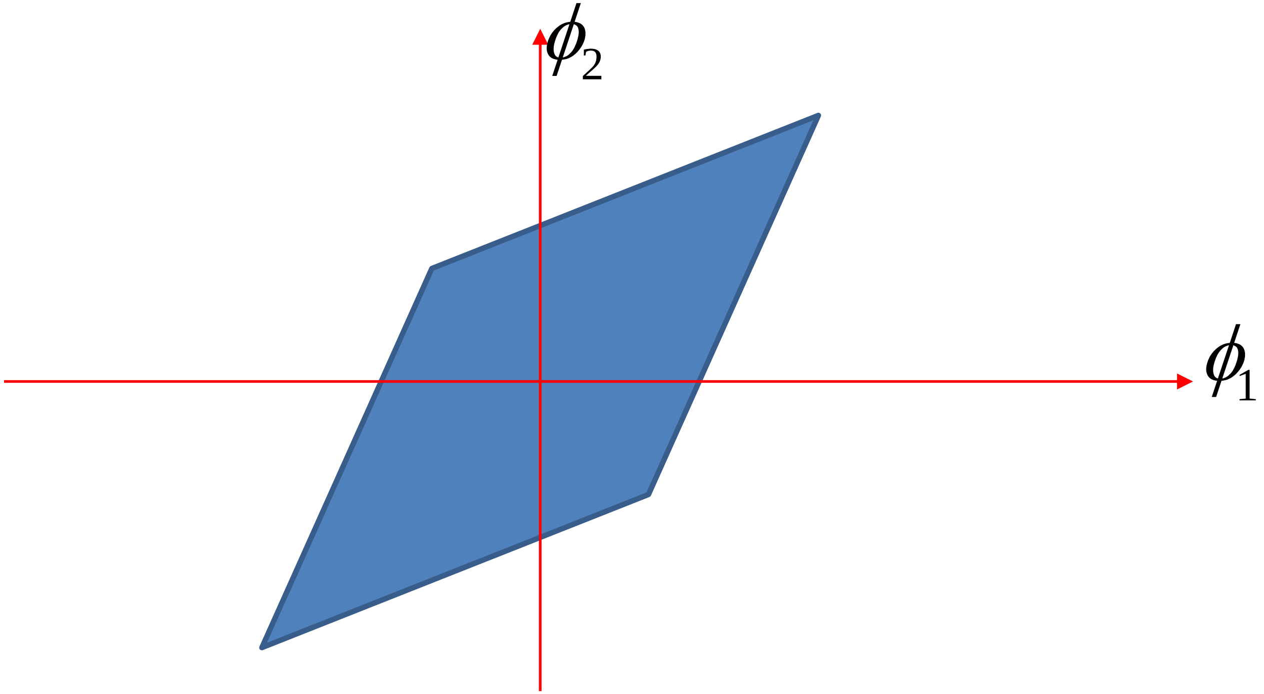
(In reality a d.p. $\phi_1 \neq \phi_2$ is unphysical in this case)

Linearized inverse solutions:

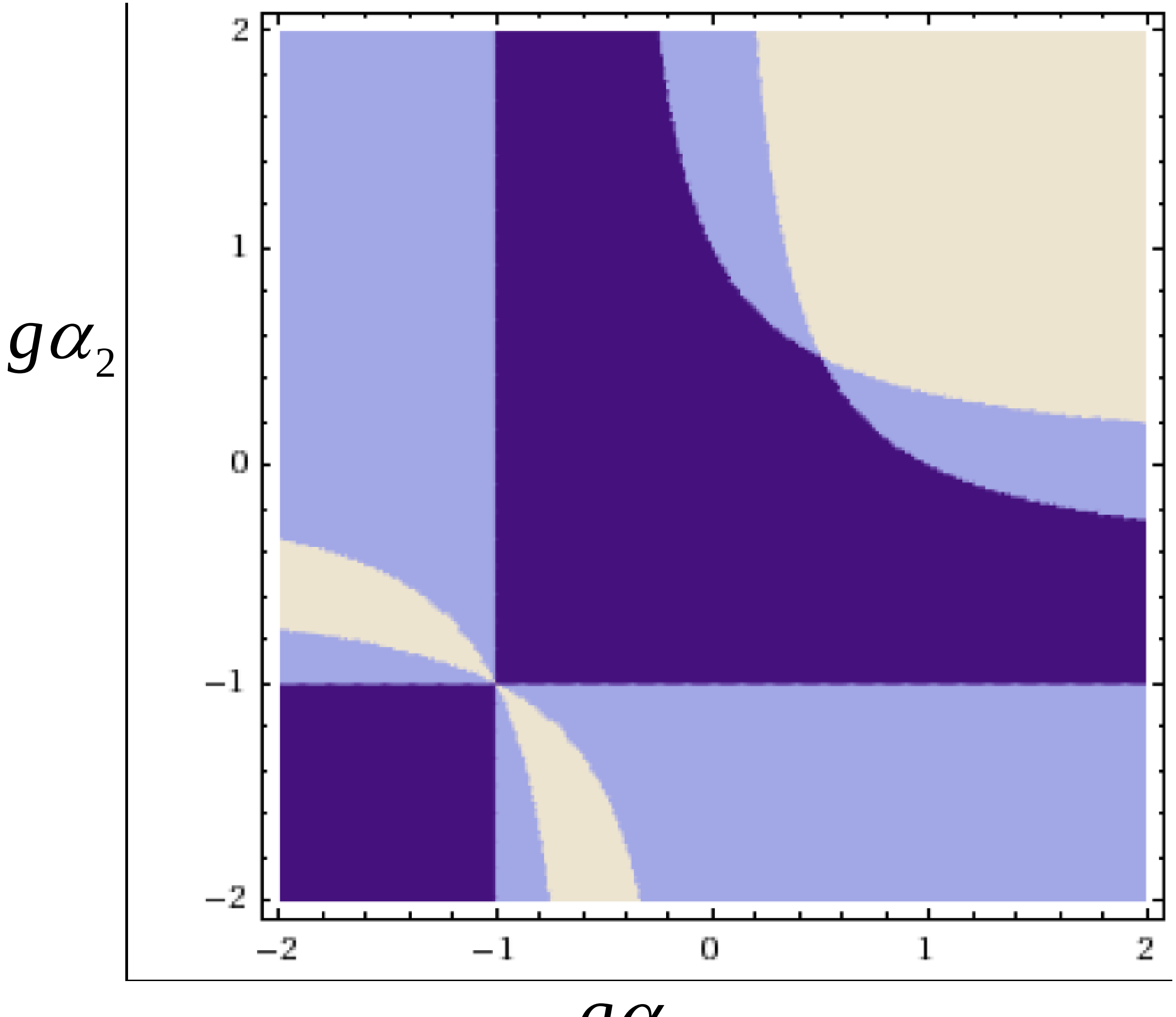
$$\begin{cases} \alpha_1^{\mathrm{lin}} = \dfrac{\beta\phi_2 - \phi_1}{\beta^2 - 1}, & \alpha_1^{\mathrm{inv}} = \dfrac{\alpha_1^{\mathrm{lin}}}{1 + g\alpha_2^{\mathrm{lin}}} \\[2ex] \alpha_2^{\mathrm{lin}} = \dfrac{\beta\phi_1 - \phi_2}{\beta^2 - 1}, & \alpha_2^{\mathrm{inv}} = \dfrac{\alpha_2^{\mathrm{lin}}}{1 + g\alpha_1^{\mathrm{lin}}} \end{cases}$$

Inverse Born series convergence condition (nec.&suff.):

$$\left|\frac{g}{\beta^2-1}\left(\beta\phi_1-\phi_2\right)\right|<1 \quad \text{AND} \quad \left|\frac{g}{\beta^2-1}\left(\beta\phi_2-\phi_1\right)\right|<1$$



Region of convergence of the inverse Born series for the model parameters



How is the T-matrix defined in this simple case?

$$\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \underbrace{\begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}}_{T} \begin{pmatrix} A \\ A \end{pmatrix}$$

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \frac{1}{AB_1} \begin{pmatrix} B_1 & B_2 \\ B_2 & B_1 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}$$

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} 1 & \beta \\ \beta & 1 \end{pmatrix} \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

What do we know about the T-matrix from the data?

$$\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} t_{11} + t_{12} + \beta(t_{21} + t_{22}) \\ \beta(t_{11} + t_{12}) + t_{21} + t_{22} \end{pmatrix}$$

$$t_{11} + t_{12} = \frac{\beta\phi_2 - \phi_1}{\beta^2 - 1} = \alpha_1^{\mathrm{lin}}$$

$$t_{21} + t_{22} = \frac{\beta\phi_1 - \phi_2}{\beta^2 - 1} = \alpha_2^{\mathrm{lin}}$$

Row-wise sums of the T-matrix elements are known from the data (and given by the linearized inversions).

Is there a general relationship between the T-matrix and the unknowns?

Define:

$$G = \begin{pmatrix} 0 & g \\ g & 0 \end{pmatrix} \qquad V = \begin{pmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{pmatrix}$$

Then

$$T = \left(I - VG\right)^{-1} V \quad = \quad V + VGV + VGVGV + \ldots\ldots$$

$$V = \left(I + TG\right)^{-1} T$$

There is a one-to-one correspondence between *T* and *V* (one matrix uniquely defines the other).

But not every *T* corresponds to a physically-meaningful (or even diagonal) *V*.

MAIN IDEA(S) OF DCTMC:

1. Use data to obtain information on the elements of the T-matrix.
   -- this is a linear problem
   -- a larger data set is factored into two much smaller data sets
   -- if we could find all elements of T, we would solve the nonlinear inverse problem immediately and exactly. But in most cases data do not allow this (requires internal measurements).

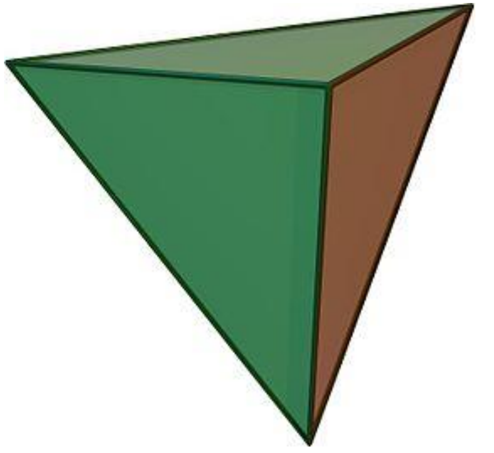
2. From the information obtain in Step 1, complete the T-matrix (find all its elements) given the constraint that the *V* is physically meaningful (e.g., diagonal).

3. It turns out that this approach helps solve even linear inverse problems with very large data sets. The factorization of a large data set into two much smaller sets works even in this case, and there are essentially no approximations.

Still a toy problem but now *N* particles....

$$d_n = \alpha_n \left[E_n + \sum_{m \neq n} g_{nm} d_m \right]$$

Let $g_{nm} = g$



Then

$$t_{nm} = \kappa_n \delta_{nm} + \frac{g}{1 - gS} \kappa_n \kappa_m$$

where $\kappa_n = \dfrac{\alpha_n}{1 + g\alpha_n}$ and $S = \sum_{n=1}^{N} \kappa_n$

We know that the functional $T[V]$ is invertible numerically…
… but in this case it is invertible analytically.

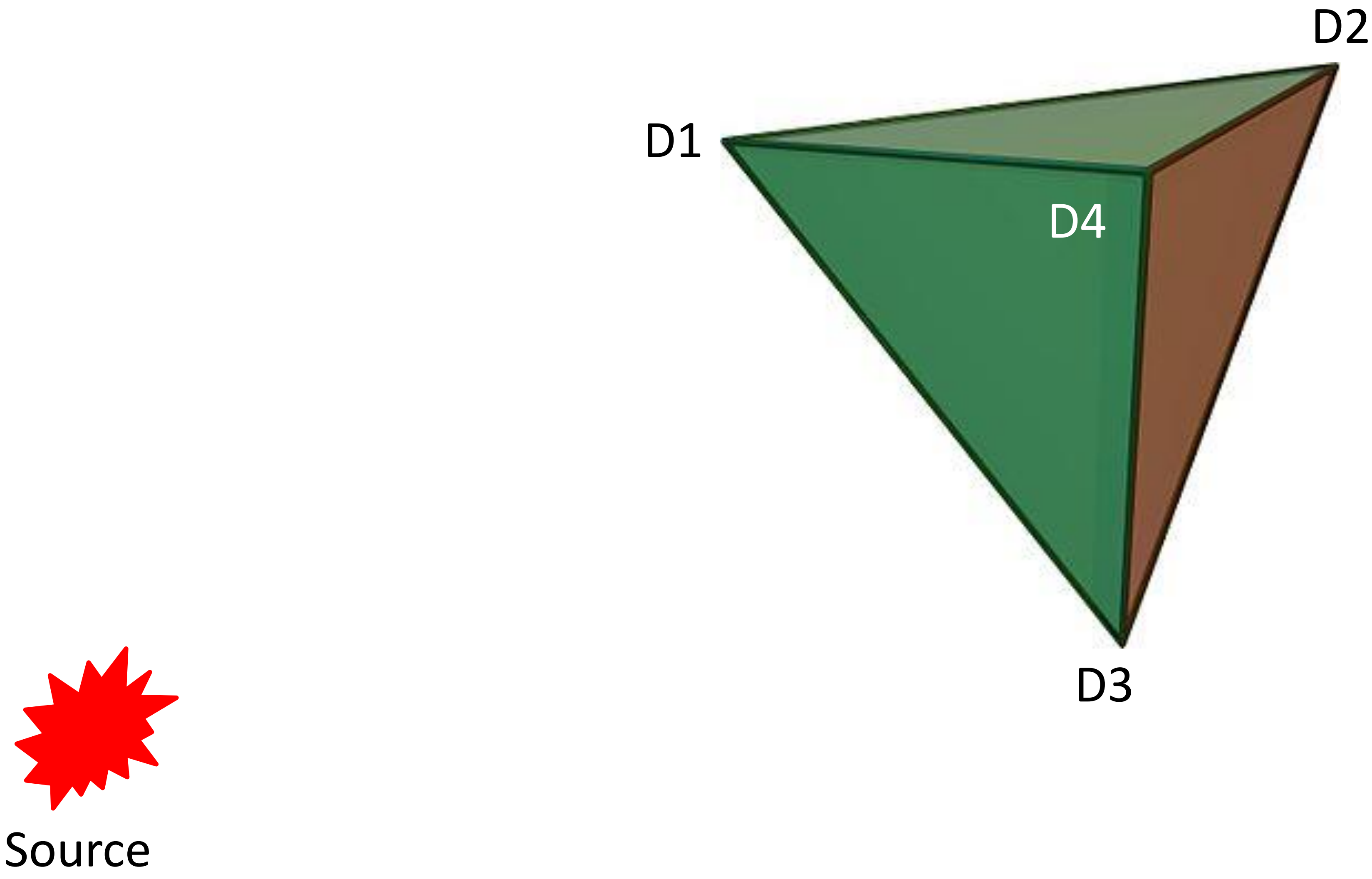
So we can gain some insight about WHAT we need to know about the T-matrix

Nonlinear inverse solution (assuming we know T):

$$\alpha_n^{\mathrm{inv}} = \frac{P_n}{1 + g\left[\sum_{m=1}^{N} P_m - P_n\right]} \qquad \text{where} \qquad P_n = \sum_{m=1}^{n} t_{nm}$$

It is sufficient to know column-wise (or row-wise) sums of the T-matrix to find V

Conjecture: it is sufficient to know ANY (or almost any) N linearly-independent combinations of the elements of the T-matrix


D2
D1
D4
D3
Source

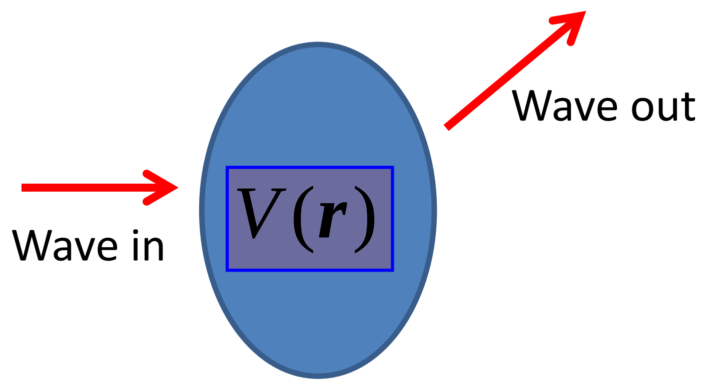
DCTMC motivation 1: Local minima

$$F[V] = \Phi$$

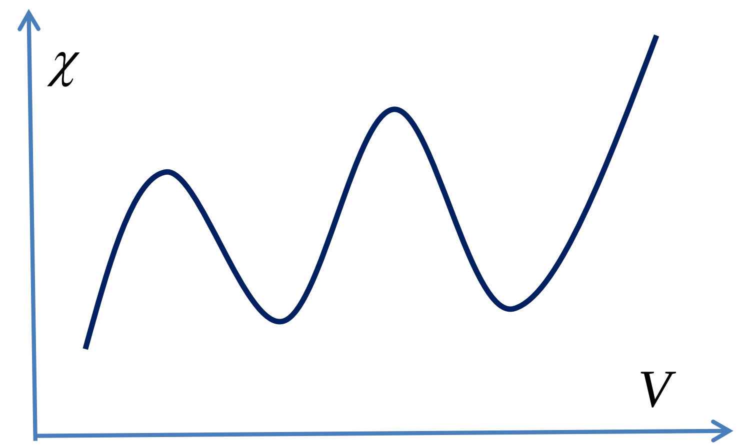
Theoretical model

Object

Data

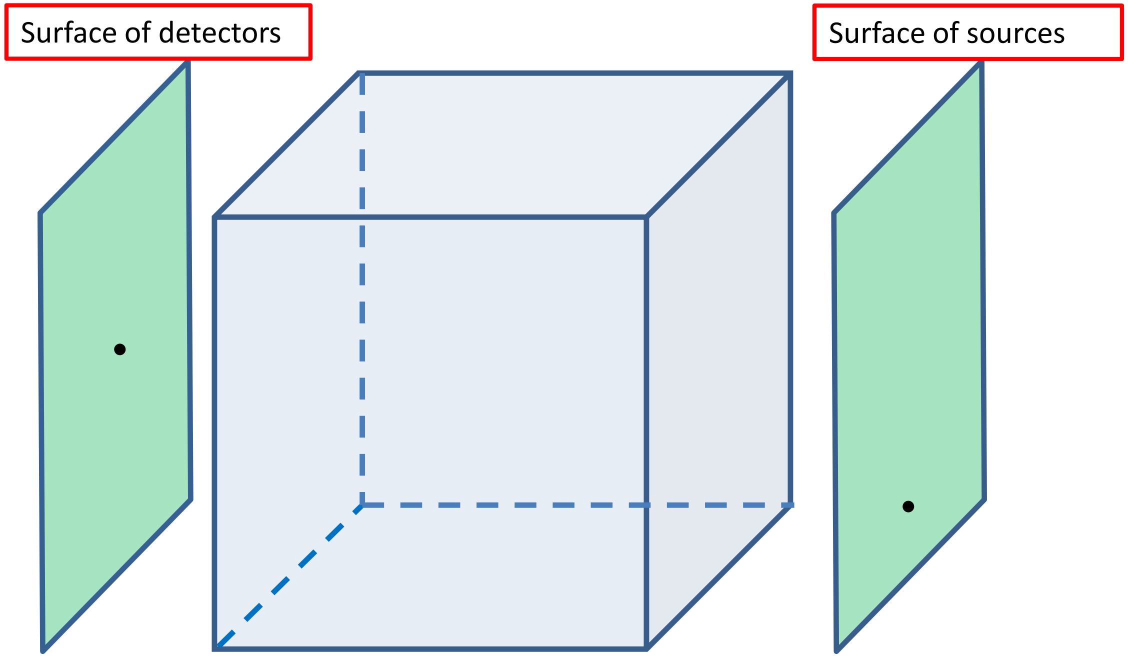


$$\chi = D\left[F[V] - \Phi\right] + \lambda^2 D\left[\Phi - \Phi_{\text{guess}}\right]$$



DCTMC motivation 2: Large data sets

$$\chi = D\left[F[V] - \Phi\right] + \lambda^2 D\left[\Phi - \Phi_{\text{guess}}\right]$$



$$N_s = L^2$$

$$N_d = L^2$$

$$N = N_s N_d = L^4$$

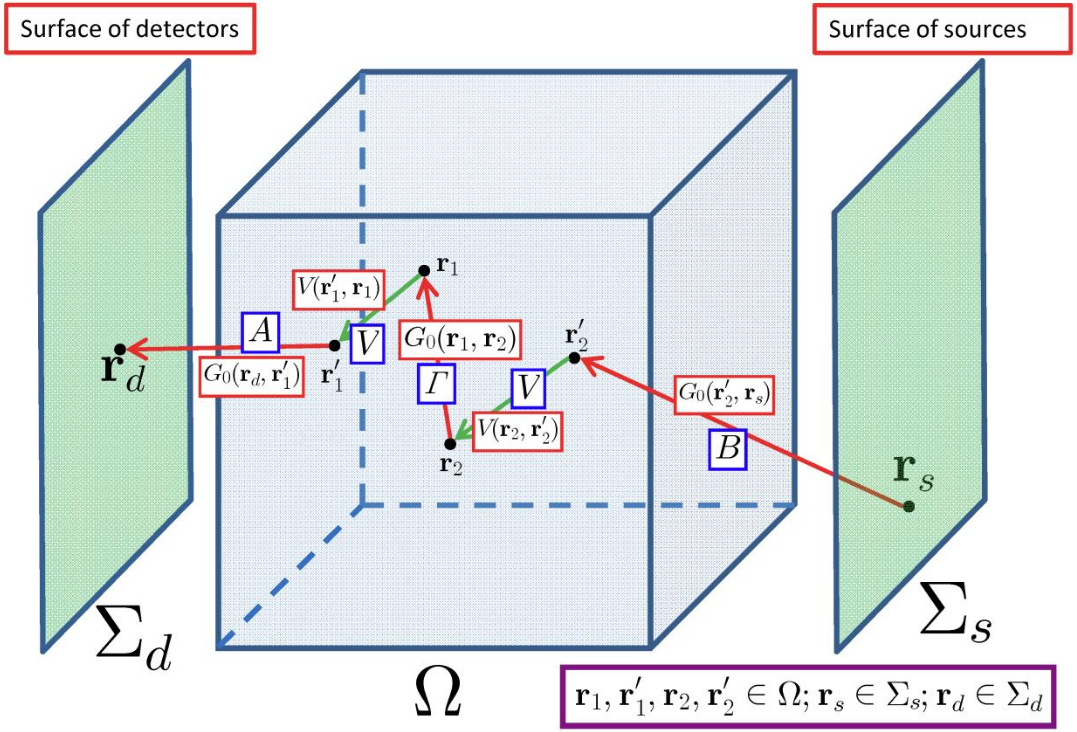
$$N_v = L^3$$

Algebraic Structure of the Inverse Problem

$$A(I - V\Gamma)^{-1}VB = \Phi$$

Every variable is a matrix!

A, B, Γ -- different restrictions of the same unperturbed Green's function G_0



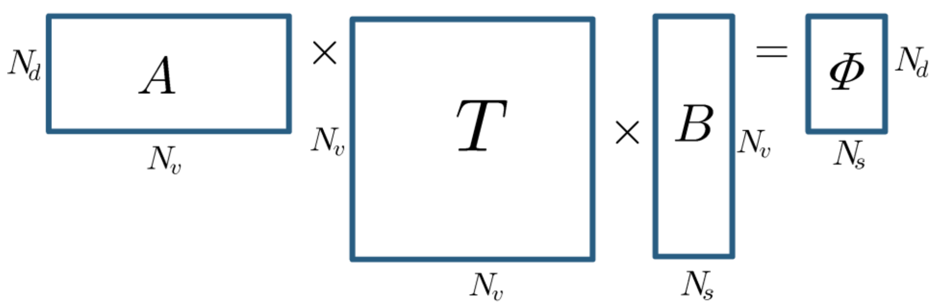
$$A(I - V\Gamma)^{-1}VB = \Phi$$

$$T[V] = (I - V\Gamma)^{-1}V$$

$$V[T] = (I + TV)^{-1}T$$

Let us view T-matrix as the fundamental unknown and use the one-to-one Correspondence between *T* and *V*

$$\boxed{ATB = \Phi}$$



The Experimental T-Matrix

$$A = \sum_{\mu=1}^{N_d} \sigma_\mu^A \left| f_\mu^A \right\rangle \left\langle g_\mu^A \right|$$

$$B = \sum_{\mu=1}^{N_s} \sigma_\mu^B \left| f_\mu^B \right\rangle \left\langle g_\mu^B \right|$$

$$\Rightarrow \quad \boxed{ATB = \Phi} \quad \Rightarrow$$

$$\sigma_\mu^A \sigma_\nu^B \tilde{T}_{\mu\nu} = \tilde{\Phi}_{\mu\nu} \ ,$$

$$1 \le \mu \le N_d$$

$$1 \le \nu \le N_s$$

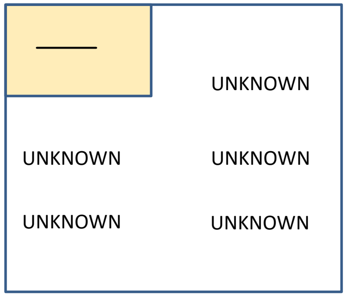
$$\tilde{\Phi}_{\mu\nu} = \left\langle f_\mu^A \middle| \Phi \middle| g_\nu^B \right\rangle , 1 \le \mu \le N_d \ ,$$

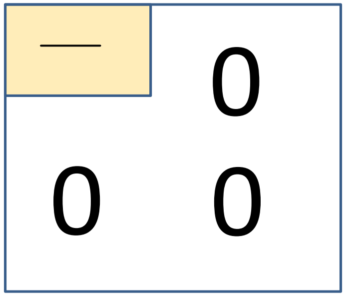
$$1 \le \nu \le N_s$$

$$\tilde{T}_{\mu\nu} = \left\langle g_\mu^A \middle| \Phi \middle| f_\nu^B \right\rangle , 1 \le \mu, \nu \le N_v$$

$$\tilde{T} = R_A^* T R_B = \mathscr{R}[T] \ , \quad T = \mathscr{R}^{-1}[\tilde{T}]$$

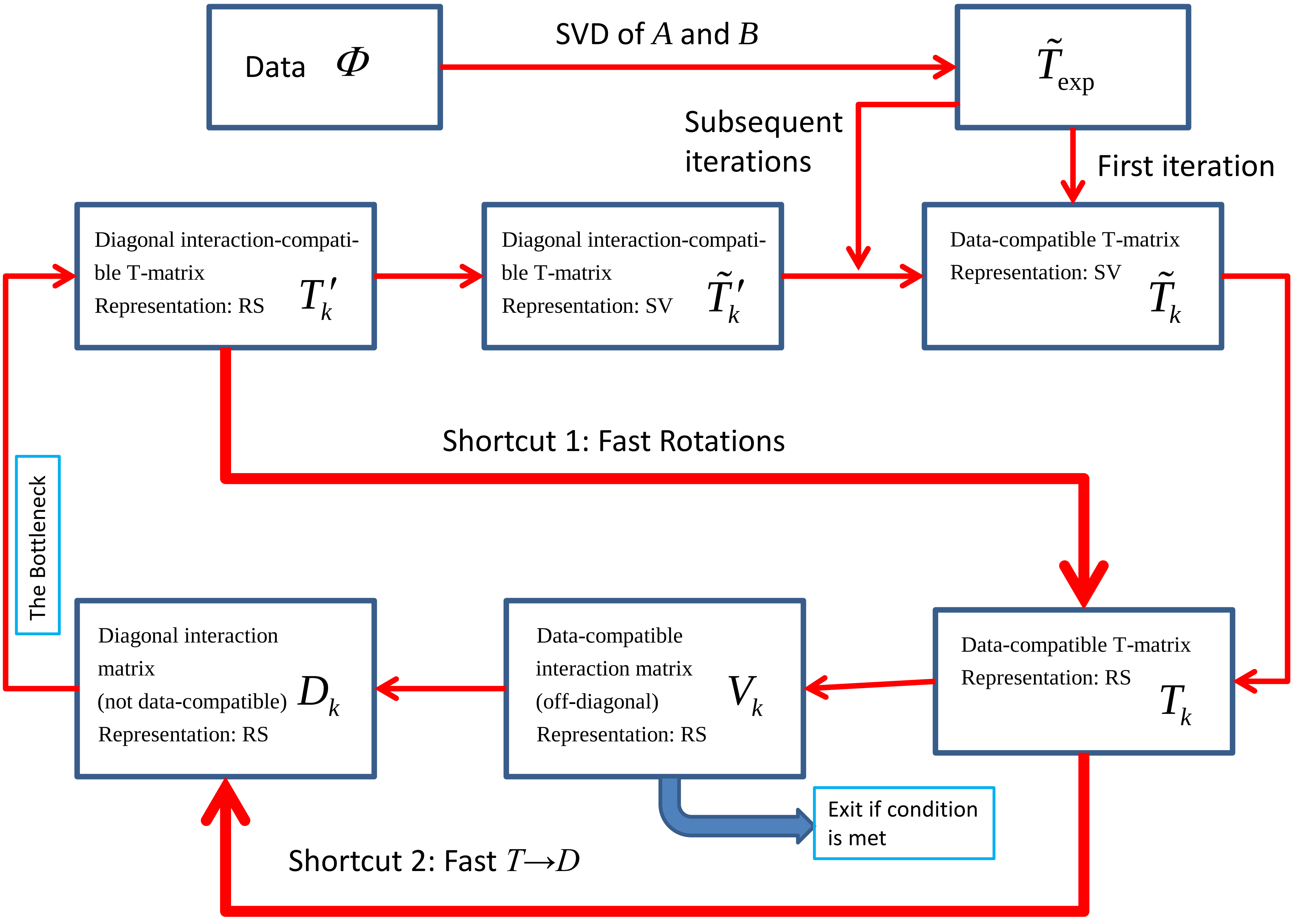
$$\tilde{T}_{\mu\nu} = \begin{cases} \dfrac{1}{\sigma_\mu^A \sigma_\nu^B} \tilde{\Phi}_{\mu\nu} \ , & \text{if} \quad \sigma_\mu^A \sigma_\nu^B > \varepsilon^2 \\ \text{Unknown} \ , & \text{otherwise} \end{cases}$$





$$T_{\exp} = A^+ \Phi B^+$$

$\sigma_1^A\sigma_1^B$	$\sigma_1^A\sigma_2^B$	$\sigma_1^A\sigma_3^B$	$\sigma_1^A\sigma_4^B$	$\sigma_1^A\sigma_5^B$	$\sigma_1^A\sigma_6^B$	$\sigma_1^A\sigma_7^B$
$\sigma_2^A\sigma_1^B$	$\sigma_2^A\sigma_2^B$	$\sigma_2^A\sigma_3^B$	$\sigma_2^A\sigma_4^B$	$\sigma_2^A\sigma_5^B$	$\sigma_2^A\sigma_6^B$	$\sigma_2^A\sigma_7^B$
$\sigma_3^A\sigma_1^B$	$\sigma_3^A\sigma_2^B$	$\sigma_3^A\sigma_3^B$	$\sigma_3^A\sigma_4^B$	$\sigma_3^A\sigma_5^B$	$\sigma_3^A\sigma_6^B$	$\sigma_3^A\sigma_7^B$
$\sigma_4^A\sigma_1^B$	$\sigma_4^A\sigma_2^B$	$\sigma_4^A\sigma_3^B$	$\sigma_4^A\sigma_4^B$	$\sigma_4^A\sigma_5^B$	$\sigma_4^A\sigma_6^B$	$\sigma_4^A\sigma_7^B$
$\sigma_5^A\sigma_1^B$	$\sigma_5^A\sigma_2^B$	$\sigma_5^A\sigma_3^B$	$\sigma_5^A\sigma_4^B$	$\sigma_5^A\sigma_5^B$	$\sigma_5^A\sigma_6^B$	$\sigma_5^A\sigma_7^B$


Data Φ
SVD of A and B
$\tilde{T}_{\exp}$
Subsequent iterations
First iteration
Diagonal interaction-compatible T-matrix
Representation: RS T'_k
Diagonal interaction-compatible T-matrix
Representation: SV $\tilde{T}'_k$
Data-compatible T-matrix
Representation: SV $\tilde{T}_k$
Shortcut 1: Fast Rotations
The Bottleneck
Diagonal interaction matrix
(not data-compatible) D_k
Representation: RS
Data-compatible interaction matrix
(off-diagonal) V_k
Representation: RS
Data-compatible T-matrix
Representation: RS T_k
Exit if condition is met
Shortcut 2: Fast $T \rightarrow D$

Computational Shortcut: Fast Rotations

5: $\tilde{T}_k' = \mathscr{R}[T_k']$

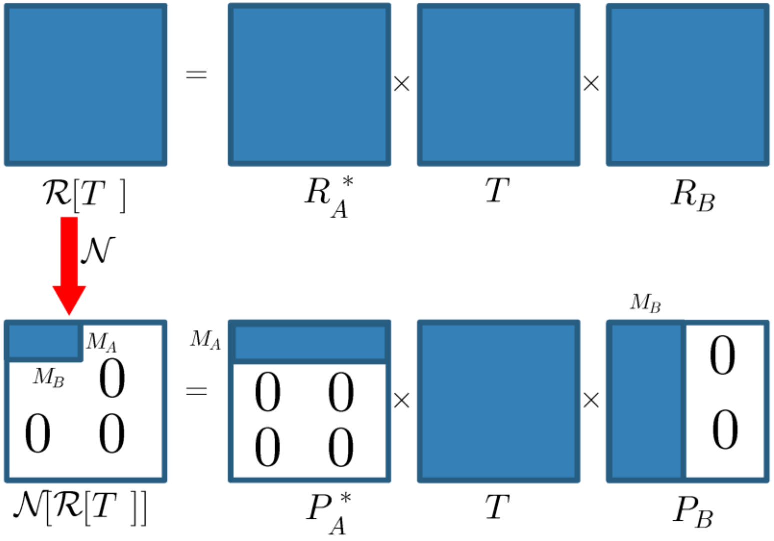
6: $\tilde{T}_{k+1} = \mathscr{O}[\tilde{T}_k']$

1: $T_{k+1} = \mathscr{R}^{-1}[\tilde{T}_{k+1}]$

$$\mathscr{O}[\tilde{T}] = \mathscr{M}[\tilde{T}] + \tilde{T}_{\exp} = \tilde{T} - \mathscr{N}[\tilde{T}] + \tilde{T}_{\exp}$$

$$\mathscr{M}[\tilde{T}] + \mathscr{N}[\tilde{T}] = \tilde{T}$$

$$T_{k+1} = \mathscr{R}^{-1}[\mathscr{O}[\mathscr{R}[T_k']]] = T_k' + T_{\exp} - \mathscr{R}^{-1}[\mathscr{N}[\mathscr{R}[T_k']]]$$



$$\mathscr{R}^{-1}[\mathscr{N}[\mathscr{R}[T]]] = P_A(P_A^* T P_B) P_B^*$$

Operation of "Diagonalization" and Linear Reconstructions

$$D = \mathscr{D}[V]$$

$$D_{ij} = \delta_{ij} \sum_j w(r_{ij}) V_{ij}$$

If $w(r_{ij}) = \delta_{ij}$, we can analyse the algorithm in the linear regime:

$$\left| v_{k+1} \right\rangle = \left| v_{\exp} \right\rangle + (I - W) \left| v_k \right\rangle$$

$$W_{ij} = (P_A^* P_A)_{ij} (P_B^* P_B)_{ji}$$

Vector containing
the diagonal part of *V*

Fixed point: $\left| v_\infty \right\rangle = W^{-1} \left| v_{\exp} \right\rangle$

Tikhonov regularization: $W \rightarrow W + \lambda^2 I$

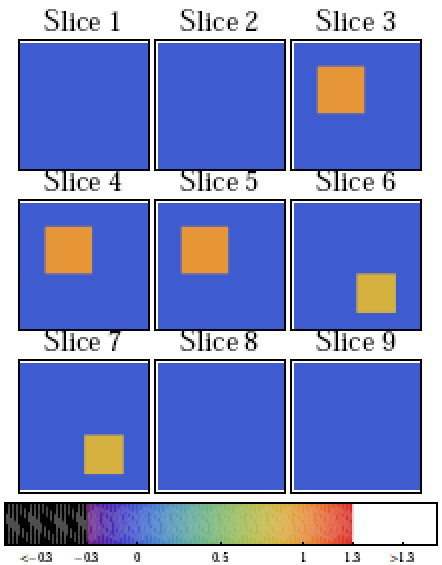
Practical tip: Richardson iteration is a very slow way to arrive at the linearized solution.
Use direct solver of CG to compute linearized solution and then use this result as an initial guess
For the nonlinear iterations.

Diffraction tomography:

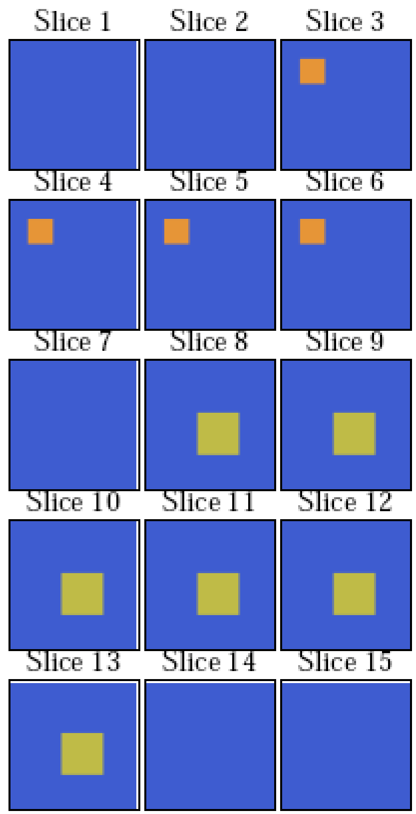
$$G_0(\boldsymbol{r},\boldsymbol{r}') = \frac{\exp\left(ik\left|\boldsymbol{r}-\boldsymbol{r}'\right|\right)}{\left|\boldsymbol{r}-\boldsymbol{r}'\right|}$$

Contrast:

$$\chi(\boldsymbol{r}) = \frac{\varepsilon(\boldsymbol{r})-1}{4\pi}$$



16x16x9
Nv=2,304
Ns*Nd=234,256



30x30x15
Nv=13,500
Ns*Nd=2,085,136

$$\chi(\boldsymbol{r}) = \chi_0 * \text{Shape}(\boldsymbol{r})$$

$$0 \le \text{Shape}(\boldsymbol{r}) \le 1$$

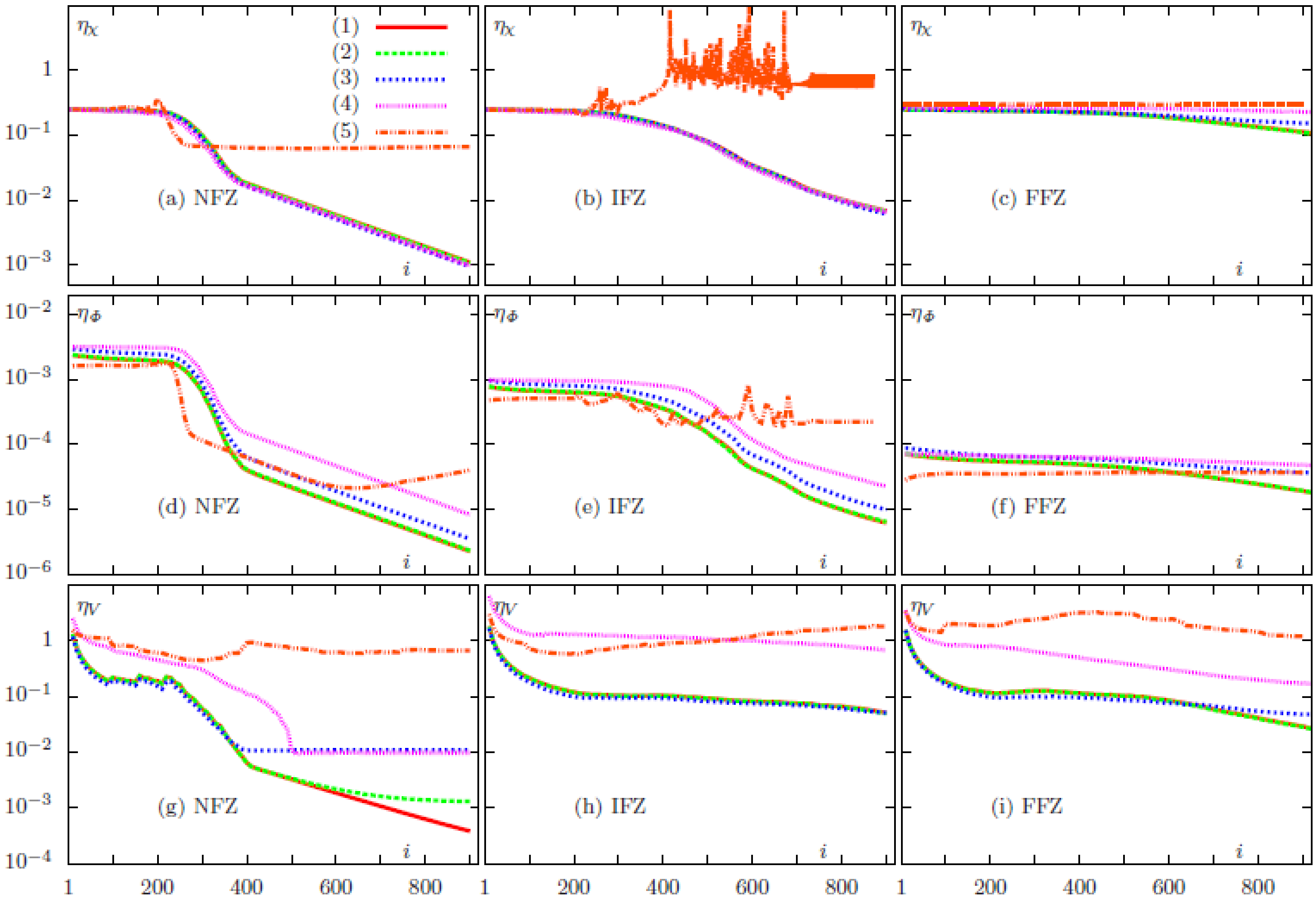
$$\chi_0 = 0.00175,$$

0.0175

0.175

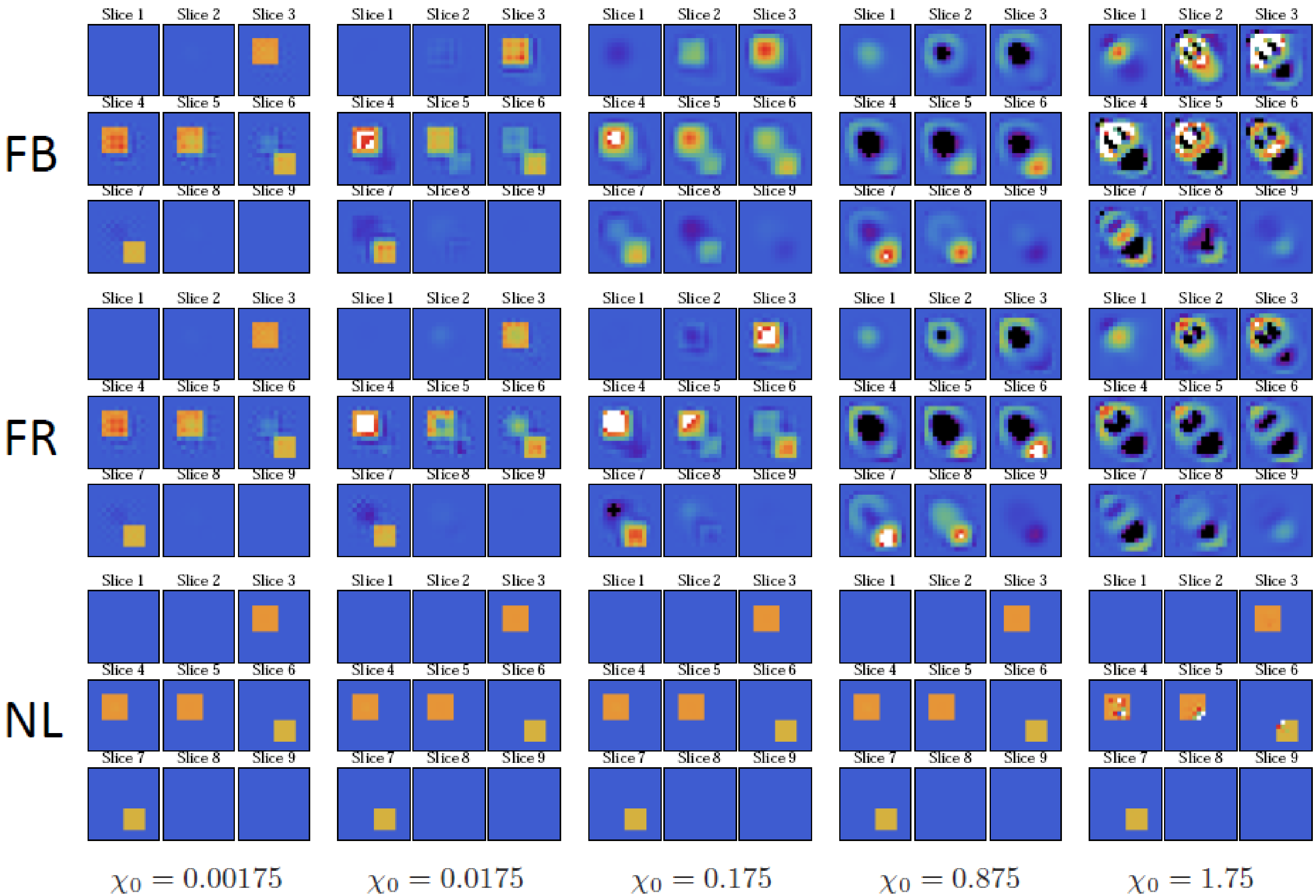
0.875

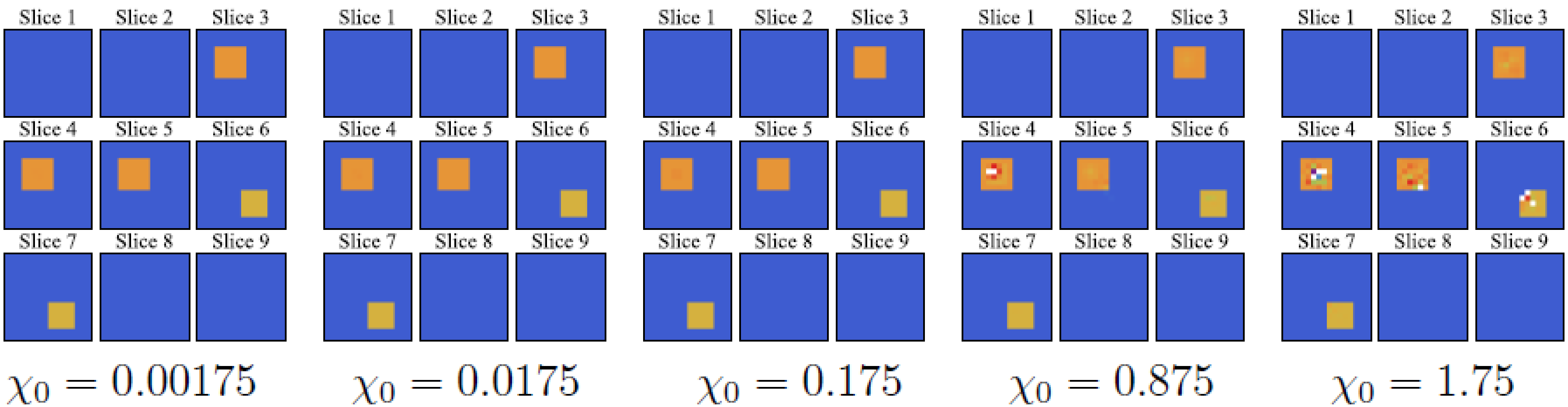
1.75



(1)
(2)
(3)
(4)
(5)
(a) NFZ
(b) IFZ
(c) FFZ
(d) NFZ
(e) IFZ
(f) FFZ
(g) NFZ
(h) IFZ
(i) FFZ
η_X
η_Φ
η_V
i
1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
200
400
600
800

No improvements: 900 iterations





With improvements: 70 iterations

- Start from linearized reconstruction (can be computed fast using our method)
- Use weighted summation to the diagonal for “force-diagonalization”
- Use reciprocity of source-detector pairs to improve symmetry of the experimental T-matrix
- Method starts to break down due to incorrect assignment of non-interacting voxels (this can be avoid altogether – not a problem of convergence)

Diffusion tomography:

$$G_0(\boldsymbol{r},\boldsymbol{r}') = \frac{\exp\left(-k\left|\boldsymbol{r}-\boldsymbol{r}'\right|\right)}{\left|\boldsymbol{r}-\boldsymbol{r}'\right|}$$

Contrast:

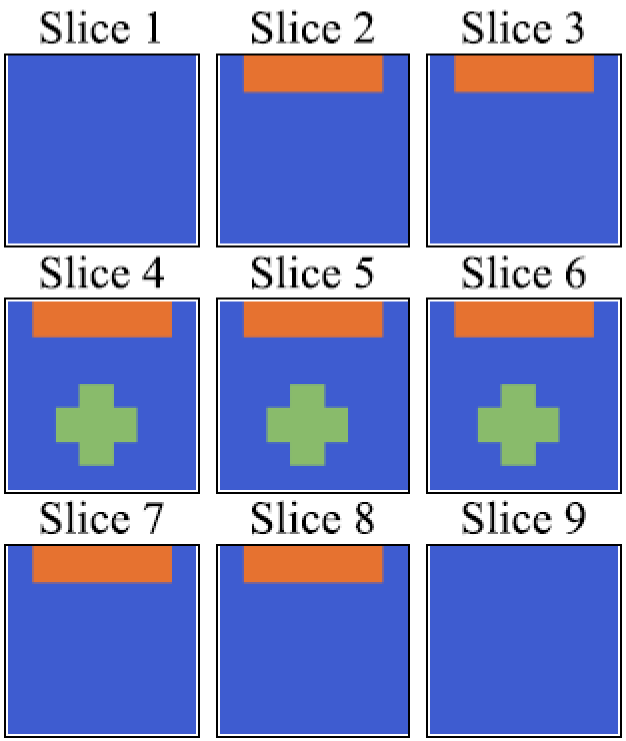
$$\delta\alpha(\boldsymbol{r}) = \frac{\alpha(\boldsymbol{r})-\alpha_0}{\alpha_0}$$

Optical depth:

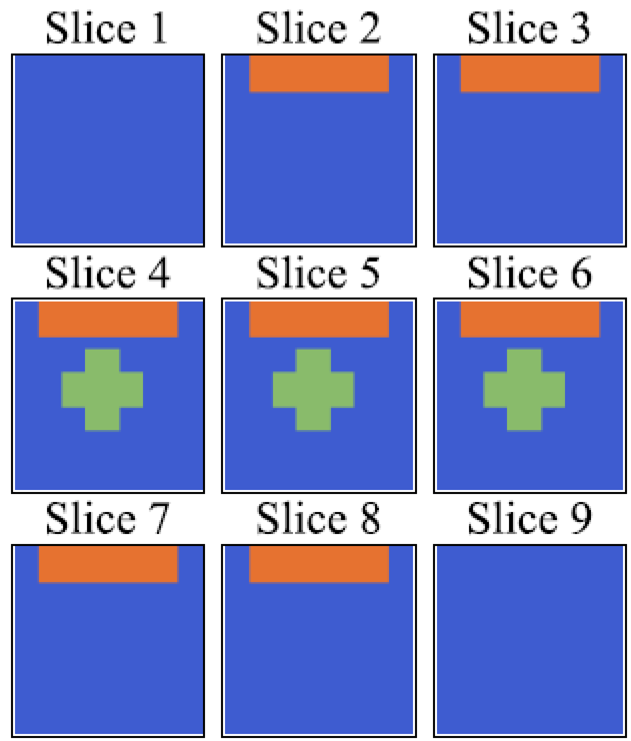
$$kL \approx 2$$

Noise:

2%

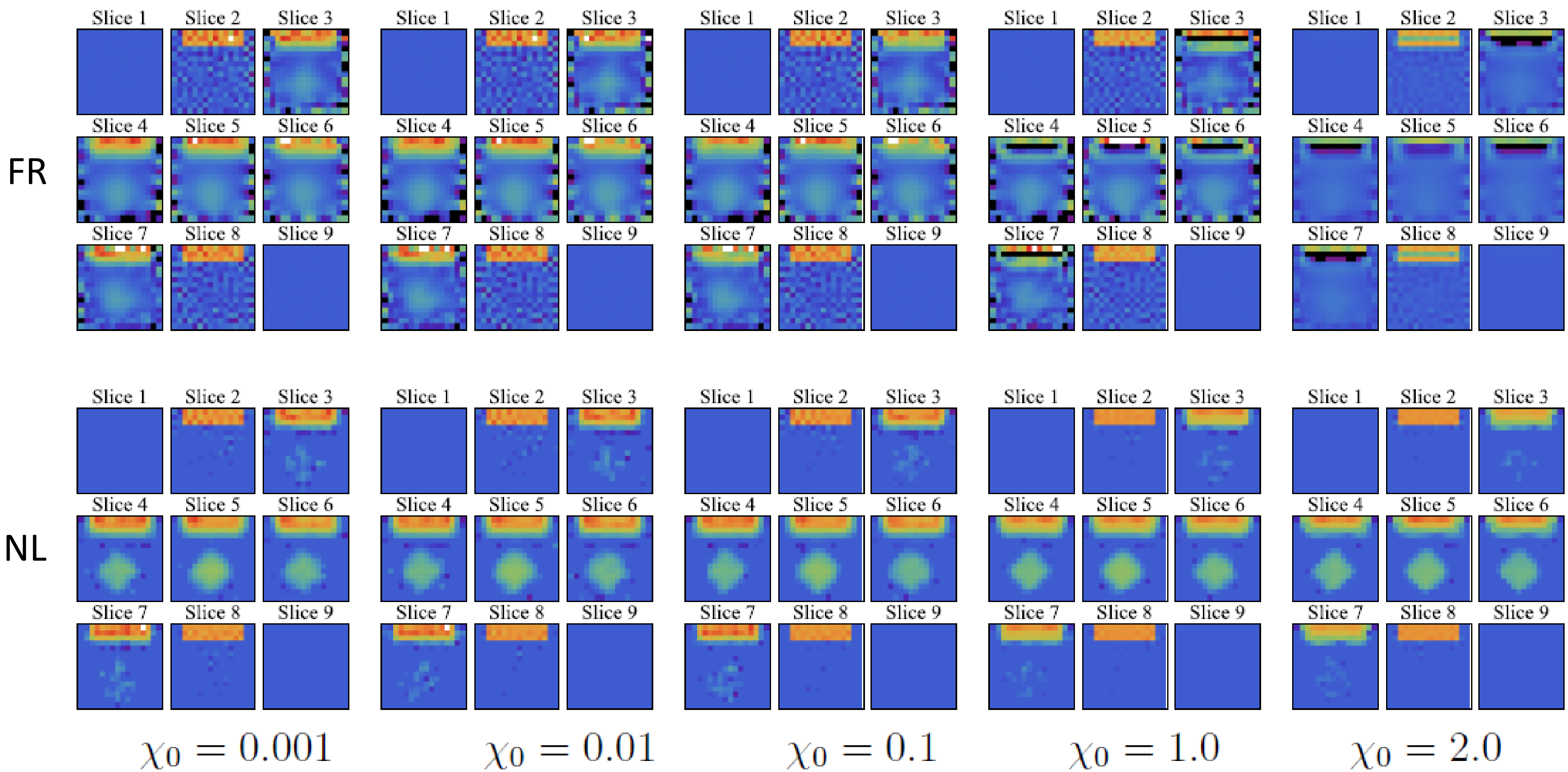


Target far

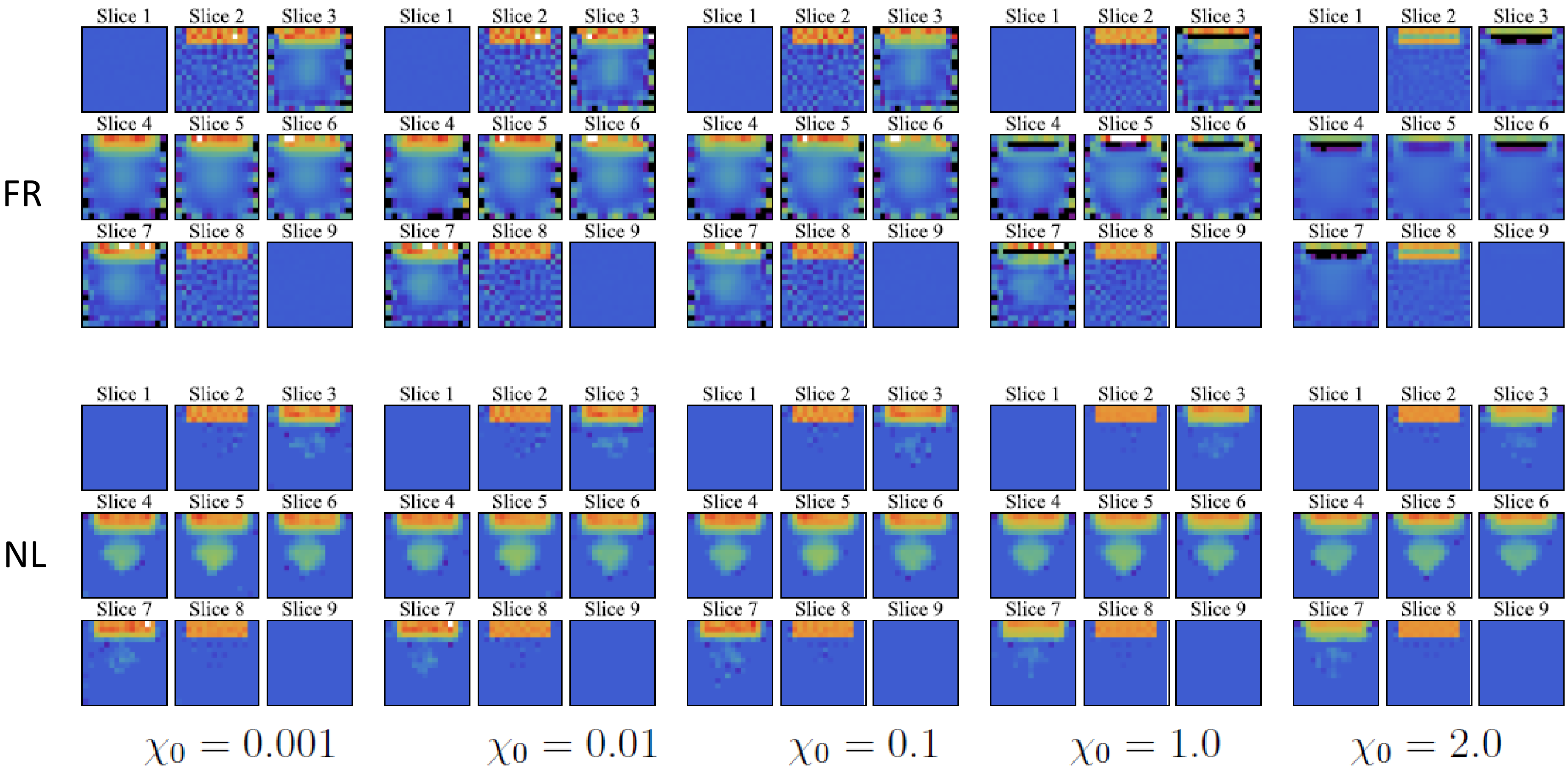


Target near

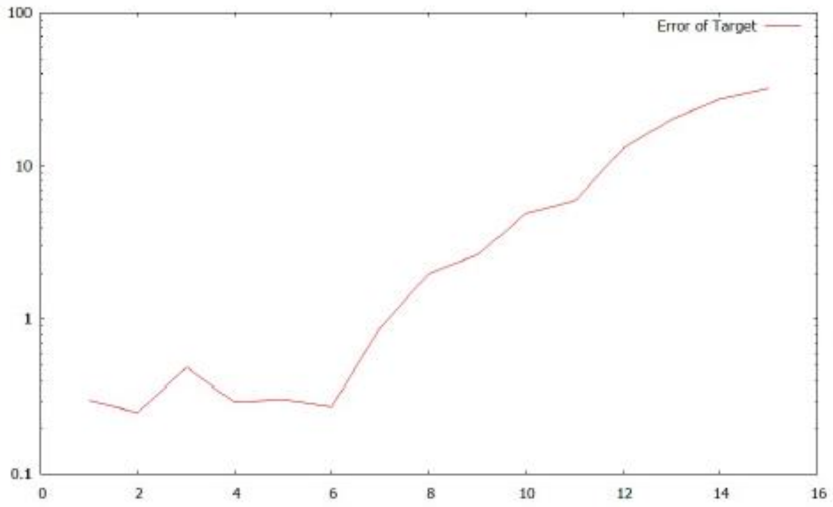
Target far

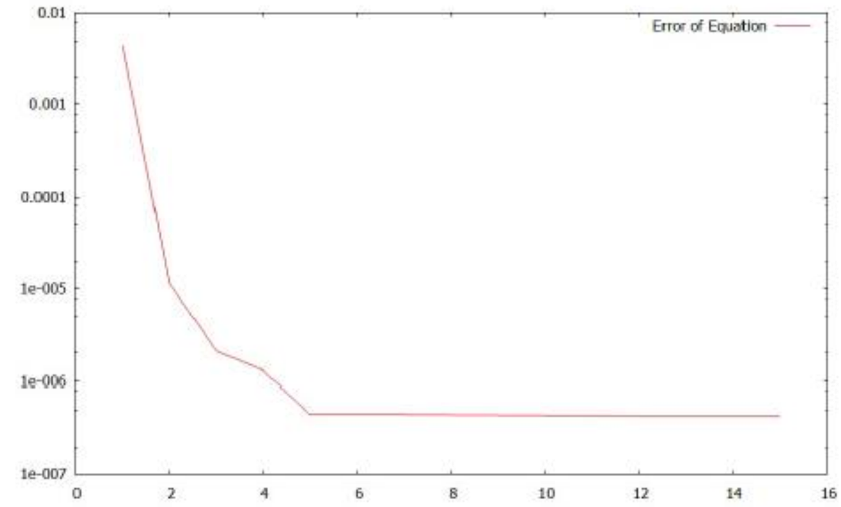

FR
NL
Slice 1
Slice 2
Slice 3
Slice 4
Slice 5
Slice 6
Slice 7
Slice 8
Slice 9
$\chi_0 = 0.001$
$\chi_0 = 0.01$
$\chi_0 = 0.1$
$\chi_0 = 1.0$
$\chi_0 = 2.0$

Target near


FR
NL
Slice 1
Slice 2
Slice 3
Slice 4
Slice 5
Slice 6
Slice 7
Slice 8
Slice 9
$\chi_0 = 0.001$
$\chi_0 = 0.01$
$\chi_0 = 0.1$
$\chi_0 = 1.0$
$\chi_0 = 2.0$

DCTMC works when Newton-Gauss fails (Convergence of Levenberg-Marquardt Iterations for the inverse diffraction problem, moderate contrast 0.0175)





CONCLUSIONS

- DCTMC works for nonlinear ISP with fairly strong nonlinearity
- DCTMC is, unfortunately, a complicated method: it requires many tweaks, attention to detail, and good programming to work
- As any other method, DCTMC breaks at some point. Not every nonlinear ISP can be solved!

H.W.Levinson and V.A.Markel, Solution of the nonlinear inverse scattering problem by T-matrix completion. I. Theory, ***Phys. Rev. E*** **94, 043317 (2016)**

H.W.Levinson and V.A.Markel, Solution of the nonlinear inverse scattering problem by T-matrix completion. II. Simulations, ***Phys. Rev. E*** **94, 043318 (2016)**