Stochastic optimization method for analytic continuation

A. S. Mishchenko

RIKEN (Institute of Physical and Chemical Research), Japan

RRC "Kurchatov Institute", Russia

Stochastic optimization method for analytic continuation

A. S. Mishchenko
RIKEN (Institute of Physical and Chemical Research), Japan
RRC "Kurchatov Institute", Russia

1. Quantities one can get by QMC – correlation functions.
2. Examples of useful correlation functions.
3. Extracting physical information: necessity of analytic continuation.
4. General problem to solve: Fredholm integral equation of kind I.
5. Where similar problems are encountered?
6. Why the naïve methods fail?
7. Tikhonov-Phillips regularization – first successful approach.
8. More sophisticated methods: MaxEnt and Stochastic sampling.
9. Stochastic optimization method (SOM) as the utmost accomplishment of stochastic methods principles.

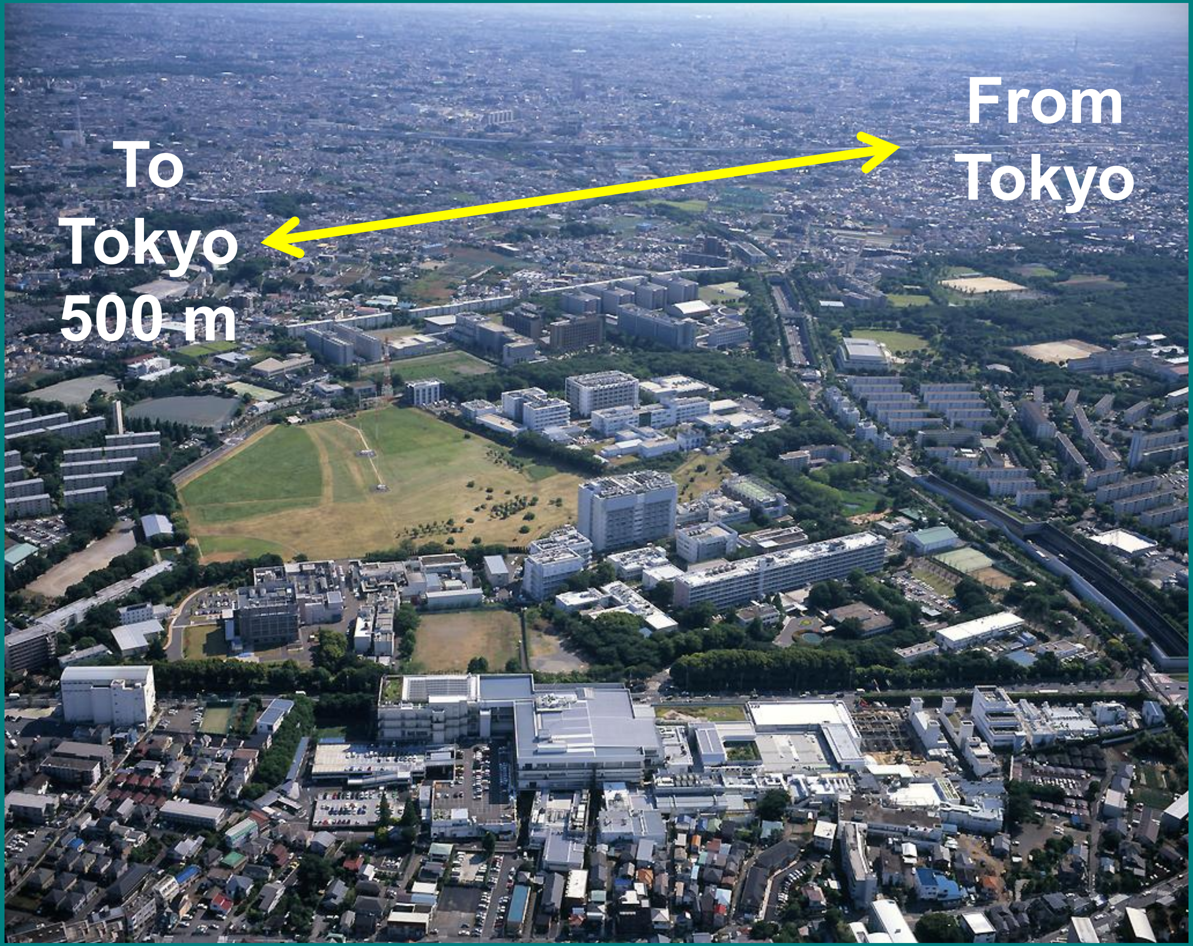
Diagrammatic Monte Carlo and new method of analytic continuation

A. S. Mishchenko

RIKEN (Institute of Physical and Chemical Research), Japan

Diagrammatic Monte Carlo and new method of analytic continuation

A. S. Mishchenko
RIKEN (Institute of Physical and Chemical Research), Japan



Examples of problems where one can get an important correlation function

Simple but not the simplest example: polaron

$$\hat{H}_0 = \sum_{\mathbf{k}} \epsilon(\mathbf{k}) a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + \sum_{\mathbf{q}} \omega_{\mathbf{q}} b^\dagger_{\mathbf{q}} b_{\mathbf{q}}$$

$$\hat{H}_{\rm int} = \sum_{\mathbf{k},\mathbf{q}} V(\mathbf{k},\mathbf{q})(b^\dagger_{\mathbf{q}} - b_{-\mathbf{q}}) a^\dagger_{\mathbf{k}-\mathbf{q}} a_{\mathbf{k}} + h.c.$$

Examples of problems where one can get an important correlation function

Simple but not the simplest example: polaron

$$\hat{H}_0 = \sum_{\mathbf{k}} \epsilon(\mathbf{k}) a^{\dagger}_{\mathbf{k}} a_{\mathbf{k}} + \sum_{\mathbf{q}} \omega_{\mathbf{q}} b^{\dagger}_{\mathbf{q}} b_{\mathbf{q}}$$

$$\hat{H}_{\text{int}} = \sum_{\mathbf{k},\mathbf{q}} V(\mathbf{k},\mathbf{q}) (b^{\dagger}_{\mathbf{q}} - b_{-\mathbf{q}}) a^{\dagger}_{\mathbf{k}-\mathbf{q}} a_{\mathbf{k}} + h.c.$$

Green function by QMC

$$G_{\mathbf{k}}(\tau) = \langle \text{vac} \mid a_{\mathbf{k}}(\tau) a^{\dagger}_{\mathbf{k}} \mid \text{vac} \rangle$$

Examples of problems where one can get an important correlation function

Simple but not the simplest example: polaron

$$\hat{H}_0 = \sum_{\mathbf{k}} \epsilon(\mathbf{k}) a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + \sum_{\mathbf{q}} \omega_{\mathbf{q}} b^\dagger_{\mathbf{q}} b_{\mathbf{q}}$$

$$\hat{H}_{\rm int} = \sum_{\mathbf{k},\mathbf{q}} V(\mathbf{k},\mathbf{q}) (b^\dagger_{\mathbf{q}} - b_{-\mathbf{q}}) a^\dagger_{\mathbf{k}-\mathbf{q}} a_{\mathbf{k}} + h.c.$$

Green function by QMC

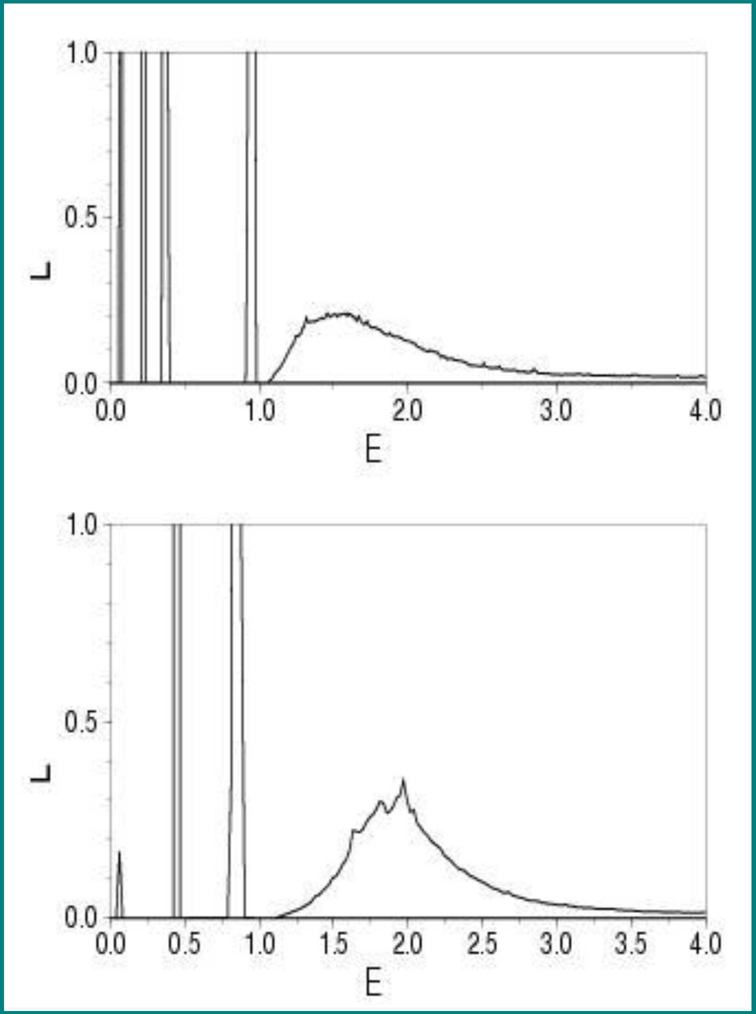
$$G_{\mathbf{k}}(\tau) = \langle \text{vac} \mid a_{\mathbf{k}}(\tau) a^\dagger_{\mathbf{k}} \mid \text{vac} \rangle$$

No simple connection to measurable properties

Physical properties under interest: Lehman function

Lehmann spectral function (LSF)

$$L_{\mathbf{k}}(\omega) = \sum_{\nu} \delta(\omega - E_{\nu}(\mathbf{k}))\, |\langle \nu | a^{\dagger}_{\mathbf{k}} | \mathrm{vac} \rangle|^2$$



LSF has poles (sharp peaks) at the energies of stable (metastable) states. It is a measurable (in ARPES) quantity.

Noninteracting one is simple:

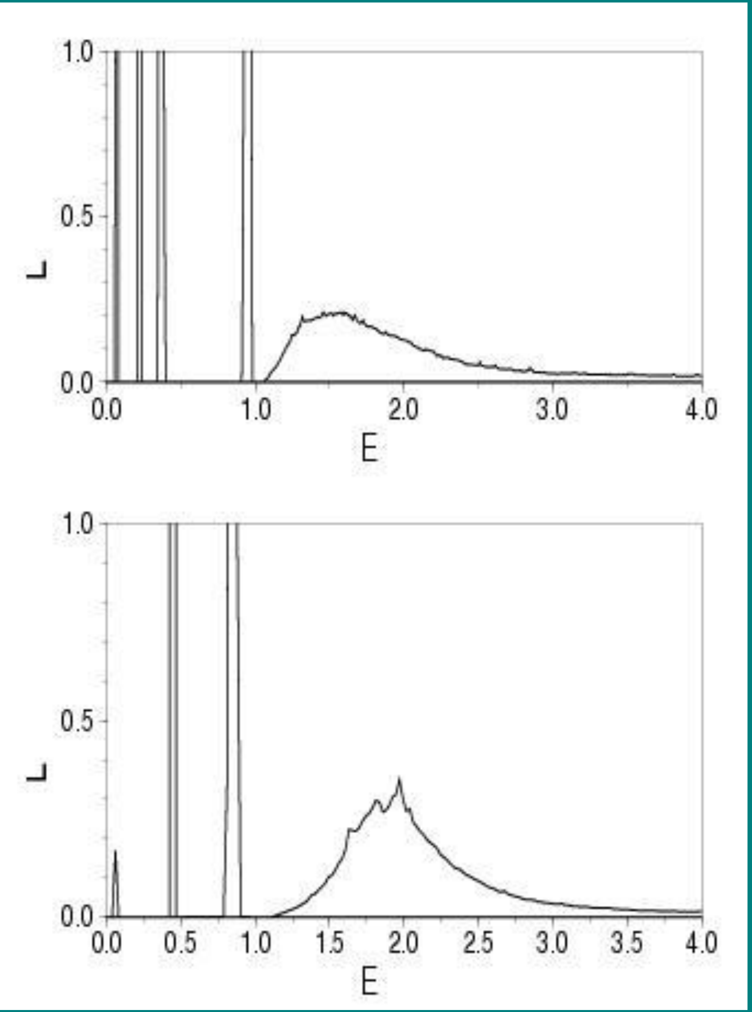
$$L^{\mathrm{NONINT}}_{\mathbf{k}}(\omega) = \delta(\omega - \epsilon(\mathbf{k}))$$

Physical properties under interest: Lehmann function.

Lehmann spectral function (LSF)

$$L_{\mathbf{k}}(\omega) = \sum_{\nu} \delta(\omega - E_{\nu}(\mathbf{k}))\, |\langle \nu | a^{\dagger}_{\mathbf{k}} | \mathrm{vac} \rangle|^2$$

LSF has poles (sharp peaks) at the energies of stable (metastable) states. It is a measurable (in ARPES) quantity.



LSF of one particle at T=0 can be determined from equation:

$$G_{\mathbf{k}}(\tau) = \int_0^{\infty} d\omega\, L_{\mathbf{k}}(\omega)\, e^{-\omega\tau}$$

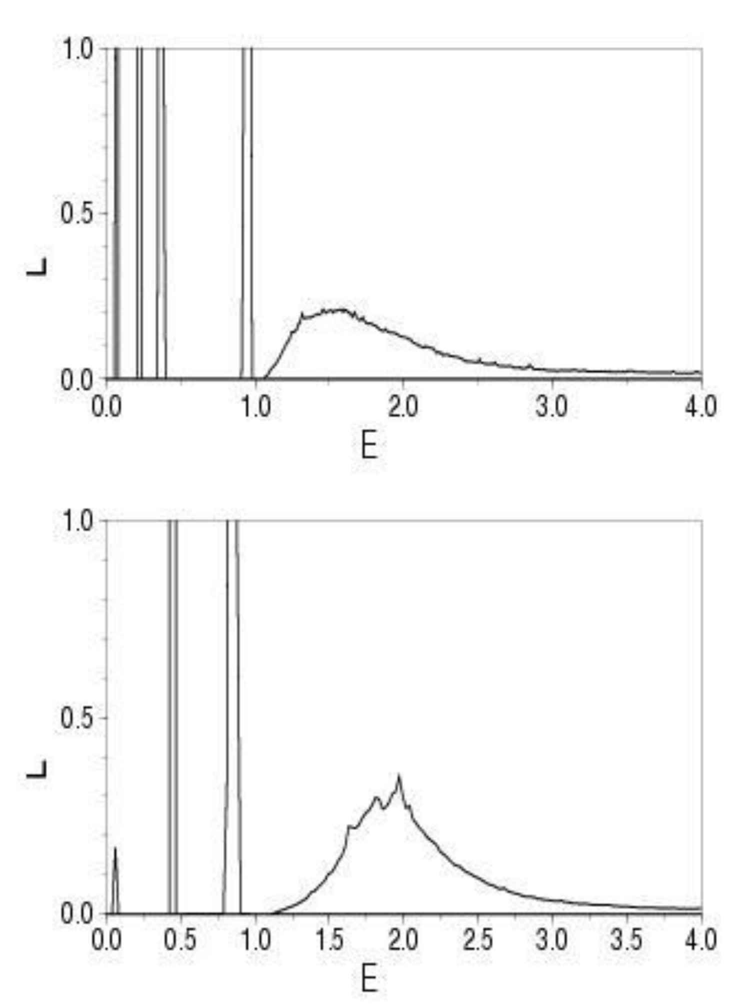
Fredholm first kind.

Physical properties under interest: Z-factor and energy

Lehmann spectral function (LSF)

$$L_{\mathbf{k}}(\omega) = \sum_{\nu} \delta(\omega - E_{\nu}(\mathbf{k}))\, |\langle \nu | a^{\dagger}_{\mathbf{k}} | \mathrm{vac} \rangle|^2$$

$$G_{\mathbf{k}}(\tau) = \int_0^{\infty} d\omega\, L_{\mathbf{k}}(\omega)\, e^{-\omega\tau}$$



If the state with the lowest energy in the sector of given momentum is stable

$$L_{\mathbf{k}}(\omega) = Z^{(\mathbf{k})}\, \delta(\omega - E(\mathbf{k})) + \ldots$$

The asymptotic behavior is

$$G_{\mathbf{k}}(\tau \gg \max[\omega^{-1}_{\mathbf{q},\kappa}]) \;\to\; Z^{(\mathbf{k})}\, \exp[-E_{\mathrm{g.s.}}(\mathbf{k})\tau]$$

Physical properties under interest: Z-factor and energy

The asymptotic behavior is

$$G_{\mathbf{k}}(\tau \gg \max[\omega_{\mathbf{q},\kappa}^{-1}]) \rightarrow Z^{(\mathbf{k})} \exp[-E_{\mathrm{g.s.}}(\mathbf{k})\tau]$$

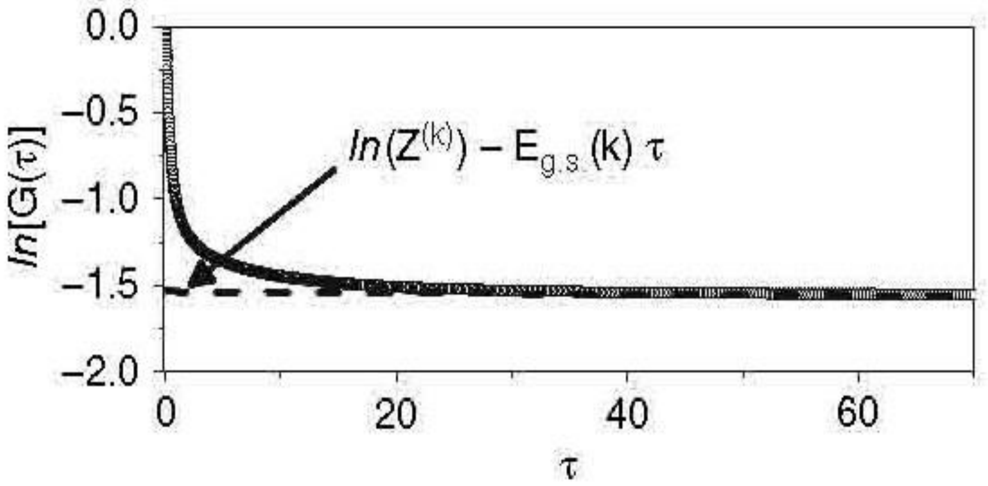


Fig. 12.1. Typical behavior of the GF of a polaron and determination of $Z^{(k)}$-factor and energy of the ground state from the fit of the linear asymptotics

Physical properties under interest: Lehmann function.

Lehmann spectral function (LSF)

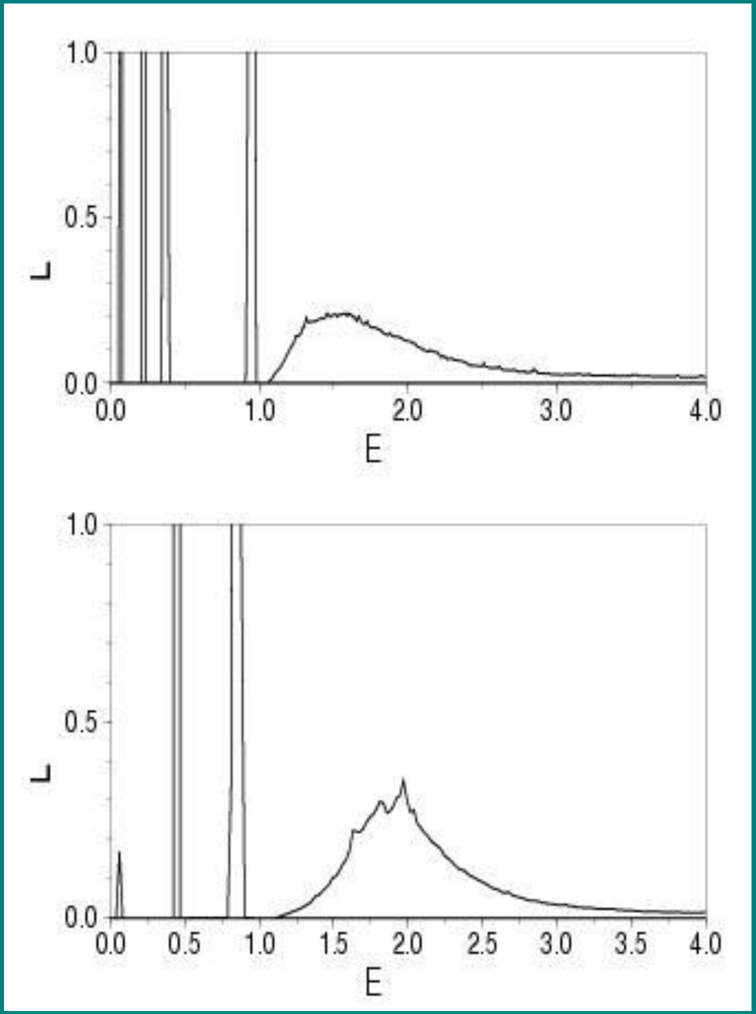
$$L_{\mathbf{k}}(\omega) = \sum_{\nu} \delta(\omega - E_{\nu}(\mathbf{k})) \, |\langle \nu | a^{\dagger}_{\mathbf{k}} | \mathrm{vac} \rangle|^2$$

LSF of one particle at T=0 can be determined from equation:

$$G_{\mathbf{k}}(\tau) = \int_0^{\infty} d\omega \, L_{\mathbf{k}}(\omega) \, e^{-\omega \tau}$$

Solving of this equation is a notoriously difficult problem

$$L_{\mathbf{k}}(\omega) = \hat{\mathcal{F}}^{-1}_{\omega} \left[G_{\mathbf{k}}(\tau) \right]$$



Physical properties under interest: Lehmann function.

Lehmann spectral function (LSF)

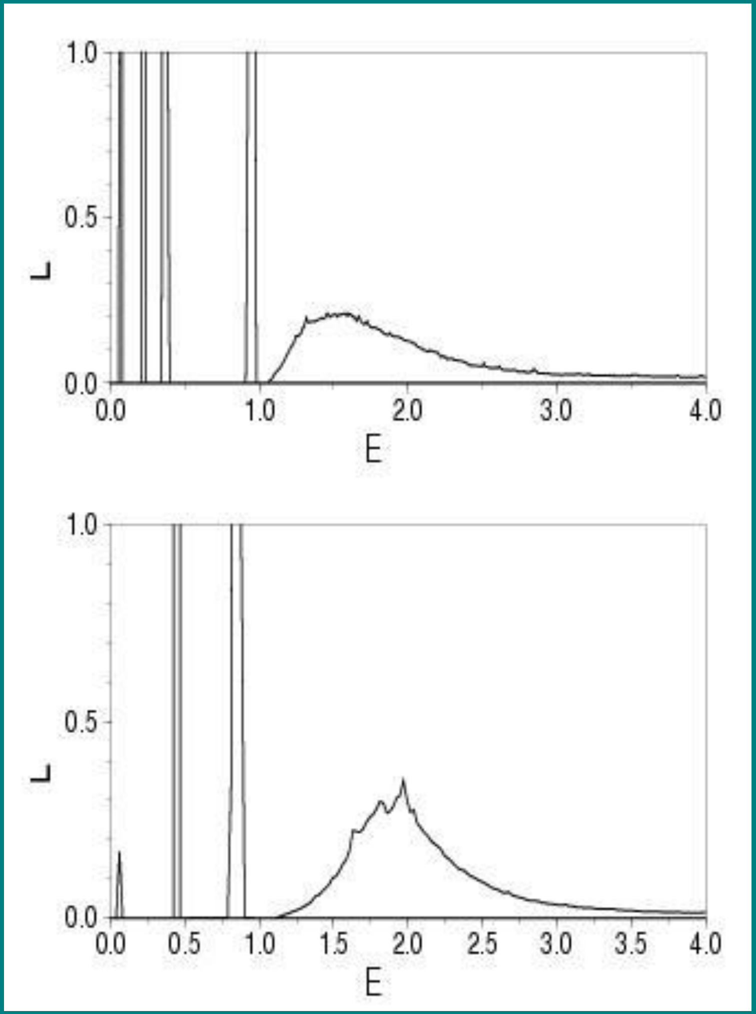
$$L_{\mathbf{k}}(\omega) = \sum_{\nu} \delta(\omega - E_{\nu}(\mathbf{k})) \, |\langle \nu | a_{\mathbf{k}}^{\dagger} | \mathrm{vac} \rangle|^2$$

LSF of one particle at T=0 can be determined from equation:

$$G(m) = \int_{-\infty}^{\infty} d\omega \; \mathcal{K}(m,\omega) \; A(\omega)$$

Solving of this equation is
a notoriously difficult problem

$$L_{\mathbf{k}}(\omega) = \hat{\mathcal{F}}_{\omega}^{-1} \left[G_{\mathbf{k}}(\tau) \right]$$



Solution of integral equation

$$G(m) = \int_{-\infty}^{\infty} d\omega \, K(m,\omega) \, A(\omega)$$

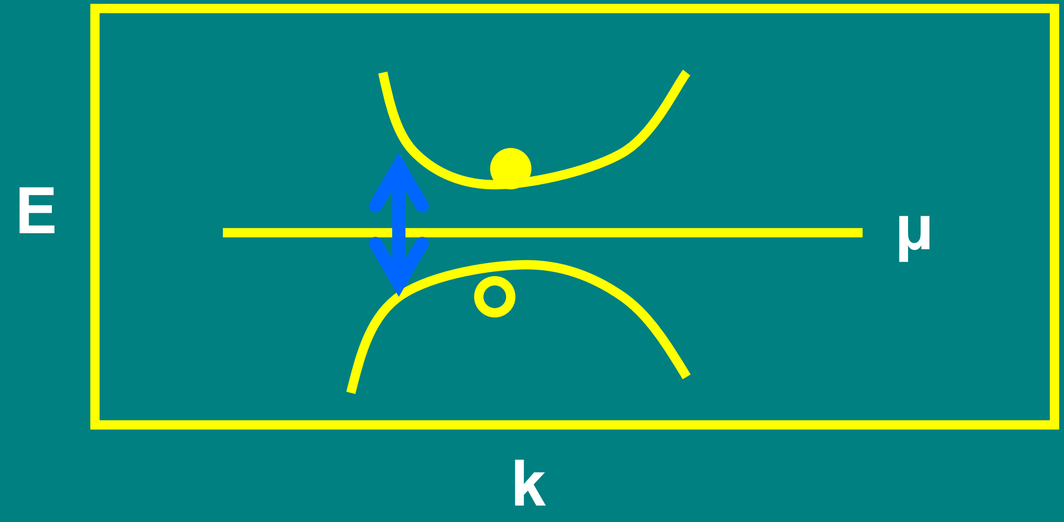
$$L_{\mathbf{K}}(\omega) = \hat{\mathcal{F}}_{\omega}^{-1}\left[G_{\mathbf{K}}(\tau)\right]$$

Examples of problems where one can get an important correlation function

Exciton

$$\hat{H}_0^{\mathrm{par}} = \sum_{\mathbf{k}} \varepsilon_a(\mathbf{k}) a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \sum_{\mathbf{k}} \varepsilon_h(\mathbf{k}) h_{\mathbf{k}} h_{\mathbf{k}}^\dagger$$

$$\hat{H}_{\text{a-h}} = -N^{-1} \sum_{\mathbf{pkk'}} \mathcal{U}(\mathbf{p},\mathbf{k},\mathbf{k'}) a_{\mathbf{p+k}}^\dagger h_{\mathbf{p-k}}^\dagger h_{\mathbf{p-k'}} a_{\mathbf{p+k'}}$$



Coulomb attraction

Examples of problems where one can get an important correlation function

Exciton

$$\hat{H}_0^{\text{par}} = \sum_{\mathbf{k}} \varepsilon_a(\mathbf{k}) a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \sum_{\mathbf{k}} \varepsilon_h(\mathbf{k}) h_{\mathbf{k}} h_{\mathbf{k}}^\dagger$$

$$\hat{H}_{\text{a-h}} = -N^{-1} \sum_{\mathbf{pkk}'} \mathcal{U}(\mathbf{p},\mathbf{k},\mathbf{k}') a_{\mathbf{p}+\mathbf{k}}^\dagger h_{\mathbf{p}-\mathbf{k}}^\dagger h_{\mathbf{p}-\mathbf{k}'} a_{\mathbf{p}+\mathbf{k}'}$$

More realistic

Exciton-polaron

$$\hat{H}_{\text{par-bos}} = i \sum_{\kappa=1}^{Q} \sum_{\mathbf{k},\mathbf{q}} (b_{\mathbf{q},\kappa}^\dagger - b_{-\mathbf{q},\kappa}) \left[\gamma_{aa,\kappa}(\mathbf{k},\mathbf{q}) a_{\mathbf{k}-\mathbf{q}}^\dagger a_{\mathbf{k}} + \gamma_{hh,\kappa}(\mathbf{k},\mathbf{q}) h_{\mathbf{k}-\mathbf{q}}^\dagger h_{\mathbf{k}} + \gamma_{ah,\kappa}(\mathbf{k},\mathbf{q}) h_{\mathbf{k}-\mathbf{q}}^\dagger a_{\mathbf{k}}\right] + h.c.$$

$$\hat{H}_{\text{bos}} = \sum_{\kappa=1}^{Q} \sum_{\mathbf{q}} \omega_{\mathbf{q},\kappa} b_{\mathbf{q},\kappa}^\dagger b_{\mathbf{q},\kappa}$$

Infinite system

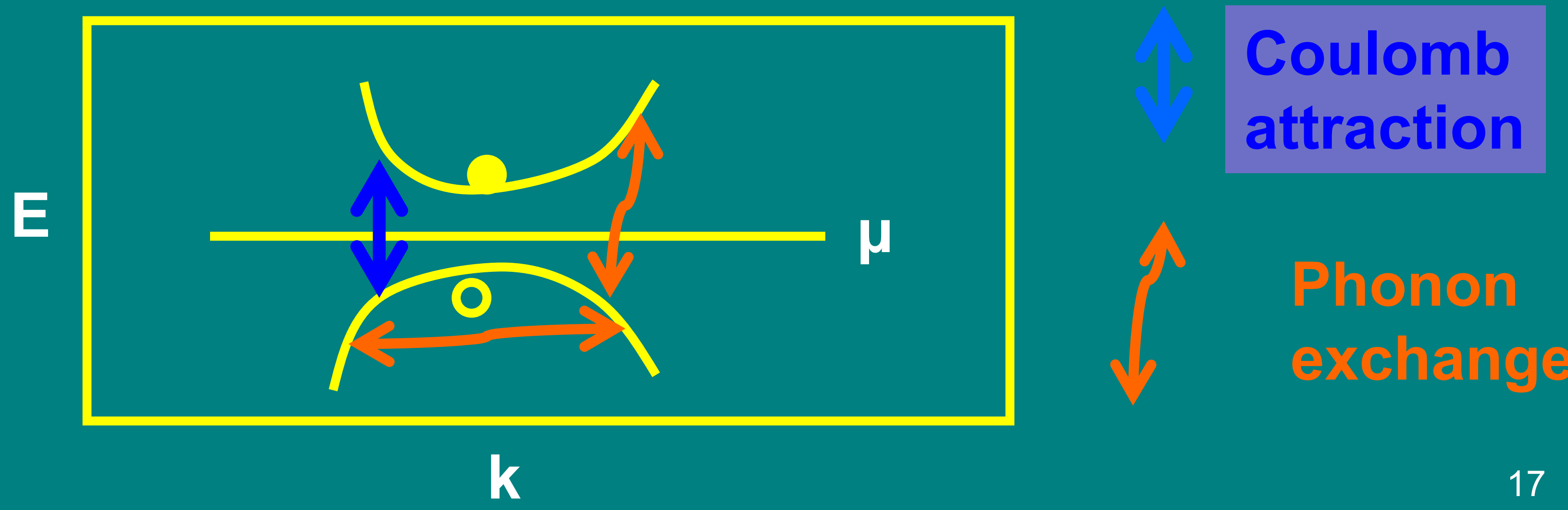
Examples of problems where one can get an important correlation function

Exciton-polaron

$$\hat{H}_0^{\mathrm{par}} = \sum_{\mathbf{k}} \varepsilon_a(\mathbf{k}) a_{\mathbf{k}}^\dagger a_{\mathbf{k}} + \sum_{\mathbf{k}} \varepsilon_h(\mathbf{k}) h_{\mathbf{k}} h_{\mathbf{k}}^\dagger$$

$$\hat{H}_{\text{a-h}} = -N^{-1} \sum_{\mathbf{pkk}'} \mathcal{U}(\mathbf{p},\mathbf{k},\mathbf{k}') a_{\mathbf{p}+\mathbf{k}}^\dagger h_{\mathbf{p}-\mathbf{k}}^\dagger h_{\mathbf{p}-\mathbf{k}'} a_{\mathbf{p}+\mathbf{k}'}$$

$+ \mathbf{H}_{\text{el-ph}}$



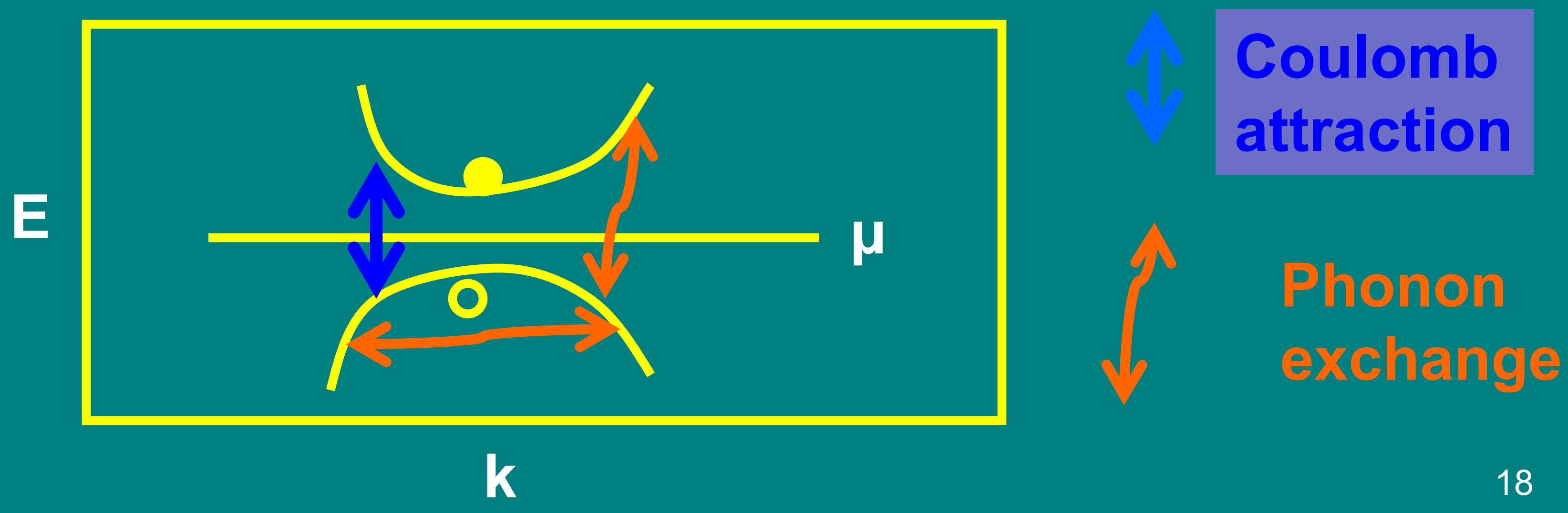
Examples of problems where one can get an important correlation function

Exciton-polaron: two-particle Green function

$$G_{\mathbf{k}}^{\mathbf{p}\mathbf{p}'}(\tau) = \langle \mathrm{vac} \mid a_{\mathbf{k}+\mathbf{p}'}(\tau) h_{\mathbf{k}-\mathbf{p}'}(\tau) h^{\dagger}_{\mathbf{k}-\mathbf{p}} a^{\dagger}_{\mathbf{k}+\mathbf{p}} \mid \mathrm{vac} \rangle$$

$$\mathcal{I}(\omega) = \hat{\mathcal{F}}_{\omega}^{-1} \left[\sum_{\mathbf{p}\mathbf{p}'} G_{\mathbf{k}=0}^{\mathbf{p}\mathbf{p}'}(\tau) \right]$$

Optical absorption

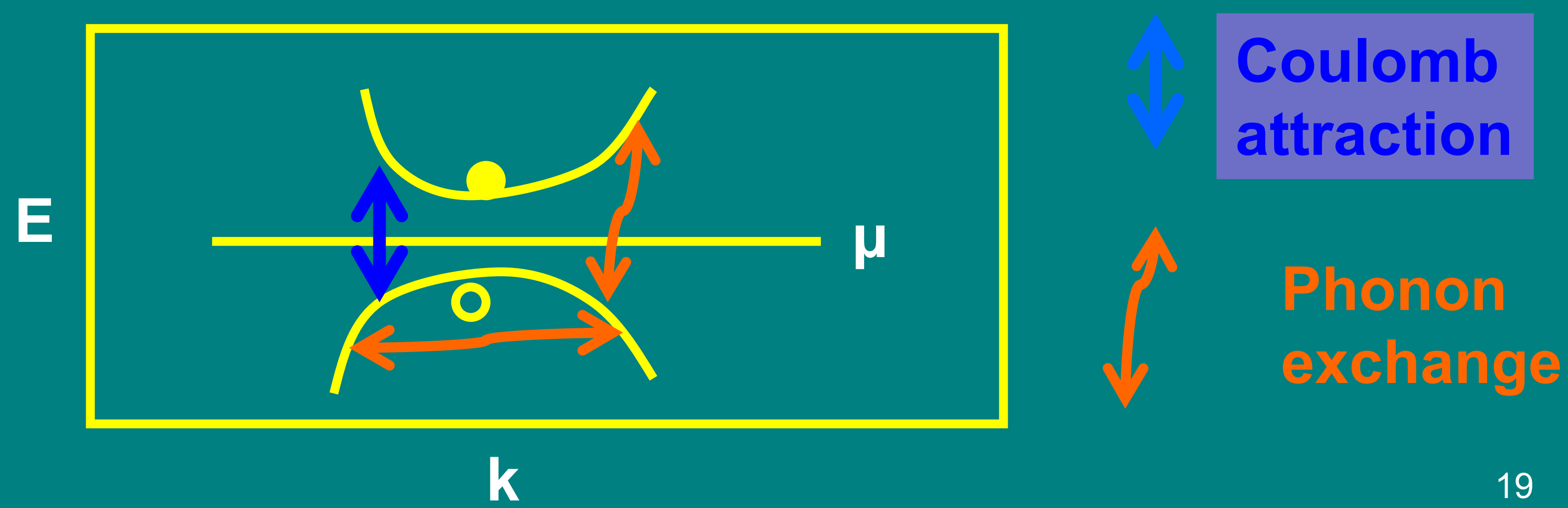


Examples of problems where one can get an important correlation function

$$\mathcal{I}(\omega) = \hat{\mathcal{F}}_{\omega}^{-1}\left[\sum_{\mathbf{pp}'} G_{\mathbf{k}=0}^{\mathbf{pp}'}(\tau)\right]$$

Optical absorption

Also Fredholm integral equation of the first kind



Exact solution for optical spectra of exciton-polaron

A. S. Mishchenko and N. Nagaosa, CMRG, RIKEN ASI

Diagrammatic Monte Carlo

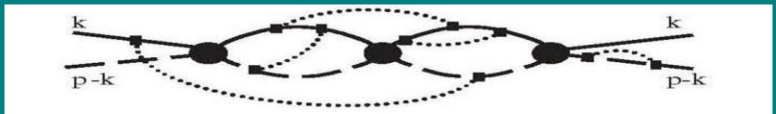
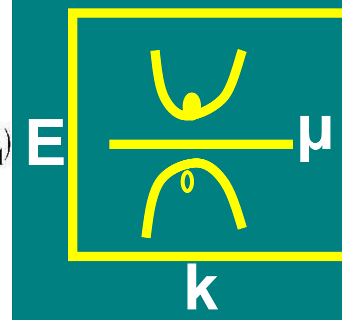


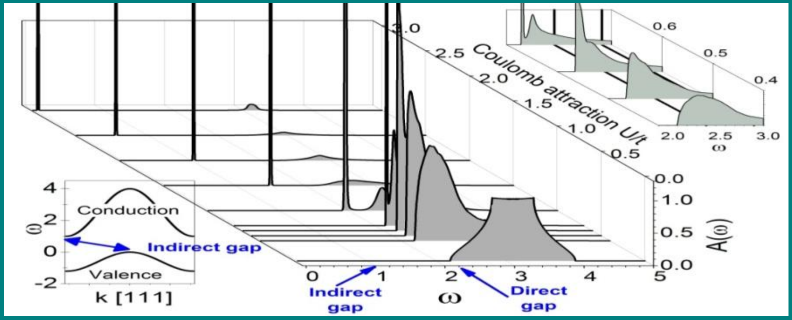
FIG. 1. A typical diagram for $G^{M=0}_{\mathbf{p},\mathbf{k}}(\tau)$. Solid (dashed) lines represent E(H) propagators, solid circles (squares) designate Coulomb (QP-phonon) interactions, and dotted lines are the phonon propagators. Imaginary time runs from left to right.

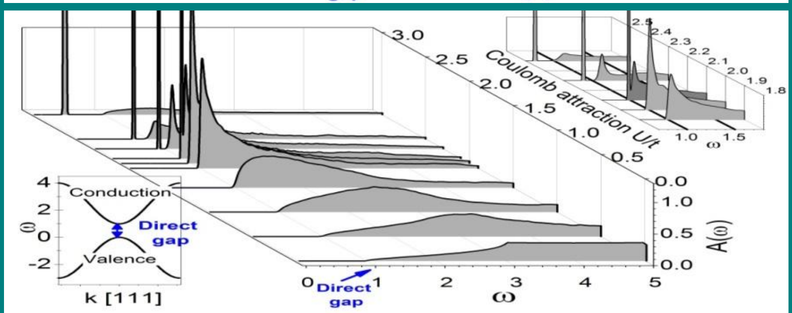
Exciton-polaron

$$H=\sum_{\mathbf{k}}\varepsilon_c(\mathbf{k})e^\dagger_{\mathbf{k}}e_{\mathbf{k}}+\sum_{\mathbf{k}}\varepsilon_v(\mathbf{k})h_{\mathbf{k}}h^\dagger_{\mathbf{k}}+\sum_{\mathbf{q}}\omega_{\mathbf{q}}b^\dagger_{\mathbf{q}}b_{\mathbf{q}}$$
$$-\sum_{\mathbf{k}\mathbf{q}}\left[\frac{g_e(\mathbf{q})}{\sqrt{N}}e^\dagger_{\mathbf{k}-\mathbf{q}}e_{\mathbf{k}}+\frac{g_h(\mathbf{q})}{\sqrt{N}}h^\dagger_{\mathbf{k}-\mathbf{q}}h_{\mathbf{k}}\right](b^\dagger_{\mathbf{q}}+b_{-\mathbf{q}})$$
$$-\sum_{\mathbf{p}\mathbf{k}\mathbf{k}'}\frac{U(\mathbf{p},\mathbf{k},\mathbf{k}')}{N}e^\dagger_{\mathbf{k}}h^\dagger_{\mathbf{p}-\mathbf{k}}h_{\mathbf{p}-\mathbf{k}'}e_{\mathbf{k}'}.$$

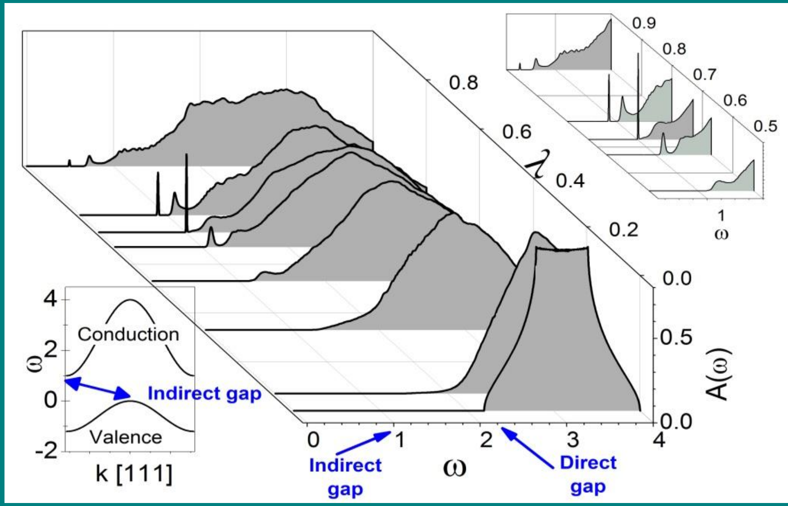


Coulomb attraction
No particle-phonon coupling





No Coulomb attraction
Particle-phonon coupling



$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

There are a lot of problems where one has to solve Fredholm integral equation of the first kind

Many-particle Fermi/Boson system in imaginary times representation

$$G(m) = \int_{-\infty}^{\infty} d\omega \; \mathcal{K}(m,\omega) \; A(\omega)$$

$$\boxed{\mathcal{K}(\tau_m,\omega) = -\frac{\exp(-\tau_m \omega)}{\exp(-\beta\omega) \pm 1}}$$

Many-particle Fermi/Boson system in Matsubara representation

$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

$$\mathcal{G}(i\omega_m) = \int_0^\beta d\tau \; e^{i\omega_m \tau} \; G(\tau)$$

$$G(\tau) = \frac{1}{\beta} \sum_{\omega_m} e^{-i\omega_m \tau} \; \mathcal{G}(i\omega_m)$$

$$\boxed{K(i\omega_m,\omega) = \pm \frac{1}{i\omega_m - \omega}}$$

Optical conductivity at finite T in imaginary times representation

$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

$$\mathcal{G}(i\omega_m) = \int_0^\beta d\tau \; e^{i\omega_m \tau} \; G(\tau)$$

$$G(\tau) = \frac{1}{\beta} \sum_{\omega_m} e^{-i\omega_m \tau} \; \mathcal{G}(i\omega_m)$$

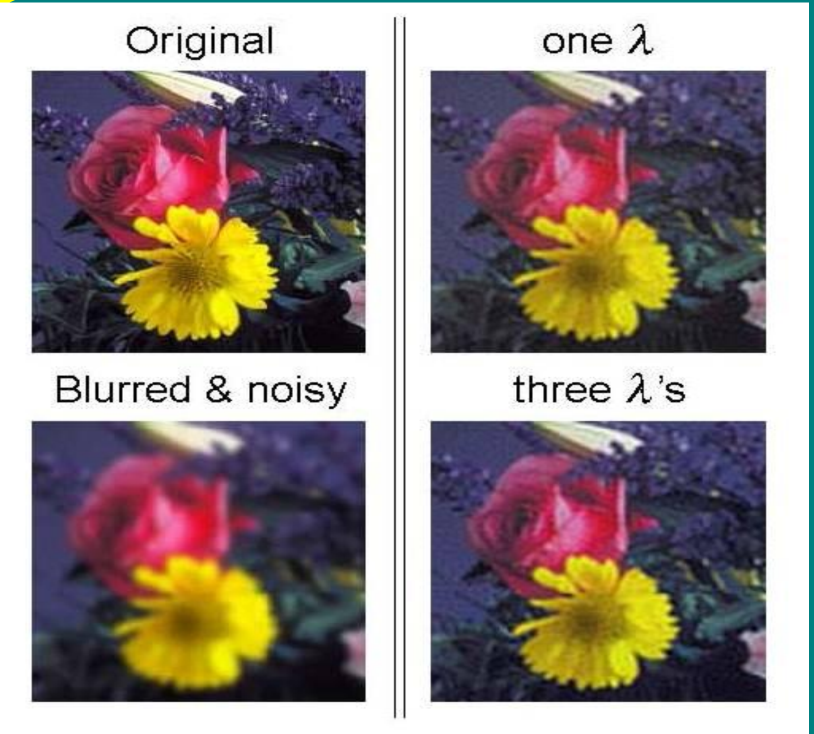
$$\boxed{K(\tau_m,\omega) = \frac{1}{\pi} \frac{\omega \exp(-\tau_m \omega)}{1 - \exp(-\beta\omega)}}$$

Image deblurring with e.g. known 2D noise K(m,ω)

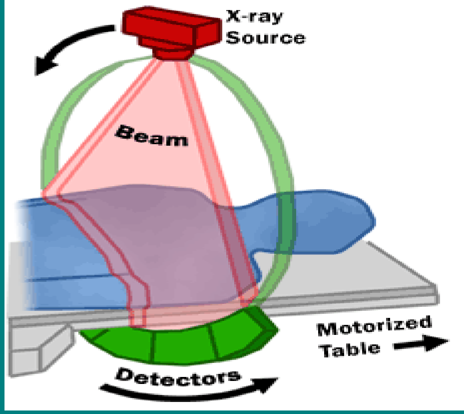
$$G(m) = \int_{-\infty}^{\infty} d\omega\; K(m,\omega)\; A(\omega)$$

m and ω are 2D vectors

K(m,ω) is a 2D x 2D noise distributon function



Tomography image reconstruction (CT scan)



$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

m and ω are 2D vectors

K(m,ω) is a 2D x 2D distribution function

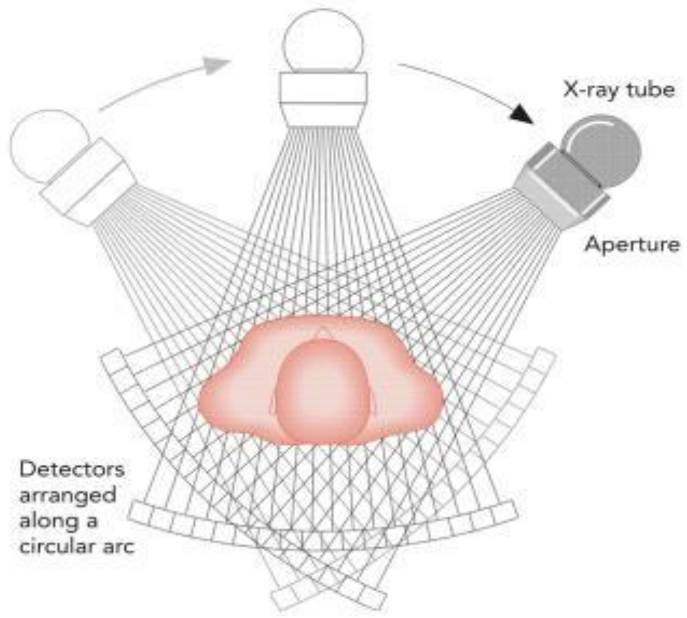


Figure 7-10 Computer tomography

$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

Aircraft stability

Nuclear reactor operation

Image deblurring

A lot of other…

What is dramatic in the problem?

ILL-POSED!

$$G(m) = \int_{-\infty}^{\infty} d\omega \; \mathcal{K}(m,\omega) \; A(\omega)$$

Aircraft stability

Nuclear reactor operation

Image deblurring

A lot of other…

What is dramatic in the problem?

ILL-POSED!

$$G(m) = \int_{-\infty}^{\infty} d\omega \; \mathcal{K}(m,\omega) \; A(\omega)$$

We cannot obtain an exact solution not because of some approximations of our approaches.

Instead, we have to admit that the exact solution does not exist at all!

ILL-POSED!

$$G(m) = \int_{-\infty}^{\infty} d\omega \; \mathcal{K}(m,\omega) \; A(\omega)$$

1. No unique solution in mathematical sense

ILL-POSED!

$$G(m) = \int_{-\infty}^{\infty} d\omega \; \mathcal{K}(m,\omega) \; A(\omega)$$

1. No unique solution in mathematical sense

2. Some additional information is required which specifies which kind of solution is expected

ILL-POSED!

$$G(m) = \int_{-\infty}^{\infty} d\omega \; \mathcal{K}(m,\omega) \; A(\omega)$$

1. ***No unique solution in mathematical sense***
   No function A to satisfy the equation

2. Some additional information is required which specifies which kind of solution is expected. In order to chose among many approximate solutions.

ILL-POSED!

How to solve?

$$G(m) = \int_{-\infty}^{\infty} d\omega \, K(m,\omega) \, A(\omega)$$

Physics department:
Max Ent.

ILL-POSED!

How to solve?

$$G(m) = \int_{-\infty}^{\infty} d\omega\ \mathcal{K}(m,\omega)\ A(\omega)$$

Physics department:
Max Ent.

Engineering department:
Tikhonov Regularization

ILL-POSED!

How to solve?

$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

Physics department:
Max Ent.

Engineering department:
Tikhonov Regularization

Statistical department:
ridge regression

ILL-POSED!

$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

Next player:
stochastic methods

Physics department:
Max Ent.

Statistical department:
ridge regression

Engineering department:
Tikhonov Regularization

Not settled!

- Still highly competitive field
- Many approaches developed, some specific ones are better for some specific cases
- Different approaches speak different languages, need some unified analysis
- Comparison of different methods, not just self-advertising, is needed

ILL-POSED!

$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

Next player:
stochastic
Methods
Since 1998

Physics
department:
Max Ent.
(Mark Jarrel)

Historically first: 1943:
Tikhonov
Regularization

$$G(m) = \int_{-\infty}^{\infty} d\omega\, \mathcal{K}(m,\omega)\, A(\omega)$$

ILL-POSED!

The easiest way to explain the problem is to turn to a discrete form of the Fredholm equation

Approximating the spectral function by its values on a finite spectral mesh of N points

$$A(\omega) = \sum_{n=1}^{N} A(\omega_n)\delta(\omega - \omega_n) \ ,$$

the integral equation (2) can be rewritten in matrix form

$$G(m) = \sum_{n=1}^{N} \mathcal{K}(m, \omega_n) A(\omega_n) \ , \ m = 1, \ldots, M \ ,$$

or equivalently presented as

$$\vec{G} = \widehat{\mathcal{K}} \vec{A} \ .$$

$$G(m) = \int_{-\infty}^{\infty} d\omega \, K(m,\omega) \, A(\omega)$$

ILL-POSED!

The easiest way to explain the problem is to turn to a discrete form of the Fredholm equation

$$G(m) = \sum_{n=1}^{N} K(m,\omega_n) A(\omega_n) \ , \ m = 1, \ldots, M$$

Because of noise present in the input data G(m) there is no unique $A(\omega_n)=A(n)$ which exactly satisfies the equation.

ILL - POSED!

$$G(m) = \int_{-\infty}^{\infty} d\omega \; \mathcal{K}(m,\omega) \; A(\omega)$$

$$G(m) = \sum_{n=1}^{N} \mathcal{K}(m,\omega_n) A(\omega_n) \;,\; m = 1,\ldots,M$$

Because of noise present in the input data G(m) there is no unique $A(\omega_n)=A(n)$ which exactly satisfies the equation.

Hence, one *can* search for the *least-square* fitted solution A(n) which minimizes:

$$\parallel \widehat{\mathcal{K}} \vec{A} - \vec{G} \parallel^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} \mathcal{K}(m,\omega_n) A(\omega_n) - G(m) \right|^2$$

$$G(m) = \int_{-\infty}^{\infty} d\omega \; \mathcal{K}(m,\omega) \; A(\omega)$$

ILL-POSED!

$$\| \widehat{\mathcal{K}}\vec{A} - \vec{G} \|^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} \mathcal{K}(m,\omega_n) A(\omega_n) - G(m) \right|^2$$

Choosing the Euclidean norm one admits the absence of unique solution because there is an infinite number of deviation norms.

$$\| \widehat{K}\vec{A} - \vec{G} \|^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} K(m,\omega_n) A(\omega_n) - G(m) \right|^2$$

ILL - POSED!

Unique solution for the least-square fit through singular values decomposition of the matrix K

$$\widehat{K} = \sum_{i=1}^{r} \sigma_i \, \vec{u}_i \otimes \vec{v}_i^{\dagger} \,,$$

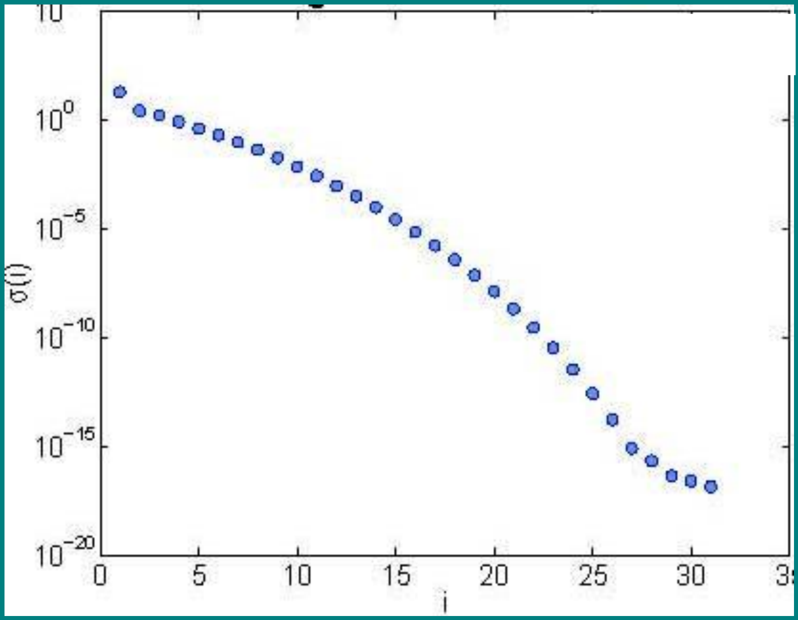
$$\Sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \\ 0 & & 0 \end{bmatrix}, \qquad \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$$

$$\| \widehat{K}\vec{A} - \vec{G} \|^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} K(m,\omega_n)A(\omega_n) - G(m) \right|^2$$

ILL-POSED!

Unique solution for the least-square fit through singular values decomposition of the matrix K

Typical singular values:



Explicit expression:

$$\vec{A} = \sum_{i=1}^{r} \frac{\vec{u}_i^{\dagger} \otimes \vec{v}_i}{\sigma_i} \vec{G}$$

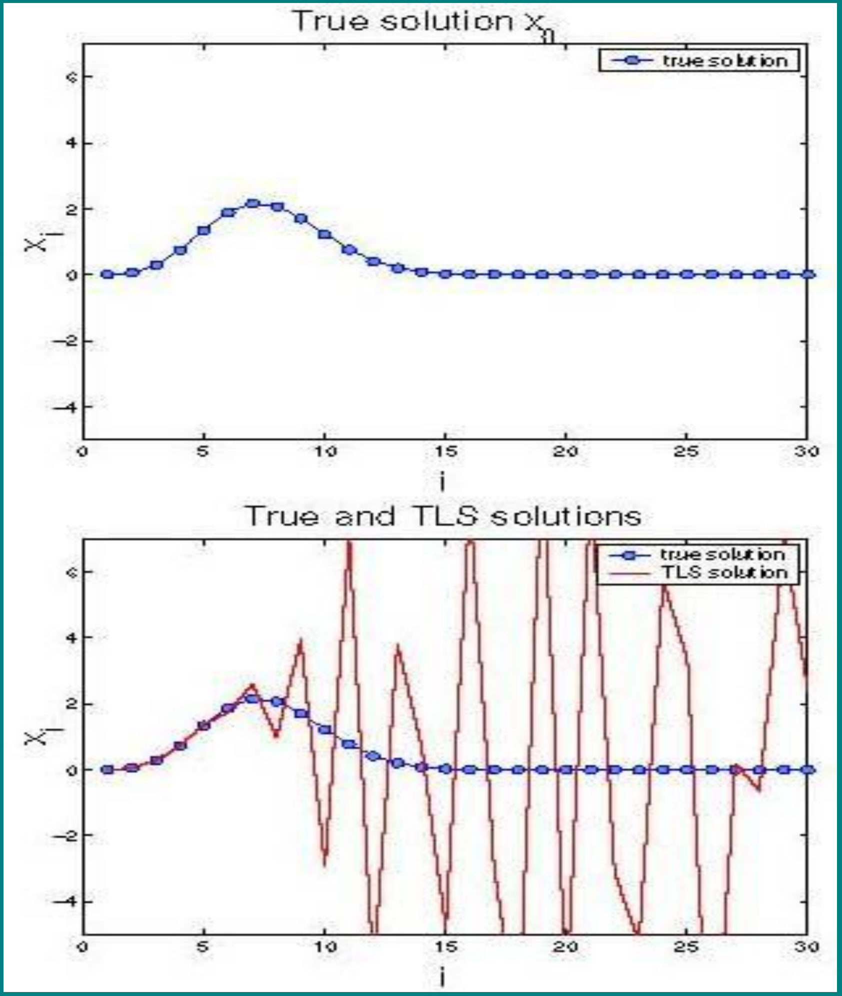
$$\| \widehat{\mathcal{K}}\vec{A} - \vec{G} \|^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} \mathcal{K}(m,\omega_n)A(\omega_n) - G(m) \right|^2$$

ILL - POSED!

Saw tooth noise

instability

due to small singular values.



Explicit expression:

$$\boxed{\vec{A} = \sum_{i=1}^{r} \frac{\vec{u}_i^{\dagger} \otimes \vec{v}_i}{\sigma_i} \vec{G}}$$

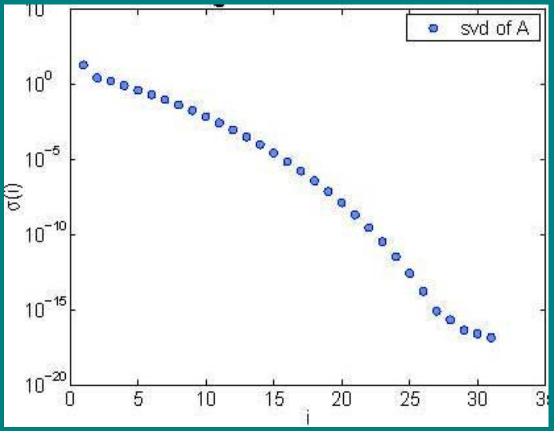
$$\| \widehat{K}\vec{A} - \vec{G} \|^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} K(m,\omega_n) A(\omega_n) - G(m) \right|^2$$

ILL - POSED!

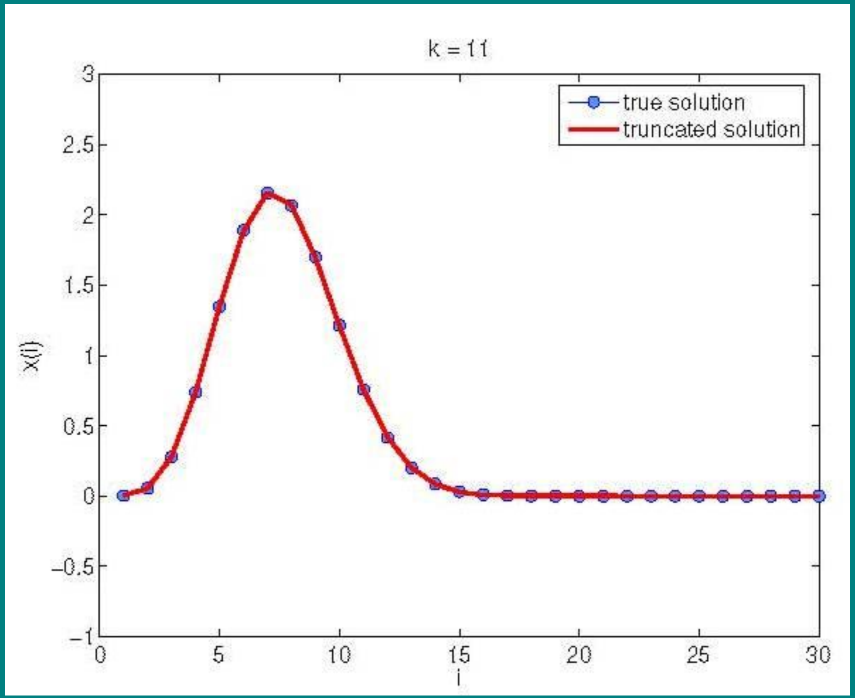
Saw tooth noise

instability

due to small singular values.



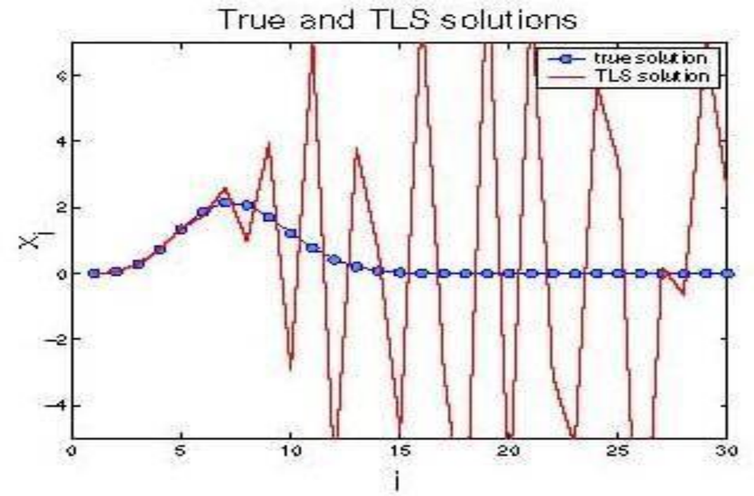
Truncating small singuar values (from 1 to 11)



$$\| \widehat{K}\vec{A} - \vec{G} \|^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} K(m,\omega_n) A(\omega_n) - G(m) \right|^2$$

ILL - POSED!

Tikhonov regularization to fight with the saw tooth noise instability.



$$\vec{A} = \sum_{i=1}^{r} \frac{\vec{u}_i^{\dagger} \otimes \vec{v}_i}{\sigma_i} \vec{G} \rightarrow \vec{A} = \sum_{i=1}^{r} \boxed{\left\{ \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \right\}} \frac{\vec{u}_i^{\dagger} \otimes \vec{v}_i}{\sigma_i} \vec{G}$$

$$\| \widehat{K}\vec{A} - \vec{G} \|^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} K(m,\omega_n) A(\omega_n) - G(m) \right|^2$$

ILL - POSED!

Tikhonov regularization to fight with the **saw tooth noise** instability.

Filter factors

$$\vec{A} = \sum_{i=1}^{r} \frac{\vec{u}_i^\dagger \otimes \vec{v}_i}{\sigma_i} \vec{G} \rightarrow \vec{A} = \sum_{i=1}^{r} \boxed{\left\{ \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \right\}} \frac{\vec{u}_i^\dagger \otimes \vec{v}_i}{\sigma_i} \vec{G}$$

$$\| \widehat{\mathcal{K}}\vec{A} - \vec{G} \|^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} \mathcal{K}(m,\omega_n)A(\omega_n) - G(m) \right|^2$$

ILL-POSED!

Tikhonov functional to minimize (Γ is unitary matrix):

$$\| \widehat{\mathcal{K}}\vec{A} - \vec{G} \|^2 + \lambda^2 \| \hat{\Gamma}\vec{A} \|^2$$

$$\vec{A} = \sum_{i=1}^{r} \frac{\vec{u}_i^\dagger \otimes \vec{v}_i}{\sigma_i}\vec{G} \quad \longrightarrow \quad \vec{A} = \sum_{i=1}^{r} \boxed{\left\{ \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \right\}} \frac{\vec{u}_i^\dagger \otimes \vec{v}_i}{\sigma_i}\vec{G}$$

ILL-POSED!

$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

General formulation of methods to solve ill-posed problems in terms of Bayesian statistical inference.

$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

Bayes theorem:

P[A|G] P[G] = P[G|A] P[A]

P[A|G] – conditional probability that the spectral function is A provided the correlation function is G

$$G(m) = \int_{-\infty}^{\infty} d\omega\, K(m,\omega)\, A(\omega)$$

Bayes theorem:

P[A|G] P[G] = P[G|A] P[A]

$\boxed{\text{P[A|G]}}$ – conditional probability that the spectral function is A provided the correlation function is G

To find it is just the analytic continuation

$$G(m) = \int_{-\infty}^{\infty} d\omega\, K(m,\omega)\, A(\omega)$$

$\boxed{P[A|G]}$ ~ P[G|A] P[A]

P[G|A] is ***easier*** problem of finding G given A: likelihood function

P[A] is prior knowledge about A:

Analytic continuation

$$G(m) = \int_{-\infty}^{\infty} d\omega \, K(m,\omega) \, A(\omega)$$

P[A|G] ~ P[G|A] P[A]

P[G|A] is easier problem of finding G given A: likelihood function

P[A] is prior knowledge about A:

All methods to solve the above problem can be formulated in terms of this relation

$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

Historically first method to solve the problem of Fredholm kind I integral equation.

Tikhonov regularization method (1943)

A.N.Tikhonov, Dokladyu Akdemii Nauk SSSR (1943)

A.N.Tikhonov, Dokladyu Akdemii Nauk SSSR (1963)
(Soviet mathematics)

Tikhonov & Arsenin, Solution of Ill-posed problems, (Washington, 1977).

$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

Historically first method to solve the problem of Fredholm kind I integral equation.

A.N.Tikhonov, Dokladyu Akdemii Nauk SSSR (1943)

Tikhonov & Arsenin, Solution of Ill-posed problems, (Washington, 1977).

The regularization method was developed not by …… in 1977, it was rediscovered….

$$G(m) = \int_{-\infty}^{\infty} d\omega\ \mathcal{K}(m,\omega)\ A(\omega)$$

Historically first method to solve the problem of Fredholm kind I integral equation.

Tikhonov regularization method (1943)

$$P[G|A] \sim \exp\{- \parallel \widehat{\mathcal{K}}\vec{A} - \vec{G} \parallel^2\}$$

$$P[A] \sim \exp\{-\lambda^2 \parallel \hat{\Gamma}\vec{A} \parallel^2\}$$

$$\parallel \widehat{\mathcal{K}}\vec{A} - \vec{G} \parallel^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} \mathcal{K}(m,\omega_n)A(\omega_n) - G(m) \right|^2$$

Tikhonov regularization method (1943)

If Γ is unit matrix:

$$\vec{A} = \sum_{i=1}^{r} \left\{ \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \right\} \frac{\vec{u}_i^{\dagger} \otimes \vec{v}_i}{\sigma_i} \vec{G}$$

$$P[G|A] \sim \exp\{- \parallel \widehat{\mathcal{K}}\vec{A} - \vec{G} \parallel^2\}$$

$$P[A] \sim \exp\{-\lambda^2 \parallel \hat{\Gamma}\vec{A} \parallel^2\}$$

$$\parallel \widehat{\mathcal{K}}\vec{A} - \vec{G} \parallel^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} \mathcal{K}(m, \omega_n) A(\omega_n) - G(m) \right|^2$$

$$\| \widehat{\mathcal{K}}\vec{A} - \vec{G} \|^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} \mathcal{K}(m,\omega_n) A(\omega_n) - G(m) \right|^2$$

ILL - POSED!

$$\boxed{\| \widehat{\mathcal{K}}\vec{A} - \vec{G} \|^2 + \lambda^2 \| \hat{\Gamma}\vec{A} \|^2}$$

$$P[G|A] \sim \exp\{- \| \widehat{\mathcal{K}}\vec{A} - \vec{G} \|^2\}$$

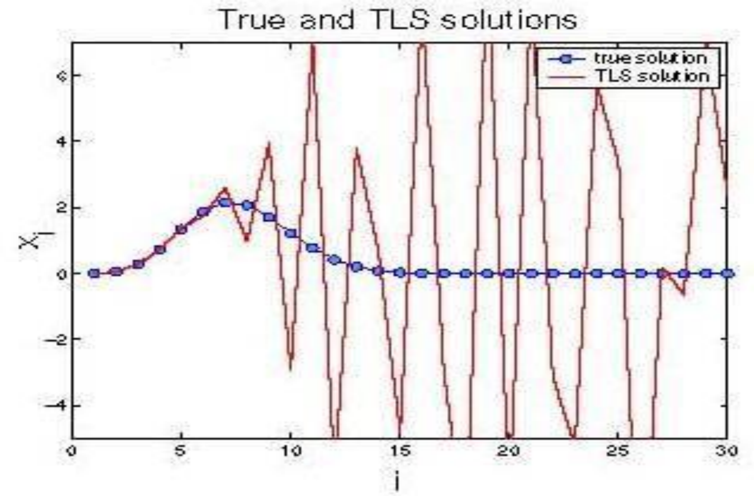
$$P[A] \sim \exp\{-\lambda^2 \| \hat{\Gamma}\vec{A} \|^2\}$$

Filter factors

$$\vec{A} = \sum_{i=1}^{r} \frac{\vec{u}_i^\dagger \otimes \vec{v}_i}{\sigma_i} \vec{G} \;\rightarrow\; \vec{A} = \sum_{i=1}^{r} \boxed{\left\{ \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \right\}} \frac{\vec{u}_i^\dagger \otimes \vec{v}_i}{\sigma_i} \vec{G}$$

$$\| \hat{K}\vec{A} - \vec{G} \|^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} K(m,\omega_n)A(\omega_n) - G(m) \right|^2$$

ILL-POSED!



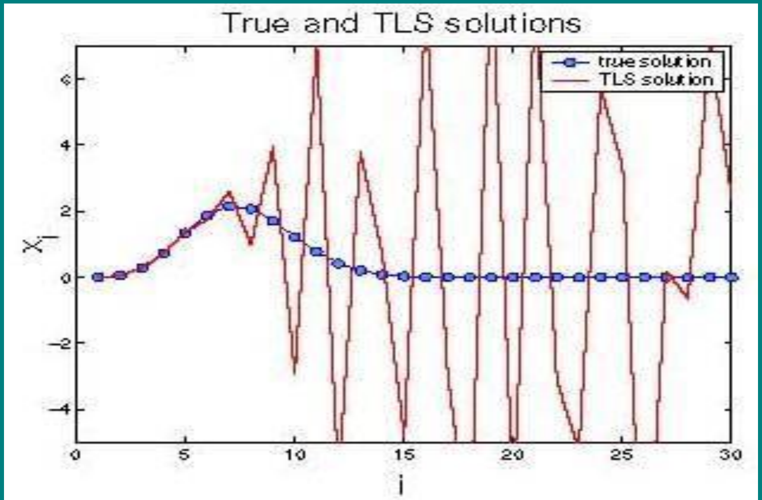
Tikhonov regularization to fight with the **saw tooth noise** instability.

$$\vec{A} = \sum_{i=1}^{r} \frac{\vec{u}_i^\dagger \otimes \vec{v}_i}{\sigma_i} \vec{G} \rightarrow \vec{A} = \sum_{i=1}^{r} \boxed{\left\{ \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \right\}} \frac{\vec{u}_i^\dagger \otimes \vec{v}_i}{\sigma_i} \vec{G}$$

Is it not too smooth???

ILL-POSED!

Tikhonov regularization to fight with the **saw tooth noise** instability.



$$\vec{A} = \sum_{i=1}^{r} \frac{\vec{u}_i^{\dagger} \otimes \vec{v}_i}{\sigma_i} \vec{G} \;\rightarrow\; \vec{A} = \sum_{i=1}^{r} \left\{ \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \right\} \frac{\vec{u}_i^{\dagger} \otimes \vec{v}_i}{\sigma_i} \vec{G}$$

Maximum entropy method

P[A|G] ~ P[G|A] P[A]

$$P[G|\widetilde{A}] = \exp\{-\chi^2[\widetilde{A}]/2\} \ ,$$

$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

Likelihood (objective) function

$$P[\quad A] = \exp\{\alpha^{-1} S[\widetilde{A}]\} \ ,$$

$$S[\widetilde{A}] = \int d\omega \ \widetilde{A}(\omega) \ln[\widetilde{A}(\omega)/D(\omega)]$$

Prior knowledge function

Maximum entropy method

P[A|G] ~ P[G|A] P[A]

D(ω) is default model

$$P[\quad A] = \exp\{\alpha^{-1} S[\widetilde{A}]\} \ ,$$

$$S[\widetilde{A}] = \int d\omega \ \widetilde{A}(\omega) \ln[\widetilde{A}(\omega)/D(\omega)]$$

Prior knowledge function

Maximum entropy method

P[A|G] ~ P[G|A] P[A]

1. **One has escaped extra smoothening.**

2. **But one has got default model as an extra price.**

$$P[\quad A] = \exp\{\alpha^{-1} S[\widetilde{A}]\}\ ,$$

$$S[\widetilde{A}] = \int d\omega\ \widetilde{A}(\omega) \ln[\widetilde{A}(\omega)/D(\omega)]$$

Prior knowledge function

Maximum entropy method

P[A|G] ~ P[G|A] P[A]

1. We want to avoid extra smoothening.

2. We want to avoid default model as an extra price.

$$P[\quad A] = \exp\{\alpha^{-1} S[\widetilde{A}]\} \, ,$$

$$S[\widetilde{A}] = \int d\omega \, \widetilde{A}(\omega) \ln[\widetilde{A}(\omega)/D(\omega)]$$

Prior knowledge function

Maximum entropy method

P[A|G] ~ P[G|A] P[A]

Both items (extra smoothening and arbitrary default model) can be somehow circumvented by the group of stochastic methods.

$$P[\quad A] = \exp\{\alpha^{-1} S[\widetilde{A}]\}\ ,$$

$$S[\widetilde{A}] = \int d\omega\ \widetilde{A}(\omega) \ln[\widetilde{A}(\omega)/D(\omega)]$$

Prior knowledge function

Stochastic methods

$$P[A|G] \sim P[G|A]\, P[A]$$

Both items (extra smoothening and arbitrary default model) can be somehow circumvented by the group of stochastic methods.

Stochastic methods

$$P[A|G] \sim P[G|A]\, P[A]$$

The main idea of the stochastic methods is:

1. Restrict the prior knowledge to the minimal possible level (positive, normalized, etc…).

2. Change the likelihood function to the likelihood functional.

Stochastic methods

$$P[A|G] \sim P[G|A]\,P[A]$$

The main idea of the stochastic methods is:

1. Restrict the prior knowledge to the minimal possible level (positive, normalized, etc…). **Avoids default model.**

2. Change the likelihood function to the likelihood functional. **Avoids saw-tooth noise.**

Stochastic methods

P[A|G] ~ P[G|A] P[A]

Change the likelihood function to the likelihood functional. Avoids sawtooth noise.

$$A = \int d\widetilde{A}\ \widetilde{A}\ P[\widetilde{A}|G]$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\widetilde{A}]/T\}$$

$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

Stochastic methods

Likelihood functional. Avoids sawtooth noise.

$$A = \int d\widetilde{A}\; \widetilde{A}\; P[\widetilde{A}|G]$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\widetilde{A}]/\mathcal{T}\}$$

$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

Sandvik, Phys. Rev. B 1998, is the first practical attempt to think stochastically.

Stochastic methods

Likelihood functional. Avoids sawtooth noise.

$$A = \int d\widetilde{A}\ \widetilde{A}\ P[\widetilde{A}|G]$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\widetilde{A}]/\mathcal{T}\}$$

$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

Sandvik, Phys. Rev. B 1998, is the first practical attempt to think stochastically.

SOM was suggested in 2000. Mishchenko et al, Appendix B in Phys. Rev. B.

Some applications of SOM:

Phys. Rev. Lett., vol. 86, 4624 (2001)
Phys. Rev. Lett., vol. 87, 186402 (2001)
Phys. Rev. Lett., vol. 91, 236401 (2003)
Phys. Rev. Lett., vol. 93, 036402 (2004)
Phys. Rev. Lett., vol. 96, 136405 (2006)
Phys. Rev. Lett., vol. 99, 226402 (2007)
Phys. Rev. Lett., vol. 100, 166401 (2008)
Phys. Rev. Lett., vol. 101, 116403 (2008)
Phys. Rev. Lett., vol. 104, 056602 (2010)
Phys. Rev. Lett., vol. 107, 076403 (2011)

Stochastic methods

Likelihood functional. Avoids sawtooth noise.

$$A = \int d\widetilde{A}\, \widetilde{A}\, P[\widetilde{A}|G]$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\widetilde{A}]/\mathcal{T}\}$$

$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

Sandvik, Phys. Rev. B 1998, is the first practical attempt to think stochastically.

SOM was suggested in 2000. Mishchenko et al, Appendix B in Phys. Rev. B.

Stochastic methods

Likelihood functional. Avoids sawtooth noise.

$$A = \int d\widetilde{A}\; \widetilde{A}\; P[\widetilde{A}|G]$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\widetilde{A}]/T\}$$

$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

Sandvik, Phys. Rev. B 1998, is the first practical attempt to think stochastically.

Beach, 2004, cond-mat

SOM was suggested in 2000.
Mishchenko et al,
Appendix B in Phys. Rev. B.

Stochastic methods

Likelihood functional. Avoids sawtooth noise.

$$A = \int d\widetilde{A}\; \widetilde{A}\; P[\widetilde{A}|G]$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\widetilde{A}]/T\}$$

$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

Sandvik, Phys. Rev. B 1998, is the first practical attempt to think stochastically.

Beach, 2004, cond-mat

SOM was suggested in 2000. Mishchenko et al, Appendix B in Phys. Rev. B.

Other variants after 2004

Stochastic methods

Likelihood functional. Avoids sawtooth noise.

$$A = \int d\widetilde{A}\; \widetilde{A}\; P[\widetilde{A}|G]$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\widetilde{A}]/T\}$$

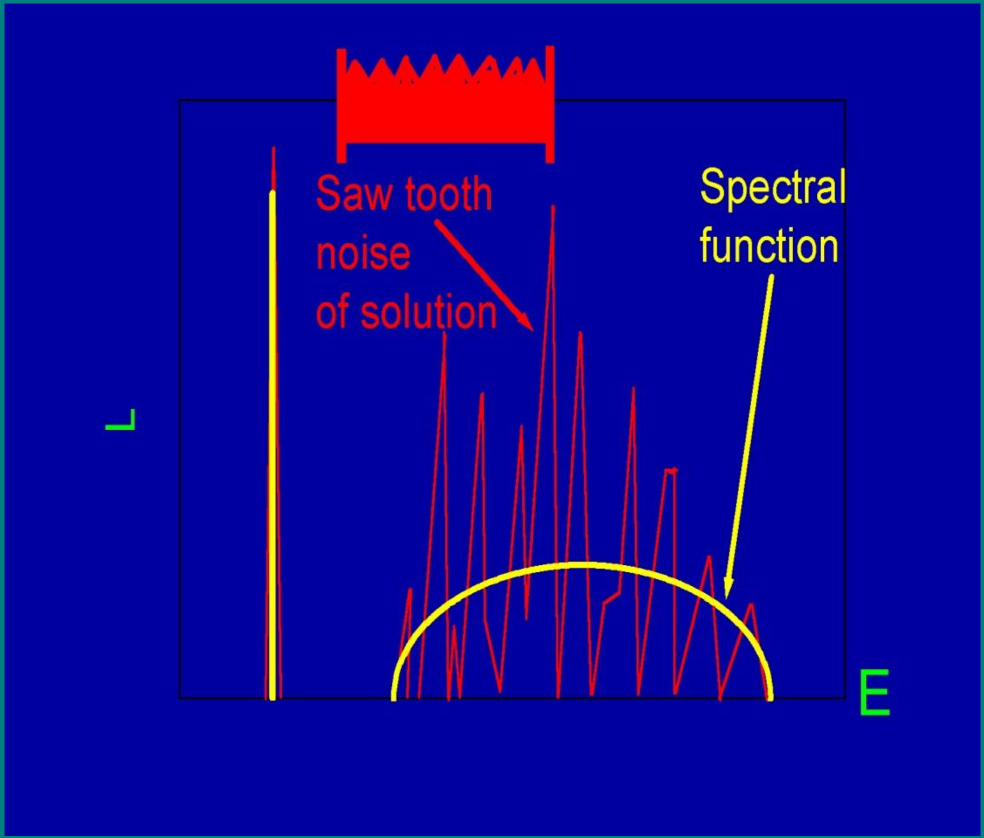
$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

What is the special need for the stochastic sampling methods?

Stochastic methods

1. Avoid saw-tooth noise.
2. Avoid over-smoothing
   of the δ-function

Typical spectrum of QP at T=0.



What is the special need for the stochastic sampling methods?

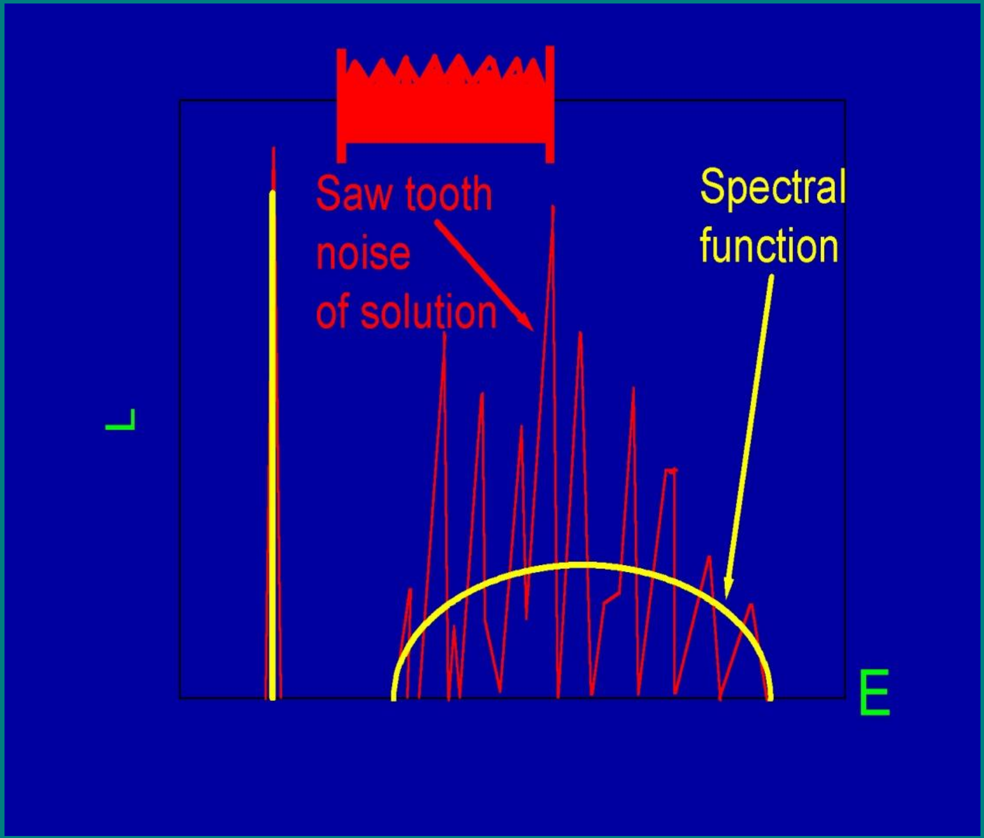
Stochastic methods

1. Avoid saw-tooth noise.
2. Avoid over-smoothing of the δ-function

Tikhonow regularization over-smoothes the δ-function.

MaxEnt – default model has to fix δ-function in advance.

Typical spectrum of QP at T=0.

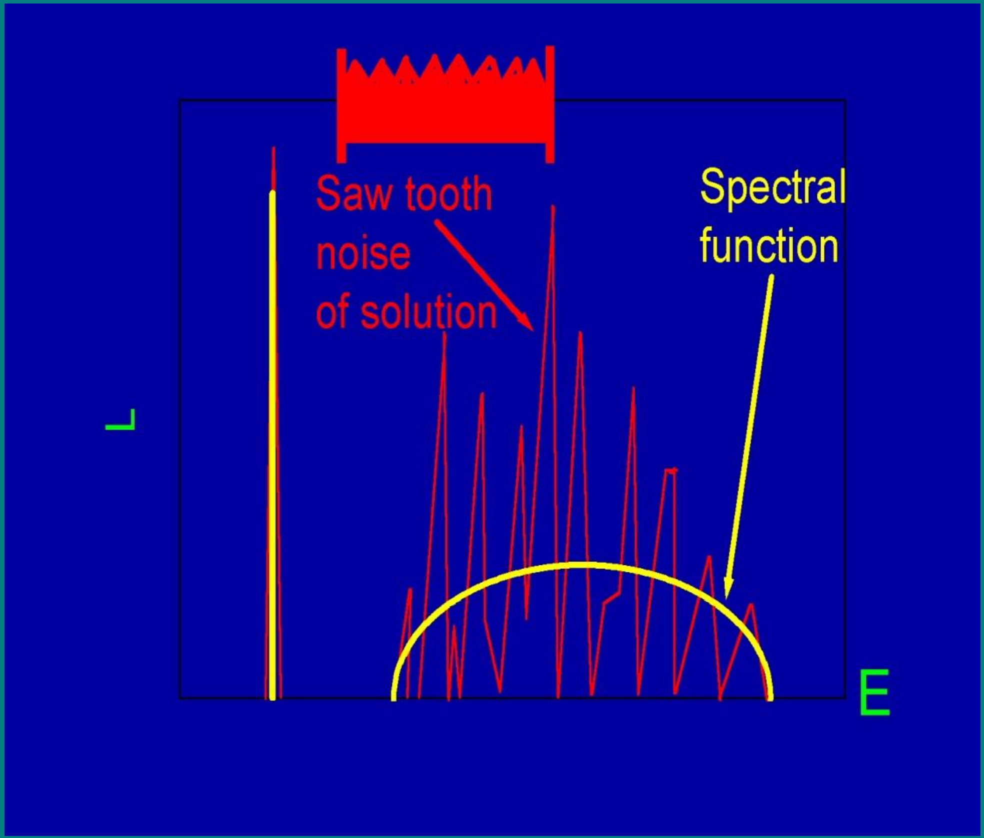


What is the special need for the stochastic sampling methods?

Stochastic methods

1. Avoid saw-tooth noise.
2. Avoid over-smoothing of the δ-function

Typical spectrum of QP at T=0.



Stochastic methods is a way to circumvent these problems.

What is the special need for the stochastic sampling methods?

Stochastic methods

$$A = \int d\widetilde{A}\ \widetilde{A}\ P[\widetilde{A}|G]$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\widetilde{A}]/T\}$$

$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

Back to Sandvik approach

Stochastic methods

$$A = \int d\widetilde{A}\ \widetilde{A}\ P[\widetilde{A}|G]$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\widetilde{A}]/T\}$$

$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

One does not search for a single solution but samples through difference "configurations" (spectral functions A) Using the likelihood function P which is characterized by fictitious "temperature" T and fictitious "energy" χ^2.

Stochastic methods

$$A = \int d\widetilde{A}\; \widetilde{A}\; P[\widetilde{A}|G]$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\widetilde{A}]/T\}$$

$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

One does not search for a single solution but samples through difference "configurations" (spectral functions A) Using the likelihood function P which is characterized by fictitious "temperature" T and fictitious "energy" χ^2.

One interprets χ^2 as an "energy" of fictitious Hamiltonian and T as a fictitious "temperature". Hence, one involves the Metropolis algorithm for Monte Carlo to sample through configurations A.

Stochastic methods

$$A = \int d\widetilde{A}\; \widetilde{A}\; P[\widetilde{A}|G]$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\widetilde{A}]/T\}$$

$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

One interprets χ^2 as an "energy" of fictitious Hamiltonian and T as a fictitious "temperature". Hence, one involves the Metropolis algorithm for Monte Carlo to sample through configurations.

1. T is not too high. Otherwise A is far from spectra which fit well the correlation function G.
2. T is not too small otherwise we are back again to the sawtooth noise problem. Ower-fitting of the noise.

Stochastic methods

$$A = \int d\widetilde{A}\ \widetilde{A}\ P[\widetilde{A}|G]\ ,$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\widetilde{A}]/T\}$$

$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

One interprets χ^2 as an "energy" of fictitious Hamiltonian and T as a fictitious "temperature". Hence, one involves the Metropolis algorithm for Monte Carlo to sample through configurations.

1. T is not too high. Otherwise A is far from spectra which fit well the correlation function G.
2. T is not too small otherwise we are back again to the sawtooth noise problem. Ower-fitting of the noise.

Simple rule T = M

Stochastic methods

$$A = \int d\tilde{A}\; \tilde{A}\; P[\tilde{A}|G]$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\tilde{A}]/T\}$$

$$\chi^2[\tilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \tilde{G}(m)]^2$$

One interprets χ^2 as an "energy" of fictitious Hamiltonian and T as a fictitious "temperature". Hence, one involves the Metropolis algorithm for Monte Carlo to sample through configurations.

Which features of Sandvik method are artificial?

1. There is no real Hamiltonian and T and, hence, one has no requirement to sample through Metropolis
2. Algorithm is not effective at low T and use the tempering procedures with sampling at different Ts.

Stochastic optimization method.

$$A = \int d\widetilde{A}\, \widetilde{A}\, P[\widetilde{A}|G]$$

1. One has to sample through solutions $\widetilde{A}(\omega)$ which fit the correlation function G well.

2. One has to make some weighted sum of these well solutions $\widetilde{A}(\omega)$.

$$A(\omega) = \sum_{j=1}^{L} \xi_j \widetilde{A}_j(\omega).$$

$$A = \int d\widetilde{A}\, \widetilde{A}\, P[\widetilde{A}|G]$$

SOM is very similar to Sandvik method but circumvent its artificial features and, as a result, turns out more effective

$$A(\omega) = \sum_{j=1}^{L} \xi_j \widetilde{A}_j(\omega).$$

Stochastic optimization method.

$$A(\omega) = \sum_{j=1}^{L} \xi_j \widetilde{A}_j(\omega).$$

One collects and averages large amount of "well" solutions and take an average.

1. What is the likelihood function (deviation measure of fit quality?
2. How the spectrum is parameterized
3. How to find one "well" solution?
4. When the number of solutions is enough?
5. Tests.

Stochastic optimization method.

$$A(\omega) = \sum_{j=1}^{L} \xi_j \widetilde{A}_j(\omega).$$

1. **What is the likelihood function (deviation measure of fit quality?**

deviation measure of SOM is given by expression

$$D[\widetilde{A}] = \sum_{m=1}^{M} |\Delta(m)| \; . \tag{29}$$

Here $\Delta(m)$ is the *deviation function*

$$\Delta(m) = \frac{G(m) - \tilde{G}(m)}{S(m)}, \tag{30}$$

which characterizes individual deviations of specific data points $G(m)$ from the values of the simulated function $\tilde{G}(m)$ defined by the particular spectral function $\widetilde{A}$ in terms of relation

$$\tilde{G}(m) = \int_{-\infty}^{\infty} d\omega \; \mathcal{K}(m,\omega) \; \widetilde{A}(\omega) \; . \tag{31}$$

Stochastic optimization method.

Parameterization of the particular solution:

We parameterize the spectral function $\tilde{A}$ as a sum

$$\tilde{A}(\omega) = \sum_{t=1}^{K} \eta_{\{P_t\}}(\omega)$$

of rectangles $\{P_t\} = \{h_t, w_t, c_t\}$

$$\eta_{\{P_t\}}(\omega) = \begin{cases} h_t \quad , & \omega \in [c_t - w_t/2, c_t + w_t/2] \, , \\ 0 \quad , & \text{otherwise} \, , \end{cases}$$

determined by height $h_t > 0$, width $w_t > 0$, and center c_t.

A configuration

$$C = \{\{P_t\}, t = 1, ..., K\}$$

with the normalization constraint

$$\sum_{t=1}^{K} h_t w_t = I,$$

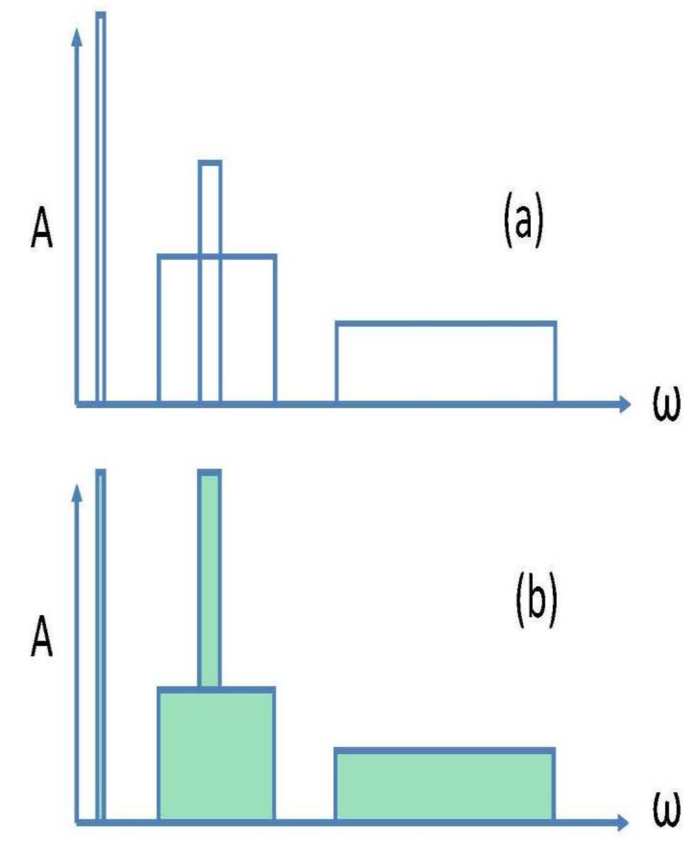


Fig. 2: *An example of a configuration with $K = 4$. Panel (b) shows how the intersection of rectangles in panel (a) is treated.*

Stochastic optimization method.

Parameterization of the particular solution:

No predefined mesh for the energy (ω) space.

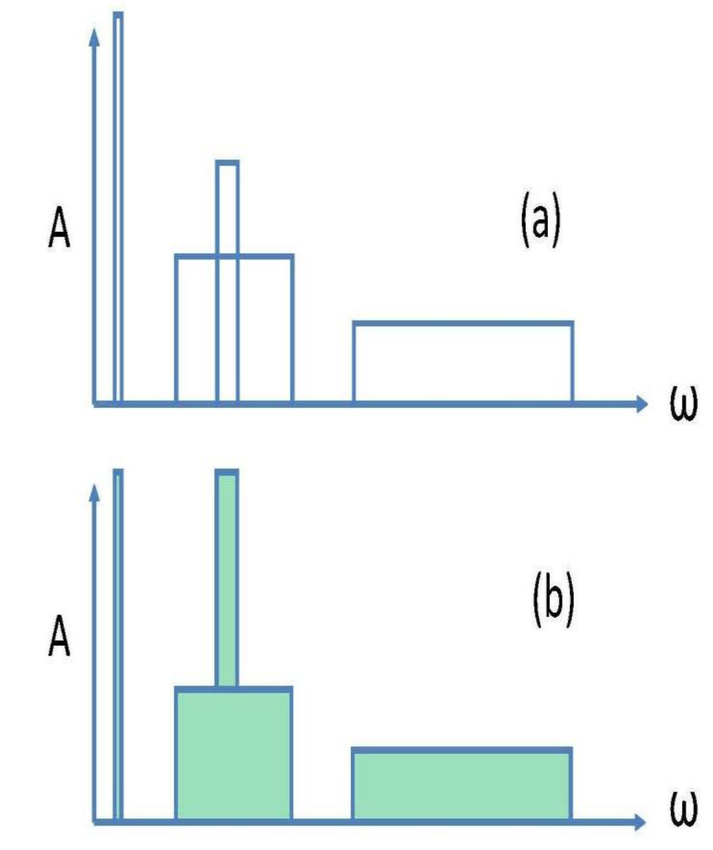


Fig. 2: *An example of a configuration with $K = 4$. Panel (b) shows how the intersection of rectangles in panel (a) is treated.*

Stochastic optimization method.

Contribution of rectangle to

$$\widetilde{G}(m) = \int_{-\infty}^{\infty} d\omega\, K(m,\omega)\, \widetilde{A}(\omega)$$

We parameterize the spectral function $\widetilde{A}$ as a sum

$$\widetilde{A}(\omega) = \sum_{t=1}^{K} \eta_{\{P_t\}}(\omega)$$

of rectangles $\{P_t\} = \{h_t, w_t, c_t\}$

$$\eta_{\{P_t\}}(\omega) = \begin{cases} h_t \quad , & \omega \in [c_t - w_t/2, c_t + w_t/2] \,, \\ 0 \quad , & \text{otherwise} \,, \end{cases}$$

determined by height $h_t > 0$, width $w_t > 0$, and center c_t.

A configuration

$$C = \{\{P_t\}, t = 1, ..., K\}$$

with the normalization constraint

$$\sum_{t=1}^{K} h_t w_t = I,$$

If no analytic expression.

One tabulates:

$$\Lambda(m,\Omega) = \int_{-\infty}^{\Omega} K(m,x)\, dx \;, m = 1, \ldots, M$$

Contribution:

$$\tilde{G}(m) = \sum_{t=1}^{K} h_t \left[\Lambda(m, c_t + w_t/2) - \Lambda(m, c_t - w_t/2)\right]$$

Stochastic optimization method.

Contribution of rectangle to

$$\widetilde{G}(m) = \int_{-\infty}^{\infty} d\omega \; \mathcal{K}(m,\omega) \; \widetilde{A}(\omega)$$

We parameterize the spectral function $\widetilde{A}$ as a sum

$$\widetilde{A}(\omega) = \sum_{t=1}^{K} \eta_{\{P_t\}}(\omega)$$

of rectangles $\{P_t\} = \{h_t, w_t, c_t\}$

$$\eta_{\{P_t\}}(\omega) = \begin{cases} h_t & , \quad \omega \in [c_t - w_t/2, c_t + w_t/2] \, , \\ 0 & , \quad \text{otherwise} \, , \end{cases}$$

determined by height $h_t > 0$, width $w_t > 0$, and center c_t.

A configuration

$$C = \{\{P_t\}, t = 1, ..., K\}$$

with the normalization constraint

$$\sum_{t=1}^{K} h_t w_t = I,$$

Particular cases.

Imaginary time, T=0:

Kernel is

$K(m,\omega) = \exp(-i\tau_m\omega)$

Contribution:

$$\tilde{G}_C(\tau_m) = \begin{cases} I & , \quad \tau_m = 0 \, , \\ 2\tau_m^{-1} \sum_{t=1}^{K} h_t e^{-c_t \tau_m} \sinh(w_t \tau_m / 2) & , \quad \tau_m \neq 0 \, . \end{cases}$$

Stochastic optimization method.

Contribution of rectangle to

$$\widetilde{G}(m) = \int_{-\infty}^{\infty} d\omega\; \mathcal{K}(m,\omega)\; \widetilde{A}(\omega)$$

We parameterize the spectral function $\widetilde{A}$ as a sum

$$\widetilde{A}(\omega) = \sum_{t=1}^{K} \eta_{\{P_t\}}(\omega)$$

of rectangles $\{P_t\} = \{h_t, w_t, c_t\}$

$$\eta_{\{P_t\}}(\omega) = \begin{cases} h_t & , \quad \omega \in [c_t - w_t/2, c_t + w_t/2]\,, \\ 0 & , \quad \text{otherwise}\,, \end{cases}$$

determined by height $h_t > 0$, width $w_t > 0$, and center c_t.

A configuration

$$C = \{\{P_t\}, t = 1, ..., K\}$$

with the normalization constraint

$$\sum_{t=1}^{K} h_t w_t = I,$$

Particular cases.

Matsubara, any T:

Kernel is

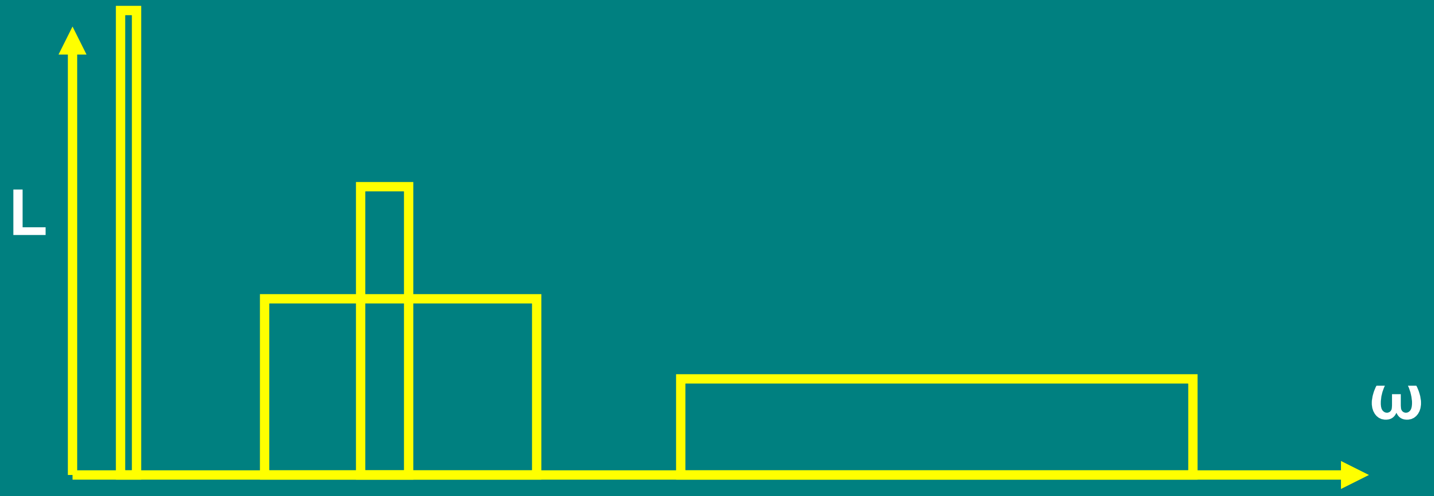
$$\mathcal{K}(i\omega_m, \omega) = \pm \frac{1}{i\omega_m - \omega}$$

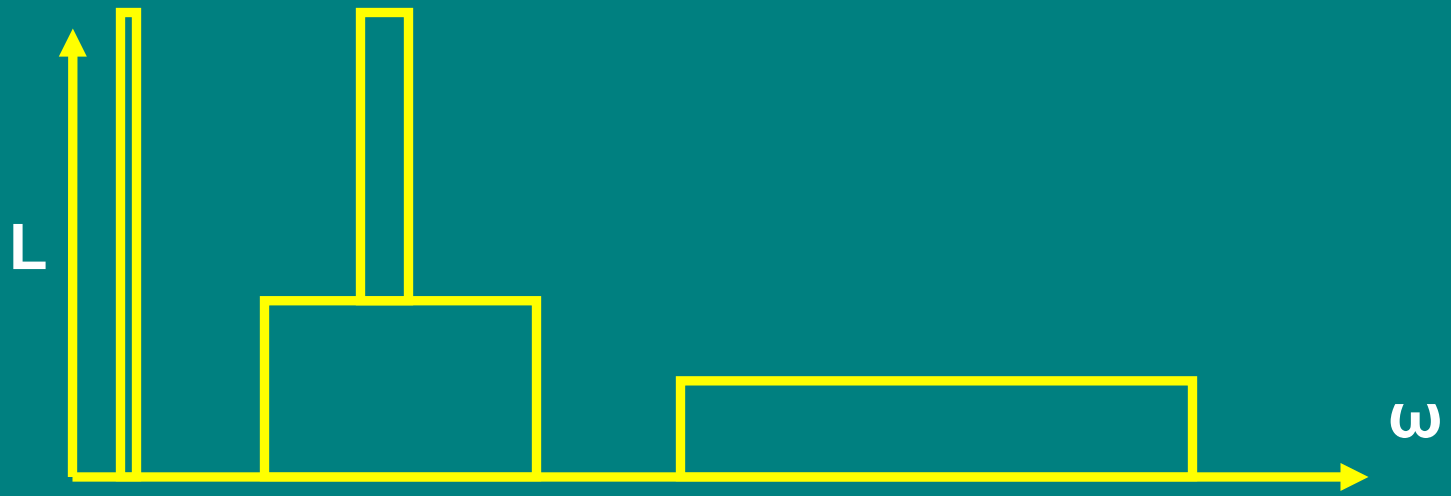
Contribution:

$$\tilde{\mathcal{G}}_C(i\omega_m) = \pm \sum_{t=1}^{K} h_t \ln \left[\frac{c_t - w_t/2 - i\omega_m}{c_t + w_t/2 - i\omega_m} \right]$$

Stochastic Optimization method.

- Particular solution $L^{(i)}(\omega)$ for LSF is presented as a sum of a number K of rectangles with some width, height and center.



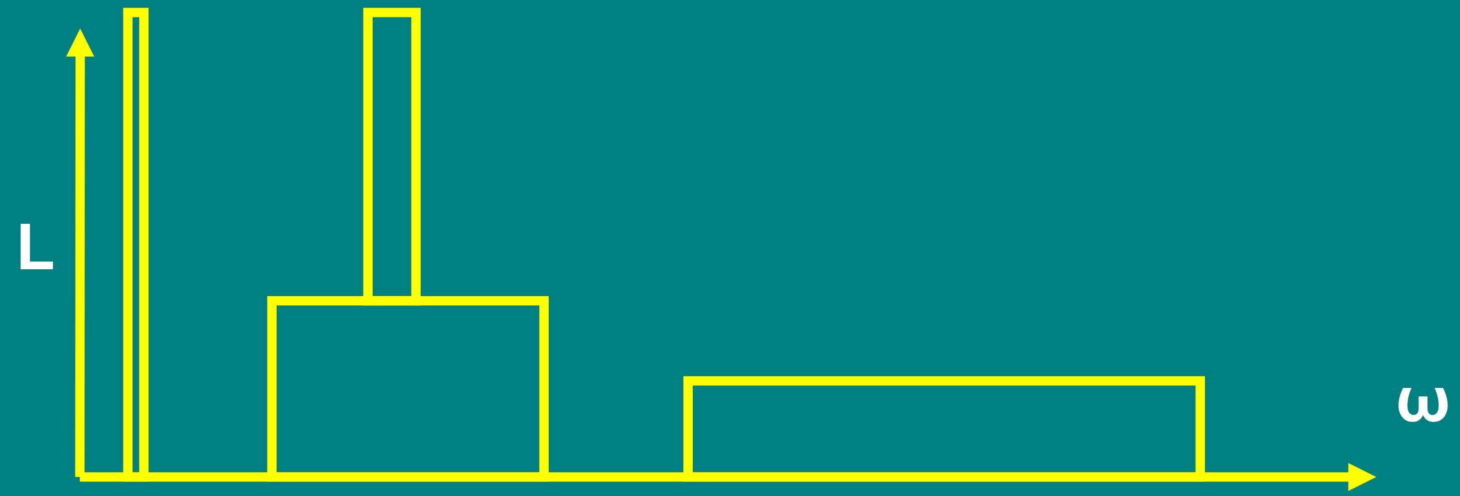


$$A(\omega) = \sum_{j=1}^{L} \xi_j \widetilde{A}_j(\omega).$$

How to find one of solutions?

Stochastic Optimization method.

- Particular solution $L^{(i)}(\omega)$ for LSF is presented as a sum of a number K of rectangles with some width, height and center.
- Initial configuration of rectangles is created by random number generator (i.e. number K and all parameters of of rectangles are randomly generated).



Stochastic Optimization method.

- Particular solution $L^{(i)}(\omega)$ for LSF is presented as a sum of a number K of rectangles with some width, height and center.
- Initial configuration of rectangles is created by random number generator (i.e. number N and all parameters of of rectangles are randomly generated).
- Each particular solution $L^{(i)(}\omega)$ is obtained by a naïve method without regularization (though, varying number K).

Deviation measure for configuration:

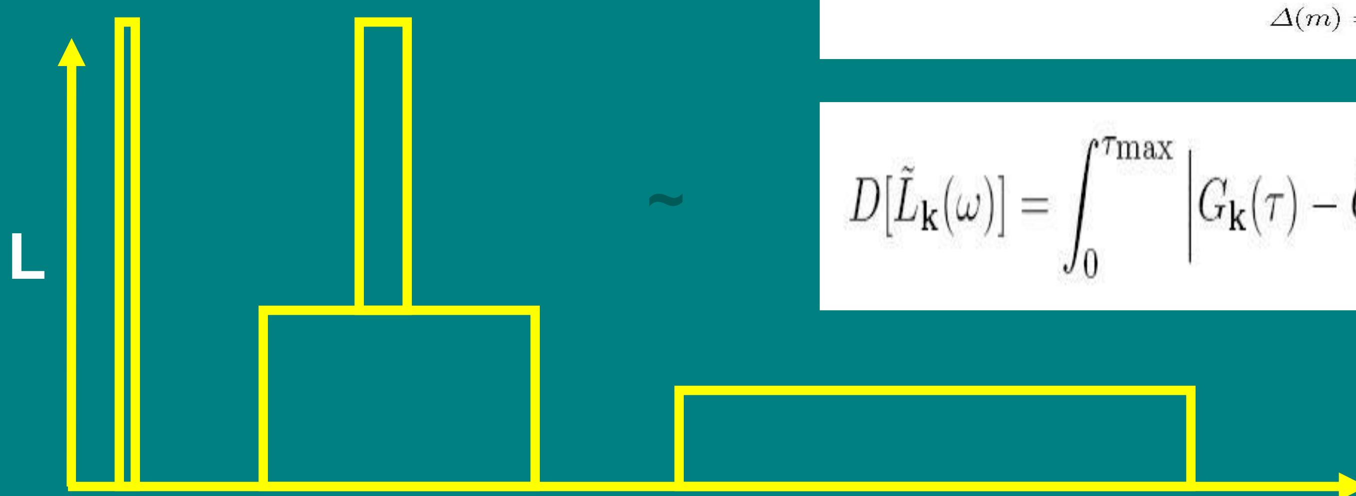
deviation measure of SOM is given by expression

$$D[\widetilde{A}] = \sum_{m=1}^{M} |\Delta(m)| \ .$$

Here $\Delta(m)$ is the *deviation function*

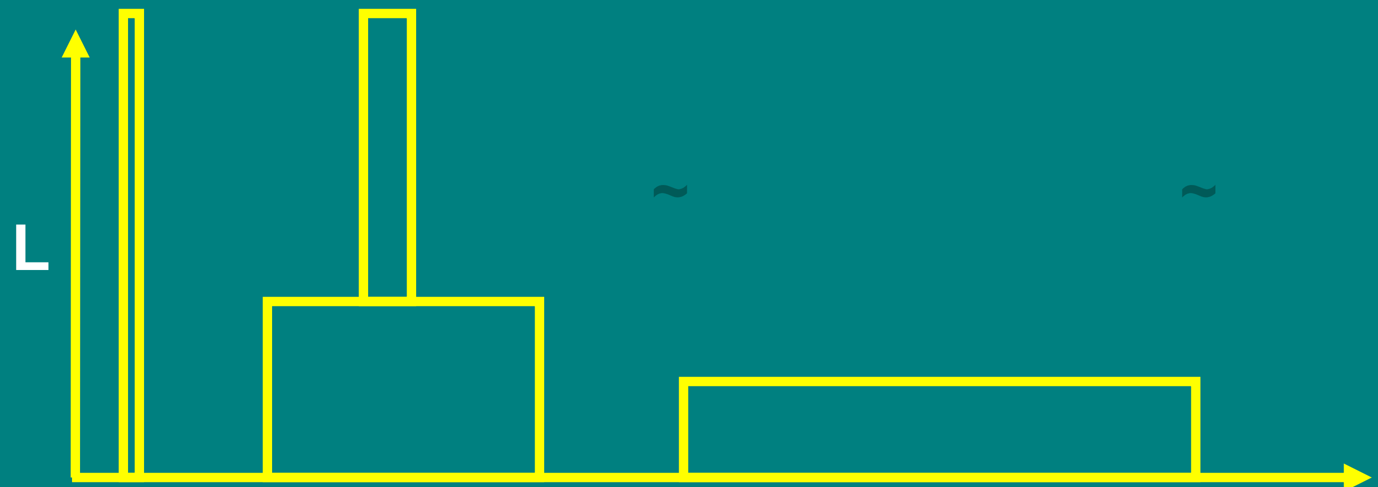
$$\Delta(m) = \frac{G(m) - \tilde{G}(m)}{S(m)},$$

$$D[\tilde{L}_{\mathrm{k}}(\omega)] = \int_0^{\tau_{\max}} \left| G_{\mathrm{k}}(\tau) - \tilde{G}_{\mathrm{k}}(\tau) \right| G_{\mathrm{k}}^{-1}(\tau) d\tau$$



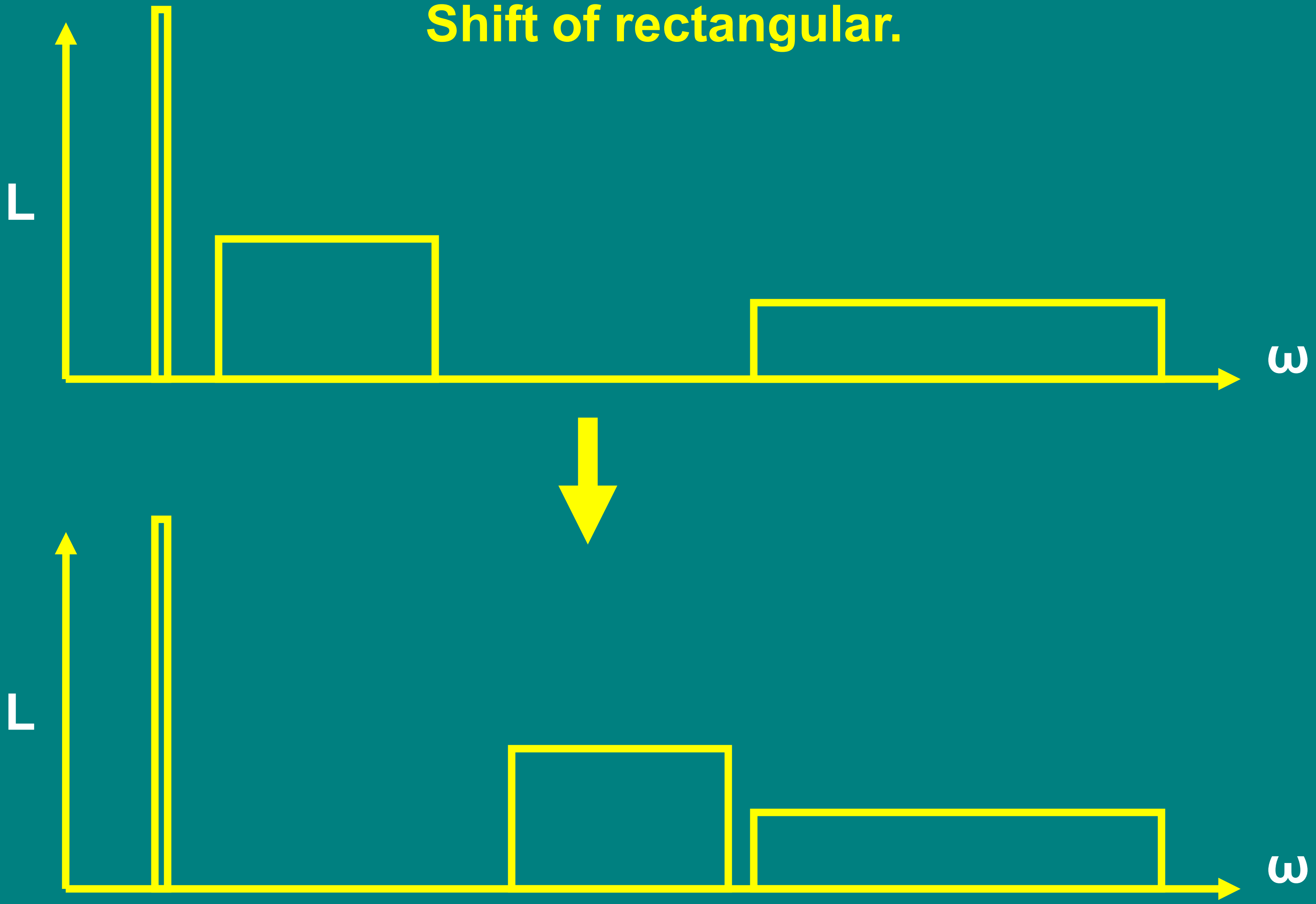
How to minimize the deviation?

Which updates?



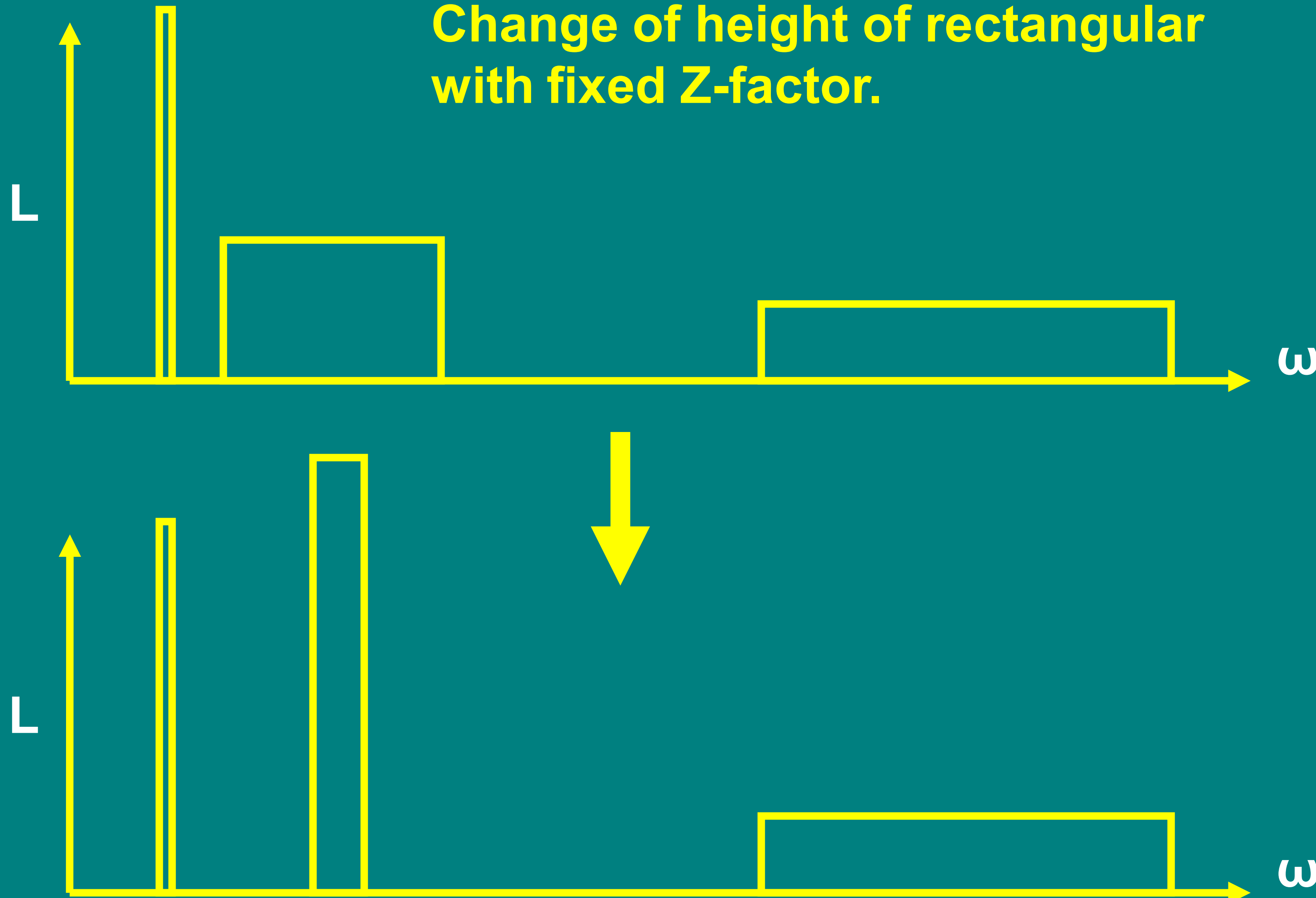
Stochastic Optimization method: update procedures.

Shift of rectangular.



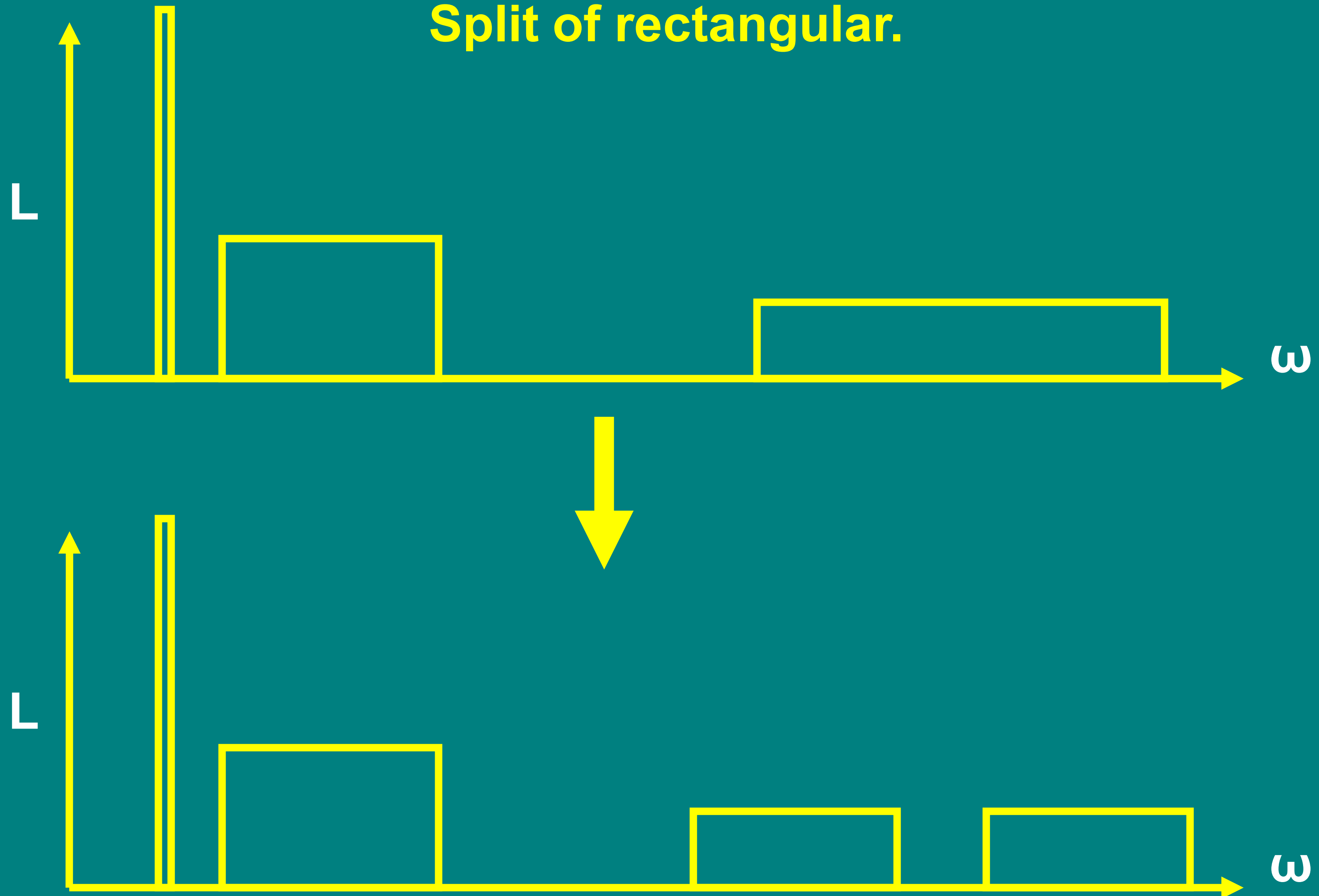
Stochastic Optimization method: update procedures.

Change of height of rectangular with fixed Z-factor.



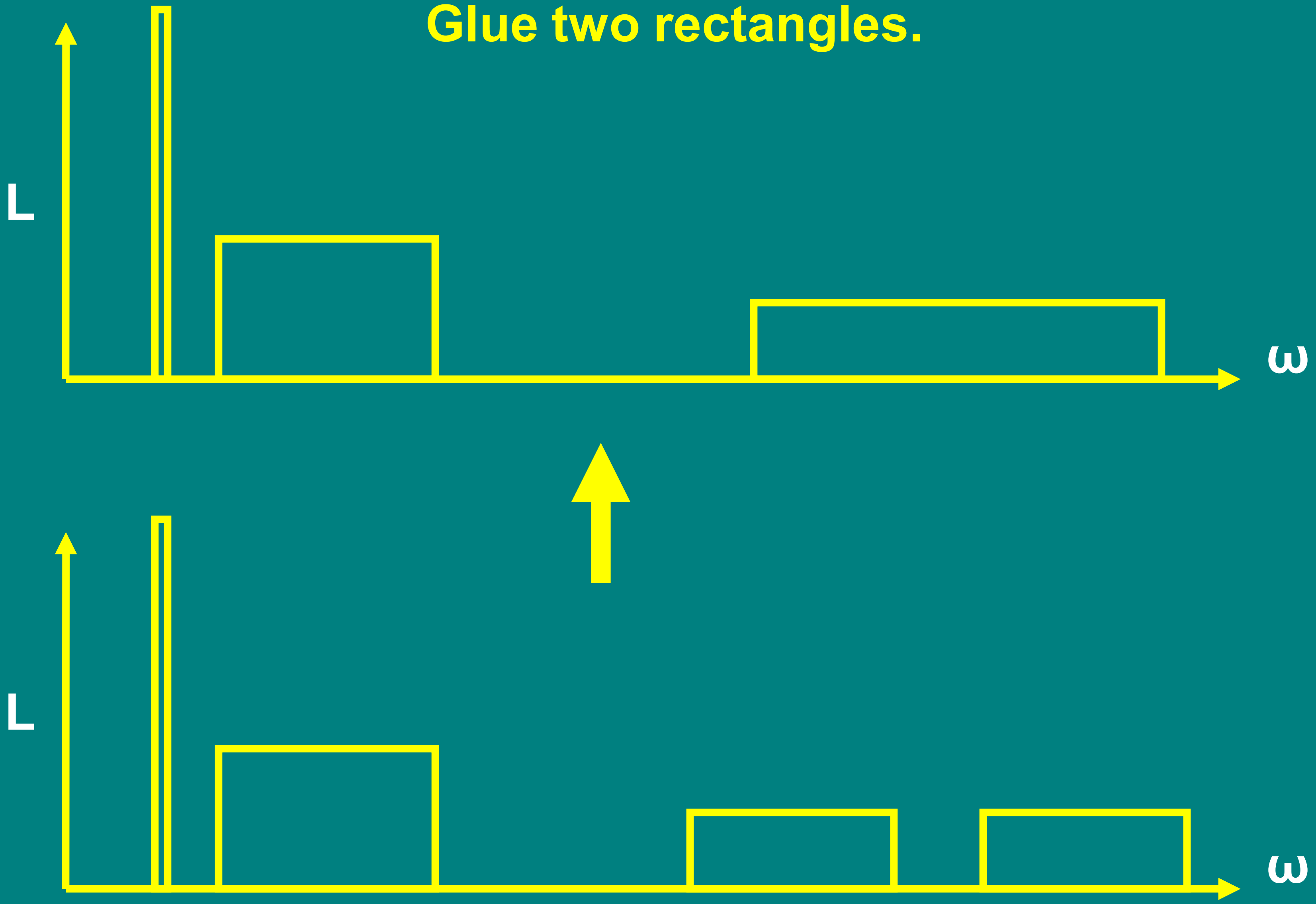
Stochastic Optimization method: update procedures.

Split of rectangular.



Stochastic Optimization method: update procedures.

Glue two rectangles.

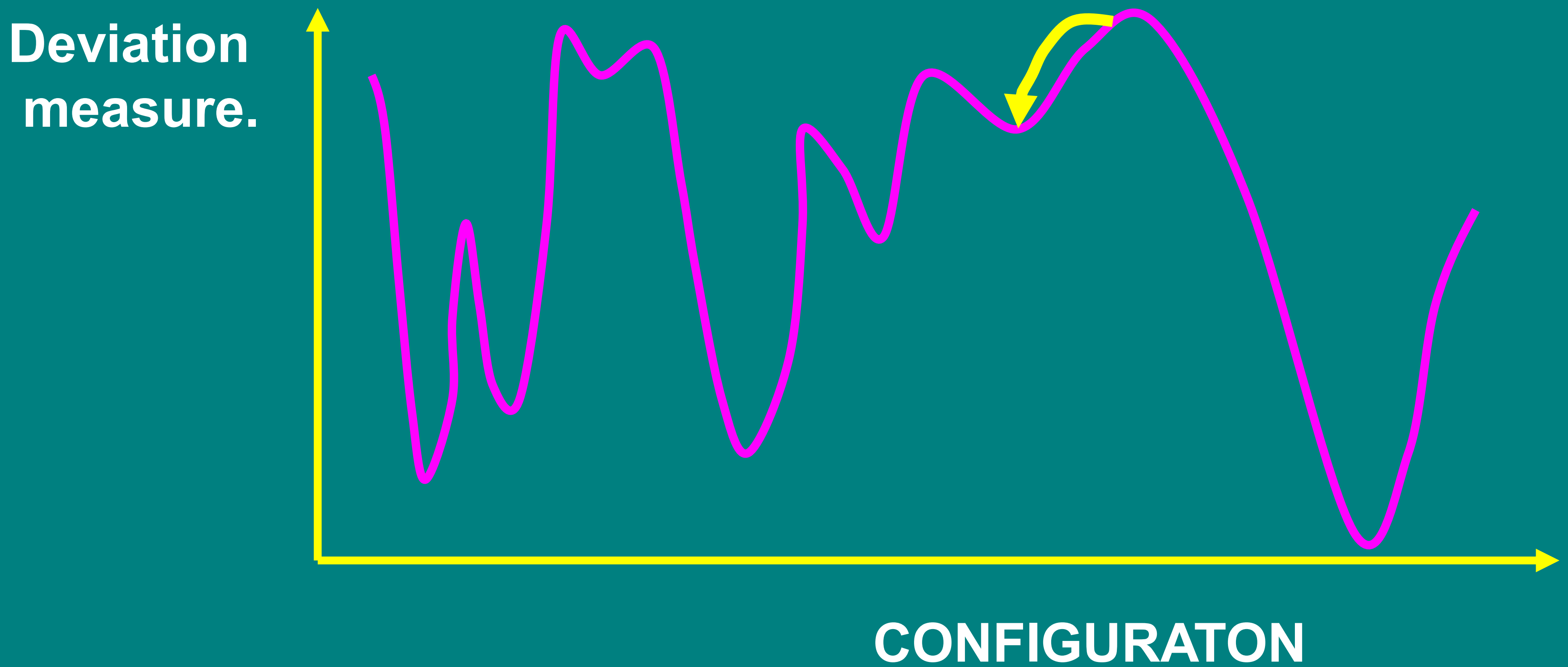


Parameters for changing are obtained by optimizing some continuous parameter making quadratic (intra)extrapolation.

For example: Measure of deviation for the shift of rectangle is calculated for distances x, 2x, and 3x and then 3 points D(x) is reproduced by parabola.

Variable x can be any other continuous parameter of the update.

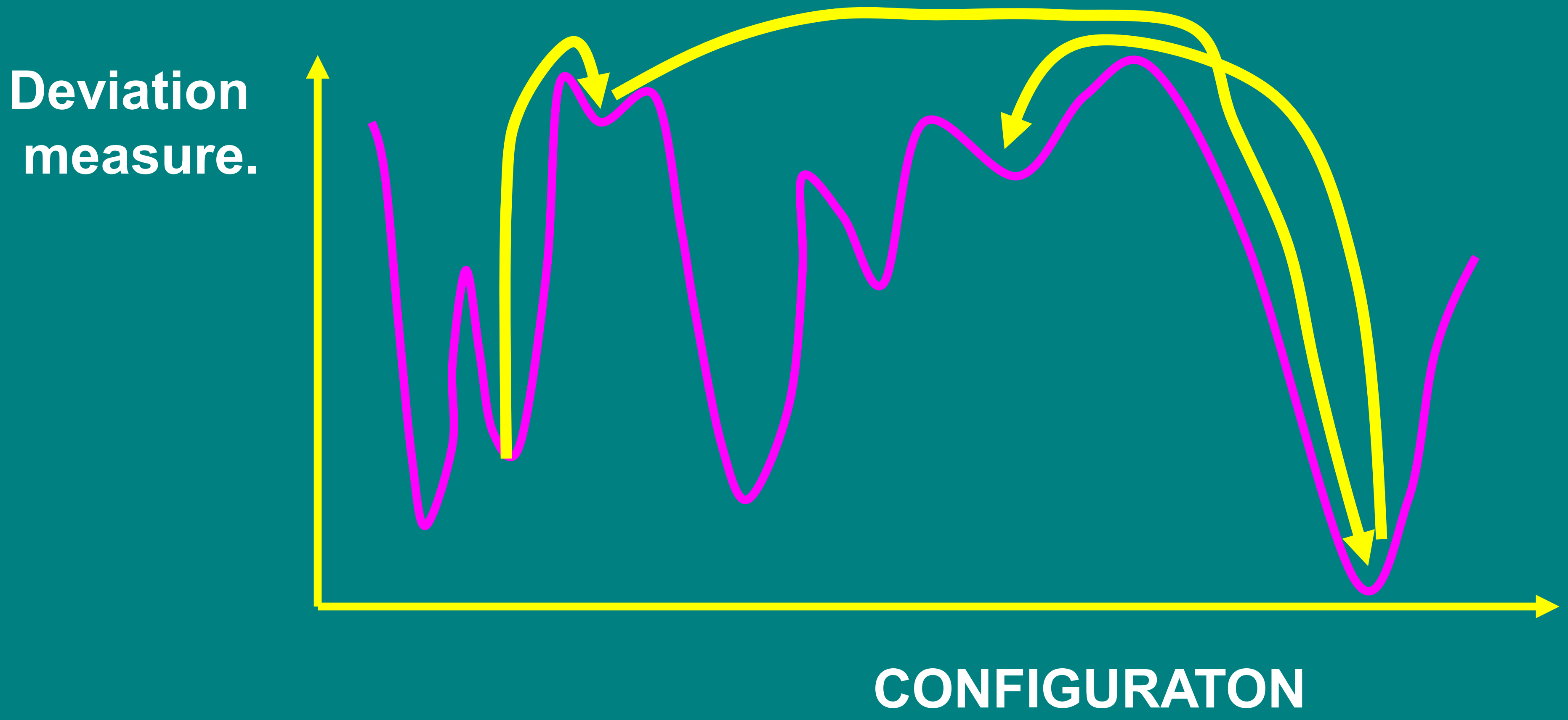
Stochastic Optimization method: update procedures.



Accept only updates which decrease the deviation measure

WRONG STRATEGY

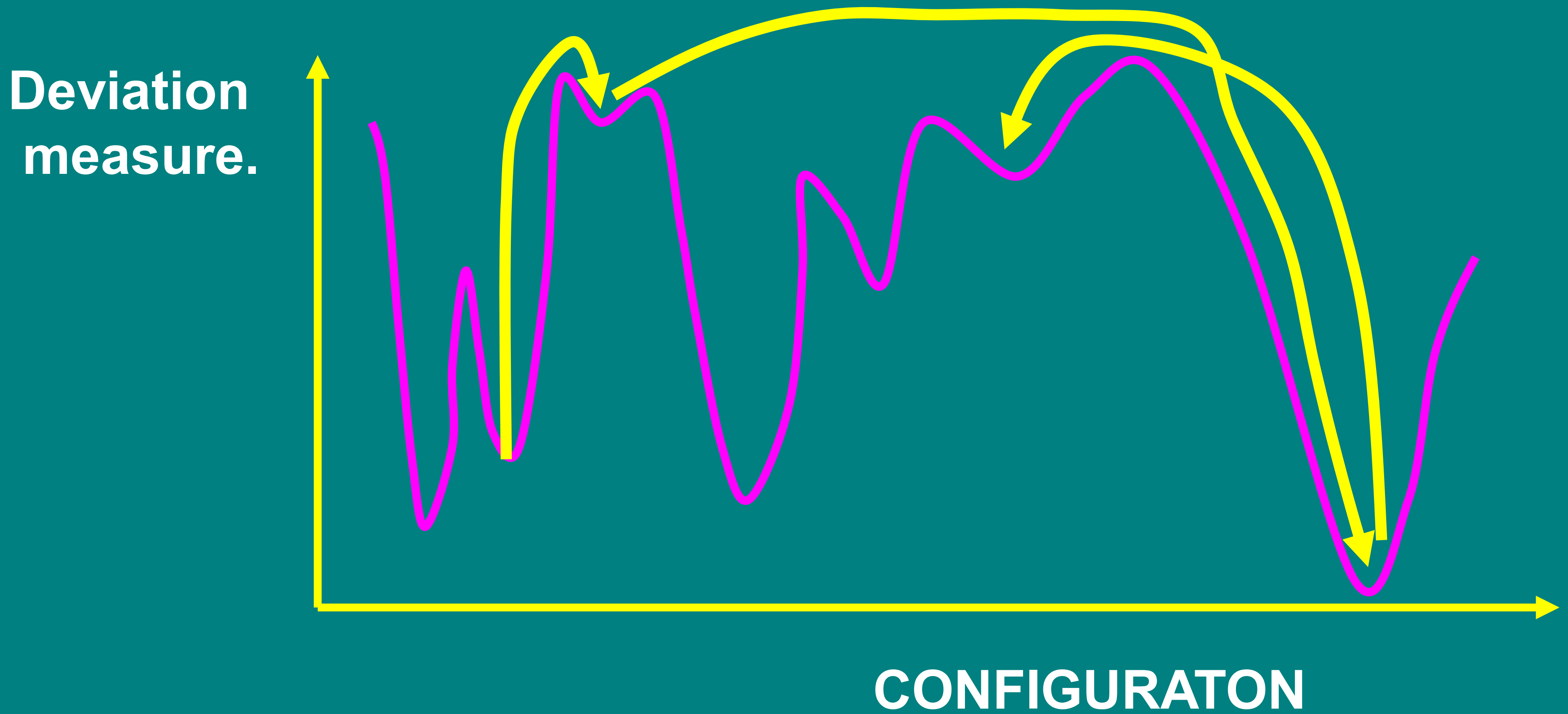
Stochastic Optimization method: update procedures.



Always accept with some probability some updates which decrease the deviation measure

WRONG STRATEGY: Sandvik 1998, Beach 2004

Stochastic Optimization method: update procedures.

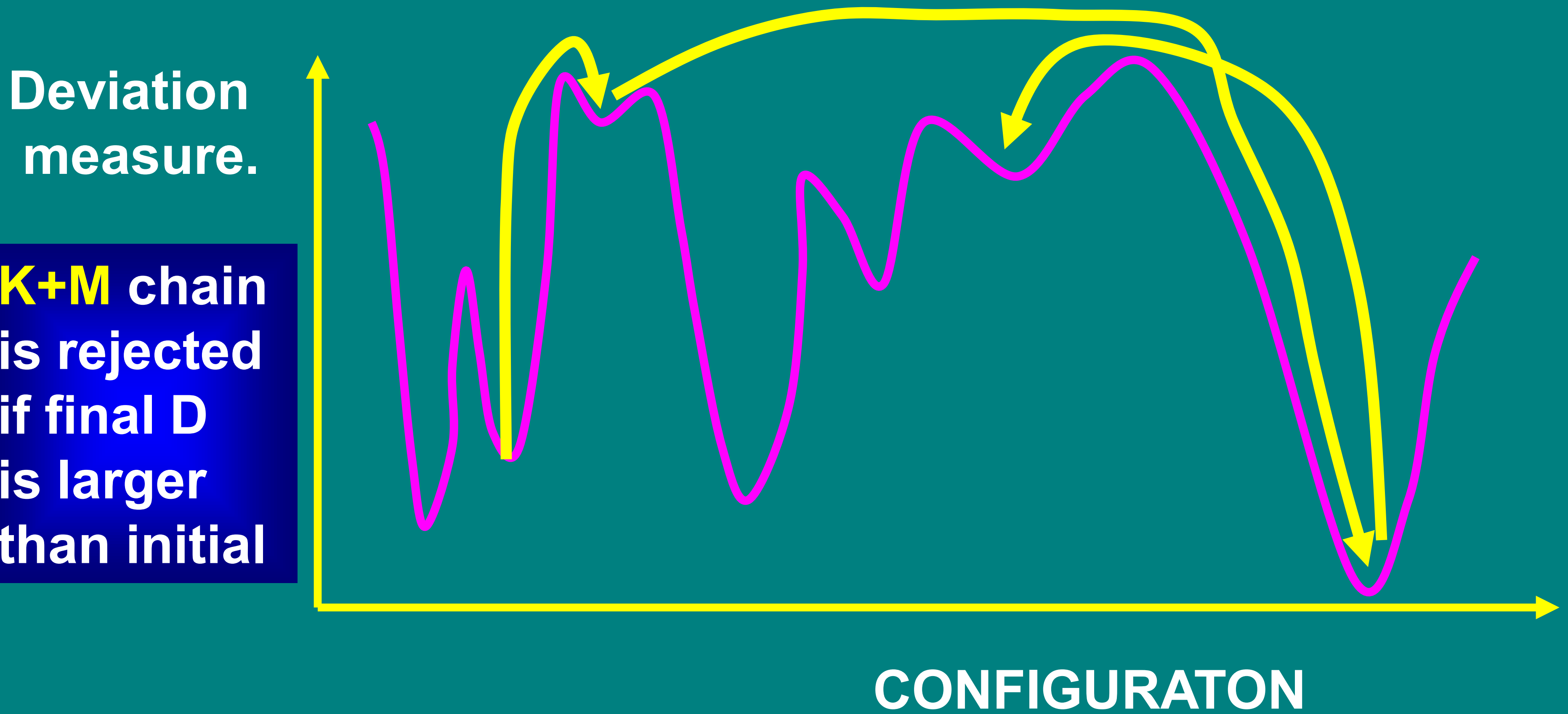


Shake-off two-step strategy:

Step 1: Increase of deviation measure is allowed during M steps with high probability

Step 2: Only decrease of deviation measure is allowed during last K steps.

Stochastic Optimization method: update procedures.



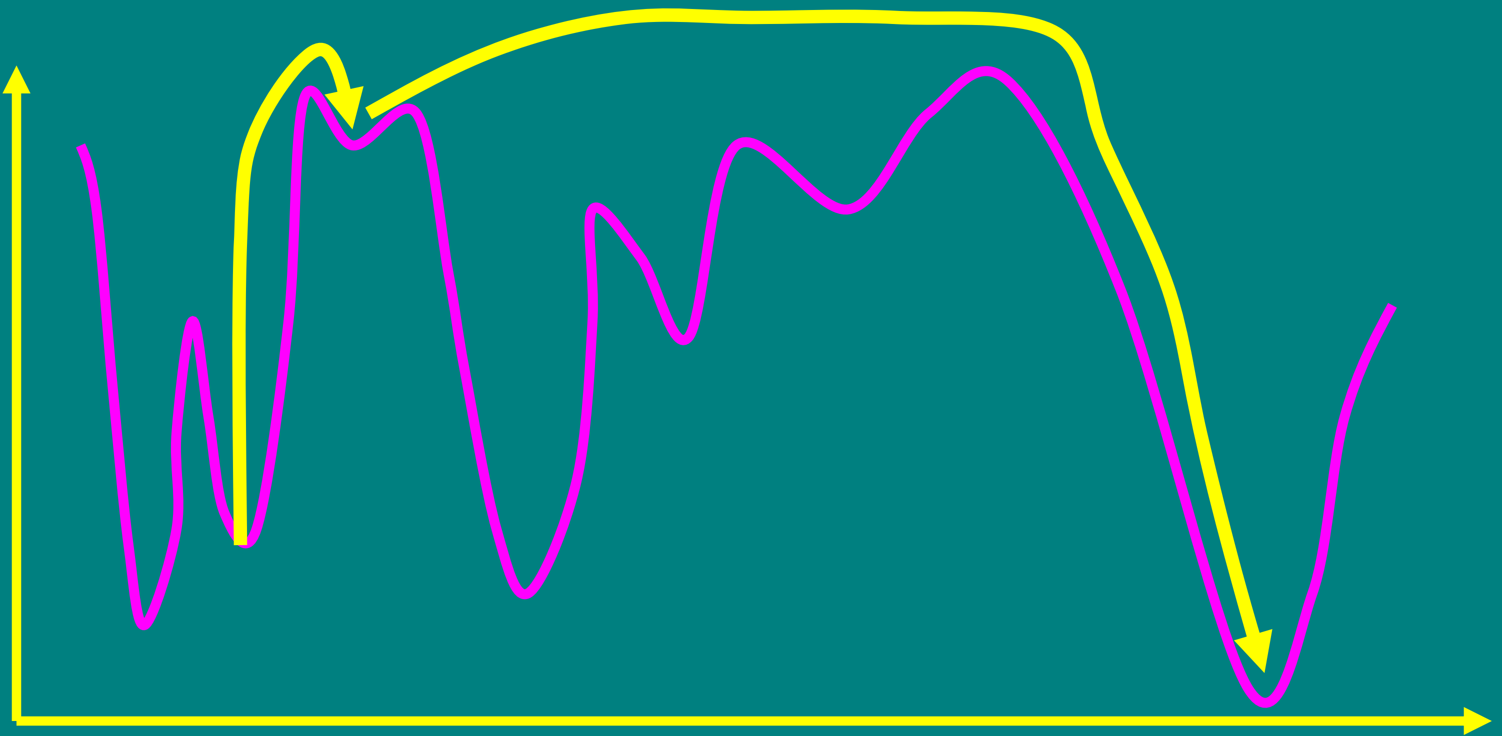
K+M chain is rejected if final D is larger than initial

Shake-off two-step strategy:

Step 1: Increase of deviation measure is allowed during M steps with high probability

Step 2: Only decrease of deviation measure is allowed during last K steps.

Stochastic Optimization method: update procedures.



K+M chain is accepted if final D is smaller than initial

Shake-off two-step strategy:

Step 1: Increase of deviation measure is allowed during M steps with high probability

Step 2: Only decrease of deviation measure is allowed during last K steps.

How to judge that one of solutions is *“GOOD”*

How to judge that one of solutions is "GOOD"

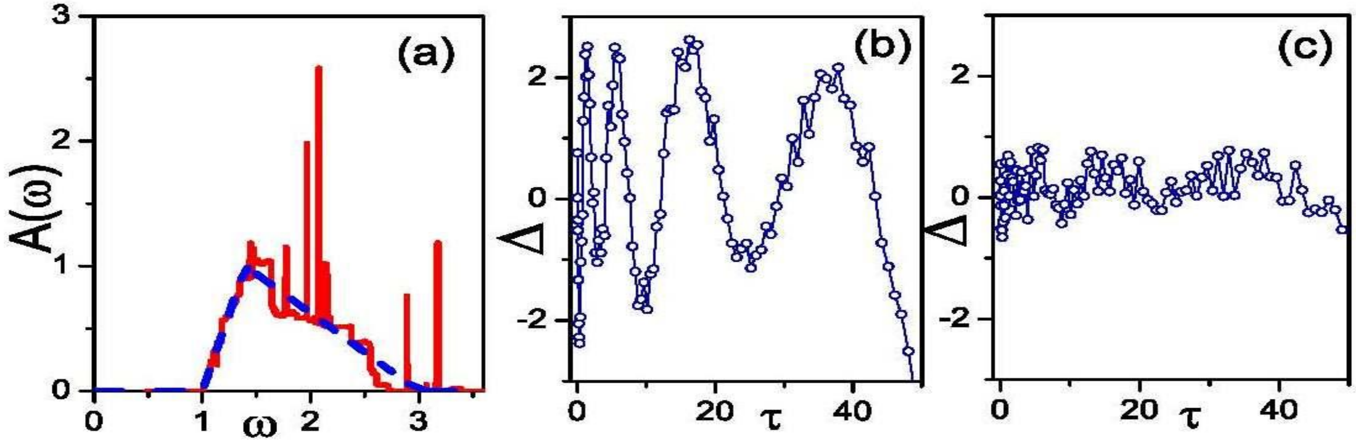


Fig. 6: *(a) Typical spectrum $\widetilde{A}_j(\omega)$ (red solid line), corresponding to a particular configuration C_j, compared to the actual spectrum (blue dashed line). Typical dependence of the deviation function $\Delta(m)$ (30) on imaginary times τ_m corresponding to a spectrum $\widetilde{A}_j(\omega)$ which (b) under-fits and (c) over-fits the uncorrelated noise of imaginary time data.*

deviation measure of SOM is given by expression

$$D[\widetilde{A}] = \sum_{m=1}^{M} |\Delta(m)| \; .$$

Here $\Delta(m)$ is the *deviation function*

$$\Delta(m) = \frac{G(m) - \tilde{G}(m)}{S(m)},$$

$$\kappa = \frac{1}{M-1} \sum_{m=2}^{M} \theta\{-\Delta(m)\Delta(m-1)\}$$

K>1/4 (Ideal limit κ=1/2)

Stochastic Optimization method.

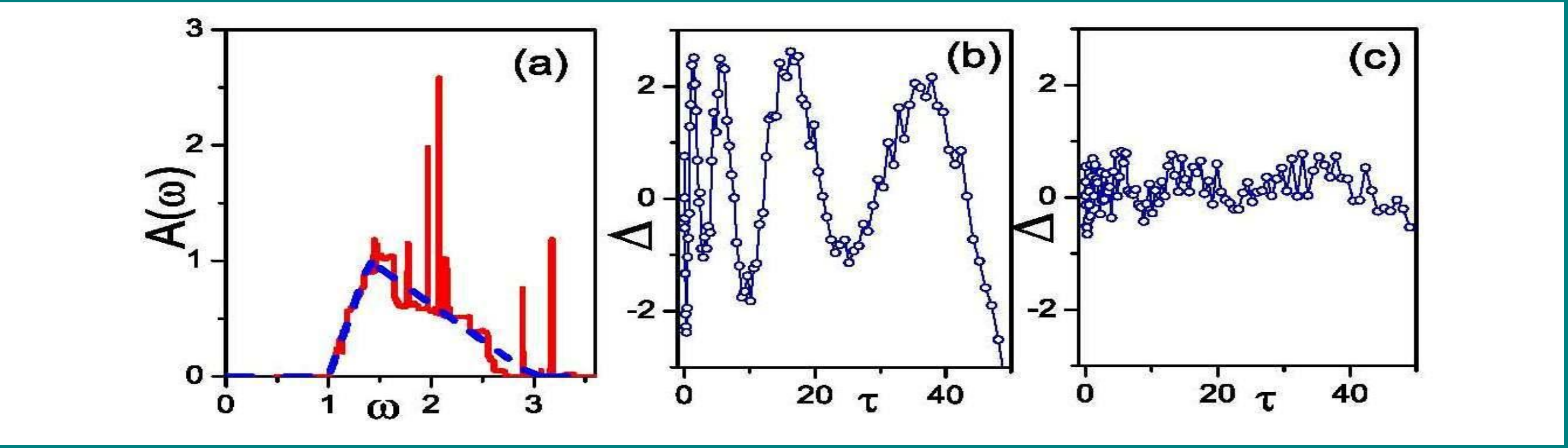
- Particular solution $L^{(i)}(\omega)$ for LSF is presented as a sum of a number K of rectangles with some width, height and center.
- Initial configuration of rectangles is created by random number generator (i.e. number K and all parameters of of rectangles are randomly generated).
- Each particular solution $L^{(i)}(\omega)$ is obtained by a naïve method without regularization (though, varying number K).

- Final solution is obtained after M steps of such procedure

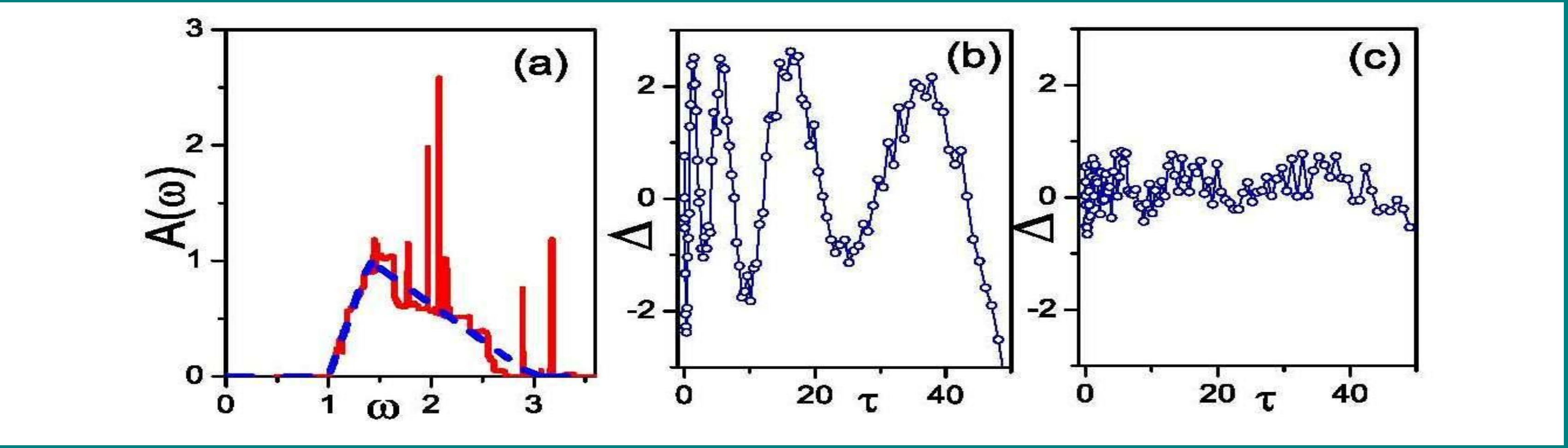
$$L(\omega) = M^{-1} \sum_i L^{(i)}(\omega)$$

- Each particular solution has saw tooth noise
- Final averaged solution $L(\omega)$ has no saw tooth noise though not regularized with sharp peaks/edges!!!!

We can find many particular solutions each of which fits the input data reasonably.



We can find many particular solutions each of which fits the input data reasonably.



Which particular solutions one has to take into account?

Self-averaging of the saw-tooth noise.

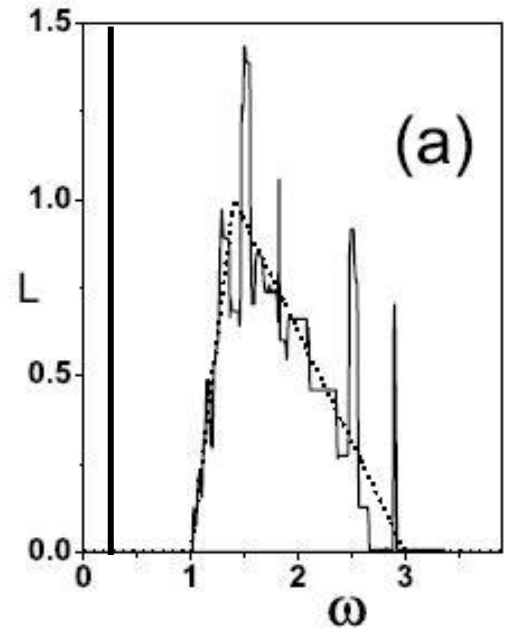


Fig. 7. Comparison of the actual spectral function (dashed line) with the results of spectral analysis after averaging over (a) $\boxed{M = 4,}$ (b) $M = 28$, and (c) $M = 500$ particular solutions.

Self-averaging of the saw-tooth noise.

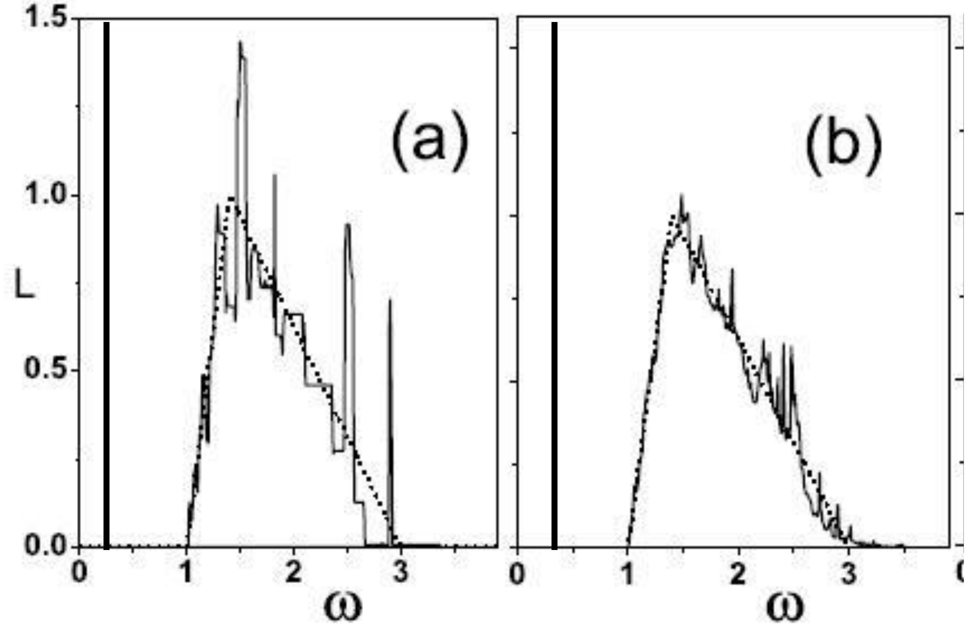


Fig. 7. Comparison of the actual spectral function (dashed line) with the results of spectral analysis after averaging over (a) $M = 4$, (b) $\boxed{M = 28,}$ and (c) $M = 500$ particular solutions.

Self-averaging of the saw-tooth noise.

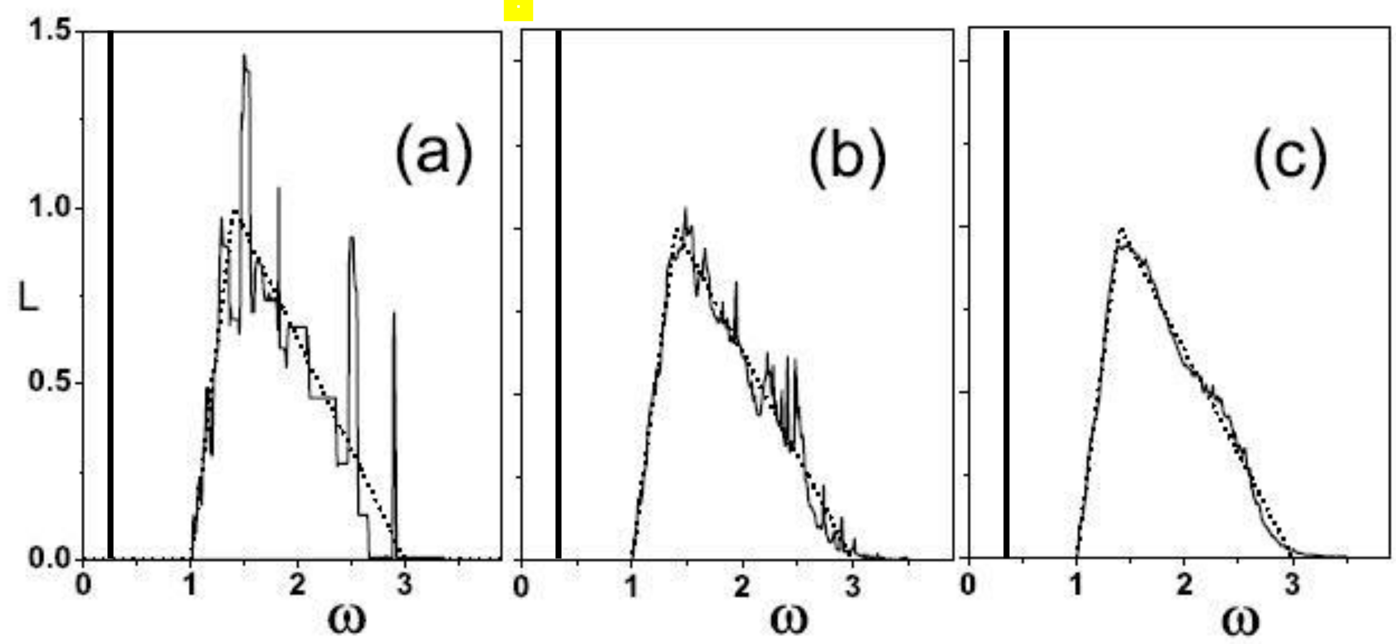


Fig. 7. Comparison of the actual spectral function (dashed line) with the results of spectral analysis after averaging over (a) $M = 4$, (b) $M = 28$, and (c) $M = 500$ particular solutions.

Which particular solutions one has to take into account?

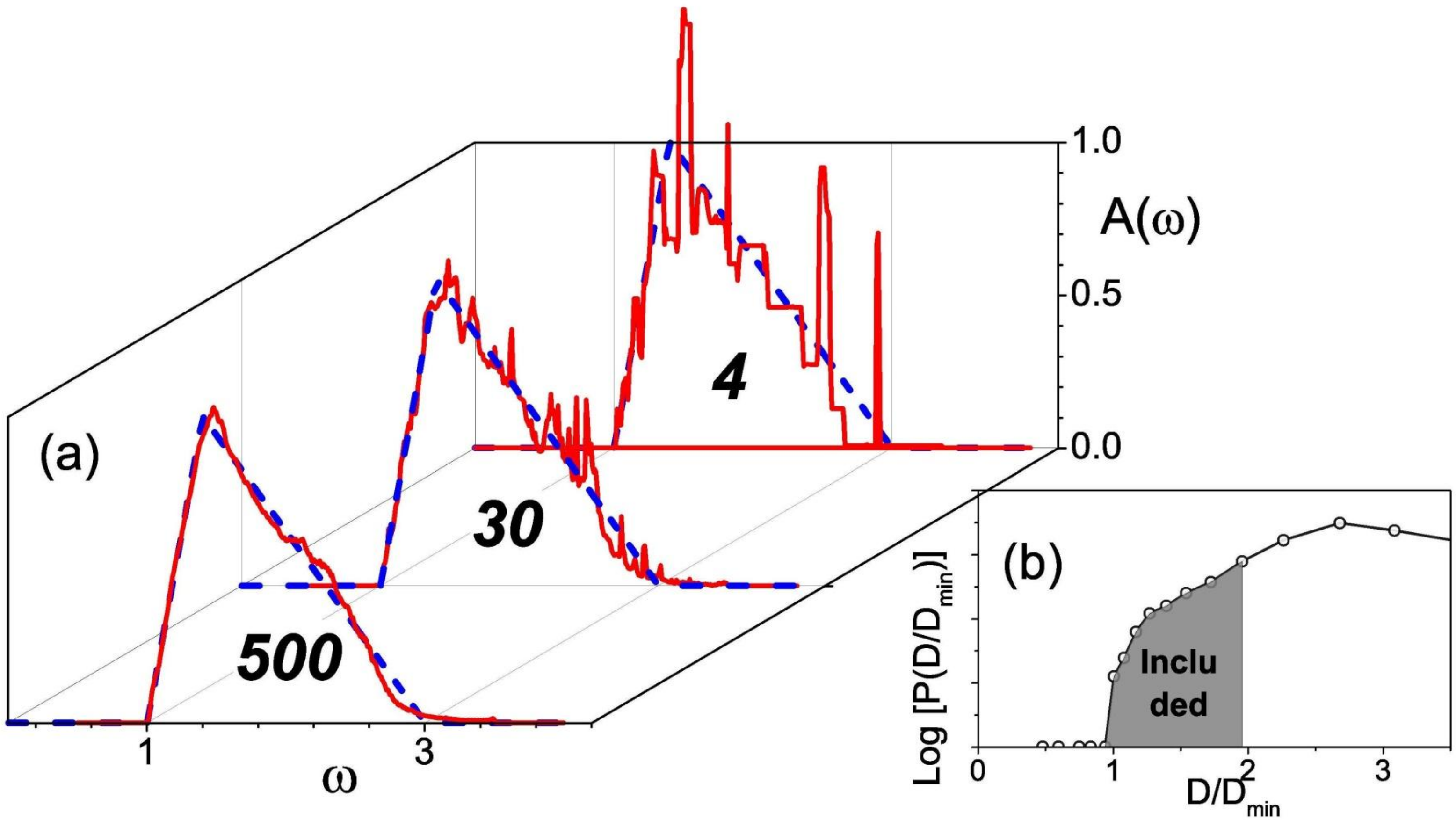


Fig. 7: *(a) Self-averaging of the sawtooth noise after summation of 4, 30, and 500 solutions. (b) Typical probability distribution $P(D/D_{min})$ of solutions with different deviation measures.*

Which particular solutions one has to take into account?

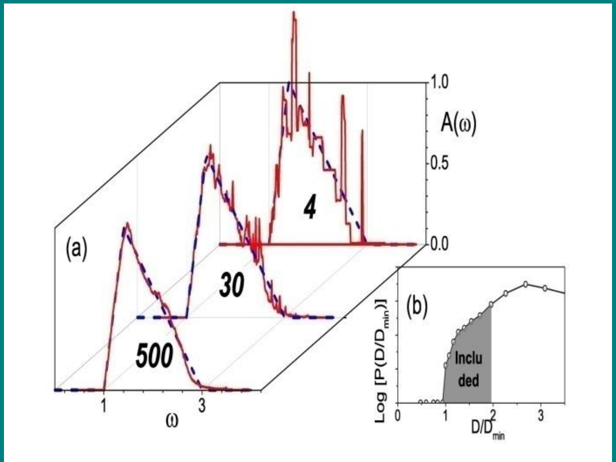


Fig. 7: *(a) Self-averaging of the sawtooth noise after summation of 4, 30, and 500 solutions. (b) Typical probability distribution $P(D/D_{min})$ of solutions with different deviation measures.*

One has to include solution with deviation measure D[A] which is less that twice of minimal MIN{D[A]}

D[A] < 2 MIN{D[A]}

deviation measure of SOM is given by expression

$$D[\widetilde{A}] = \sum_{m=1}^{M} |\Delta(m)| \ .$$

Here $\Delta(m)$ is the *deviation function*

$$\Delta(m) = \frac{G(m) - \tilde{G}(m)}{S(m)},$$

Sandvik method

$$A = \int d\widetilde{A}\, \widetilde{A}\, P[\widetilde{A}|G]$$

$$\mathcal{P}[A|G] = \exp\{-\chi^2[\widetilde{A}]/T\}$$

$$\chi^2[\widetilde{A}] = \sum_{m=1}^{M} \mathcal{E}^{-1}(m)[G(m) - \widetilde{G}(m)]^2$$

1. T is not too high. Otherwise A is far from spectra which fit well the correlation function G.
2. T is not too small otherwise we are back again to the sawtooth noise problem. Ower-fitting of the noise.

Simple rule T = M

$$\| \widehat{\mathcal{K}}\vec{A} - \vec{G} \|^2 = \sum_{m=1}^{M} \left| \sum_{n=1}^{N} \mathcal{K}(m,\omega_n) A(\omega_n) - G(m) \right|^2$$

Tikhonov functional:

similar strategy for choice of λ

$$\| \widehat{\mathcal{K}}\vec{A} - \vec{G} \|^2 + \lambda^2 \| \hat{\Gamma}\vec{A} \|^2$$

$$\vec{A} = \sum_{i=1}^{r} \frac{\vec{u}_i^\dagger \otimes \vec{v}_i}{\sigma_i} \vec{G} \quad \rightarrow \quad \vec{A} = \sum_{i=1}^{r} \boxed{\left\{ \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \right\}} \frac{\vec{u}_i^\dagger \otimes \vec{v}_i}{\sigma_i} \vec{G}$$

Max Ent

Similar strategy everywhere:
equate noise contribution
with regularization contribution

Avoid over-fitting

Similar strategy everywhere: equate noise contribution with regularization contribution

Tikhonov & Arsenin, Solution of Ill-posed problems, (Washington, 1977).

Arsenin (1986):

the art of finding solution for ill posed problem lies in an intuition which tells us when to stop improve the deviation before the noise of input data overruns the information contained in the input data.

Which particular solutions one has to take into account?

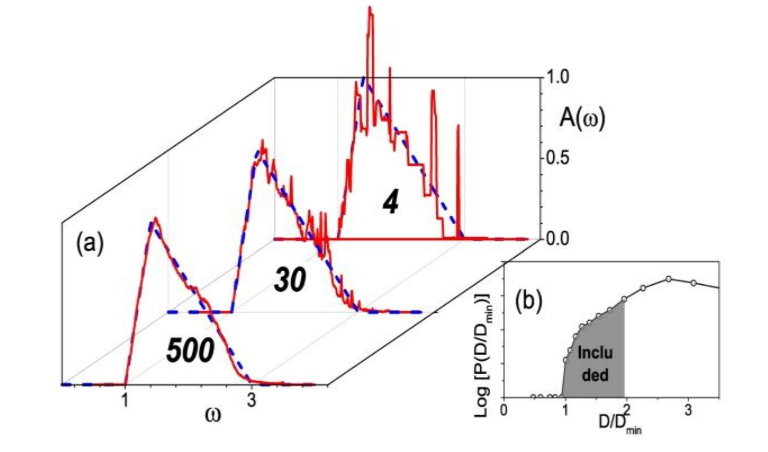


Fig. 7: *(a) Self-averaging of the sawtooth noise after summation of 4, 30, and 500 solutions. (b) Typical probability distribution $P(D/D_{min})$ of solutions with different deviation measures.*

One has to include solution with deviation measure D[A] which is less that twice of minimal MIN{D[A]}

D[A] < 2 MIN{D[A]}

deviation measure of SOM is given by expression

$$D[\widetilde{A}] = \sum_{m=1}^{M} |\Delta(m)| \ .$$

Here $\Delta(m)$ is the *deviation function*

$$\Delta(m) = \frac{G(m) - \tilde{G}(m)}{S(m)},$$

Some tests

Some tests

$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

$$\left\{ \tilde{G}(m) \left[1 + \frac{\mathcal{B}}{2} \mathcal{R} \right] \; , \; m = 1, M \right\}$$

$$\mathcal{R} \in [-1, 1]$$

***B*=10^-4^**

$B = 10^{-4}$

Some tests

$$G(m) = \int_{-\infty}^{\infty} d\omega \, K(m,\omega) \, A(\omega)$$

Particular cases.

Imaginary time, T=0:

Kernel is

$K(m,\omega) = \exp(-\tau_m \omega)$

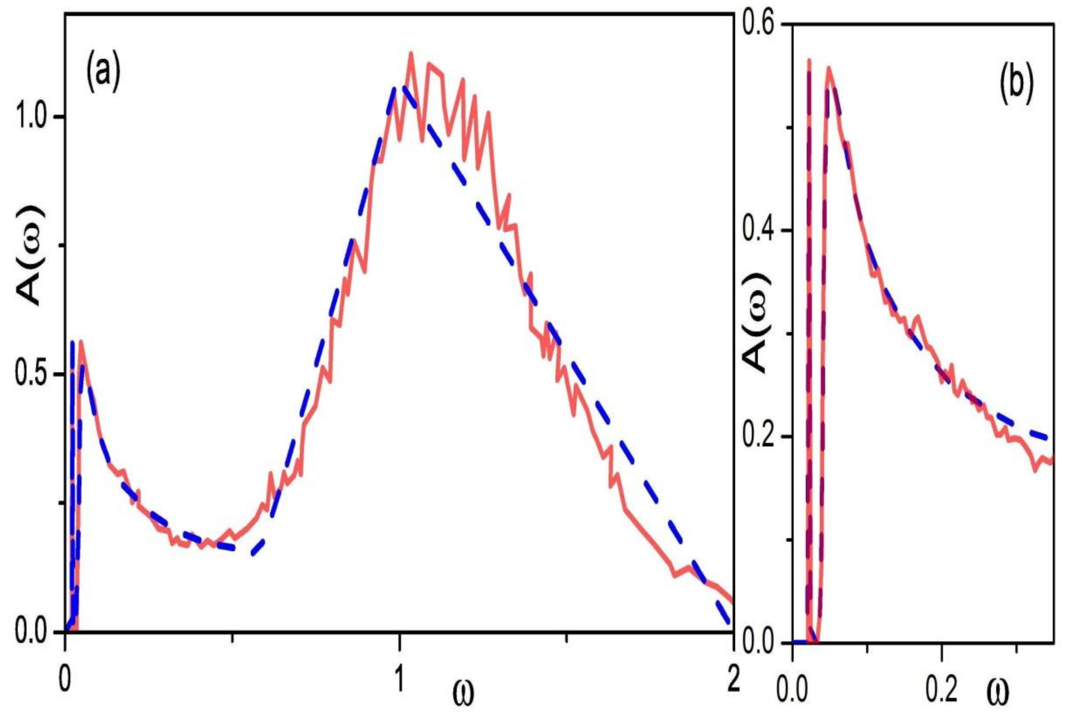


Fig. 8: *The test spectrum (dashed blue line) and the spectrum obtained by SOM (solid red line). Panels (a) and (b) show the whole spectrum and its low energy part, respectively.*

Some tests

$$G(m) = \int_{-\infty}^{\infty} d\omega\, \mathcal{K}(m,\omega)\, A(\omega)$$

Particular cases.

Imaginary time, finite T, fermions

Kernel is

$$\mathcal{K}(\tau_m,\omega) = -\frac{\exp(-\tau_m\omega)}{\exp(-\beta\omega)\pm 1}$$

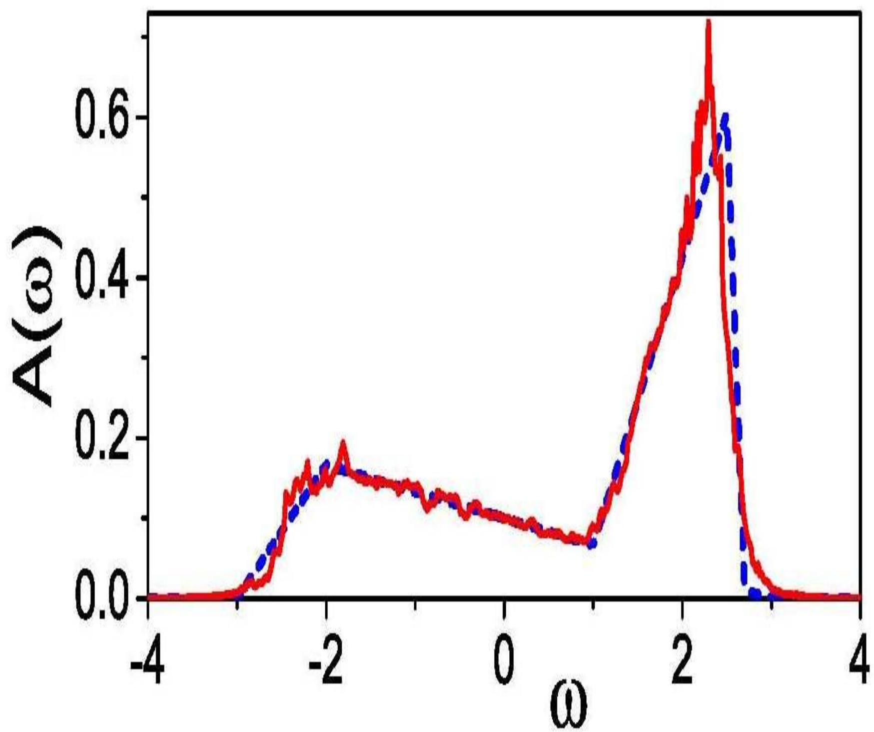


Fig. 9: *The test spectrum (dashed blue line) and the spectrum obtained by SOM (solid red line) for the Lehmann spectral function of fermions at finite temperature.*

Some tests

$$G(m) = \int_{-\infty}^{\infty} d\omega\, K(m,\omega)\, A(\omega)$$

Particular cases.

Imaginary time, finite T, optical conductivity

Kernel is

$$K(\tau_m,\omega) = \frac{1}{\pi}\frac{\omega \exp(-\tau_m \omega)}{1-\exp(-\beta\omega)}$$

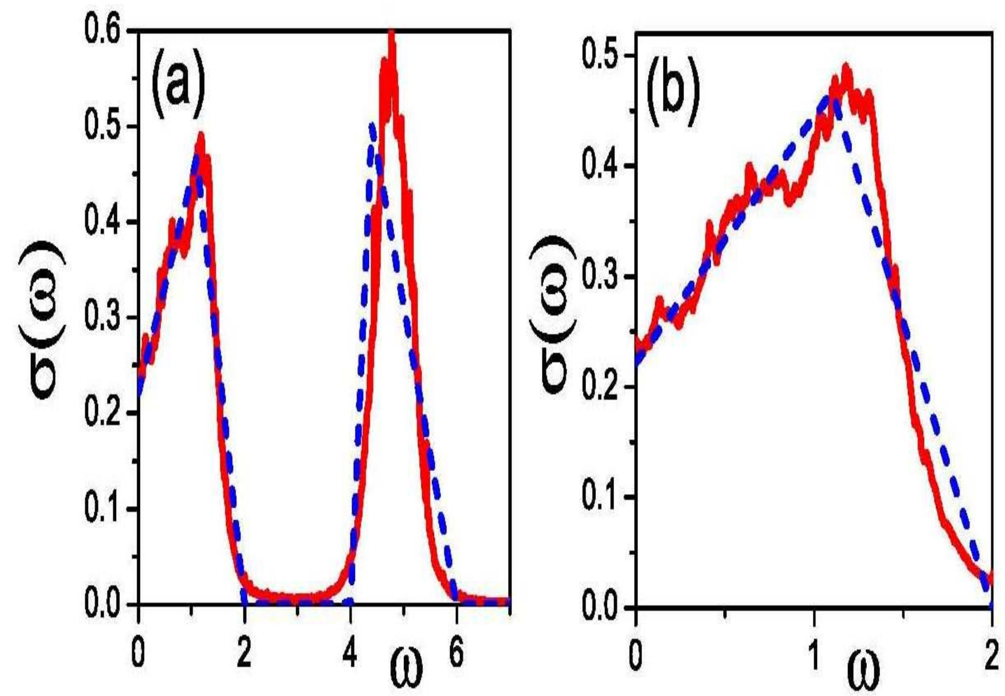


Fig. 10: *The test spectrum (dashed blue line) and the spectrum obtained by SOM (solid red line) for optical conductivity at finite temperature. Panels (a) and (b) show the whole range and low energy part, respectively.*

Some tests

$$G(m) = \int_{-\infty}^{\infty} d\omega\, K(m,\omega)\, A(\omega)$$

Particular cases.

Matsubara frequencies, finite T, fermions

Kernel is

$$K(i\omega_m, \omega) = \pm \frac{1}{i\omega_m - \omega}$$

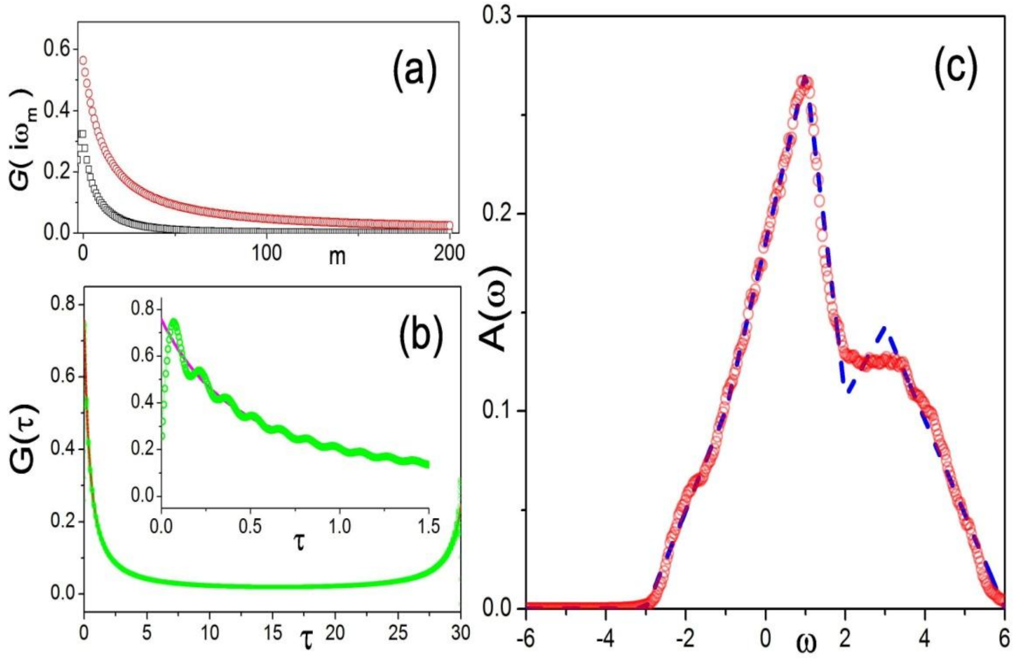


Fig. 11: *(a) First 200 Fourier components of the real (red circles) and imaginary (black squares) part of the GF in Matsubara representation obtained from the GF in imaginary time. (b) Imaginary time GF (solid line) and imaginary time GF obtained from first the $M = 200$ GFs in Matsubara representation. The inset shows low imaginary times. (c) Actual spectrum (dashed blue line) and that restored from 200 Matsubara components (red solid line).*

Back to optical conductivity. Let us compare MaxEnt and Stochastic.

$$G(m) = \int_{-\infty}^{\infty} d\omega \; K(m,\omega) \; A(\omega)$$

Particular cases.

Imaginary time, finite T, optical conductivity

Kernel is

$$K(\tau_m, \omega) = \frac{1}{\pi} \frac{\omega \exp(-\tau_m \omega)}{1 - \exp(-\beta \omega)}$$

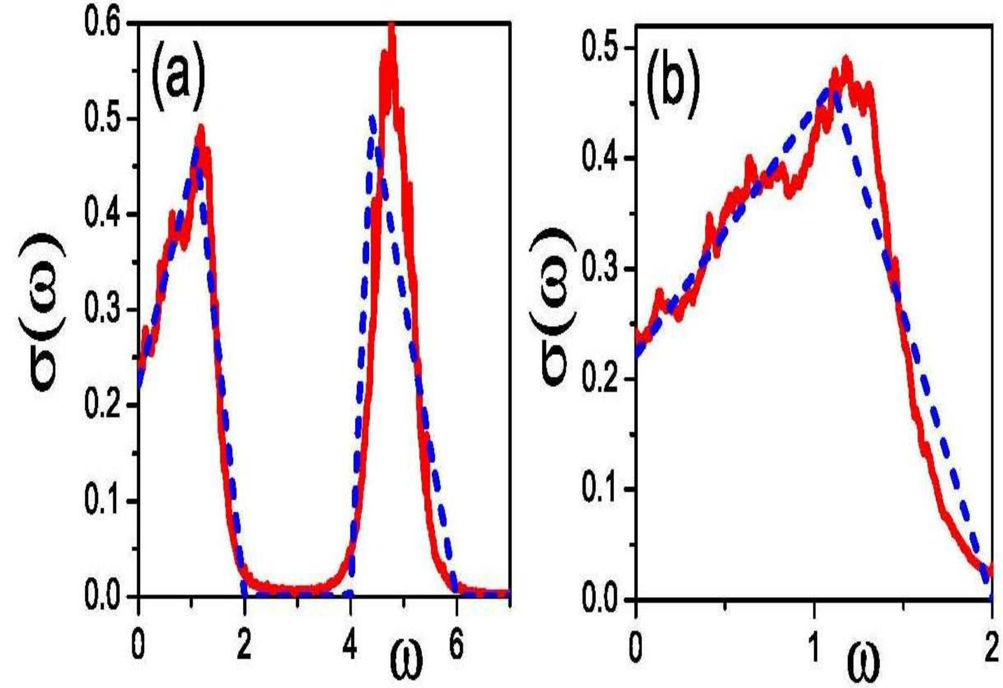


Fig. 10: *The test spectrum (dashed blue line) and the spectrum obtained by SOM (solid red line) for optical conductivity at finite temperature. Panels (a) and (b) show the whole range and low energy part, respectively.*

PHYSICAL REVIEW B **76**, 035115 (2007)

Analytical continuation of spectral data from imaginary time axis to real frequency axis using statistical sampling

K. Vafayi and O. Gunnarsson
Max-Planck-Institut für Festkörperforschung, D-70506 Stuttgart, Germany
(Received 27 February 2007; revised manuscript received 2 May 2007; published 19 July 2007)

We present a method for performing analytical continuation of spectral data from imaginary time to real frequencies based on a statistical sampling method. Compared with the maximum entropy method (MEM), an advantage is that no default model needs to be introduced. For the problems studied here, the statistical sampling method gives comparable or slightly better results than MEM using quite accurate default models.

DOI: 10.1103/PhysRevB.76.035115 PACS number(s): 72.15.Eb, 02.70.Ss

ity in Eq. (7) as a weight function. Comparing with the maximum entropy method (MEM), an advantage is that there is no need to provide a default model, which influences the MEM results if the method is close to its limit of applicability. For the problems considered here, the statistical sampling method gives comparable or slightly better results than MEM using default models close to the exact result. The

Stochastic Optimization method.

- Particular solution $L^{(i)}(\omega)$ for LSF is presented as a sum of a number K of rectangles with some width, height and center.
- Initial configuration of rectangles is created by random number generator (i.e. number K and all parameters of of rectangles are randomly generated).
- Each particular solution $L^{(i)(}\omega)$ is obtained by a naïve method without regularization (though, varying number K).
- Final solution is obtained after M steps of such procedure

$$L(\omega) = M^{-1} \sum_i L^{(i)(}\omega)$$

- Each particular solution has saw tooth noise
- Final averaged solution $L(\omega)$ has no saw tooth noise though not regularized with sharp peaks/edges!!!!

Conclusions:

1. Analytic continuation is ill posed problem.
2. Similar Fredholm I integral equatuion problem in many applications.
3. Long history of the methods: Tikhonov -> MaxEnt -> stochastic.
4. All methods bear similar strategy of regularization: not to over-fit the noise
5. Each method is the best in each particular case. There is no universal method which is “the best” for all cases.
6. We are still on the way to improve the analytic continuation.
7. Combinations of methods might help.

Questions?

New Method for Low Temperature analysis of the ESR spectra

Andrey Mishchenko
CMRG, RIKEN

Collaborations:

Tatsuo Hasegawa (AIST)
Hiroyuki Matsui (AIST)

Phys Rev. Lett. 104, 056602 (2010)

New Method for Low Temperature analysis of the ESR spectra

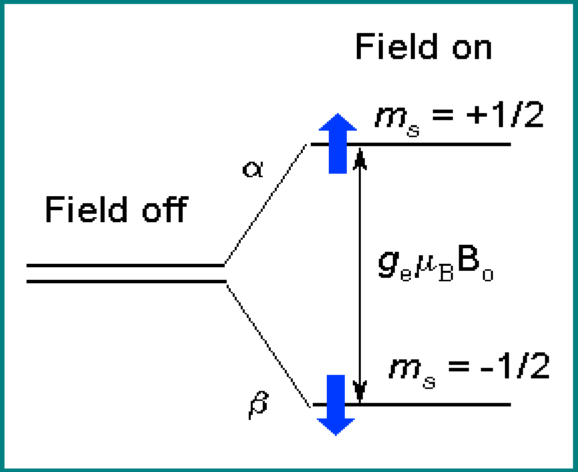
1. Nature of the inhomogeneous ESR lineshape and line narrowing
3. Analysis of the lineshape of an electron trapped by an impurity
4. Analysis of the fine structure of the ESR line can give a complete information on the distribution of the traps versus localization parameters

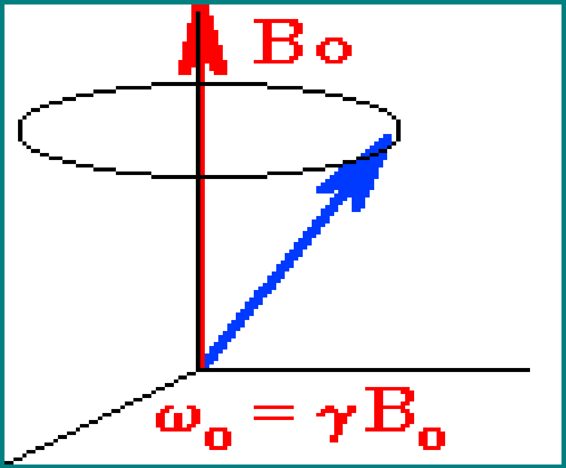
Basics of ESR

Transition between Zeeman split levels under the influence of the electromagnetic field.

For example, the frequency is fixed and magnetic field B is varied. Then, the intensity of signal I(B) is

$$I(B) \sim \delta(B-B_0)$$



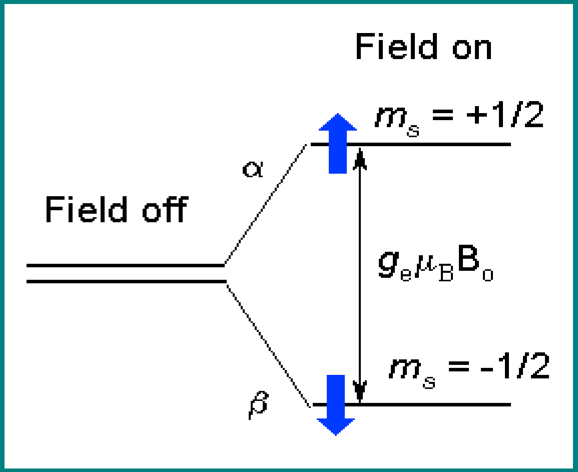


Basics of ESR

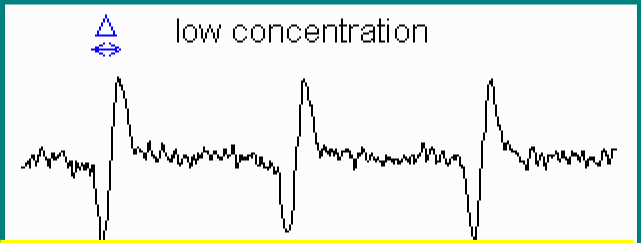
Transition between Zeeman split levels under the influence of the electromagnetic field.

For example, the frequency is fixed and magnetic field B is varied. Then, the intensity of signal I(B) is

$$I(B) \sim \delta(B-B_0)$$



Hyperfine splitting

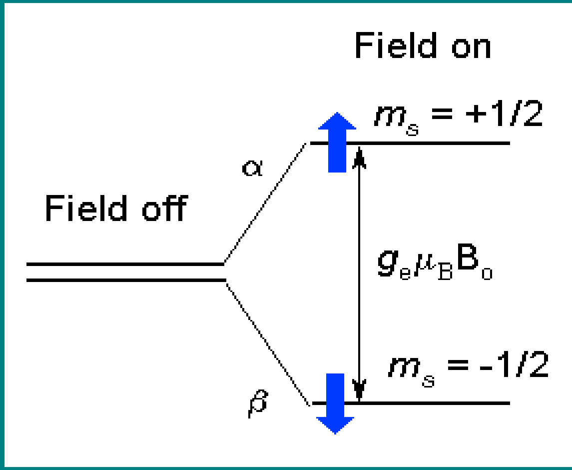


Basics of ESR

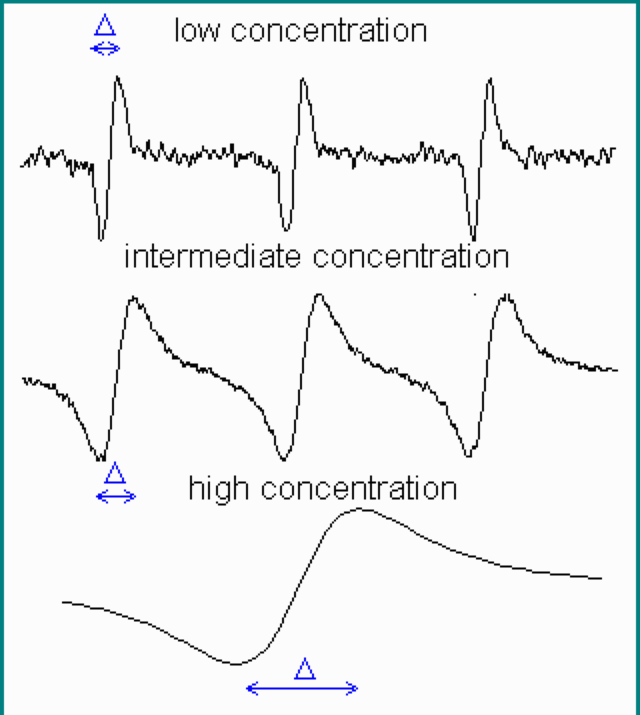
Transition between Zeeman split levels under the influence of the electromagnetic field.

For example, the frequency is fixed and magnetic field B is varied. Then, the intensity of signal I(B) is

$$I(B) \sim \delta(B-B_0)$$



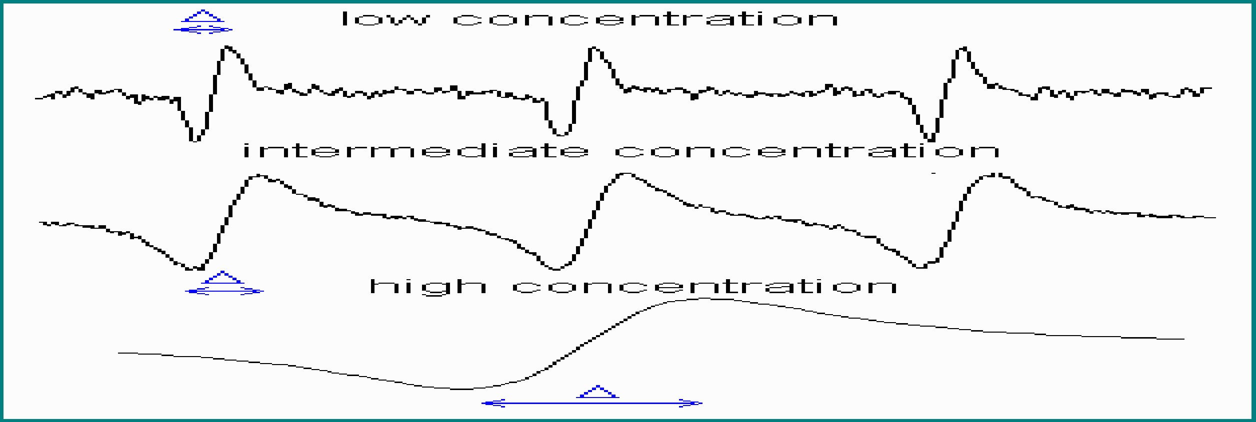
Hyperfine splitting



In complex system at low temperatures the lineshape is set by the sum of random contributions coming from hyperfine and superhyperfine interactions.

Basics of ESR

Hyperfine splitting



In complex system at low temperatures the lineshape is set by the sum of random contributions coming from hyperfine and superhyperfine interactions.

The distribution of the sum of random variable is Gaussian:

$$G(B) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(B-B_0)^2}{2\sigma_0^2}\right]$$

$$S(B) = dG(B) / dB$$

Basics of ESR

The distribution of the sum of random variable is Gaussian:

$$G(B) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(B - B_0)^2}{2\sigma_0^2}\right]$$

If the electron is spread over N molecules, one has a distribution over sum of random variables. Then, according to the Central Limit Theorem the distribution is Gaussian with more narrow dispersion σ:

$$< B > = \frac{1}{N}(B_1 + B_2 + \ldots + B_N)$$

$$\sigma = \frac{\sigma_0}{\sqrt{N}}$$

Basics of ESR

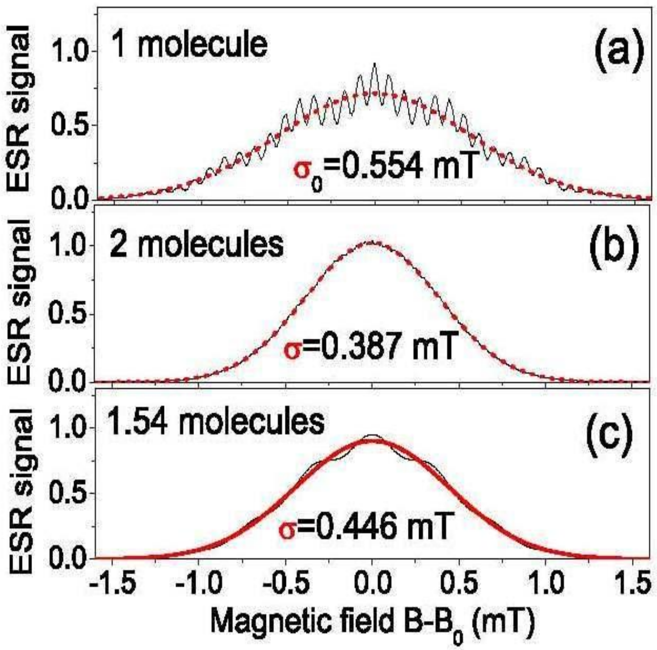


FIG. 1: (color online) Simulated ESR spectrum (black solid lines) of 1, 2, and 1.54 pentacene molecules based on the experimental hyperfine splitting of one pentacene molecule and fits by Gauss distribution (red dashed lines).

The ESR signal from single molecule in this case is Gaussian. The standard expression describing the hyperfine structure of one molecule reads[10]

$$f(B) = \sum_{m_1=-n_1I_1}^{n_1I_1} \cdots \sum_{m_k=-n_kI_k}^{n_kI_k} P(m_1,\ldots,m_k) \times \frac{1}{\pi} \frac{\Gamma}{\left(B - \sum_{i=1}^{k} A_i m_i\right)^2 + \Gamma^2}. \quad (1)$$

Here k is the number of the groups of equivalent nuclei, n_i is the number of the equivalent nuclei in the ith group, I_i is nuclear spin in the ith group, Γ is the linewidth of each peak, P is the intensity of each peak and B is magnetic field. If protons ($I = 1/2$) are the only paramagnetic nuclei, as it is, e.g, in the case for pentacene molecule, P is given as

$$P(m_1,\ldots,m_k) = \Pi_{i=1}^{k} \frac{C_{2n_iI_i}^{m_i+n_iI_i}}{(2I_i+1)^{-n_i}}, \quad (2)$$

where $C_{2n_iI_i}^{m_i+n_iI_i}$ are binomial coefficients.

N molecules, $A_i=A_i/N$, $n_i=n_i/N$

If the electron is spread over N molecules, one has a distribution over sum jf random variables. Then, according to the Central Limit Theorem the distribution is Gaussian with more narrow dispersion σ:

$$< B >= \frac{1}{N}(B_1 + B_2 + \ldots + B_N)$$

$$\sigma = \frac{\sigma_0}{\sqrt{N}}$$

Above knowledge is from the theory of inhomogeneous lineshape in molecules. In solids? When an electron is localized on the trap, there is a charge distribution f(i) and one needs to look at the distribution of different variable <B>:

If the electron is spread over N molecules, one has a distribution over sum jf random variables. Then, according to the Central Limit Theorem the distribution is Gaussian with more narrow dispersion σ:

$$< B >= \frac{1}{N}(B_1 + B_2 + \ldots + B_N)$$

$$\sigma = \frac{\sigma_0}{\sqrt{N}}$$

Above knowledge is from the theory of inhomogeneous lineshape in molecules. In solids? When an electron is localized on the trap, there is a charge distribution f(i) and one needs to look at the distribution of different variable <B>:

Experiment reveals non-Gaussian signal. Maybe this is the reason.

If the electron is spread over N molecules, one has a distribution over sum of random variables. Then, according to the Central Limit Theorem the distribution is Gaussian with more narrow dispersion σ:

$$< B >= \frac{1}{N}(B_1 + B_2 + \ldots + B_N)$$

$$\sigma = \frac{\sigma_0}{\sqrt{N}}$$

Above knowledge is from the theory of inhomogeneous lineshape in molecules. In solids? When an electron is localized on the trap, there is a charge distribution f(i) and one needs to look at the distribution of variables <B>:

$$\sum_i f(i) = 1 \qquad\qquad < B >= \sum_i f(i) B_i$$

If the electron is spread over N molecules, one has a distribution over sum jf random variables. Then, according to the Central Limit Theorem the distribution is Gaussian with more narrow dispersion σ:

$$< B > = \frac{1}{N}(B_1 + B_2 + \ldots + B_N)$$

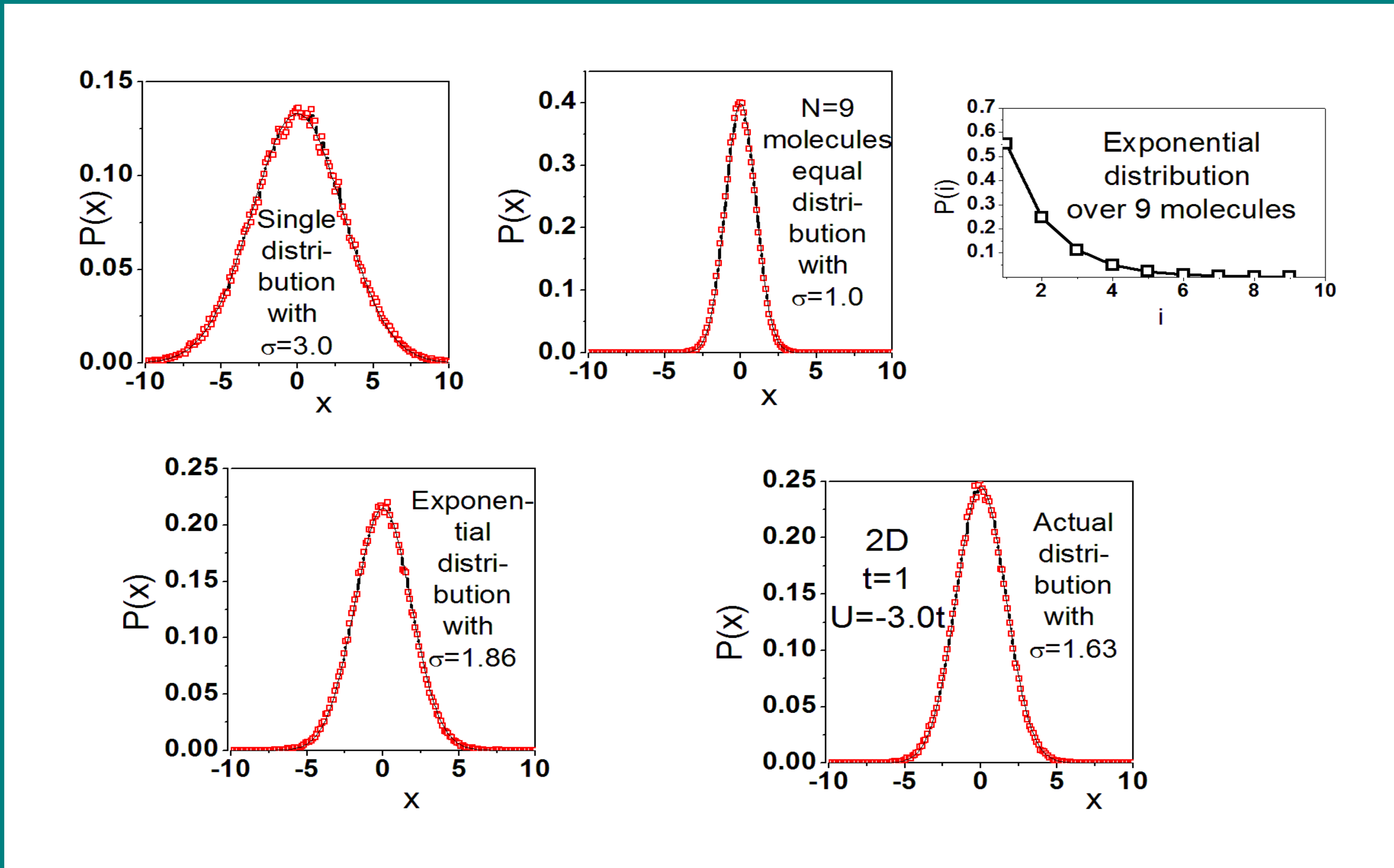
$$\sigma = \frac{\sigma_0}{\sqrt{N}}$$

Above knowledge is from the theory of inhomogeneous lineshape in molecules. In solids? When an electron is localized on the trap, there is a charge distribution f(i) and one needs to look at the distribution of variables <B>:

Numerical simulations show that (although CLT does not work in this case) the distribution is still Gaussian.

$$\sum_i f(i) = 1 \qquad < B > = \sum_i f(i) B_i$$

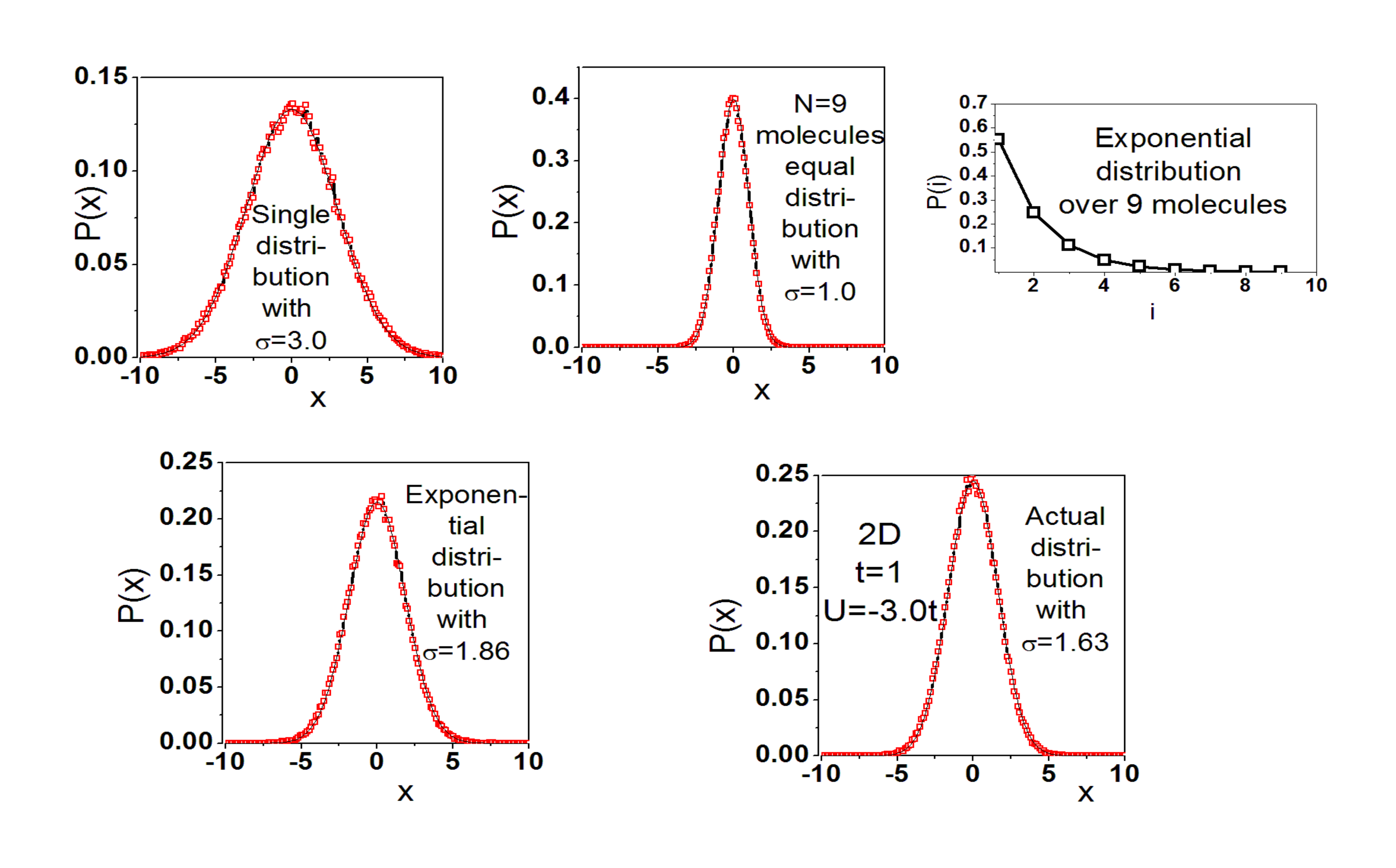
CLT for non-uniformly distributed variables: nontrivial!



The uniform distribution is the best case for narrowing

In extreme limit of localized case $N_{eff} \rightarrow 1$.

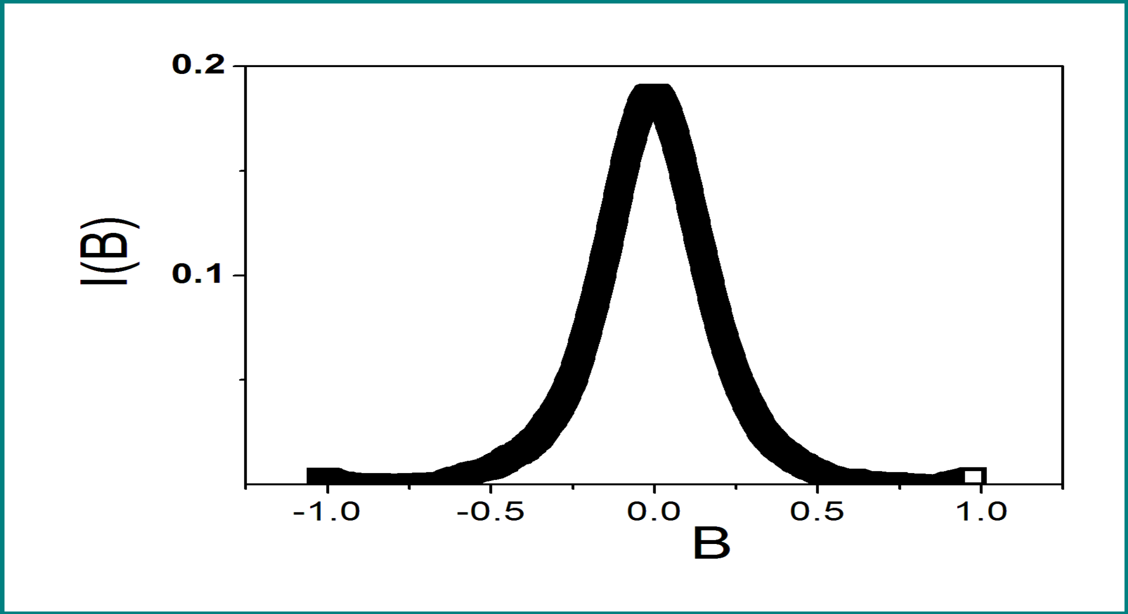
CLT for non-uniformly distributed variables: nontrivial!



$$N_{eff} = [\ \Sigma_i\ p(i)^2\ ]^{-1/2}$$

Experimental ESR in pentacene

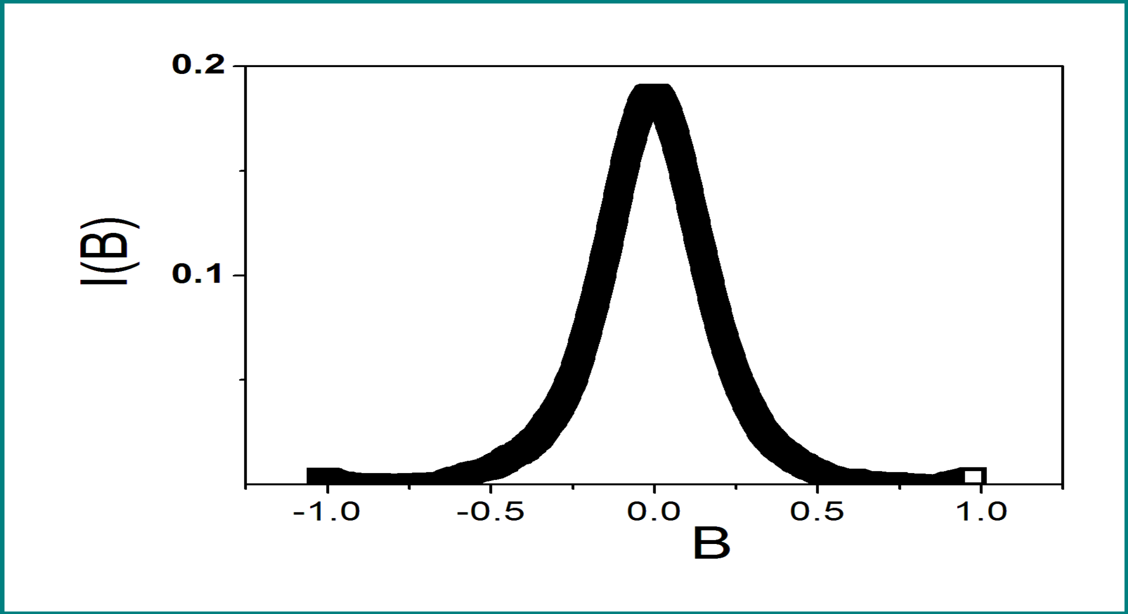
However, experimental signal is not Gaussian which means that there is no traps with some given value of N_{eff} which dominate.



For T<50K the wave saturation experiment shows that all carriers are localized and no broadening except nonhomogeneous one is expected!

Experimental ESR in pentacene

However, experimental signal is not Gaussian which means that there is no traps with some given value of N_{eff} which dominate.

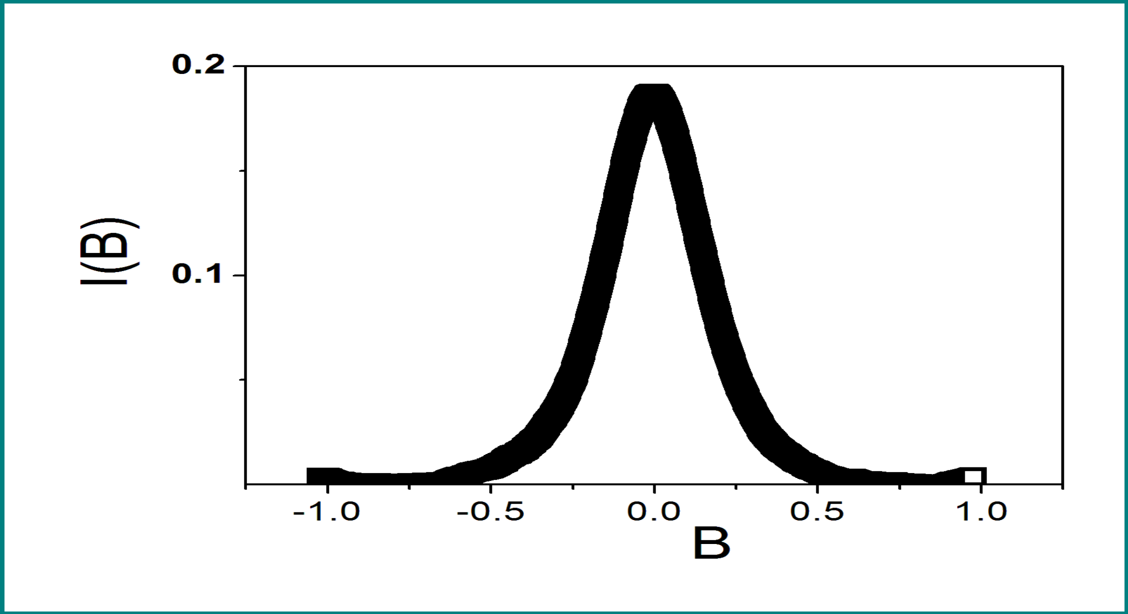


Two kinds of traps:

$$I(B) = \alpha\, G(N^1_{eff}, B-B_0) + \beta\, G(N^2_{eff}, B-B_0)$$

Experimental ESR in pentacene

However, experimental signal is not Gaussian which means that there is no traps with some given value of N_{eff} which dominate.



Two kinds of traps:

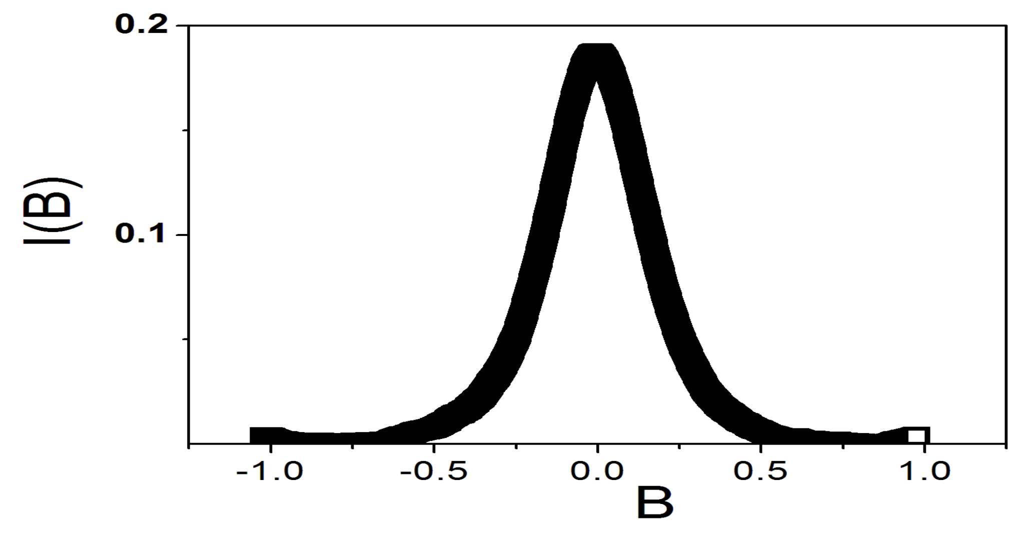
$$I(B) = \alpha\, G(N^1_{eff}, B-B_0) + \beta\, G(N^2_{eff}, B-B_0)$$

Three kinds of traps:

$$I(B) = \alpha\, G(N^1_{eff}, B-B_0) + \beta\, G(N^2_{eff}, B-B_0) + \gamma\, G(N^3_{eff}, B-B_0)$$

Broader view: **distribution of traps**

Experimental ESR in pentacene



$$G_{N_{eff}}(B) = \frac{1}{\sqrt{2\pi[\sigma_0/\sqrt{N_{eff}}]^2}} \exp\left[\frac{(B-B_0)^2}{2[\sigma_0/\sqrt{N_{eff}}]^2}\right]$$

$$I(B) = \int_1^{+\infty} D(N_{eff}) G_{N_{eff}}(B{-}B_0) dN_{eff} \quad S(B) = \int_1^{+\infty} D(N_{eff}) \left[\frac{d}{dB} G_{N_{eff}}(B-B_0)\right] dN_{eff}$$

$$G(\tau) = \int_0^{+\infty} \rho(\omega) K[\tau, \omega]$$

Experimental ESR in pentacene

Broader view: **distribution of traps**

Note, such interpretation requires that all molecules of molecular crystal are equally oriented with respect to surface.

$$G_{N_{eff}}(B) = \frac{1}{\sqrt{2\pi[\sigma_0/\sqrt{N_{eff}}]^2}} \exp\left[\frac{(B-B_0)^2}{2[\sigma_0/\sqrt{N_{eff}}]^2}\right]$$

$$I(B) = \int_1^{+\infty} D(N_{eff}) G_{N_{eff}}(B-B_0) dN_{eff} \quad S(B) = \int_1^{+\infty} D(N_{eff}) \left[\frac{d}{dB} G_{N_{eff}}(B-B_0)\right] dN_{eff}$$

$$G(\tau) = \int_0^{+\infty} \rho(\omega) K[\tau, \omega]$$

Fredholm integral equation of the 1-st kind: so called ill posed problem

$$G_{N_{eff}}(B) = \frac{1}{\sqrt{2\pi[\sigma_0/\sqrt{N_{eff}}]^2}} \exp\left[\frac{(B-B_0)^2}{2[\sigma_0/\sqrt{N_{eff}}]^2}\right]$$

$$I(B) = \int_1^{+\infty} D(N_{eff}) G_{N_{eff}}(B-B_0)dN_{eff} \quad S(B) = \int_1^{+\infty} D(N_{eff}) \left[\frac{d}{dB} G_{N_{eff}}(B-B_0)\right] dN_{eff}$$

$$G(\tau) = \int_0^{+\infty} \rho(\omega) K[\tau, \omega]$$

Previously (A.S. Mishchenko et al, PRB v. 62, 6317 (2000)) a method more flexible and less capricious than MEM was developed for solving the analytic continuation problem -

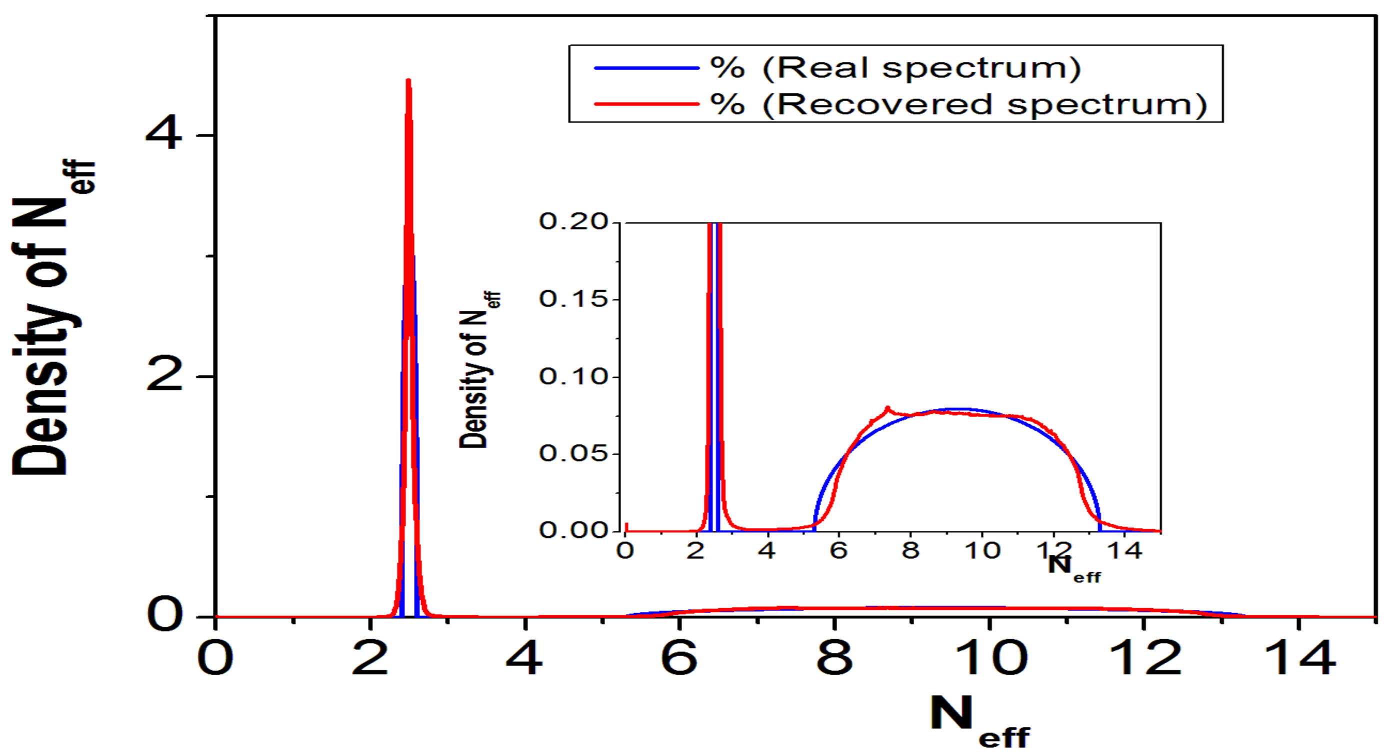
Stochastic Optimization method

$$G(\tau) = \int_0^{+\infty} \rho(\omega) e^{-\tau\omega}$$

$$G_{N_{eff}}(B) = \frac{1}{\sqrt{2\pi[\sigma_0/\sqrt{N_{eff}}]^2}} \exp\left[\frac{(B-B_0)^2}{2[\sigma_0/\sqrt{N_{eff}}]^2}\right]$$

$$I(B) = \int_1^{+\infty} D(N_{eff}) G_{N_{eff}}(B-B_0) dN_{eff} \quad S(B) = \int_1^{+\infty} D(N_{eff}) \left[\frac{d}{dB} G_{N_{eff}}(B-B_0)\right] dN_{eff}$$

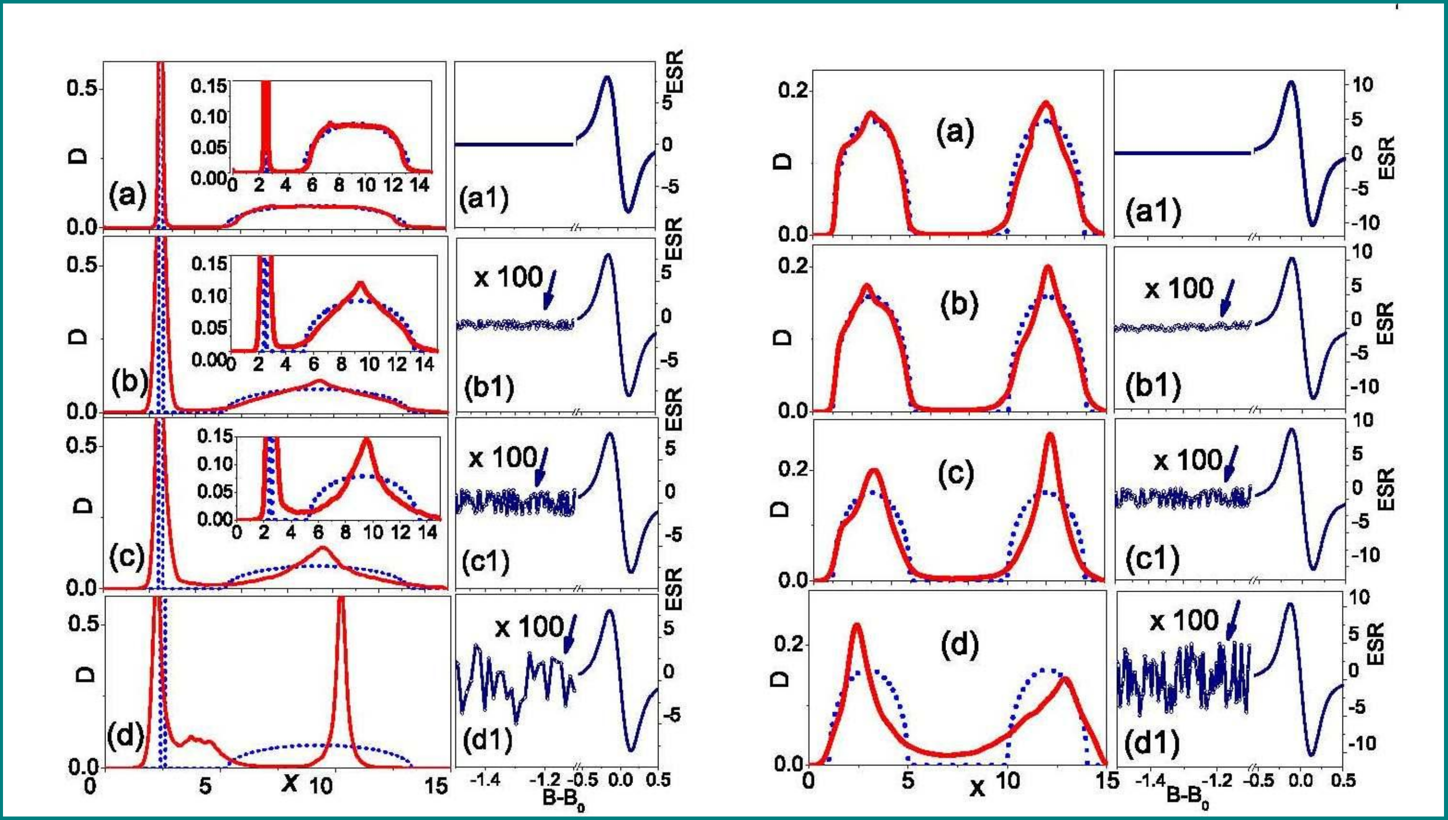
$$G(\tau) = \int_0^{+\infty} \rho(\omega) K[\tau, \omega]$$



$$G_{N_{eff}}(B) = \frac{1}{\sqrt{2\pi[\sigma_0/\sqrt{N_{eff}}]^2}} \exp\left[\frac{(B-B_0)^2}{2[\sigma_0/\sqrt{N_{eff}}]^2}\right]$$

$$I(B) = \int_1^{+\infty} D(N_{eff}) G_{N_{eff}}(B-B_0) dN_{eff} \quad S(B) = \int_1^{+\infty} D(N_{eff}) \left[\frac{d}{dB} G_{N_{eff}}(B-B_0)\right] dN_{eff}$$

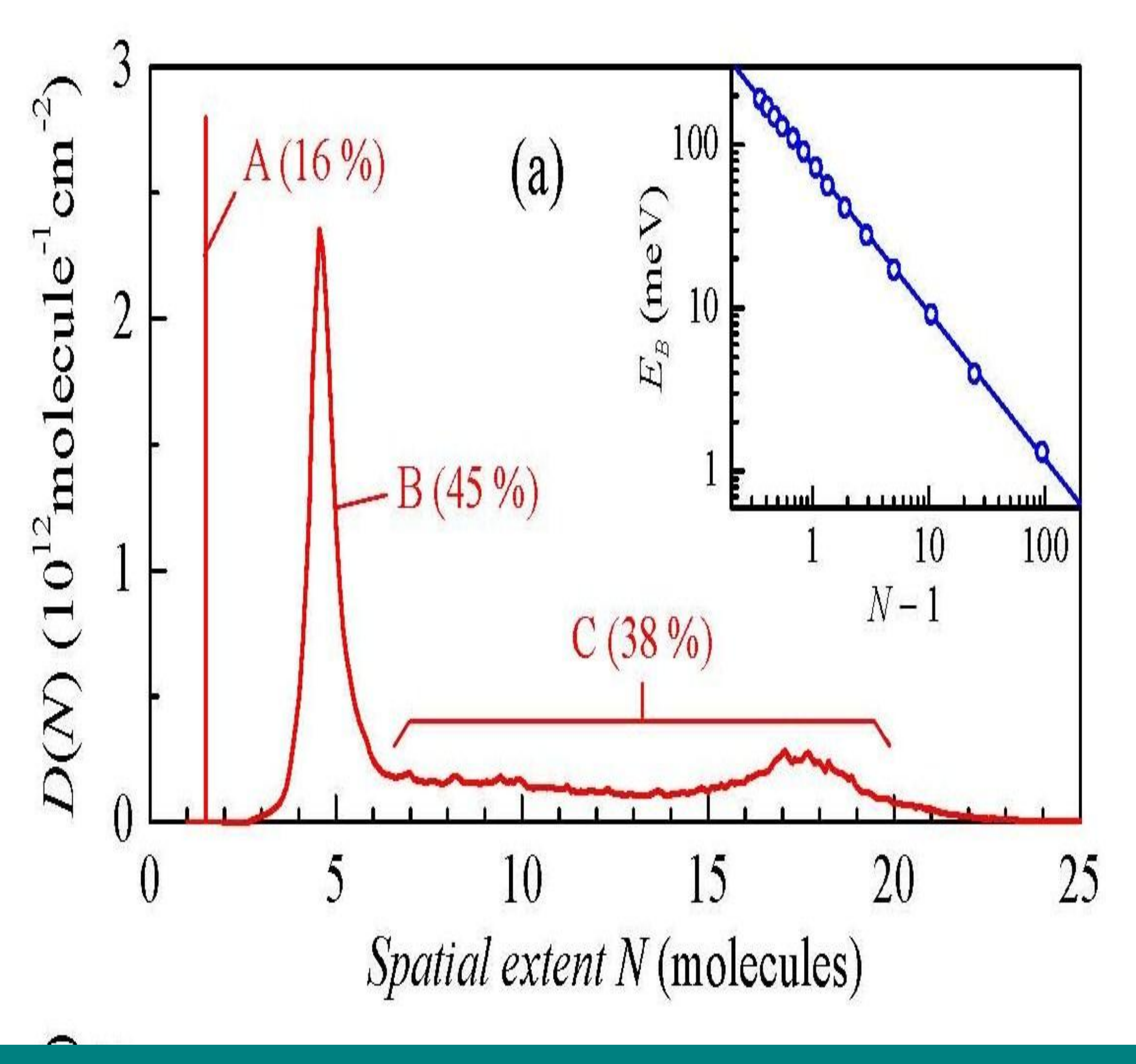
$$G(\tau) = \int_0^{+\infty} \rho(\omega) K[\tau, \omega]$$



ESR spectrum of organic FET

There are 100-s methods to fit the signal.

We say for the 1-st time that we do not need any fit!!!



ESR spectrum of organic FET

Reliability of result: