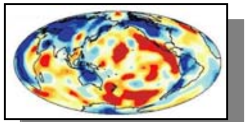
$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Simulated Annealing



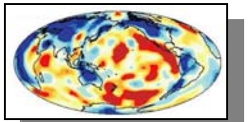
- What is simulated annealing?
- Simulated annealing and probabilistic inversion
- Examples

The problem: we need to efficiently search a possibly multi-modal function in order to either sample the function or find the maximum-likely point of of that function.

We draw on an analogy from solid state physics: the annealing process.

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Simulated Annealing

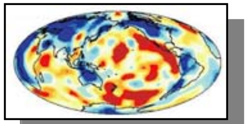


Annealing is the process of heating a solid until thermal stresses are released. Then, in cooling it very slowly to the ambient temperature until perfect crystals emerge. The quality of the results strongly depends on the cooling temperature. The final state can be interpreted as an energy state (crystaline potential energy) which is lowest if a perfectly crystal emerged.

But where's the connection to inverse problems?

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Simulated Annealing



Our goal is to sample a multi-modal function efficiently. We use an analogy between the physical process of annealing and the mathematical problem of obtaining a global minimum of a function.

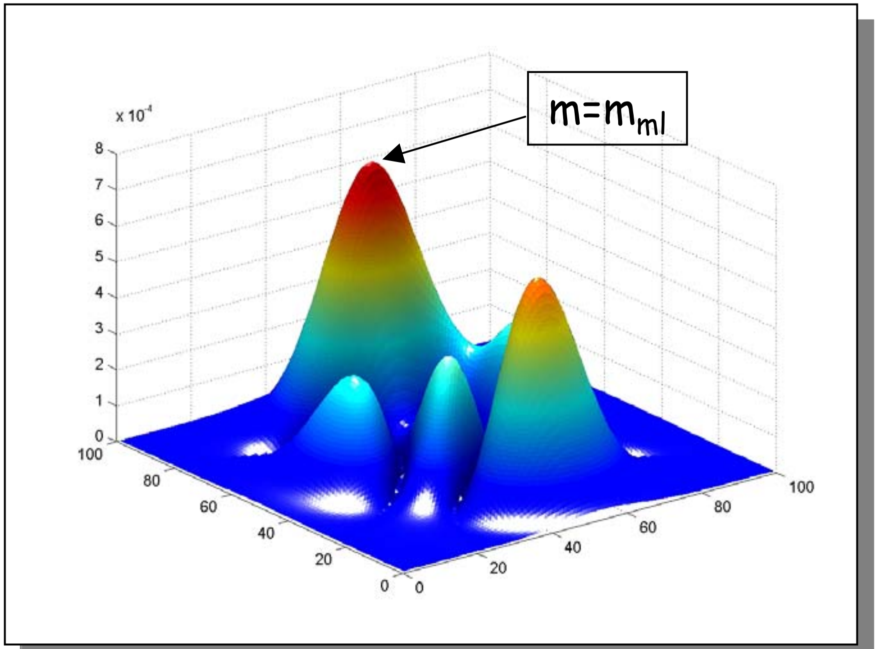
... it may also be turned around so that we try and find the maximum of – for example – a probability density ...

Simulated annealing really aims at finding the maximum likelihood point m_{ml}

$\sigma_M(m)$ Maximum for $m=m_{ml}$

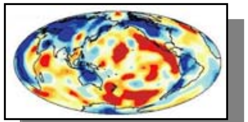
We first define an energy function:

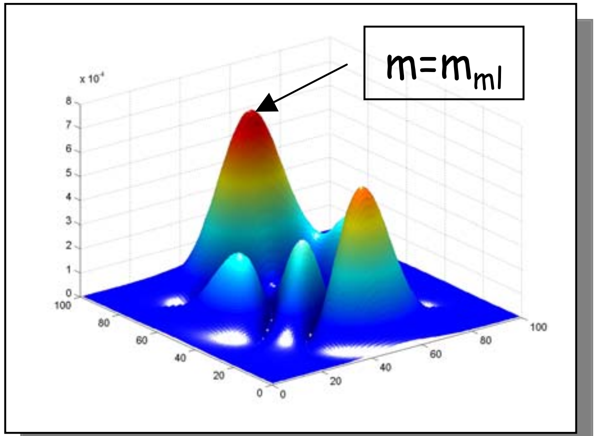
$$S(m) = -T_0 \log\left[\frac{\sigma_M(m)}{\mu_M(m)}\right]$$



$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Simulated Annealing





T_0 is a fixed positive number termed the ambient temperature (e.g. T=1). We obtain the probability density function

$$\sigma_M(m,T) = \rho_M(m)\exp\left(-\frac{S(m)}{T}\right)$$

written out ...

$$\sigma_M(m,T) = \rho_M(m)\exp\left(-\frac{-T_0\log\dfrac{\sigma_M(m)}{\rho_M(m)_o}}{T}\right)$$

The probability thus defined has interesting properties:

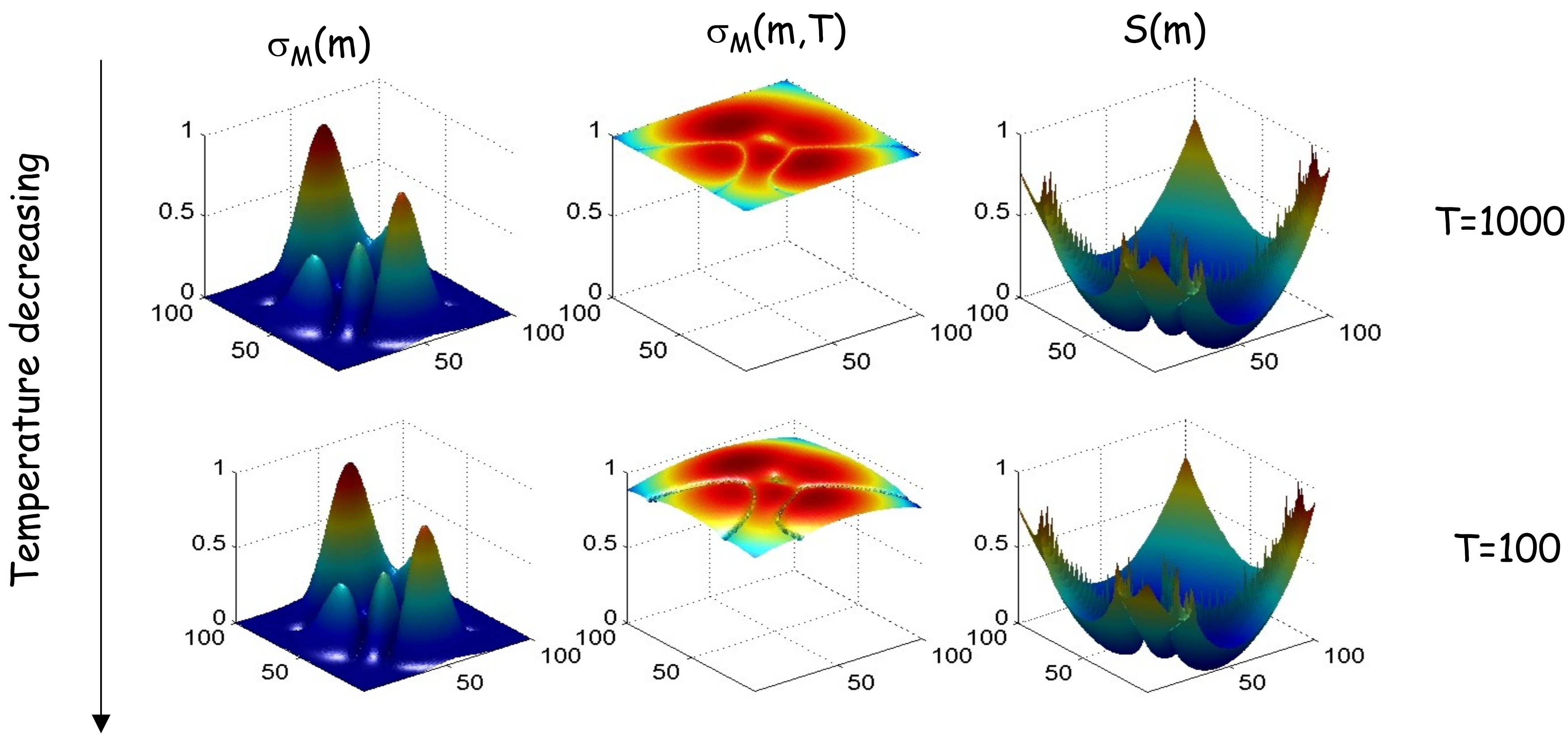
$$\sigma_M(m,T=\infty) = \rho_M(m)$$
$$\sigma_M(m,T=T_0) = \sigma_M(m)$$
$$\sigma_M(m,T=0) = const\,\delta(m-m_{ml})$$

$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

The Heat Bath

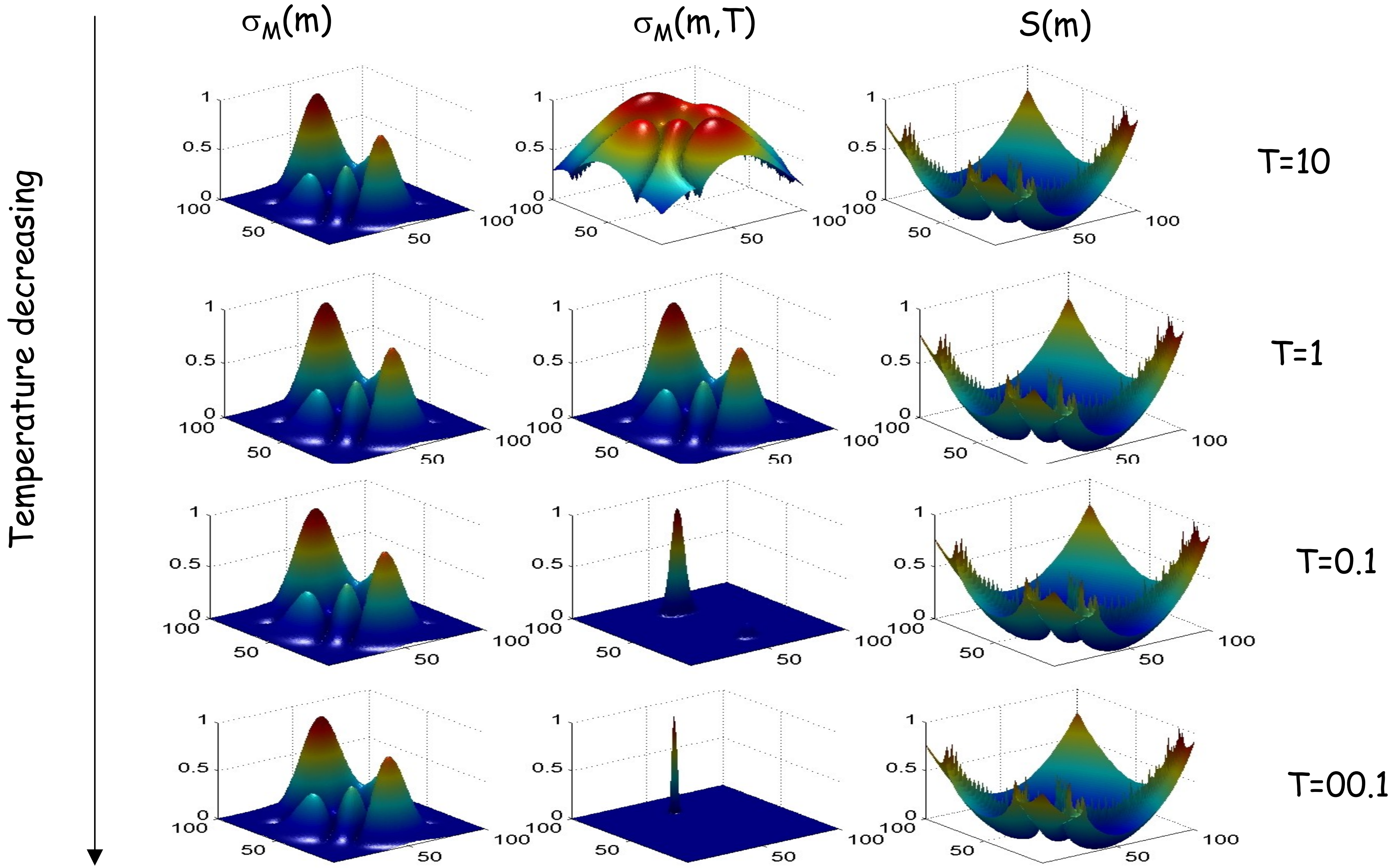
$$\sigma_M(m,T) = \rho_M(m)\exp\left(-\frac{-T_0 \log\dfrac{\sigma_M(m)}{\rho_M(m)_o}}{T}\right)$$

$\sigma_M(m)$ -> „ peaks „ (as pdf)
$\rho_M(m)$ -> constant
T_0=1



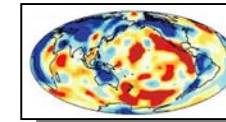
$$\sigma(d,m) = k\,\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

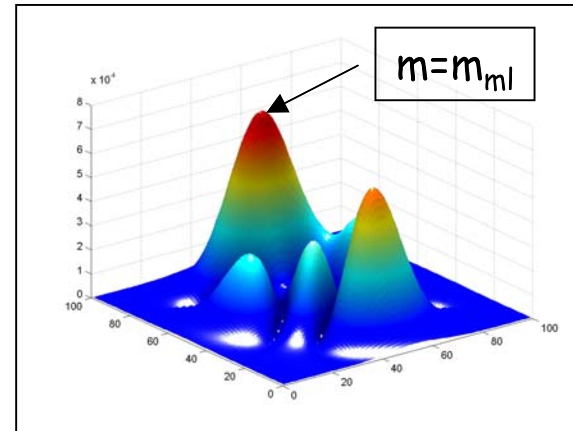
The Heat Bath



$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Simulated Annealing





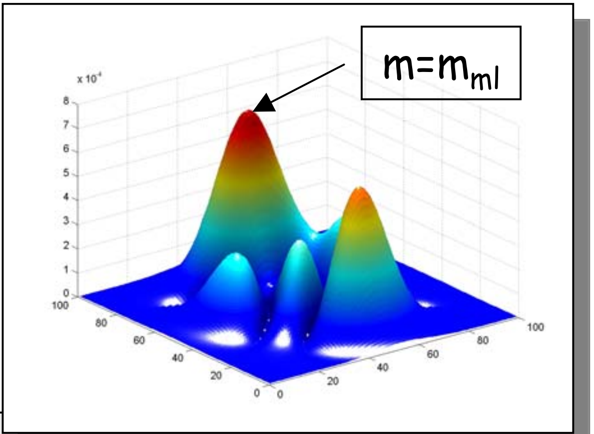
For constant prior denstiy this function resembles The Gibbs distribution, giving the probability of State m with energ S(m) of a statistical system at temperature T.

The procedure would be:

1. Take a high temperature T (heat the system) and generate random models this means we re effectively sampling the prior distribution
2. Cool the system slowly while continuing to generate random models until T=0. you should now be in the global minimum.

The efficienvcy strongly depends on the cooling procedure. If too fast you may end up in scondary minima. If too slow you will waste a lot of forward calculations.

Simulated Annealing



Here is a pseudo-code. It is only a slight modifiction to the Metripolis algorithm

Simulated annealing:

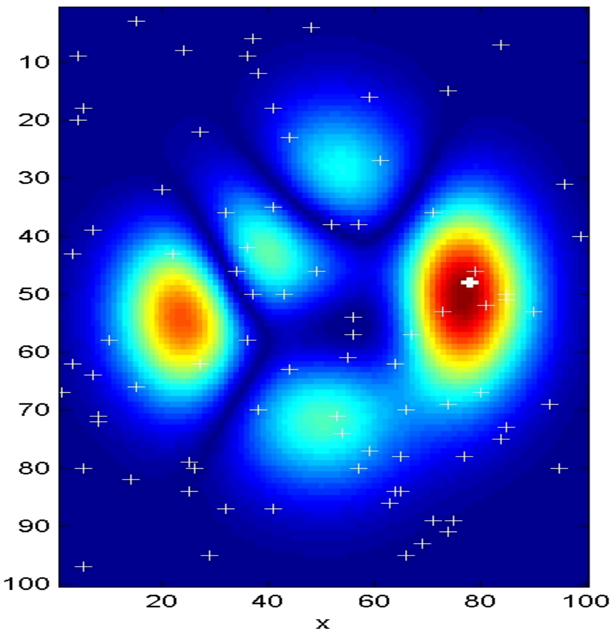
```
Define a high temperature T
Define a cooling schedule T(it), e.g. T=alpha T
Define an energy function S
Define  current_model initial state

        While (not converged)
                        new_model = random
                        Delta_S =  S(new_model)-S(current_model)
                        If (Delta_S < 0) current_model = new_model
        Else with probability P = e^(-Delta_S/T) : current_model = new_model
        T=alpha T
```

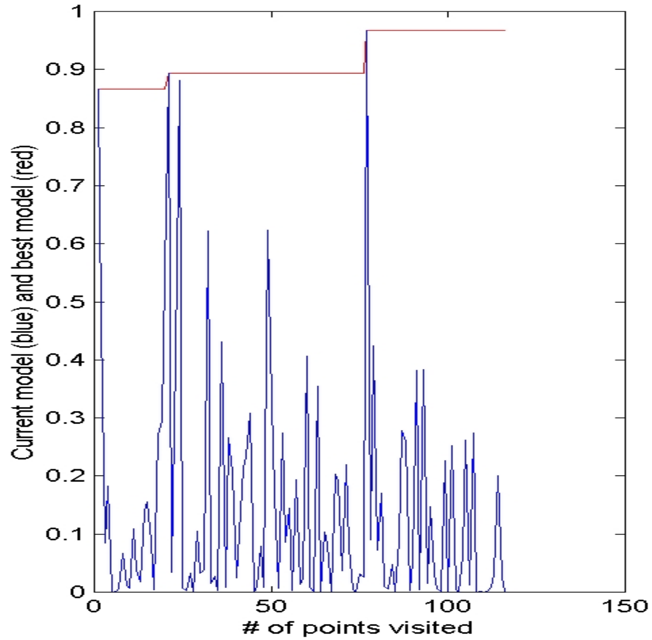
$$\sigma(d,m) = k\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Simulated Annealing

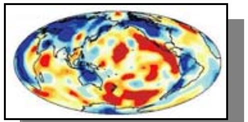
Cooling schedules could be:

- alpha(T) = a T, 0.8 < a < 0.995
- alpha(T) = T+bT, b close to 0
- alpha(T) = c/log(1+k) , k is the iteration number and c is a constant





Simulated Annealing



We will adopt a special approach (Rothmann, 1986):

Define a high temperature T
Define a cooling schedule T(it), e.g. T=alpha T
Define an energy function S and the associated pdf
Define current_model initial state

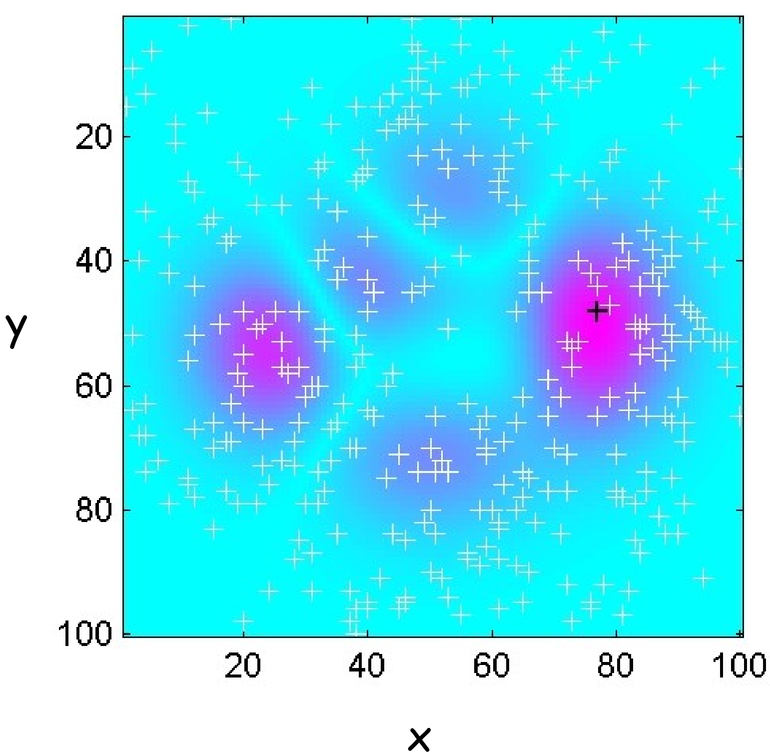
```
While (not converged)
        new_model = random
        calculate P(new_model)
        generate random number x in (0,1)
        accept with Metroplis rule
update T
```

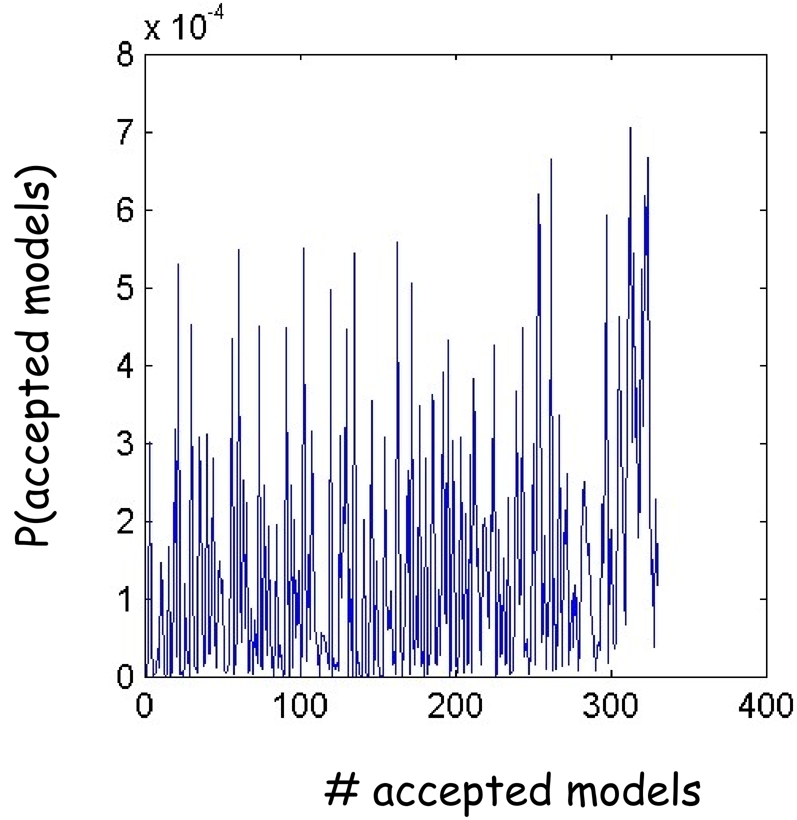
$$\sigma(d,m) = k\,\frac{\rho(d,m)\theta(d,m)}{\mu(d,m)}$$

Simulated Annealing

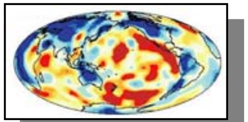
Example with peaks function: T_0=1, T_a=50
1000 iterations (ca. 340 accepted model updates)

A posteriori probability





Summary



Simulated annealing is an mathematical analogy to a cooling system which can be used to sample highly nonlinear, multidimensional functions.

There are many flavors around and the efficiency strongly depends on the particular function to sample. Therefore it is extremely difficult to make general statements as to what parameters work best.