Inverse Problems in Geophysics
Part 11: Probability and Likelihood

2. MGPY+MGIN

Thomas Günther

thomas.guenther@geophysik.tu-freiberg.de

TUBAF
Die Ressourcenuniversität.
Seit 1765.

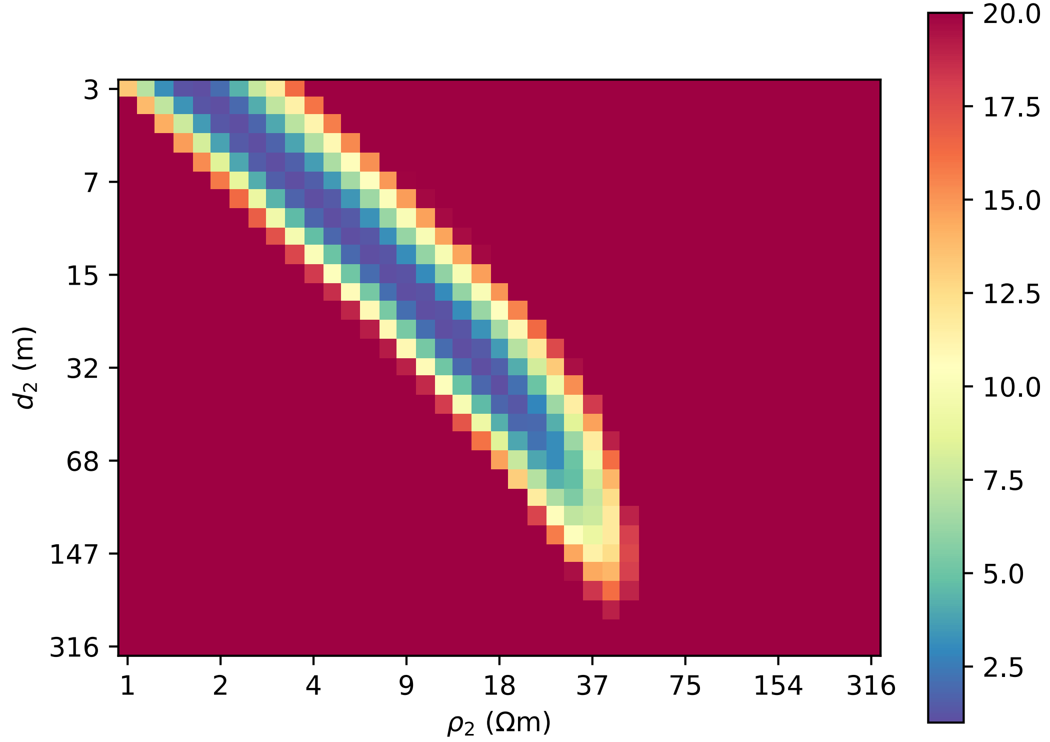
Recap non-linear

undirected search methods

- grid search
- random sampling (Monte Carlo)
- simulated annealing

directed search methods

- gradients (steepest descent)
- Newtons method (linearization)



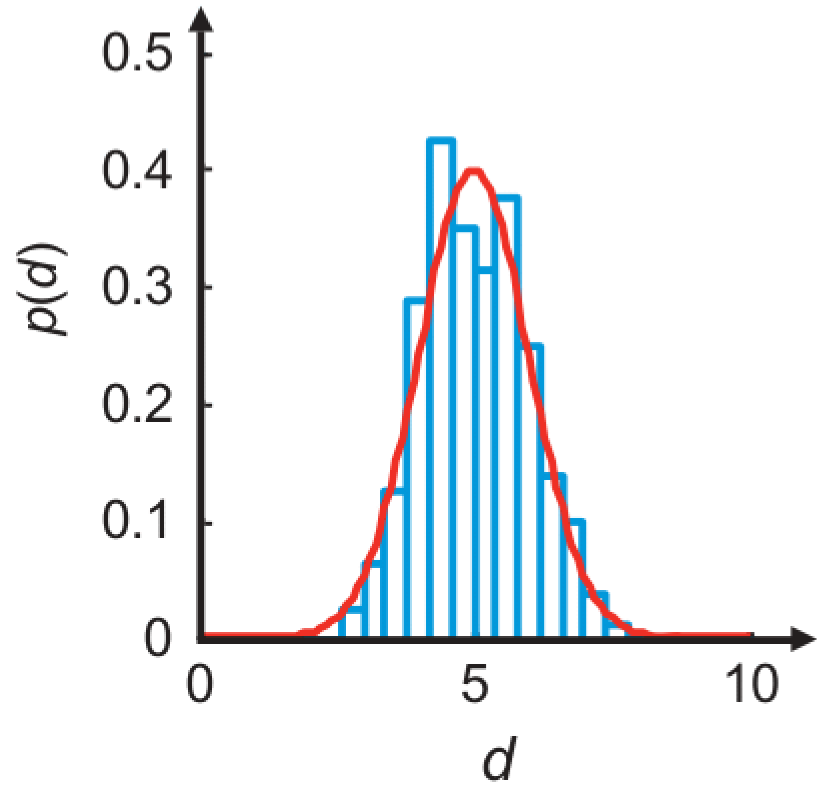
Objective function for VES

Probability and likelihood

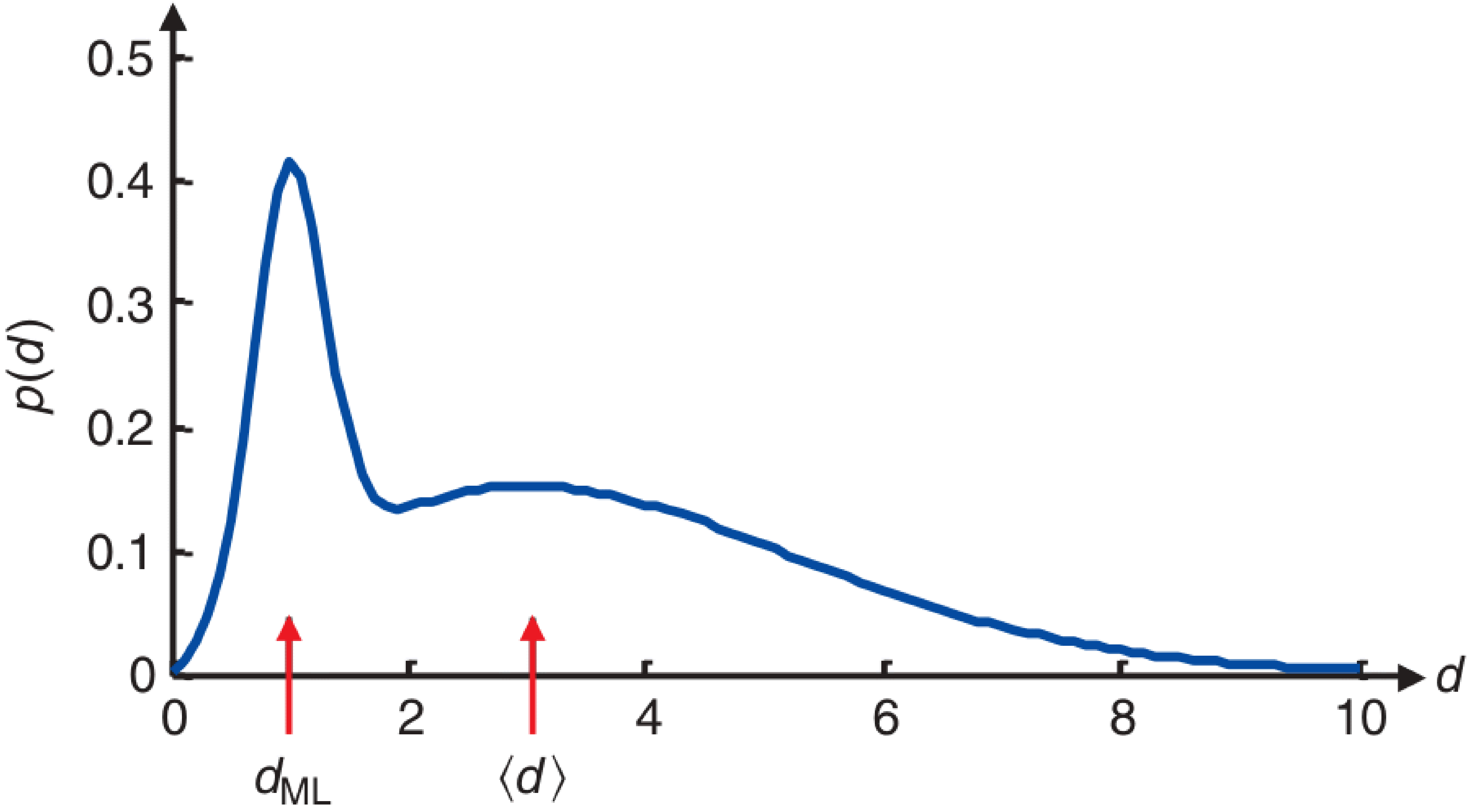
random variables: probability through many repetitions

maximum $p(d)$ is most likely

expectation: $\langle d \rangle = \int d \cdot p(d)\,\mathrm{d}d$



Probability density function



Variance

$$\sigma^2 = \int (d - \langle d \rangle)^2 p(d)\, \mathrm{d}d$$

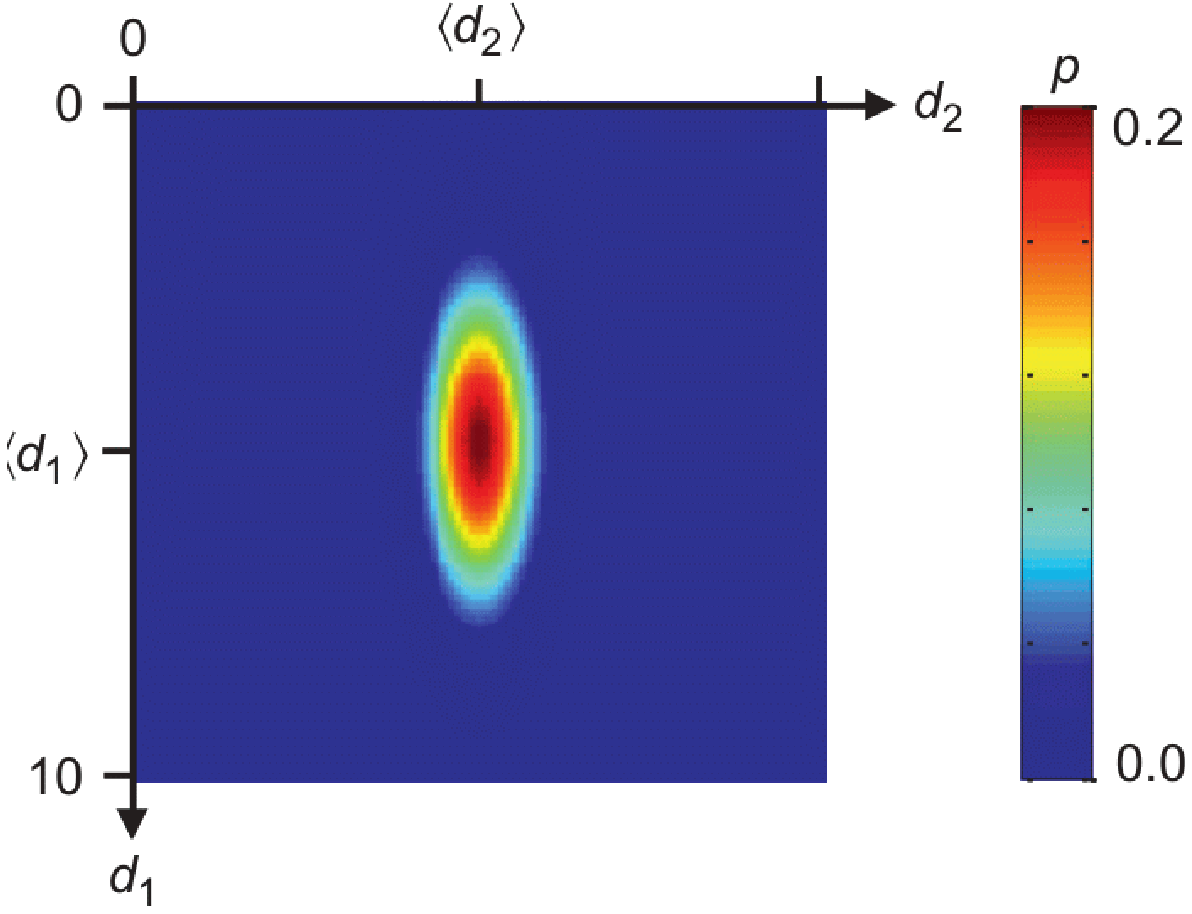
σ is a measure of the width of the distribution

related to standard deviation and mean of sampling

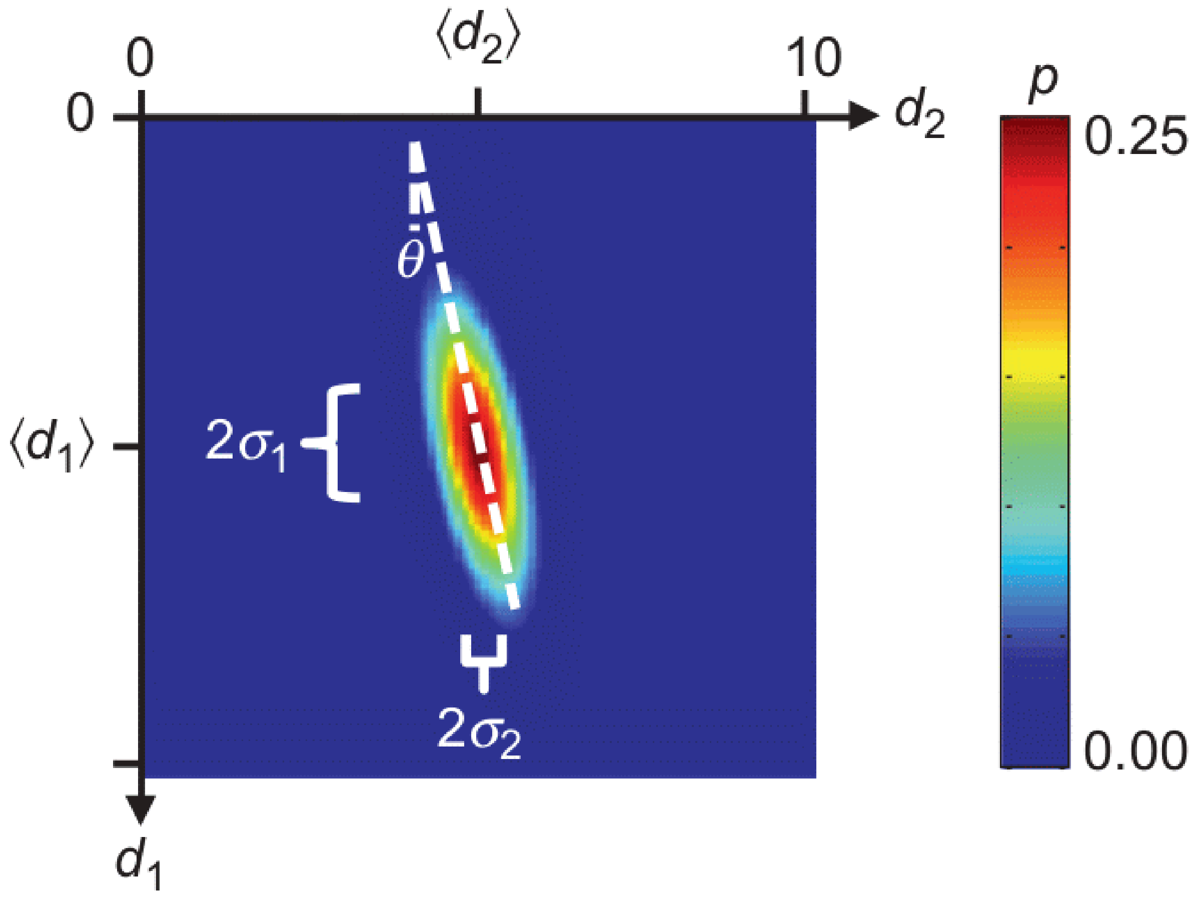
$$\sigma_{est}^2 = \frac{1}{N-1} \sum_{i=1}^{N} (d_i - \langle d \rangle)^2 \quad \text{with} \quad \langle d \rangle = \frac{1}{N} \sum_{i=1}^{N} d_i$$

Correlated data

independent: $p(\mathbf{d}) = p(d_1)p(d_2)\ldots p(d_N)$



uncorrelated data (Menke, 2012)



correlated data (Menke, 2012)

Covariance

(measure of correlation between data)

$$\mathrm{cov}(d_1, d_2) = \int\int (d_1 - \langle d_1 \rangle)(d_2 - \langle d_2 \rangle) p(d_1, d_2)\, \mathrm{d}d_1\, \mathrm{d}d_2$$

$$\langle d_i \rangle = \int \dots \int d_i p(\mathbf{d})\, \mathrm{d}d_1 \dots \mathrm{d}d_N$$

Covariance propagation

Linear problem $\mathbf{m} = \mathbf{M}\mathbf{d}$, e.g. $\mathbf{m} = \mathbf{G}^{\dagger}\mathbf{d}$

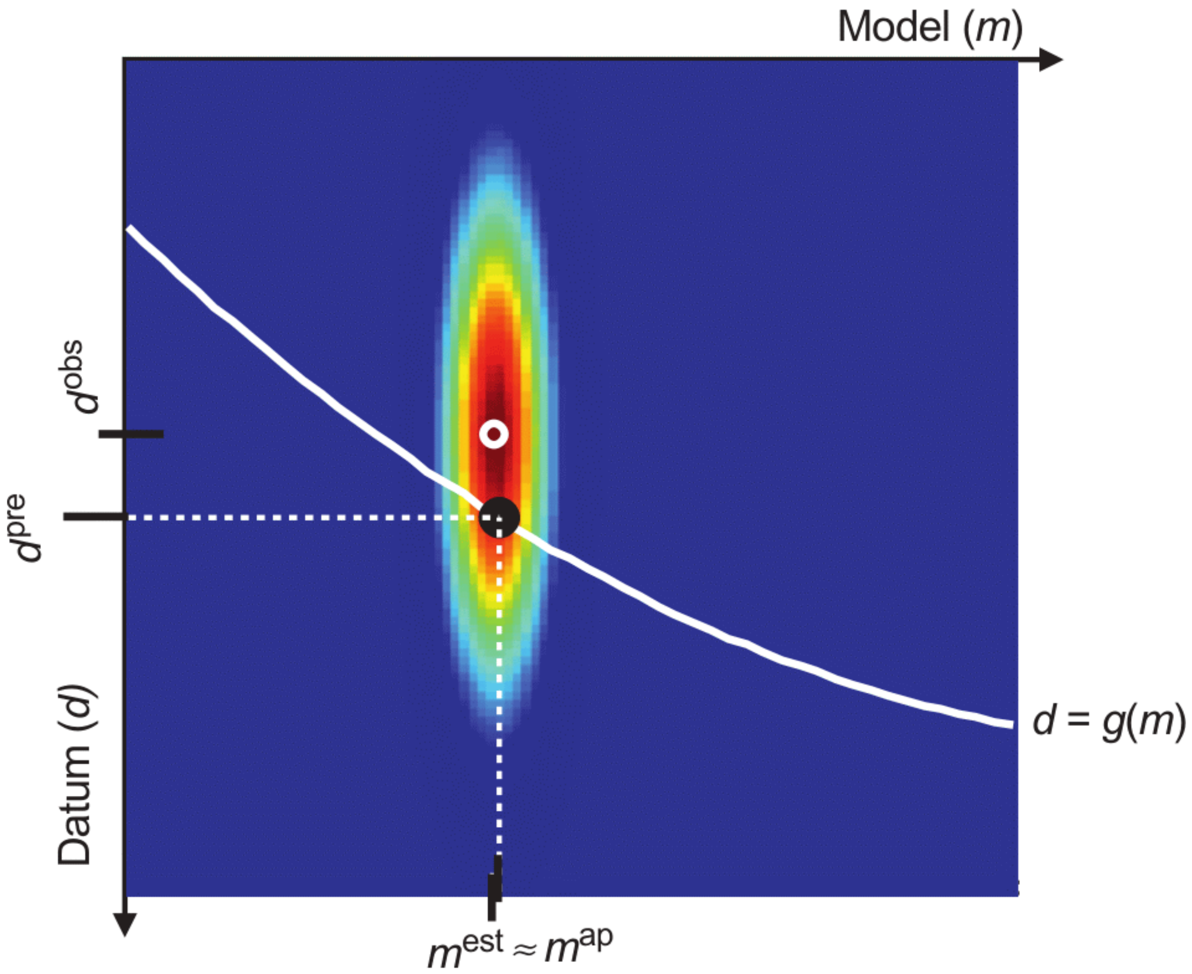
Mean value $\langle \mathbf{m} \rangle = \mathbf{M} \langle \mathbf{d} \rangle + \mathbf{n}$ and covariance

$$\mathrm{cov}(\mathbf{m}) = \mathbf{M}\mathrm{cov}(\mathbf{d})\mathbf{M}^T$$

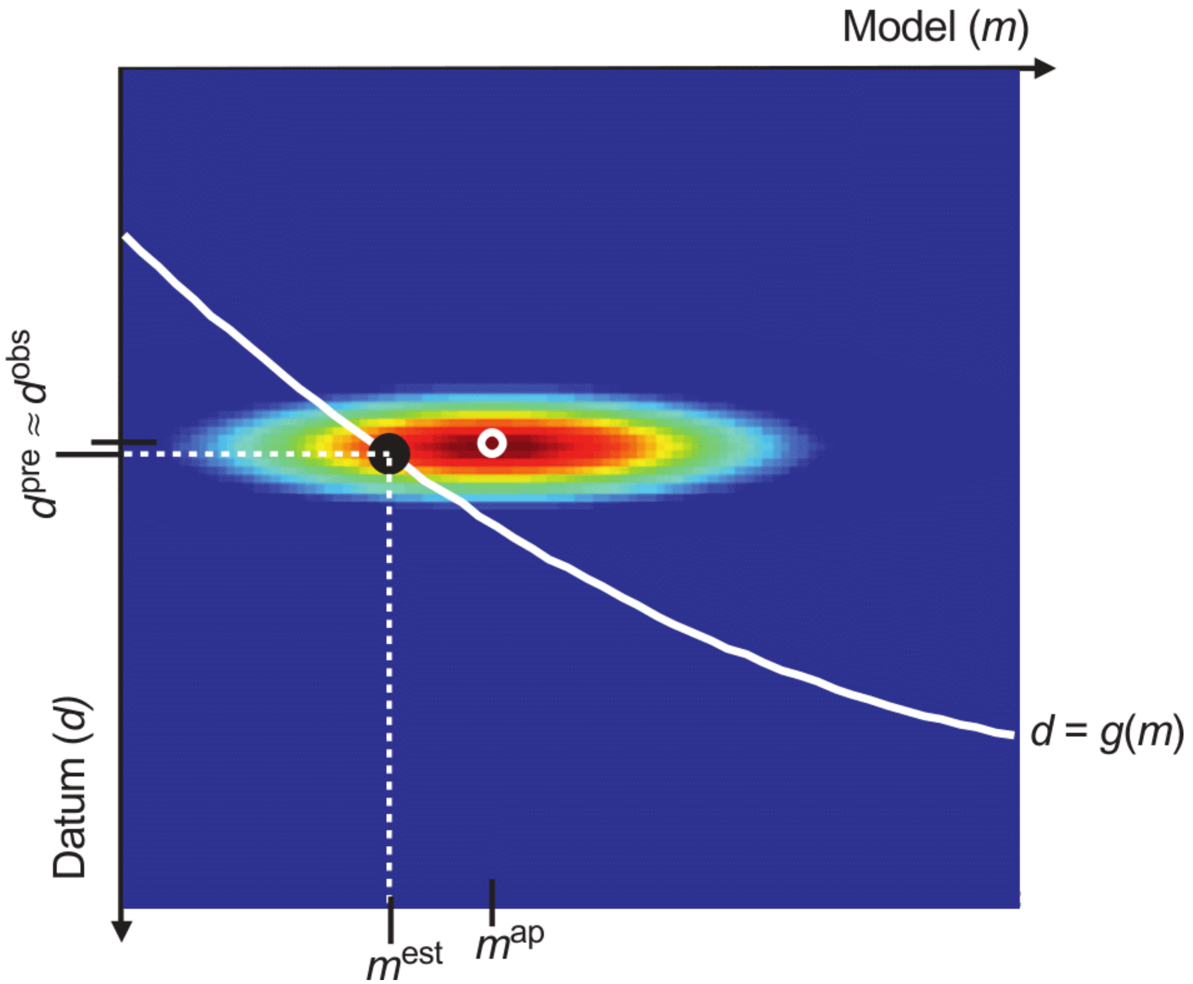
Least-squares: $\mathbf{M} = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T$, uncorrelated data: $\mathrm{cov}(\mathbf{d}) = \sigma_d^2\mathbf{I}$

$$\Rightarrow \mathrm{cov}(\mathbf{m}) = (\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\sigma_d^2\mathbf{I}((\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T)^T = \sigma_d^2(\mathbf{G}^T\mathbf{G})^{-1}$$

A priori knowledge



accurate prior model (Menke, 2012)



accurate data (Menke, 2012)

Bayes' theorem

Conditional probability

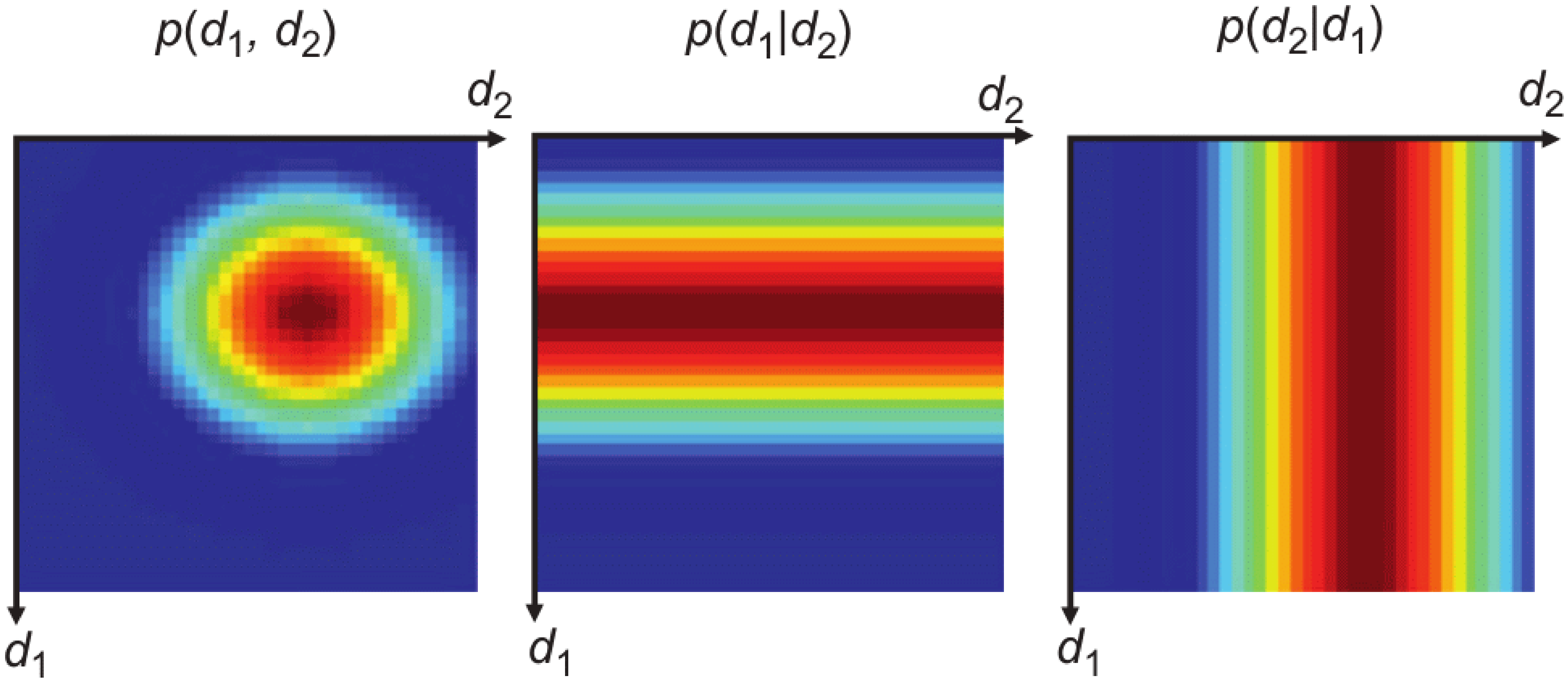
$$p(d_1|d_2) = p(d_1, d_2)/p(d_2)$$

$$p(\mathbf{m}|\mathbf{d})p(\mathbf{d}) = p(\mathbf{d}|\mathbf{m})p(\mathbf{m})$$

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})}$$

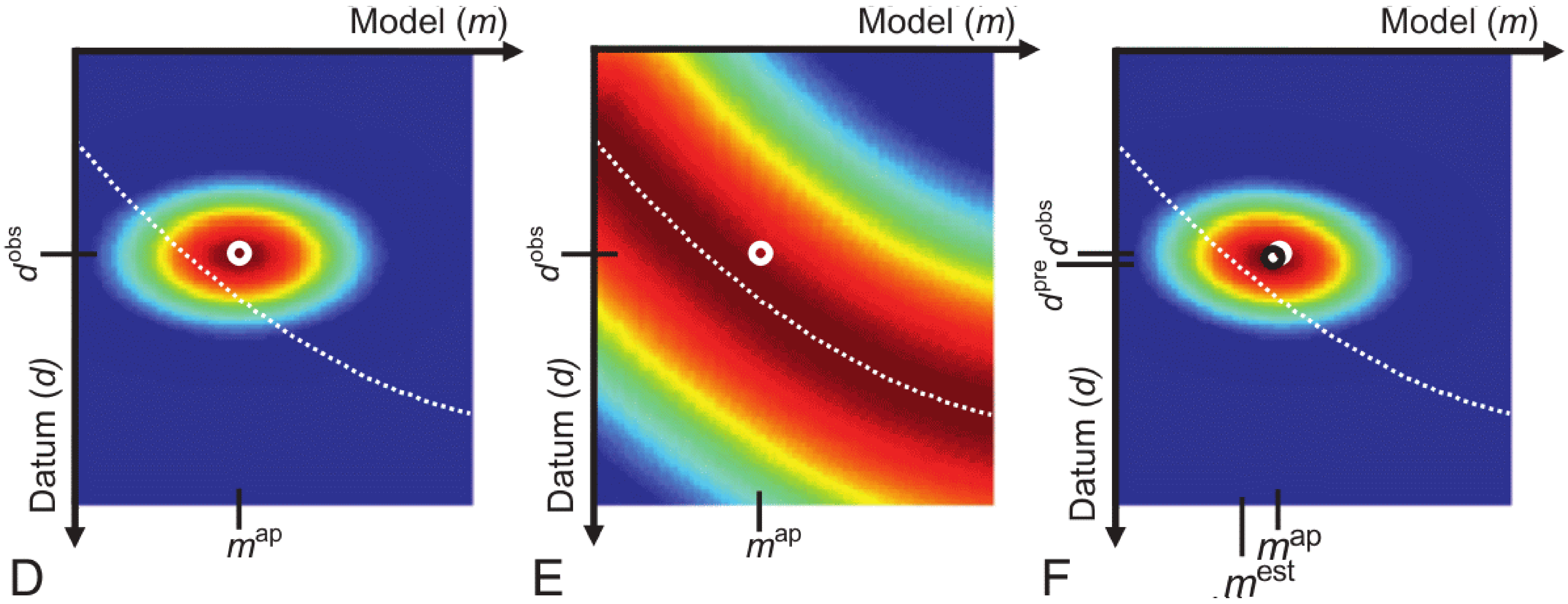
posterior distribution $\propto$ likelihood x prior distribution

Bayes theorem simple example



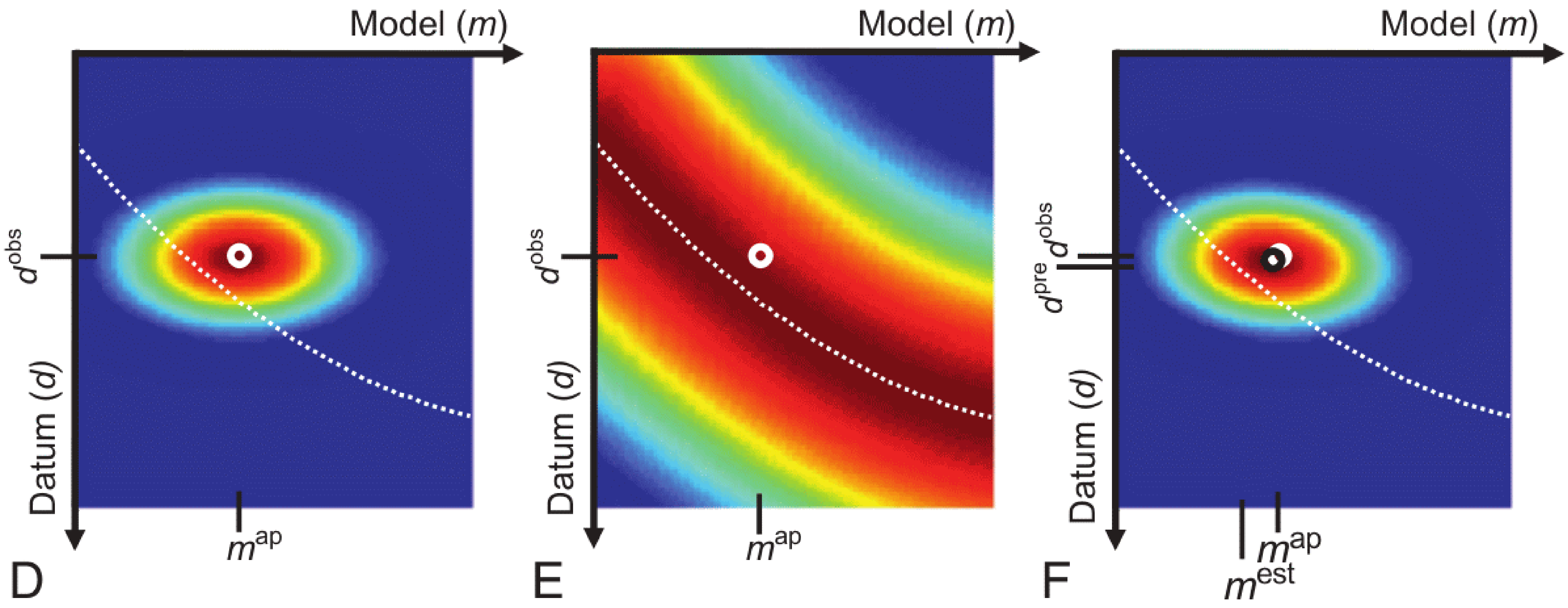
joint probability & conditional probabilities (Menke, 2015)

A priori (Menke, 2012)

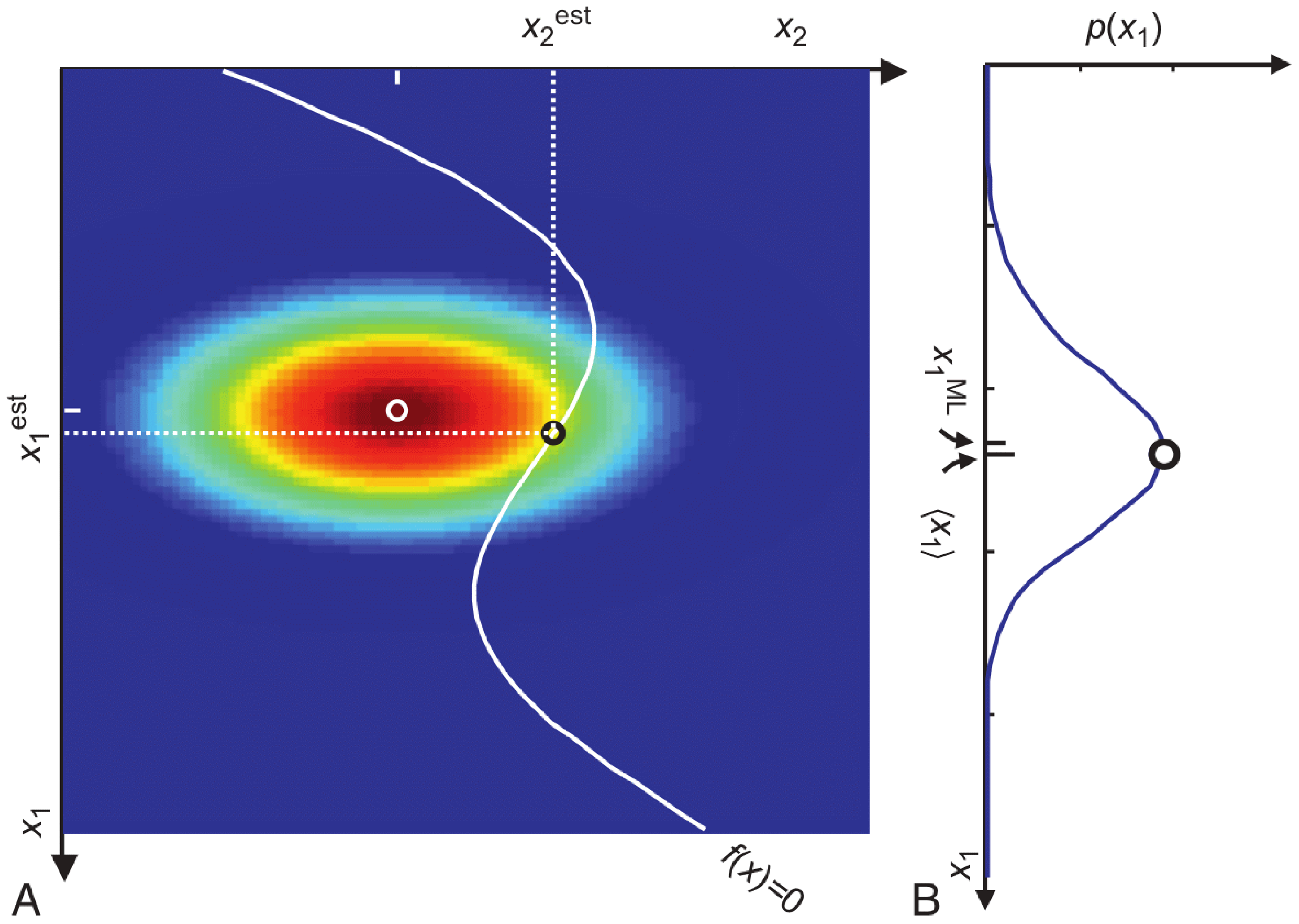


A: *a priori pdf* $p_a(\mathbf{m},\mathbf{d})$, B: *conditional pdf* $p_g(\mathbf{m},\mathbf{d})$, C: product $p_t(\mathbf{m},\mathbf{d}) = p_a(\mathbf{m},\mathbf{d})p_g(\mathbf{m},\mathbf{d})$, white: theory

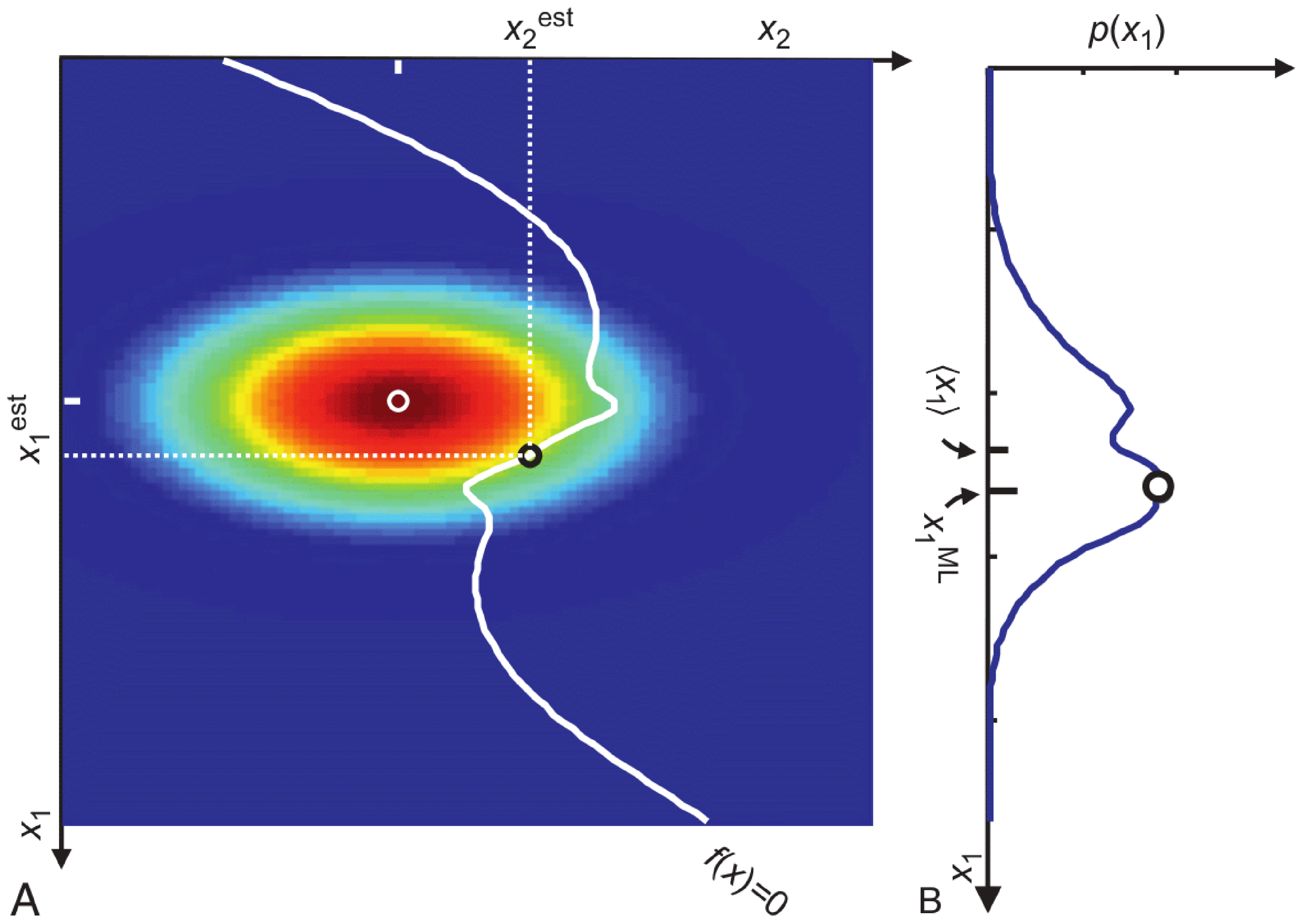
A priori and likelihood



Bayes view in nonlinear problems

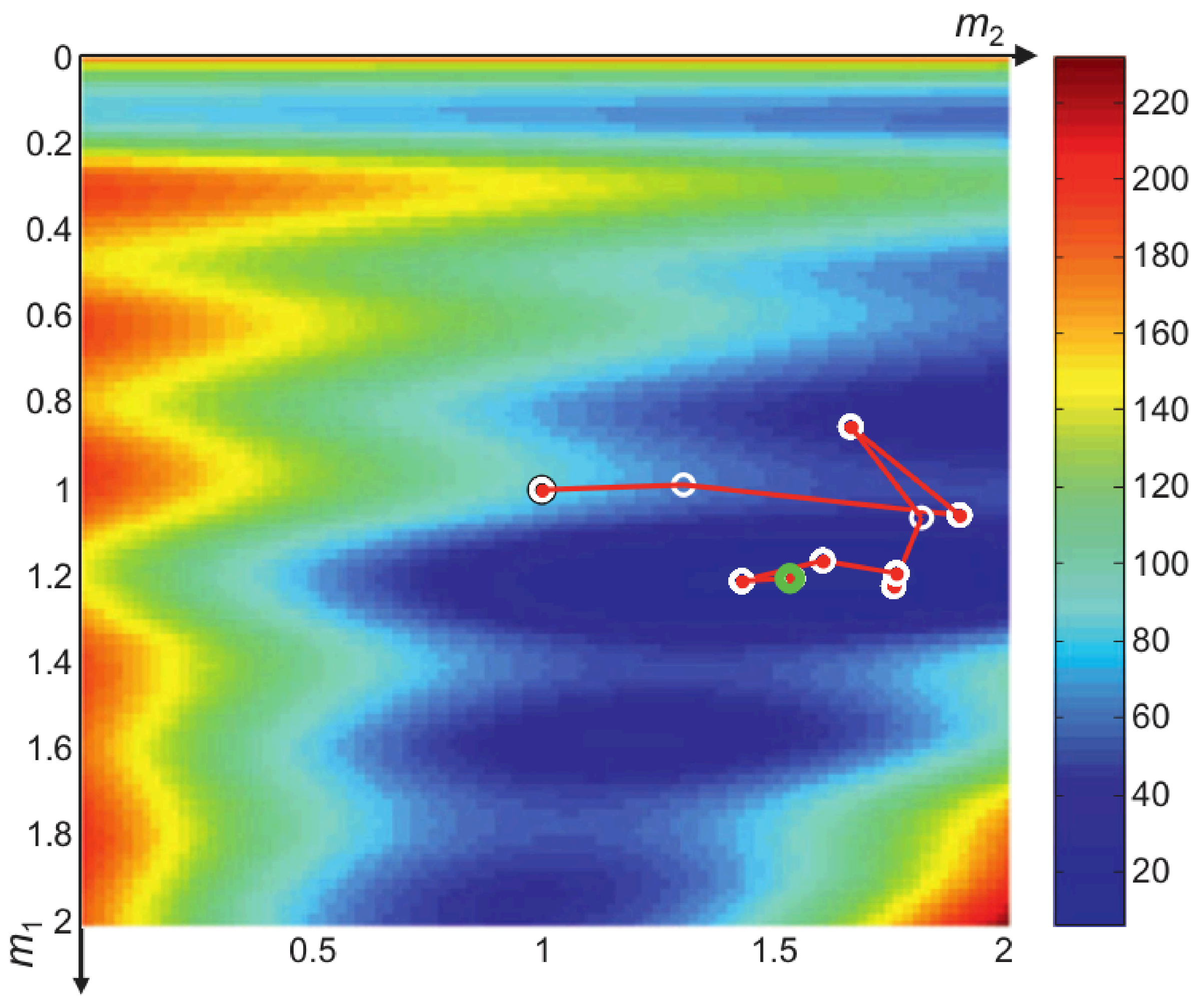


Highly nonlinear problems



Monte Carlo methods

- Monte Carlos search: randomly draw solutions from grid
- accept solution only if better than old
- Markow-Chain-Monte-Carlo
- Metropolis-Hastings (Metropolis et al., 1953; Hastings, 1970)

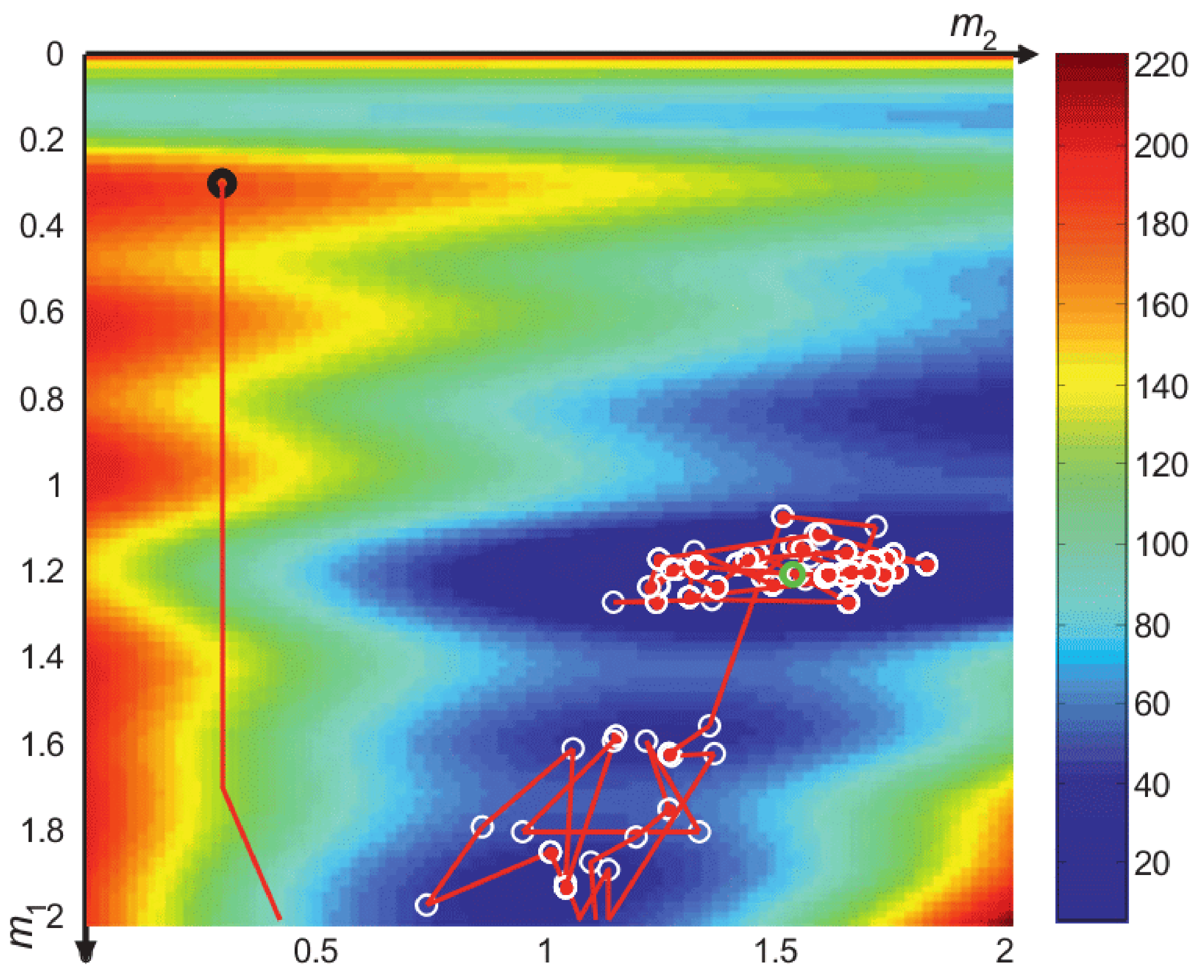


Monte Carlo method

Simulate Annealing

Test parameter

$$t = e^{-(\Phi(\mathbf{m}) - \Phi(\mathbf{m}^p))/T}$$



Simulated Annealing

Particle swarm optimization

Alternatives to grid search

Monte Carlo search

draw random samples and accept them if the error is improved

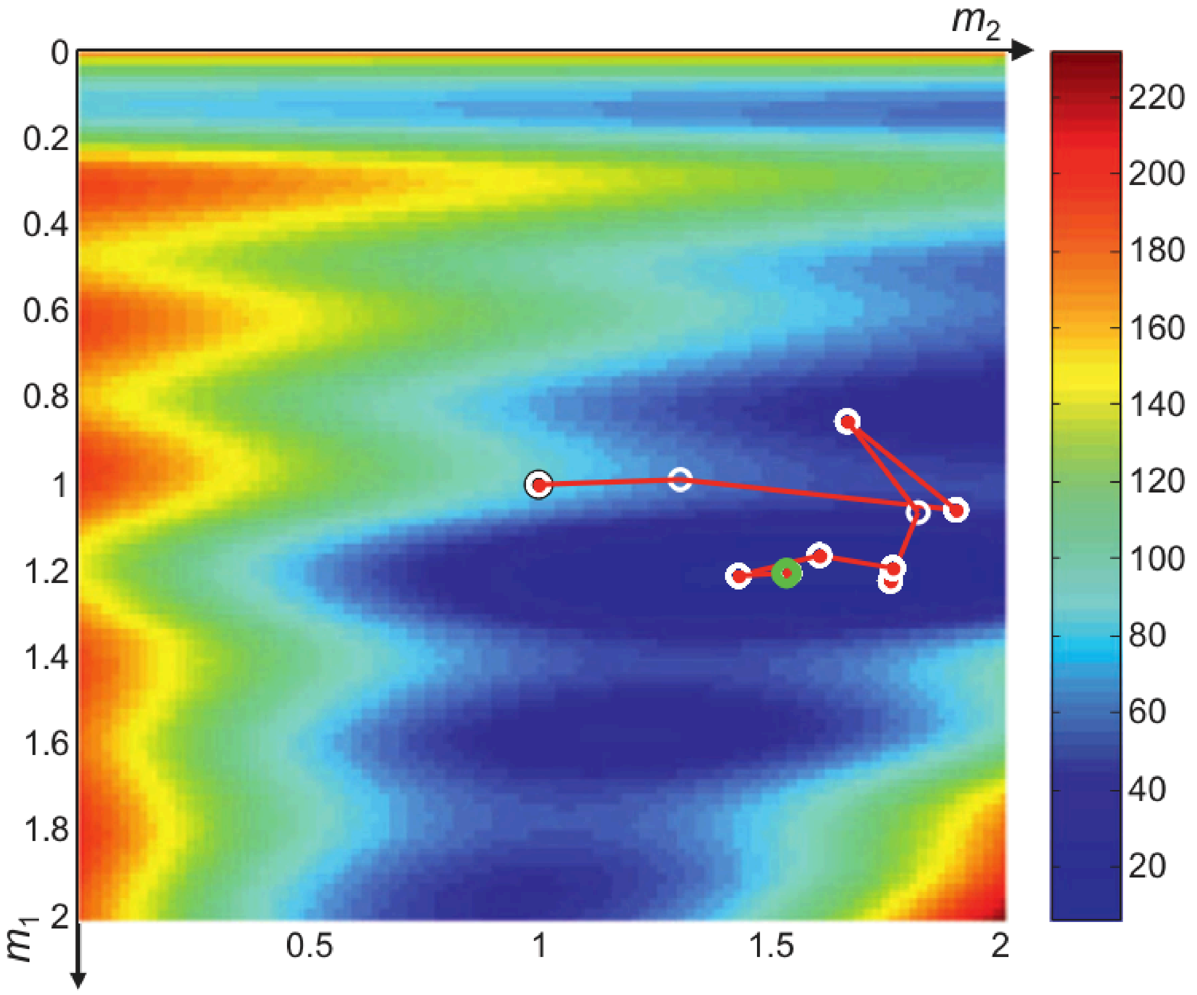
undirected search (Newtons method is directed)

Simulated annealing

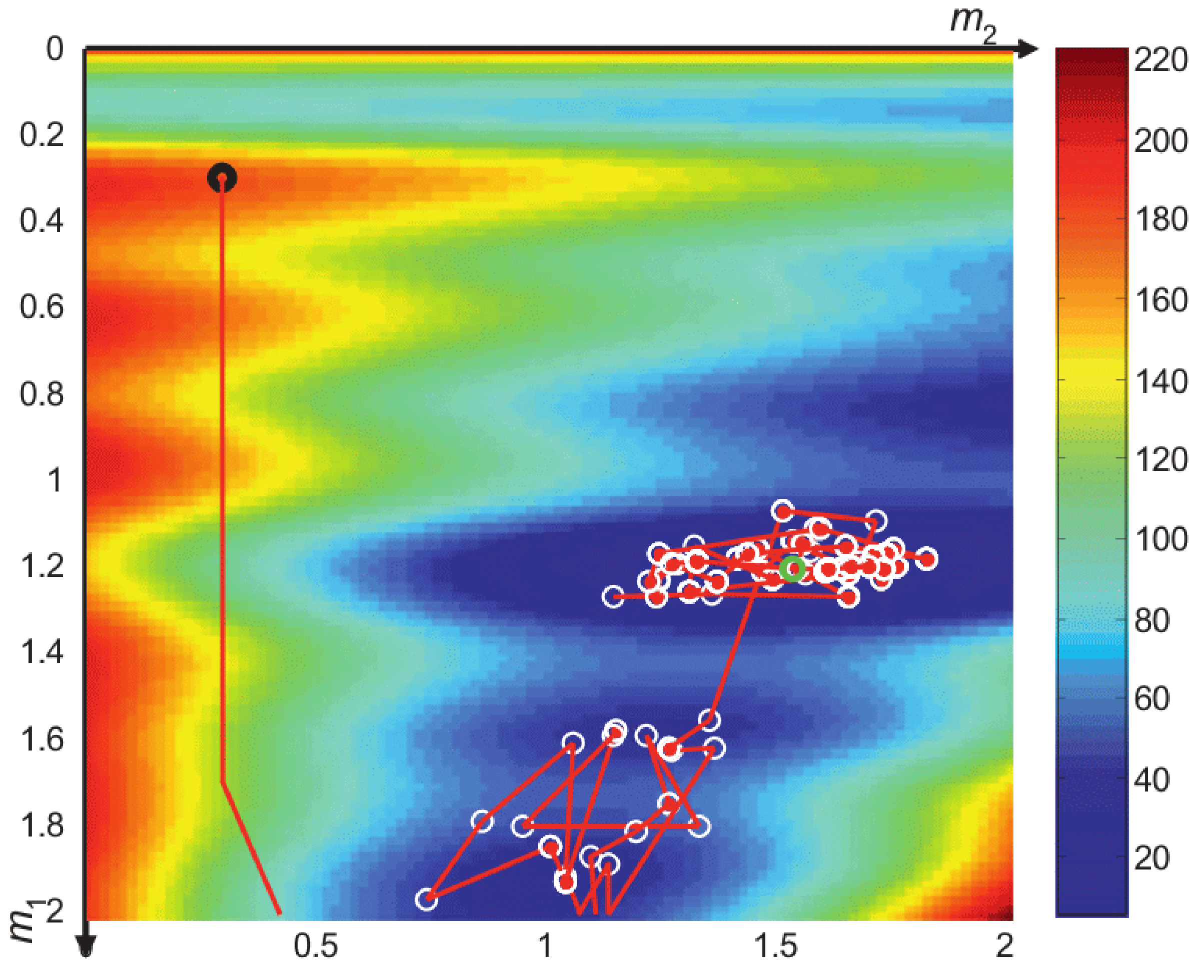
decrease temperature controlling particle movements:

high T: undirected, low T: search in vicinity of current model

Monte Carlo vs. Simulated Annealing

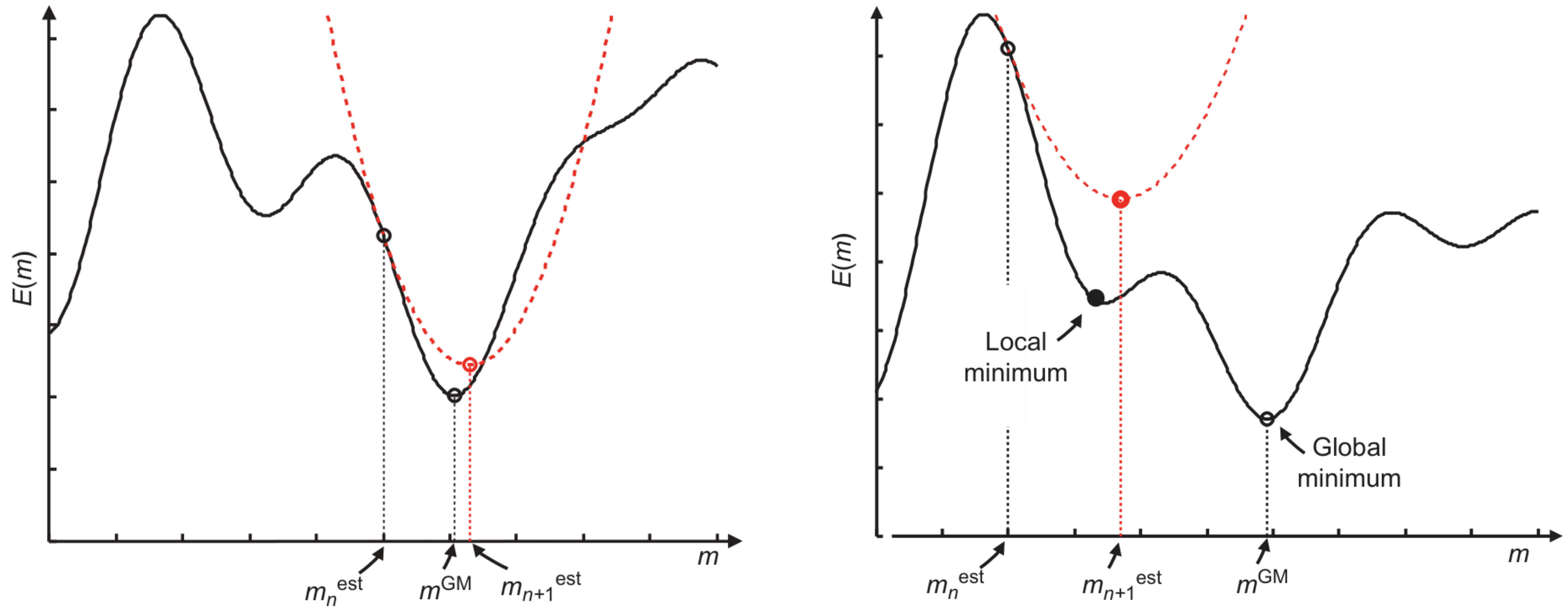


Monte Carlo method



Simulated Annealing

Newtons method (Menke, 2012)



linearize with value, slope and curvature of Φ_d