Correlated Knowledge Gradients for Ranking and Selection of Many Alternatives

Peter Frazier

Presenting joint work with Warren Powell, Savas Dayanik, Diana Negoescu, and Warren Scott

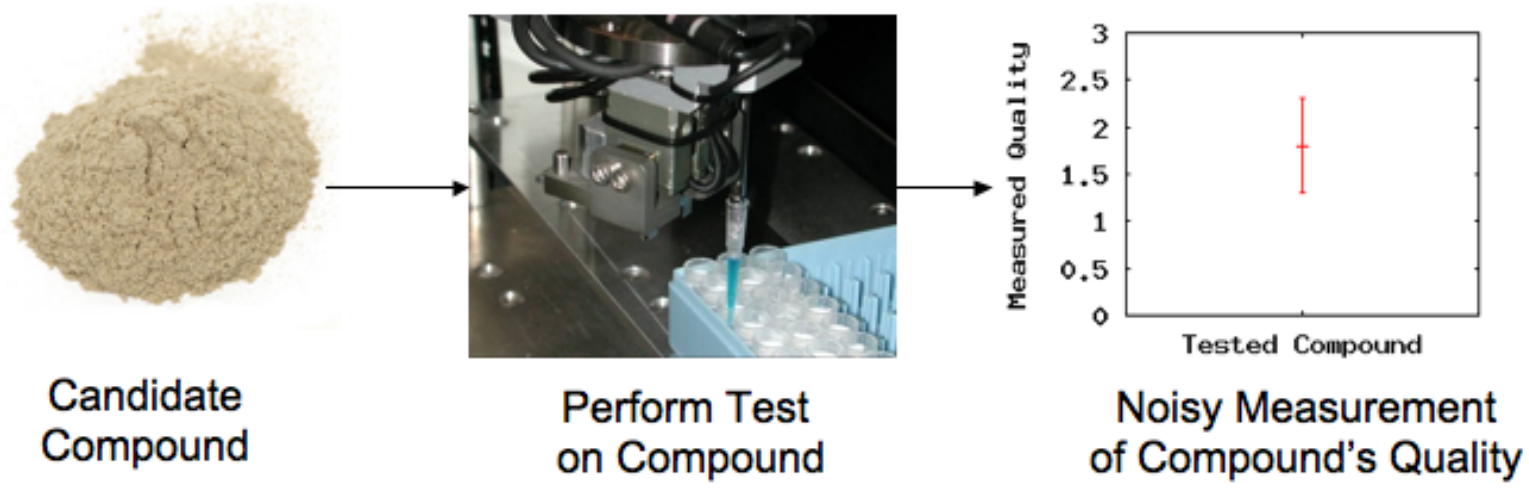
Operations Research & Information Engineering, Cornell University

Thursday February 4, 2010
Department of Applied Mathematics & Statistics
Johns Hopkins University

Outline

Drug Discovery

- We have a large number ($10^4 - 10^6$) of chemical compounds.
- We want to find a compound that performs well in a laboratory test.
- We only have time and money to perform a limited number of tests.
  - Typically, the budget allows for between 10^2 and 10^3 tests.
- Question: which tests should we perform?



Sources: http://www.paa.co.uk/img/labauto/inst_highres/ssi/mini_dispenser.jpg,

http://www.kalyx.com/store/images/Images_SW/SW_201442-51.jpg

Simulation Optimization

- We have a large discrete event simulator with which we can estimate the consequences of a future real-world decision. For example:
  - Designs of a queuing network.
  - Inventory policies for a supply chain.
  - Pricing strategies for a revenue management problem.
  - ...
- We would to find a real-world decision that will work well, according to the simulator.
- Our simulator requires significant time to accurately characterize a decision, and we do not have enough time to do so for each one.
- Question: which real-world decisions should we simulate, and how accurately?

Ranking & Selection (R&S)

1. Nature fixes $\theta \in \mathbb{R}^k$. θ_x is the true value of alternative x.
2. For $n = 0$ up to $N - 1$, where our sampling horizon is N,
   1. Choose x_{n+1} based on $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$.
   2. Observe $y_{n+1} \sim \text{Normal}(\theta_{x_{n+1}}, \lambda_{x_{n+1}})$, where $\lambda \in \mathbb{R}_+^k$ is known.
3. Choose x^* based on $x_1, \ldots, x_N$ and $y_1, \ldots, y_N$.
4. Receive a reward θ_{x^*}.

Goal: Maximize the reward received by finding the best alternative.

Bayesian Ranking & Selection

- We assume that nature has chosen θ at random from some "population of potential truths" whose distribution is $\mathbb{P}_0$.
- We call $\mathbb{P}_0$ our prior.
- Let π be a policy, which is a method for choosing $x_1, \ldots, x_N$ and x^* based upon the appropriate inputs.
- The expected reward under policy π when the truth is θ is $\mathbb{E}^\pi[\theta_{x^*} \mid \theta]$.
- Goal: maximize the expected reward received:

$$\sup_\pi \int \mathbb{E}^\pi[\theta_{x^*} \mid \theta] \, \mathbb{P}_0(d\theta) = \sup_\pi \mathbb{E}^\pi[\theta_{x^*}]$$

Bayesian Ranking & Selection

- We call $\mathbb{P}_n = \mathbb{P}_0\{\cdot \mid x_{1:n}, y_{1:n}\}$ our posterior.
- Define $\mu_x^n = \mathbb{E}_n[\theta_x]$.
- The Bayes-optimal implementation decision is $x^* \in \arg\max_i \mu_i^N$.
- Assuming that x^* is chosen in this way, the problem may be written

$$\sup_{\pi} \mathbb{E}^{\pi}[\max_i \mu_i^N]$$

"Traditional" Bayesian Ranking & Selection

- A common choice for $\mathbb{P}_0$ is an independent normal prior.
- Under $\mathbb{P}_0$, $\theta_x \sim \text{Normal}(\mu_i^0, (\sigma_i^0)^2)$ and is independent of $\theta_{x'}$.
- The independence means that measuring alternative x teaches us nothing about alternative x'.
- In order to do reasonably well, we need $N \geq k$.

Outline

Correlated Bayesian Ranking & Selection

- Recall that, traditionally, θ_x is independent of $\theta_{x'}$ under $\mathbb{P}_0$.
- However, in many situations, we believe that **similar alternatives often have similar values**.
- In such situations, our belief on θ may be more appropriately modeled by a multivariate normal prior.

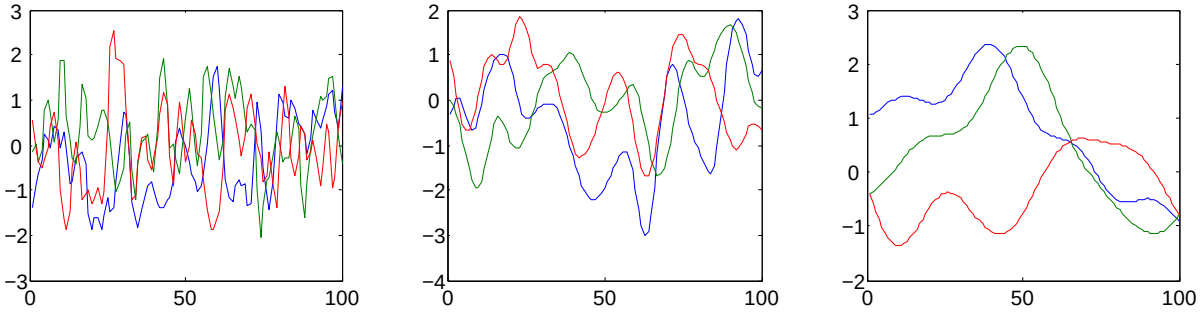
$$\theta \sim \text{Normal}(\mu^0, \Sigma^0).$$

Simulation Optimization

- Suppose the underlying true value of the real-world decision varies smoothly in that decision.
- Example: Revenue often varies smoothly with price.
- In such situations, two alternatives x and x' that are "close" should be more likly under the prior to have similar values θ_x and $\theta_{x'}$.
- To accomplish this, we may take the correlation under our prior between θ_x and $\theta_{x'}$ to be larger when x and x' are close.

Simulation Optimization

- Suppose the alternatives correspond to points in a continuous domain, e.g., prices, or inventory levels.
- Let $||x - x'||$ be the distance between alternatives x and x' in this domain.
- A common choice from spatial statistics is $Cov_0[\theta_x, \theta_{x'}] = \alpha \exp(-\beta(||x - x'||)^p)$ for constants α, β, and p.
- By varying p and β we may specify different beliefs about the smoothness of $x \mapsto \theta_x$.



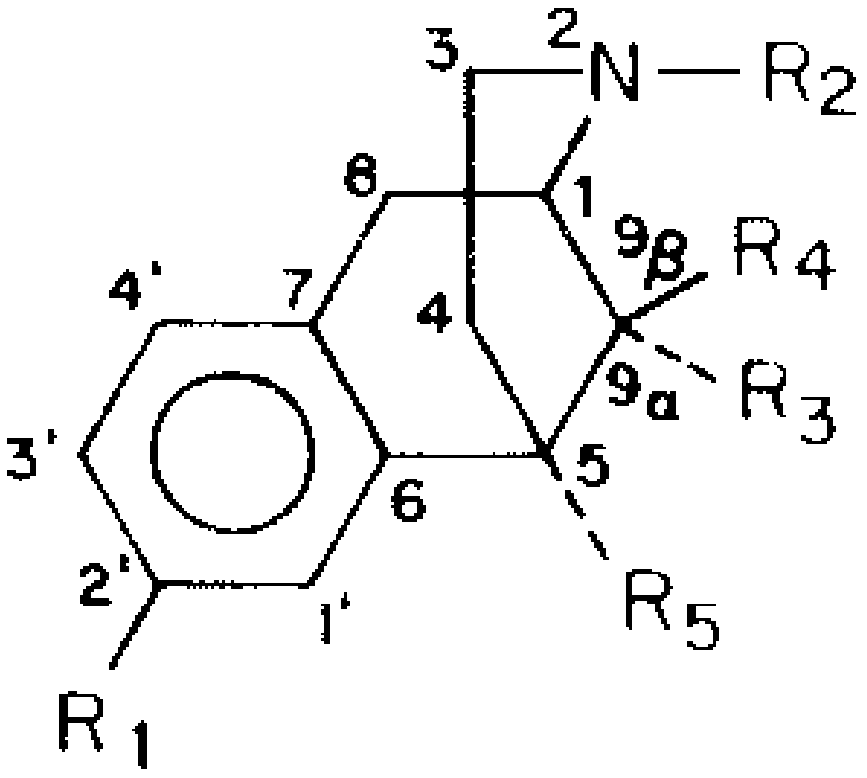
Drug Discovery

- Several models from medicinal chemistry specify the value of a compound as a linear combination of a set of chemical features $a_1, \ldots, a_J$, plus a compound specific term b_x.
- The value of compound x is given by

$$\theta_x = a_0 + \sum_{j=1}^{J} c_{xj} a_j + b_x.$$

  where c_{xj} is some known quantity that tells us how much compound x exhibits feature j.
- In the Free-Wilson model [Free, Wilson 1964], the features indicate the presence or absence of particular substituents at particular locations,
  e.g. $c_{xj} = 1$ if compound x has chlorine at location R1, and $c_{xj} = 0$ otherwise.
- In the Hansch analysis [Hansch, Fujita 1964], the features are chemical or physical properties, e.g., solubility.

Drug Discovery



Substituent
$R_1 = F$
$R_1 = OH$
$R_1 = OCOCH_3$
$R_1 = OCOCH_2CH_3$
$R_1 = OCOCH_2CH_2CH_3$
$R_1 = OCOCH_2(CH_2)_2CH_3$
$R_1 = OCH_3$
$R_1 = NO_2$
$R_1 = OCOAr(3\text{-}C_5H_4N)$
$R_1 = Cl$

Source: Katz, Osborne, Ionescu 1977

Left: Base benzomorphan molecule, with locations R1-R5 available for substitution.
Right: Set of possible substituents at location R1.

Drug Discovery

- If we take an independent normal prior on the a_j and b_x,

$$a_0 \sim \text{Normal}(\mu_0, \sigma_0^2),$$
$$a_j \sim \text{Normal}(0, \sigma_a^2),$$
$$b_x \sim \text{Normal}(0, \sigma_b^2),$$

  then what results is a multivariate normal prior on θ.
- Under this prior, two compounds that share similar features will be more likely to have similar values.
- We may take a hierarchical prior on the parameters of this model, and then use adaptively updated maximum a posteriori estimates in what follows.

Correlated Bayesian Ranking & Selection

- Let $\mathbb{P}_0$ be a multivariate normal prior.
- Under $\mathbb{P}_0$, $\theta \sim \mathcal{N}(\mu^0, \Sigma^0)$, where Σ^0 is allowed to be non-diagonal.
- Define $\mu^n = \mathbb{E}_n[\theta]$, $\Sigma^n = Var_n[\theta]$.
- We have a new prior, but the same goal:

$$\sup_{\pi} \mathbb{E}^{\pi} \left[\max_i \mu_i^N \right].$$

- We may do well under this prior even if $k >> N$.
- i.e., we can do well even if the number of alternatives is much larger than the number of measurements we can perform.

Correlated Bayesian Ranking & Selection

- With a multivariate normal prior, the posterior is also multivariate normal.
- $\theta \sim \mathcal{N}(\mu^n, \Sigma^n)$ under the posterior $\mathbb{P}_n$, where μ^n and Σ^n may be computed recursively as

$$\mu^{n+1} = \mu^n + \frac{y_{n+1} - \mu_x^n}{\lambda_x + \Sigma_{xx}^n} \Sigma^n e_x,$$
$$\Sigma^{n+1} = \Sigma^n - \frac{\Sigma^n e_x e_x' \Sigma^n}{\lambda_x + \Sigma_{xx}^n}.$$

  where $x = x_{n+1}$, and e_x is the vector of all 0s, with a 1 at index x.
- For drug discovery, μ^n and Σ^n are defined implicitly in terms of the multivariate normal posterior on $a_0, a_1, \ldots, a_J$, and $(b_{x_m})_{m \le n}$.
- We need not store the belief on those b_x we have not measured, since our posterior on these is independent $\mathcal{N}(0, \sigma_b^2)$.

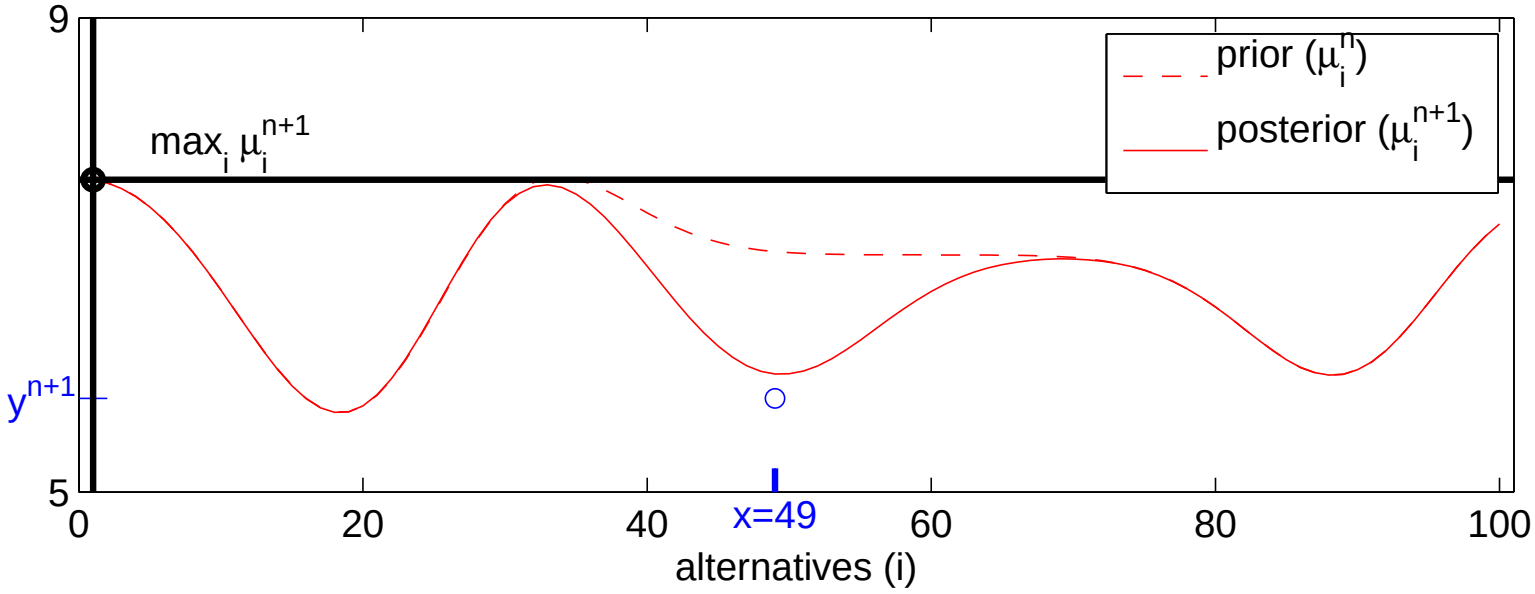
Knowledge-Gradient Policy

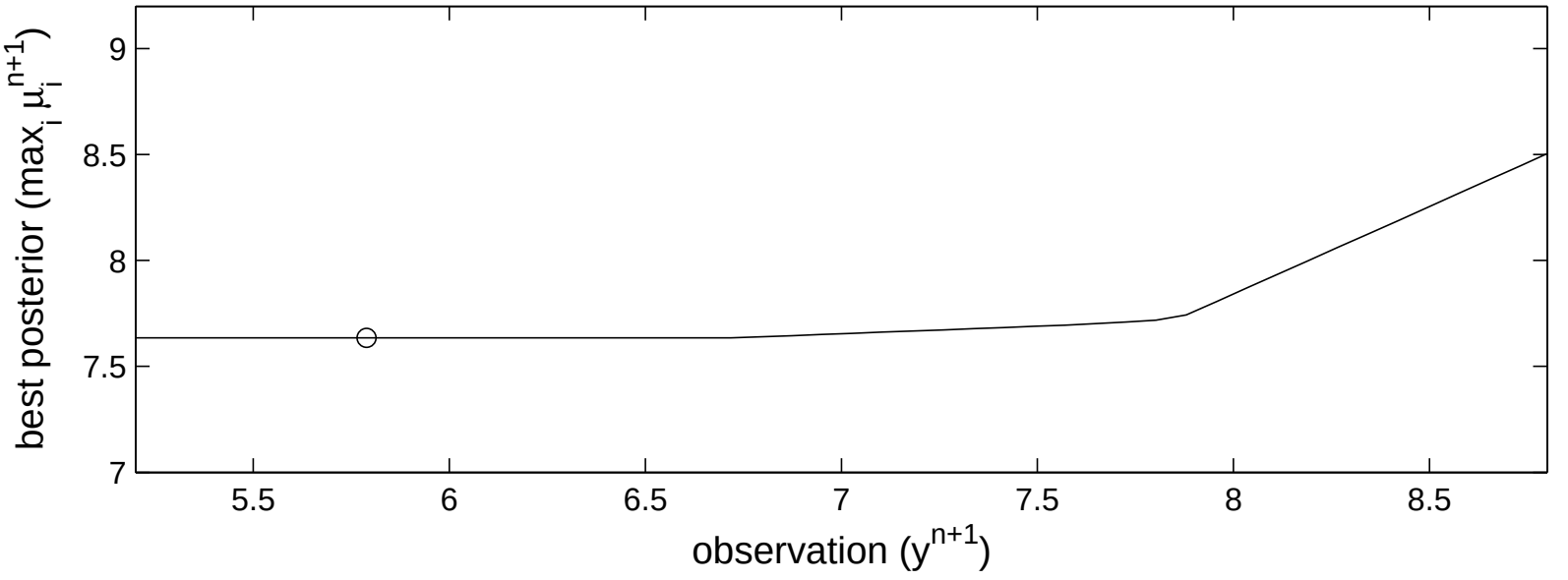
The **knowledge-gradient policy** is defined to be the policy that chooses x_{n+1} to maximize the KG factor,

$$\mathrm{KG}(x) = \mathbb{E}_n \left[\max_i \mu_i^{n+1} \mid x_{n+1} = x \right] - \max_i \mu_i^n.$$

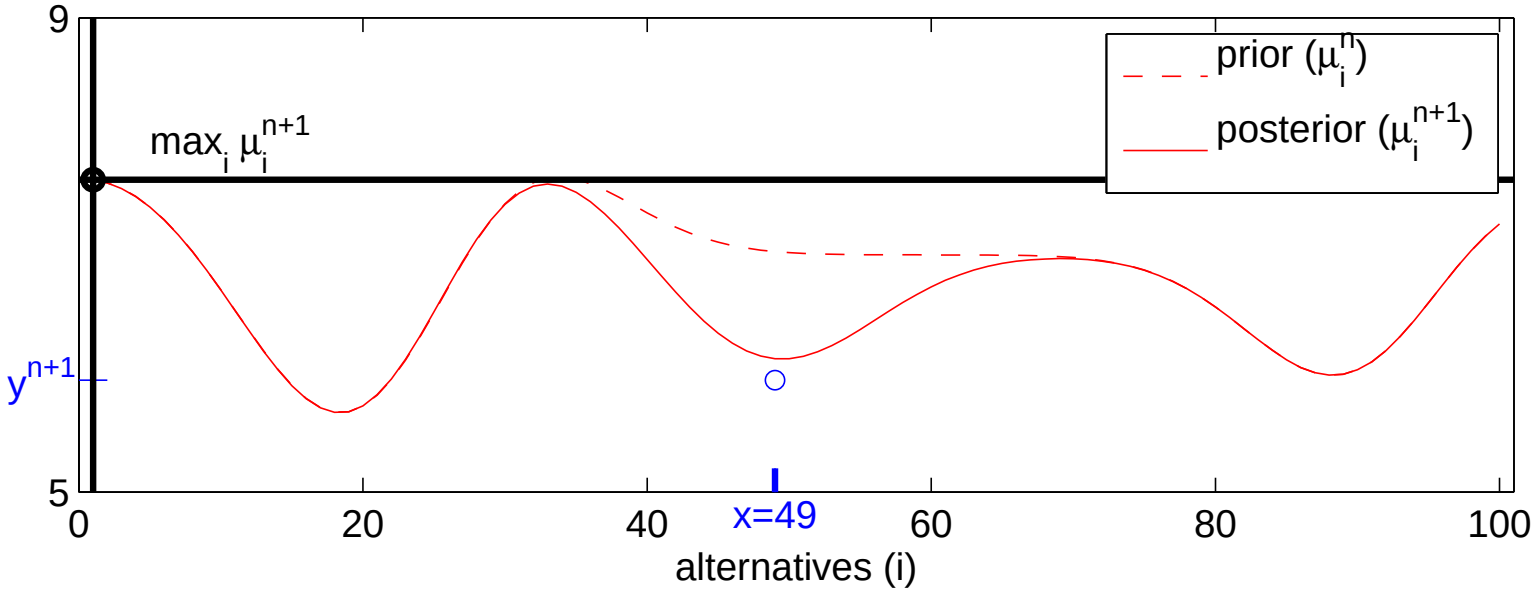
- μ_i^n is the expected value of alternative i given what we know at time n.
- $\max_i \mu_i^n$ is the best we can do given what we know at n.
- $\max_i \mu_i^{n+1}$ is the best we will be able to do given what we know at n and what we learn from our measurement x_{n+1}.
- The KG factor is similar to the expected improvement [Jones et al. 1998], and is the expected value of sampling information [Howard 1998]. When Σ^0 is diagonal, the KG policy is the (R1,. . . ,R1) policy [Gupta & Miescke 1996].

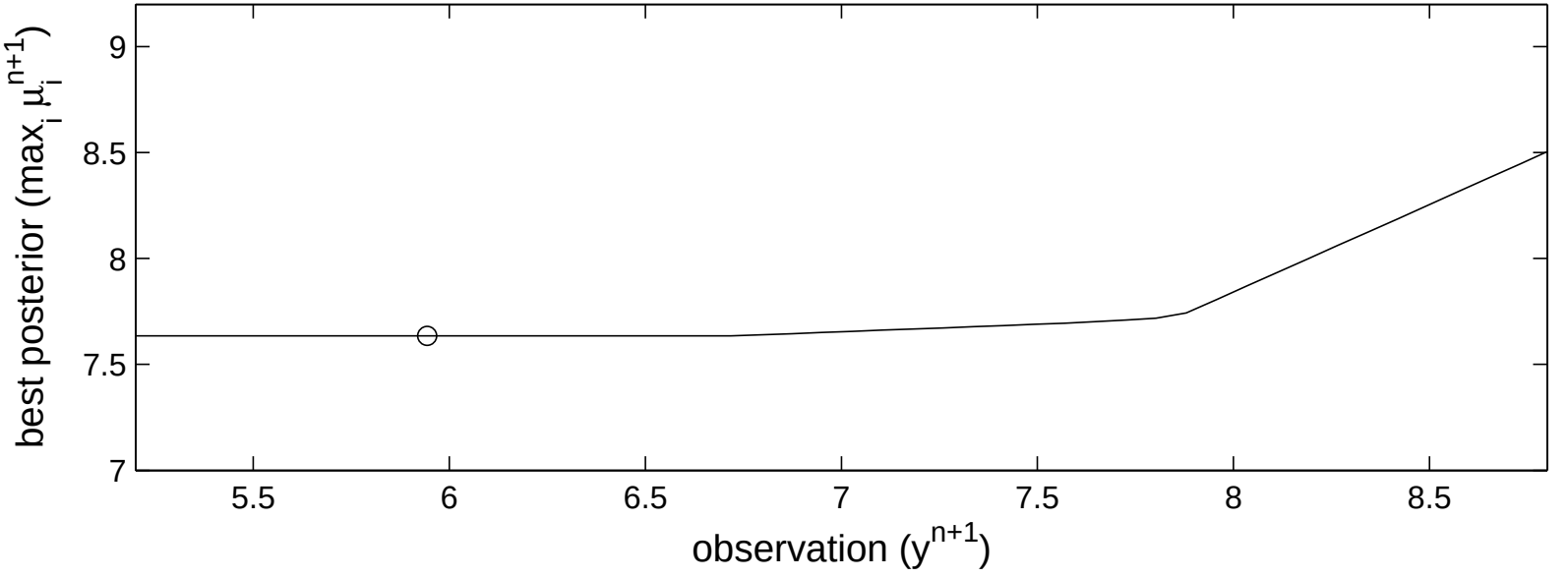
Computing the Knowledge-Gradient



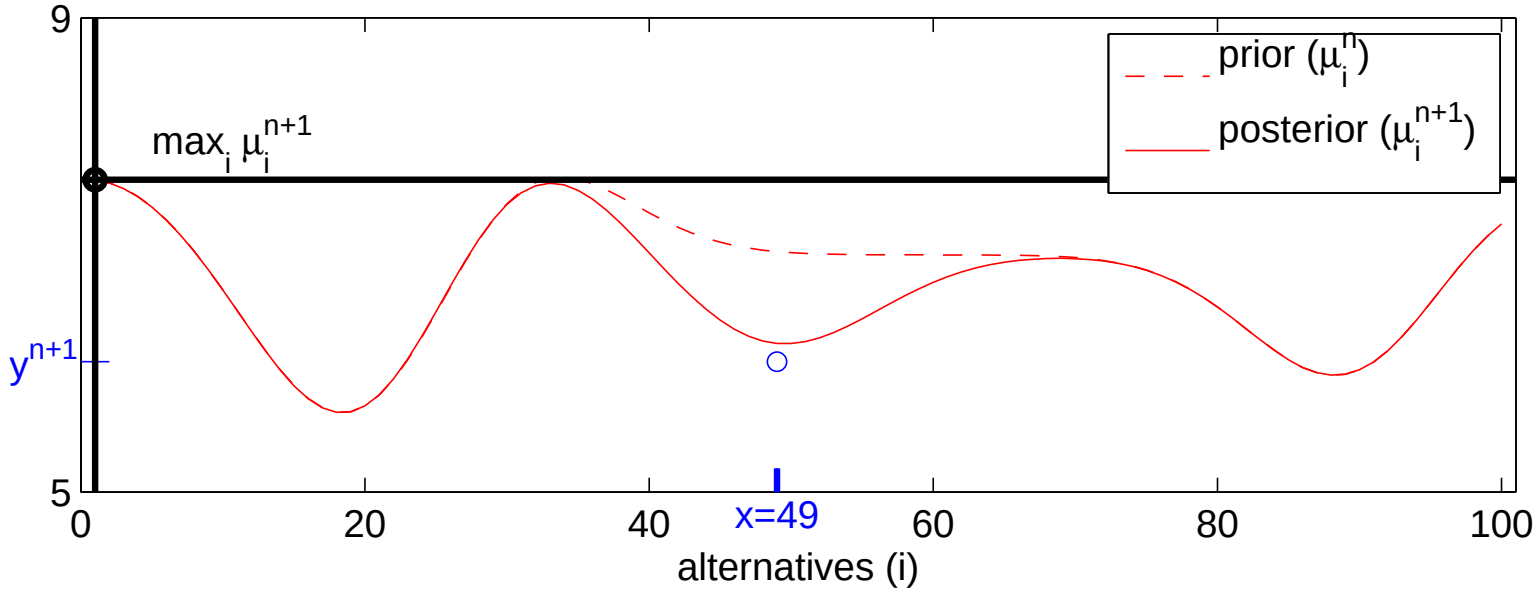


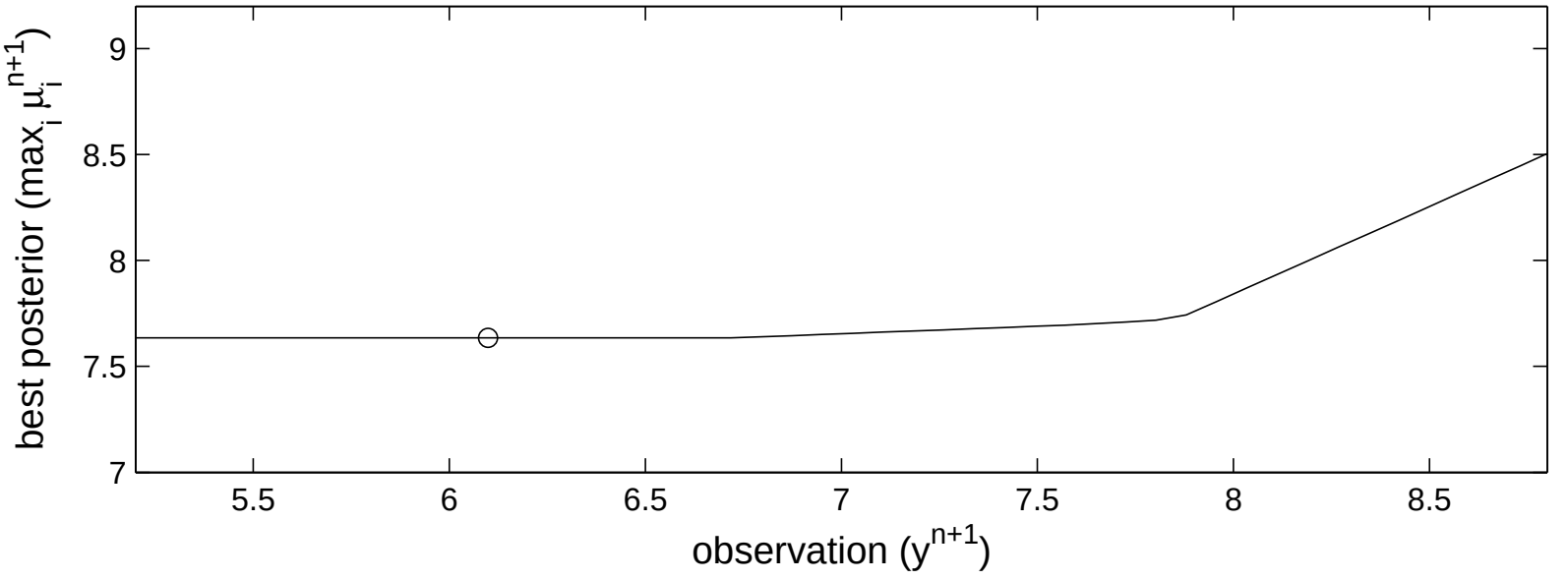
Computing the Knowledge-Gradient



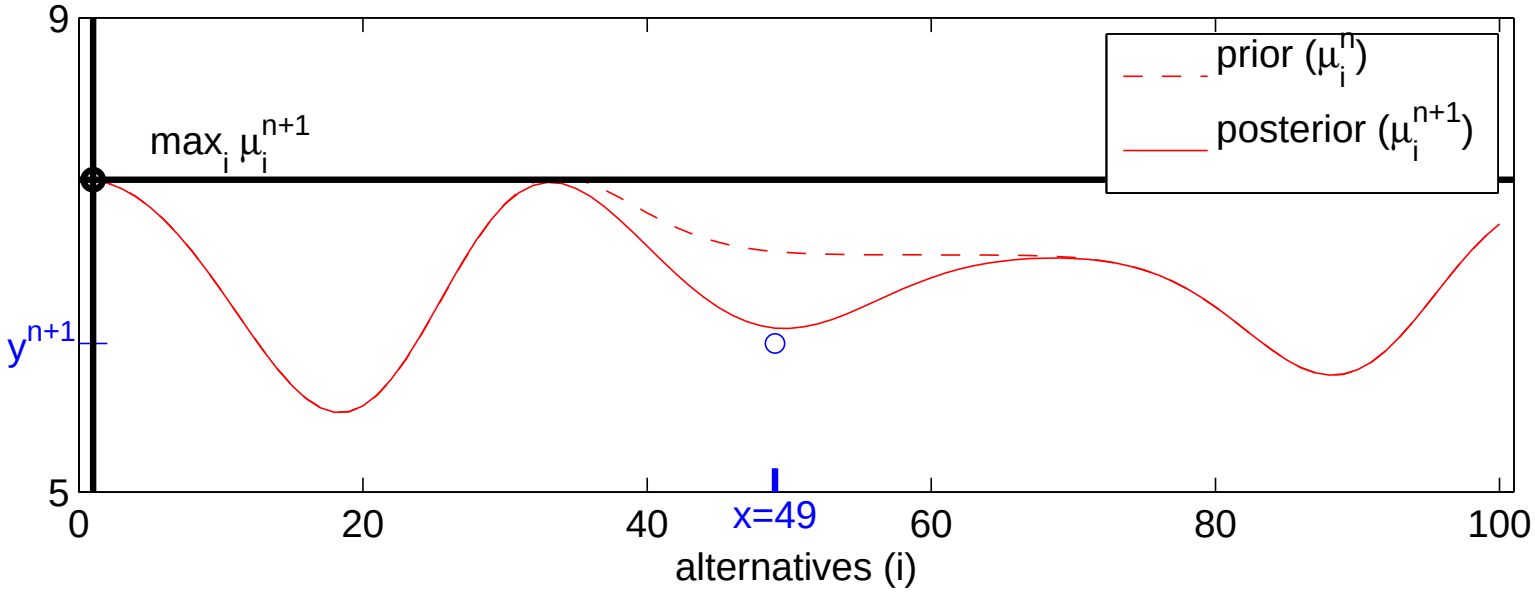


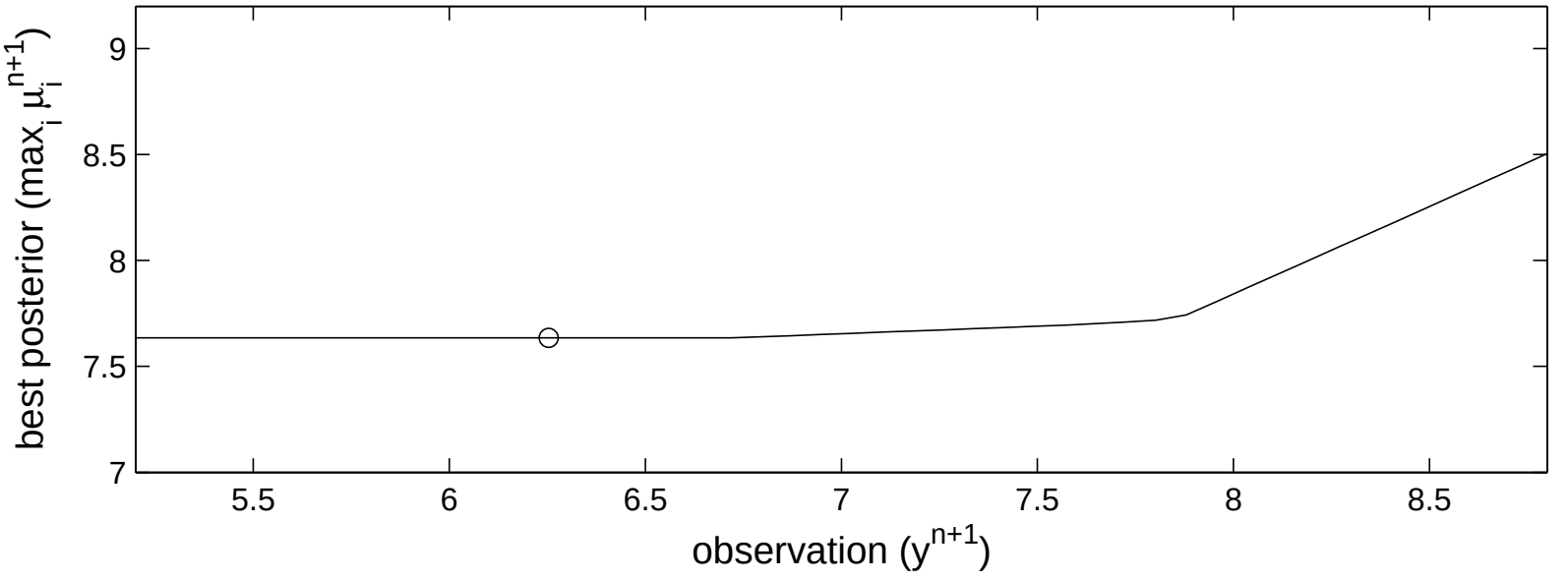
Computing the Knowledge-Gradient



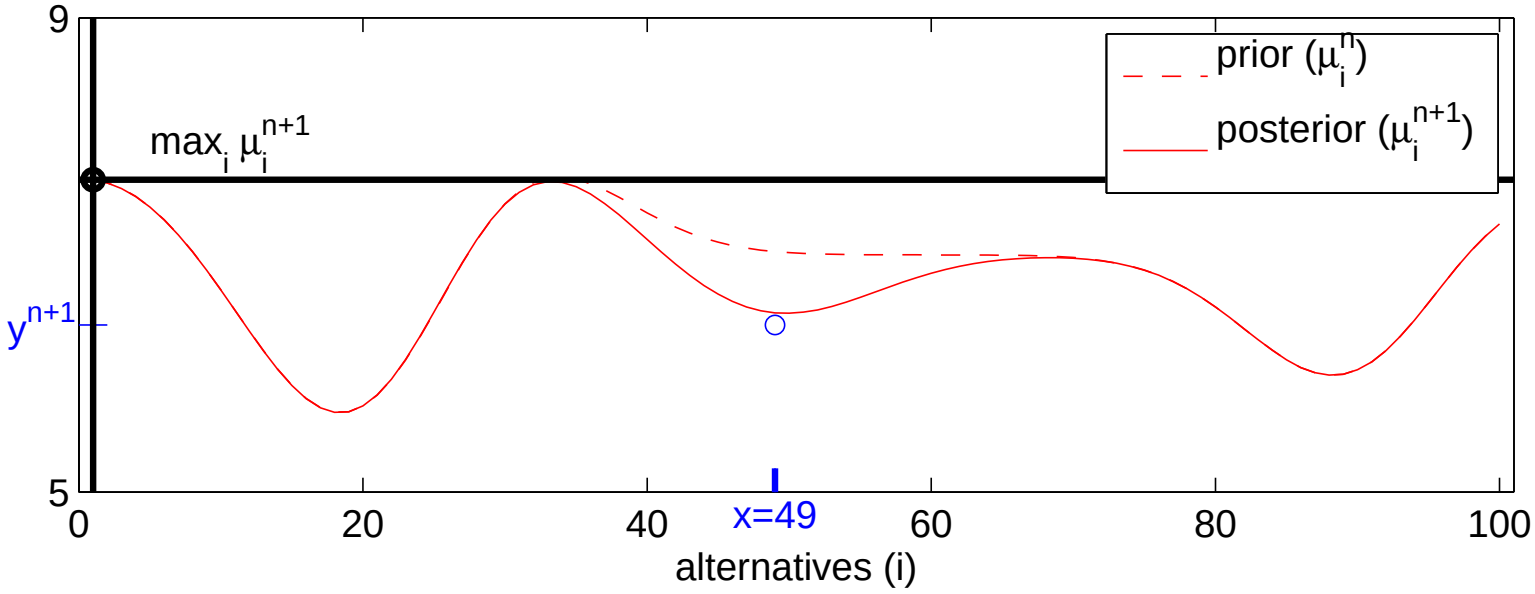


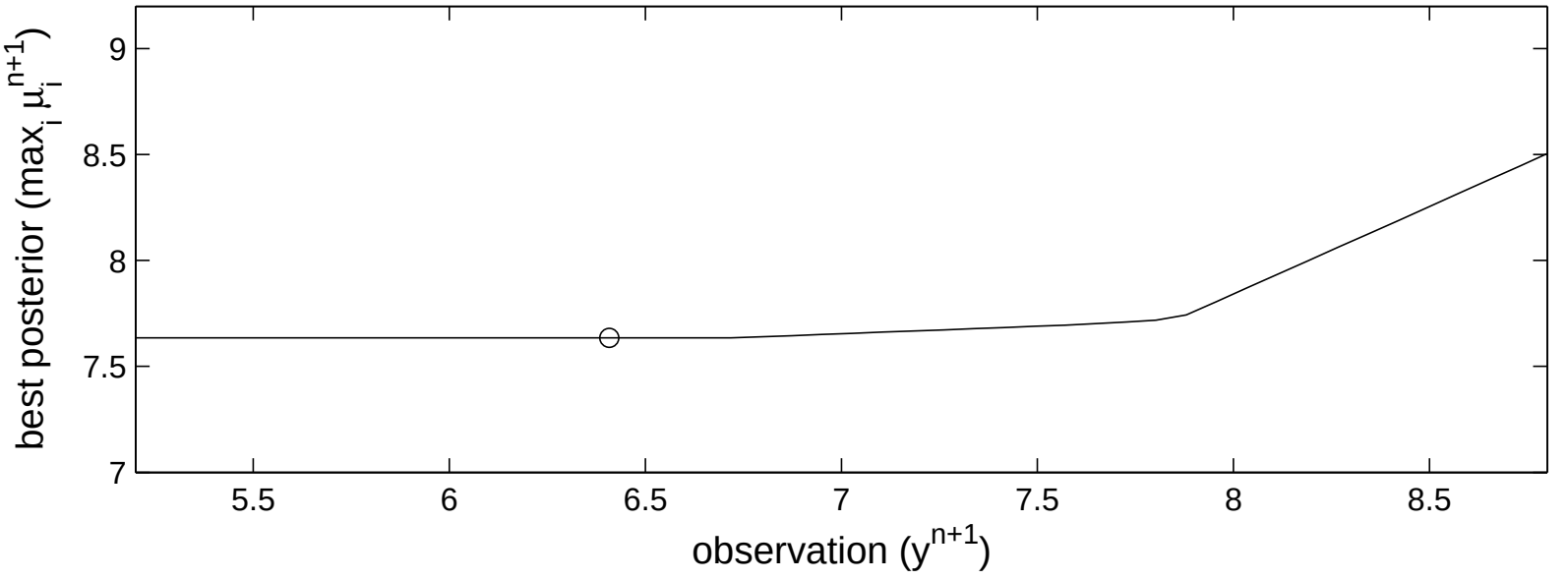
Computing the Knowledge-Gradient



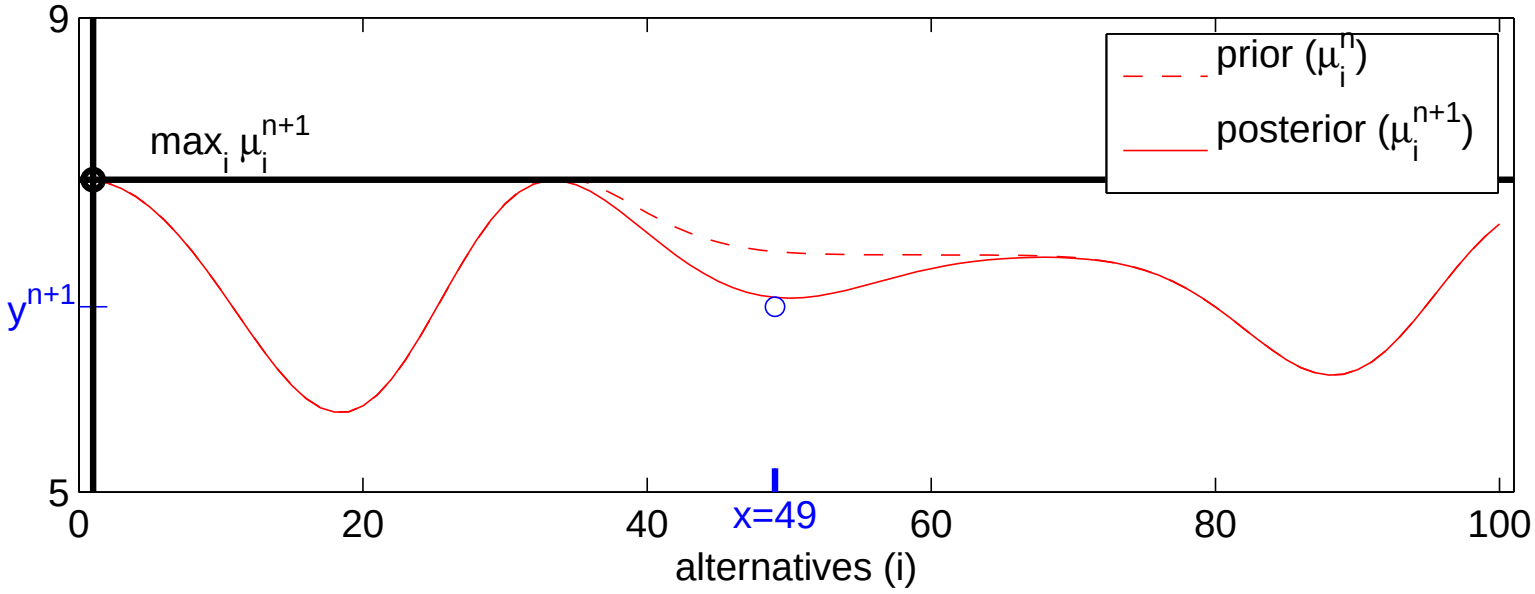


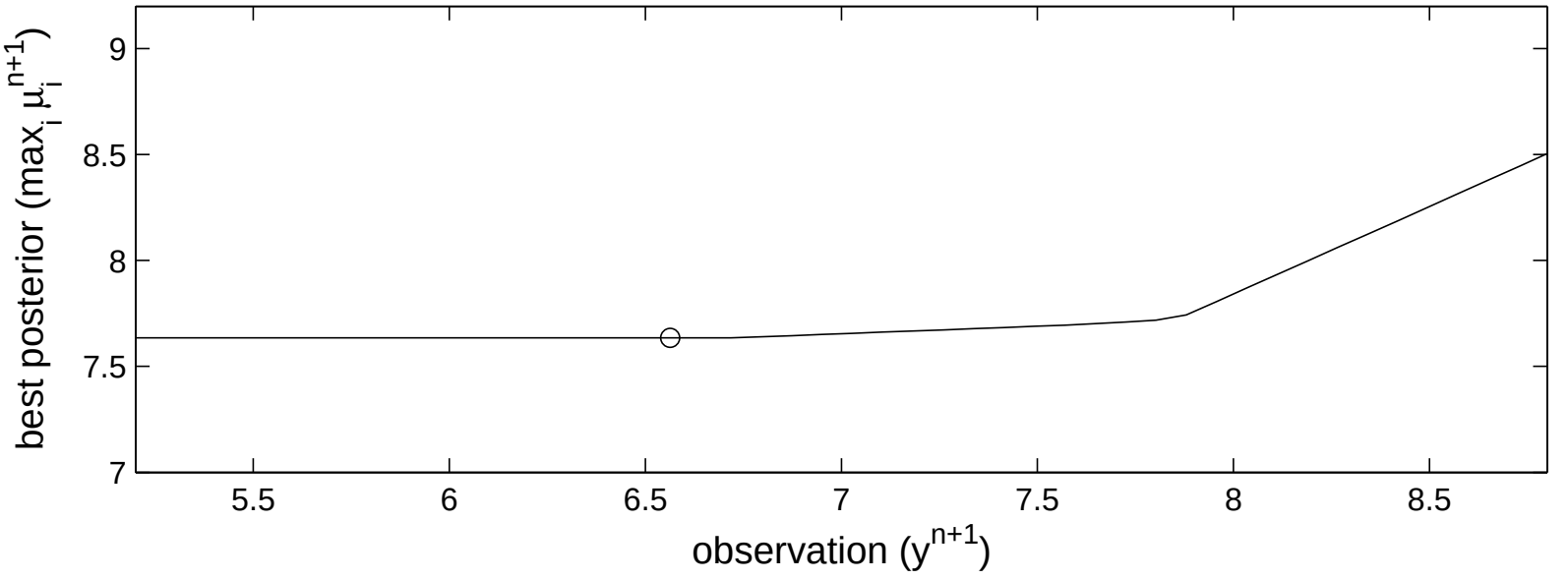
Computing the Knowledge-Gradient



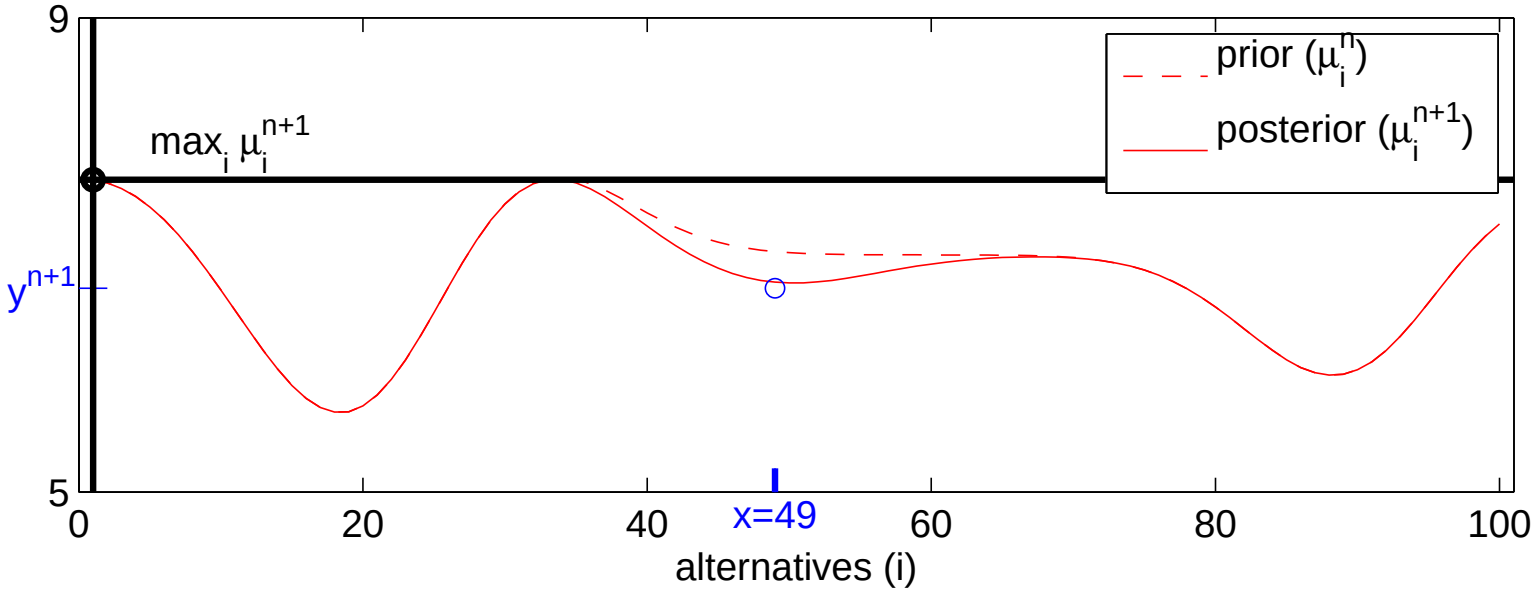


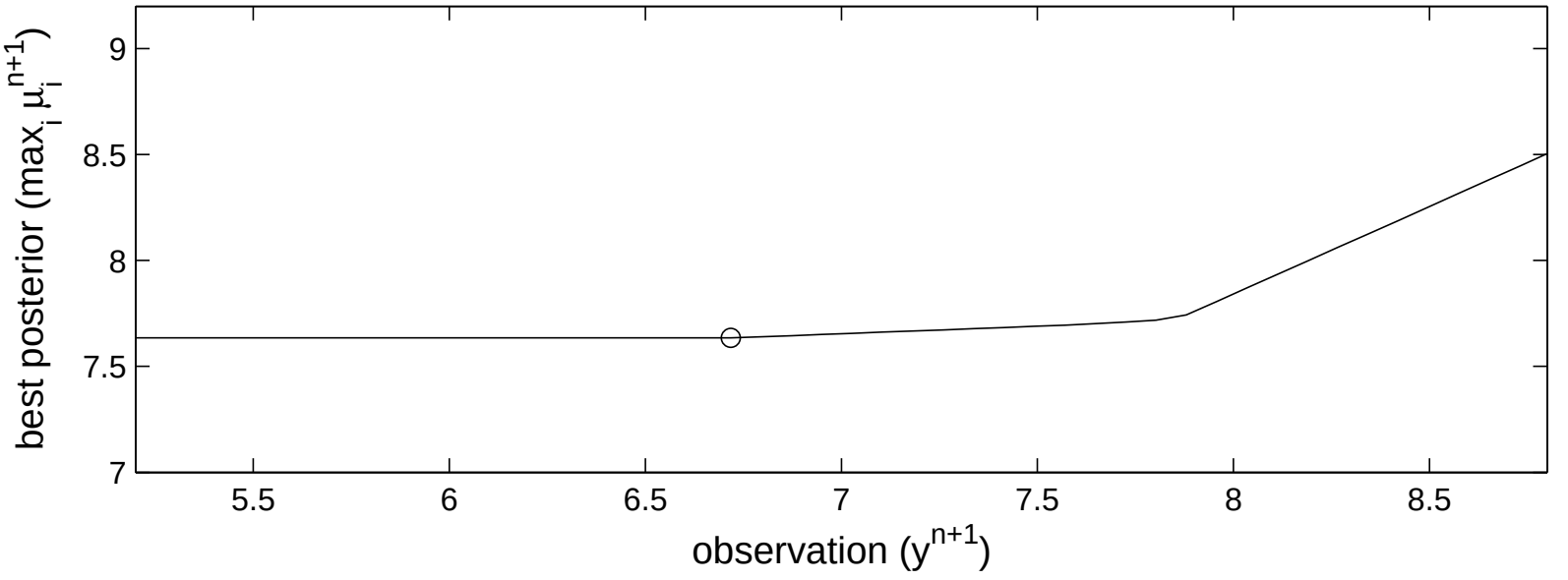
Computing the Knowledge-Gradient



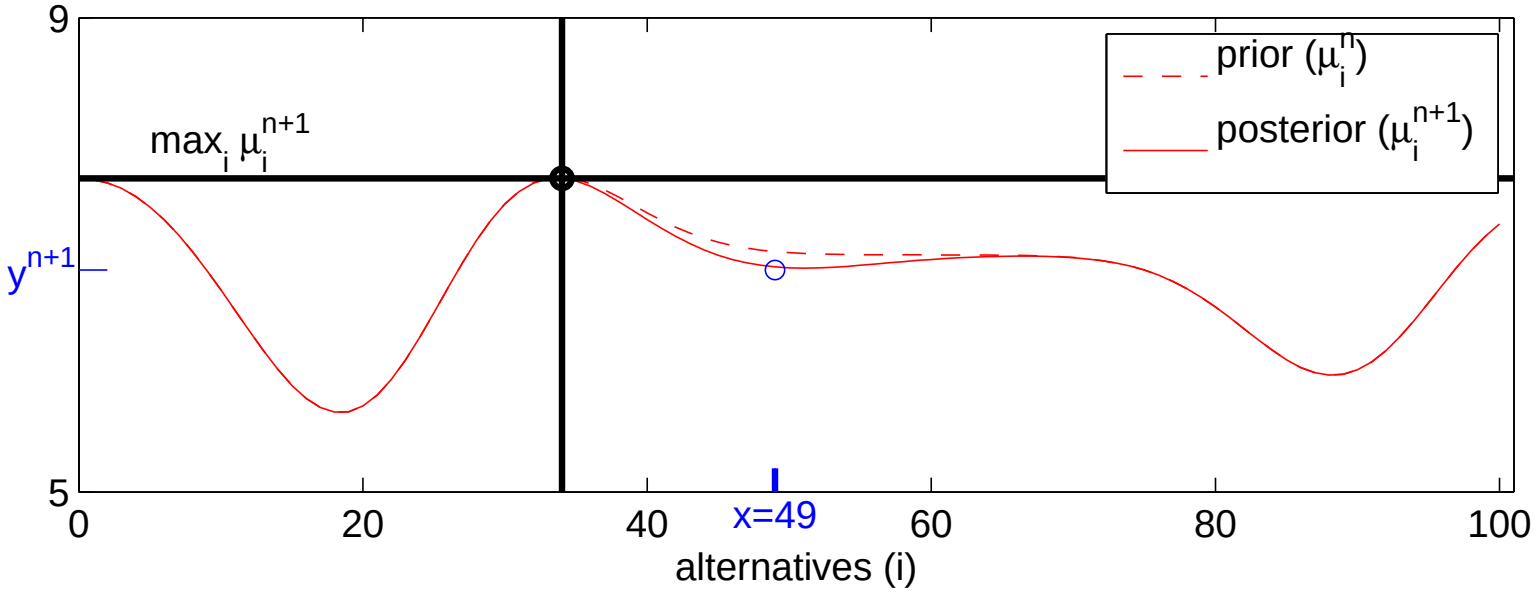


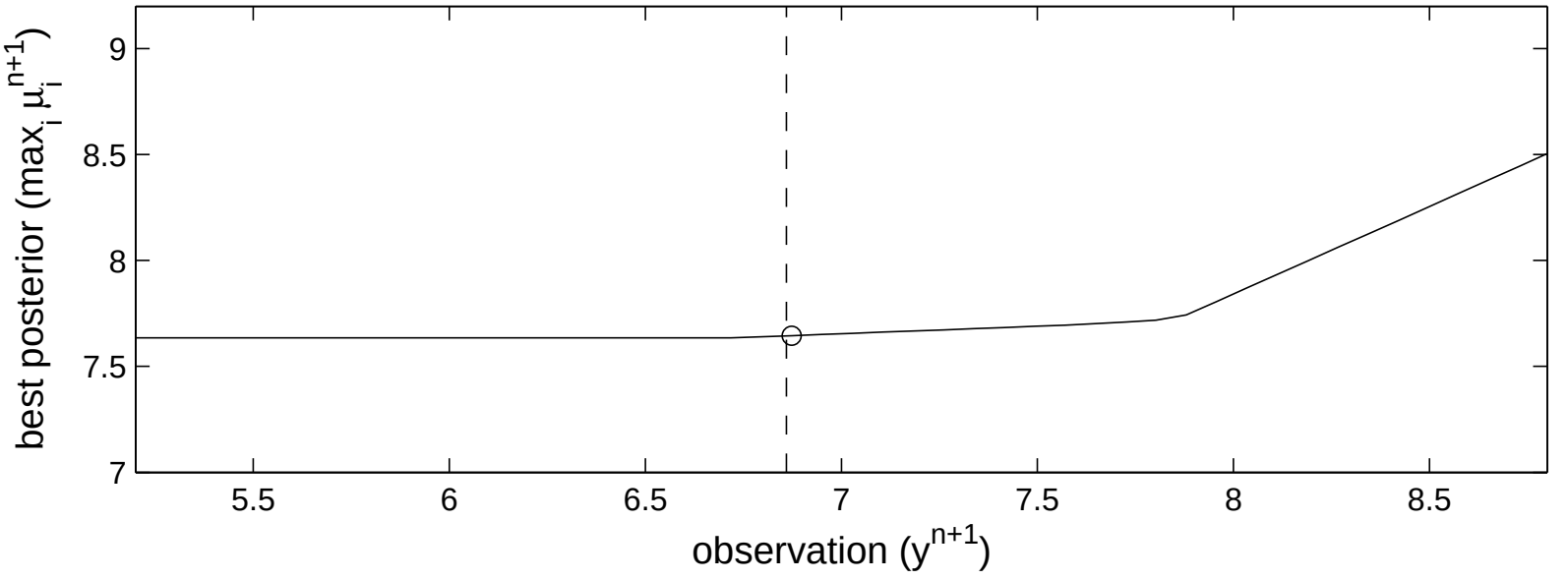
Computing the Knowledge-Gradient



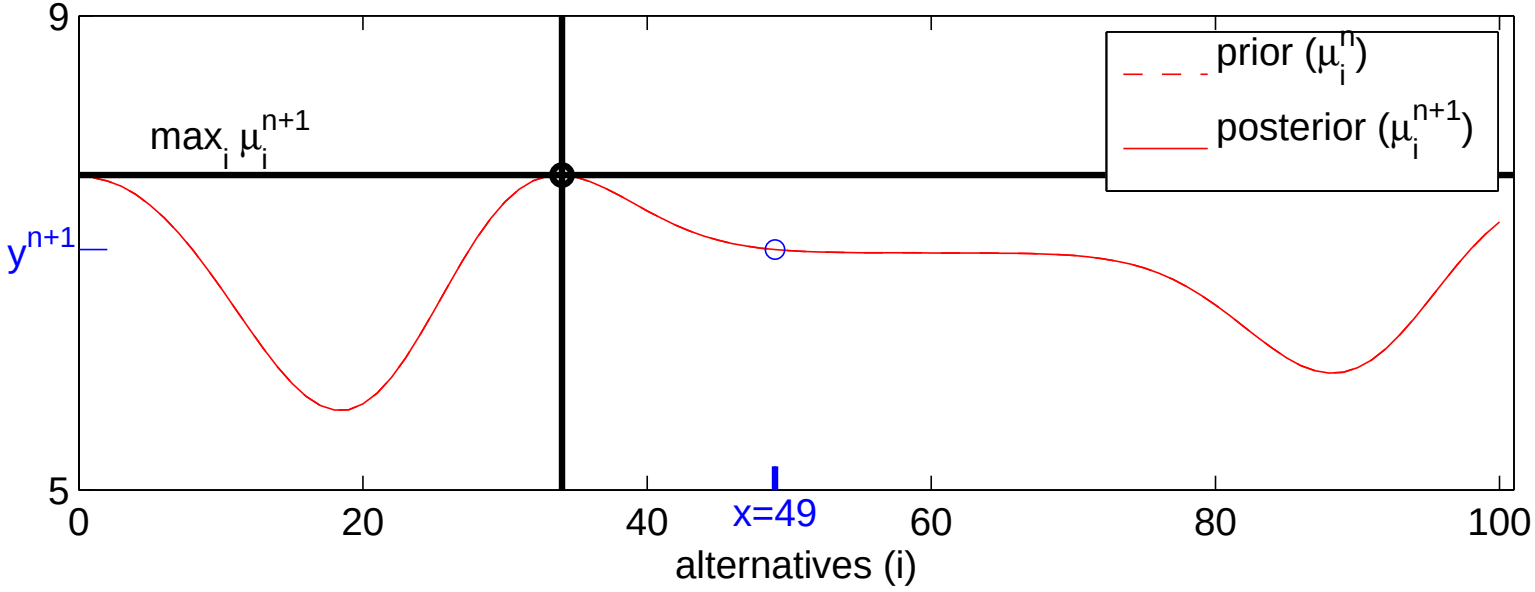


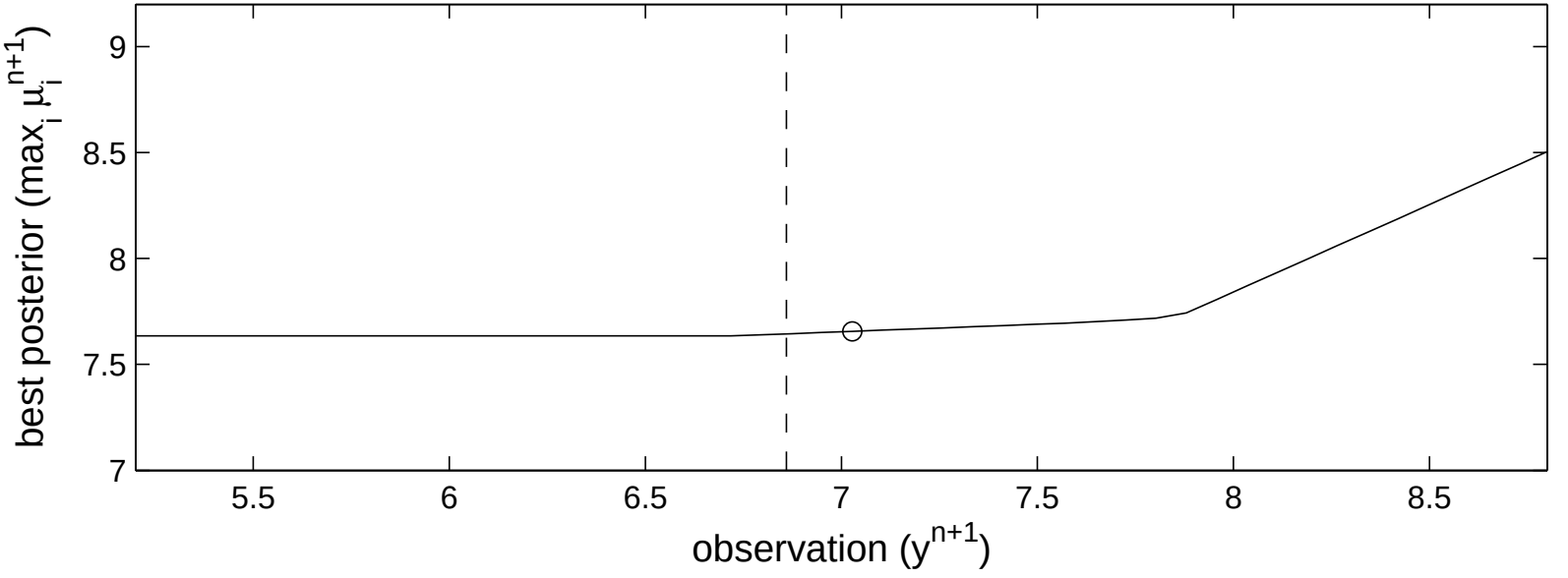
Computing the Knowledge-Gradient



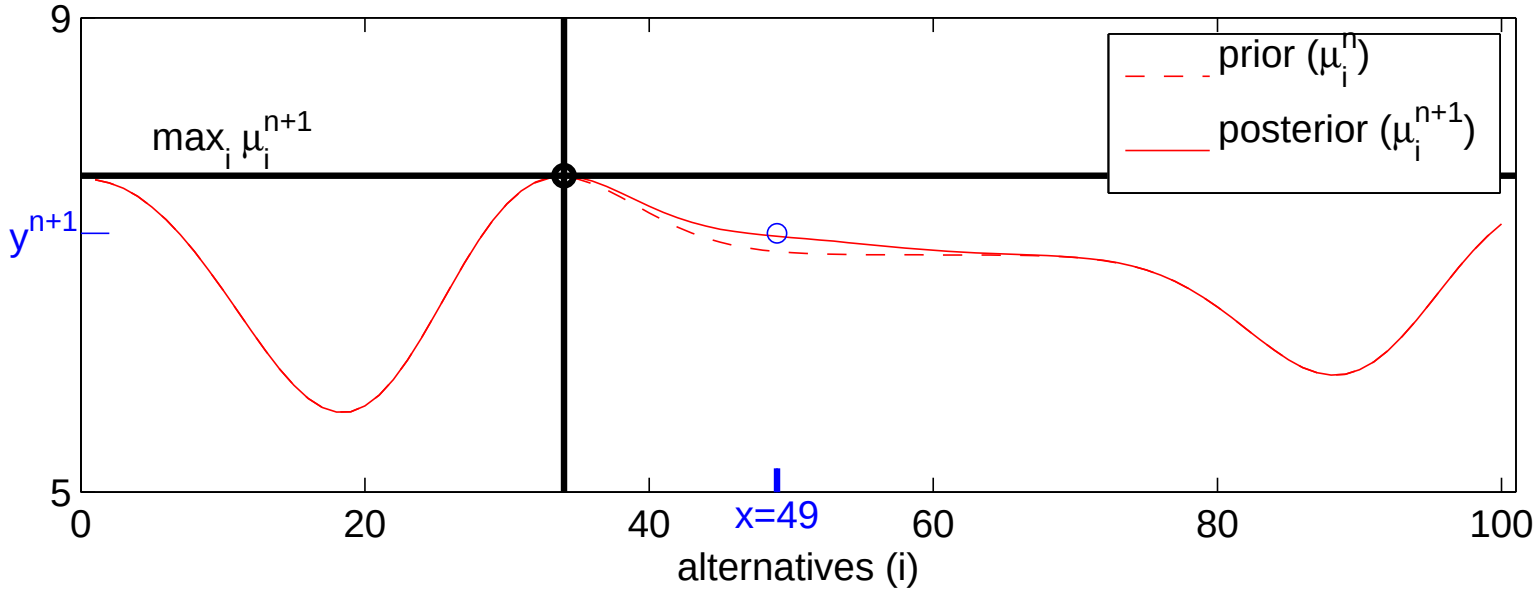


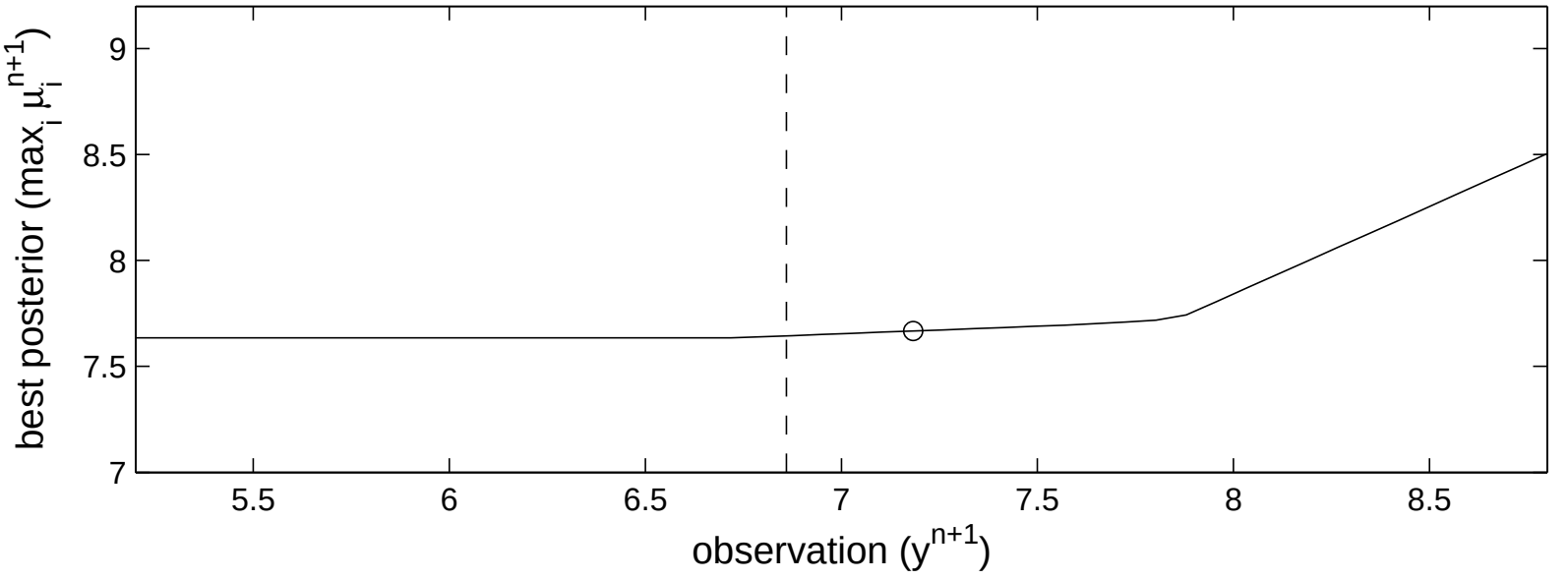
Computing the Knowledge-Gradient



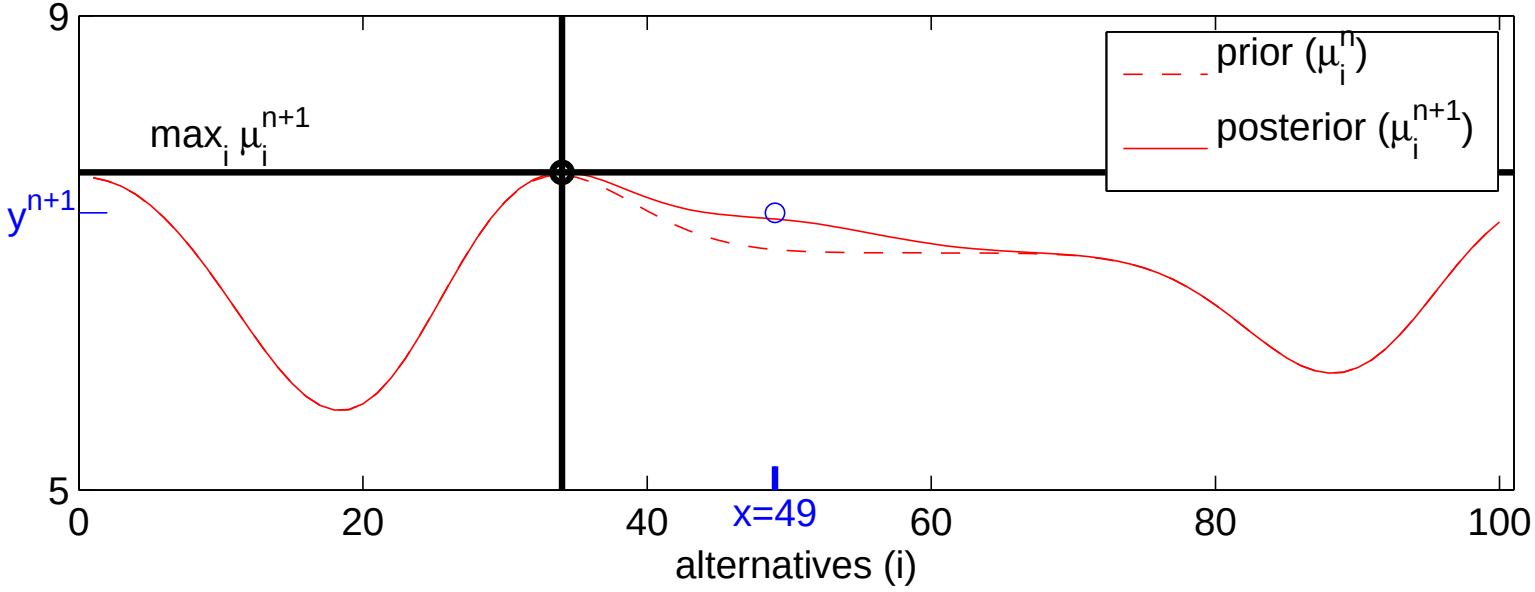


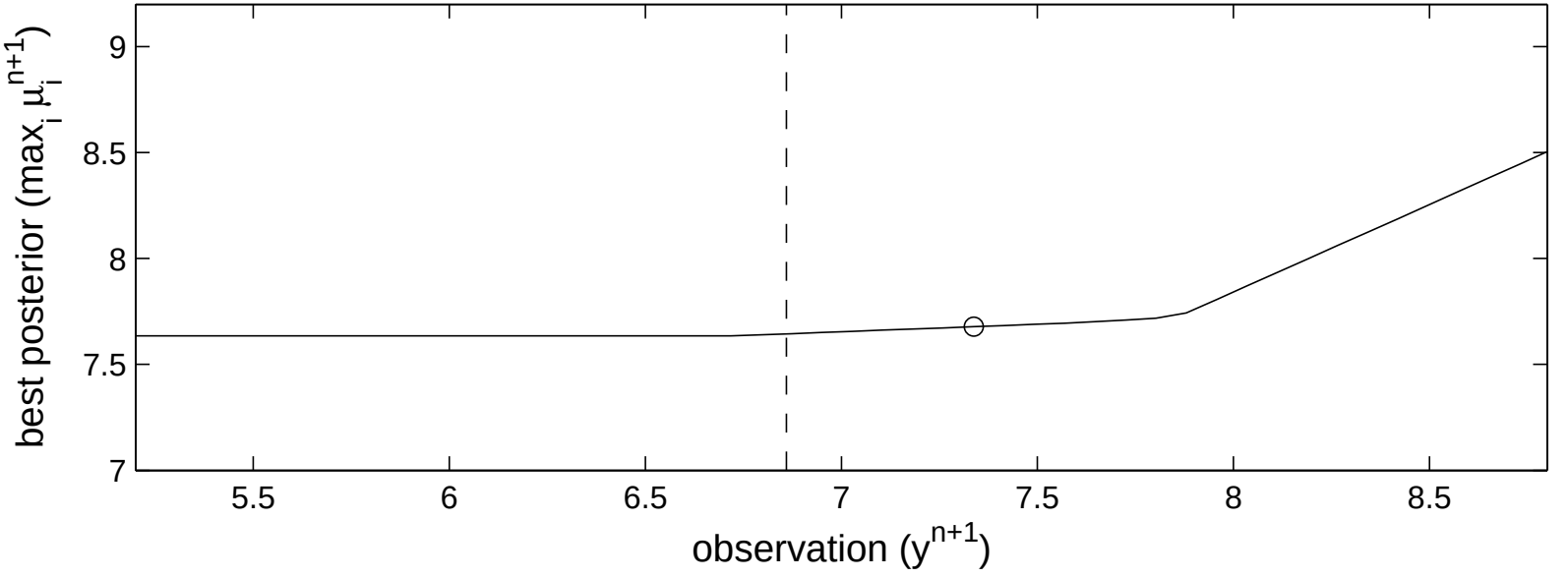
Computing the Knowledge-Gradient



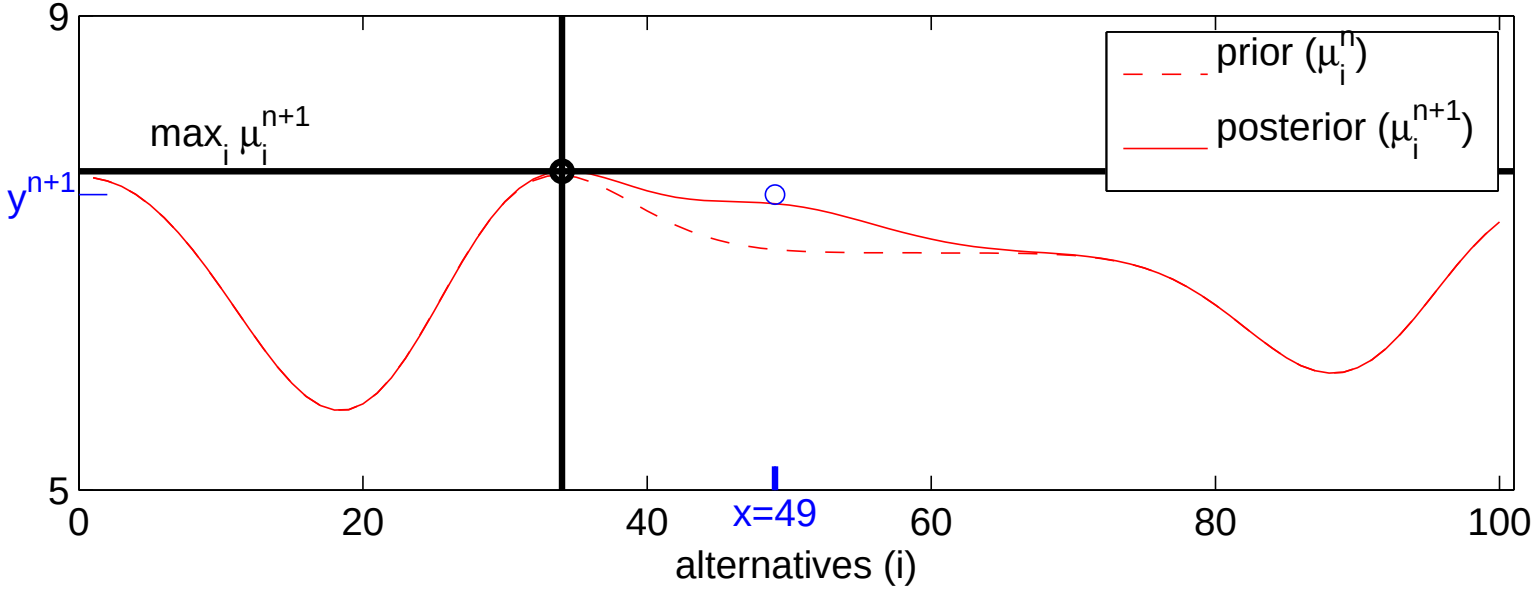


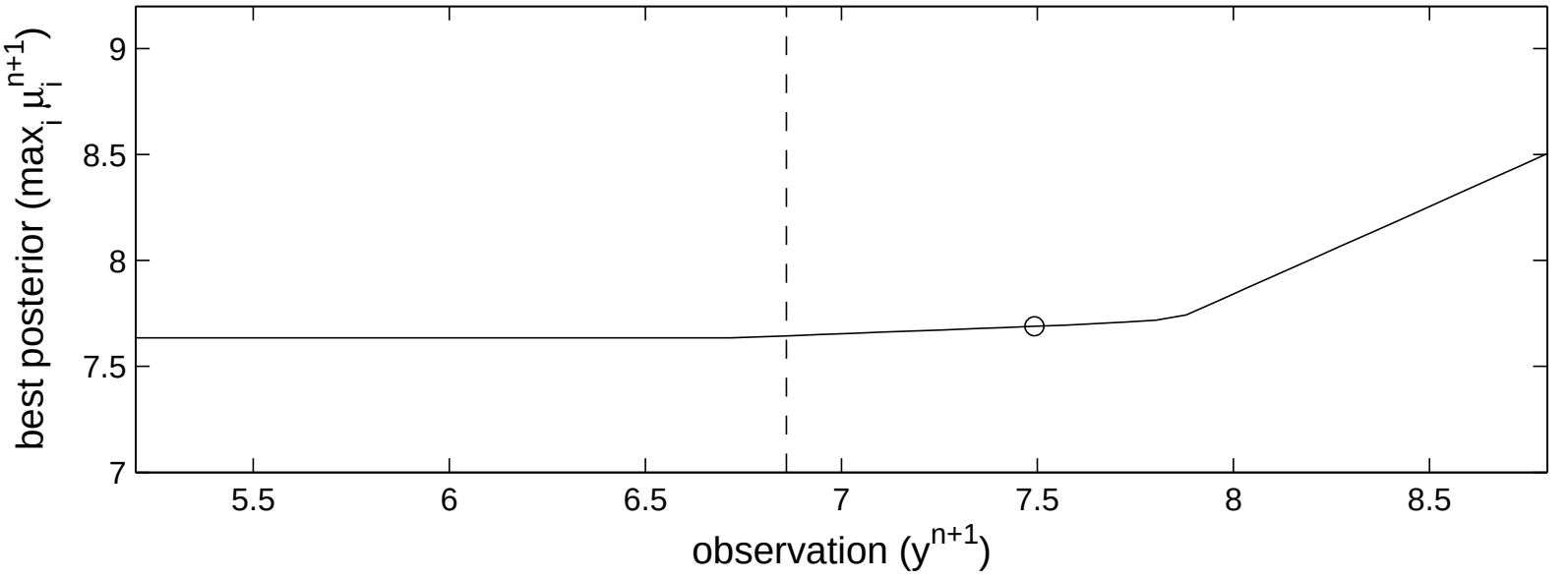
Computing the Knowledge-Gradient



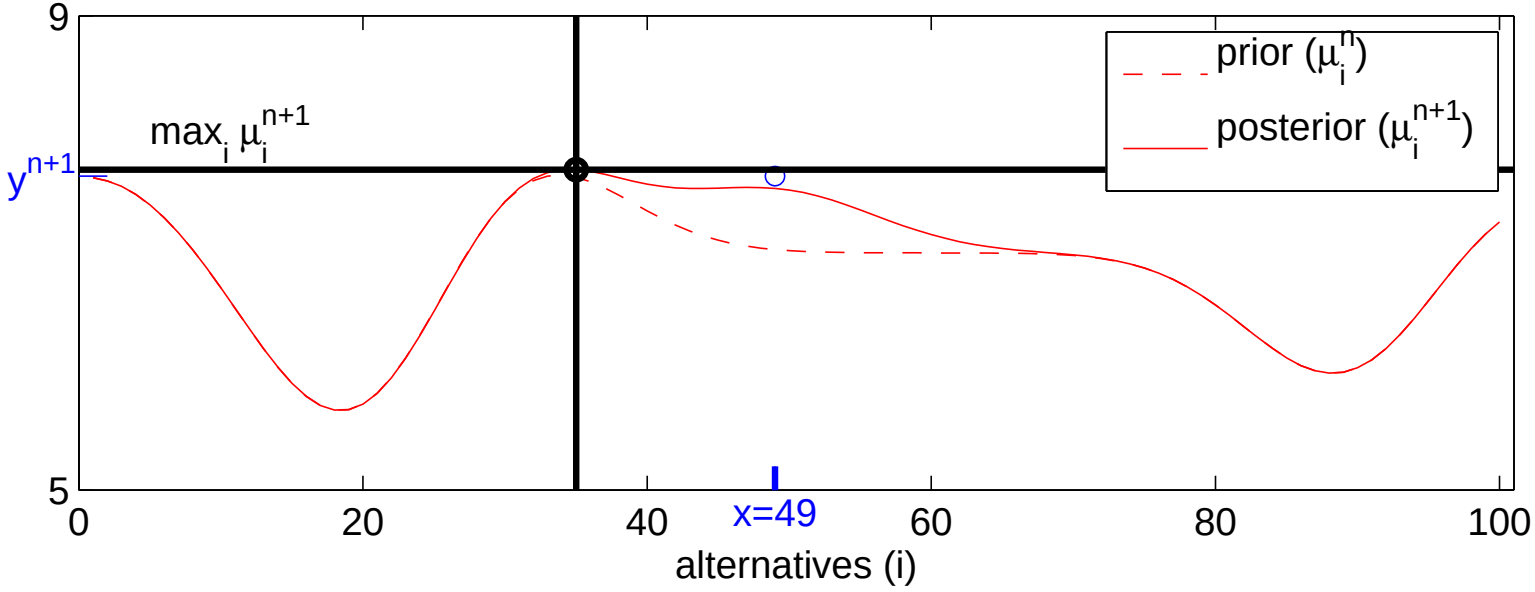


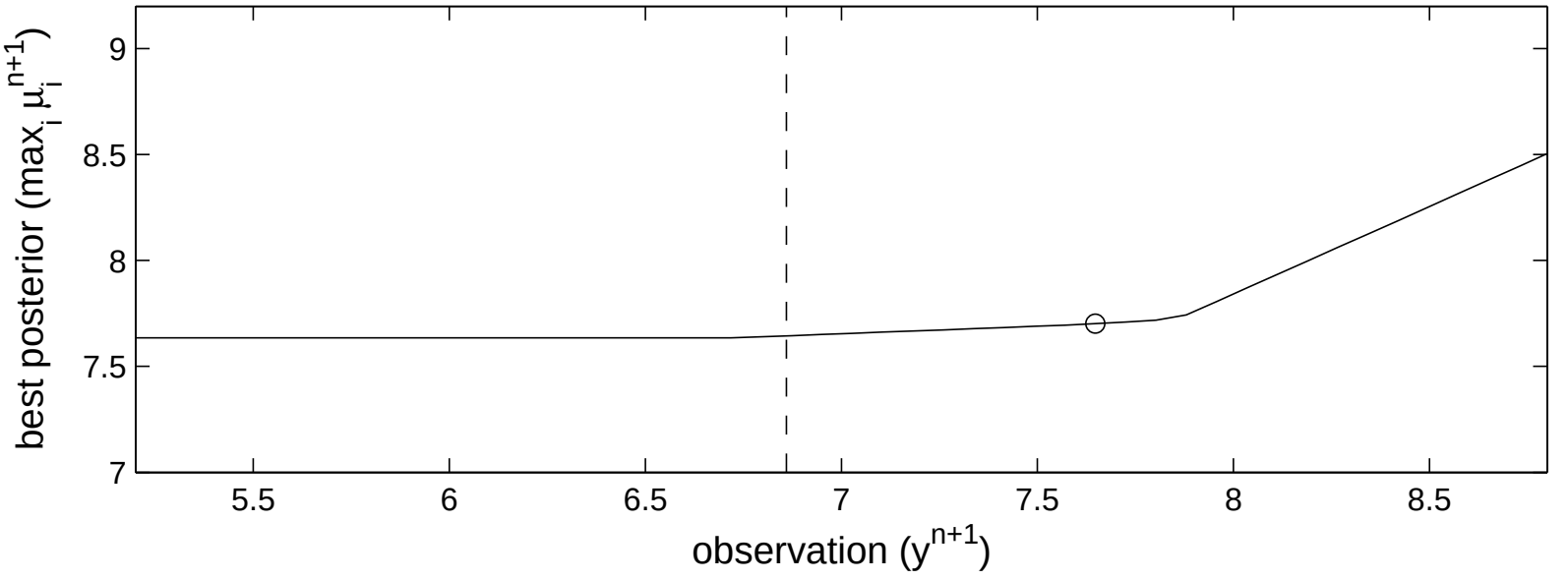
Computing the Knowledge-Gradient



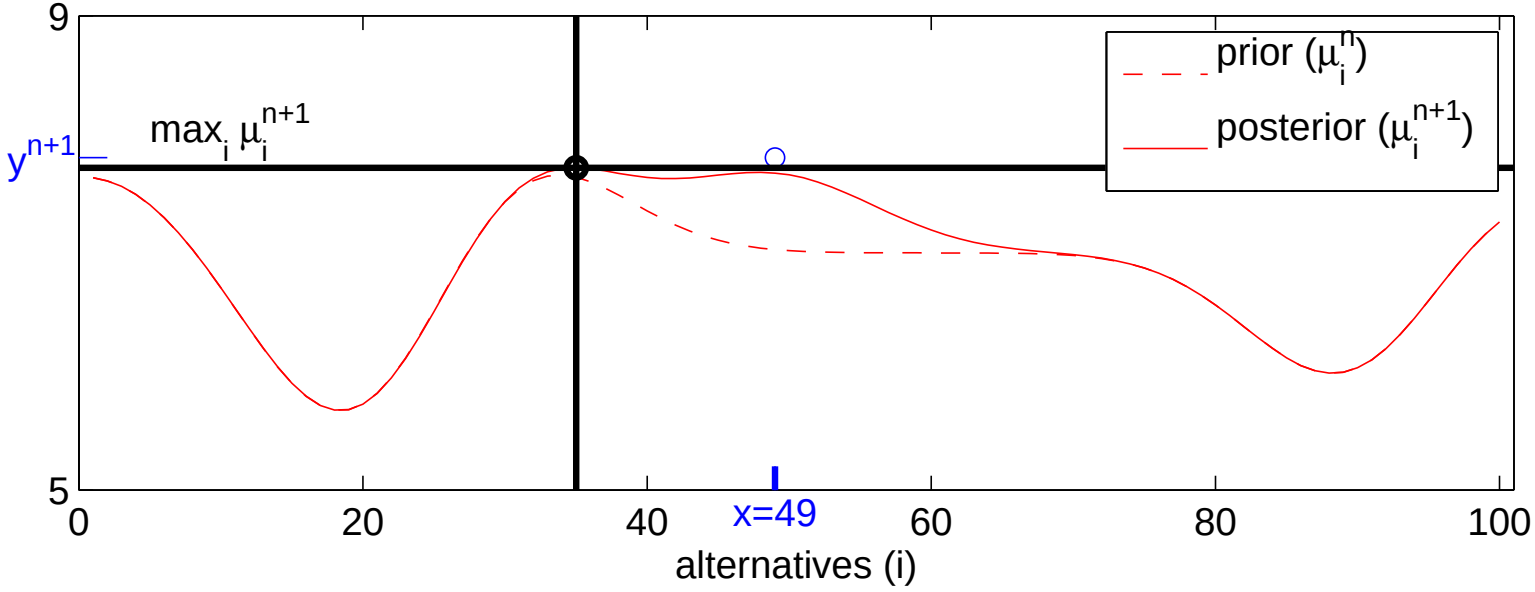


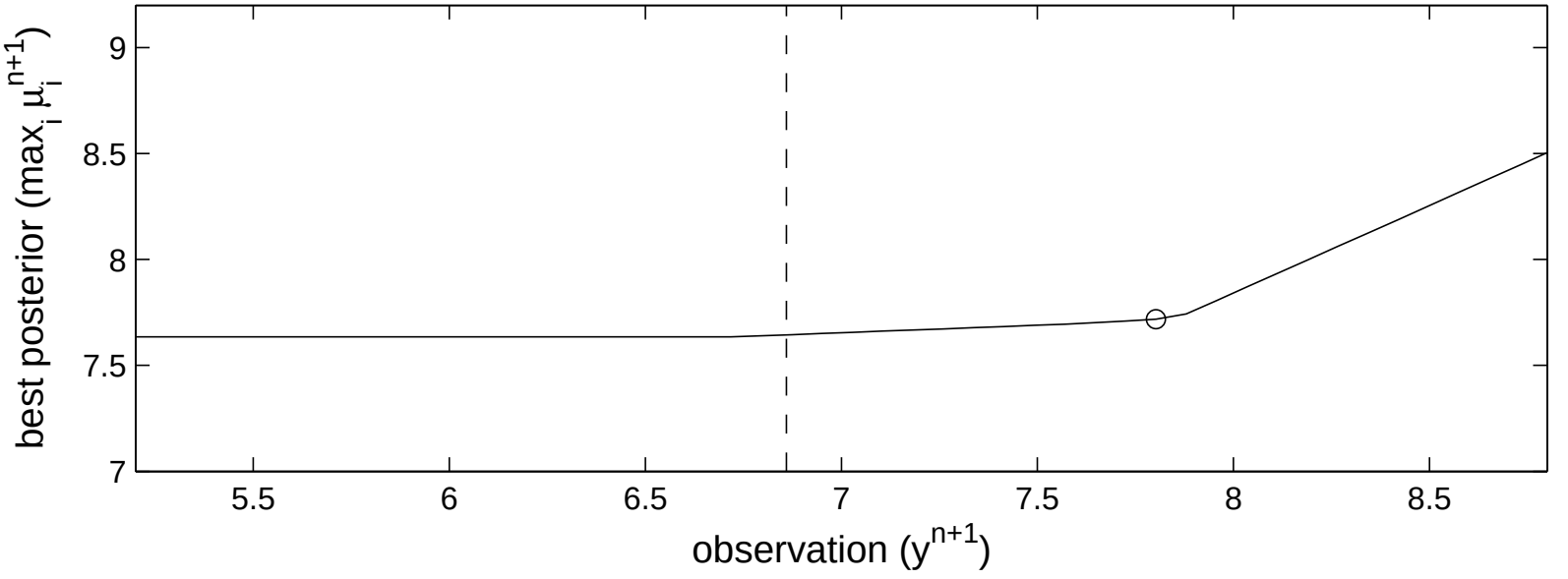
Computing the Knowledge-Gradient



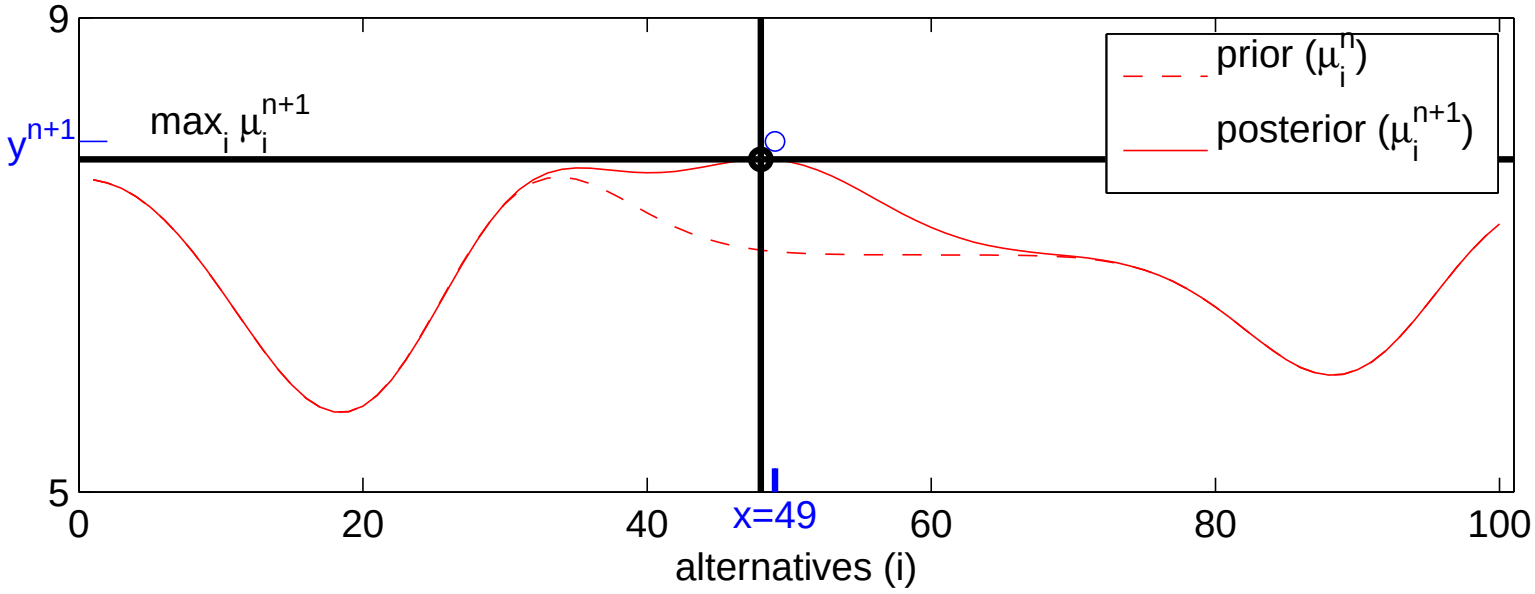


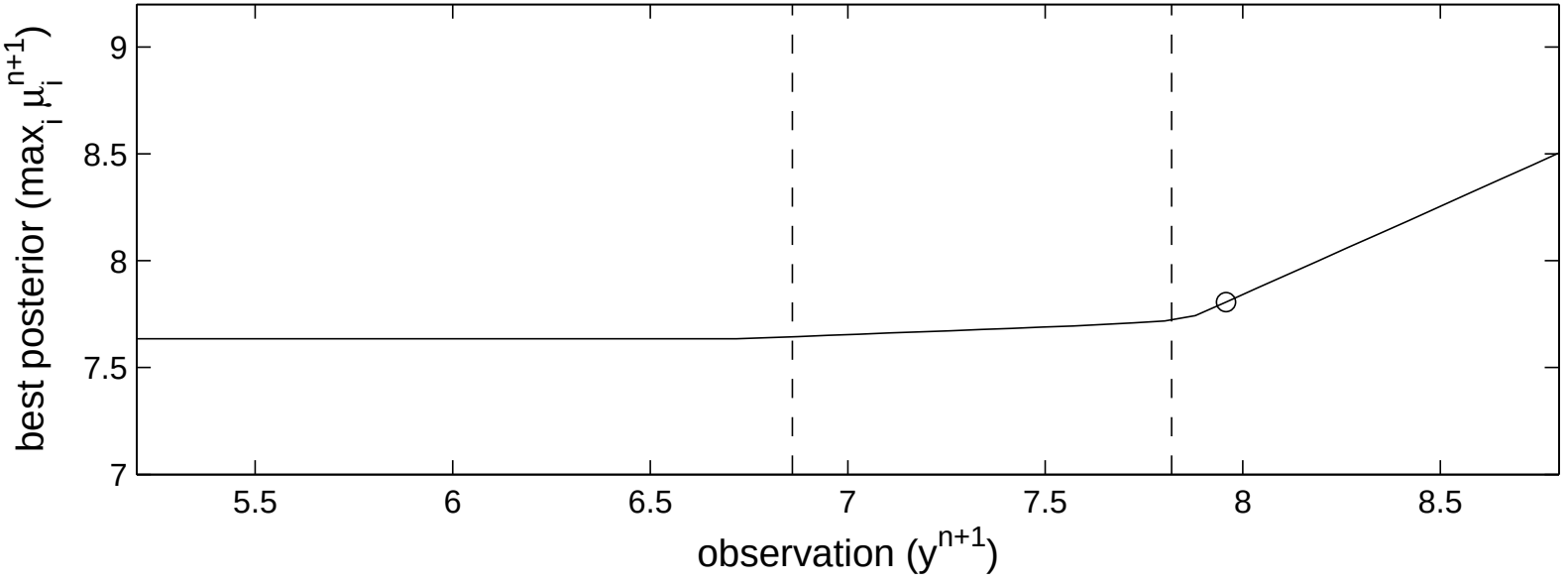
Computing the Knowledge-Gradient



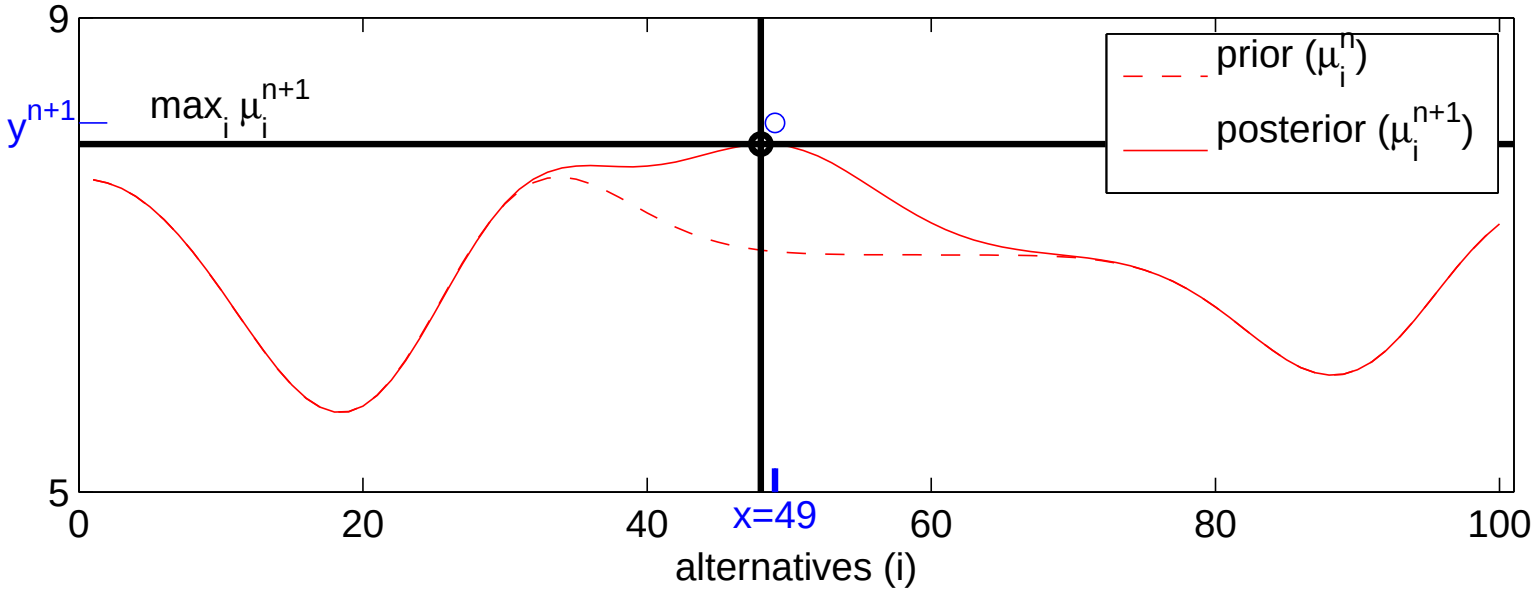


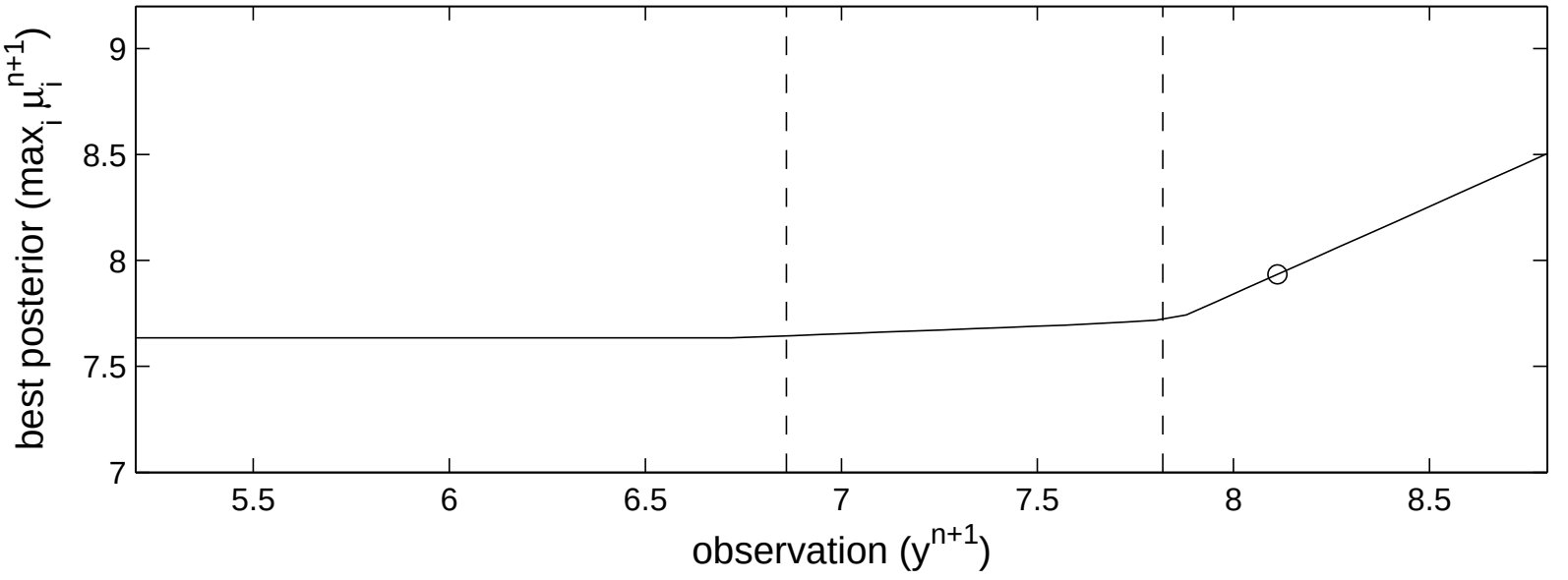
Computing the Knowledge-Gradient



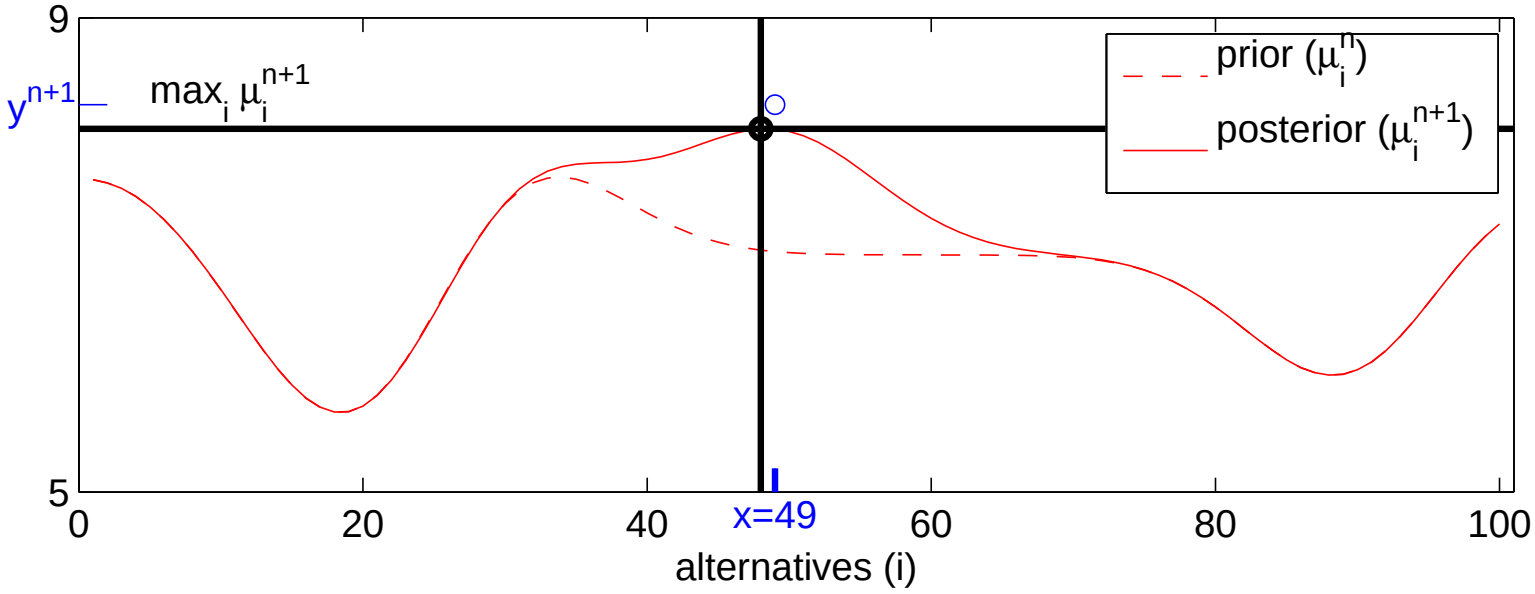


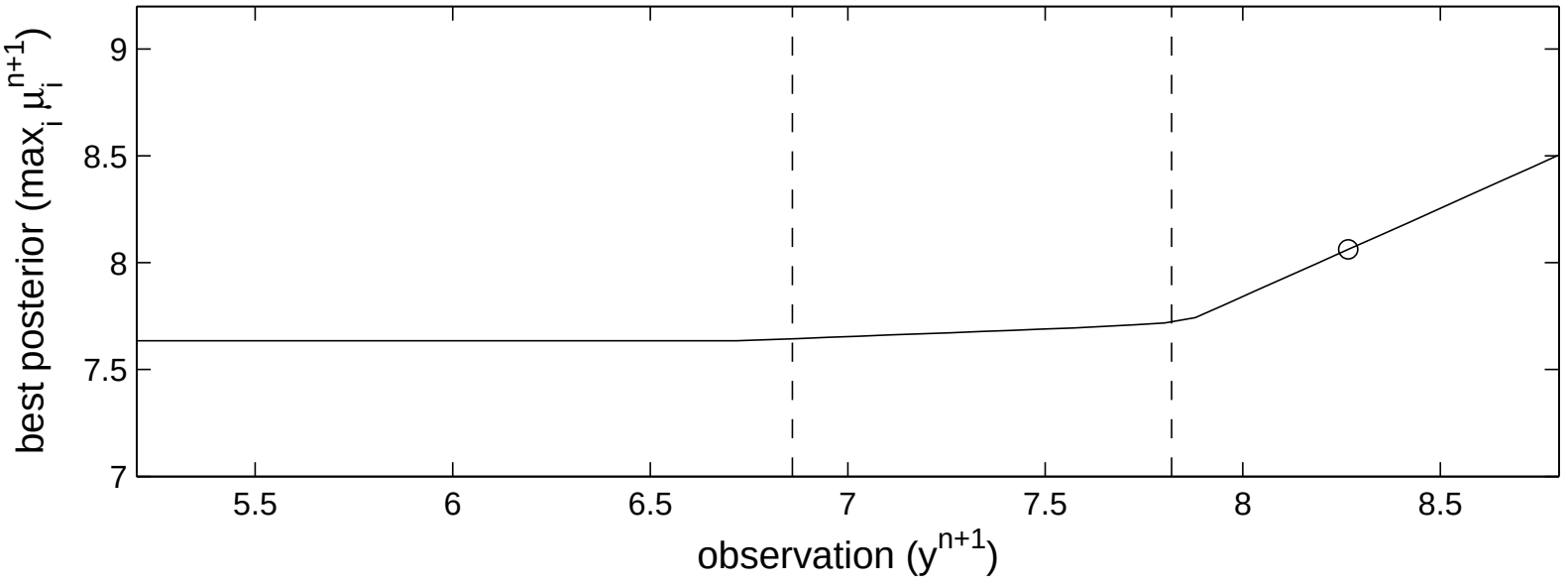
Computing the Knowledge-Gradient



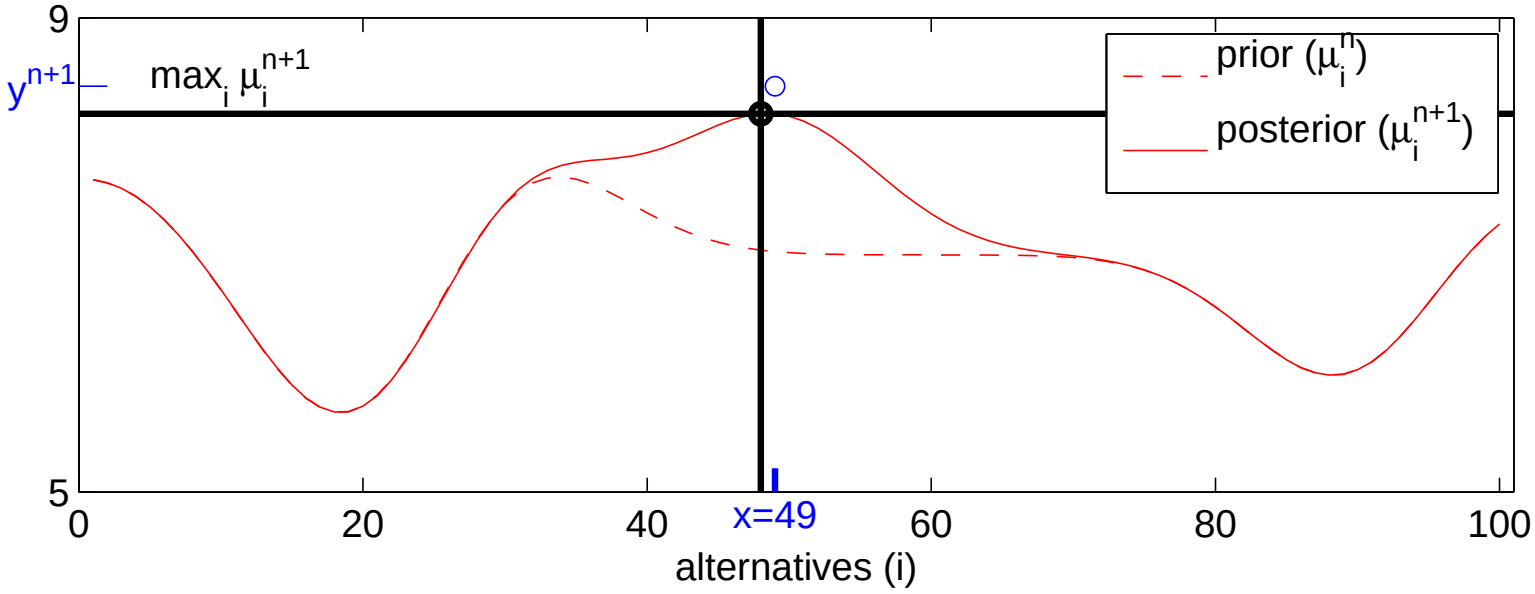


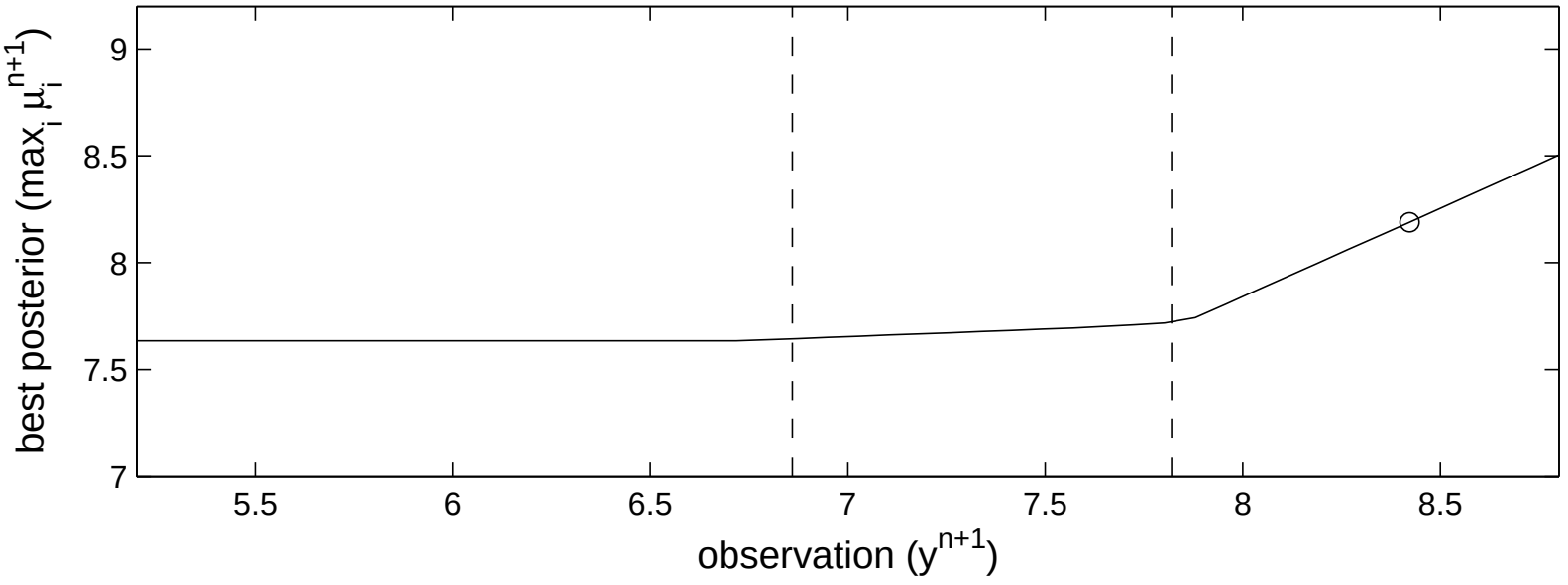
Computing the Knowledge-Gradient



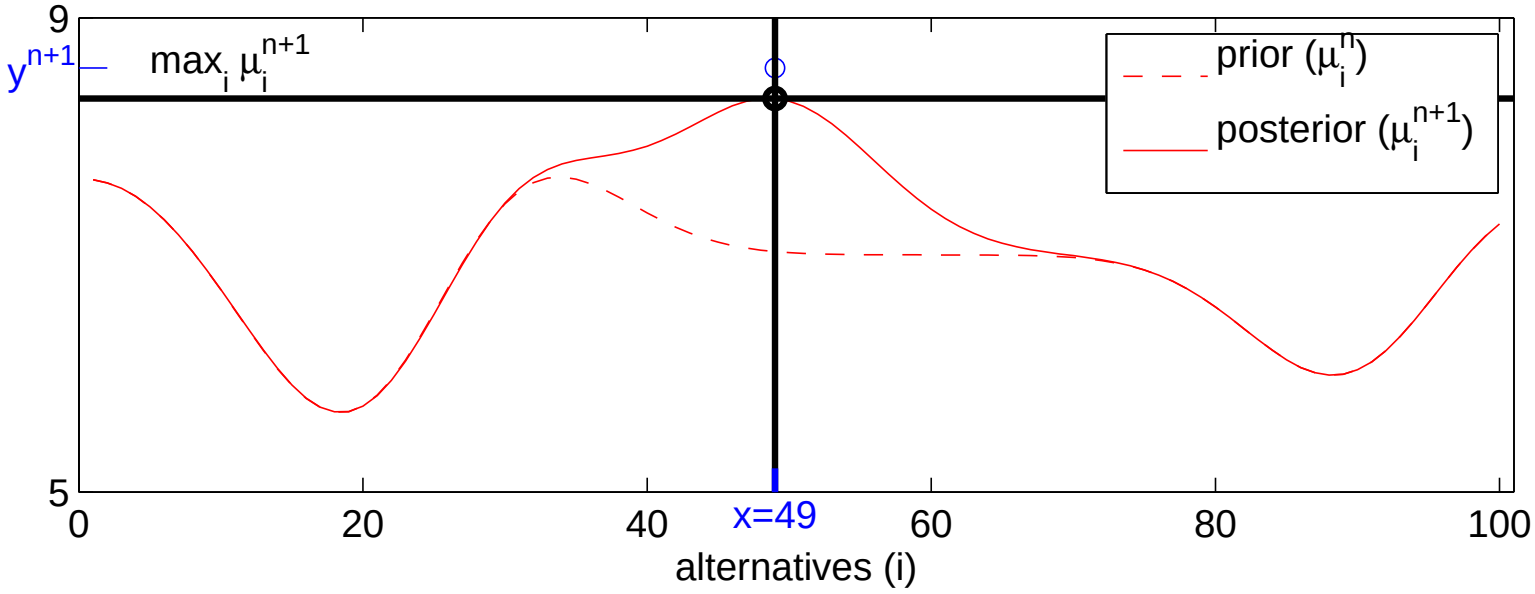


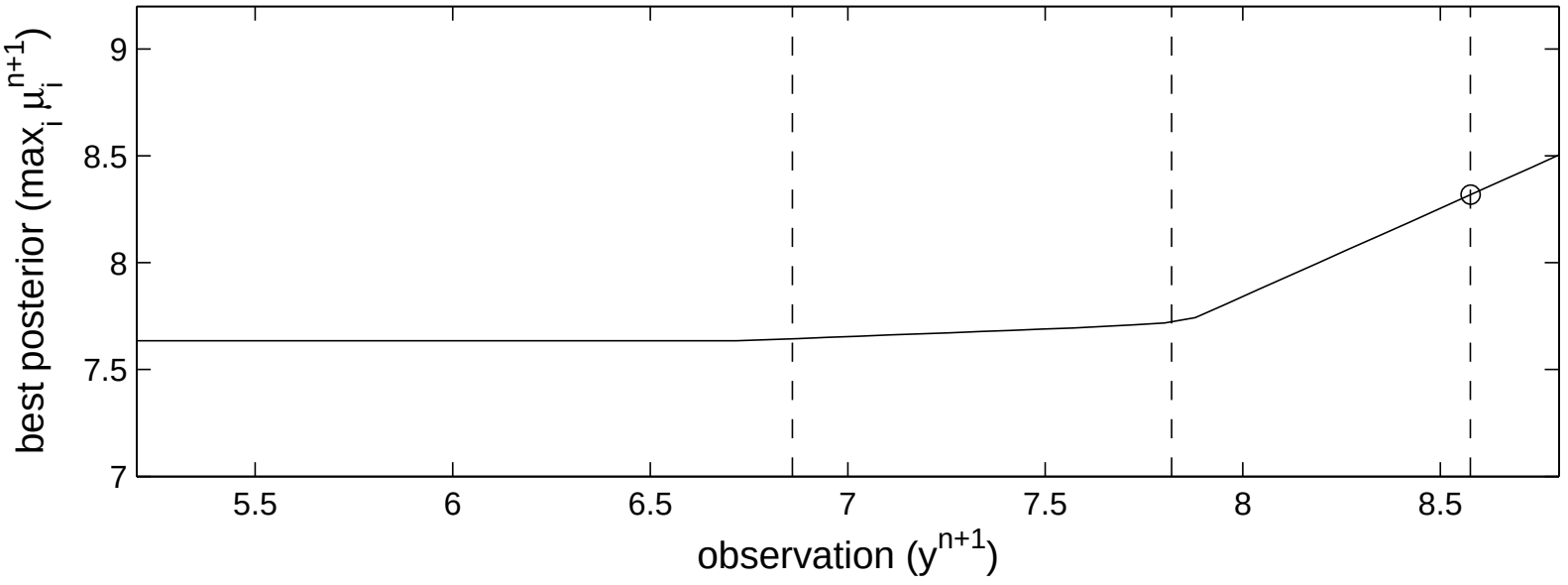
Computing the Knowledge-Gradient



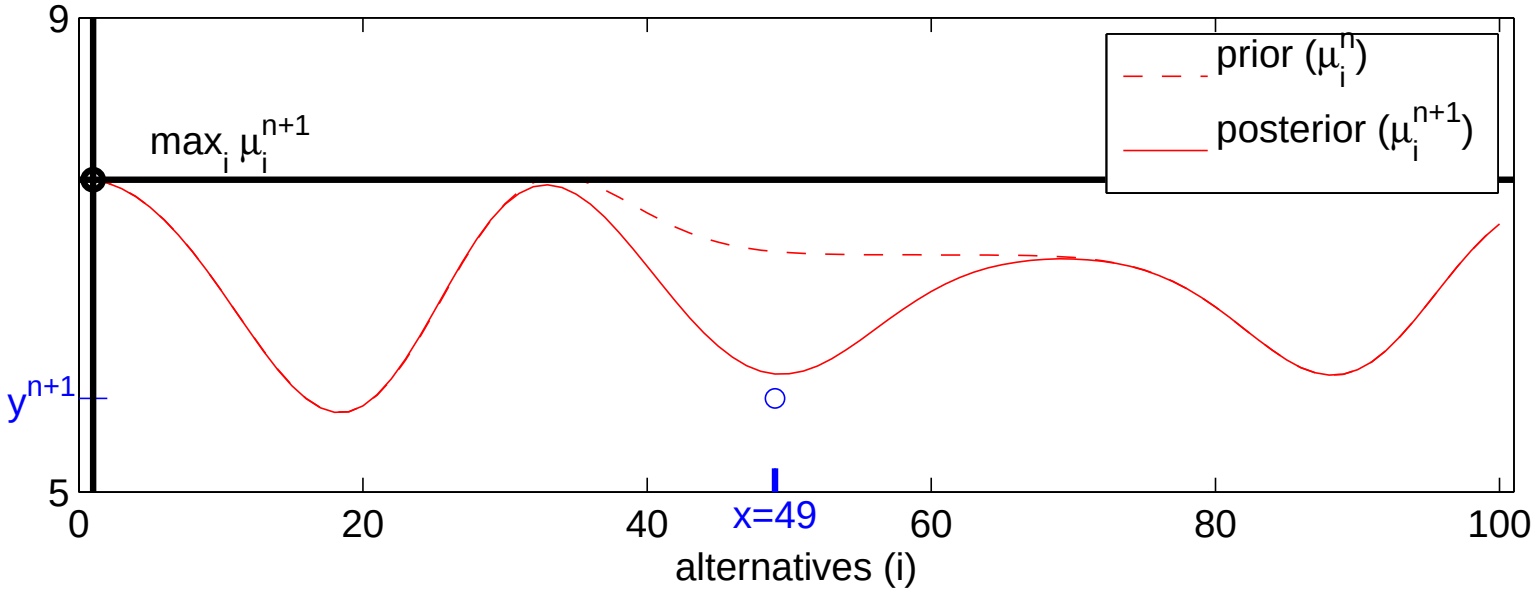


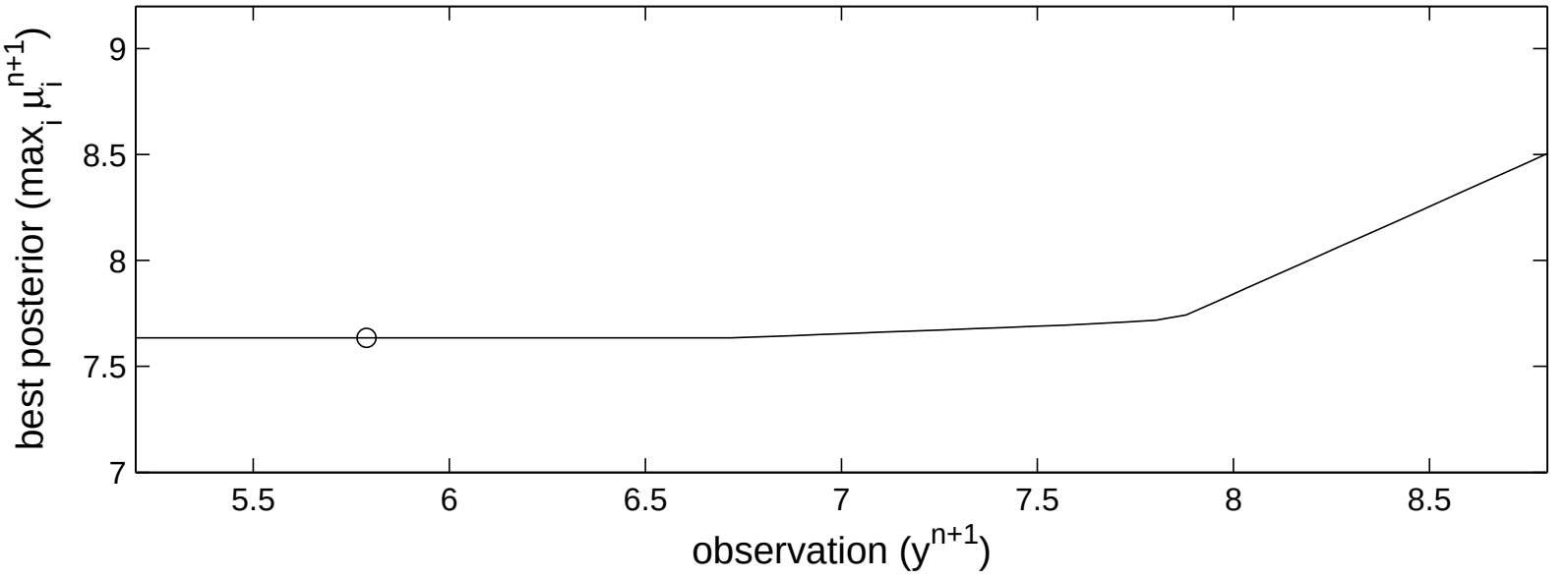
Computing the Knowledge-Gradient



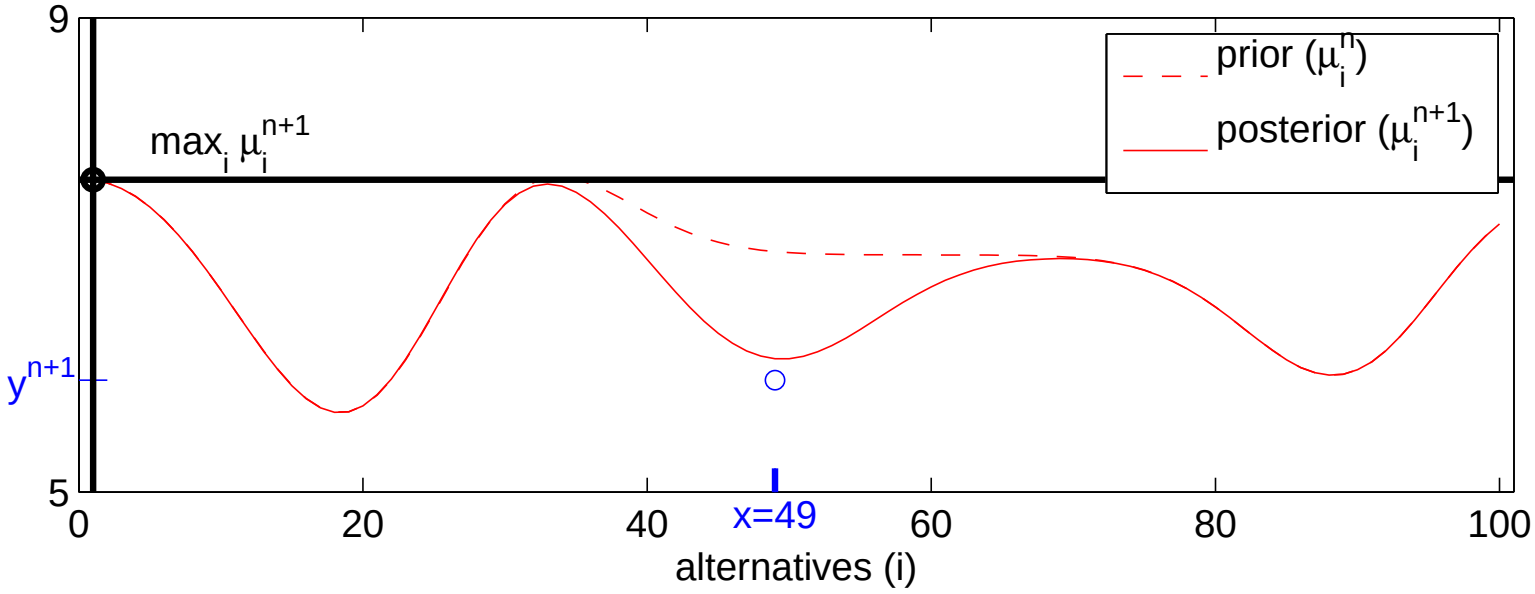


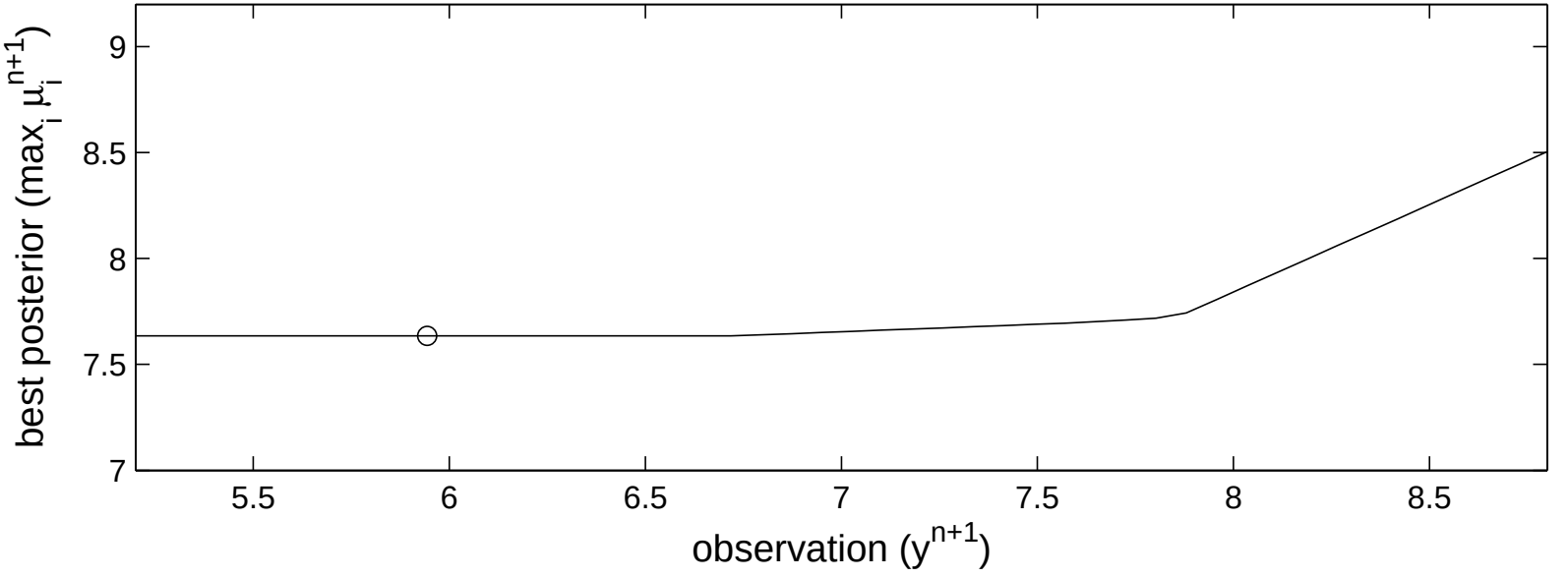
Computing the Knowledge-Gradient



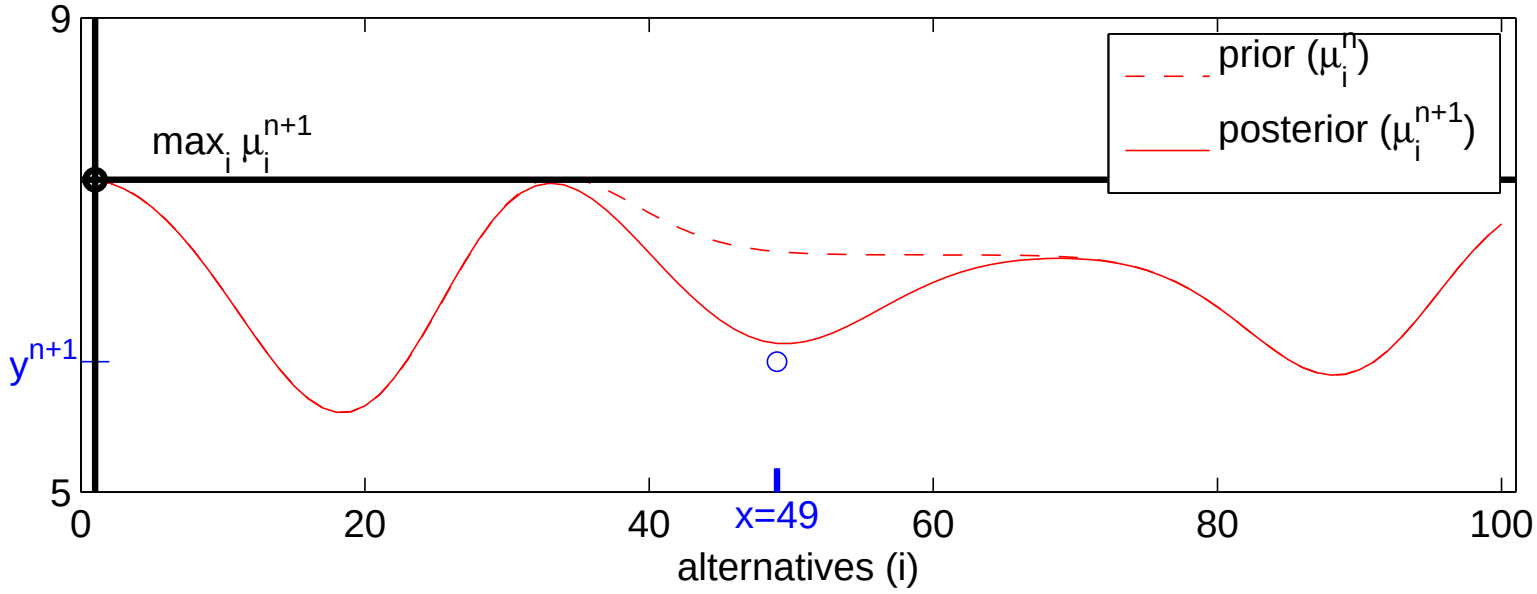


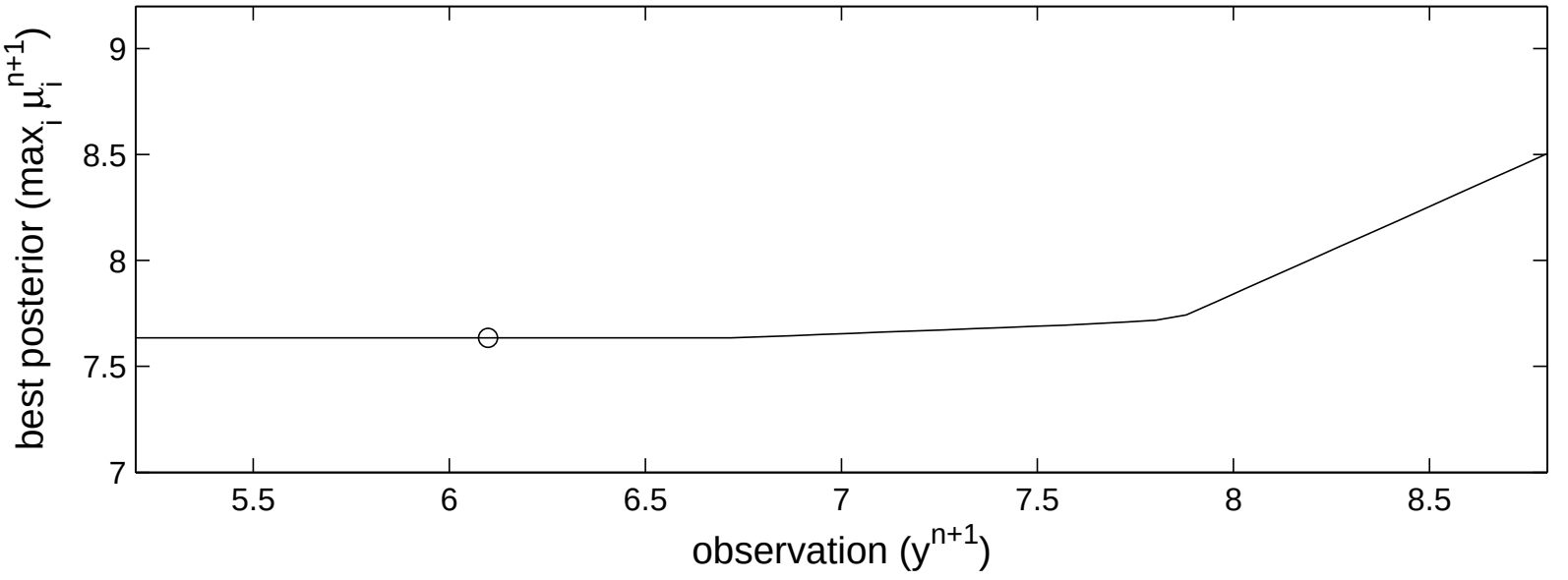
Computing the Knowledge-Gradient



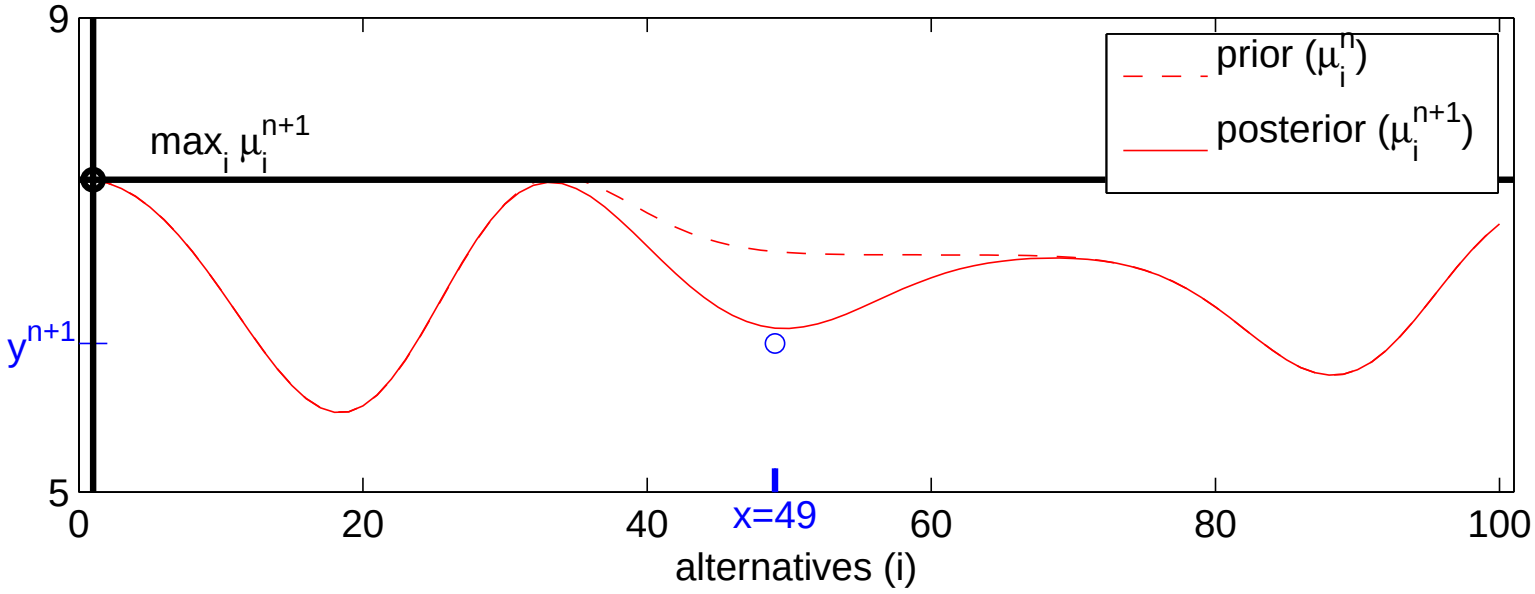


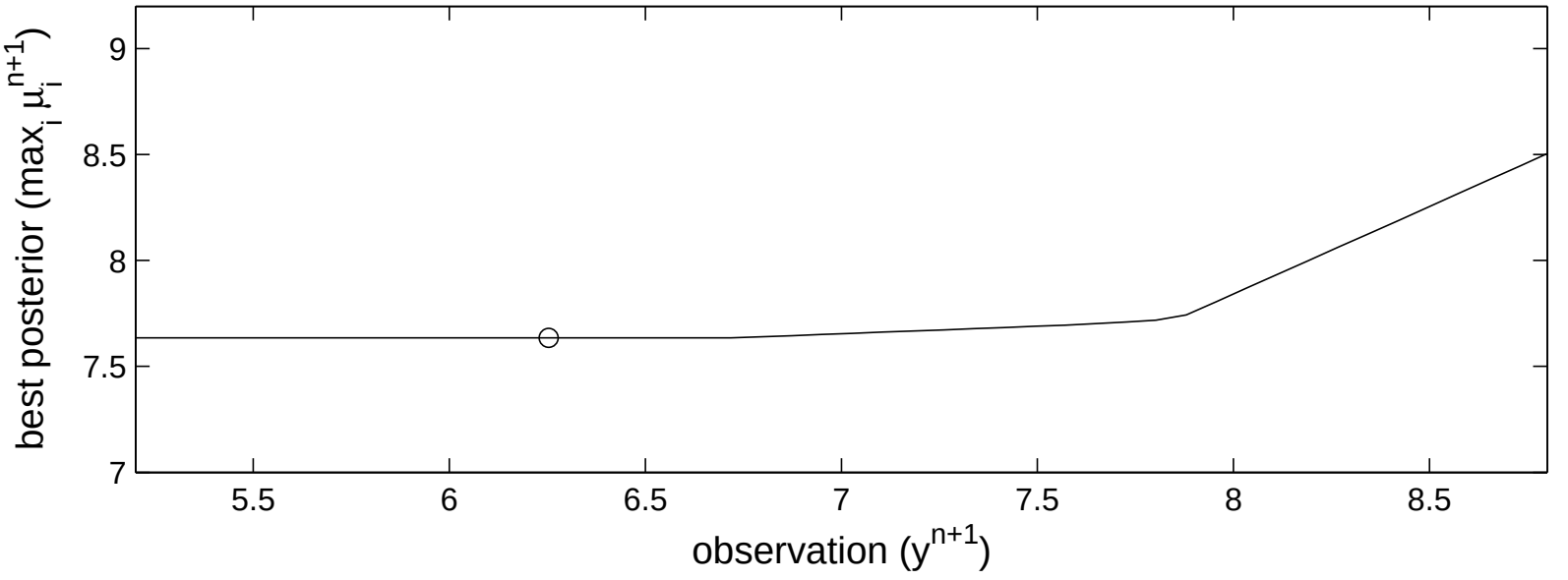
Computing the Knowledge-Gradient



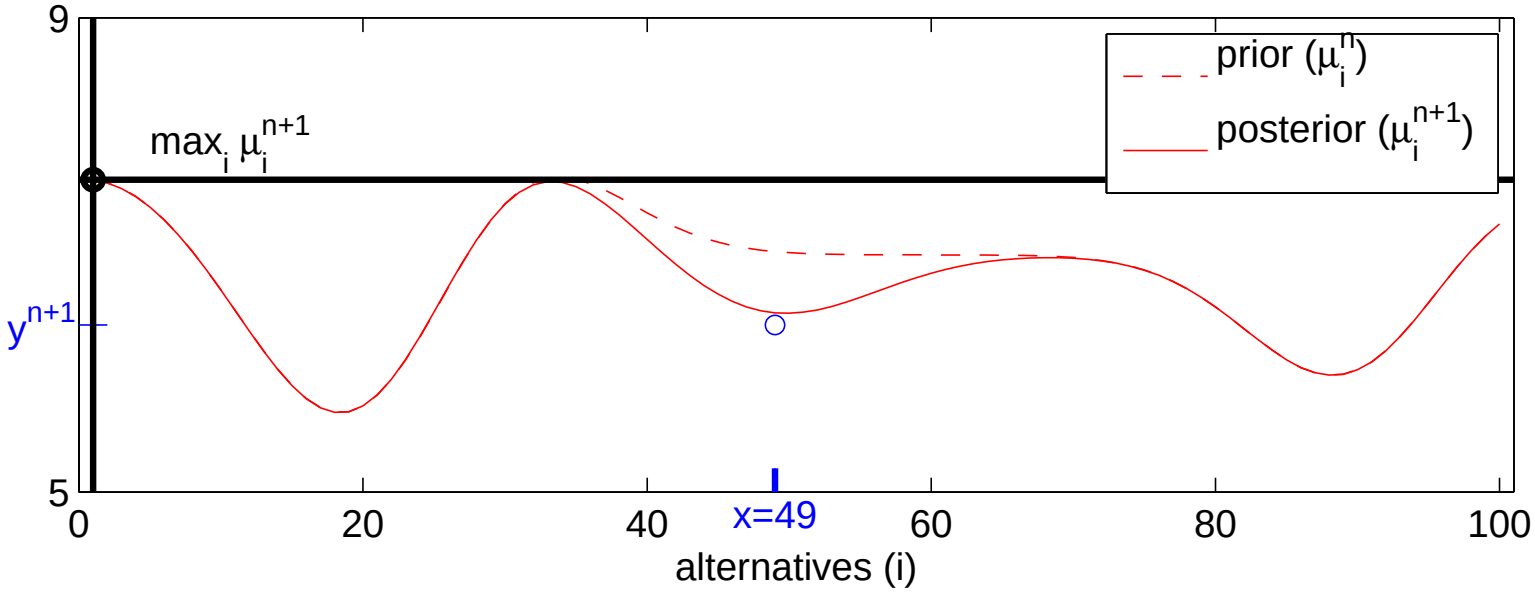


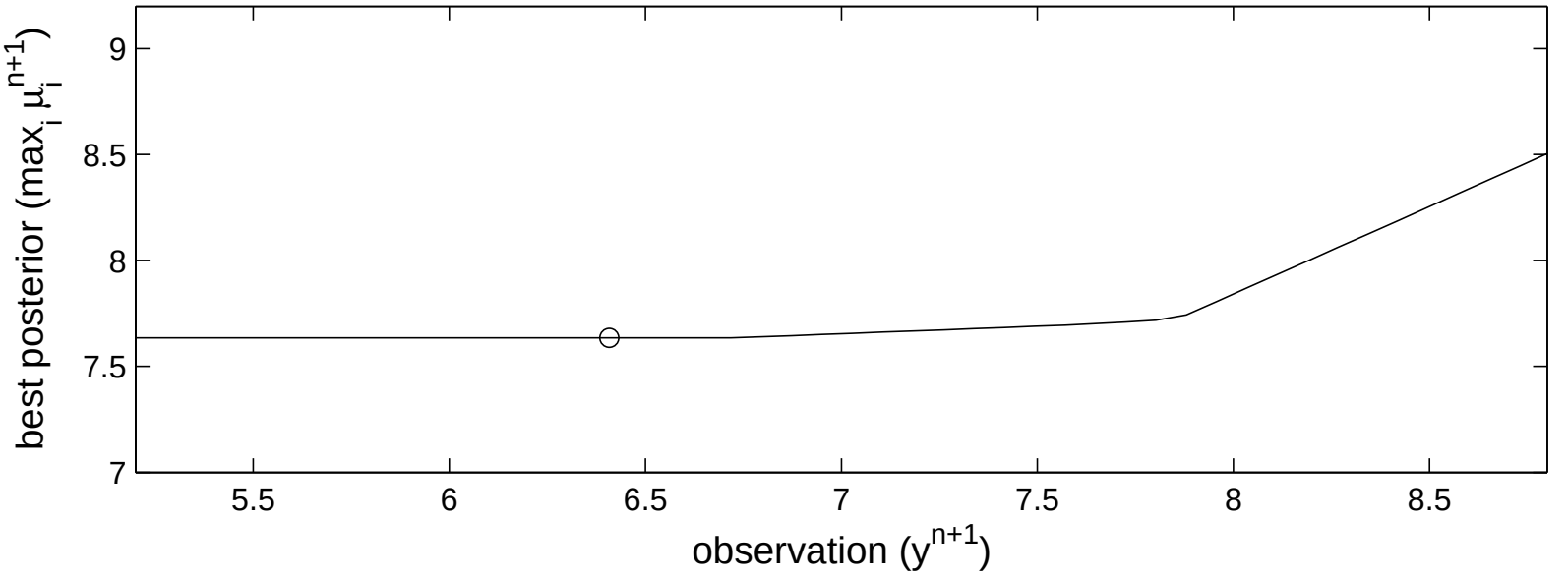
Computing the Knowledge-Gradient



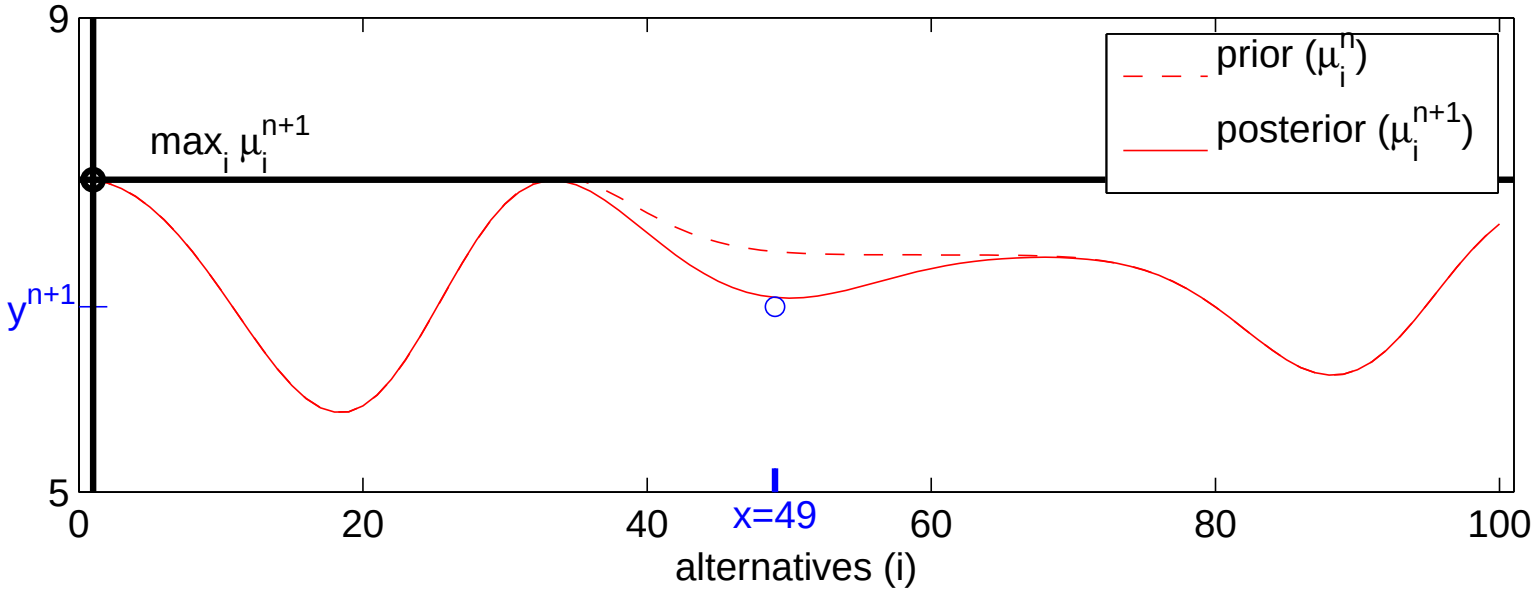


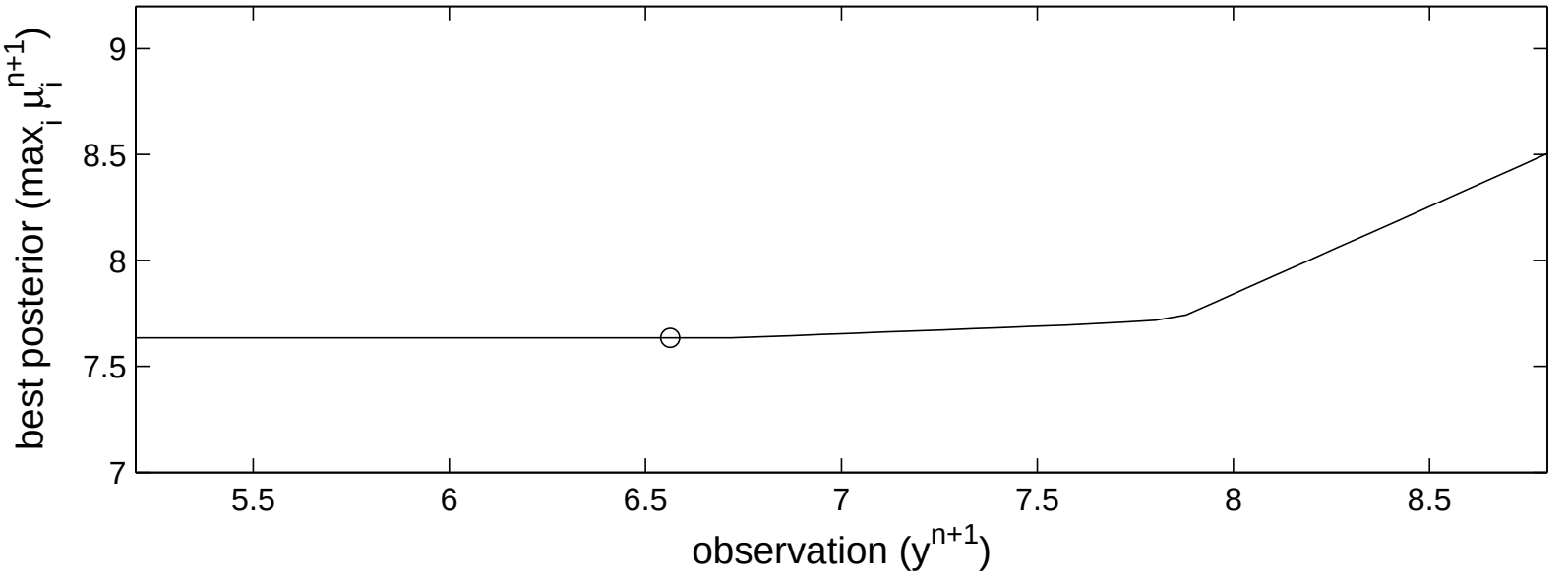
Computing the Knowledge-Gradient



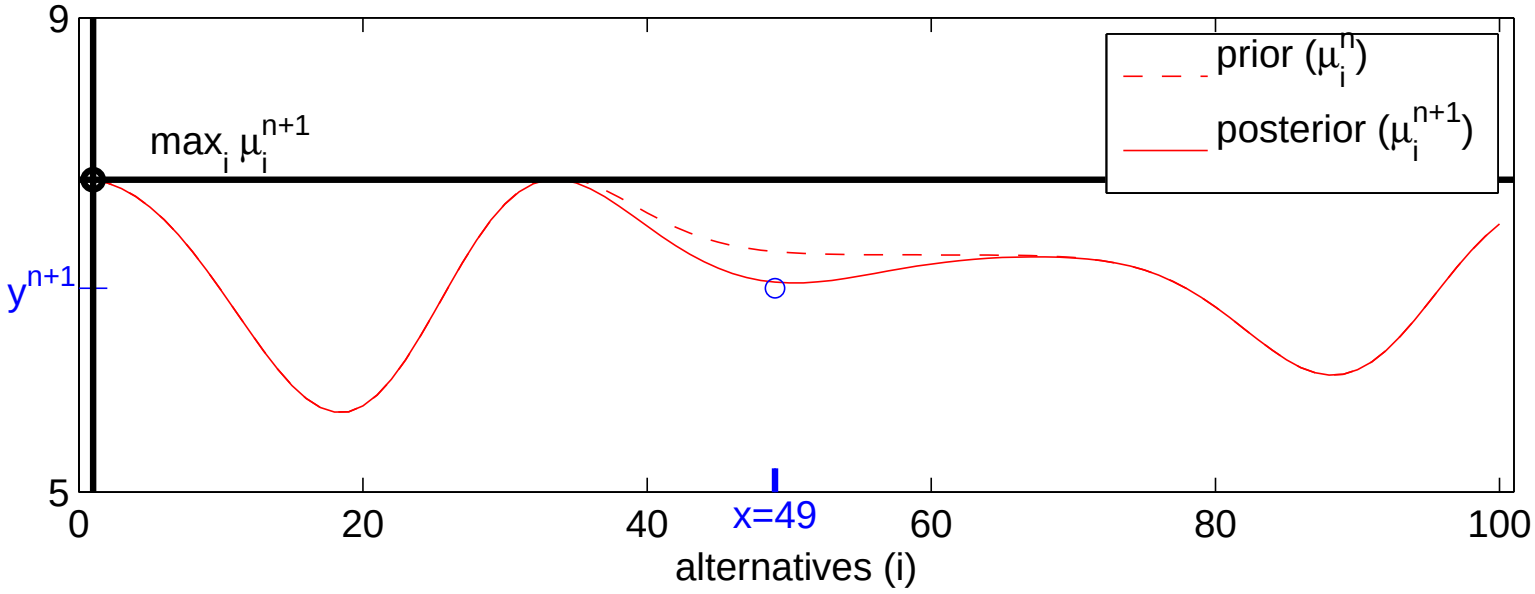


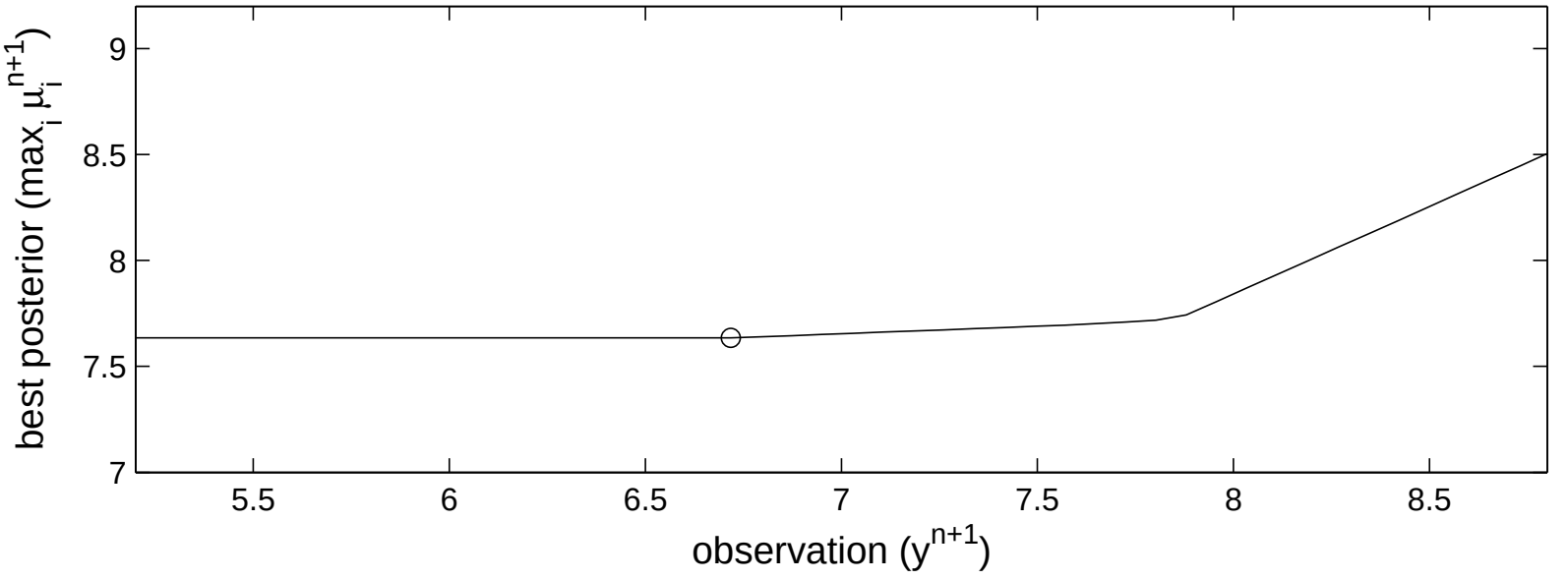
Computing the Knowledge-Gradient



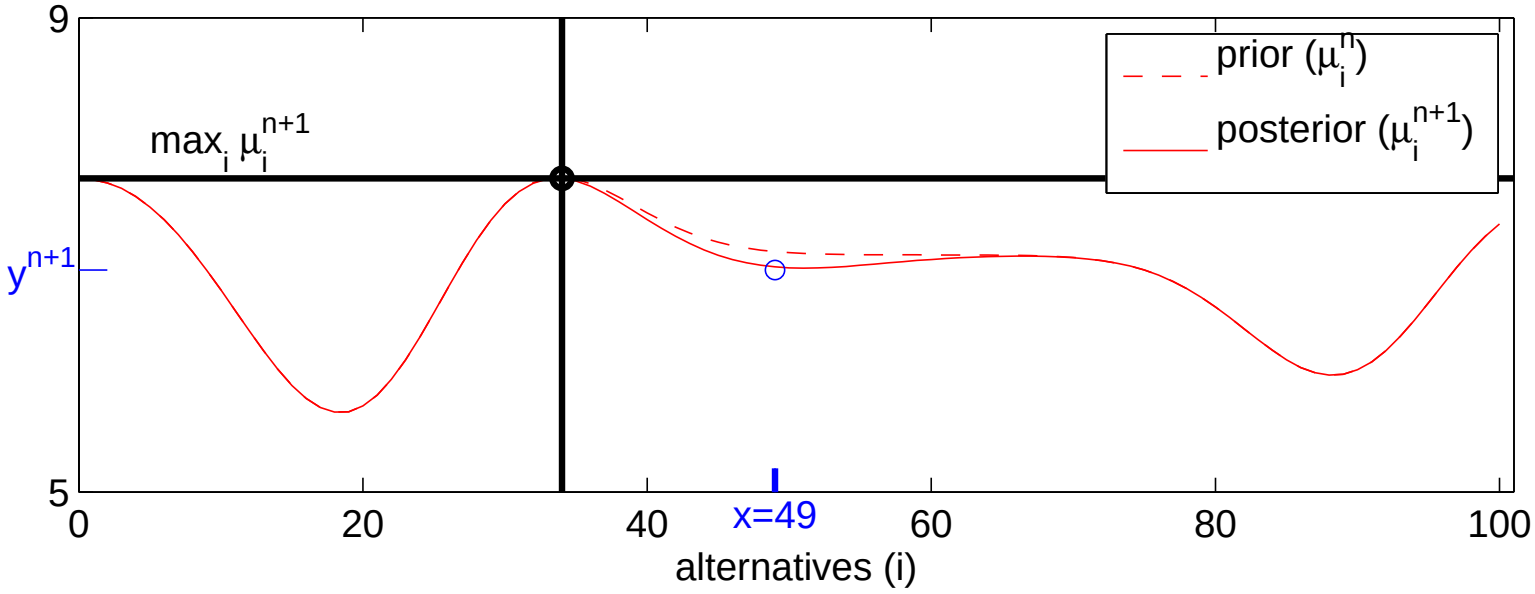


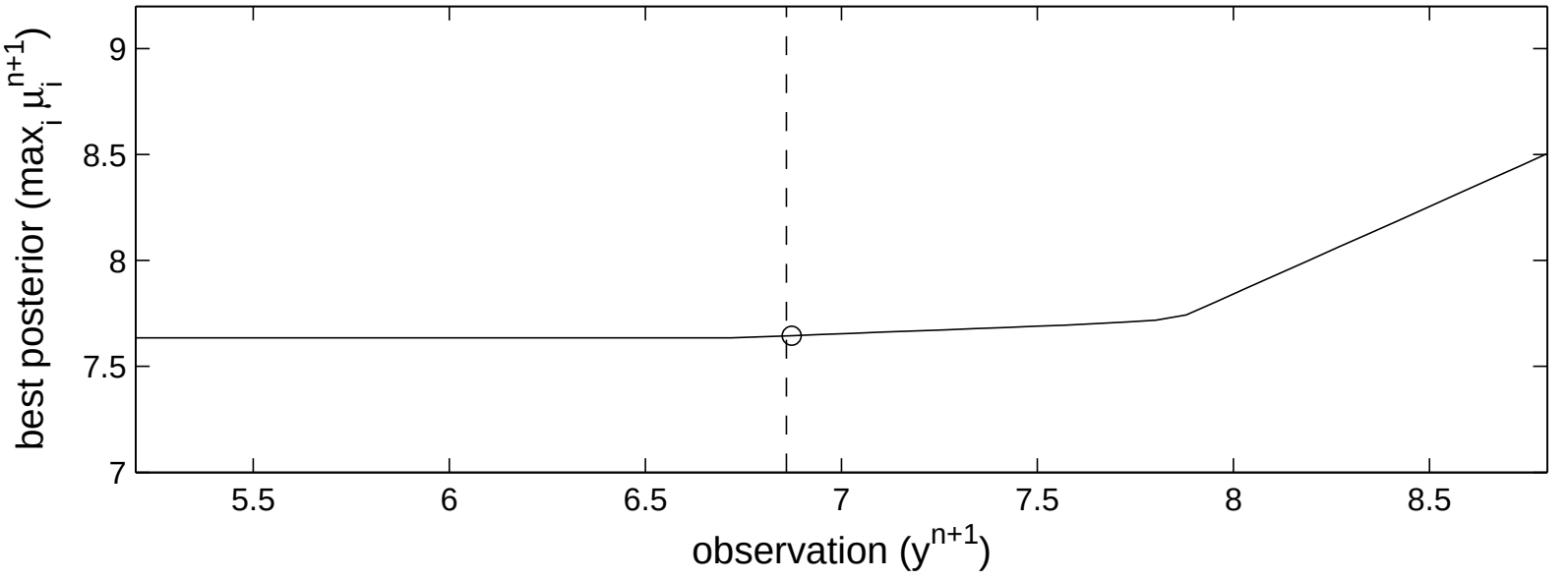
Computing the Knowledge-Gradient



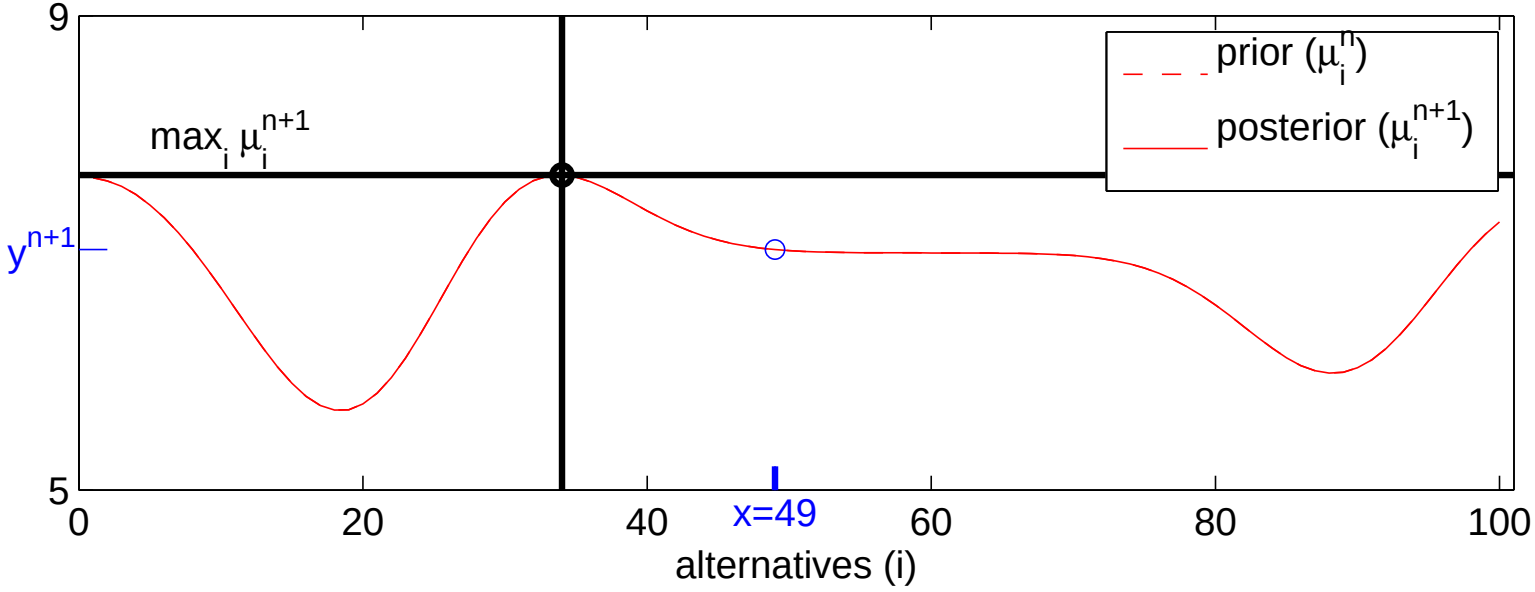


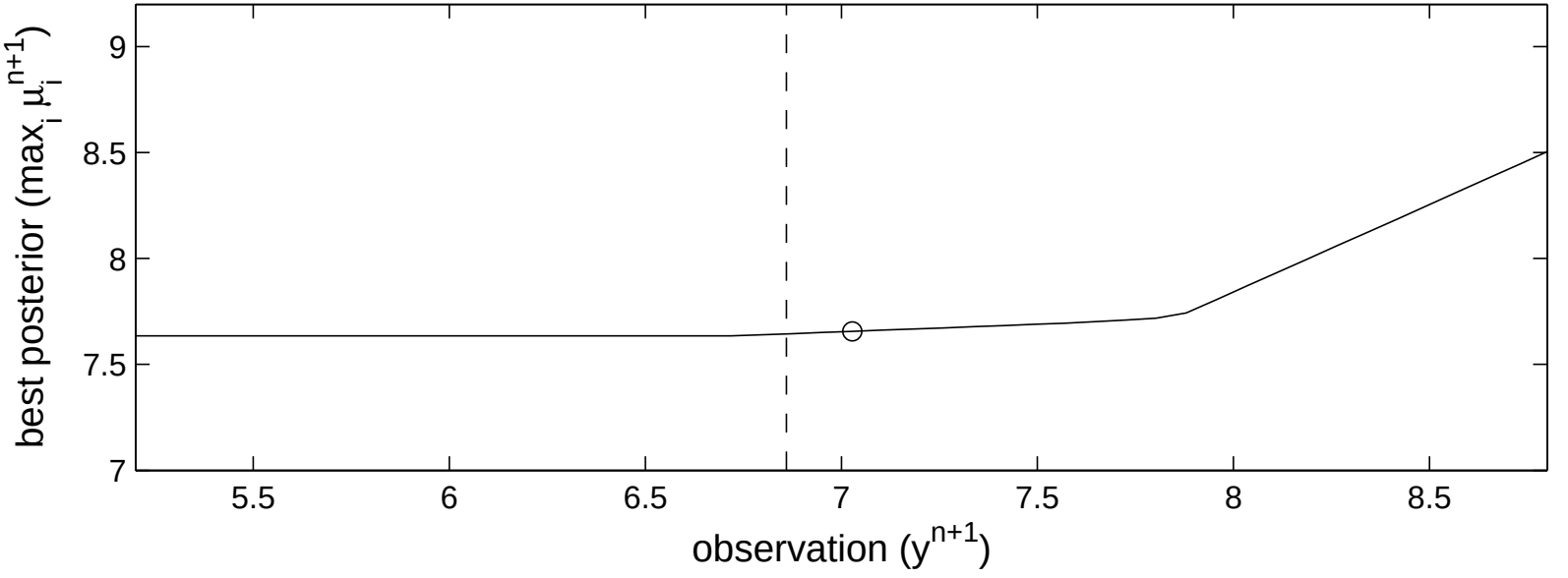
Computing the Knowledge-Gradient



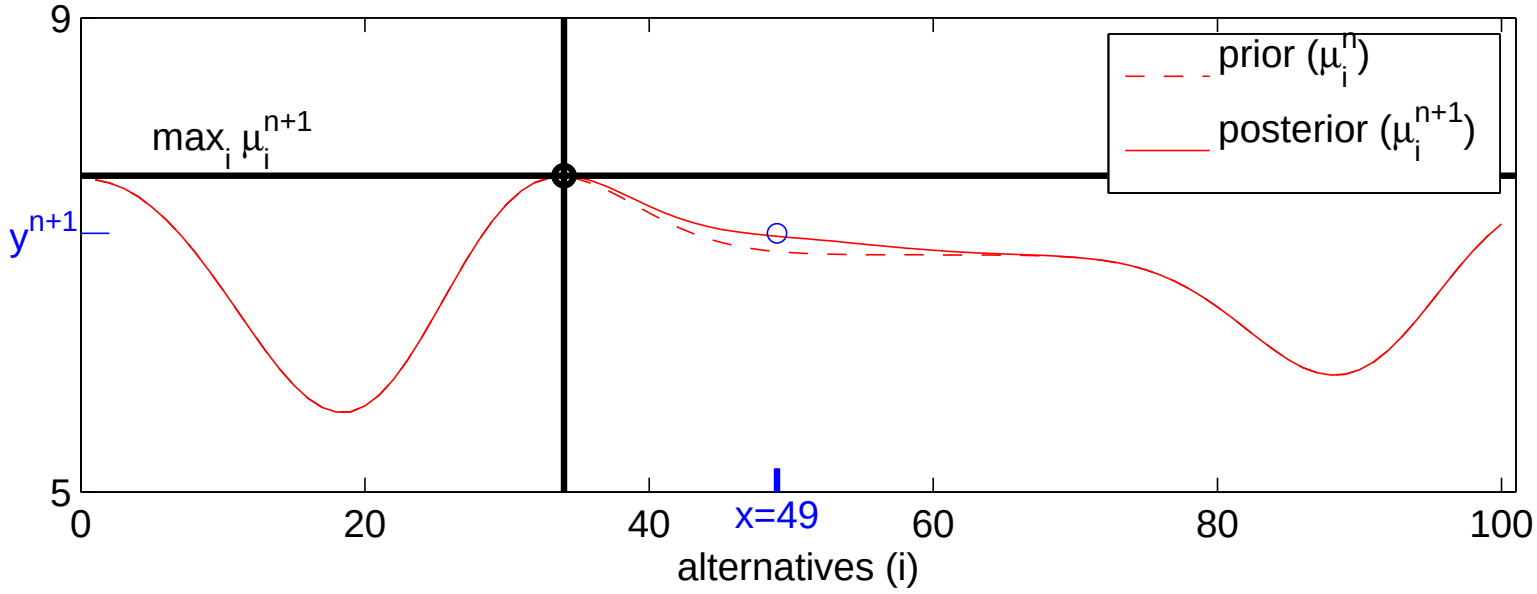


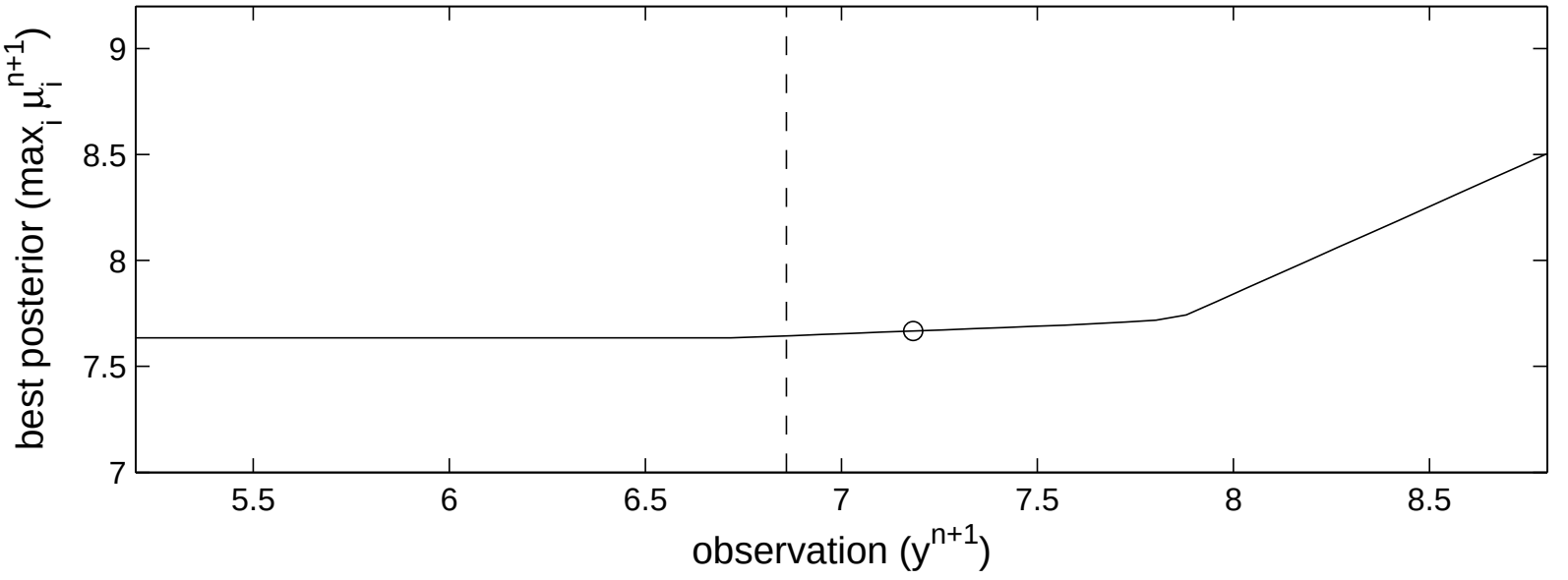
Computing the Knowledge-Gradient



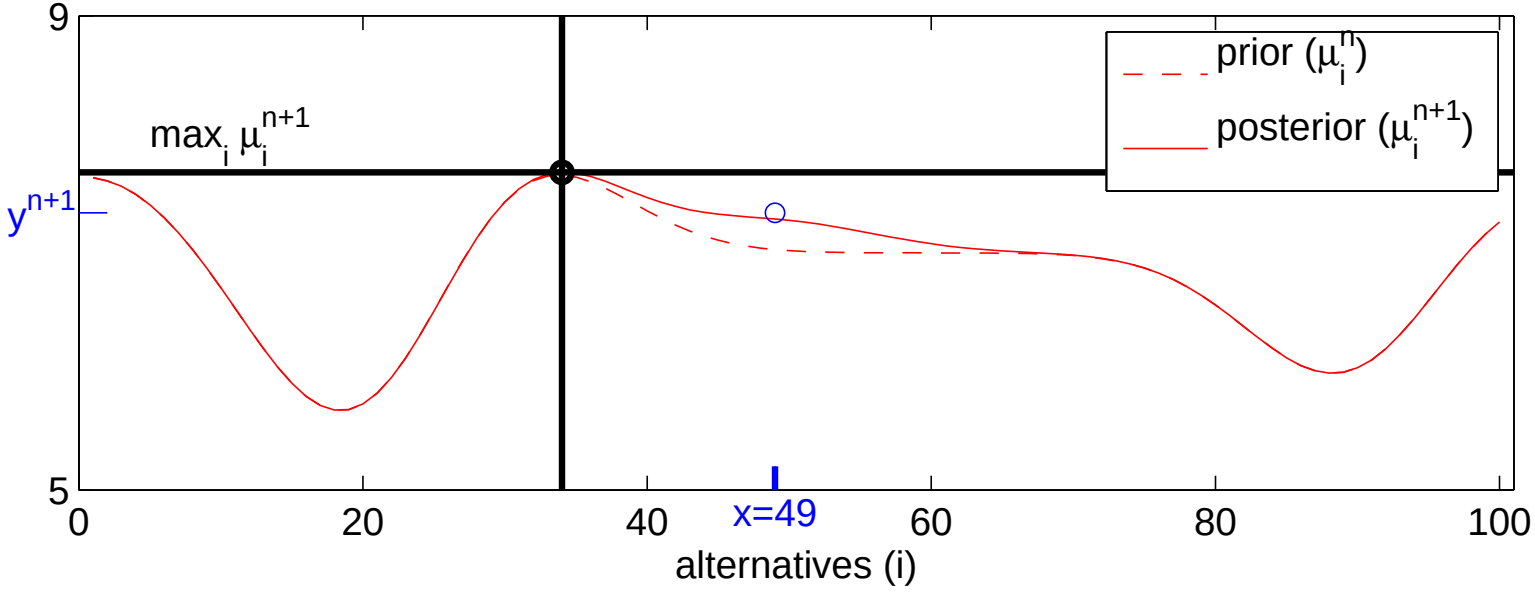


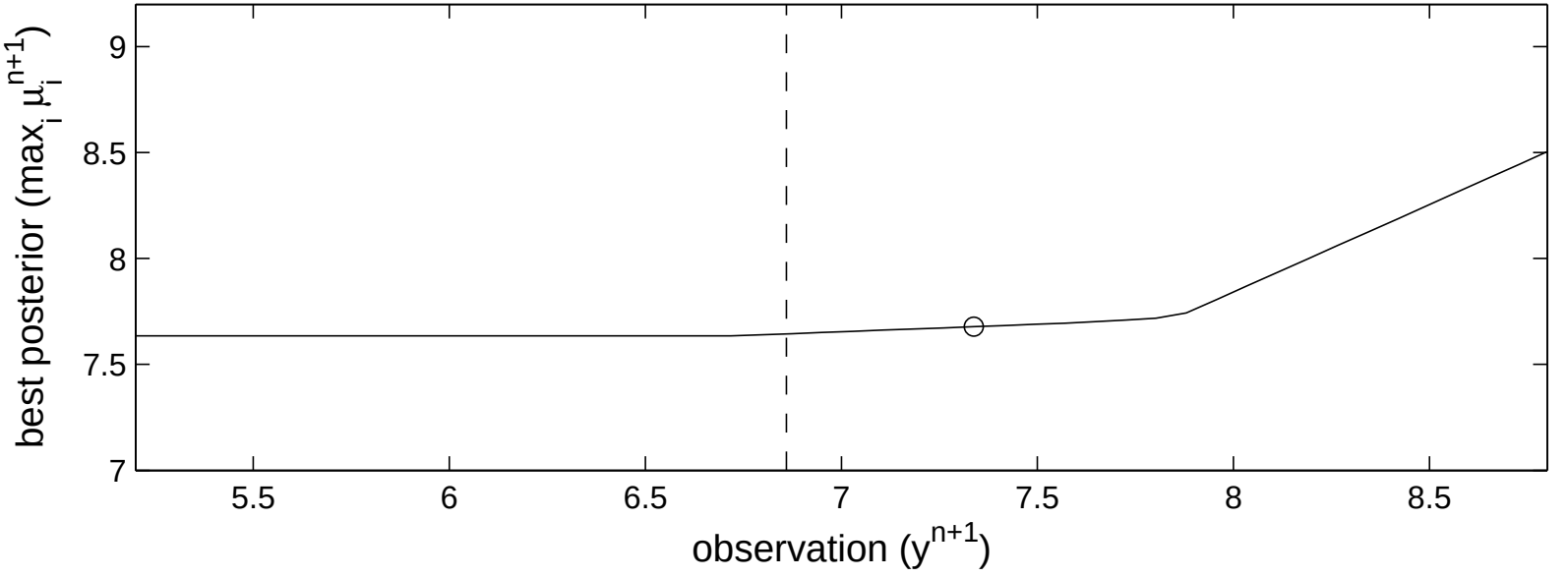
Computing the Knowledge-Gradient



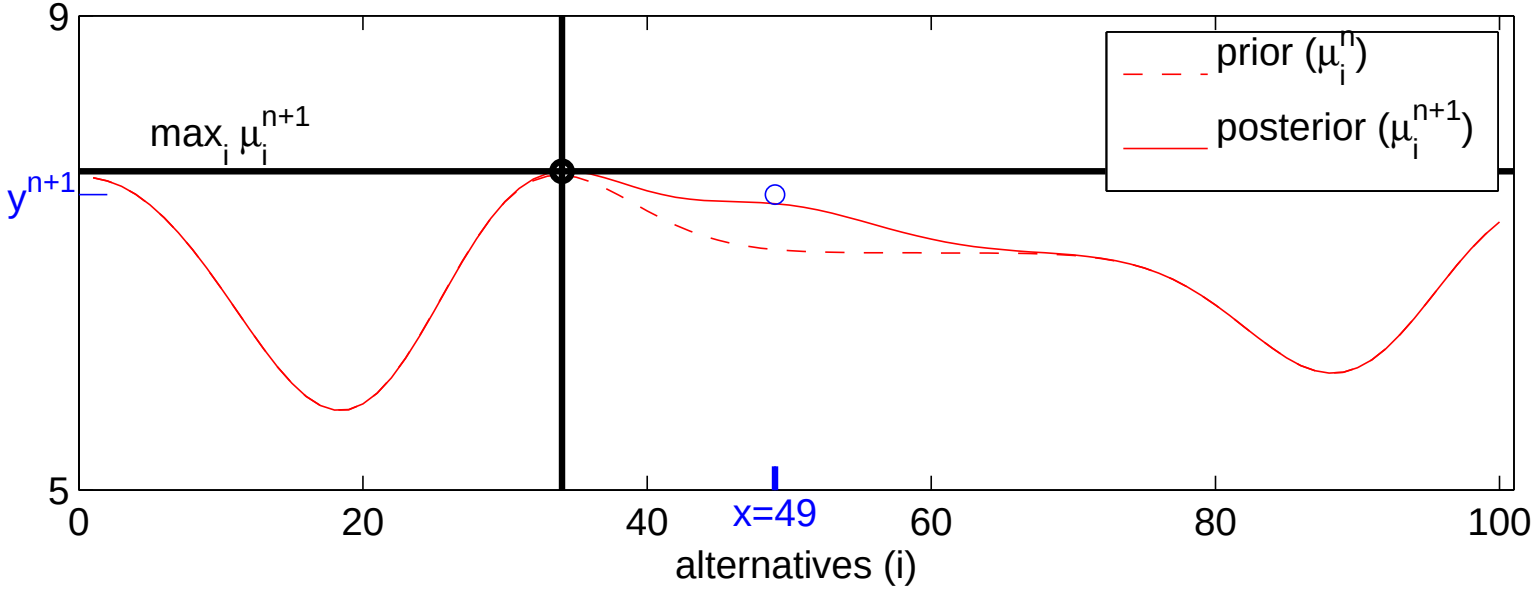


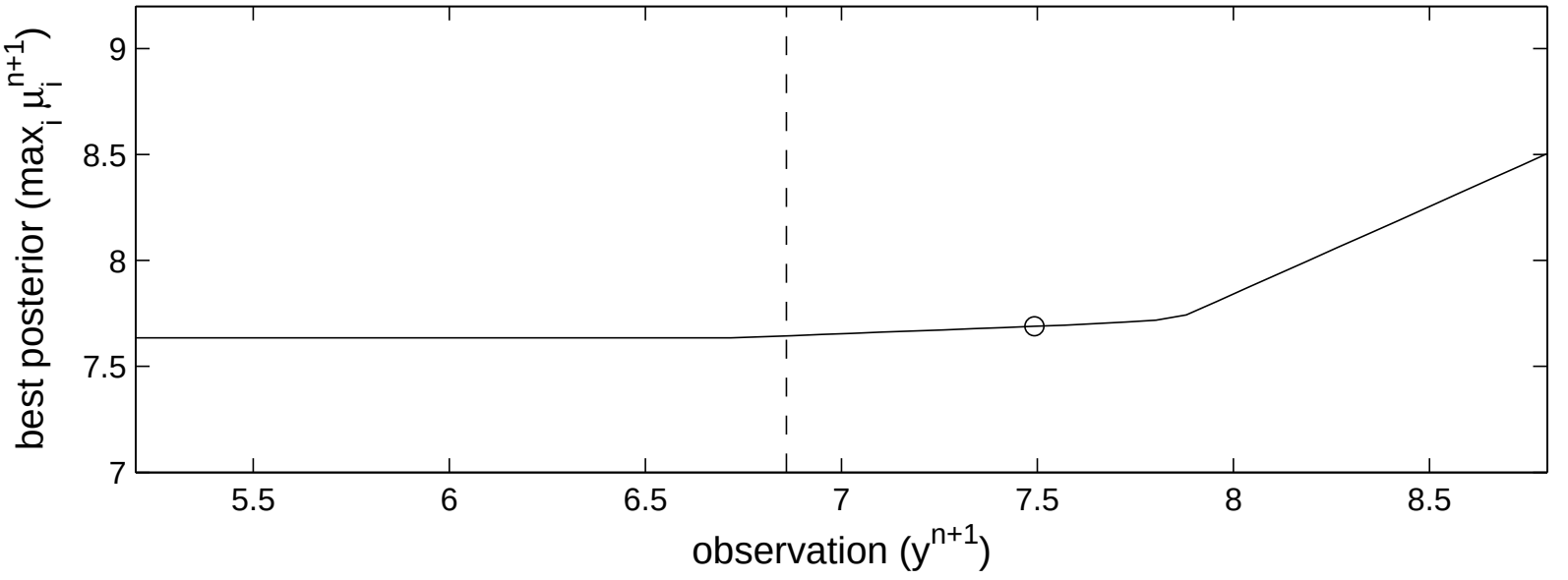
Computing the Knowledge-Gradient



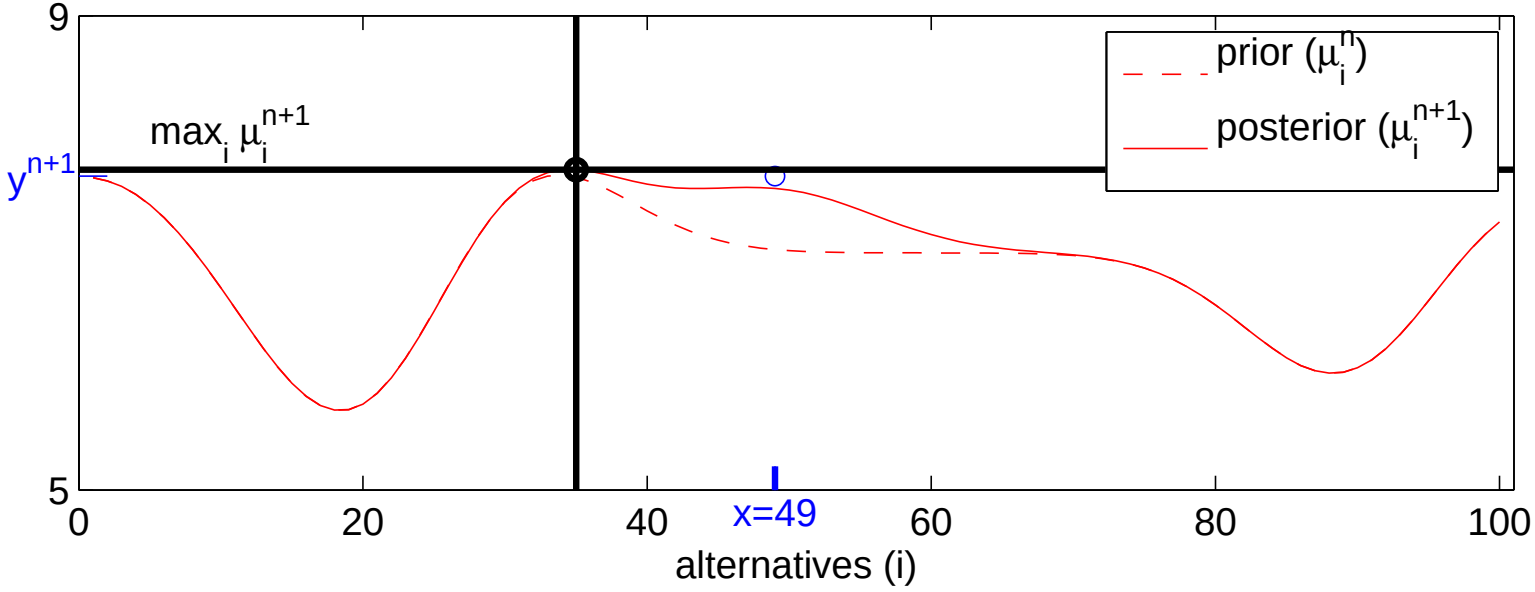


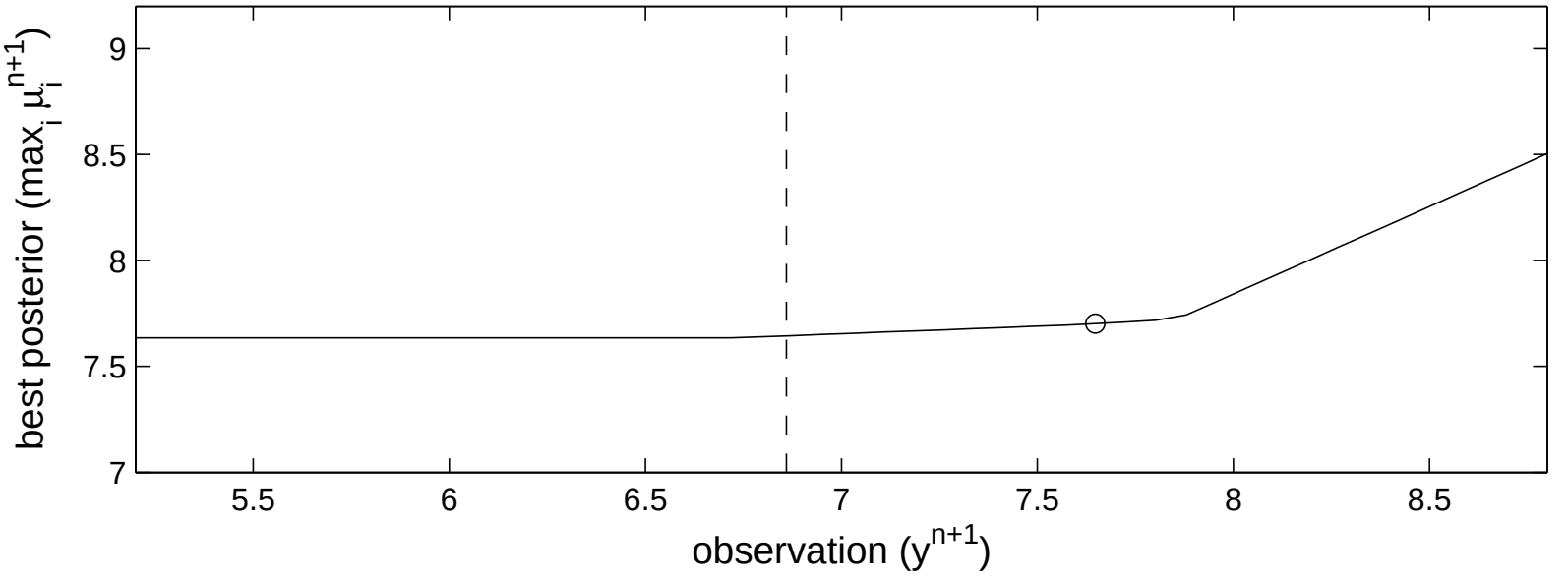
Computing the Knowledge-Gradient



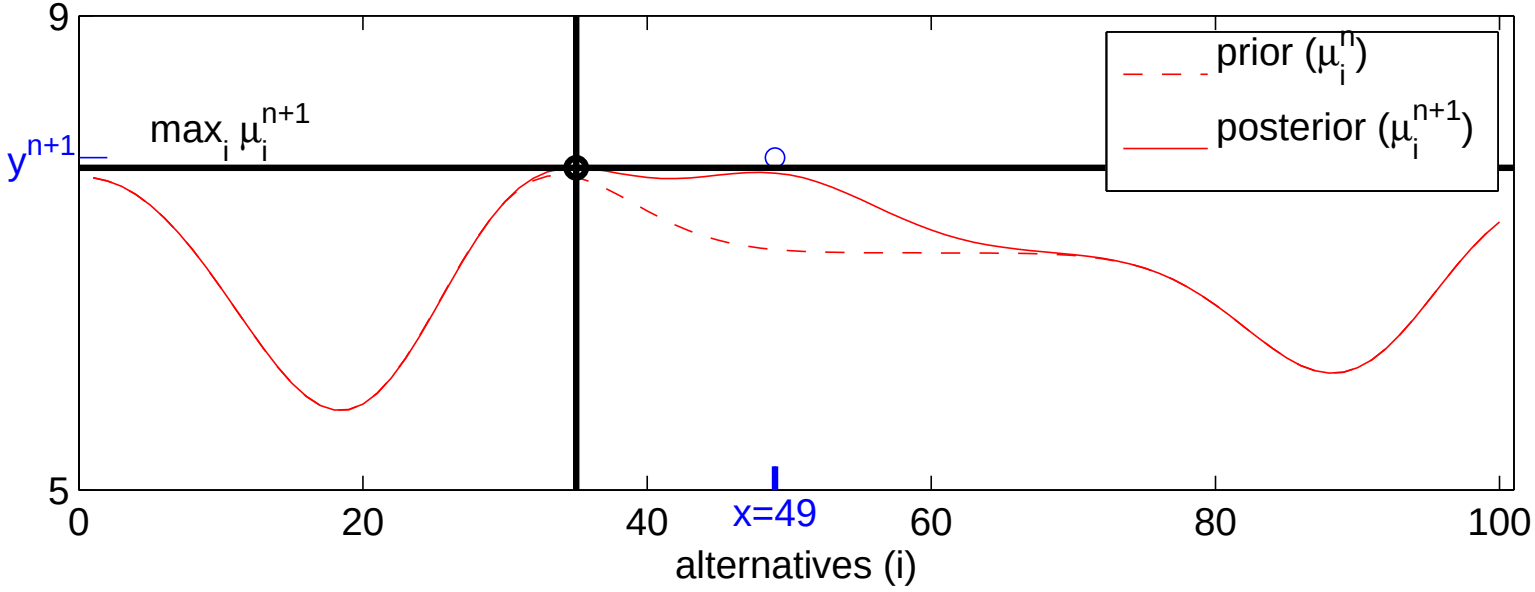


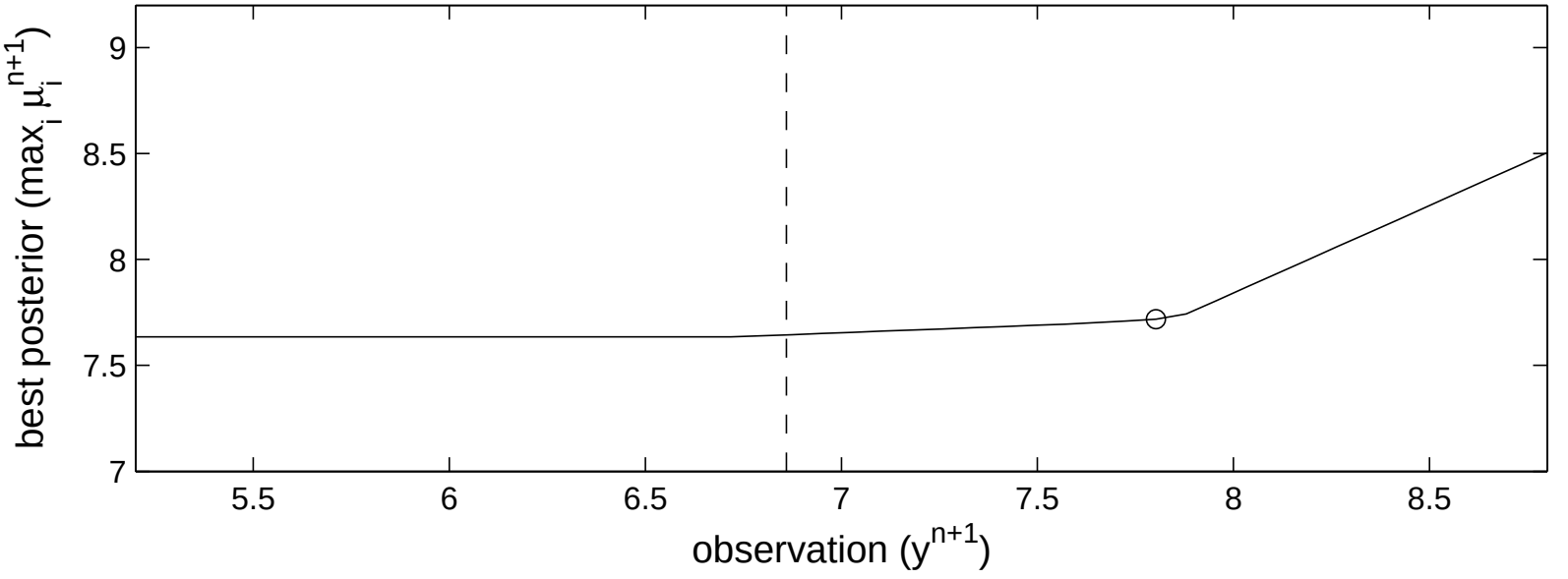
Computing the Knowledge-Gradient



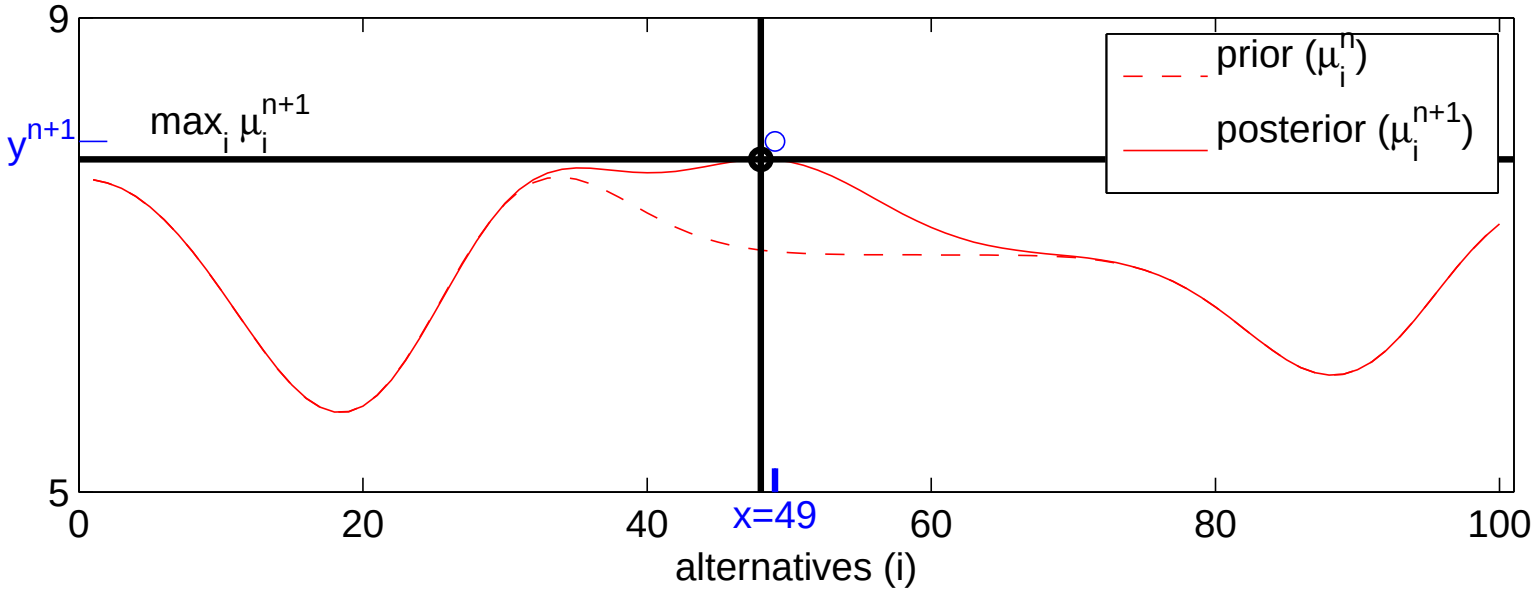


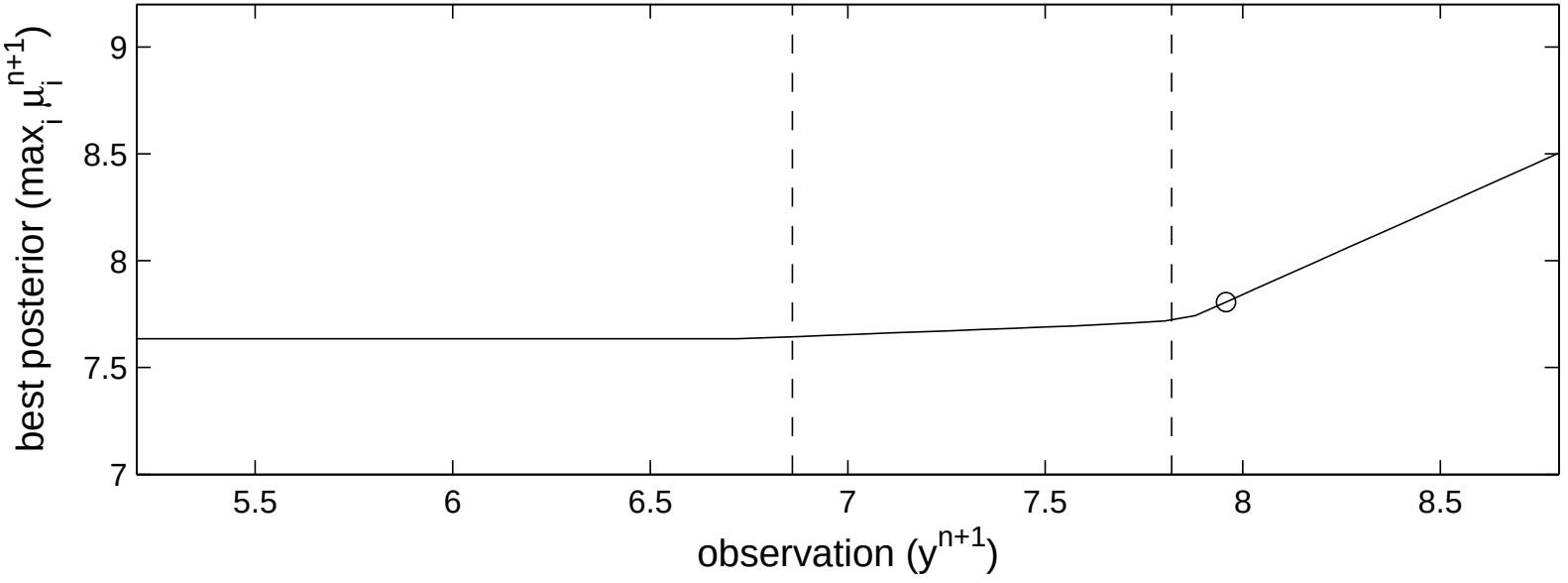
Computing the Knowledge-Gradient



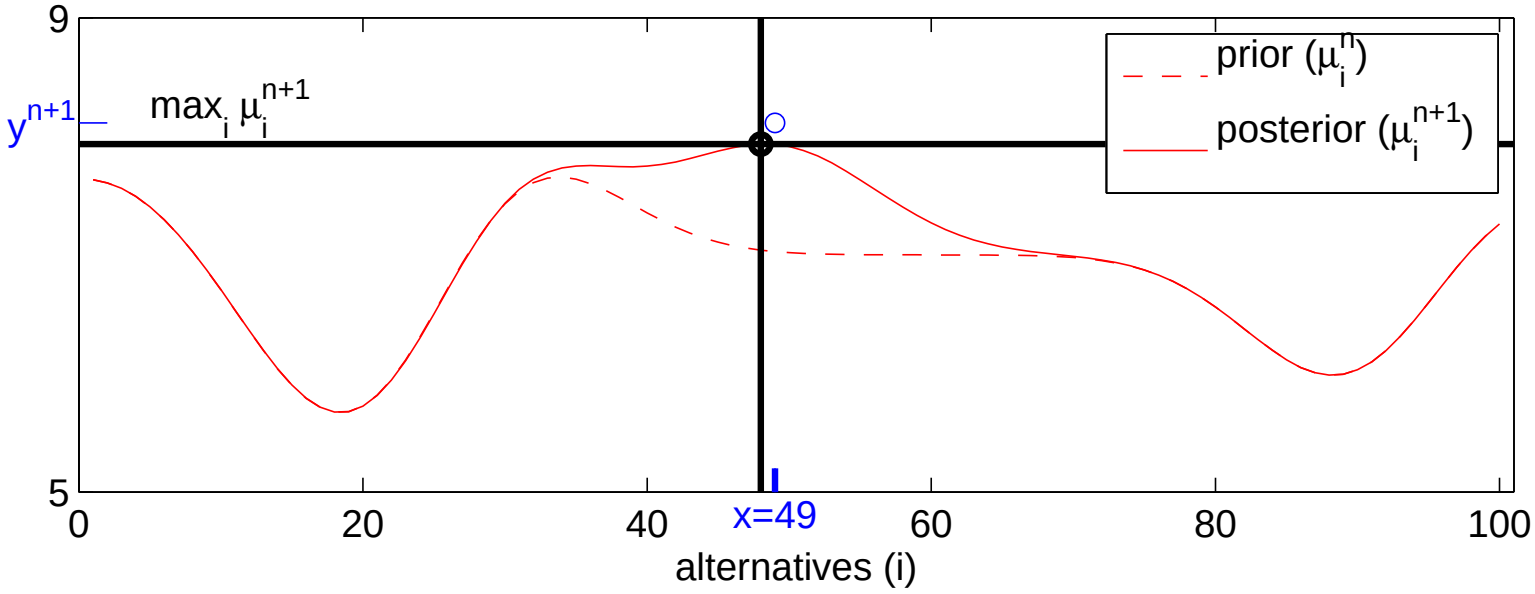


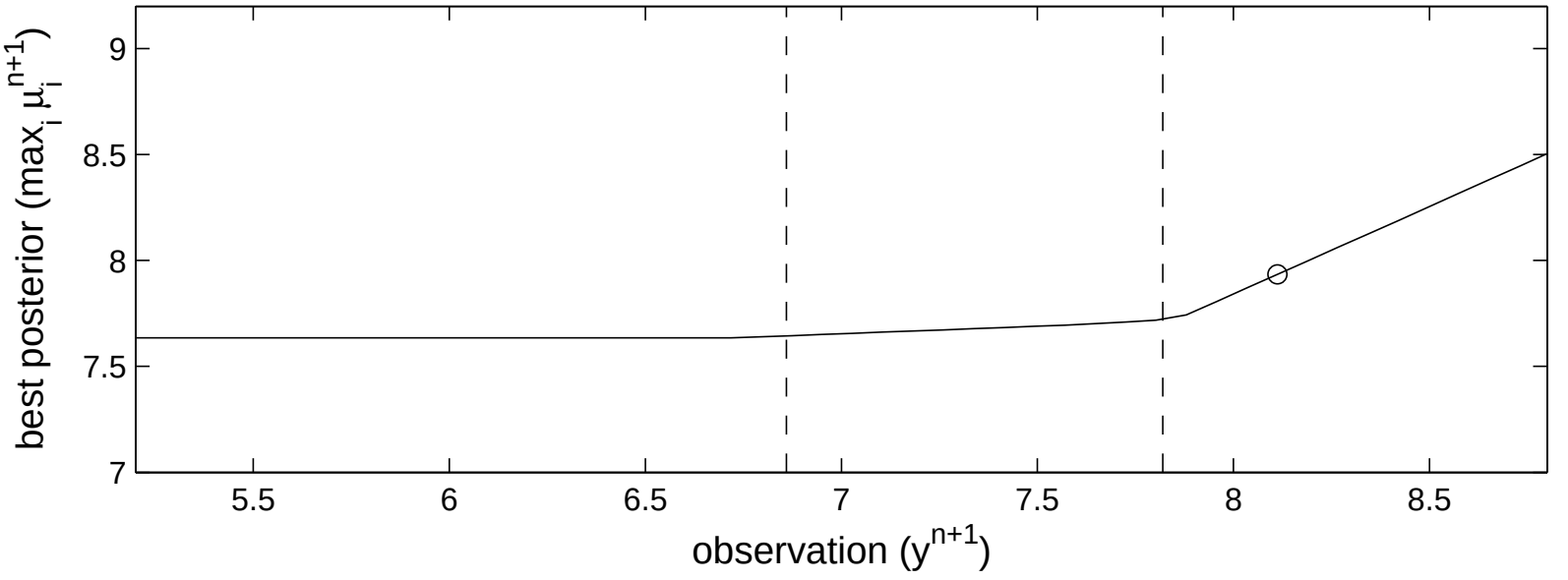
Computing the Knowledge-Gradient



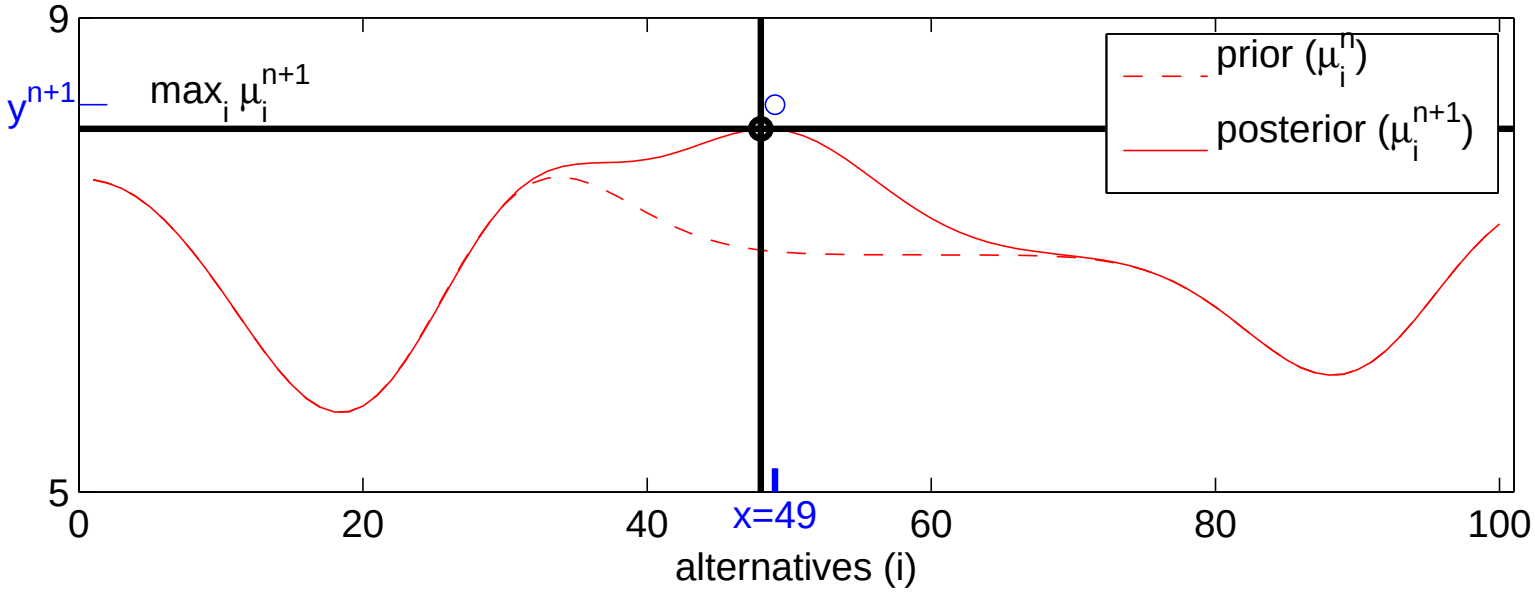


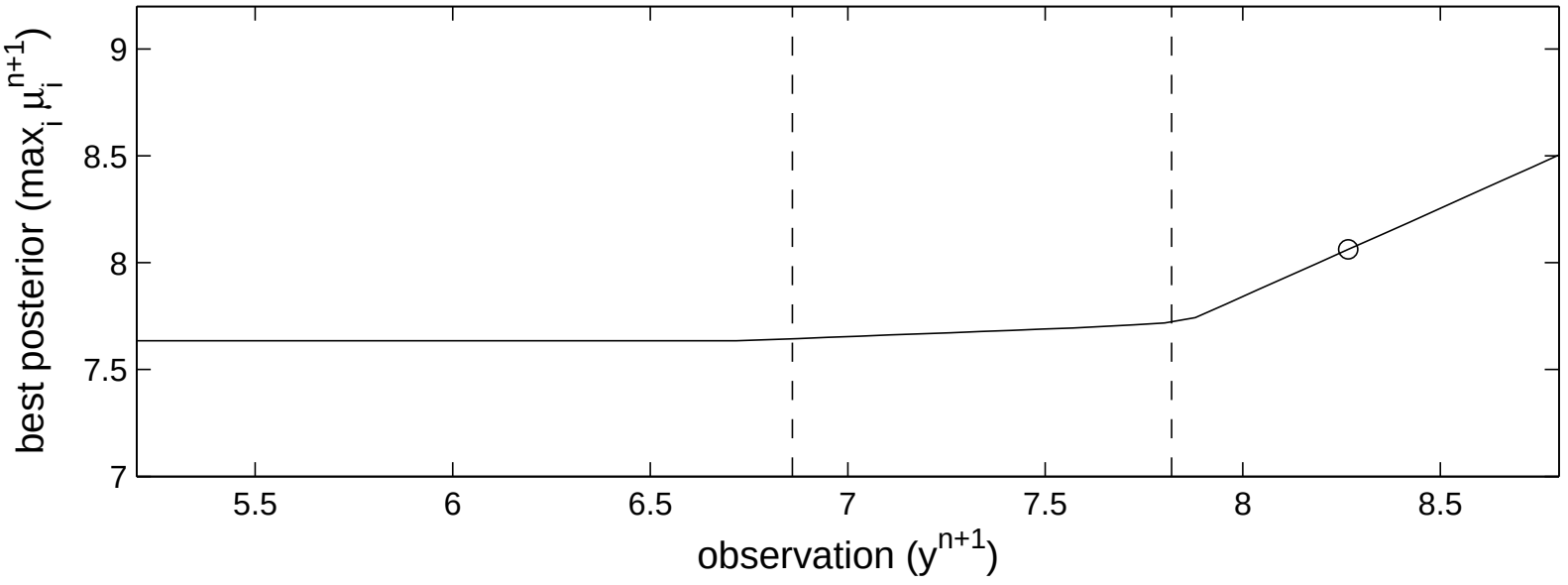
Computing the Knowledge-Gradient



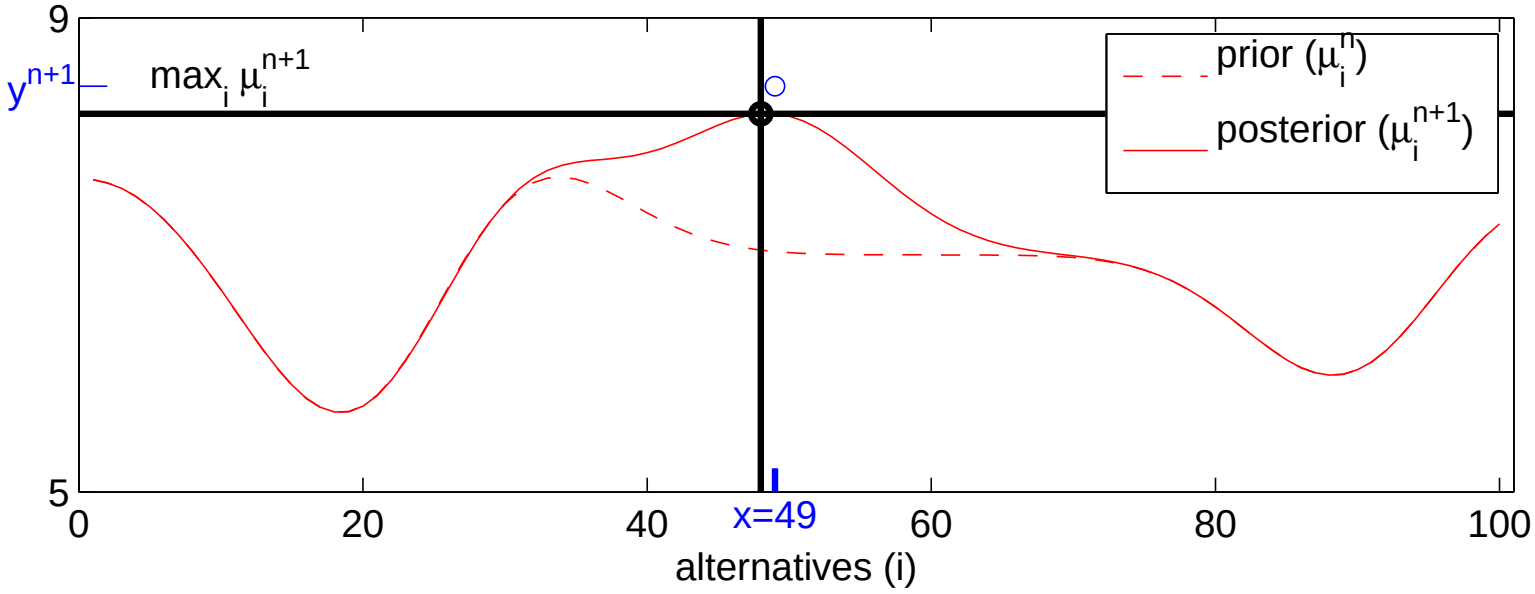


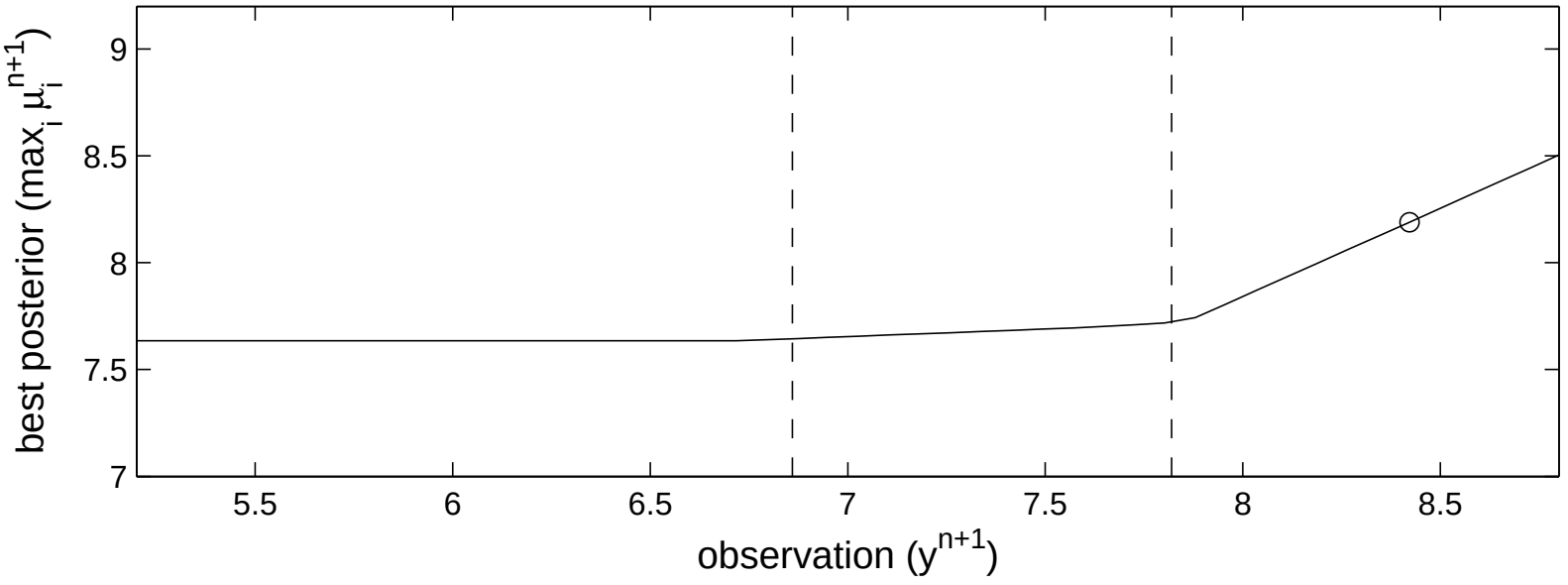
Computing the Knowledge-Gradient



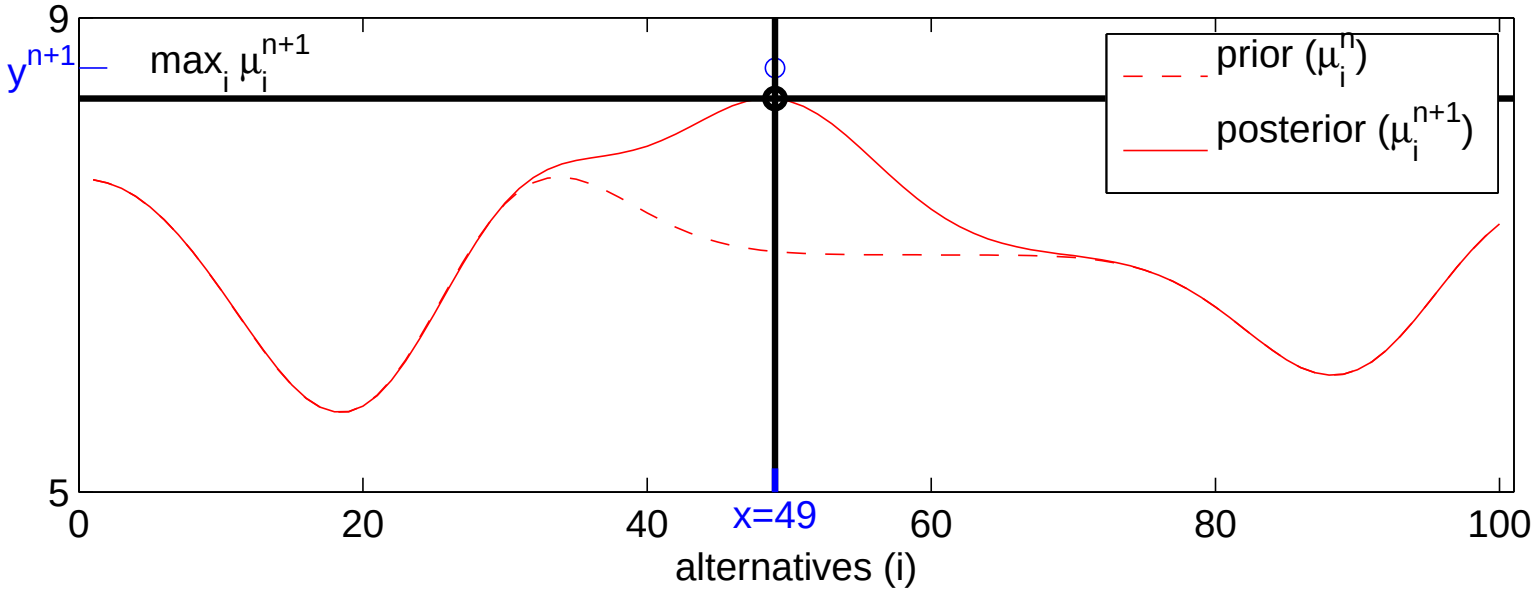


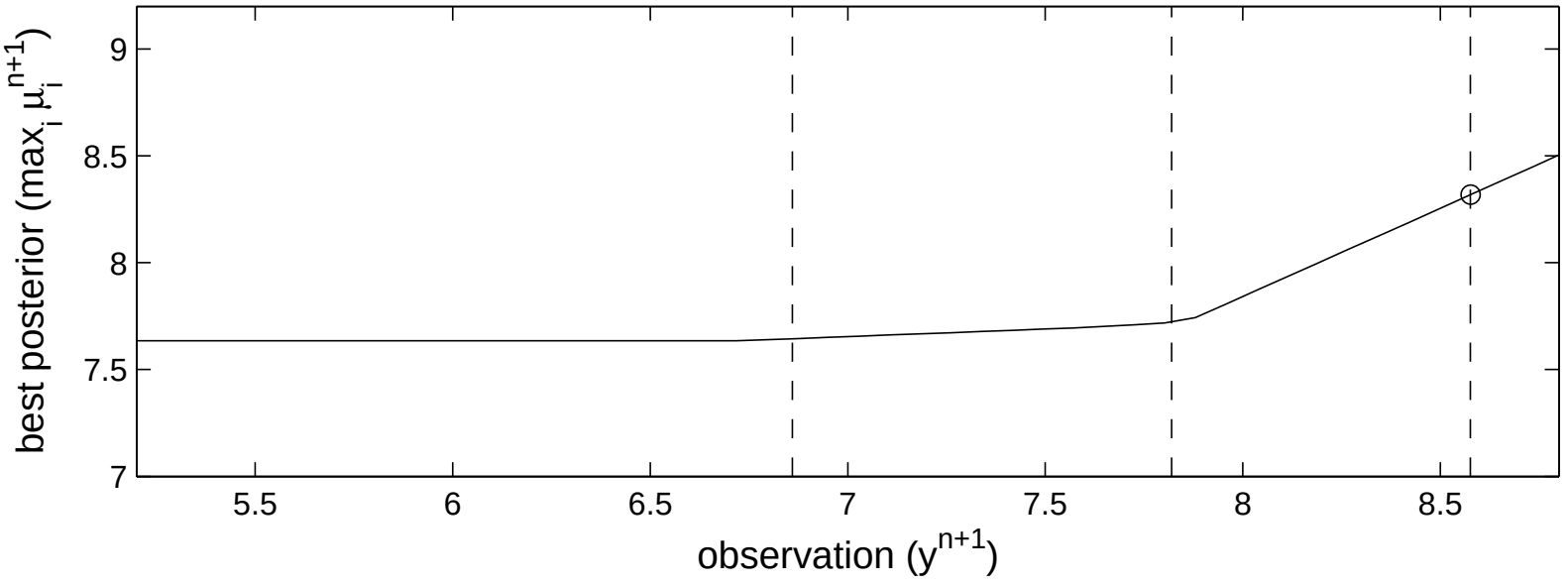
Computing the Knowledge-Gradient





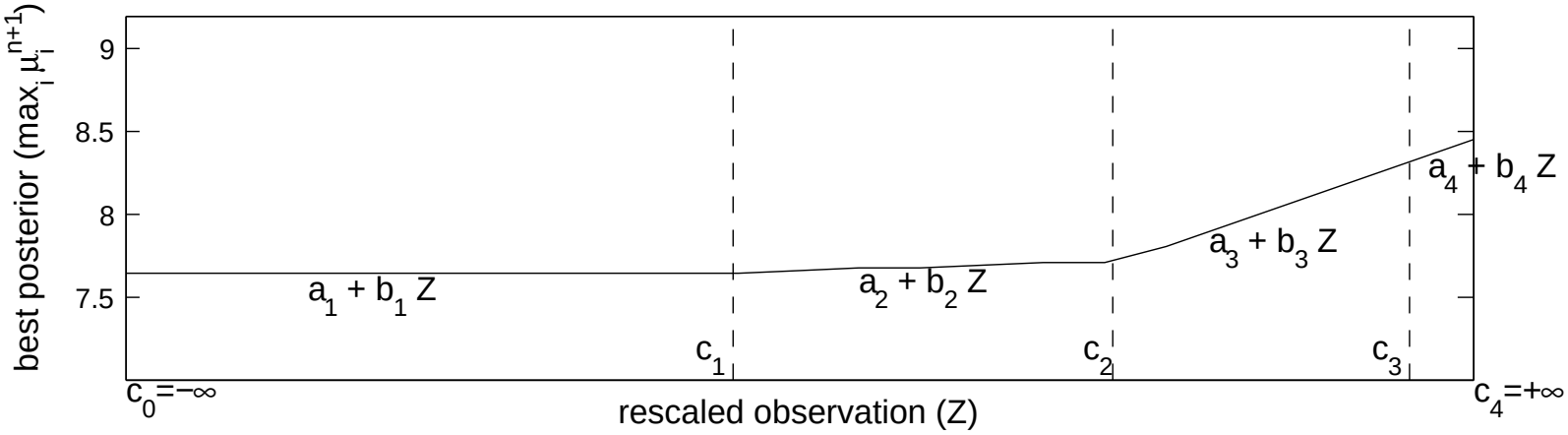
Computing the Knowledge-Gradient





Computing the Knowledge Gradient

Recall that the KG policy chooses the measurement that maximizes the KG factor $\mathbb{E}_n\left[\left(\max_x \mu_x^{n+1}\right)\right] - \max_x \mu_x^n$.
How do we compute this expectation?



We rescale the observation $Z = (y^{n+1} - \mathbb{E}_n[y^{n+1}])/\sqrt{Var_n[y^{n+1}]}$, note that Z is standard normal, and compute the expectation as

$$\mathbb{E}_n\left[\left(\max_x \mu_x^{n+1}\right)\right] = \sum_{j=1}^{4} \mathbb{E}_n\left[(a_j + b_j Z)\mathbf{1}_{\{c_{j-1} \le Z < c_j\}}\right].$$

Computing the Knowledge Gradient

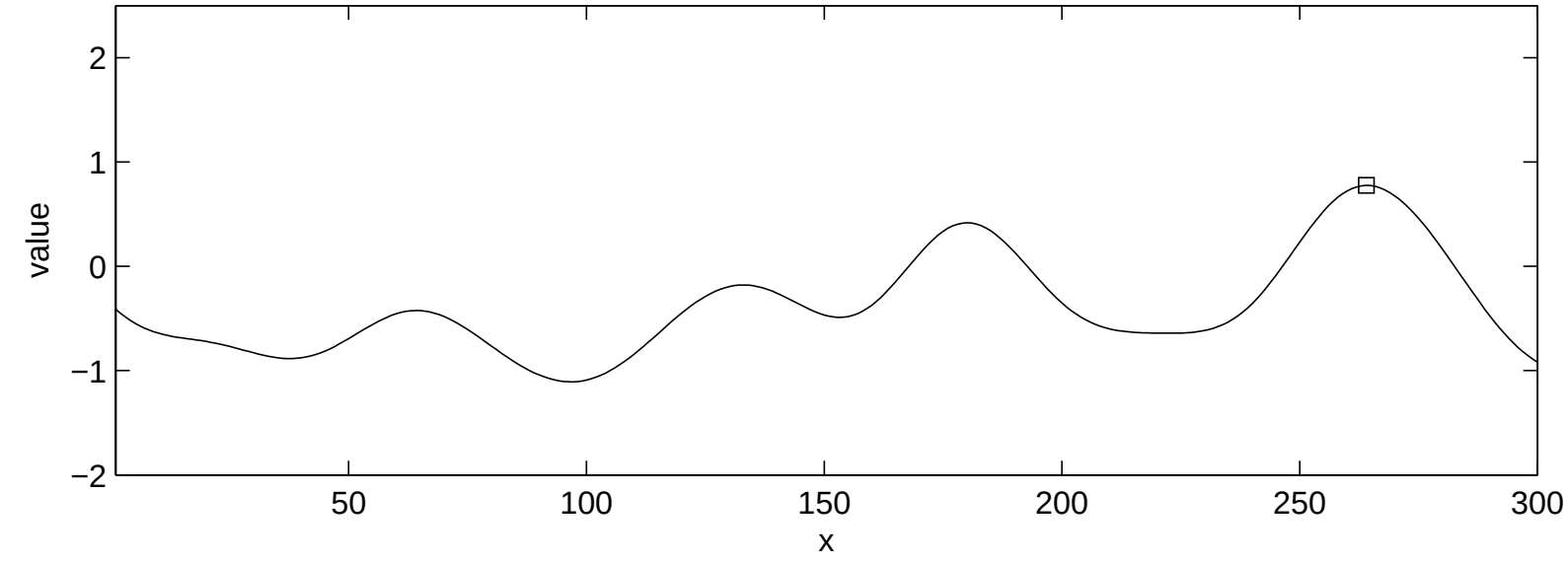
In general, to compute the KG factor for a candidate measurement:

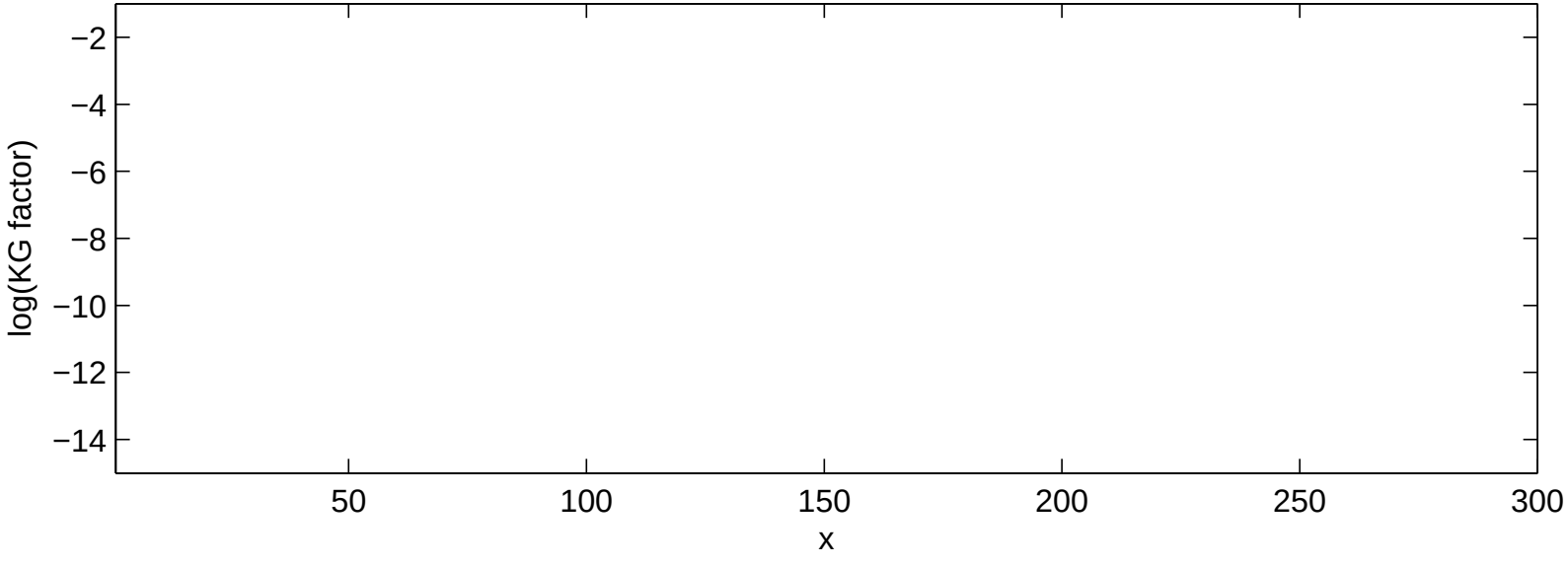
- Let A contain those alternatives that are best under the posterior with nonzero probability.
- Let $[j]$ denote the j^{th} entry in A.
- Let $a_j = \mu^n_{[j]}$ and $b_j = \sqrt{\text{Var}_n\left[\mu^{n+1}_{[j]}\right]}$.
- Sort A in order of increasing b_j.
- The KG factor is

$$\sum_{j=1}^{|A|-1} (b_{i+1} - b_i) f\left(\frac{-|a_{i+1} - a_i|}{b_{i+1} - b_i}\right),$$

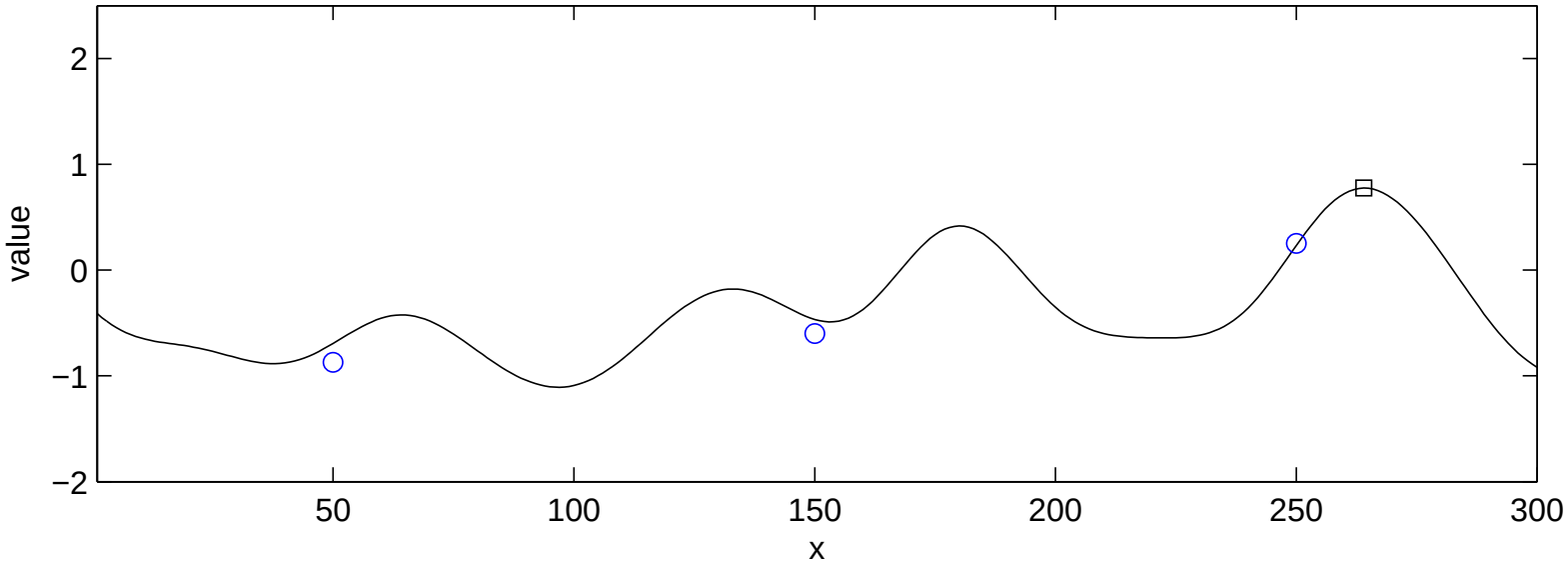
where $f(z) = \varphi(z) + z\Phi(z)$, φ is the normal pdf and Φ is the normal cdf.

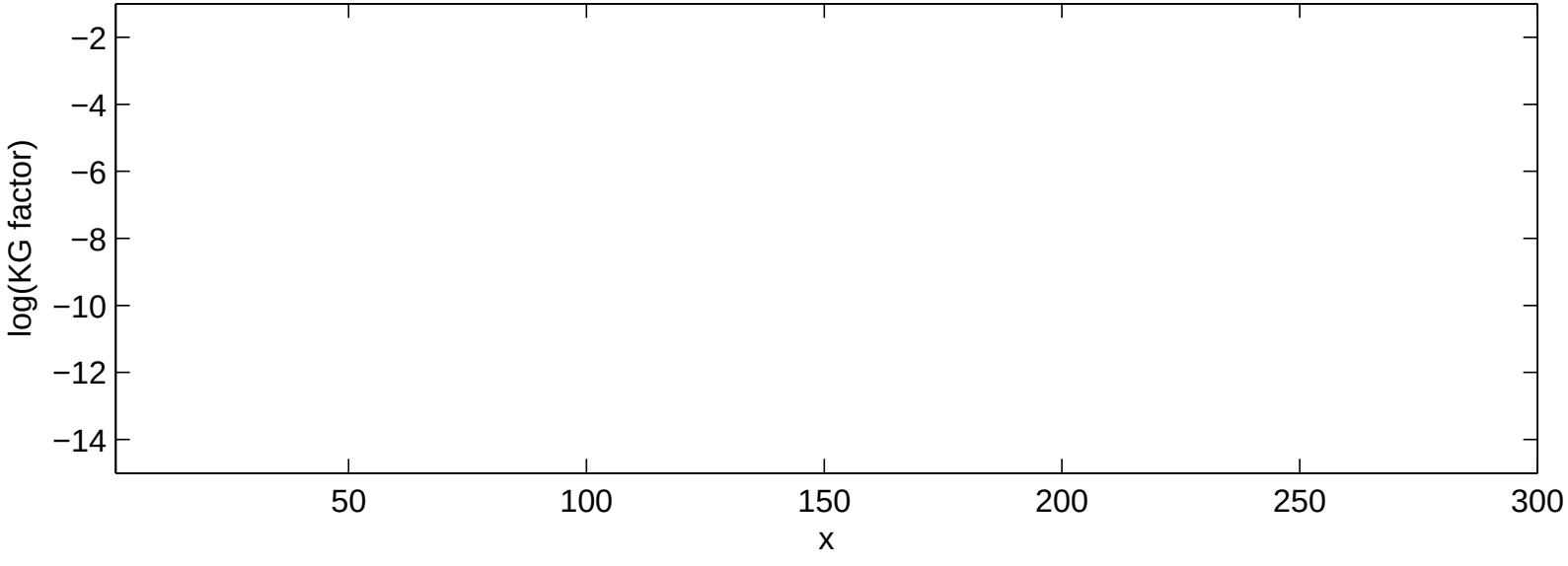
Global Optimization Example



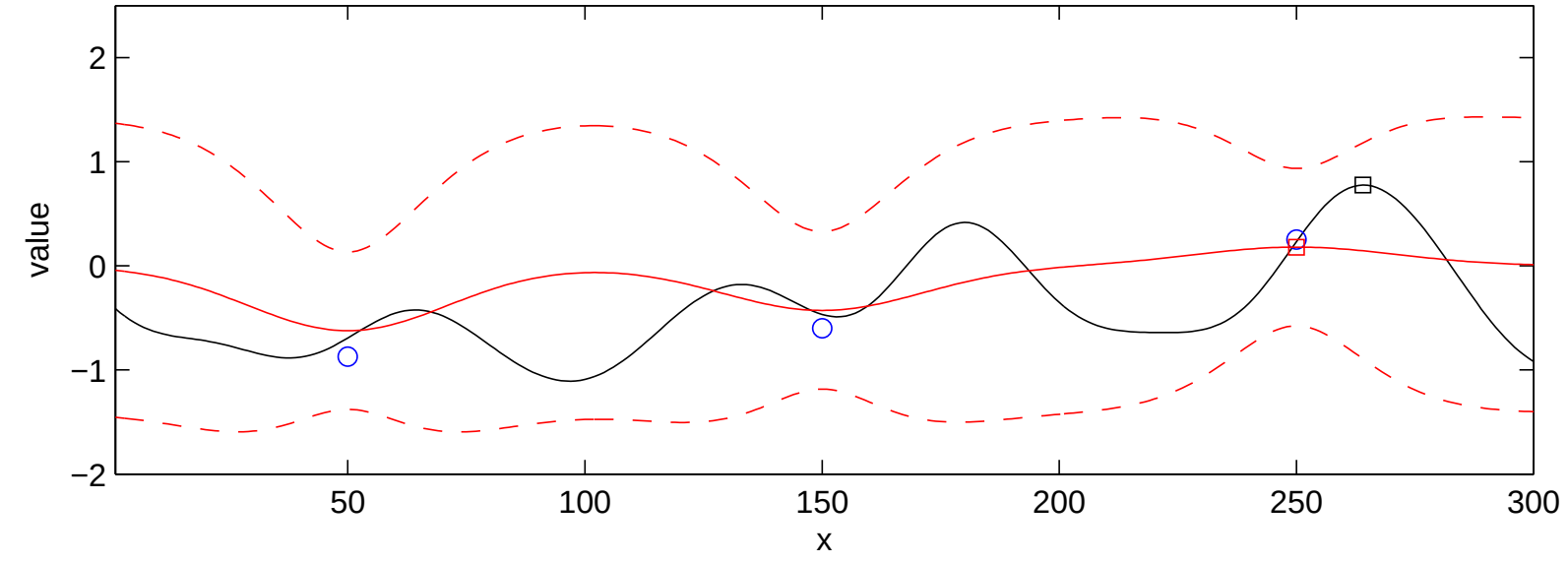


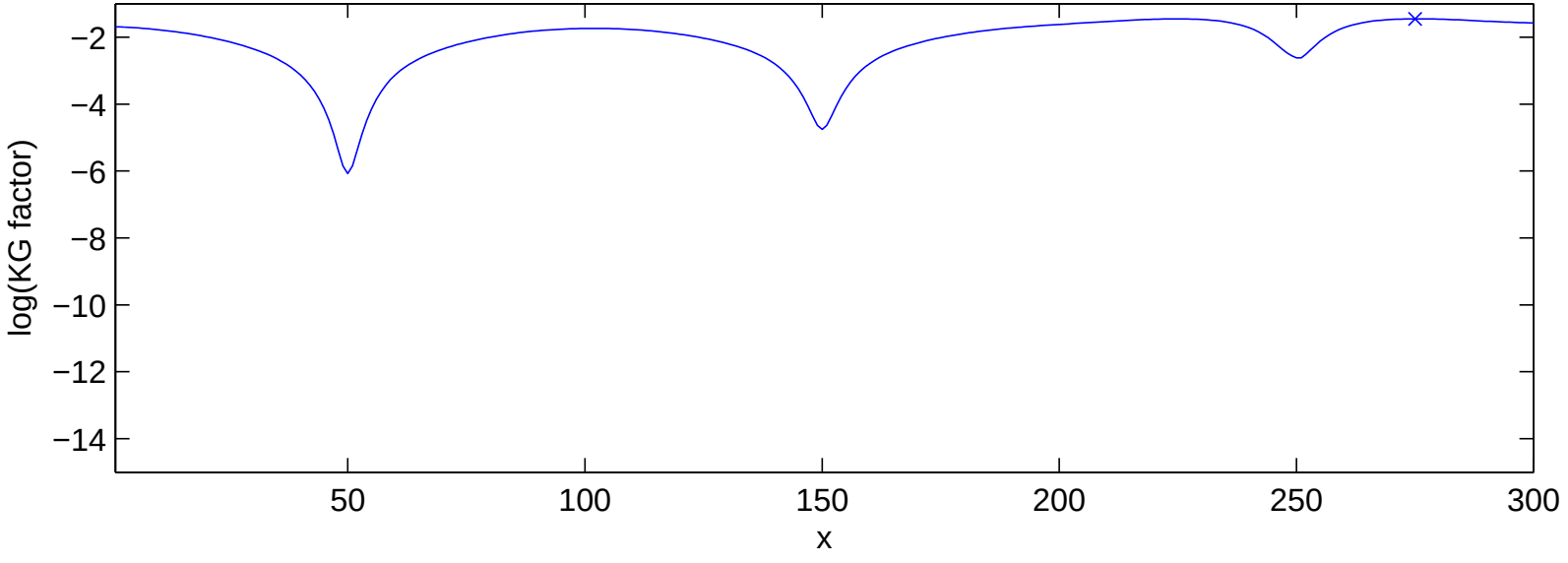
Global Optimization Example



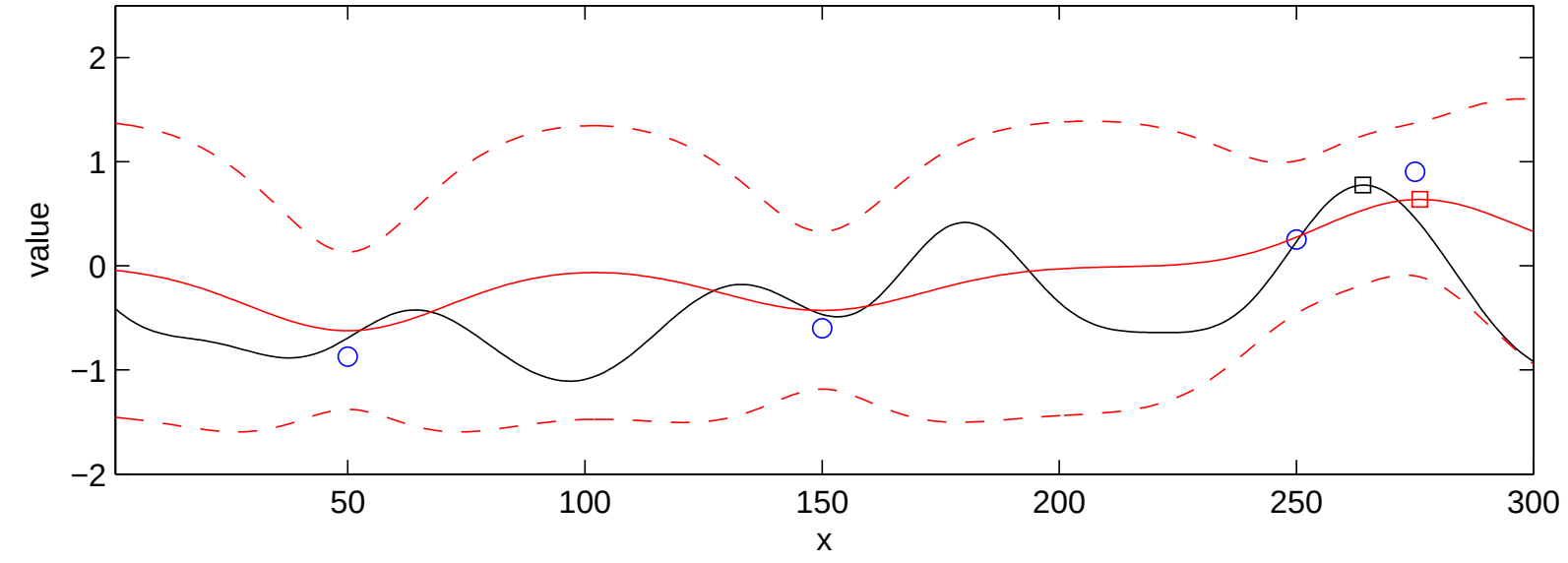


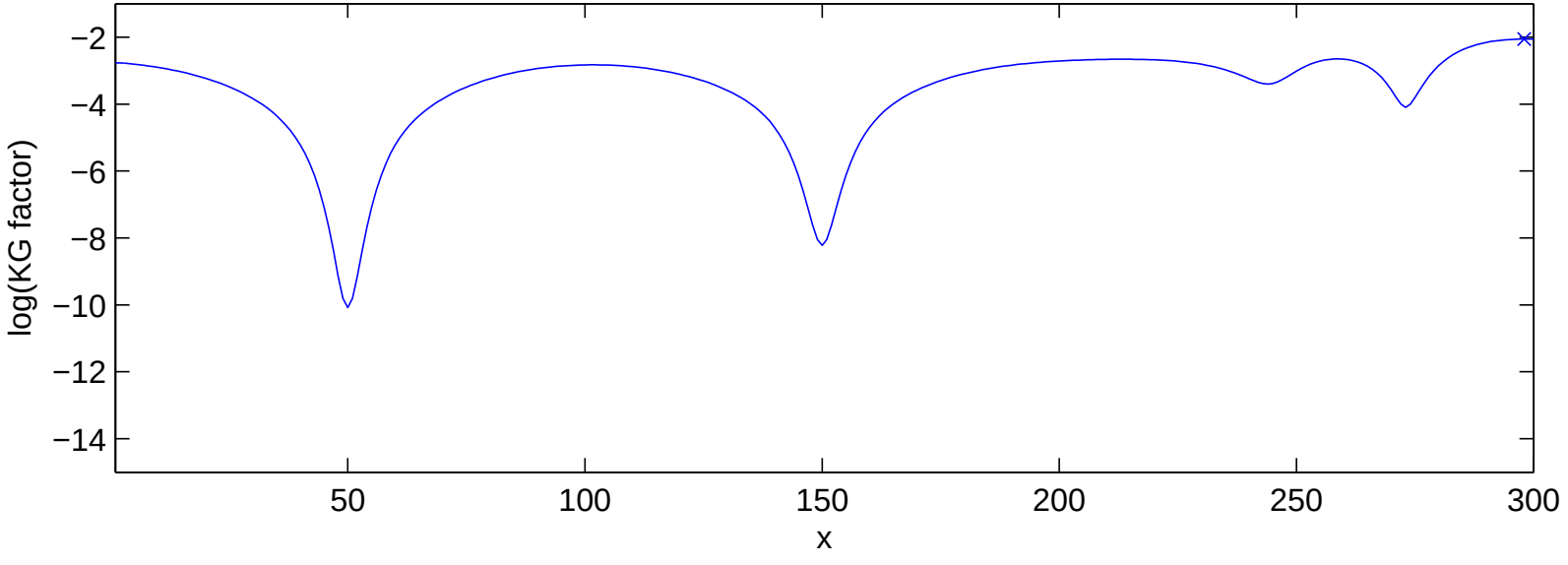
Global Optimization Example



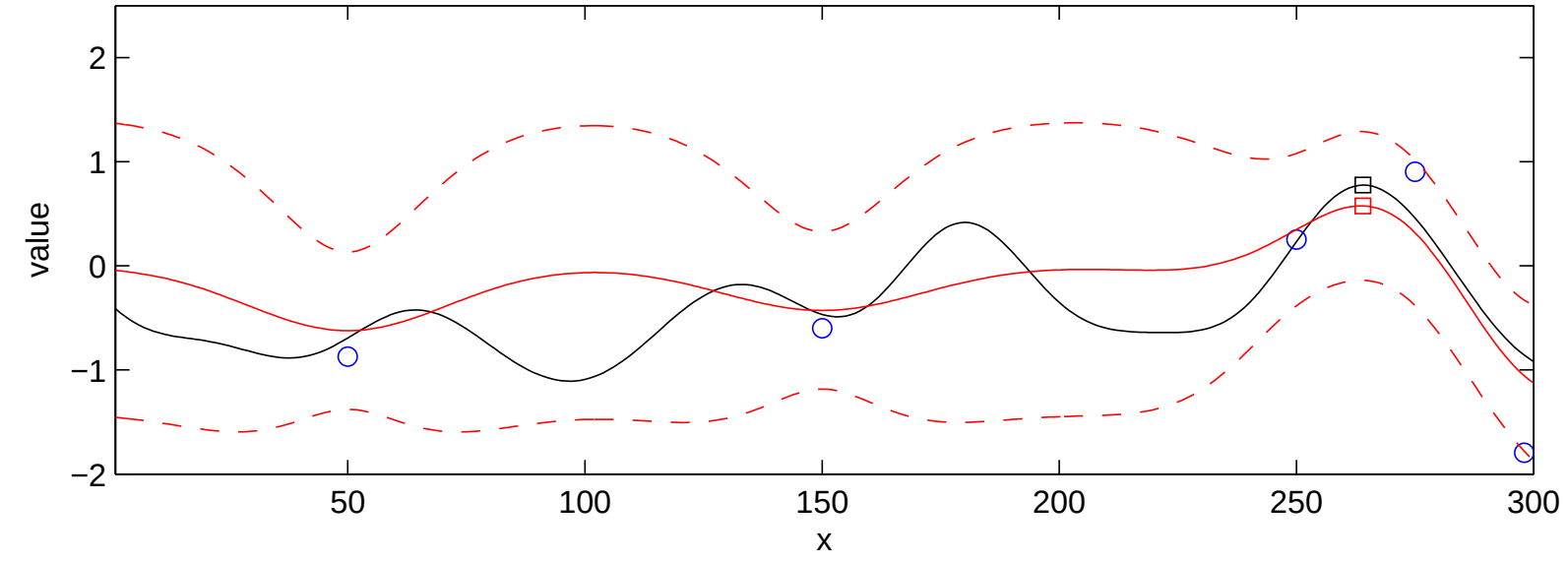


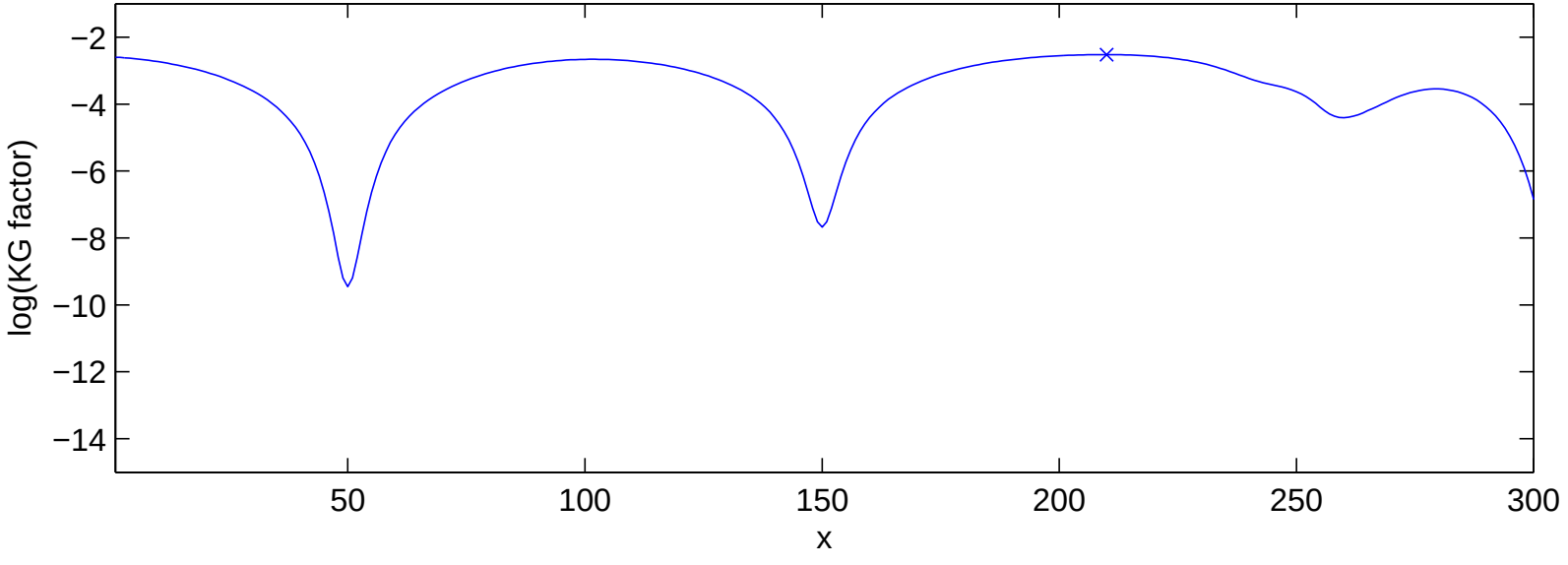
Global Optimization Example



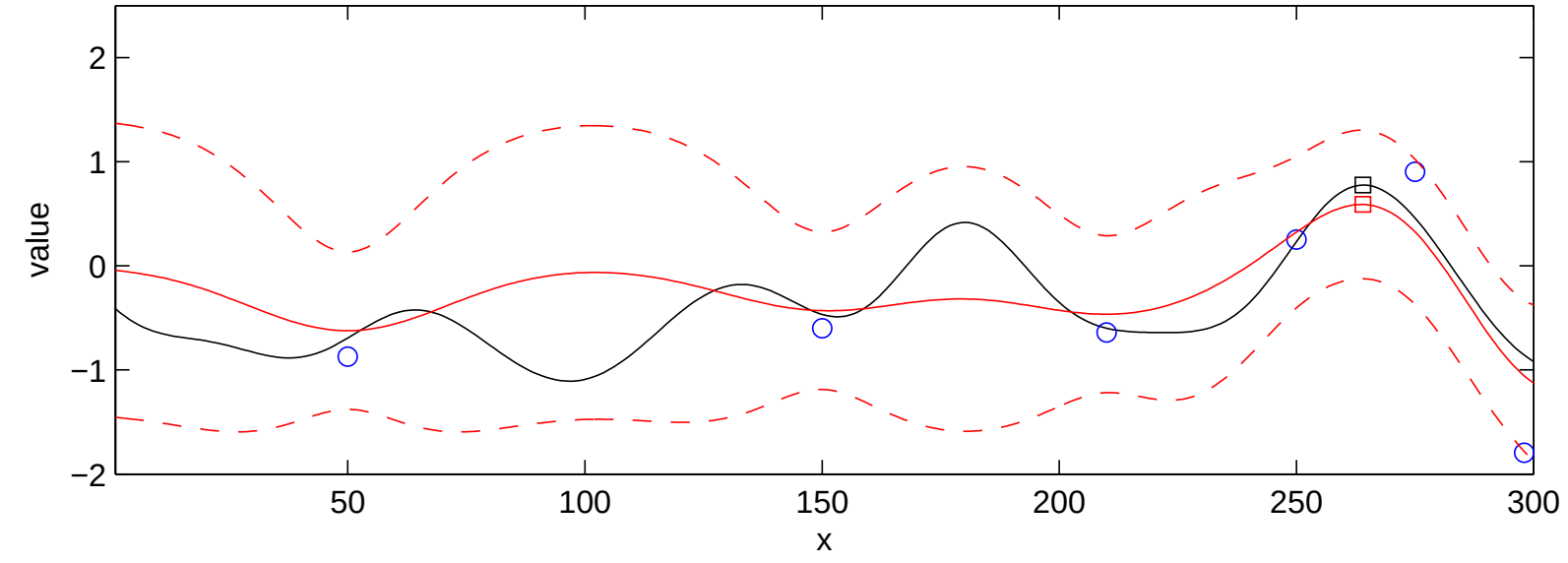


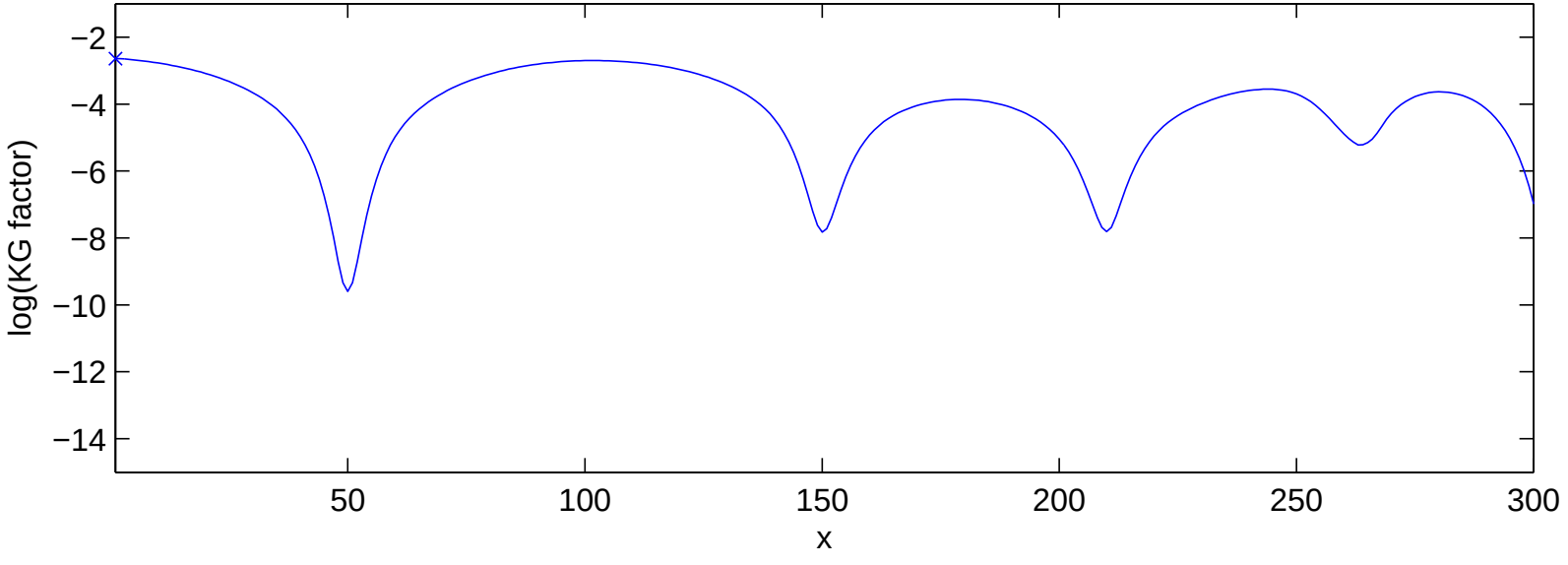
Global Optimization Example



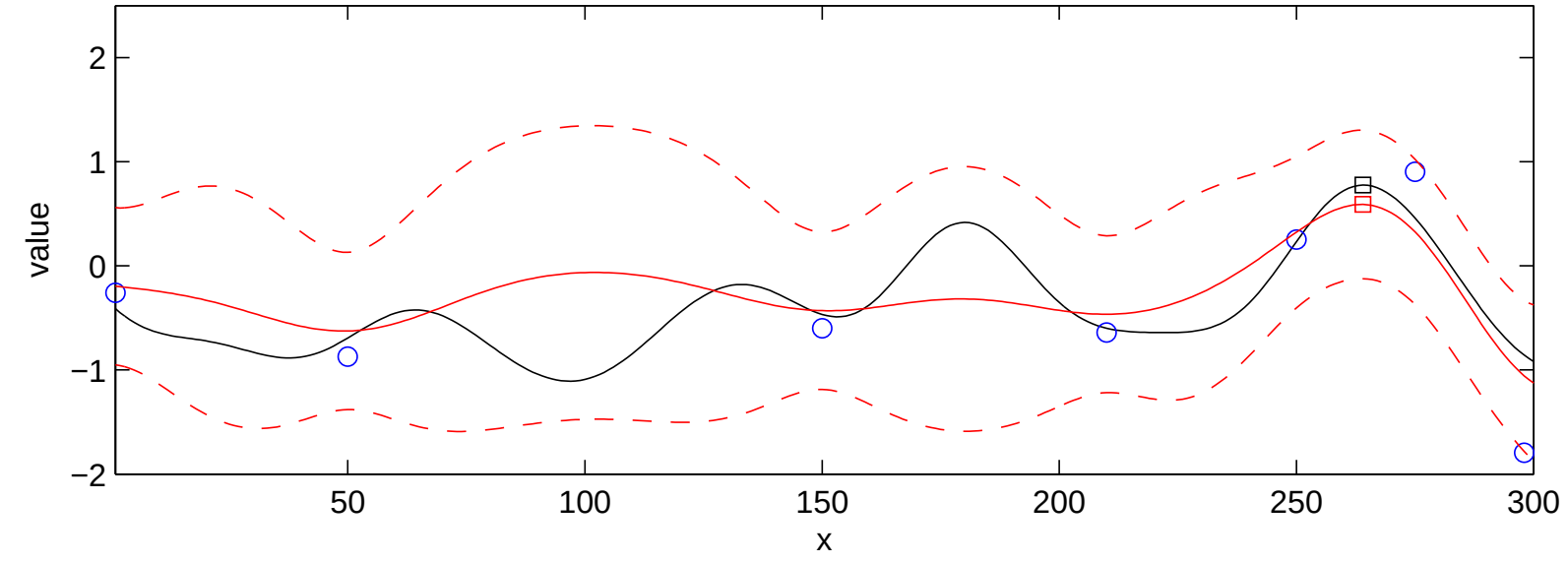


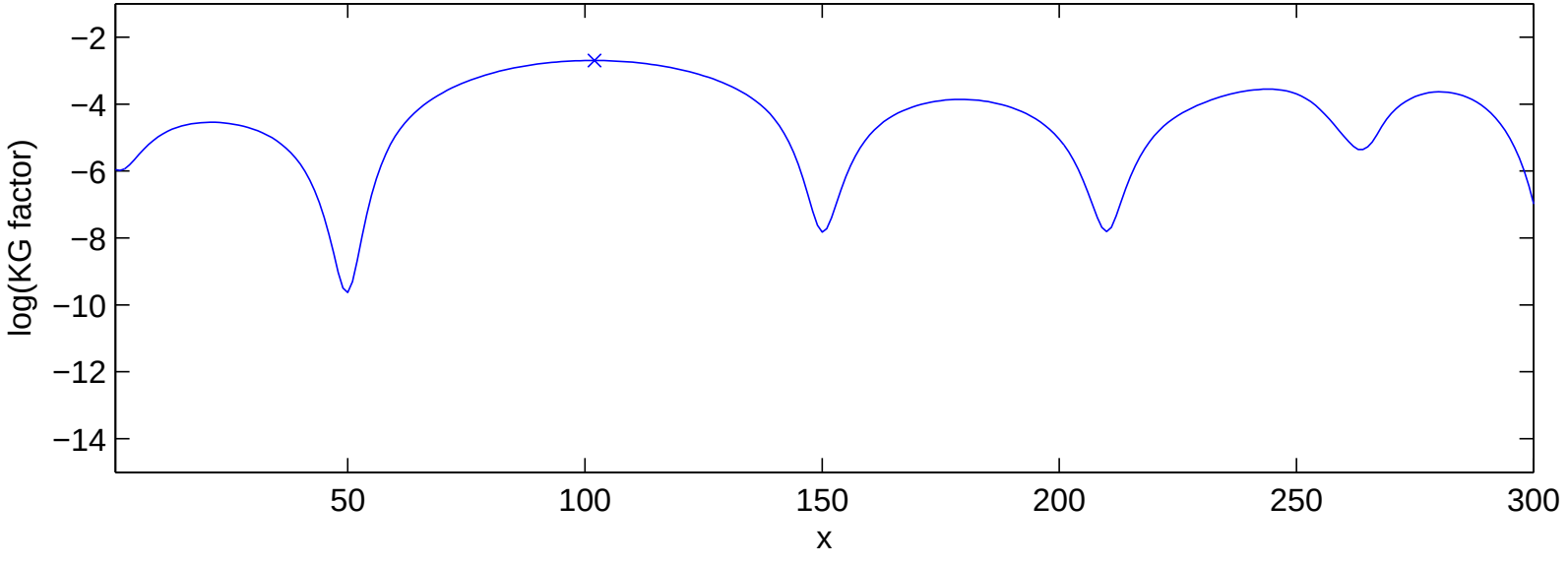
Global Optimization Example



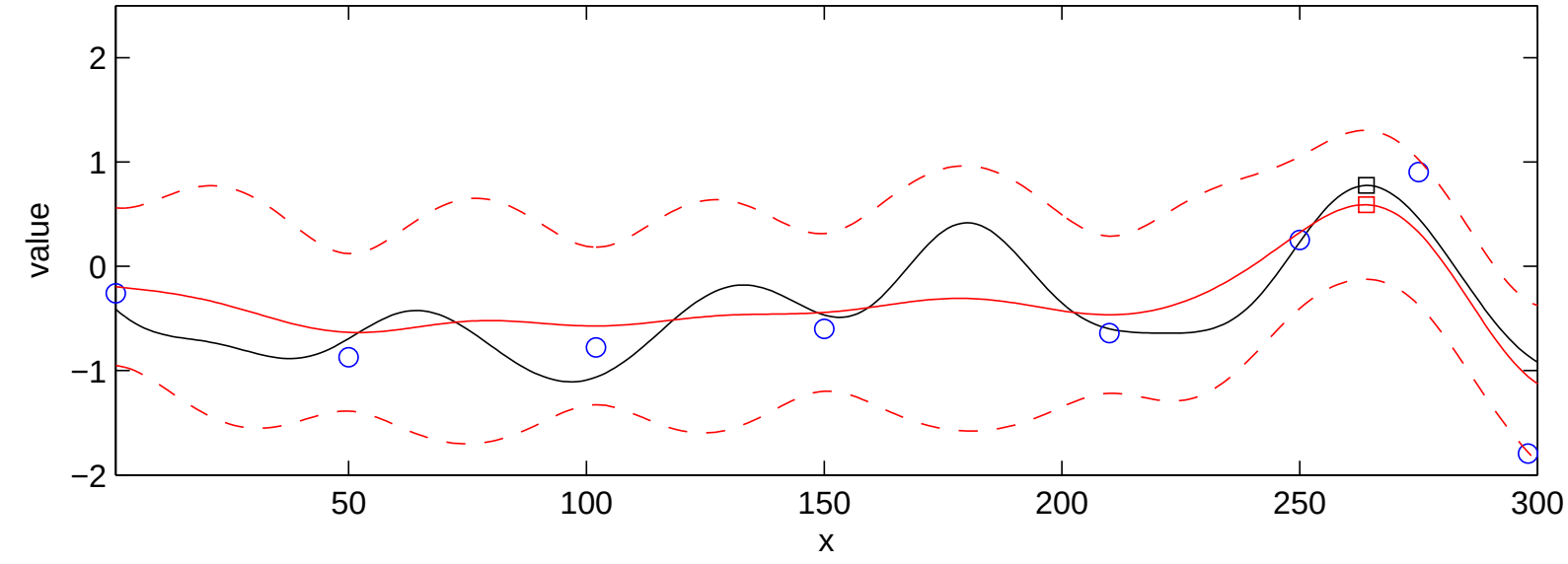


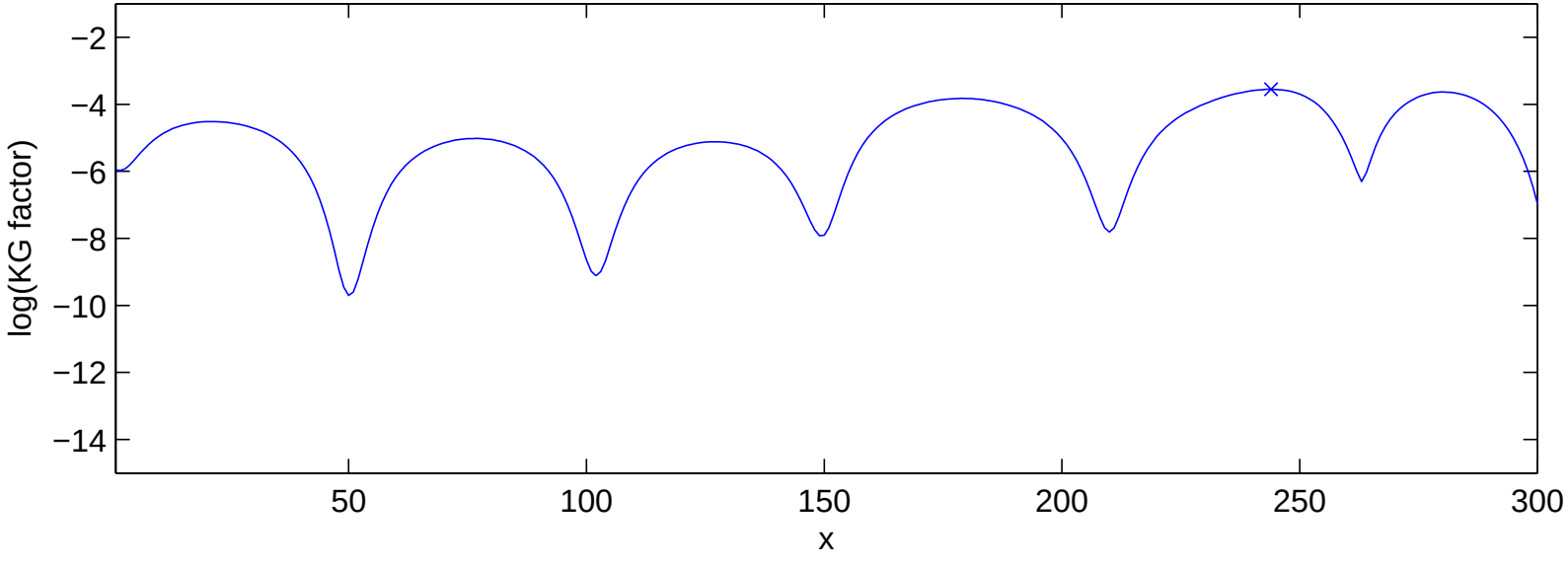
Global Optimization Example



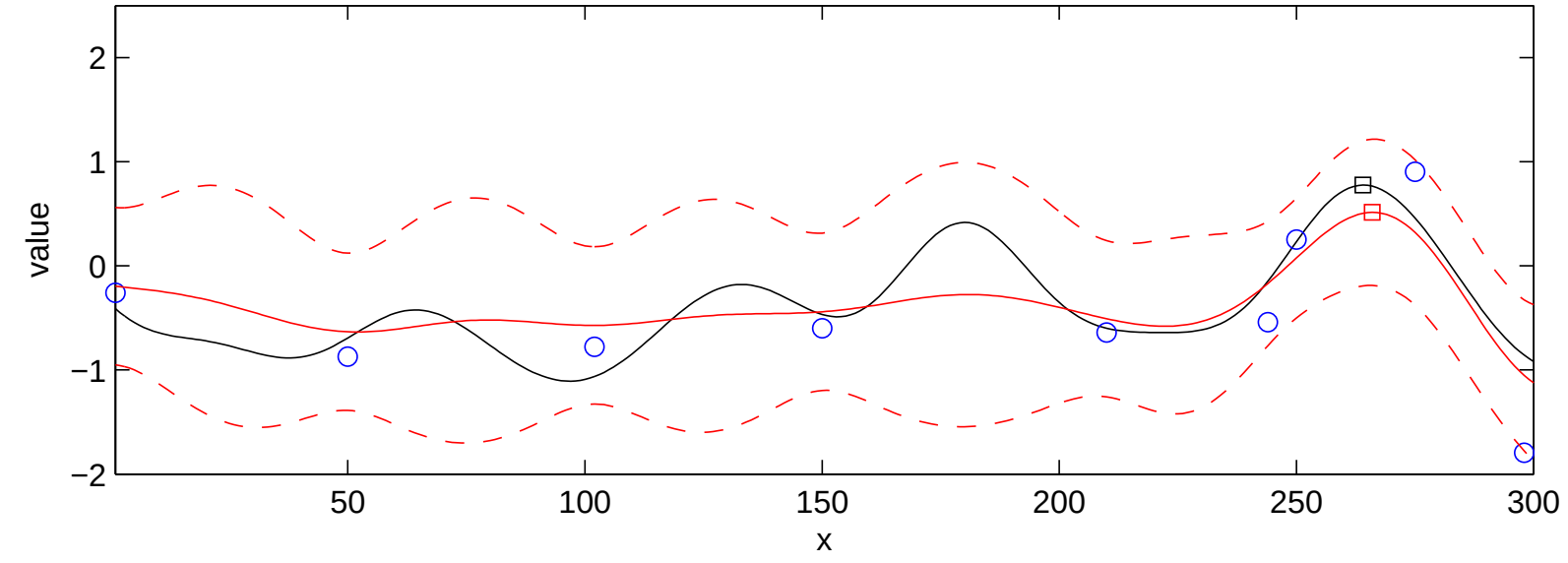


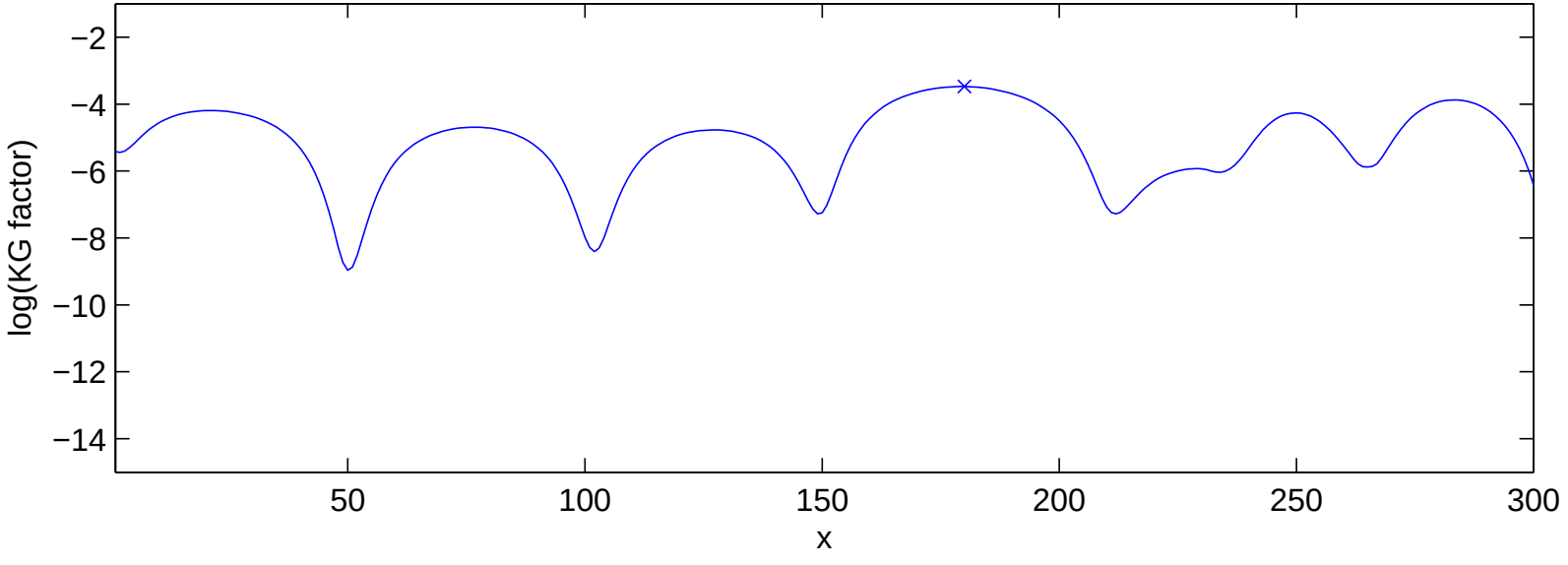
Global Optimization Example



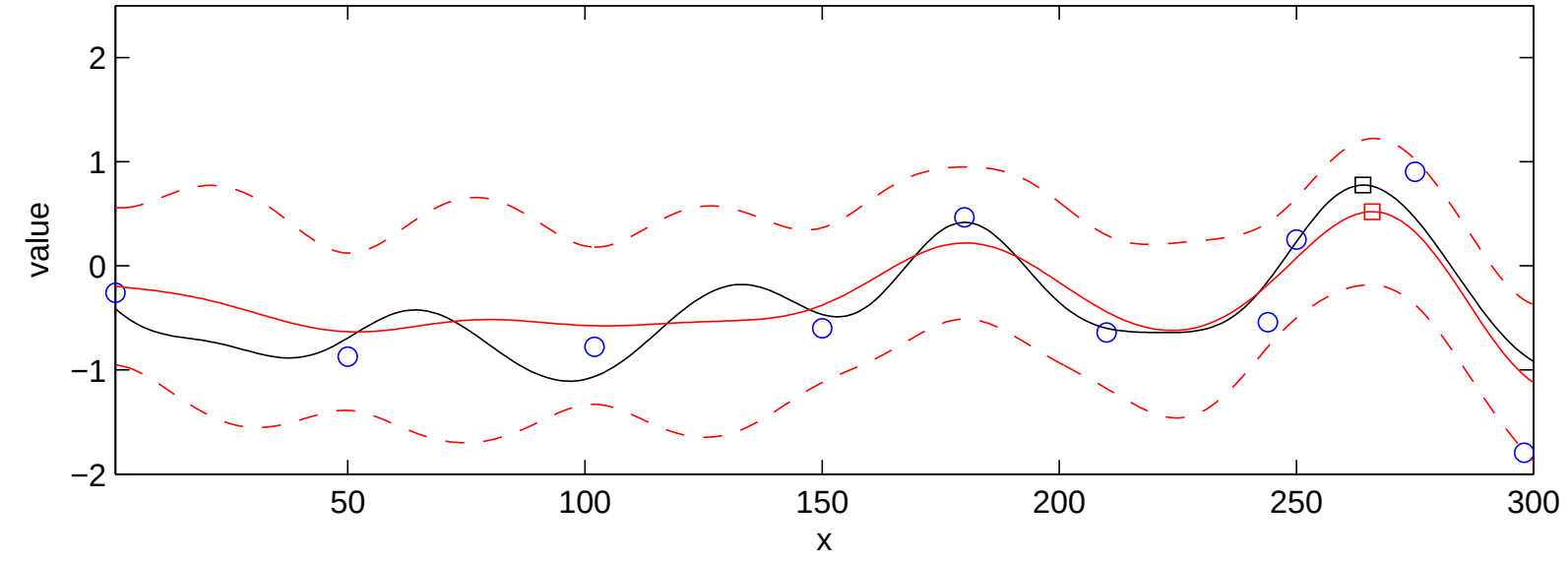


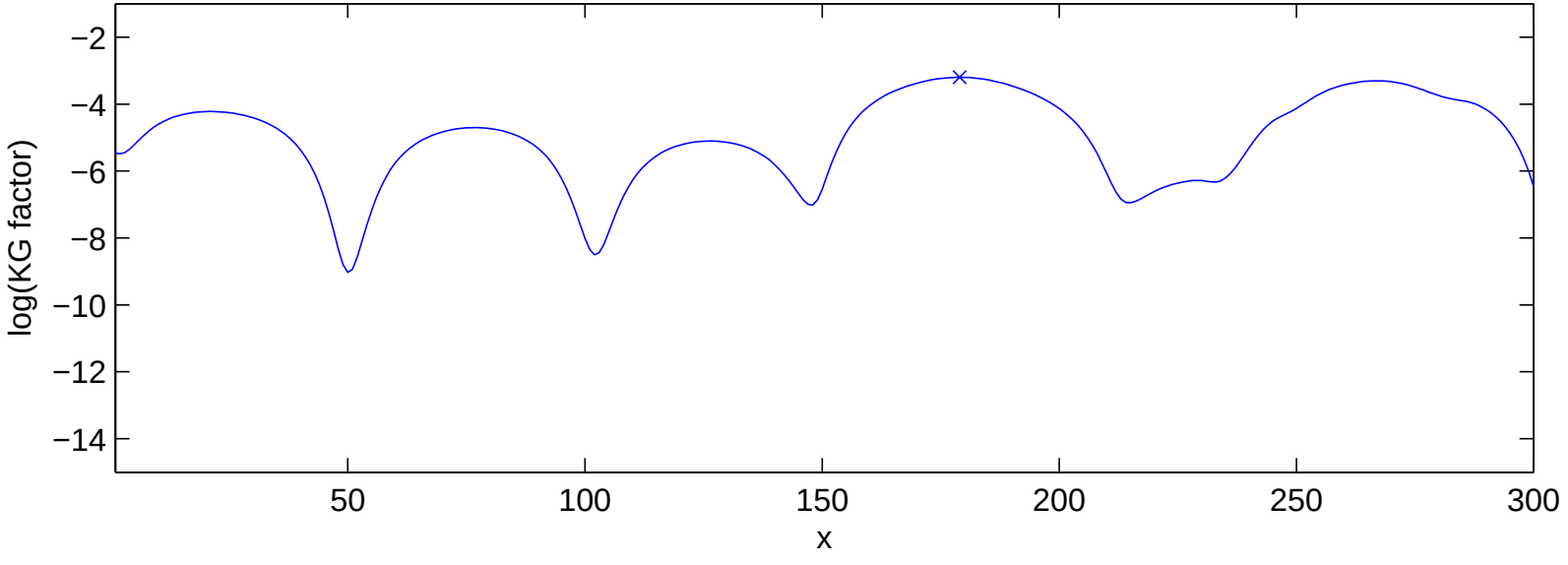
Global Optimization Example



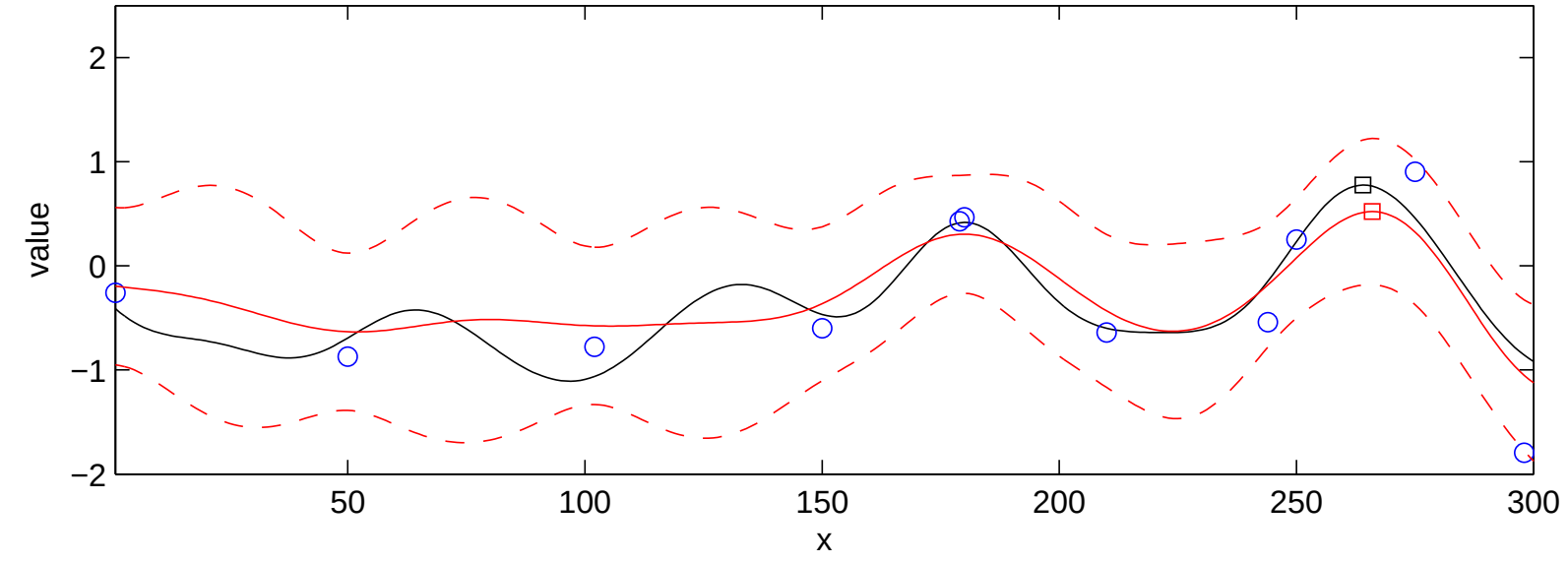


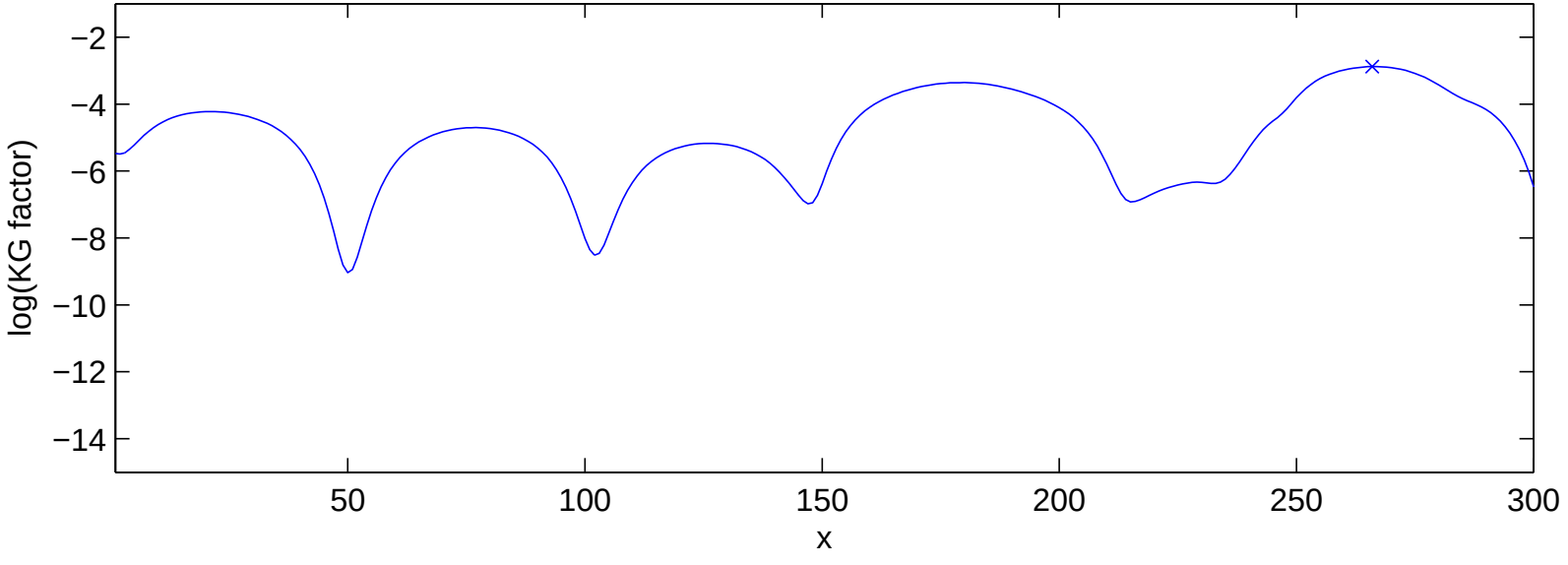
Global Optimization Example



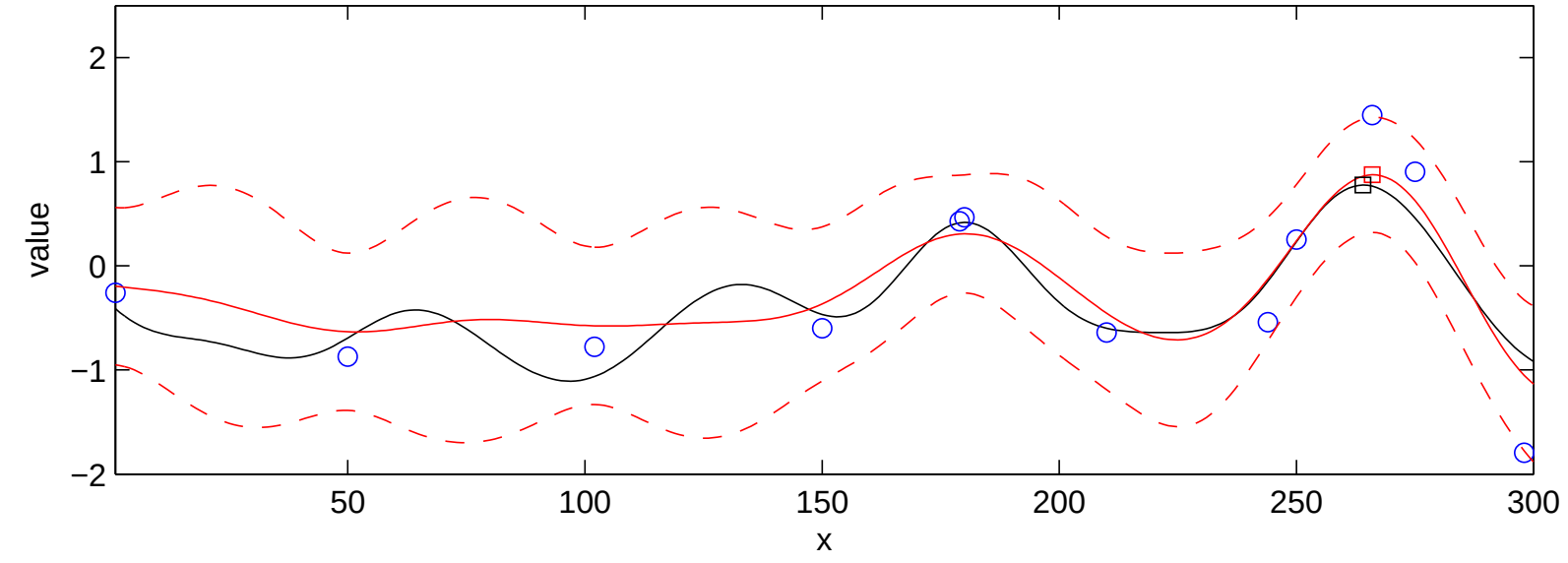


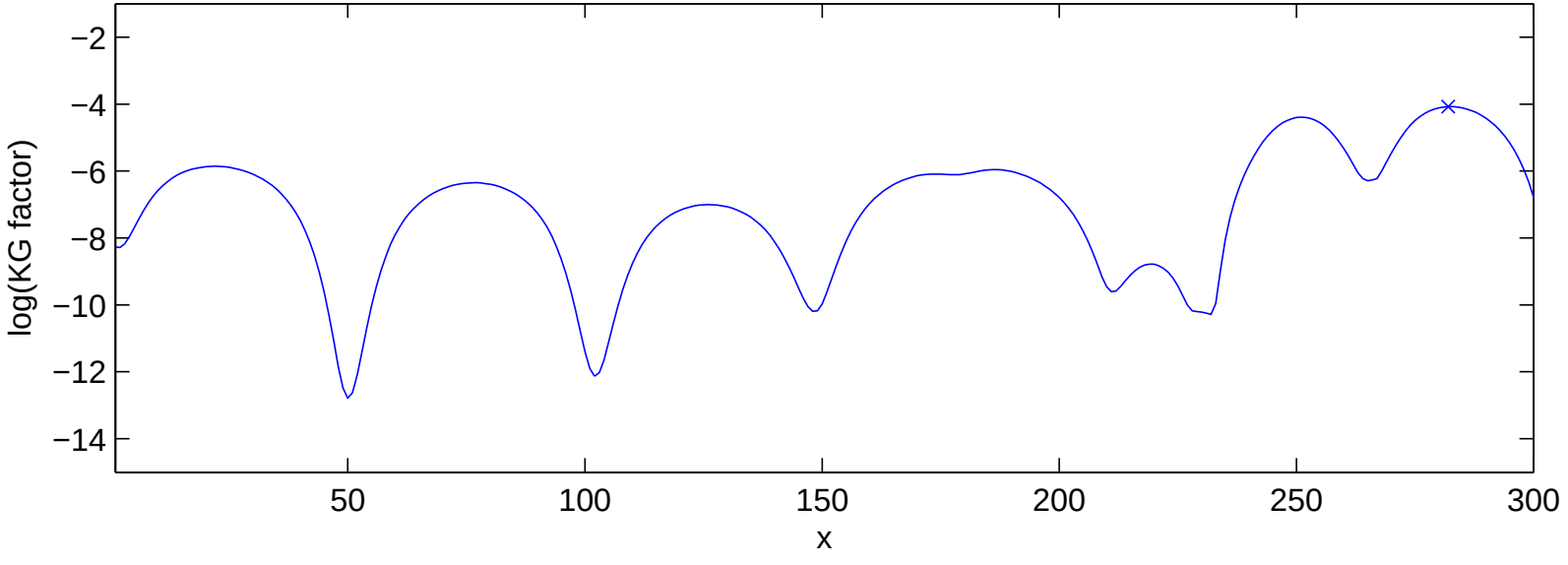
Global Optimization Example



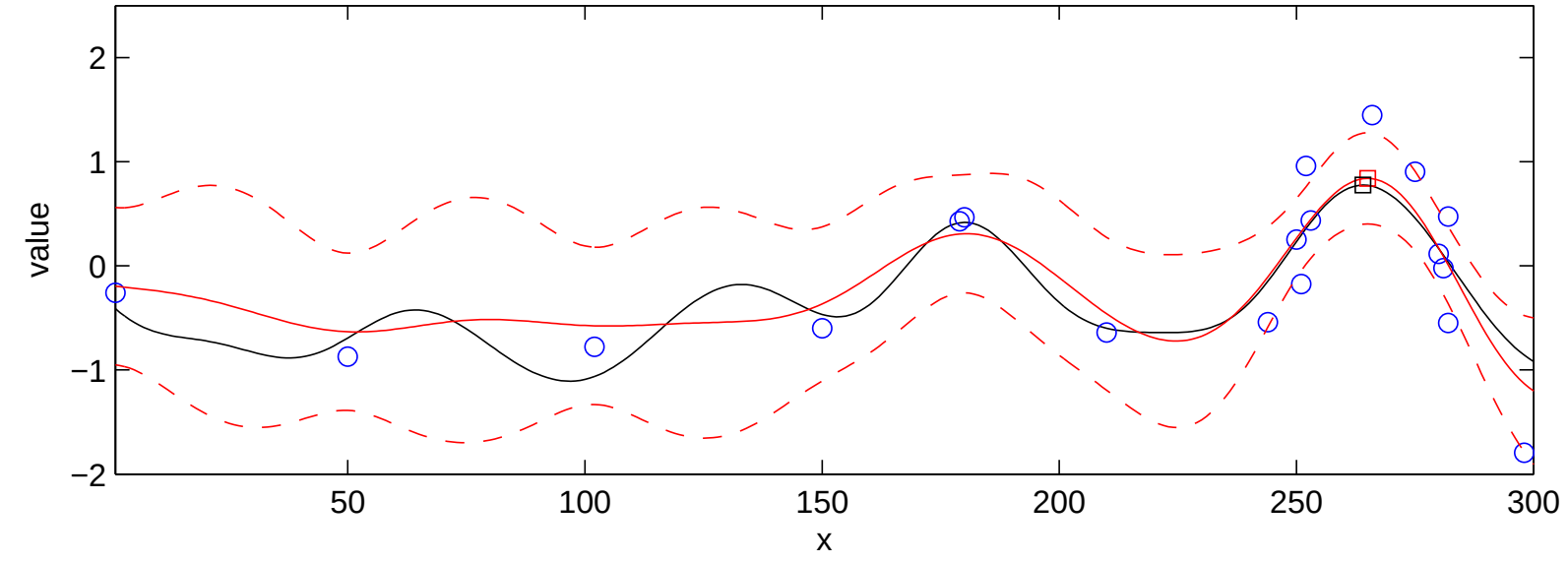


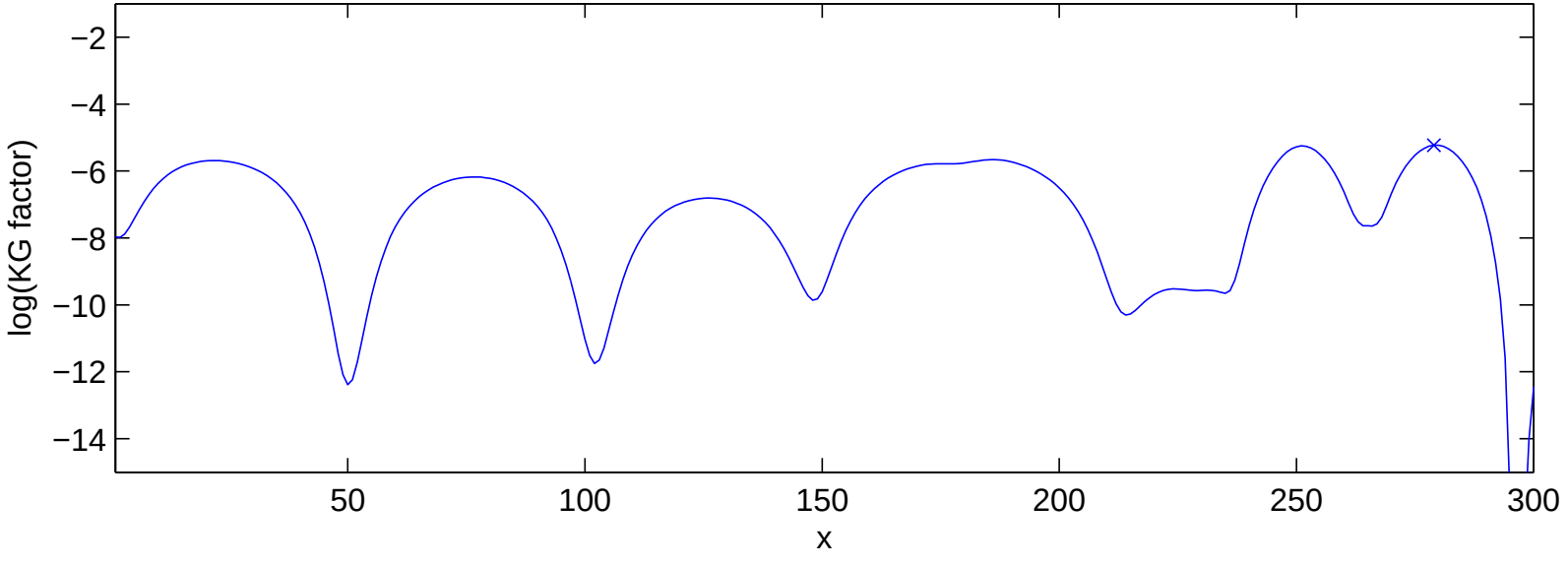
Global Optimization Example





Global Optimization Example





KG on Continuous Domains

- Instead of finitely many alternatives $\{1, \ldots, k\}$, suppose we have a continuous domain $[0,1]^d$.
- Then θ is a function $\theta : [0,1]^d \mapsto \mathbb{R}$.
- We may take a Gaussian process prior $\mathbb{P}_0$ on θ.
- Then μ^n is a function $\mu^n(x) = \mathbb{E}_n[\theta(x)]$.
- The KG factor is

$$\mathrm{KG}(x) = \mathbb{E}_n \left[\sup_{x' \in [0,1]^d} \mu^{n+1}(x') \mid x_{n+1} = x \right] - \sup_{x' \in [0,1]^d} \mu^n(x').$$

- The KG policy chooses its measurements

$$x_{n+1} \in \arg\max_x \mathrm{KG}(x).$$

KG on Continuous Domains

- Recall that the KG factor is

$$\mathrm{KG}(x) = \mathbb{E}_n \left[\sup_{x' \in [0,1]^d} \mu^{n+1}(x') \mid x_{n+1} = x \right] - \sup_{x' \in [0,1]^d} \mu^n(x').$$

- The KG factor is approximated by

$$\overline{\mathrm{KG}}(x) = \mathbb{E}_n \left[\max_{x' \in A} \mu^{n+1}(x') \mid x_{n+1} = x \right] - \max_{x' \in A} \mu^n(x'),$$

  where A is a judiciously chosen set.
- As A gets bigger, the approximation improves.
- We may compute the gradient $\nabla_x \overline{\mathrm{KG}}(x)$ and use it in a first order method (e.g., multistart gradient ascent) to approximate

$$\arg\max_{x \in [0,1]^d} \overline{\mathrm{KG}}(x).$$

- This provides an approximation to the continuous KG policy.

Other Approaches to Continuous Global Optimization

- Many other derivative-free noise-tolerant global optimization methods exist, e.g.,
  - pattern search, e.g., Nelder-Mead
  - stochastic approximation, e.g., SPSA [Spall 1992].
  - evolutionary algorithms, simulated annealing, tabu search
- The KG method is a Bayesian global optimization (BGO) method because it places a Bayesian prior distribution on the underlying but unknown function. (See [Jones et al. 1998]).
- The posterior may understood as a surrogate model for the unknown function, and in this way BGO methods are similar to radial basis function methods [Gutmann 2001, Regis & Shoemaker 2007] and response surface methods [Myers & Montgomery 2002].
- Surrogate-based methods require more computation to decide where to evaluate next, but often require fewer evaluations to find global extrema.

Outline

Bayesian R&S as a Stochastic Optimization Problem

- Our Bayesian R&S problem is a stochastic optimization problem:

$$\sup_{\pi} \mathbb{E}^{\pi} \left[\max_{i} \mu_i^N \right].$$

- The optimal solution may be computed using through dynamic programming using Bellman's recursion: $V_N(\mathbb{P}_N) = \max_i \mu_i^N$, $V_n(\mathbb{P}_n) = \max_{x_n} \mathbb{E}_n [V_{n+1}(\mathbb{P}_{n+1})]$.
- ...but the space of possible posteriors is so large that computation is impossible except for the smallest problems.

Optimality when $N = 1$

Remark

The knowledge-gradient policy is optimal when $N = 1$.

This follows by the construction of the KG policy.

Optimality Results

Theorem

$\lim_{N\to\infty} \arg\max_i \mu_i^N = \arg\max_i \theta_i$ *almost surely.*

This states that the global optimum is eventually discovered. The proof uses a martingale convergence argument to show that every alternative is sampled infinitely often.

Corollary

$\lim_{N\to\infty} V^0(S^0; N) - V^{KG,0}(S^0; N) = 0.$

This may be understood as a weak form of asymptotic optimality, although many policies that perform poorly also possess this property, e.g., equal allocation.

Independent Normal Optimality Results

Suppose that Σ^0 is diagonal. Then:

1. If there are exactly 2 alternatives (k=2), the knowledge-gradient policy is optimal. In this case, the optimal policy reduces to

$$x_{n+1} = \arg\max_x \Sigma^n_{xx}.$$

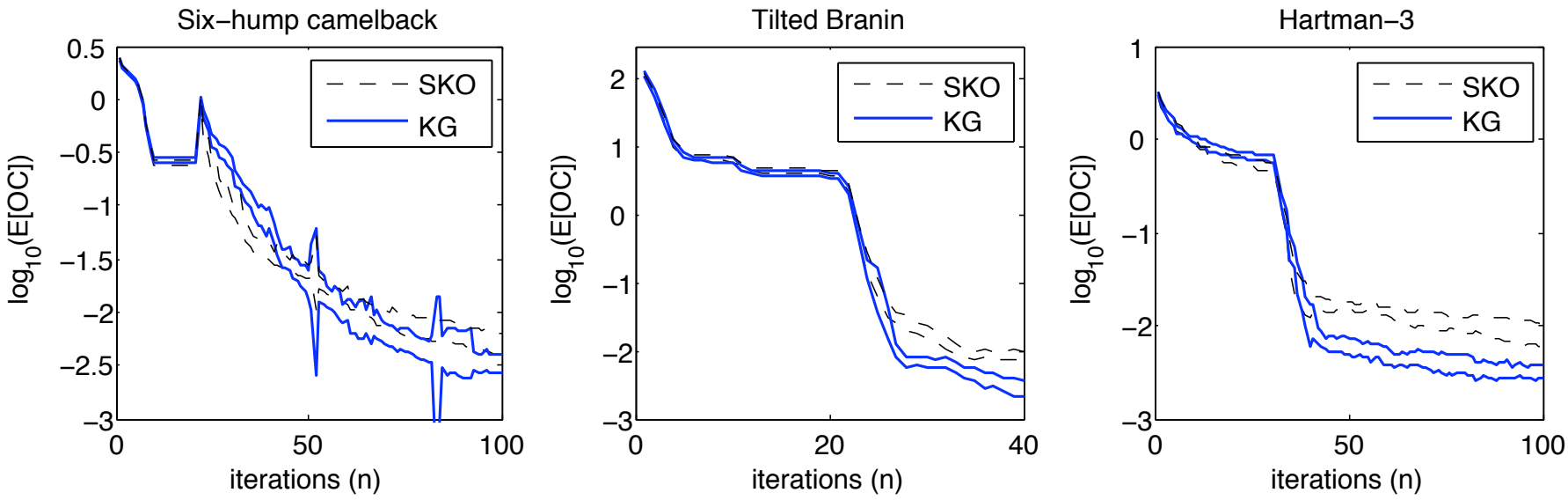
2. If there is no measurement noise, and alternatives may be reordered so that

$$\mu^0_1 \geq \mu^0_2 \geq \ldots \geq \mu^0_M$$
$$\Sigma^0_{11} \geq \Sigma^0_{22} \geq \ldots \geq \Sigma^0_{MM},$$

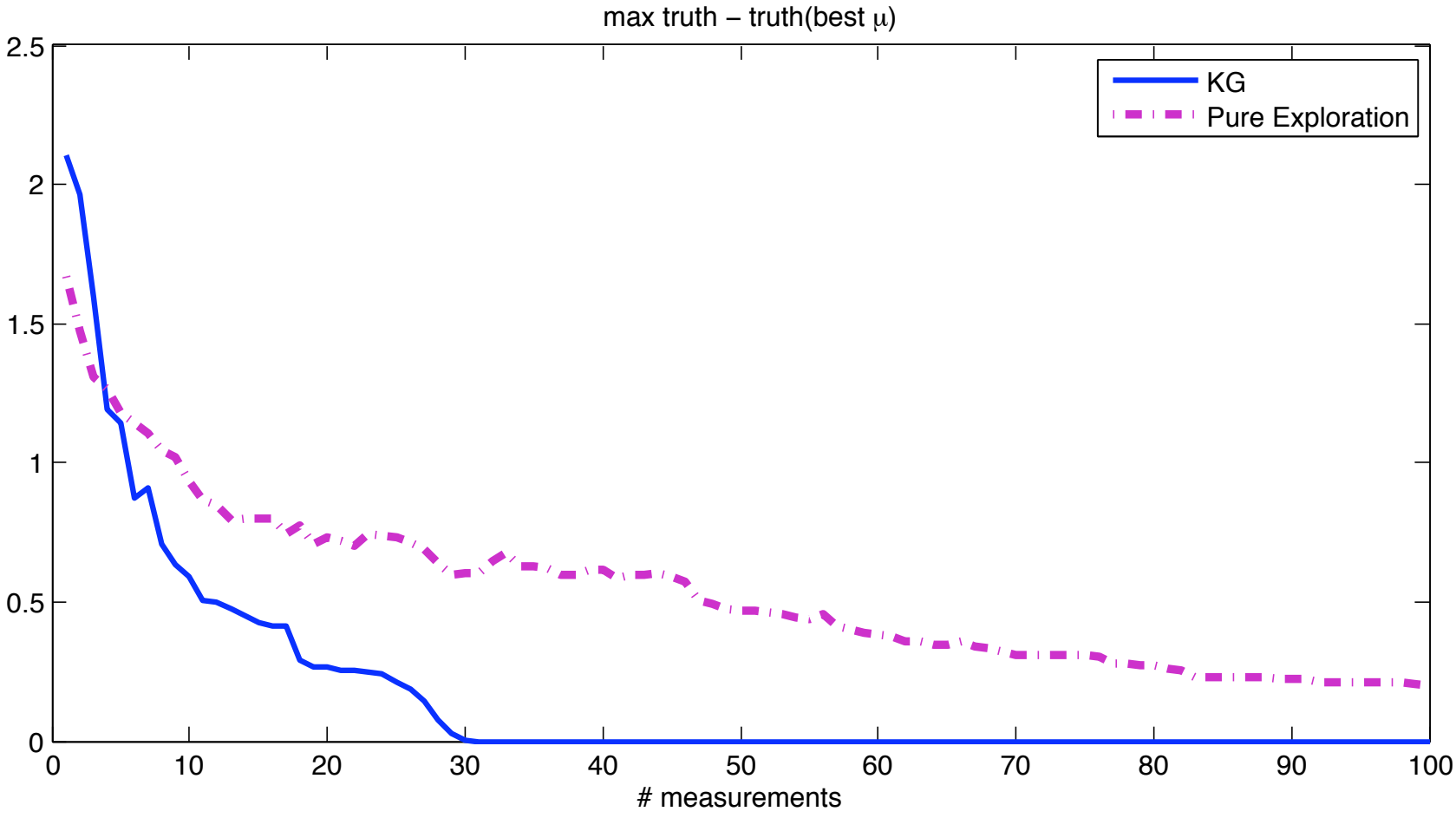
then the knowledge-gradient policy is optimal.

Outline

Global Optimization: Numerical Results

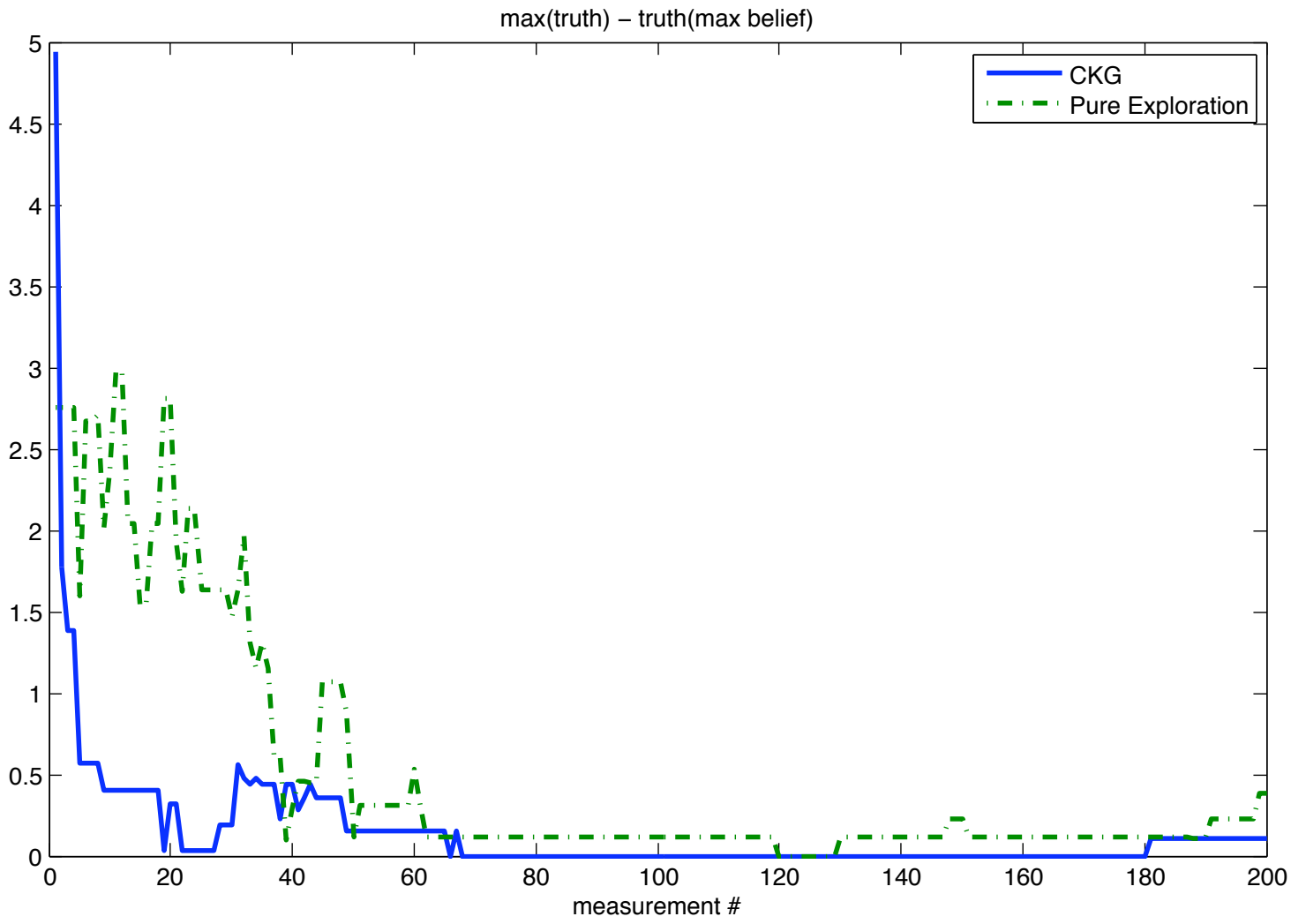


Drug Discovery: Numerical Results



Average over 100 sample paths on randomly selected subsets of benzomorphan compounds of size 99.

Drug Discovery: Numerical Results



One sample path on the full set of 87,120 benzomorphan compounds.

Conclusion

- Correlated beliefs allow us to perform well even with **large numbers of alternatives**.
- The knowledge-gradient policy:
  - is **"easy" to compute** (compared to the optimal policy),
  - **performs well in numerical experiments**,
  - and is **provably optimal in certain special cases**.

For Details, please see. . .

Frazier, P., Powell, W. B., and Dayanik, S. (2008).

A knowledge gradient policy for sequential information collection.

SIAM Journal on Control and Optimization.

Frazier, P., Powell, W. B., and Dayanik, S. (2009).

The knowledge gradient policy for correlated normal beliefs.

INFORMS Journal on Computing.

Negoescu, D., Frazier, P., and Powell, W. B. (2009).

The knowledge gradient algorithm for sequencing experiments in drug discovery.

in review.

Scott, W., Frazier, P. and Powell, W. B. (2010).

The correlated knowledge gradient with continuous decision variables applied to maximizing expensive smooth functions with noisy observations.

in preparation.

Thank You

Any questions?