# Appendix

### Bayesian Quadrature with Gaussian Processes

In previous literature, $p(x)$ is assumed to be Gaussian. For the Genz functions, $p(x)$ is uniform. Following Osborne (Osborne, 2010) and Duvenaud[1] we write:

$$Z = \int f(x)p(x)dx, \tag{1}$$

where we have a GP prior over $f$:

$$f \sim \mathcal{GP}(0, k(\cdot, \cdot)). \tag{2}$$

Given samples $x_1, \ldots, x_n$ we can calculate (Rasmussen & Williams, 2005):

$$E_{\textbf{GP}}\left[Z|f(x_1), \ldots, f(x_n)\right] = z^\top K^{-1}\boldsymbol{f}, \tag{3}$$

where $z = [z_1, \ldots, z_n]^\top$ and $z_i = \int k(x, x_i)p(x)dx$, and $\boldsymbol{f} = (f(x_1), .., f(x_n))^\top$. Sequential Bayesian Quadrature proceeds by finding the point $x$ that minimizes the variance of our estimate of $Z$:

$$Var_{\textbf{GP}}\left[Z|f(x_1), \ldots, f(x_n)\right] = \int\int k(x, x')p(x)p(x')dxdx' - z^\top K^{-1}z. \tag{4}$$

This does not depend on the values of the function at the points. For a given $k(x, x')$ and $p(x)$, we choose the point $x^*$:

$$\text{argmax}_{x^*} Var_{\textbf{GP}}\left[Z|f(x_1), \ldots, f(x_n), f(x^*)\right] \tag{5}$$

We will thus also need to calculate this double integral for $p(x)$ uniform.

Let us assume that we use the Gaussian kernel:

$$k(x, x') = exp(-\|x - x'\|^2/2). \tag{6}$$

For starters, we assume we are in one dimension with $p(x) = 1$ on the unit interval. Thus we can calculate:

$$\begin{aligned} z_i &= \int_0^1 k(x, x_i)dx \quad (7)\\ &= \int_0^1 \exp(-\|x - x_i\|^2/2)dx, \quad (8)\\ & \quad (9) \end{aligned}$$

which is computable using `pmvnorm` (in the `mvtnorm` package) in R.

For the double integral

$$\begin{aligned} &\int k(x, x')p(x)p(x')dxdx' \quad (10)\\ &= \int_0^1\int_0^1 \exp(-\|x - x'\|^2/2)dxdx', \quad (11)\\ & \quad (12) \end{aligned}$$

and we have

$$\begin{aligned} &\text{argmin}_{x^*} Var_{\textbf{GP}}\left[Z|f(x_1), \ldots, f(x_n), f(x^*)\right] \quad (13)\\ &= \text{argmin}_{x^*} \int_0^1\int_0^1 \exp(-\|x - x'\|^2/2)dxdx' - z^{*\top}K^{-1}z^* \quad (14)\\ &= \text{argmin}_{x^*} - z^{*\top}K^{-1}z^*, \quad (15) \end{aligned}$$

[1] https://www.cs.toronto.edu/~duvenaud/talks/intro_bq.pdf

where $z^* = \int_0^1 k(x, x^*)dx$, and Eq. (15) follows as the double integral does not depend on $x^*$. In practice, although the function values obtained are noise-free, we add a jitter of value $10^{-6}$ to $K$ when inverting the matrix for numerical stability.

Note that the double integral in the one-dimensional case (although not required when implementing the sequential design) can be computed by noticing that (wol)

$$\int_0^1 \int_0^1 \exp(-\|x - x'\|^2/2)dxdx' = \sqrt{\pi} \cdot \text{erf}(1) - 1 + \frac{1}{e} \approx 0.861528, \tag{16}$$

where erf stands for the error function

$$\text{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t^2}. \tag{17}$$

Note that for $\boldsymbol{x}, \boldsymbol{x}^{'} \in [0, 1]^N$, the integration becomes:

$$\int_{[0,1]^N} \int_{[0,1]^N} \exp(-\|\boldsymbol{x} - \boldsymbol{x}'\|^2/2)d\boldsymbol{x}d\boldsymbol{x}' = \left[\int_0^1 \int_0^1 \exp(-\|x - x'\|^2/2)dxdx'\right]^N, \tag{18}$$

which can be computed quickly given the one-dimensional result in Eq. (16).

### Analytical Formulae of the Genz Integrals

In our experiment, we tested the BQ with BART algorithm on six Genz integrand families (Genz, 1984), whose integral over $[0, 1]^d$ can be readily obtained analytically. Here, we provide the analytical formulae for two Genz families whose derivations require a little more effort than the others, namely the Corner Peak and the Oscillatory integrands. Derivations for the rest are straightforward. Note however that the Gaussian peak family is essentially a Gaussian kernel and does not have an analytically tractable integral; one can use the `pmvnorm` function in the R package `mvtnorm` to obtain the numerical values.

#### Oscillatory Integrand Family

For a given dimension $d \in \boldsymbol{N}$, we are interested in deriving the $d$-dimensional integral

$$I_d = \int_{\boldsymbol{x} \in [0,1]^d} f(\boldsymbol{x})d\boldsymbol{x}, \tag{19}$$

where $u$, $a \in \mathbb{R}$ are function parameters, $\boldsymbol{x} = (x_1, \ ..., \ x_d)^T$ and

$$f(\mathbf{x}) = \cos(2\pi u + a \sum_{i=1}^{d} x_i). \tag{20}$$

Using mathematical induction (by considering all cases) one can show that, for $d \in \mathbb{N}$,

$$I_d = \frac{1}{a^d} \sum_{i=0}^{d} \binom{d}{i} (-1)^{d-i} \phi_d(2\pi u + na), \tag{21}$$

where

$$\phi_d(\cdot) = \begin{cases} \sin(\cdot) & d \equiv 1 \bmod 4, \\ -\cos(\cdot) & d \equiv 2 \bmod 4, \\ -\sin(\cdot) & d \equiv 3 \bmod 4, \\ \cos(\cdot) & d \equiv 0 \bmod 4. \end{cases} \tag{22}$$

### Corner Peak Integrand Family

For the Corner Peak family, we proceed as before with the function

$$f(\mathbf{x}) = \left(1 + a \sum_{i=1}^{d} x_i\right)^{-(d+1)}, \tag{23}$$

where $a \in \mathbb{R}$ is a parameter. Again using induction, we obtained the following expression

$$I_d = \frac{1}{a^d d!} \sum_{i=1}^{d} \binom{d-1}{i-1} \frac{(-1)^{i-1}}{(5i+1)(5i-4)}. \tag{24}$$

## Plots of Estimates of the Genz Integrals

### Continuous Integrand Family

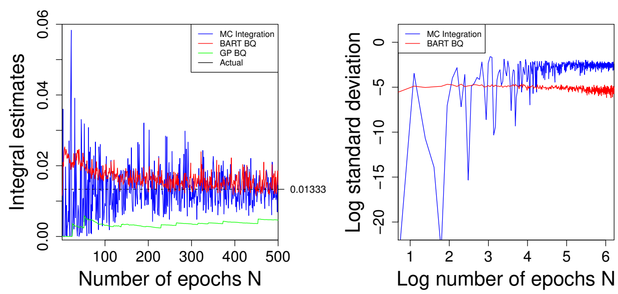


*Figure 1.* Dimension 1

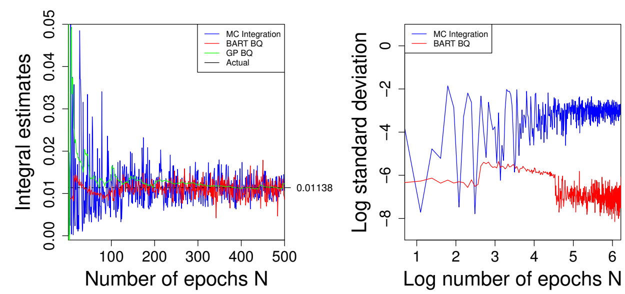


*Figure 2.* Dimension 2

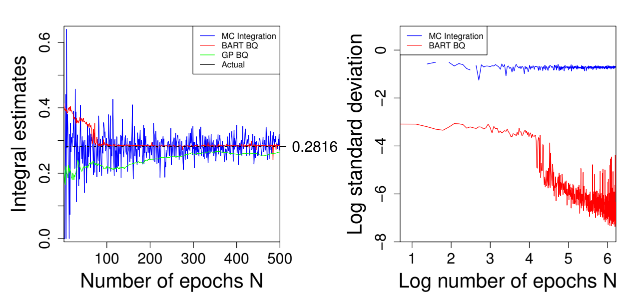


*Figure 3.* Dimension 3

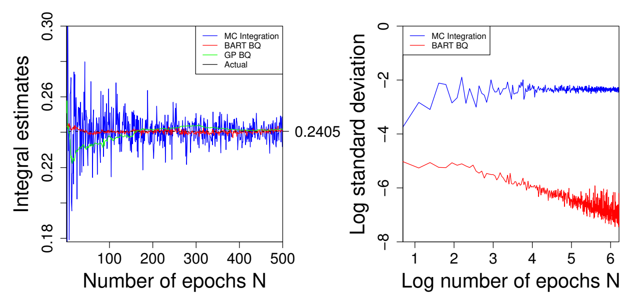


*Figure 4.* Dimension 5

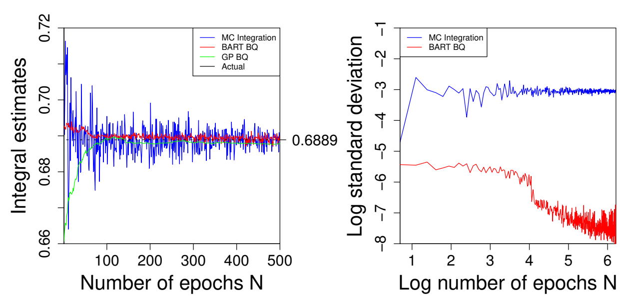


*Figure 5.* Dimension 10

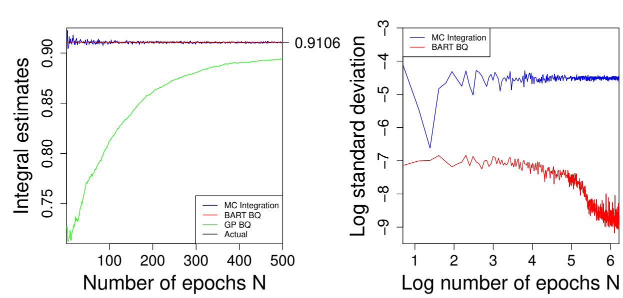


*Figure 6.* Dimension 20

## CORNER PEAK INTEGRAND FAMILY

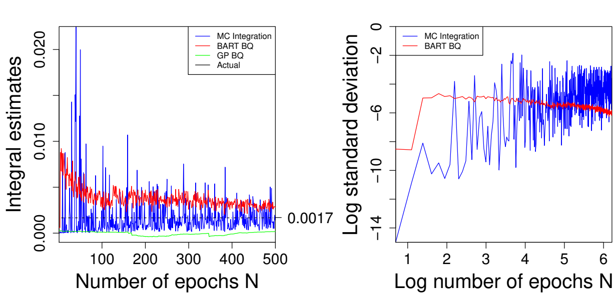


*Figure 7.* Dimension 1

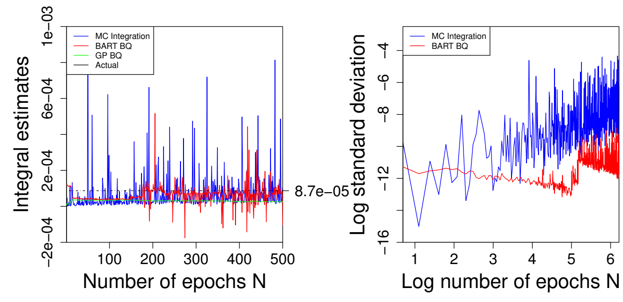


*Figure 8.* Dimension 2

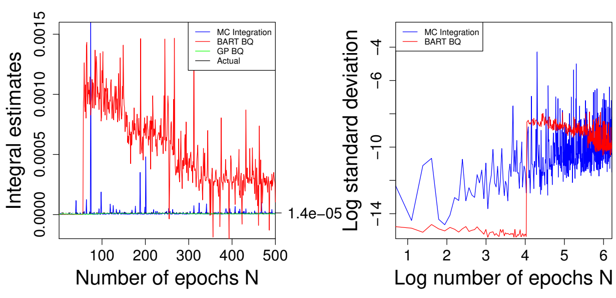


*Figure 9.* Dimension 3

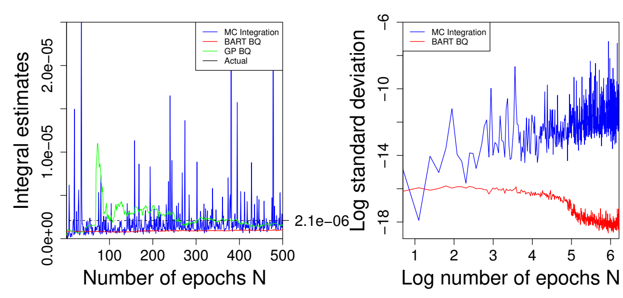


*Figure 10.* Dimension 5

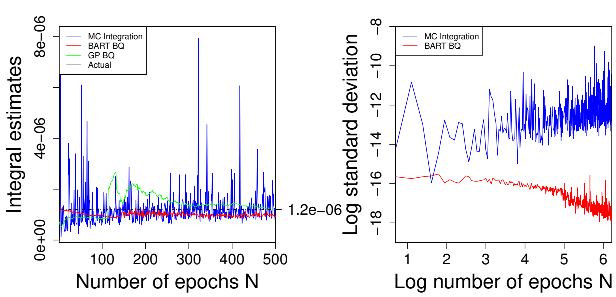


*Figure 11.* Dimension 10

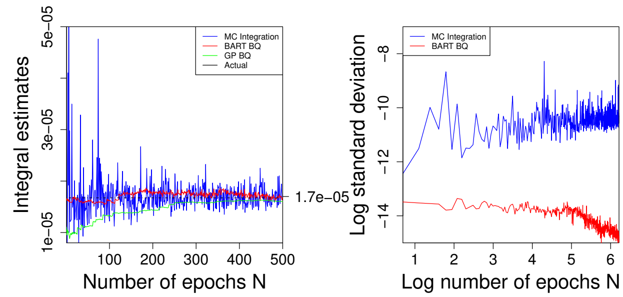


*Figure 12.* Dimension 20

## DISCONTINUOUS INTEGRAND FAMILY

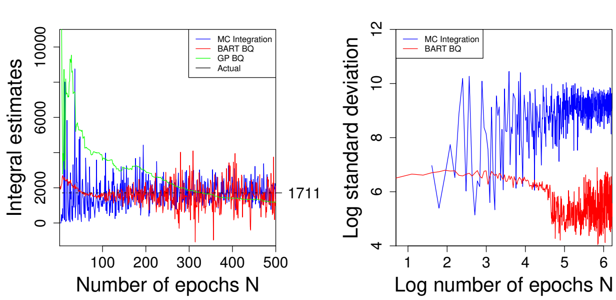


*Figure 13.* Dimension 2

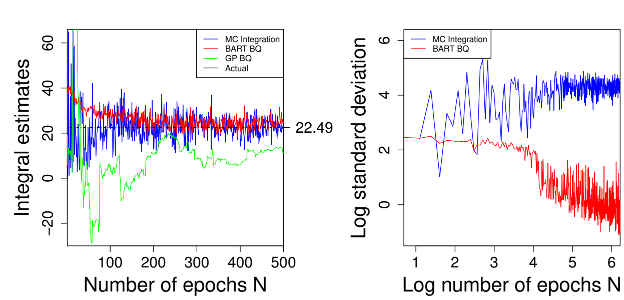


*Figure 14.* Dimension 3

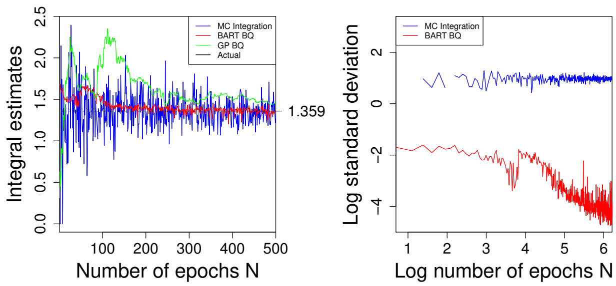


*Figure 15.* Dimension 5

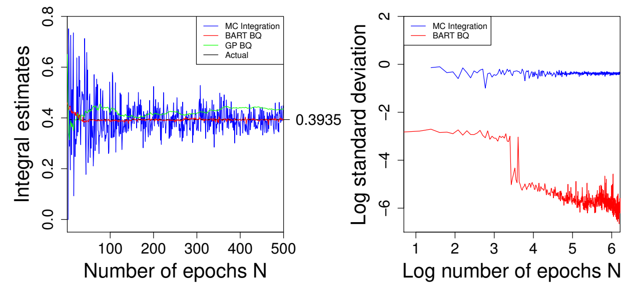


*Figure 16.* Dimension 10

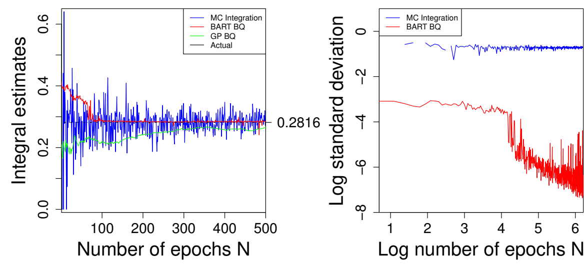


*Figure 17.* Dimension 20

### GAUSSIAN PEAK INTEGRAND FAMILY

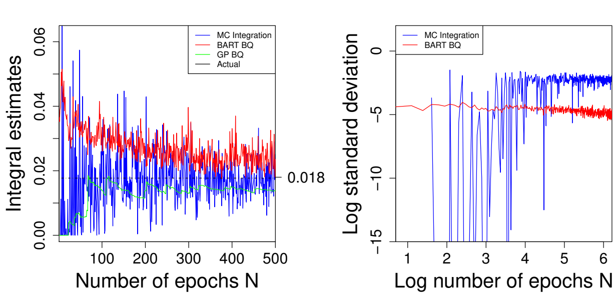


*Figure 18.* Dimension 1

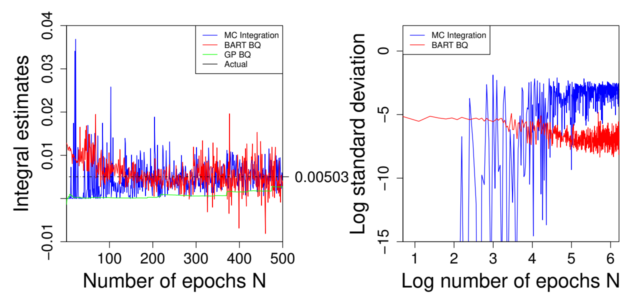


*Figure 19.* Dimension 2

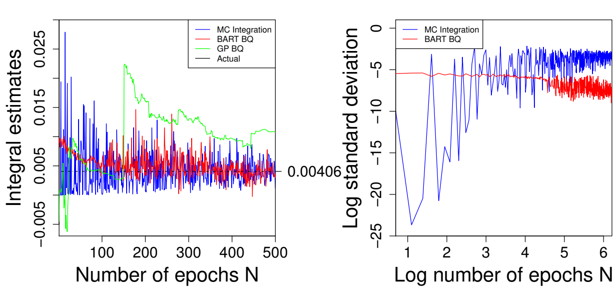


*Figure 20.* Dimension 3

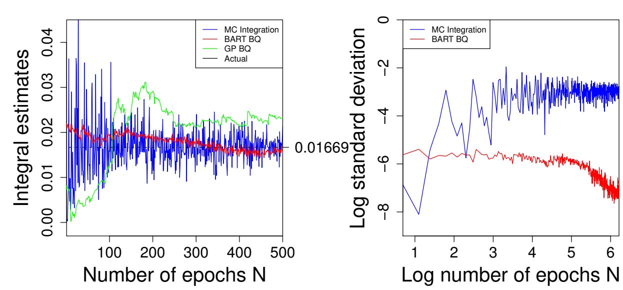


*Figure 21.* Dimension 5

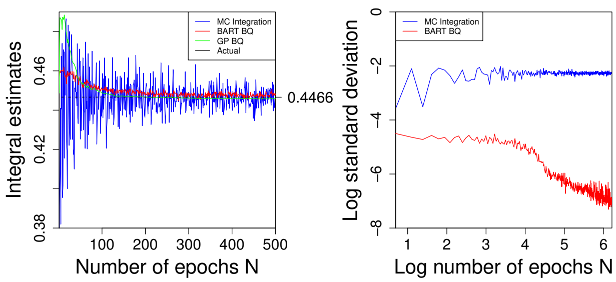


*Figure 22.* Dimension 10

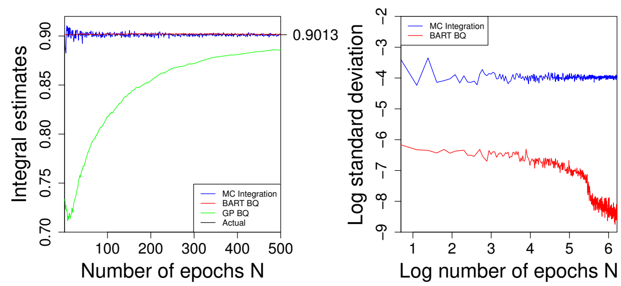


*Figure 23.* Dimension 20

## OSCILLATORY INTEGRAND FAMILY

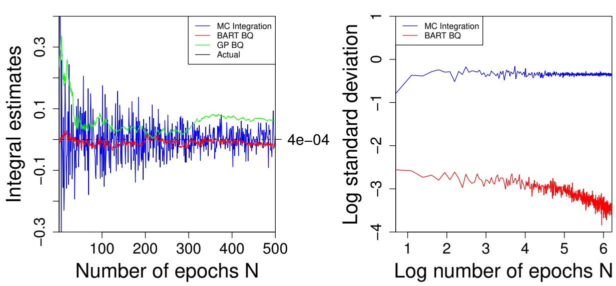


*Figure 24.* Dimension 1

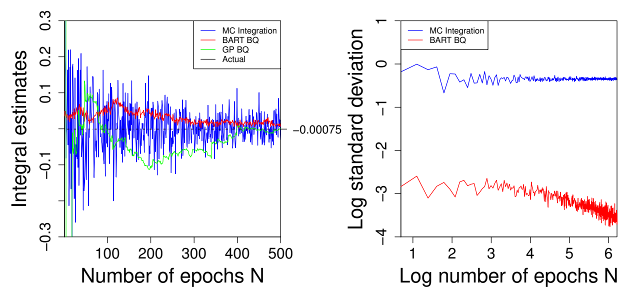


*Figure 25.* Dimension 2

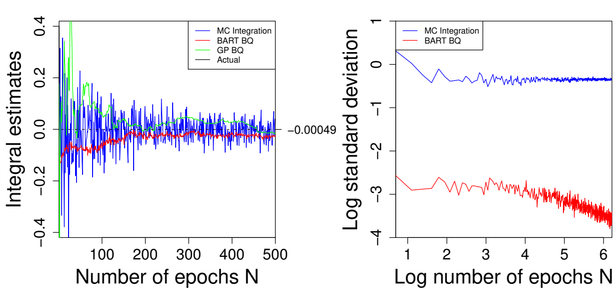


*Figure 26.* Dimension 3

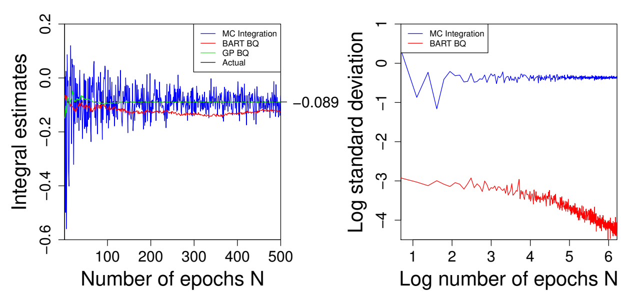


*Figure 27.* Dimension 5

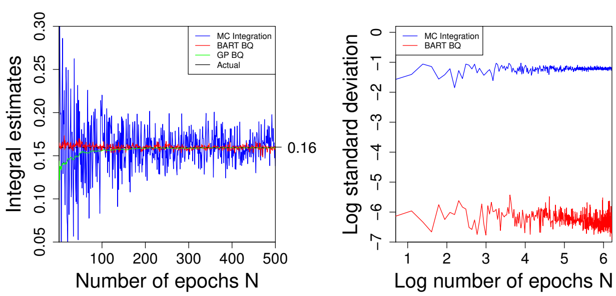


*Figure 28.* Dimension 10

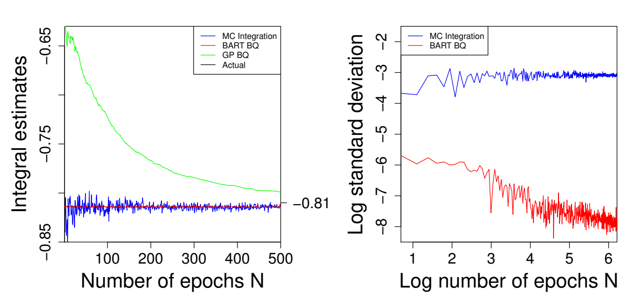


*Figure 29.* Dimension 20

**PRODUCT PEAK INTEGRAND FAMILY**

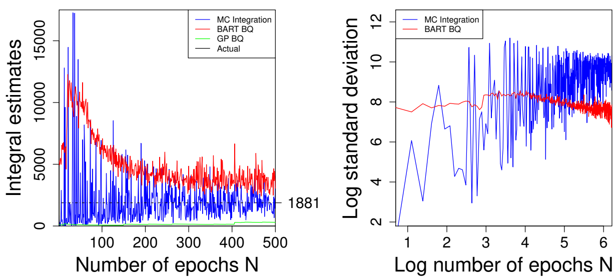


*Figure 30.* Dimension 1.

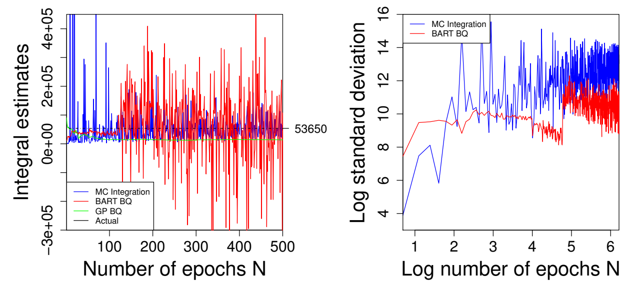


*Figure 31.* Dimension 2.

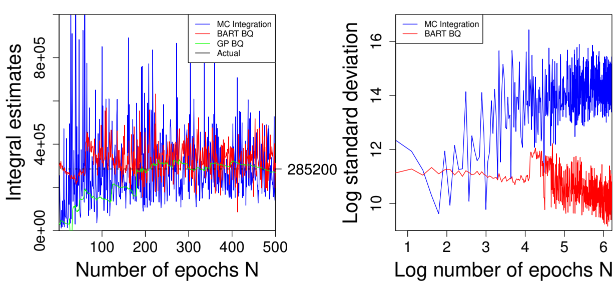


*Figure 32.* Dimension 3.

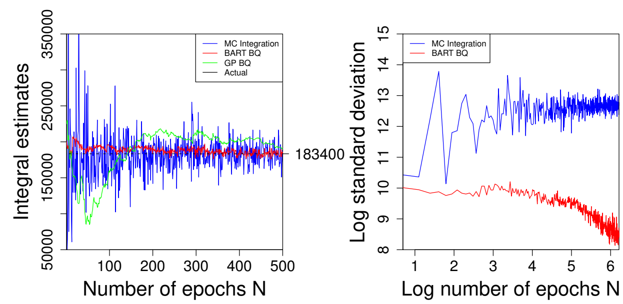


*Figure 33.* Dimension 5.

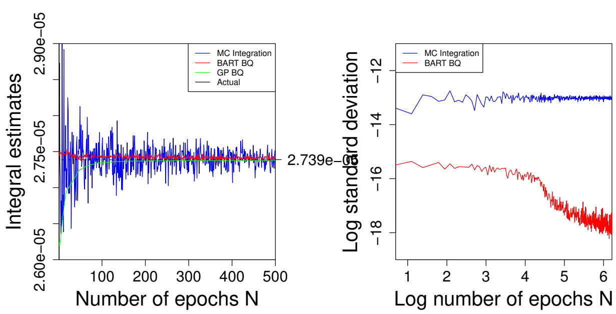


*Figure 34.* Dimension 10.

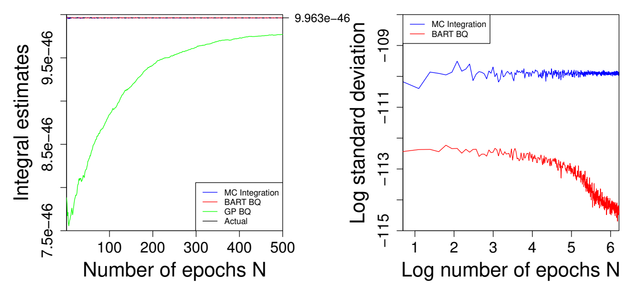


*Figure 35.* Dimension 20.

## References


Wolfram alpha: Computational intelligence. https://www.wolframalpha.com/. Accessed: 2018-01-23.

Genz, A. Testing multidimensional integration routines. In *Proc. Of International Conference on Tools, Methods and Languages for Scientific and Engineering Computation*, pp. 81–94, New York, NY, USA, 1984. Elsevier North-Holland, Inc. ISBN 0-444-87570-0. URL http://dl.acm.org/citation.cfm?id=2837.2842.

Osborne, M. A. *Bayesian Gaussian Processes for Sequential Prediction, Optimisation and Quadrature*. PhD thesis, Oxford University, 2010.

Rasmussen, C. E. and Williams, C. K. I. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005. ISBN 026218253X.