

A Bayes–Sard Cubature Method

Toni Karvonen¹, Chris. J. Oates^{2,3} & Simo Särkkä¹

¹Aalto University, Finland | ²Newcastle University, UK | ³Alan Turing Institute, UK

Bayes–Sard Cubature

Problem Statement

Numerical approximation of the integral

$$I(f^\dagger) := \int_D f^\dagger(x) d\nu(x), \quad D \subset \mathbb{R}^d, \quad \nu \text{ a probability measure on } D,$$

of a *deterministic* function $f^\dagger: D \rightarrow \mathbb{R}$.

Gaussian Process Model and Bayes–Sard Cubature

Model f^\dagger as a Gaussian process

$$f(x) | \gamma \sim \text{GP}(s(x), k(x, x')), \quad s(x) = \sum_{j=1}^Q \gamma_j p_j(x), \quad \gamma \sim \mathcal{N}(0, \Sigma).$$

The functions p_j span a Q -dimensional linear function space π (**typically polynomials**). Then

1. obtain data $\mathcal{D} = \{(x_i, f^\dagger(x_i))\}_{i=1}^n$ consisting of function evaluations at points $X = \{x_i\}_{i=1}^n \subset D$;
2. marginalise out γ and consider the *flat prior limit* $\Sigma^{-1} \rightarrow \mathbf{0}$ [1, Sec. 4.1.2];
3. the posterior for the integral $I(f)$ is Gaussian:

$$I(f) | \mathcal{D} \sim \mathcal{N}(\mu_X(f^\dagger), \sigma_X^2), \quad \mu_X(f^\dagger) = \sum_{i=1}^n w_{k,i} f^\dagger(x_i).$$

The posterior mean $\mu_X(f^\dagger)$ is an estimate for the integral $I(f^\dagger)$. We call this **Bayes–Sard cubature** (BSC). The variance σ_X^2 is supposed to quantify associated numerical uncertainty.

Properties of Bayes–Sard Cubature

Property 1. If $f^\dagger \in \pi = \text{span}\{p_1, \dots, p_Q\}$, then $\mu_X(f^\dagger) = I(f^\dagger)$ (but, in general, $\sigma_X^2 > 0$). Functions in π are integrated exactly. BSC is thus a mixture of conventional GP-based and classical cubatures.

Property 2. If $\dim(\pi) = Q = n$, the weights $w_{k,i}$ are completely determined by π . However, σ_X equals the worst-case error (WCE) in $H(k)$. If π consists of, e.g., polynomials, we recover classical Gaussian cubature rules and can interpret their WCE as a posterior std.

An Alternative

Related Bayesian cubature methods have been proposed before using improper priors [2,3]. This yields a Student- t posterior over the integral with $n - Q$ degrees of freedom. *Probabilistic interpretation of classical cubature rules is not possible as the posterior does not exist if $Q = n$.*

Summary

- BSC is more robust against improper prior specifications than standard Bayesian cubature.
- Can be easily used in high dimensions even with isotropic kernels.
- Allows for endowing *any* cubature rule with a meaningful probabilistic output.

References

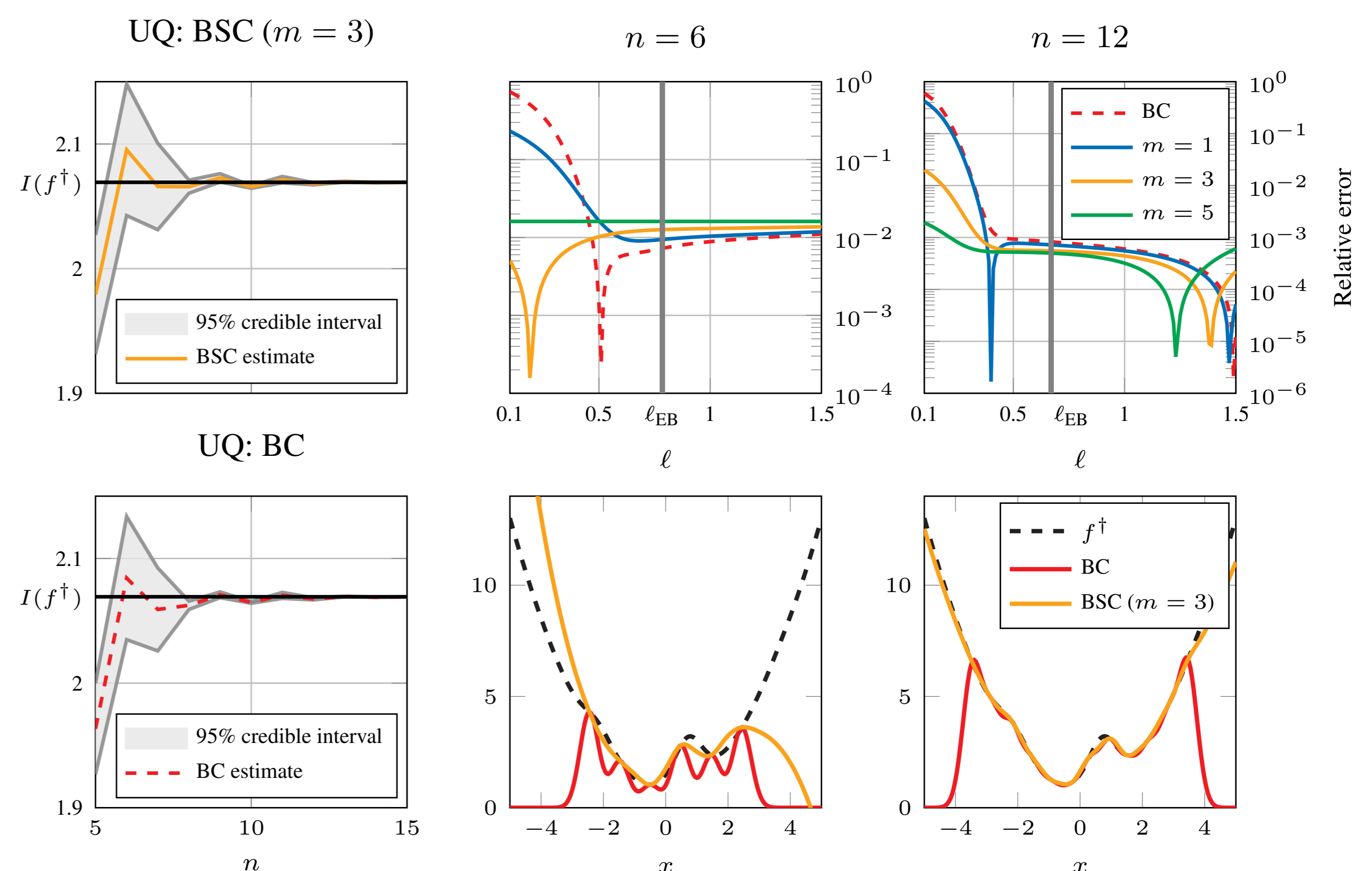
- [1] T. J. Santner, B. J. Williams & W. I. Notz, *The Design and Analysis of Computer Experiments*. Springer, 2003.
- [2] A. O'Hagan, "Bayes–Hermite quadrature." *Journal of Statistical Planning and Inference*, vol. 29, no. 3, pp. 245–260, 1991.
- [3] M. Kennedy, "Bayesian quadrature with non-normal approximating functions." *Statistics and Computing*, vol. 8, no. 4, pp. 365–375, 1998.
- [4] T. Karvonen, S. Särkkä & C. J. Oates, "Symmetry exploits for Bayesian cubature methods." Preprint, arXiv:1809.10227, 2018.

Numerical Examples

In all examples π is, for some $m \in \mathbb{N}_0$, the space of d -variate polynomials of degree at most m .

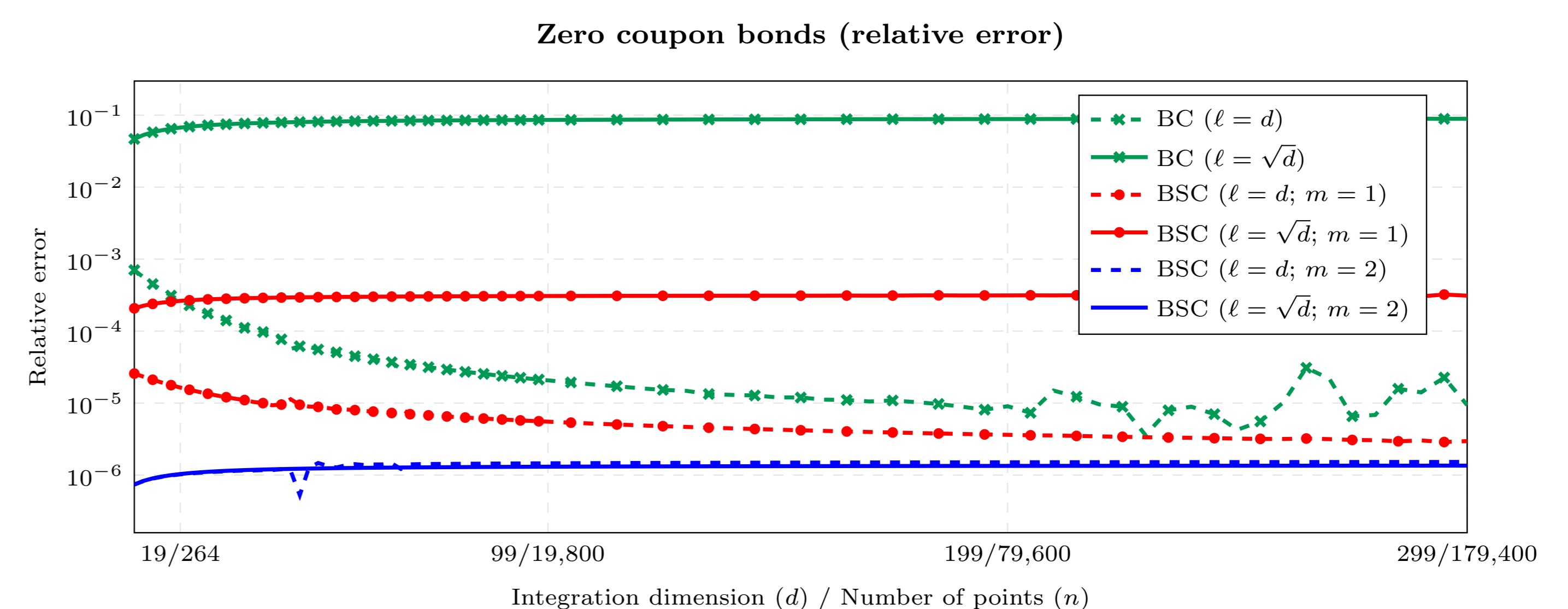
One-Dimensional Toy Example

Computation of Gaussian integral of a one-dimensional toy function.



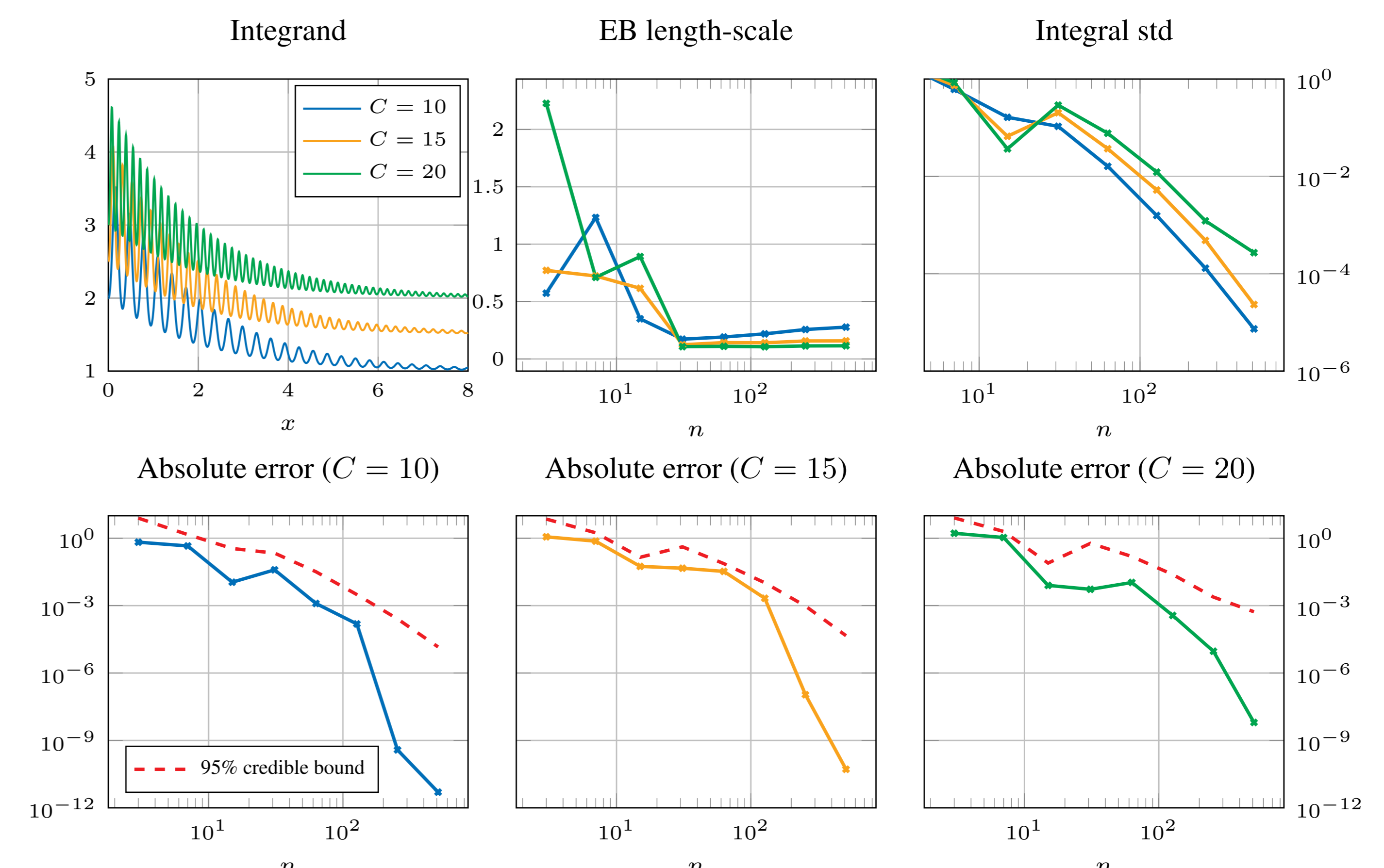
High-Dimensional Financial Example

Computation of the Gaussian integral over \mathbb{R}^d of a high-dimensional financial functions arising from discretisations of an SDE. Symmetric set are used for fast computations [4]. **The integrand varies with d .**



Uncertainty Quantification for Classical Quadrature

Uncertainty quantification for *Gauss–Patterson quadrature* using BSC with the Matérn 5/2 kernel $Q = n$. Kernel length-scale ℓ fitted using empirical Bayes and magnitude λ using the improper prior $p(\lambda) \propto 1/\lambda$.



Contact information for comments and questions:

Toni Karvonen (Aalto University, Finland)
email: toni.karvonen@aalto.fi
web: <https://users.aalto.fi/~karvont2/>