



HAL
open science

Sequential sampling for optimal weighted least squares approximations in hierarchical spaces

Benjamin Arras, Markus Bachmayr, Albert Cohen

► **To cite this version:**

Benjamin Arras, Markus Bachmayr, Albert Cohen. Sequential sampling for optimal weighted least squares approximations in hierarchical spaces. *SIAM Journal on Mathematics of Data Science*, 2019, 1, pp.189-207. hal-01802977

HAL Id: hal-01802977

<https://hal.science/hal-01802977>

Submitted on 30 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sequential sampling for optimal weighted least squares approximations in hierarchical spaces*

Benjamin Arras[†], Markus Bachmayr[‡] and Albert Cohen[§]

May 28, 2018

Abstract

We consider the problem of approximating an unknown function $u \in L^2(D, \rho)$ from its evaluations at given sampling points $x^1, \dots, x^n \in D$, where $D \subset \mathbb{R}^d$ is a general domain and ρ a probability measure. The approximation is picked in a linear space V_m where $m = \dim(V_m)$ and computed by a weighted least squares method. Recent results show the advantages of picking the sampling points at random according to a well-chosen probability measure μ that depends both on V_m and ρ . With such a random design, the weighted least squares approximation is proved to be stable with high probability, and having precision comparable to that of the exact $L^2(D, \rho)$ -orthonormal projection onto V_m , in a near-linear sampling regime $n \sim m \log m$. The present paper is motivated by the adaptive approximation context, in which one typically generates a nested sequence of spaces $(V_m)_{m \geq 1}$ with increasing dimension. Although the measure $\mu = \mu_m$ changes with V_m , it is possible to recycle the previously generated samples by interpreting μ_m as a mixture between μ_{m-1} and an update measure σ_m . Based on this observation, we discuss sequential sampling algorithms that maintain the stability and approximation properties uniformly over all spaces V_m . Our main result is that the total number of computed sample at step m remains of the order $m \log m$ with high probability. Numerical experiments confirm this analysis.

MSC 2010: 41A10, 41A65, 62E17, 65C50, 93E24

1 Introduction

Least squares approximations are ubiquitously used in numerical computation when trying to reconstruct an unknown function u defined on some domain $D \subseteq \mathbb{R}^d$ from its observations y^1, \dots, y^n at a limited amount of points $x^1, \dots, x^n \in D$. In its simplest form the method amounts to minimizing the least squares fit

$$\frac{1}{n} \sum_{i=1}^n |y^i - v(x^i)|^2, \quad (1.1)$$

over a set of functions v that are subject to certain constraints expressing a prior on the unknown function u . There are two classical approaches for imposing such constraints:

*Benjamin Arras is supported by the European Research Council under grant ERC AdG 338977 BREAD. Markus Bachmayr acknowledges support by the Hausdorff Center of Mathematics, University of Bonn. Albert Cohen is supported by the Institut Universitaire de France and by the European Research Council under grant ERC AdG 338977 BREAD.

[†]Sorbonne Universités, UPMC Univ Paris 06, CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, 4 place Jussieu, 75005 Paris, France (arras@ljl.math.upmc.fr)

[‡]Hausdorff Center for Mathematics & Institute for Numerical Simulation, Wegelerstr. 6, 53115 Bonn, Germany (bachmayr@ins.uni-bonn.de)

[§]Sorbonne Universités, UPMC Univ Paris 06, CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, 4 place Jussieu, 75005 Paris, France (cohen@ljl.math.upmc.fr)

- (i) Add a penalty term $P(v)$ to the least squares fit. Classical instances include norms of reproducing kernel Hilbert spaces or ℓ^1 norms that promote sparsity of v when expressed in a certain basis of functions.
- (ii) Limit the search of v to a space V_m of finite dimension $m \leq n$. Classical instances include spaces of algebraic or trigonometric polynomials, wavelets, or splines.

The present paper is concerned with the second approach, in which approximability by the space V_m may be viewed as a prior on the unknown function. We measure accuracy in the Hilbertian norm

$$\|v\| = \left(\int_D |v(x)|^2 d\rho \right)^{1/2} = \|v\|_{L^2(D,\rho)}, \quad (1.2)$$

where ρ is a probability measure over D . We denote by $\langle \cdot, \cdot \rangle$ the associated inner product. The error of best approximation is defined by

$$e_m(u) := \min_{v \in V_m} \|u - v\|, \quad (1.3)$$

and is attained by $P_m u$, the $L^2(D, \rho)$ -orthogonal projection of u onto V_m . Since the least squares approximation \tilde{u} is picked in V_m , it is natural to compare $\|u - \tilde{u}\|$ with $e_m(u)$. In particular, the method is said to be near-optimal (or instance optimal with constant C) if the comparison

$$\|u - \tilde{u}\| \leq C e_m(u), \quad (1.4)$$

holds for all u , where $C > 1$ is some fixed constant.

The present paper is motivated by applications where the sampling points x^i are not prescribed and can be chosen by the user. Such a situation typically occurs when evaluation of u is performed either by a computer simulation or a physical experiment, depending on a vector of input parameters x that can be set by the user. This evaluation is typically costly and the goal is to obtain a satisfactory *surrogate model* \tilde{u} from a minimal number of evaluations

$$y^i = u(x^i). \quad (1.5)$$

For a given probability measure ρ and approximation space V_m of interest, a relevant question is therefore whether instance optimality can be achieved with sample size n that is moderate, ideally linear in m .

Recent results of [10, 16] for polynomial spaces and [6] in a general approximation setting show that this objective can be achieved by certain random sampling schemes in the more general framework of *weighted least squares* methods. The approximation \tilde{u} is then defined as the solution to

$$\min_{v \in V_m} \frac{1}{n} \sum_{i=1}^n w(x_i) |y^i - v(x^i)|^2, \quad (1.6)$$

where w is a positive function and the x^i are independently drawn according to a probability measure μ , that satisfy the constraint

$$w d\mu = d\rho. \quad (1.7)$$

The choice of a sampling measure μ that differs from the error norm measure ρ appears to be critical in order to obtain instance optimal approximations with an optimal sampling budget.

In particular, it is shown in [10, 6], that there exists an optimal choice of (ρ, μ) such that the weighted least squares is stable with high probability and instance optimal in expectation,

under the near-linear regime $n \sim m$ up to logarithmic factors. The optimal sampling measure and weights are given by

$$d\mu_m = \frac{k_m}{m} d\rho \quad \text{and} \quad w_m = \frac{m}{k_m}, \quad (1.8)$$

where k_m is the so-called Christoffel function defined by

$$k_m(x) = \sum_{j=1}^m |\varphi_j(x)|^2, \quad (1.9)$$

with $\{\varphi_1, \dots, \varphi_m\}$ any $L^2(D, \rho)$ -orthonormal basis of V_m .

In many practical applications, the space V_m is picked within a family $(V_m)_{m \geq 1}$ that has the nestedness property

$$V_1 \subset V_2 \subset \dots \quad (1.10)$$

and accuracy is improved by raising the dimension m . The sequence $(V_m)_{m \geq 1}$ may either be a priori defined, or adaptively generated, which means that the way V_m is refined into V_{m+1} may depend on the result of the least squares computation. Example of such hierarchical adaptive or non-adaptive schemes include in particular:

- (i) Mesh refinement in low-dimension performed by progressive addition of hierarchical basis functions or wavelets, which is relevant for approximating piecewise smooth functions, such as images or shock profiles, see [3, 8, 9].
- (ii) Sparse polynomial approximation in high dimension, which is relevant for the treatment of certain parametric and stochastic PDEs, see [4].

In this setting, we are facing the difficulty that the optimal measure μ_m defined by (1.8) varies together with m .

In order to maintain an optimal sampling budget, one should avoid the option of drawing a new sample

$$S_m = \{x_m^1, \dots, x_m^n\} \quad (1.11)$$

of increasing size $n = n(m)$ at each step m . In the particular case where V_m are the univariate polynomials of degree $m - 1$, and ρ a Jacobi type measure on $[-1, 1]$, it is known that μ_m converges weakly to the equilibrium measure defined by

$$d\mu^*(y) = \frac{dy}{\pi\sqrt{1-y^2}}, \quad (1.12)$$

with a uniform equivalence

$$c_1\mu^* \leq \mu_m \leq c_2\mu^*, \quad (1.13)$$

see [13, 17]. This suggests the option of replacing all μ_m by the single μ^* , as studied in [16]. Unfortunately, such an asymptotic behaviour is not encountered for most general choices of spaces $(V_m)_{m \geq 1}$. An equivalence of the form (1.13) was proved in [14] for sparse multivariate polynomials, however with a ratio c_2/c_1 that increases exponentially with the dimension, which theoretically impacts in a similar manner the sampling budget needed for stability.

In this paper, we discuss sampling strategies that are based on the observation that the optimal measure μ_m enjoys the mixture property

$$\mu_{m+1} = \left(1 - \frac{1}{m+1}\right)\mu_m + \frac{1}{m+1}\sigma_{m+1}, \quad \text{where} \quad d\sigma_m := |\varphi_m|^2 d\rho. \quad (1.14)$$

As noticed in [12], this leads naturally to sequential sampling strategies, where the sample S_m is recycled for generating S_{m+1} . The main contribution of this paper is to analyze such a sampling strategy and prove that the two following properties can be jointly achieved in an expectation or high probability sense:

1. Stability and instance optimality of weighted least squares hold uniformly over all $m \geq 1$.
2. The total sampling budget after m step is linear in m up to logarithmic factors.

The rest of the paper is organized as follows. We recall in §2 stability and approximation estimates from [10, 6] concerning the weighted least squares method in a fixed space V_m . We then describe in §3 a sequential sampling strategy based on (1.14) and establish the above optimality properties 1) and 2). These optimality properties are numerically illustrated in §4 for the algorithm and two of its variants.

2 Optimal weighted least squares

We denote by $\|\cdot\|_n$ the discrete Euclidean norm defined by

$$\|v\|_n^2 := \frac{1}{n} \sum_{i=1}^n w(x^i) |v(x^i)|^2, \quad (2.1)$$

and by $\langle \cdot, \cdot \rangle_n$ the associated inner product. The solution $\tilde{u} \in V_m$ to (1.6) may be thought of as an orthogonal projection of u onto V_m for this norm. Expanding

$$\tilde{u} = \sum_{j=1}^m c_j \varphi_j, \quad (2.2)$$

in the basis $\{\varphi_1, \dots, \varphi_m\}$ of V_m , the coefficient vector $\mathbf{c} = (c_1, \dots, c_m)^T$ is solution to the linear system

$$\mathbf{G}_m \mathbf{c} = \mathbf{d}, \quad (2.3)$$

where \mathbf{G}_m is the Gramian matrix for the inner product $\langle \cdot, \cdot \rangle_n$ with entries

$$\mathbf{G}_{j,k} := \langle \varphi_j, \varphi_k \rangle_n = \frac{1}{n} \sum_{i=1}^n w(x^i) \varphi_j(x^i) \varphi_k(x^i), \quad (2.4)$$

and the vector \mathbf{d} has entries $\mathbf{d}_k = \frac{1}{n} \sum_{i=1}^n w(x^i) y^i \varphi_k(x^i)$. The solution \mathbf{c} always exists and is unique when \mathbf{G}_m is invertible.

Since x^1, \dots, x^n are drawn independently according to $d\mu$, the relation (1.7) implies that $\mathbb{E}(\langle v_1, v_2 \rangle_n) = \langle v_1, v_2 \rangle$, and in particular

$$\mathbb{E}(\mathbf{G}_m) = \mathbf{I}. \quad (2.5)$$

The stability and accuracy analysis of the weighted least squares method can be related to the amount of deviation between \mathbf{G}_m and its expectation \mathbf{I} measured in the spectral norm. Recall that for $m \times m$ matrices \mathbf{M} , this norm is defined as $\|\mathbf{M}\|_2 = \sup_{\|v\|_2=1} \|\mathbf{M}v\|_2$. This deviation also describe the closeness of the norms $\|\cdot\|$ and $\|\cdot\|_n$ over the space V_m , since one has

$$\|\mathbf{G}_m - \mathbf{I}\|_2 \leq \delta \iff (1 - \delta)\|v\|^2 \leq \|v\|_n^2 \leq (1 + \delta)\|v\|^2, \quad v \in V_m. \quad (2.6)$$

Note that this closeness also implies a bound

$$\kappa(\mathbf{G}_m) \leq \frac{1 + \delta}{1 - \delta}, \quad (2.7)$$

on the condition number of \mathbf{G}_m .

Following [6], we use the particular value $\delta = \frac{1}{2}$, and define \tilde{u} as the solution to (1.6) when $\|\mathbf{G}_m - \mathbf{I}\|_2 \leq \frac{1}{2}$ and set $\tilde{u} = 0$ otherwise. The probability of the latter event can be estimated by a matrix tail bound, noting that

$$\mathbf{G}_m = \frac{1}{n} \sum_{i=1}^n \mathbf{X}^i, \quad (2.8)$$

where the \mathbf{X}^i are n independent realizations of the rank one random matrix

$$\mathbf{X} := w(x)(\varphi_j(x)\varphi_k(x))_{j,k=1,\dots,m}, \quad (2.9)$$

where x is distributed according to μ . The matrix Chernoff bound, see Theorem 1.1 in [18], gives

$$\Pr\left(\|\mathbf{G}_m - \mathbf{I}\|_2 \geq \frac{1}{2}\right) \leq 2m \exp(-\gamma n/K), \quad \gamma := \frac{1}{2}(1 - \ln 2). \quad (2.10)$$

where $K := \|wk_m\|_{L^\infty}$ is an almost sure bound for $\|\mathbf{X}\|$. Since $\int wk_m d\mu = \int k_m d\rho = m$, it follows that K is always larger than m . With the choice $\mu = \mu_m$ given by (1.8) for the sampling measure, one has exactly $K = m$, which leads to the following result.

Theorem 2.1. *Assume that the sampling measure and weight function are given by (1.8). Then for any $0 < \varepsilon < 1$, the condition*

$$n \geq cm(\ln(2m) - \ln(\varepsilon)), \quad c := \gamma^{-1} = \frac{2}{1 - \ln 2}, \quad (2.11)$$

implies the following stability and instance optimality properties:

$$\Pr\left(\|\mathbf{G}_m - \mathbf{I}\|_2 \geq \frac{1}{2}\right) \leq \varepsilon, \quad m \geq 1. \quad (2.12)$$

The probabilistic inequality (2.12) induces an instance optimality estimate in expectation, since in the event where $\|\mathbf{G}_m - \mathbf{I}\|_2 \leq \frac{1}{2}$, one has

$$\begin{aligned} \|u - \tilde{u}\|^2 &= \|u - P_m u\|^2 + \|\tilde{u} - P_m u\|^2 \\ &\leq e_m(u)^2 + 2\|\tilde{u} - P_m u\|_n^2 \leq e_m(u)^2 + 2\|u - P_m u\|_n^2, \end{aligned} \quad (2.13)$$

and therefore

$$\mathbb{E}(\|u - \tilde{u}\|^2) \leq e_m(u)^2 + 2\mathbb{E}(\|u - P_m u\|_n^2) + \varepsilon\|u\|^2 = 3e_m(u)^2 + \varepsilon\|u\|^2. \quad (2.14)$$

By a more refined reasoning, the constant 3 can be replaced by $1 + \frac{c}{\ln(2m) - \ln(\varepsilon)}$ which tends to 1 as m/ε becomes large, see [6].

In summary, when using the optimal sampling measure μ_m defined by (1.8), stability and instance optimality can be achieved in the near linear regime

$$n = n_\varepsilon(m) := \lceil cm(\ln(2m) - \ln \varepsilon) \rceil, \quad (2.15)$$

where $\varepsilon \in]0, 1[$ controls the probability of failure. To simplify notation later, we set $n_\varepsilon(0) := 0$.

3 An optimal sequential sampling procedure

In the following analysis of a sequential sampling scheme, we assume a sequence of nested spaces $V_1 \subset V_2 \subset \dots$ and corresponding basis functions $\varphi_1, \varphi_2, \dots$ to be given such that for each m , $\{\varphi_1, \dots, \varphi_m\}$ is an orthonormal basis of V_m . In practical applications, such spaces may either be fixed in advance, or adaptively selected, that is, the choice of φ_{m+1} depends on the computation of the weighted least squares approximation (1.6) for V_m . In view of the previous result, one natural objective is to generate sequences of samples $(S_m)_{m \geq 1}$ distributed according to the different measures $(\mu_m)_{m \geq 1}$, with

$$\#(S_m) = n_\varepsilon(m), \quad (3.1)$$

for some prescribed $\varepsilon > 0$.

The simplest option for generating such sequences would be directly drawing samples from μ_m for each $m = 1, 2, \dots$ separately. Since we ask that $n_\varepsilon(m)$ is proportional to m up to logarithmic factors, this leads to a total cost C_m after m step given by

$$C_m = \sum_{k=1}^m \#(S_k), \quad (3.2)$$

which increases faster than quadratically with m . Instead, we want to recycle the existing samples S_m in generating S_{m+1} to arrive at a scheme such that the total cost C_m remains comparable to $n_\varepsilon(m)$, that is, close to linear in m .

To this end, we use the mixture property (1.14). In what follows, we assume a procedure for sampling from each update measure $d\sigma_j := |\varphi_j|^2 d\rho$ to be available. In the univariate case, standard methods are inversion transform sampling or rejection sampling. These methods may in turn serve in the multivariate case when the φ_j are tensor product basis functions on a product domain and ρ is itself of tensor product type, since σ_j are then product measures that can be sampled via their univariate factor measures. We first observe that, in order to draw x distributed according to μ_m , for some fixed m , we can proceed as follows:

$$\text{Draw } j \text{ uniformly distributed in } \{1, \dots, m\}, \text{ then draw } x \text{ from } \sigma_j. \quad (3.3)$$

Let now $(n(m))_{m \geq 1}$ be an increasing sequence representing the prescribed size of the samples $(S_m)_{m \geq 1}$. Suppose that we are given $S_m = \{x_m^1, \dots, x_m^{n(m)}\}$ i.i.d. according to μ_m . In order to obtain the new sample $S_{m+1} = \{x_{m+1}^1, \dots, x_{m+1}^{n(m+1)}\}$ i.i.d. according to μ_{m+1} , we can proceed as stated in the following Algorithm 1.

Algorithm 1 requires a fixed number $n(m+1) - n(m)$ of samples from μ_{m+1} and an additional number $\tilde{n}(m)$ of samples from σ_{m+1} . The latter is a random variable that can be expressed as

$$\tilde{n}(m) = \sum_{i=1}^{n(m)} b_{m+1}^i, \quad (3.4)$$

where for each fixed $m \geq 1$, the $(b_m^i)_{i=1, \dots, n(m)}$ are i.i.d. Bernoulli random variables with $\Pr(b_m^i = 1) = \frac{1}{m}$. Moreover, $\{b_m^i : i = 1, \dots, n(m), m \geq 1\}$ is a collection of independent random variables. This immediately gives an expression for the total cost C_m after m successive applications of Algorithm 1, beginning with $n(1)$ samples from μ_1 , as the random variable

$$C_m := n(m) + s(m), \quad s(m) := \sum_{k=1}^{m-1} \tilde{n}(k) = \sum_{k=1}^{m-1} \sum_{i=1}^{n(k)} b_{k+1}^i. \quad (3.5)$$

Algorithm 1 Sequential sampling

input: sample $S_m = \{x_m^1, \dots, x_m^{n(m)}\}$ from μ_m
output: sample $S_{m+1} = \{x_{m+1}^1, \dots, x_{m+1}^{n(m+1)}\}$ from μ_{m+1}

```
for  $i = 1, \dots, n(m)$  do
  draw  $a_i$  uniformly distributed in  $\{1, \dots, m + 1\}$ 
  if  $a_i = m + 1$  then
    draw  $x_{m+1}^i$  from  $\sigma_{m+1}$ 
  else
    set  $x_{m+1}^i := x_m^i$ 
  end if
end for
for  $i = n(m) + 1, \dots, n(m + 1)$  do
  draw  $x_{m+1}^i$  from  $\mu_{m+1}$  by (3.3)
end for
```

We now focus on the particular choice

$$n(m) := n_\varepsilon(m), \quad (3.6)$$

as in (2.15), for a prescribed $\varepsilon \in]0, 1[$. This particular choice ensures that, for all $m \geq 1$,

$$\Pr(\|\mathbf{G}_m - \mathbf{I}\|_2 \geq \frac{1}{2}) \leq \varepsilon, \quad (3.7)$$

where \mathbf{G}_m denotes the Gramian for V_m according to (2.4).

We first estimate $\mathbb{E}(s(m))$ for this choice of $n(m)$. For this purpose, we note that $\mathbb{E}(\tilde{n}(k)) = \frac{n(k)}{k+1}$ and use the following lemma.

Lemma 3.1. For $m \geq 1$ and $\varepsilon > 0$,

$$\frac{1}{2}n(m) - 2c \leq \sum_{k=1}^m \frac{n(k)}{k+1} \leq n(m) + 1. \quad (3.8)$$

Proof: For the upper bound, we note that $n(k) \leq ck(\ln(2k) - \ln \varepsilon) + 1$ and

$$\begin{aligned} \sum_{k=1}^m \frac{1}{k+1} (ck(\ln(2k) - \ln \varepsilon) + 1) &\leq c \sum_{k=1}^m (\ln(2k) - \ln \varepsilon) + \sum_{k=1}^m \frac{1}{k+1} \\ &\leq c(m \ln 2 + \ln m! - m \ln \varepsilon) + \ln(m+1), \end{aligned} \quad (3.9)$$

By the Stirling bound $k! \leq ek^{k+\frac{1}{2}}e^{-k}$ for $k \geq 1$,

$$\ln m! \leq m(\ln m - 1) + \frac{1}{2} \ln m + 1. \quad (3.10)$$

Combining this with (3.9) gives

$$\begin{aligned} \sum_{k=1}^m \frac{n(k)}{k+1} &\leq cm(\ln 2m - \ln \varepsilon) - c(m-1) + \frac{c}{2} \ln m + \ln(m+1) \\ &\leq n(m) + 1, \end{aligned} \quad (3.11)$$

where the inequality $-c(m-1) + \frac{c}{2} \ln m + \ln(m+1) \leq 1$ for $m \geq 1$ is verified for the choice of c in (2.11) by direct evaluation for $m = 1$ and monotonicity. For the lower bound, we estimate

$$\sum_{k=1}^m \frac{n(k)}{k+1} \geq \frac{cm}{2}(\ln 2 - \ln \varepsilon) + c \sum_{k=1}^m \frac{k}{k+1} \ln k. \quad (3.12)$$

Using monotonicity and integration by parts with $\frac{d}{dx}(x - \ln(1+x)) = \frac{x}{x+1}$, we obtain

$$\begin{aligned} \sum_{k=1}^m \frac{k}{k+1} \ln k &\geq \int_1^m \frac{x}{x+1} \ln x \, dx \\ &= - \int_1^m \frac{x - \ln(x+1)}{x} \, dx + [(x - \ln(x+1)) \ln x]_1^m \\ &= -(m-1) + \int_1^m \frac{\ln(x+1)}{x} \, dx + (m - \ln(m+1)) \ln m. \end{aligned} \quad (3.13)$$

Moreover,

$$\int_1^m \frac{\ln(x+1)}{x} \, dx = \int_0^{\ln m} \ln(1+e^t) \, dt \geq \int_0^{\ln m} t \, dt = \frac{1}{2} \ln^2 m, \quad (3.14)$$

and using this in (3.13) gives

$$\sum_{k=1}^m \frac{k}{k+1} \ln k \geq \frac{1}{2} m \ln m + \frac{1}{2c} + R(m), \quad (3.15)$$

where

$$R(m) = \frac{1}{2} m (\ln m - 2) - \ln(m+1) \ln m + \frac{1}{2} \ln^2 m + 1 - \frac{1}{2c} > -2. \quad (3.16)$$

The latter inequality can be directly verified for the first few values of m and then follows for $m \geq 1$ by monotonicity. Thus, we obtain

$$\sum_{k=1}^m \frac{n(k)}{k+1} \geq \frac{1}{2} [cm(\ln 2m - \ln \varepsilon) + 1] - 2c \geq \frac{1}{2} n(m) - 2c, \quad (3.17)$$

which concludes the proof of the lemma. \square

As an immediate consequence, we obtain an estimate of the total cost in expectation by

Theorem 3.2. *With $n(m) = n_\varepsilon(m)$, the total cost after m steps of Algorithm 1 satisfies*

$$n(m) + \frac{1}{2} n(m-1) - 2c \leq \mathbb{E}(C_m) \leq n(m) + n(m-1) + 1. \quad (3.18)$$

We next derive estimates for the probability that the upper bound in the above estimate is exceeded substantially by C_m . The following Chernoff inequality can be found in [1]. For convenience of the reader, we give a standard short proof following [15].

Lemma 3.3. *Let $N \geq 1$, $(p_i)_{1 \leq i \leq N} \in [0, 1]^N$ and $(X_i)_{1 \leq i \leq N}$ be a collection of independent Bernoulli random variables with $\Pr(X_i = 1) = p_i$ for all $1 \leq i \leq N$. Set $\bar{X} = \mathbb{E}(X_1 + \dots + X_N) = p_1 + \dots + p_N$. Then, for all $\tau \geq 0$,*

$$\Pr(X_1 + \dots + X_N \geq (1 + \tau)\bar{X}) \leq (1 + \tau)^{-(1+\tau)\bar{X}} e^{\tau\bar{X}}.$$

In particular, for all $\tau \in [0, 1]$,

$$\Pr(X_1 + \dots + X_N \geq (1 + \tau)\bar{X}) \leq e^{-\tau^2 \bar{X}/3}.$$

Proof: For $t \geq 0$ and $Y := X_1 + \dots + X_N$, $\Pr(Y \geq (1 + \tau)\bar{X}) \leq e^{-t(1+\tau)\bar{X}} \mathbb{E}(e^{tY})$ by Markov's inequality, and using that $1 + a \leq e^a$, $a \in \mathbb{R}$,

$$\mathbb{E}(e^{tY}) = \prod_{i=1}^N \mathbb{E}(e^{tX_i}) = \prod_{i=1}^N (p_i e^t + (1 - p_i)) \leq \prod_{i=1}^N e^{p_i(e^t - 1)} = e^{(e^t - 1)\bar{X}}.$$

Now take $t = \ln(1 + \tau)$. Moreover, for $\tau \in [0, 1]$, one has $\tau - (1 + \tau) \ln(1 + \tau) \leq -\frac{1}{3}\tau^2$. \square

We now apply this result to the random part $s(m)$ of the total sampling costs C_m as defined in (3.5), and obtain the following probabilistic estimate.

Theorem 3.4. *With $n(m) = n_\varepsilon(m)$, the total cost after m step of Algorithm 1 satisfies, for any $\tau \in [0, 1]$,*

$$\begin{aligned} \Pr(C_m \geq n(m) + (1 + \tau)(n(m-1) + 1)) &\leq M_\tau e^{-\frac{\tau^2}{6}n(m-1)} \\ &\leq M_\tau \left(\frac{2(m-1)}{\varepsilon} \right)^{-\frac{\tau^2 \varepsilon}{6}(m-1)} \end{aligned} \quad (3.19)$$

with $M_\tau := e^{\frac{2c\tau^2}{3}}$.

Proof: We apply Proposition 3.3 to the independent variable b_k^i which appear in $s(m)$, which gives

$$\Pr(C_m \geq n(m) + (1 + \tau)\mathbb{E}(s(m))) = \Pr(s(m) \geq (1 + \tau)\mathbb{E}(s(m))) \leq e^{-\tau^2 \mathbb{E}(s(m))/3}. \quad (3.20)$$

The result follows by using the lower bound on $\mathbb{E}(s(m))$ in Lemma 3.1. \square

With the choice $n(m) = n_\varepsilon(m)$ we are ensured that, for each value of m separately, the failure probability is bounded by ε . When the intermediate results in each step are used to drive an adaptive selection of the sequence of basis functions, however, it will typically be of interest to ensure the stronger uniform statement that with high probability, $\|\mathbf{G}_m - \mathbf{I}\|_2 \leq \frac{1}{2}$ for *all* m . To achieve this, we now consider a slight modification of the above results with m -dependent choice of failure probability to ensure that \mathbf{G}_m remains well-conditioned with high probability jointly for all m . We define the sequence of failure probabilities

$$\varepsilon(m) = \frac{6\varepsilon_0}{(\pi m)^2}, \quad m \geq 1, \quad (3.21)$$

for a fixed $\varepsilon_0 \in]0, 1[$, and now analyze the repeated application of Algorithm 1, now using

$$n(m) := n_{\varepsilon(m)}(m) = \lceil c m (\ln(2m) + 2 \ln m - \ln(6\varepsilon_0/\pi^2)) \rceil \quad (3.22)$$

samples for V_m . Note that $n(m)$ differs only by a further term of order $\log m$ from $n_{\varepsilon_0}(m)$.

Lemma 3.5. *For $m \geq 1$, let $n(m)$ be defined as in (3.22) with $\varepsilon(m)$ as in (3.21). Then*

$$\frac{1}{2}n(m) - 6c \leq \sum_{k=1}^m \frac{n(k)}{k+1} \leq n(m) + 1. \quad (3.23)$$

Proof: For the upper bound in (3.23), note that $n(k) \leq ck(\ln(2k) + \ln k^2 - \ln(6\varepsilon_0/\pi^2)) + 1$. Using (3.10),

$$\sum_{k=1}^m \ln k^2 \leq m \ln m^2 + 2(1 - m + \frac{1}{2} \ln m) \leq m \ln m^2. \quad (3.24)$$

Thus, for all $m \geq 1$,

$$\begin{aligned} \sum_{k=1}^m \frac{n(k)}{k+1} &\leq c \sum_{k=1}^m \frac{k}{k+1} \left(\ln(2k) - \ln\left(\frac{6\varepsilon_0}{\pi^2}\right) \right) + \sum_{k=1}^m \frac{1}{k+1} + c \sum_{k=1}^m \frac{k \ln k^2}{1+k} \\ &\leq cm(\ln 2m - \ln(6\varepsilon_0/\pi^2)) + 1 + cm \ln m^2 \\ &\leq n(m) + 1, \end{aligned}$$

where we have used (3.9), (3.11) together with (3.24). Let us deal with the lower bound. For all $m \geq 1$,

$$\begin{aligned} \sum_{k=1}^m \frac{n(m)}{k+1} &\geq \sum_{k=1}^m \frac{1}{k+1} \left(ck \left(\ln(2k) - \ln\left(\frac{6\varepsilon_0}{\pi^2}\right) + \ln k^2 \right) \right) \\ &\geq \frac{1}{2} \left(cm \left(\ln(2m) - \ln\left(\frac{6\varepsilon_0}{\pi^2}\right) \right) + 1 \right) - 2c + \sum_{k=1}^m c \frac{k}{k+1} \ln k^2. \end{aligned} \quad (3.25)$$

Moreover using (3.15) and (3.16)

$$c \sum_{k=1}^m \frac{k}{k+1} \ln k^2 \geq 2c \left(\frac{m}{2} \ln m + \frac{1}{2c} - 2 \right).$$

Thus, combining the previous bound together with (3.25) and the definition of $n(m)$ concludes the proof of the lemma. \square

As a consequence, analogously to (3.18) we have

$$\mathbb{E}(C_m) \leq n(m) + n(m-1) + 1. \quad (3.26)$$

Using the above lemma as in Theorem 3.4 combined with a union bound, we arrive at the following uniform stability result.

Theorem 3.6. *Let $\varepsilon(m)$ be defined as in (3.21). Then applying Algorithm 1 with $n(m)$ as in (3.22), one has*

$$\Pr(\exists m \in \mathbb{N}: \|\mathbf{G}_m - \mathbf{I}\|_2 \geq \frac{1}{2}) \leq \varepsilon_0,$$

and for any $\tau \in [0, 1]$ and all $m \geq 1$, the random variable C_m satisfies

$$\begin{aligned} \Pr(C_m \geq n(m) + (1 + \tau)(n(m-1) + 1)) &\leq M_\tau e^{-\frac{\tau^2}{6} n(m-1)} \\ &\leq M_\tau \left(\frac{\pi^2(m-1)}{3\varepsilon_0} \right)^{-\frac{\tau^2 c}{2}(m-1)} \end{aligned} \quad (3.27)$$

with $M_\tau := e^{2c\tau^2}$.

In summary, applying Algorithm 1 successively to generate the samples S_1, S_2, \dots , we can ensure that $\|\mathbf{G}_m - \mathbf{I}\|_2 \leq \frac{1}{2}$ holds uniformly for all steps with probability at least $1 - \varepsilon_0$. The corresponding total costs for generating S_1, \dots, S_m can exceed a fixed multiple of $n(m)$ only with a probability that rapidly approaches zero as m increases.

Remark 3.7. Algorithm 1 and the above analysis can be adapted to create samples from μ_{m+q} , $q \geq 2$, using those for μ_m , which corresponds to adding q basis functions in one step. In this case,

$$\mu_{m+q} = \frac{m}{m+q}\mu_m + \frac{q}{m+q} \sum_{j=m+1}^{m+q} \frac{1}{q}\sigma_j, \quad (3.28)$$

where samples from $\sum_{j=m+1}^{m+q} \frac{1}{q}\sigma_j$ can be obtained by mixture sampling as in (3.3).

4 Numerical illustration

In our numerical tests, we consider two different types of orthonormal bases and corresponding target measures ρ on $D \subseteq \mathbb{R}$:

- (i) On the one hand, we consider the case where D is equal to \mathbb{R} , ρ is the standard Gaussian measure on \mathbb{R} and V_m the vector space spanned by the Hermite polynomials normalized to one with respect to the norm $\|\cdot\|$ up to degree $m-1$, for all $m \geq 1$. This case is an instance of polynomial approximation method where the function u and its approximants from V_m might be unbounded in L^∞ , as well as the Christoffel function $k_m(x) := \sum_{j=1}^m |H_{j-1}(x)|^2$.
- (ii) On the other hand, we consider the case where $D = [0, 1]$ with ρ the uniform measure on $[0, 1]$ and the approximation spaces $(V_m)_{m \geq 1}$ are generated by Haar wavelets refinement. The Haar wavelets are of the form $\psi_{l,k} = 2^{l/2}\psi(2^l - k)$ with $l \geq 0$ and $k = 0, \dots, 2^l - 1$, and $\psi := \chi_{[0,1/2[} - \chi_{[1/2,1[}$. In adaptive approximation, the spaces V_m are typically generated by including as the scale level l grows, the values of k such that the coefficient of $\psi_{l,k}$ for the approximated function is expected to be large. This can be described by growing a finite tree within the hierarchical structure induced by the dyadic indexing of the Haar wavelet family: the indices $(l+1, 2k)$ or $(l+1, 2k+1)$ can be selected only if (l, k) has already been. In our experiment, we generate the spaces V_m by letting such a tree grow at random, starting from $V_1 = \text{span}\{\psi_{0,0}\}$. This selection of $(V_m)_{m \geq 1}$ is done once and used for all further tests. The resulting sampling measures $(\mu_m)_{m \geq 1}$ exhibit the local refinement behavior of the corresponding approximation spaces $(V_m)_{m \geq 1}$.

As seen further, although the spaces and measures are quite different, the cost of the sampling algorithm behaves similarly in these two cases. While we have used univariate domains D for the sake of numerical simplicity, we expect a similar behaviour in multivariate cases, since our results are also immune to the spatial dimension d . As already mentioned, the sampling method is then facilitated when the functions φ_j are tensor product functions and ρ is a product measure.

The sampling measures μ_m are shown in Figure 1. In case (ii), the measures σ_j are uniform measures on dyadic intervals in $]0, 1[$, from which we can sample directly, using $\mathcal{O}(1)$ operations per sample. In case (i), we have instead $d\sigma_j = |H_{j-1}|^2 d\rho$. Several strategies for sampling from these densities are discussed in [6]. For our following tests, we use inverse transform sampling: we draw z uniformly distributed in $[0, 1]$, and obtain a sample x from σ_j by solving $\Phi_j(x) := \int_{-\infty}^x |H_{j-1}|^2 d\rho = z$. To solve these equations, in order to ensure robustness, we use a simple bisection scheme. Each point value of Φ_j can be obtained using $\mathcal{O}(j)$ operations, exploiting the fact that for $j \geq 2$,

$$\Phi_j = \Phi - \sum_{k=1}^{j-1} \frac{H_k H_{k-1}}{\sqrt{k}} g,$$

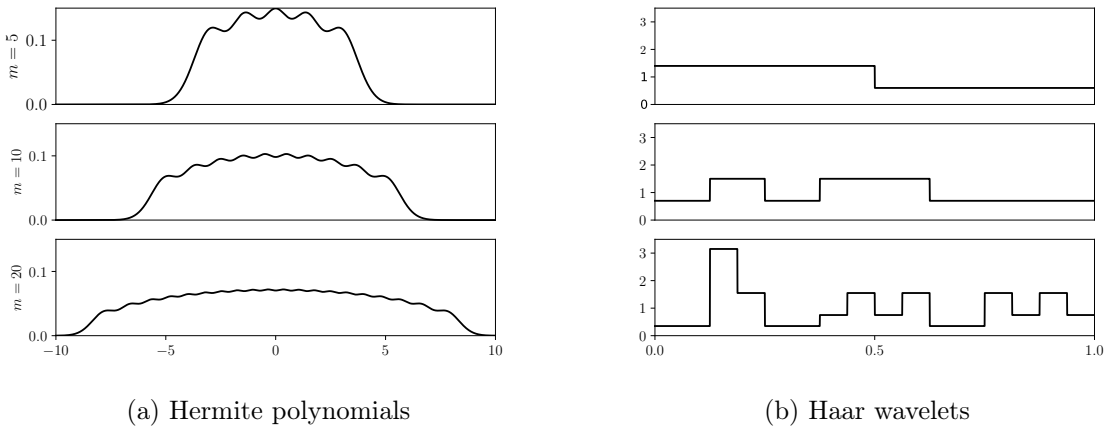


Figure 1: Sampling densities μ_m for (a) Hermite polynomials of degrees $0, \dots, m-1$, (b) subsets of Haar wavelet basis selected by random tree refinement.

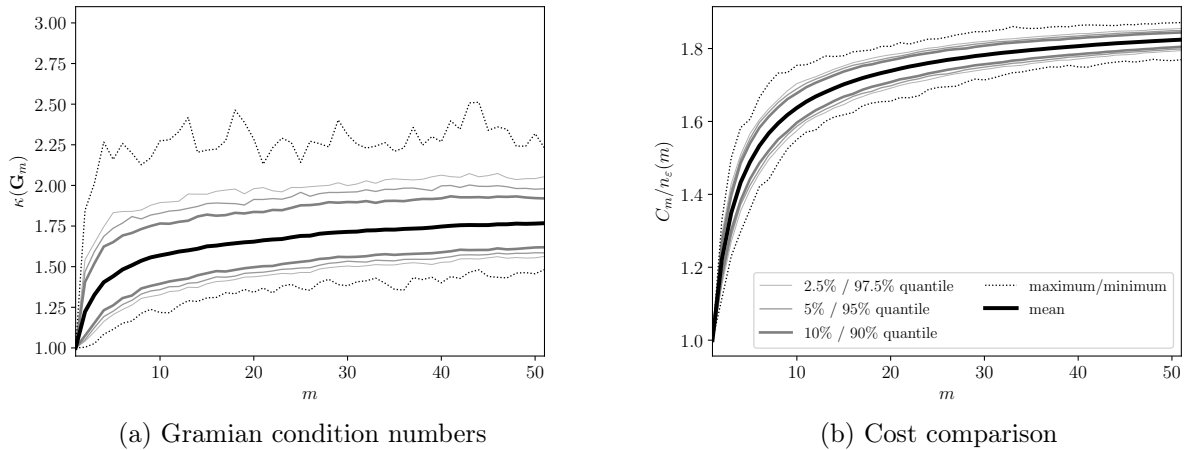


Figure 2: Results for Algorithm 1 with $n(m) = n_\epsilon(m)$ as in (3.6), $\epsilon = 10^{-2}$, applied to Hermite polynomials of degrees $0, \dots, m-1$.

where g is the standard Gaussian density and Φ its cumulative distribution function. The bisection thus takes $\mathcal{O}(j \log \epsilon)$ operations to converge to accuracy ϵ . For practical purposes, the sampling for case (i) can be accelerated by precomputation of interpolants for σ_j .

In each of our numerical tests, we consider the distributions of $\kappa(\mathbf{G}_m)$ and C_m resulting from the sequential generation of sample sets S_m for $m = 1, \dots, 50$ by Algorithm 1. In each case, the quantiles of the corresponding distributions are estimated from 1000 test runs. Figure 2 shows the results for Algorithm 1 in case (i) with $n(m) = n_\epsilon(m)$ as in (3.6), where we choose $\epsilon = 10^{-2}$. We know from Theorem 2.1 that, for each m , one has $\kappa(\mathbf{G}_m) \leq 3$ with probability greater than $1 - \epsilon$. In fact, this bound appears to be fairly pessimistic, since no sample with $\kappa(\mathbf{G}_m) > 3$ is encountered in the present test. In Figure 2(b), we show the ratio $C_m/n_\epsilon(m)$. Recall that C_m is defined in 3.5 as the total number of samples used in the repeated application of Algorithm 1 to produce S_1, \dots, S_m , and thus $C_m/n_\epsilon(m)$ provides a comparison to the costs of directly sampling S_m from μ_m . The results are in agreement with Theorems 3.2 and 3.4, which show in particular that $C_m < 2n_\epsilon(m)$ with very high probability.

Using the same setup with $n(m) = n_{\epsilon(m)}(m)$ as in (3.22), where $\epsilon_0 = 10^{-2}$, leads to the

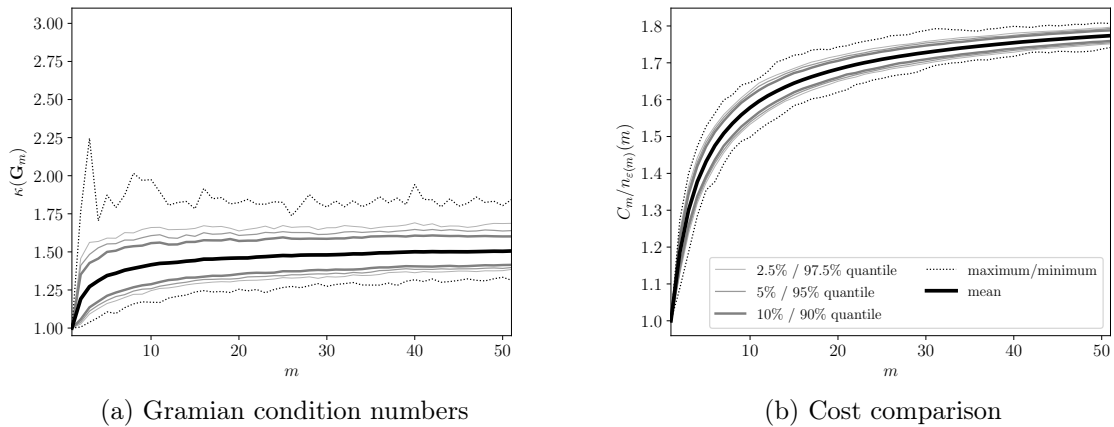


Figure 3: Results for Algorithm 1 with $n(m) = n_{\varepsilon(m)}(m)$ as in (3.22), $\varepsilon_0 = 10^{-2}$, applied to Hermite polynomials of degrees $0, \dots, m-1$.

expected improved uniformity in $\kappa(\mathbf{G}_m)$. The corresponding results are shown in Figure 3, with sampling costs that are in agreement with (3.26) and Theorem 3.6. Since the effects of replacing $n_{\varepsilon}(m)$ by $n_{\varepsilon(m)}(m)$ are very similar in all further tests, we only show results for $n(m) = n_{\varepsilon}(m)$ with $\varepsilon = 10^{-2}$ in what follows.

While the simple scheme in Algorithm 1 already ensures near-optimal costs with high probability, there are some practical variants that can yield better quantitative performance. A first such variant is given in Algorithm 2. Instead of deciding for each previous sample separately whether it will be re-used, here a queue of previous samples is kept, from which these are extracted in order until the previous sample set S_m is exhausted. Clearly, the costs of this scheme are bounded from above by those of Algorithm 1.

Algorithm 2 Sequential sampling with sample queue

input: sample $S_m = \{x_m^1, \dots, x_m^{n(m)}\}$ from μ_m
output: sample $S_{m+1} = \{x_{m+1}^1, \dots, x_{m+1}^{n(m+1)}\}$ from μ_{m+1}

```

j := 1
for i = 1, ..., n(m+1) do
  draw  $a_i$  uniformly distributed in  $\{1, \dots, m+1\}$ 
  if  $a_i = m+1$  then
    draw  $x_{m+1}^i$  from  $\sigma_{m+1}$ 
  else if  $j \leq n(m)$  then
     $x_{m+1}^i := x_m^j$ 
     $j \leftarrow j + 1$ 
  else
    draw  $x_{m+1}^i$  from  $\mu_m$  by (3.3)
  end if
end for

```

As expected, in the results for Algorithm 2 applied in case (i), which are shown in Figure 4, we find an estimate of the distribution of $\kappa(\mathbf{G}_m)$ that is essentially identical to the one for Algorithm 1 in Figure 2(a). The costs, however, are substantially more favorable than the ones in Figure 2(b): using Algorithm 2, the successive sampling of S_1, \dots, S_m uses only a small fraction

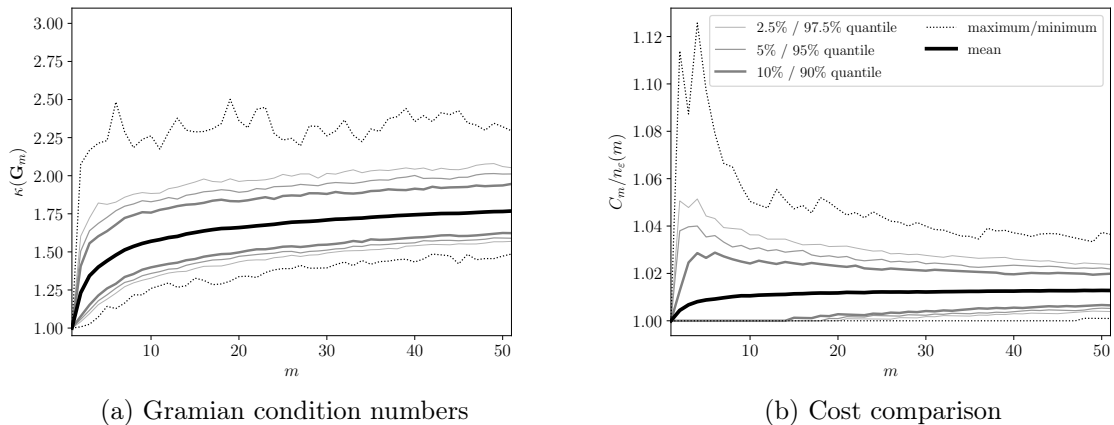


Figure 4: Results for Algorithm 2 with $n(m) = n_\varepsilon(m)$ as in (3.6), $\varepsilon = 10^{-2}$, applied to Hermite polynomials of degrees $0, \dots, m - 1$.

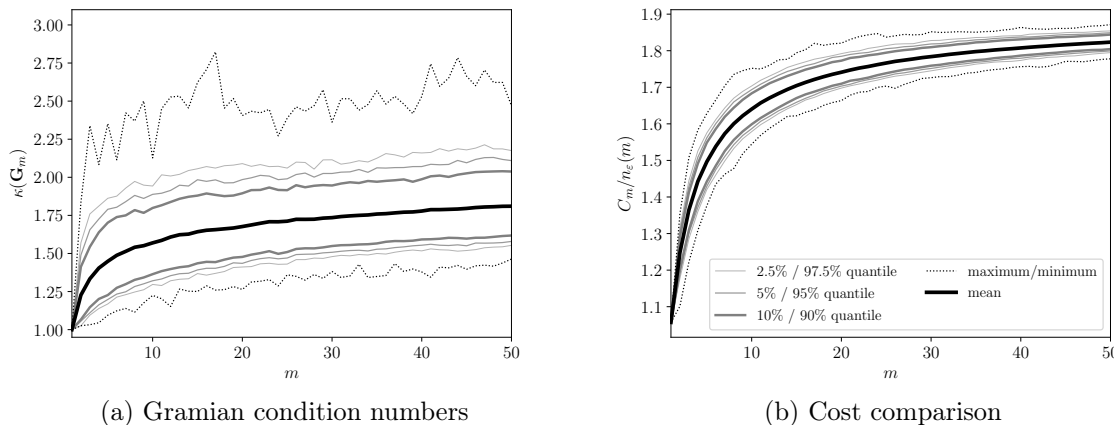


Figure 5: Results for Algorithm 1 with $n(m) = n_\varepsilon(m)$ as in (3.6), $\varepsilon = 10^{-2}$, applied to subset of Haar basis obtained by random tree refinement.

of additional samples when compared to directly sampling only S_m .

Figures 5 and 6 show the analogous comparison of Algorithms 1 and 2 applied to case (ii), which leads to very similar results. This is not surprising, considering that the bounds in the general Theorem 2.1 on optimal least squares sampling, as well as those in §3, are all independent of the chosen L^2 -space and of the corresponding orthonormal basis. Our numerical results thus indicate that one can indeed expect a rather minor effect of this choice also in practice.

A further algorithmic variant consists in applying the inner loop of Algorithm 2 until one is ensured that the stability criterion $\|\mathbf{G}_m - \mathbf{I}\| \leq \frac{1}{2}$ is met, so that in particular $\kappa(\mathbf{G}_m) \leq 3$ holds with certainty. This procedure is described in Algorithm 3. Note that here the size \hat{n}_m of the sample S_m is not fixed a priori.

Since the stability criterion is ensured with certainty by this third Algorithm, we only need to study the total sampling cost C_m . This is illustrated in Figure 7 in the case (i) of Hermite polynomials. For a direct comparison to Algorithms 1 and 2, we compare the costs to $n_\varepsilon(m)$ as before, although this value plays no role in Algorithm 3. Figure 7(a) shows that with high probability, one has $C_m/n_\varepsilon(m) < 1$ for the considered range of m and $\varepsilon = 10^{-2}$, although this

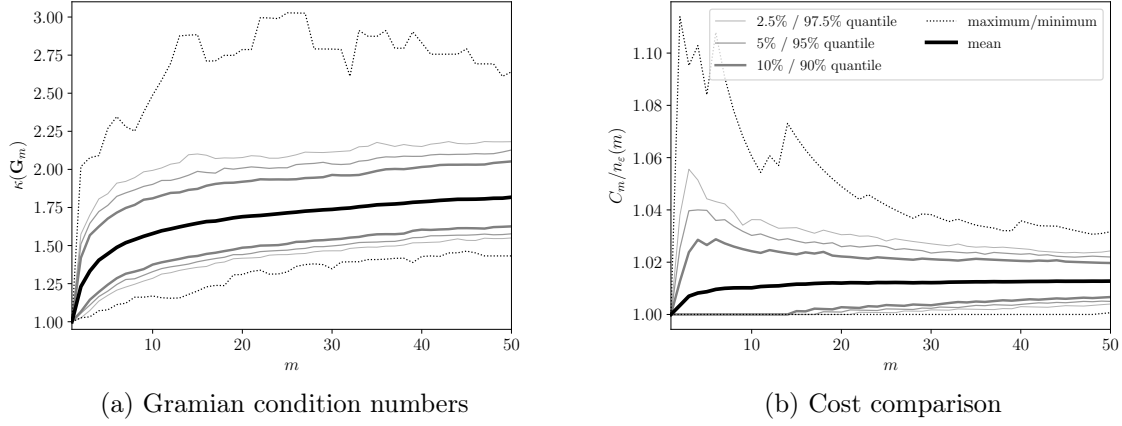


Figure 6: Results for Algorithm 2 with $n(m) = n_\varepsilon(m)$ as in (3.6), $\varepsilon = 10^{-2}$, applied to subset of Haar basis obtained by random tree refinement.

Algorithm 3 Sequential sampling with guaranteed condition number

input: sample $S_m = \{x_m^1, \dots, x_m^{\hat{n}_m}\}$ from μ_m

output: sample $S_{m+1} = \{x_{m+1}^1, \dots, x_{m+1}^{\hat{n}_{m+1}}\}$ from μ_{m+1}

$j := 1, i := 1, \lambda := 1$

repeat

draw a_i uniformly distributed in $\{1, \dots, m+1\}$

if $a_i = m+1$ **then**

draw x_{m+1}^i from σ_{m+1}

else if $j \leq \hat{n}_m$ **then**

$x_{m+1}^i := x_m^j$

$j \leftarrow j+1$

else

draw x_{m+1}^i from μ_m by (3.3)

end if

if $i \geq m+1$ **then**

Assemble $\tilde{\mathbf{G}}_{m+1}^i$ according to (2.4) using $\{x_{m+1}^1, \dots, x_{m+1}^i\}$

$\lambda \leftarrow \|\tilde{\mathbf{G}}_{m+1}^i - \mathbf{I}\|$

end if

$i \leftarrow i+1$

until $\lambda \leq \frac{1}{2}$

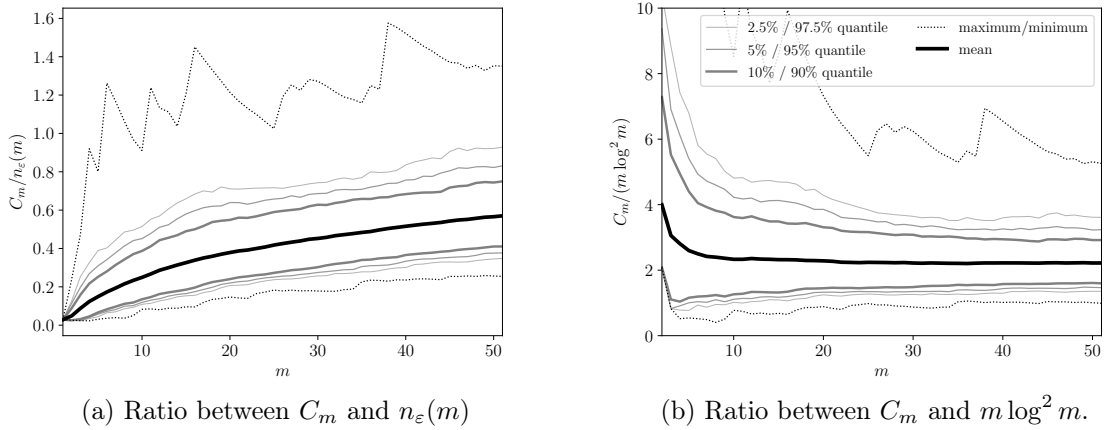


Figure 7: Results for Algorithm 3 applied to Hermite polynomials of degrees $0, \dots, m-1$.

ratio can be seen to increase approximately logarithmically. A closer inspection shows that C_m tends to behave like $m \log^2 m$, as illustrated on Figure 7(b). This hints that an extra logarithmic factor is needed for ensuring stability with certainty for all values of m .

References

- [1] D. Angluin and L. G. Valiant, *Fast Probabilistic algorithms for Hamiltonian circuits and mappings*, Journal of Computer and System Sciences **18**, 155-193, 1979.
- [2] A. Chkifa, A. Cohen, G. Migliorati, F. Nobile, and R. Tempone, *Discrete least squares polynomial approximation with random evaluations - application to parametric and stochastic PDEs*, M2AN , **49**, 815-837, 2015.
- [3] A. Cohen *Numerical analysis of wavelet methods*, Studies in mathematics and its applications, Elsevier, Amsterdam, 2003
- [4] A. Cohen and R. DeVore, *Approximation of high-dimensional PDEs*, Acta Numerica, **24**, 1-159, 2015.
- [5] A. Cohen, R. DeVore, and C. Schwab, *Analytic regularity and polynomial approximation of parametric and stochastic PDEs*, Analysis and Applications, **9**, 11-47, 2011.
- [6] A. Cohen and G. Migliorati, *Optimal weighted least squares methods*, SMAI Journal of Computational Mathematics **3**, 181–203, 2017.
- [7] A. Cohen and G. Migliorati, *Multivariate approximation in downward closed polynomial spaces*, to appear in Ian Sloan’s 80 Festschrift, 2018.
- [8] W. Dahmen *Wavelet and multiscale methods for operator equations*, Acta Numer., 55-228, 1997.
- [9] R. DeVore *Nonlinear approximation*, Acta Numer., 51–150, 1998.
- [10] A. Doostan and J. Hampton, *Coherence motivated sampling and convergence analysis of least squares polynomial Chaos regression*, Computer Methods in Applied Mechanics and Engineering **290**, 73-97, 2015.

- [11] A. Doostan and M. Hadigol, *Least squares polynomial chaos expansion: A review of sampling strategies*, Computer Methods in Applied Mechanics and Engineering, **332**, 382-407, 2018.
- [12] A. Doostan and J. Hampton, *Basis Adaptive Sample Efficient Polynomial Chaos (BASE-PC)*, preprint, arXiv:1702.01185, 2017
- [13] T. Erdélyi, A.P. Magnus and P. Nevai, *Generalized Jacobi weights, Christoffel functions, and Jacobi polynomials*, SIAM Journal on Mathematical Analysis, **25**, 602-614, 1994.
- [14] A.L. Haji-Ali, F. Nobile, R. Tempone and S. Wölfer *Multilevel weighted least squares polynomial approximation*, preprint, arXiv:1707.00026, 2017
- [15] T. Hagerup and C. Rüb, *A guided tour of Chernoff bounds*, Information Processing Letters **33**, 305–308, 1990.
- [16] J.D. Jakeman, A. Narayan, and T. Zhou, *A Christoffel function weighted least squares algorithm for collocation approximations*, Math. Comp., **86**, 1913-1947, 2017.
- [17] G. Mastroianni and V. Totik, *Weighted polynomial inequalities with doubling and a_∞ weights*, Constructive approximation **16**, 37-71, 2000.
- [18] J. Tropp, *User-Friendly tail bounds for sums of random matrices*, Found. Comput. Math. **12**, 389-434, 2012.