

# REDUCTION OF MULTIVARIATE MIXTURES AND ITS APPLICATIONS

GREGORY BEYLKIN, LUCAS MONZÓN AND XINSHUO YANG

ABSTRACT. We consider fast deterministic algorithms to identify the “best” linearly independent terms in multivariate mixtures and use them to compute, up to a user-selected accuracy, an equivalent representation with fewer terms. Importantly, the multivariate mixtures do not have to be a separated representation of a function. One algorithm employs a pivoted Cholesky decomposition of the Gram matrix constructed from the terms of the mixture to select what we call skeleton terms and the other uses orthogonalization for the same purpose. Both algorithms require  $\mathcal{O}(r^2N + p(d)rN)$  operations, where  $N$  is the initial number of terms in the multivariate mixture,  $r$  is the number of selected linearly independent terms, and  $p(d)$  is the cost of computing the inner product between two terms of a mixture in  $d$  variables. For general Gaussian mixtures  $p(d) \sim d^3$  since we need to diagonalize a  $d \times d$  matrix, whereas for separated representations  $p(d) \sim d$  (there is no need for diagonalization). Due to conditioning issues, the resulting accuracy is limited to about one half of the available significant digits for both algorithms. We also describe an alternative algorithm that is capable of achieving higher accuracy but is only applicable in low dimensions or to multivariate mixtures in separated form.

We describe a number of initial applications of these algorithms to solve partial differential and integral equations and to address several problems in data science. For data science applications in high dimensions, we consider the kernel density estimation (KDE) approach for constructing a probability density function (PDF) of a cloud of points, a far-field kernel summation method and the construction of equivalent sources for non-oscillatory kernels (used in both, computational physics and data science) and, finally, show how to use the new algorithm to produce seeds for subdividing a cloud of points into groups.

## 1. INTRODUCTION

We present (what we call) reduction algorithms for computing with multivariate mixtures that allow us to obtain solutions of PDEs in high dimensions as well as to address several problems in data science. We describe a new approach for solving partial differential and integral equations in a functional form, consider a far-field kernel summation method and the construction of equivalent sources for non-oscillatory kernels. As an illustration of data science applications, we present examples of using these reduction algorithms for kernel density estimation (KDE) to construct a probability density function (PDF) of a cloud of points and to generate seeds for subdividing a cloud of points into groups.

We use the well-known pivoted Cholesky factorization<sup>1</sup> as well as a version of modified Gram-Schmidt orthogonalization to identify the “best” linearly independent terms from a collection of functions. The renewed interest in this problem is due to two observations: (i) the approach can be used for more general multivariate mixtures than the separated representations in [8, 9] and (ii) multivariate Gaussian mixtures can achieve any target accuracy when approximating functions since a multiresolution analysis can employ a Gaussian as an approximate scaling function [13]. The first observation makes our approach so far the only choice for reduction of general multivariate mixtures while the second assures that a multivariate Gaussian mixture (and its modifications that e.g. include polynomial factors) is sufficient to represent an arbitrary function while allowing us to exploit the fact that integrals involving Gaussian mixtures can be evaluated explicitly. We expand further on these observations below.

---

<sup>1</sup>Pivoted Cholesky decomposition of the Gram matrix was used by Martin Mohlenkamp (Ohio University) and G.B. for the reduction of the number of terms in separated representations (see comments in [15]).

We consider multivariate functions that can be approximated via a linear combination of multivariate atoms,

$$(1.1) \quad u(\mathbf{x}) = \sum_{l=1}^N c_l g_l(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d,$$

such that the inner product between the atoms,

$$(1.2) \quad \langle g_l, g_{l'} \rangle = \int_{\mathbb{R}^d} g_l(\mathbf{x}) g_{l'}(\mathbf{x}) d\mathbf{x},$$

can be computed efficiently. We normalize the atoms so that they have unit  $L^2$ -norm

$$\|g_l\|_2 = \sqrt{\langle g_l, g_l \rangle} = 1.$$

A particularly important example are multivariate Gaussian atoms yielding a multivariate Gaussian mixture. In this case the atoms are

$$(1.3) \quad \begin{aligned} g_l(\mathbf{x}, \mu_l, \Sigma_l) &= (\det(4\pi\Sigma_l))^{-\frac{1}{4}} N(\mathbf{x}, \mu_l, \Sigma_l) \\ &= \frac{1}{(\det(\pi\Sigma_l))^{1/4}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu_l)^T \Sigma_l^{-1} (\mathbf{x} - \mu_l)\right), \end{aligned}$$

where  $N(\mathbf{x}, \mu_l, \Sigma_l)$  is the standard multivariate Gaussian distribution with mean  $\mu_l$ , symmetric positive definite covariance matrix  $\Sigma_l \in \mathbb{R}^{d \times d}$ , and  $\|g_l\|_2 = 1$ . Already in early quantum chemistry computations, Gaussian mixtures were used because integrals involving them can be computed explicitly (see e.g. [18, 35, 44]). Indeed, integrals with Gaussians (and (1.2) in particular), can be evaluated explicitly in any dimension. Multivariate Gaussian mixtures (as well as other multivariate atoms) are used within the Radial Basis Functions (RBF) approach (see e.g. [23] and references therein). A method for approximating smooth functions via Gaussians and a number of its applications have been developed in [37].

As was demonstrated previously, a number of key operators of mathematical physics can be efficiently represented via Gaussians leading to their separated representations (see e.g. [8, 9, 11, 12, 6]) and, as a consequence, to practical algorithms (see [29, 47, 48, 28]). Importantly, as it was demonstrated recently, for any finite but arbitrary accuracy a Gaussian can serve as a scaling function of an approximate multiresolution analysis [13].

These considerations combined with algorithms of this paper to reduce the number of terms in (1.1) suggest a new type of numerical algorithms. The basic idea of such algorithms is simple: in the process of iteratively solving equations, we represent both *operators and functions* via Gaussians and, at each iteration step, compute the required integrals explicitly. The difficulty of this approach is then how to deal with a rapid proliferation of terms in the resulting Gaussian mixtures. For example, if an integral involves three Gaussian mixtures with 100 terms each, the resulting Gaussian mixture has  $10^6$  terms. However, in most practical applications most of these terms are close to be linearly dependent and, thus, to represent the result, we only need a fast algorithm to find the “best” linearly independent subset of the terms. We describe reduction algorithms to maintain a reasonable number of terms in intermediate computations when using these representations.

The multivariate mixtures that we consider (and construct algorithms for) can be significantly more general than the separated representations introduced in [8, 9] for the purpose of computing in higher dimensions by avoiding “the curse of dimensionality”. Recall that a separated representation is a natural extension of the usual separation of variables as we seek an approximation

$$(1.4) \quad f(x_1, \dots, x_d) = \sum_{l=1}^r s_l \phi_1^{(l)}(x_1) \cdots \phi_d^{(l)}(x_d) + \mathcal{O}(\varepsilon),$$

where the functions  $\phi_j^{(l)}(x_j)$  are normalized by the standard  $L^2$ -norm,  $\|\phi_j^{(l)}\|_2 = 1$  and  $s_l > 0$  are referred to as  $s$ -values. In this approximation the functions  $\phi_j^{(l)}(x_j)$  are not fixed in advance but are optimized as to achieve the accuracy goal with (ideally) a minimal separation rank  $r$ . Importantly, a separated representation is not a projection onto a subspace, but rather a nonlinear method to track a function in a high-dimensional space while using a small number of parameters. The key to obtaining useful separated representations is to use the Alternating Least Squares (ALS) algorithm to reduce the separation rank while maintaining an acceptable error. ALS is one of the key tools in numerical multilinear algebra and was originally introduced for data fitting as PARAFAC model (PARAllel FACtor analysis) [30] and CANDECOMP (Canonical Tensor Decomposition) [20]. ALS has been used extensively in data analysis of (mostly) three-way arrays (see e.g. the reviews [45, 19], [32] and references therein). We note that any discretization of  $f$  in (1.4) leads to a  $d$ -dimensional tensor  $\mathcal{U} \in \mathbb{R}^{M_1 \times \dots \times M_d}$  yielding a canonical tensor decomposition of separation rank  $r$ ,

$$(1.5) \quad \mathcal{U}_{i_1, \dots, i_d} = \sum_{l=1}^r \sigma_l \prod_{j=1}^d u_{i_j}^{(l)},$$

where the  $s$ -values  $\sigma_l$  are chosen so that each vector  $\mathbf{u}_j^{(l)} = \left\{ u_{i_j}^{(l)} \right\}_{i_j=1}^{M_j}$  has unit Frobenius norm  $\|\mathbf{u}_j^{(l)}\|_F = 1$  for all  $j, l$ . However, the ALS algorithm relies heavily on the separated form (1.4-1.5) and is not available for general multivariate mixtures.

In this paper we detail and use algorithms to reduce the number of terms in multivariate mixtures that do not necessarily admit a separated representation. In spite their accuracy limitations, these algorithms have several advantages as their complexity depends mildly on the number of variables, the dimension  $d$ . The ‘‘mild’’ dependence should be understood in the context of the ‘‘curse of dimensionality’’ arising when fast algorithms in low dimensions are extended to high dimensions in a straightforward manner. Fast reduction algorithms of this paper require  $\mathcal{O}(r^2N + p(d)rN)$  operations, where  $N$  is the initial number of terms in a multivariate mixture,  $r$  is the number of selected terms and  $p(d)$  is the cost of computing the inner product between two terms of the mixture. We also describe an algorithm capable of achieving higher accuracy by avoiding the loss of precision due to conditioning issues but which currently is only applicable in low dimensions or to multivariate mixtures in separated form.

We start by describing reduction algorithms in Section 2. We then present several examples of using a reduction algorithm for solving equations in Section 3. In Section 4 we show how to apply our approach to construct the PDF for a cloud of points in high dimensions using the KDE approach. We then turn to far-field summation in high dimensions in Section 5 and demonstrate how to use a reduction algorithm in this problem and for the problem of finding equivalent sources. We also describe how to use a reduction algorithm for splitting a cloud of points into groups (potentially in a hierarchical manner) and briefly mention properties of such subdivisions. Conclusions are presented in Section 6 and some key identities for multivariate Gaussians in the Appendix.

## 2. REDUCTION ALGORITHMS

In this section we describe fast deterministic algorithms to reduce the number of terms of a linear combination of multivariate functions of  $d$  variables by selecting the ‘‘best’’ subset of these functions that can, within a target accuracy, approximate the rest of them. The first algorithm is based on a pivoted Cholesky decomposition of the Gram matrix of the terms of the multivariate mixture; we assume that the entries of this matrix, i.e. the inner product of these functions, can be computed efficiently. This algorithm was mentioned in the discussion of tensor interpolative decomposition (tensor ID) of the canonical tensor representation in [15]. Due to the use of a Gram matrix, the accuracy of this approach is limited to about one half of the available significant digits (e.g.  $10^{-7} \sim 10^{-8}$  when using double precision arithmetic). The second algorithm is based on a pivoted modified Gram-Schmidt orthogonalization and

achieves the same accuracy (due to conditioning issues) as the first algorithm. Nevertheless, these algorithms appear advantageous in high dimensions since their complexities depend mildly on dimension and, if desired, full accuracy can be restored by performing some evaluations in higher precision. We also describe an alternative approach that achieves full precision, but so far is limited to low dimensions or mixtures in separated form. For this alternative reduction algorithm, we need access to the Fourier transform of the functions in the mixture. Since the Fourier transform is readily available for Gaussian atoms, we present this algorithm for the case of Gaussian mixtures and note that it can be used for any functional form that allows a rapid computation of the integrals involved.

**2.1. Cholesky reduction.** We start with a linear combination of atoms of the form

$$(2.1) \quad u(\mathbf{x}) = \sum_{l=1}^N c_l g_l(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d$$

and, within a user-selected accuracy  $\varepsilon$ , seek a representation of the same form but with fewer terms. More precisely, we look for a partition of indices  $I = [\widehat{I}, \widetilde{I}]$ , where  $\widehat{I} = [i_1, i_2, \dots, i_r]$  and  $\widetilde{I} = [i_{r+1}, i_{r+2}, \dots, i_N]$ , and new coefficients  $\tilde{c}_{i_m}, m = 1, \dots, r$ , such that the function

$$(2.2) \quad \tilde{u}(\mathbf{x}) = \sum_{m=1}^r \tilde{c}_{i_m} g_{i_m}(\mathbf{x}), \quad r \ll N,$$

approximates  $u$ ,

$$(2.3) \quad u(\mathbf{x}) \approx \tilde{u}(\mathbf{x}).$$

We present an algorithm based on a partial, pivoted Cholesky decomposition of the Gram matrix constructed using the atoms  $g_l$  in (2.1) and provide an estimate for the error in (2.3).

By analogy with the matrix Interpolative Decomposition (matrix-ID) (see e.g. [27]), we call the subset  $\{g_{i_m}\}_{m=1}^r$  the skeleton terms and  $\{g_{i_m}\}_{m=r+1}^N$  the residual terms. In order to identify the “best” subset of linear independent terms, we compute a pivoted Cholesky decomposition of the Gram matrix of the atoms of the multivariate mixture in (2.1). If the number of terms,  $N$ , is large then the cost of the full Cholesky decomposition is prohibitive. However, we show that we can terminate the Cholesky decomposition once the pivots are below a selected threshold. As a result, the complexity of the algorithm is  $\mathcal{O}(r^2N + p(d)rN)$ , where  $N$  is the initial number of terms,  $r$  is the number of selected (skeleton) terms and  $p(d)$  is the cost of computing the inner product between two terms of a mixture in  $d$  variables. In fact, the final result will be the same as if we were to perform the full decomposition and then keep only the significant terms. This property is a consequence of the following lemma that can be found in e.g. [31, p.434, problem 7.1.P1].

**Lemma 1.** *Let  $\mathbf{B} \in \mathbb{C}^{n \times n}$  be positive semi-definite and self-adjoint, i.e.  $\mathbf{x}^* \mathbf{B} \mathbf{x} \geq 0$  and  $\mathbf{B} = \mathbf{B}^*$  for any  $\mathbf{x} \in \mathbb{C}^n$ . Then its diagonal entries  $b_{ii}$  are non-negative and the entries  $b_{ij}$  of  $\mathbf{B}$  satisfy*

$$(2.4) \quad |b_{ij}| \leq \sqrt{b_{ii} b_{jj}}.$$

*In particular, assuming that the first  $i$  diagonal entries are in descending order and are greater or equal than the remaining diagonal entries,*

$$b_{11} \geq b_{22} \geq \dots \geq b_{ii} \geq b_{i+1, i+1}, b_{i+2, i+2}, \dots, b_{nn},$$

*we have*

$$(2.5) \quad |b_{ij}| \leq b_{ii}, \text{ for all } j \geq i.$$

*Proof.* Let  $\{\mathbf{e}_i\}_{1 \leq i \leq n}$  be the standard basis vectors, that is,  $(\mathbf{e}_i)_j = \delta_{ij}$ . The diagonal entries are non-negative since, for any index  $i$ ,  $b_{ii} = \mathbf{e}_i^* \mathbf{B} \mathbf{e}_i \geq 0$ . We now use the same approach to estimate the size of an off-diagonal entry  $b_{ij} = |b_{ij}| e^{i\theta_{ij}}$ . For the vector  $\mathbf{x} = x_i \mathbf{e}_i + x_j \mathbf{e}_j$  we have

$$(2.6) \quad 0 \leq \mathbf{x}^* \mathbf{B} \mathbf{x} = b_{ii} x_i \bar{x}_i + b_{ij} x_i \bar{x}_j + \overline{b_{ij}} \bar{x}_i x_j + b_{jj} x_j \bar{x}_j.$$

Setting

$$x_i = \left(\frac{b_{jj}}{b_{ii}}\right)^{1/4} \quad \text{and} \quad x_j = -e^{i\theta_{ij}} \left(\frac{b_{ii}}{b_{jj}}\right)^{1/4}$$

in (2.6), we obtain

$$0 \leq b_{ii} \left(\frac{b_{jj}}{b_{ii}}\right)^{1/2} - |b_{ij}| - |b_{ij}| + b_{jj} \left(\frac{b_{ii}}{b_{jj}}\right)^{1/2} = 2\sqrt{b_{ii}b_{jj}} - 2|b_{ij}|.$$

Thus, we arrive at

$$(2.7) \quad |b_{ij}| \leq \sqrt{b_{ii}b_{jj}} \leq \frac{b_{ii} + b_{jj}}{2}.$$

For the second part of the lemma, selecting  $i \leq j$  implies that  $b_{ii} \geq b_{jj}$  and, thus, (2.5) follows from the last inequality in (2.7).  $\square$

Lemma 1 implies

**Corollary 2.** *Let  $G$  be a self-adjoint positive semi-definite matrix such that its Cholesky decomposition has monotonically decaying diagonal entries. If we write its Cholesky decomposition as*

$$G = \begin{pmatrix} L_r & 0 \\ W & Q \end{pmatrix} \begin{pmatrix} L_r^* & W^* \\ 0 & Q^* \end{pmatrix},$$

where  $L_r$  is an  $r \times r$  lower triangular matrix with the smallest diagonal entry  $\varepsilon > 0$ , then a partial Cholesky decomposition is of the form

$$G = \begin{pmatrix} L_r & 0 \\ W & 0 \end{pmatrix} \begin{pmatrix} L_r^* & W^* \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & QQ^* \end{pmatrix},$$

where all entries of the matrix  $QQ^*$  are less than  $\varepsilon$ .

*Proof.* After applying  $r$  steps of Cholesky decomposition, the matrix  $W$  does not change in the consecutive steps. The remaining matrix  $QQ^*$  is self-adjoint positive semi-definite and, due to the decay of the pivots, all of its diagonal entries are less than  $\varepsilon$ . Using Lemma 1, we conclude that all entries of  $QQ^*$  are less than  $\varepsilon$ .  $\square$

Let us organize the collection of atoms in (2.1) as

$$A = [g_1(\mathbf{x}), g_2(\mathbf{x}), \dots, g_N(\mathbf{x})].$$

We can view  $A$  as a matrix with a gigantic number of rows resulting from a discretization of the argument  $\mathbf{x} \in \mathbb{R}^d$ . If we replace operations that require row-wise summation by the inner product between the atoms, then we can (and do) use matrix notation in the sequel. Without loss of generality, to simplify notation, we assume that the first  $r$  atoms in  $A$  form the skeleton, that is,  $A = (A_s | A_{ns})$ , where  $A_s$  denotes the  $r$  skeleton atoms and  $A_{ns}$  the  $N - r$  non-skeleton atoms.

Given the vector of coefficients  $c = [c_1, c_2, \dots, c_N]^T$  in (2.1), we want to find new coefficients  $\tilde{c} = [\tilde{c}_1, \tilde{c}_2, \dots, \tilde{c}_r]^T$  to approximate  $u(\mathbf{x})$  by

$$(2.8) \quad \tilde{u}(\mathbf{x}) = \sum_{l=1}^r \tilde{c}_l g_l(\mathbf{x}), \quad r \ll N,$$

and estimate the error  $\|u - \tilde{u}\|_2$  of the approximation. Note that we identify  $u = Ac$  and  $\tilde{u} = A_s \tilde{c}$ .

We first seek an approximation of all atoms via the skeleton atoms,

$$g_k(\mathbf{x}) \approx \sum_{i=1}^r p_{ik} g_i(\mathbf{x}), \quad k = 1, \dots, N.$$

Selecting the coefficients  $p_{ik}$  as the solutions of the least squares problem,  $p_{ik}$  satisfy the normal equations,

$$(2.9) \quad \sum_{i=1}^r p_{ik} \langle g_i, g_{i'} \rangle = \langle g_k, g_{i'} \rangle, \quad k = 1, \dots, N, \quad i' = 1, 2, \dots, r.$$

Introducing the matrix  $P = \{p_{ik}\}_{\substack{i=1, \dots, r \\ k=1, \dots, N}}$ , we write (2.9) as

$$(2.10) \quad A_s^* A_s P = A_s^* A$$

and observe that  $P = (I_r | S)$ , where  $I_r$  is the  $r \times r$  identity matrix and  $S$ , an  $(N-r) \times r$  matrix, which satisfies

$$(2.11) \quad A_s^* A_s S = A_s^* A_{ns}.$$

Setting

$$\tilde{c}_i = \sum_{k=1}^N p_{ik} c_k,$$

or

$$\tilde{c} = Pc,$$

we obtain from (2.10) that the coefficients  $\tilde{c}$  solve the system of normal equations

$$(2.12) \quad A_s^* A_s \tilde{c} = A_s^* A c.$$

**Theorem 3.** *Let the Gram matrix  $G$ ,*

$$G = \begin{pmatrix} A_s^* A_s & A_s^* A_{ns} \\ A_{ns}^* A_s & A_{ns}^* A_{ns} \end{pmatrix},$$

*be such that its partial Cholesky decomposition has monotonically decaying pivots and is of the form*

$$G = \begin{pmatrix} L_r & 0 \\ W & 0 \end{pmatrix} \begin{pmatrix} L_r^* & W^* \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & QQ^* \end{pmatrix}$$

*where  $L_r$  is an  $r \times r$  lower triangular matrix with the smallest diagonal entry  $\varepsilon > 0$ . If the coefficients of the skeleton terms are computed via (2.12), then the difference between  $u$  in (2.1) and its approximation (2.8) can be estimated as*

$$(2.13) \quad \|u - \tilde{u}\|_2 = \|A_s \tilde{c} - A c\|_2 \leq \|c\|_2 \sqrt{N-r} \varepsilon^{1/2}.$$

*Proof.* We have

$$(2.14) \quad \|A_s \tilde{c} - A c\|_2^2 = \|A_s P c - A c\|_2^2 = \langle c, (P^* A_s^* - A^*) (A_s P - A) c \rangle,$$

where the coefficient matrix  $P$  solves the normal equations (2.10) and, therefore,

$$(2.15) \quad A_s^* (A_s P - A) = 0,$$

as well as

$$(2.16) \quad A_s^* (A_s S - A_{ns}) = 0.$$

Using (2.15), we obtain

$$(2.17) \quad (P^* A_s^* - A^*) (A_s P - A) = A^* A - A^* A_s P$$

and proceed to compute  $A^* A_s P$ . We have  $A^* = \begin{pmatrix} A_s^* \\ A_{ns}^* \end{pmatrix}$  and  $A_s P = (A_s | A_s S)$ , so that

$$A^* A_s P = \begin{pmatrix} A_s^* A_s & A_s^* A_s S \\ A_{ns}^* A_s & A_{ns}^* A_s S \end{pmatrix}.$$

Thus, we have

$$A^*A - A^*A_sP = \begin{pmatrix} 0 & A_s^*(A_{ns} - A_sS) \\ 0 & A_{ns}^*(A_{ns} - A_sS) \end{pmatrix}.$$

and, using (2.16), arrive at

$$(2.18) \quad A^*A - A^*A_sP = \begin{pmatrix} 0 & 0 \\ 0 & A_{ns}^*A_{ns} - A_{ns}^*A_sS \end{pmatrix}.$$

Equating the two expressions of the Gram matrix in the statement of the Theorem,

$$G = \begin{pmatrix} A_s^*A_s & A_s^*A_{ns} \\ A_{ns}^*A_s & A_{ns}^*A_{ns} \end{pmatrix} = \begin{pmatrix} L_r & 0 \\ W & 0 \end{pmatrix} \begin{pmatrix} L_r^* & W^* \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & QQ^* \end{pmatrix},$$

we obtain that

$$(2.19) \quad A_s^*A_s = L_rL_r^*,$$

$$(2.20) \quad A_{ns}^*A_s = WL_r^*, \quad A_sA_{ns}^* = L_rW^*,$$

and

$$(2.21) \quad A_{ns}^*A_{ns} = WW^* + QQ^*.$$

We observe that from (2.11) using (2.19) and (2.20), we arrive at

$$(2.22) \quad L_rL_r^*S = L_rW^*.$$

Next we show that the non-zero block of the matrix in the right hand side of (2.18) coincides with  $QQ^*$ . Using (2.20) and (2.22), we have

$$\begin{aligned} A_{ns}^*A_sS &= WL_r^*S \\ &= W(L_r^{-1}L_r)L_r^*S \\ &= WL_r^{-1}(L_rL_r^*S) \\ &= WL_r^{-1}(L_rW^*) = WW^*, \end{aligned}$$

where we used that  $L_r$  is non-singular. Hence, combining the last identity with (2.21), we obtain

$$(2.23) \quad A_{ns}^*A_{ns} - A_{ns}^*A_sS = WW^* + QQ^* - WW^* = QQ^*.$$

By (2.14), (2.17), (2.18), and (2.23) we obtain

$$\begin{aligned} \|A_s\tilde{c} - Ac\|_2^2 &\leq \|c\|_2^2 \|(P^*A_s^* - A^*)(A_sP - A)\|_2 \\ &= \|c\|_2^2 \|A_{ns}^*A_{ns} - A_{ns}^*A_sS\|_2 \\ &= \|c\|_2^2 \|QQ^*\|_2. \end{aligned}$$

Using Corollary 2, we estimate  $\|QQ^*\|_2$  by its Frobenius norm,

$$\|QQ^*\|_2 \leq \|QQ^*\|_F \leq (N - r) \varepsilon$$

which yields the desired estimate (2.13).  $\square$

*Remark 4.* The estimate (2.13) is tighter than the one obtained in [15, Theorem 3.1] since the dependence on the number of terms is  $\mathcal{O}(N^{1/2})$  instead of  $\mathcal{O}(N^{3/4})$ . Yet, the estimate is still pessimistic since  $\|QQ^*\|_2$  is usually significantly smaller than the Frobenius norm  $\|QQ^*\|_F$ . In practice, for  $N$  in the range  $10^5 - 10^6$ , we did not observe the reduction of accuracy suggested by the factor  $\mathcal{O}(N^{1/2})$ .

We present pseudo-code for reduction as Algorithm 1. Note that this algorithm is dimension independent except for the cost of computing the inner product which we always assume to be reasonable by a judicious choice of the functions in the mixture. As a consequence of Theorem 3 and Lemma 1, it is sufficient to generate only  $N \times r$  entries of the Cholesky decomposition of the Gram matrix  $G$  which requires  $p(d)rN$  operations, where  $p(d)$  is the cost of computing the inner product between two terms of a mixture. Therefore, the overall computational cost of Algorithm 1 is  $\mathcal{O}(r^2N + p(d)rN)$  operations (see Table 2 for some examples). For a general Gaussian mixture  $p(d)$  is proportional to  $d^3$  since we need to diagonalize a  $d \times d$  matrix in order to evaluate the explicit expression for the inner product in Appendix 7.0.4 whereas, for separated representations, no diagonalization is needed so that  $p(d) \sim d$ .

**2.2. Reduction via a rank-revealing modified Gram-Schmidt algorithm.** For ordinary matrices, Rank Revealing QR (RRQR) algorithms (see e.g. [22, 26] and references therein) select a set of linear independent columns that one can use to represent the other columns (within a certain accuracy). All matrix RRQR algorithms routinely use vector addition whereas, in our problem, the sum of two or more terms of multivariate mixture does not simplify and, therefore, a sum must to be maintained as a linear combination. Moreover, as far as we know, there is no analogue of the Householder reflection or Givens rotation (as tools for orthogonalization) unless we discretize the argument  $\mathbf{x} \in \mathbb{R}^d$  of the terms  $\{g_i(\mathbf{x})\}_{i=1}^N$  (note that for reduction in low dimensions we use a discretization of the Fourier transform of the terms of the mixture as described in Section 2.3). As long as we rely on the rapid evaluation of inner products, the reduction algorithm mimics the modified Gram-Schmidt (MGS) orthogonalization with the caveat that the sum of terms is maintained as a linear combination. Unfortunately, this algorithm (even for ordinary matrices) loses one half of significant digits (see [16] and discussion in [25, Section 5.2.9]) as does Algorithm 1 (it remains an open question if there exists a modification of MGS for our problem that maintains full accuracy). As we show below, the reduction algorithm via orthogonalization for multivariate mixtures has formal complexity  $\mathcal{O}(r^3 + r^2N + p(d)rN)$ , which is somewhat worse than the complexity of the reduction via Cholesky decomposition. However, since in a typical reduction  $N \gg r$ , such algorithm can be considered to be of the same complexity as Algorithm 1.

To introduce notation, we consider a set of multivariate atoms  $\{g_i(\mathbf{x})\}_{i=1}^N$  and write

$$\begin{aligned}
 \tilde{\psi}_1(\mathbf{x}) &= g_1(\mathbf{x}), & \psi_1(\mathbf{x}) &= \frac{\tilde{\psi}_1(\mathbf{x})}{\|\tilde{\psi}_1\|} \\
 \tilde{\psi}_2(\mathbf{x}) &= g_2(\mathbf{x}) - t_{21}\tilde{\psi}_1(\mathbf{x}), & \psi_2(\mathbf{x}) &= \frac{\tilde{\psi}_2(\mathbf{x})}{\|\tilde{\psi}_2\|} \\
 \tilde{\psi}_3(\mathbf{x}) &= g_3(\mathbf{x}) - t_{31}\tilde{\psi}_1(\mathbf{x}) - t_{32}\tilde{\psi}_2(\mathbf{x}), & \psi_3(\mathbf{x}) &= \frac{\tilde{\psi}_3(\mathbf{x})}{\|\tilde{\psi}_3\|} \\
 & \vdots & & \vdots \\
 \tilde{\psi}_N(\mathbf{x}) &= g_N(\mathbf{x}) - \left( \sum_{j=1}^N t_{Nj}\tilde{\psi}_j(\mathbf{x}) \right), & \psi_N(\mathbf{x}) &= \frac{\tilde{\psi}_N(\mathbf{x})}{\|\tilde{\psi}_N\|}
 \end{aligned}
 \tag{2.24}$$

where the functions  $\tilde{\psi}_i$  are orthogonal,  $\psi_i$  are orthonormal, and

$$t_{ij} = \frac{\langle \tilde{\psi}_j, g_i \rangle}{\langle \tilde{\psi}_j, \tilde{\psi}_j \rangle}, \quad 1 \leq j < i, \quad i = 1, \dots, N.$$



**Algorithm 1** Reduction algorithm using a Gram matrix

**Inputs:** Atoms  $g_l$  and coefficients  $c_l$  in the representation of  $u(\mathbf{x}) = \sum_{l=1}^N c_l g_l(\mathbf{x})$  and error tolerance,  $\varepsilon$ ,  $10^{-14} \leq \varepsilon < 1$ . We assume that a subroutine to compute the inner product  $\langle g_l, g_m \rangle$  is available.

**Outputs:** A pivot vector  $I = [\widehat{I}, \widetilde{I}]$ , where  $\widehat{I} = [i_1, \dots, i_r]$  contains indices of  $r$  skeleton terms and  $\widetilde{I} = [i_{r+1}, \dots, i_N]$  indices of terms being removed from the final representation and the coefficients  $\widetilde{c}_m, m = 1, \dots, r$ , such that  $\|u - \sum_{m=1}^r \widetilde{c}_m g_{i_m}\|_2 \leq \|c\|_2 \sqrt{N-r} \varepsilon^{1/2}$  due to Theorem 3.

**Stage 1: Pivoted Cholesky decomposition of the Gram matrix.****Initialization.**

We maintain the diagonal  $[d_1, d_2, \dots, d_N]$  of the Cholesky factor  $L$  separately and initialize it as  $[1, 1, \dots, 1]$  (since all atoms have unit  $L^2$ -norm).

Set  $r = 0$  and initialize a pivot vector as  $I = [1, 2, \dots, N]$ .

**for**  $l = 1, N$

(1) Find the largest element of the diagonal and its index  $i_j = \{i_j : d_{i_j} \geq d_{i_k}, k = l, \dots, N\}$

**if**  $d_{i_j} < \varepsilon$  **goto** **Stage 2**

(2) Swap indices  $i_j$  and  $i_l$  in the pivot vector  $I$ .

(3) Set the diagonal element of the matrix  $L_{i,l} = (d_{i_l})^{1/2}$

**for**  $j = l + 1, N$

$$L_{i,j} = \left( \langle g_{i_j}, g_{i_l} \rangle - \sum_{k=1}^{l-1} L_{i,k} L_{i_j,k} \right) / L_{i,l}$$

$$d_{i_j} = d_{i_j} - L_{i,j}^2$$

**end**

update:  $r = r + 1$

**end**

**Stage 2: Find new coefficients**  $\widetilde{c}_m, m = 1, \dots, r$ 

(1) Form a vector  $b$  such that its  $j$ -th element is the inner product of  $\sum_{i_m \in \widetilde{I}} c_{i_m} g_{i_m} = \sum_{m=r+1}^N c_{i_m} g_{i_m}$  and  $g_{i_j}$ .

**for**  $j = 1, r$

$$b(j) = \sum_{m=r+1}^N c_{i_m} \left( \sum_{k=1}^j L_{i_j,k} L_{i_m,k} \right)$$

**end**

(2) Solve the linear system  $\widehat{G}\widetilde{c} = b$ , where  $\widehat{G}_{jl} = \langle g_{i_j}, g_{i_l} \rangle = \sum_{k=1}^N L_{i_j,k} L_{i_l,k}$  and  $\widetilde{c} = [\widetilde{c}_1, \widetilde{c}_1, \dots, \widetilde{c}_r]^T$  using forward and backward substitution.

(3) Add the original coefficients of the skeleton terms  $c_{i_m}$  to  $\widetilde{c}_{i_m}$  to get the new coefficients

**for**  $m = 1, r$

$$\widetilde{c}_{i_m} = \widetilde{c}_{i_m} + c_{i_m}$$

**end**

We can also rewrite (2.24) as

$$\begin{aligned}
 (2.25) \quad g_1(\mathbf{x}) &= r_{11} \psi_1(\mathbf{x}), \\
 g_2(\mathbf{x}) &= r_{21} \psi_1(\mathbf{x}) + r_{22} \psi_2(\mathbf{x}), \\
 g_3(\mathbf{x}) &= r_{31} \psi_1(\mathbf{x}) + r_{32} \psi_2(\mathbf{x}) + r_{33} \psi_3(\mathbf{x}), \\
 &\vdots \\
 g_N(\mathbf{x}) &= \sum_{j=1}^N r_{Nj} \psi_j(\mathbf{x}),
 \end{aligned}$$

where  $r_{ij} = t_{ij} \|\tilde{\psi}_j\|$  and  $r_{ii} = \|\tilde{\psi}_i\|$ ,  $i = 1, \dots, N$ . We can consider a set of coefficients  $s_{ij}$ ,  $1 \leq j \leq i \leq N$  such that

$$(2.26) \quad \begin{aligned} \psi_1(\mathbf{x}) &= s_{11}g_1(\mathbf{x}), \\ \psi_2(\mathbf{x}) &= s_{21}g_1(\mathbf{x}) + s_{22}g_2(\mathbf{x}), \\ \psi_3(\mathbf{x}) &= s_{31}g_1(\mathbf{x}) + s_{32}g_2(\mathbf{x}) + s_{33}g_3(\mathbf{x}), \\ &\vdots \\ \psi_N(\mathbf{x}) &= \sum_{j=1}^N s_{Nj}g_j(\mathbf{x}). \end{aligned}$$

To compute the coefficients  $r_{ij}$  in (2.25) and  $s_{ij}$  in (2.26), we need to evaluate the inner product between the multivariate atoms  $g_i$  and  $g_j$  which we denote as  $G_{ij} = \langle g_i, g_j \rangle$ .

In order to estimate the computational cost, let us assume that the first  $k-1$  steps have been accomplished so that the coefficients  $r_{ij}$ ,  $1 \leq i \leq N$ ,  $1 \leq j \leq k-1$  and  $j \leq i$  in (2.25), the coefficients  $s_{ij}$ ,  $1 \leq j \leq i \leq k$  in (2.26), the norms of  $\tilde{\psi}_i = g_i - \sum_{j=1}^{k-1} r_{ij}\psi_j$ ,  $i = 1, \dots, N$  as well as the partial Gram matrix  $G_{ij}$ ,  $1 \leq i \leq N$ ,  $1 \leq j \leq k-1$  and  $j \leq i$  are already available. We then select the term,  $\tilde{\psi}_i = g_i - \sum_{j=1}^{k-1} r_{ij}\psi_j$ ,  $i = k, \dots, N$  with the largest norm which we then swap to become the term with index  $k$ . For simplicity of indexing we assume that such term was already in the position  $k$ . Thus, our pivoting strategy uses a greedy algorithm by selecting terms with the largest norm (we note that alternative pivoting strategies may be possible but we do not explore them here).

Note that the norms  $\|\tilde{\psi}_i\| = \left\| g_i - \sum_{j=1}^{i-1} r_{ij}\psi_j \right\| = r_{ii}$ ,  $i = 1, \dots, k$  do not change in the steps that follow. At the  $k$ -th step, we first need to compute the coefficients in  $s_{kj}$ ,  $j = 1, \dots, k$  such that

$$\psi_k(\mathbf{x}) = \sum_{j=1}^k s_{kj}g_j(\mathbf{x})$$

as in (2.26). Since the norm  $r_{kk} = \|\tilde{\psi}_k\|$  is available we use (2.24) to evaluate

$$\begin{aligned} r_{kk}\psi_k(\mathbf{x}) &= g_k(\mathbf{x}) - \sum_{i=1}^{k-1} r_{ki}\psi_i(\mathbf{x}) = g_k(\mathbf{x}) - \sum_{i=1}^{k-1} r_{ki} \left( \sum_{j=1}^i s_{ij}g_j(\mathbf{x}) \right) \\ &= g_k(\mathbf{x}) - \sum_{j=1}^{k-1} \left( \sum_{i=j}^{k-1} r_{ki}s_{ij} \right) g_j(\mathbf{x}) = - \sum_{j=1}^{k-1} \left( \sum_{i=j}^{k-1} r_{ki}s_{ij} \right) g_j(\mathbf{x}) + g_k(\mathbf{x}) \end{aligned}$$

and, therefore, obtain

$$(2.27) \quad s_{kj} = -\frac{1}{r_{kk}} \sum_{i=j}^{k-1} r_{ki}s_{ij}, \quad 1 \leq j \leq k-1, \quad s_{kk} = \frac{1}{r_{k,k}}.$$

Computing the coefficients  $s_{kj}$  take  $\mathcal{O}(k^2)$  operations at this step. Next, we compute the coefficients in (2.25) and update the term  $\tilde{\psi}_i = g_i - \sum_{j=1}^{k-1} r_{ij}\psi_j$  by subtracting  $r_{ik}\psi_k$  and evaluate the resulting norms for  $i = k+1, \dots, N$ . Specifically, for  $i = k+1, \dots, N$ , we compute  $G_{ik} = \langle g_i, g_k \rangle$ , and

$$(2.28) \quad r_{ik} = \langle g_i, \psi_k \rangle = \left\langle g_i, \sum_{j=1}^k s_{kj}g_j \right\rangle = \sum_{j=1}^k s_{kj} \langle g_i, g_j \rangle = \sum_{j=1}^k s_{kj}G_{ij}.$$

We update the norm of  $\tilde{\psi}_i$  as follows,

$$\begin{aligned}
(2.29) \quad \|\tilde{\psi}_i\|^2 &= \left\| g_i - \sum_{j=1}^{k-1} r_{ij} \Psi_j - r_{ik} \Psi_k \right\|^2 = \left\| g_i - \sum_{j=1}^{k-1} r_{ij} \Psi_j \right\|^2 + r_{ik}^2 - 2 \left\langle g_i - \sum_{j=1}^{k-1} r_{ij} \Psi_j, r_{ik} \Psi_k \right\rangle \\
&= \left\| g_i - \sum_{j=1}^{k-1} r_{ij} \Psi_j \right\|^2 + r_{ik}^2 - 2 \langle g_i, r_{ik} \Psi_k \rangle = \left\| g_i - \sum_{j=1}^{k-1} r_{ij} \Psi_j \right\|^2 + r_{ik}^2 - 2 r_{ik} \langle g_i, \Psi_k \rangle \\
&= \left\| g_i - \sum_{j=1}^{k-1} r_{ij} \Psi_j \right\|^2 - r_{ik}^2,
\end{aligned}$$

where we used (2.28). At this step, computing  $G_{ik}$ , (2.28) and (2.29) are performed for each index  $i$  so that it requires  $\mathcal{O}(p(d)N + kN)$  operations, where  $p(d)$  is the cost of computing the inner product between two terms of the mixture.

Once the skeleton terms  $\{g_i(\mathbf{x})\}_{i=1}^r$  are identified in the process of orthogonalization, we compute the coefficients  $\{\tilde{c}_i\}_{i=1}^r$  of the new representation of the multivariate mixture via these terms, i.e. approximate  $\sum_{i=1}^N c_i g_i(\mathbf{x})$  as  $\sum_{i=1}^r \tilde{c}_i g_i(\mathbf{x})$ . Using the fact that for  $j > r$

$$g_j(\mathbf{x}) \approx \sum_{k=1}^r r_{jk} \Psi_k(\mathbf{x}) = \sum_{k=1}^r r_{jk} \left( \sum_{i=1}^k s_{ki} g_i(\mathbf{x}) \right) = \sum_{i=1}^r \left( \sum_{k=i}^r r_{jk} s_{ki} \right) g_i(\mathbf{x}),$$

we compute

$$\begin{aligned}
\sum_{i=1}^N c_i g_i(\mathbf{x}) &= \sum_{i=1}^r c_i g_i(\mathbf{x}) + \sum_{j=r+1}^N c_j g_j(\mathbf{x}) \\
&\approx \sum_{i=1}^r c_i g_i(\mathbf{x}) + \sum_{j=r+1}^N c_j \left( \sum_{i=1}^r \left( \sum_{k=i}^r r_{jk} s_{ki} \right) g_i(\mathbf{x}) \right) \\
&= \sum_{i=1}^r c_i g_i(\mathbf{x}) + \sum_{i=1}^r \left( \sum_{j=r+1}^N c_j \left( \sum_{k=i}^r r_{jk} s_{ki} \right) \right) g_i(\mathbf{x}) \\
&= \sum_{i=1}^r \tilde{c}_i g_i(\mathbf{x})
\end{aligned}$$

so that

$$(2.30) \quad \tilde{c}_i = c_i + \sum_{j=r+1}^N c_j \left( \sum_{k=i}^r r_{jk} s_{ki} \right).$$

If the number of skeleton terms is  $r$  then, combining complexity estimates for all steps, the resulting algorithm has a complexity  $\mathcal{O}(r^3 + r^2N + p(d)rN)$ . Since this algorithm is designed to be used when  $N \gg r$ , we conclude that the overall cost is  $\mathcal{O}(r^2N + p(d)rN)$ . Pseudo-code for this reduction algorithm is presented as Algorithm 2.

**2.3. Alternative reduction algorithms.** Using Algorithm 1 or 2, half of the significant digits are lost due to poor conditioning (see examples in [15] and [41]). In order to identify “best” linear independent terms we can design a matrix with a better condition number if instead of the functions of the mixture we use a “dual” family for computing inner products. In the case of Gaussians (which are well localized), a natural set of such “dual” functions are exponentials with purely imaginary exponents (which are global functions); computing the inner product with them reduces to computing their Fourier transform. Therefore, as representatives of Gaussian atoms we can then use frequency vectors, i.e. samples of their Fourier transforms. Such sampling strategy should be sufficient to differentiate between all Gaussian atoms; it is fairly straightforward to achieve this in low dimensions or if the functions admit a separated

**Algorithm 2** Reduction algorithm using modified Gram-Schmidt orthogonalization

**Inputs:** Atoms  $g_l$  and coefficients  $c_l$  in the representation of  $u(\mathbf{x}) = \sum_{l=1}^N c_l g_l(\mathbf{x})$  and error tolerance,  $\varepsilon$ ,  $10^{-14} \leq \varepsilon < 1$ . We assume that a subroutine to compute the inner product  $\langle g_l, g_{l'} \rangle$  is available.

**Outputs:** The pivot vector  $I = [\hat{I}, \tilde{I}]$ , where  $\hat{I} = [i_1, \dots, i_r]$  contains indices of  $r$  skeleton terms and  $\tilde{I} = [i_{r+1}, \dots, i_N]$  indices of terms being removed from the final representation and the coefficients  $\tilde{c}_{i_m}, m = 1, \dots, r$ , such that  $|u(\mathbf{x}) - \sum_{m=1}^r \tilde{c}_{i_m} g_{i_m}(\mathbf{x})| = \mathcal{O}(\varepsilon^{1/2})$ .

**Stage 1: Pivoted MGS applied to the atoms**  $g_l, l = 1, \dots, N$ .

Set  $\tilde{\psi}_l(\mathbf{x}) = g_l(\mathbf{x})$  for  $l = 1, \dots, N$ . Note that, initially, all norms  $\|\tilde{\psi}_l\| = 1$  since all atoms have unit  $L^2$ -norm.

Set rank  $r = 0$  and initialize a pivot vector as  $I = [1, 2, \dots, N]$ .

**for**  $k = 1, N$

(1) Find the largest norm  $\|\tilde{\psi}_{i_j}\|$  and its index  $i_j = \{i_j : \|\tilde{\psi}_{i_j}\| \geq \|\tilde{\psi}_{i_l}\|, l = k, \dots, N\}$

**if**  $\|\tilde{\psi}_{i_j}\| < \varepsilon$  **goto** Stage 2

(2) Swap indices  $i_j$  and  $i_k$  in the pivot vector  $I$ .

(3) **set**  $r_{i_k, k} = \|\tilde{\psi}_{i_k}\|$

**for**  $j = 1, \dots, k-1$

$$s_{i_k, j} = -\frac{1}{r_{i_k, k}} \sum_{l=j}^{k-1} r_{i_k, l} s_{i_l, j}$$

**end**

$$s_{i_k, k} = \frac{1}{r_{i_k, k}}$$

**for**  $j = k+1, N$

$$G_{i_j, k} = \langle g_{i_j}, g_{i_k} \rangle$$

$$r_{i_j, k} = \sum_{l=1}^k s_{i_k, l} G_{i_j, l}$$

$$\text{update } \|\tilde{\psi}_{i_j}\| = \left( \|\tilde{\psi}_{i_j}\|^2 - r_{i_j, k}^2 \right)^{1/2}$$

**end**

update:  $r = r + 1$

**end**

**Stage 2: Find new coefficients**  $\tilde{c}_{i_m}, m = 1, \dots, r$ .

Compute new coefficients  $\tilde{c}_{i_m} = c_{i_m} + \sum_{j=r+1}^N c_{i_j} \left( \sum_{k=m}^r r_{i_j, k} s_{i_k, m} \right), m = 1, \dots, r$ .

representation. Currently, we do not know how to do it efficiently in high dimensions. Naively it appears to require the construction of a sample matrix with  $\mathcal{O}(N \times N d r)$  entries and additional work is required to understand how to lower this complexity. Alternatively, in dimensions  $d = 1, 2, 3$  it is sufficient to use  $\mathcal{O}(N \times r^d)$  samples if we were to use the straightforward generalization of the algorithm in dimension  $d = 1$  described below. In all cases, the last step in this approach is to compute the matrix ID of the sample matrix.

We present a deterministic algorithm in dimension  $d = 1$  and note that its extension to functions in separated form in high dimensions can follow the approach in [15]. We consider a univariate Gaussian mixture

$$(2.31) \quad u(x) = \sum_{l=1}^N c_l g_l(x), \quad x \in \mathbb{R},$$

where

$$g_l(x) = \frac{1}{\pi^{\frac{1}{4}} \sigma_l^{\frac{1}{2}}} e^{-\frac{(x-\mu_l)^2}{2\sigma_l^2}}, \quad \|g_l\|_2 = 1,$$

and seek the best linear independent subset as in (2.2). Defining the Fourier transform of  $f$  as

$$\hat{f}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ix\xi} dx,$$

we obtain

$$\hat{g}_l(\xi) = \frac{\sigma_l^{\frac{1}{2}}}{\pi^{\frac{1}{4}}} e^{-\frac{\sigma_l^2 \xi^2}{2}} e^{-i\mu_l \xi}.$$

We set the highest frequency  $\xi_{high}$  of  $\{\hat{g}_l(\xi)\}_{l=1}^N$  as

$$\xi_{high} = \sqrt{\frac{-2 \log \frac{\pi^{\frac{1}{4}} 10^{-16}}{\sigma_l^{\frac{1}{2}}}}{\sigma^2}}, \quad \sigma = \min_{l=1, \dots, N} \sigma_l,$$

such that  $|\hat{g}_l(\xi)| < 10^{-16}$  for  $\xi > \xi_{high}$  and  $l = 1, \dots, N$ . We also set the lowest frequency  $\xi_{low} = 10^{-2} \sim 10^{-3}$ , a positive value obtained experimentally. We then sample the interval  $[\xi_{low}, \xi_{high}]$  using frequencies  $\xi_k, k = 1, \dots, r_p$ , equally spaced on a logarithmic scale,

$$\xi_k = e^{\left( \log \xi_{low} + k \frac{\xi_{high} - \xi_{low}}{r_p} \right)},$$

where  $r_p$  is the number of samples. We choose  $r_p > r$ , where  $r$  is the expected final number of terms. In our setup, the column

$$\begin{bmatrix} \hat{g}_l(\xi_1) \\ \hat{g}_l(\xi_2) \\ \vdots \\ \hat{g}_l(\xi_{r_p}) \end{bmatrix}$$

serves as a representative of the Gaussian  $g_l$ . In this way, we reduce the problem to that of using the matrix ID. Specifically, given frequencies  $\xi_k, k = 1, \dots, r_p$ , we construct a  $r_p \times N$  sample matrix  $Y$ ,

$$Y = \begin{bmatrix} \hat{g}_1(\xi_1) & \hat{g}_2(\xi_1) & \cdots & \hat{g}_N(\xi_1) \\ \hat{g}_1(\xi_2) & \hat{g}_2(\xi_2) & \cdots & \hat{g}_N(\xi_2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{g}_1(\xi_{r_p}) & \hat{g}_2(\xi_{r_p}) & \cdots & \hat{g}_N(\xi_{r_p}) \end{bmatrix},$$

and compute its matrix ID (see e.g. [27]). We obtain a partition of indices  $I = [\hat{I}, \tilde{I}]$ , where  $\hat{I} = [i_1, \dots, i_r]$  and  $\tilde{I} = [i_{r+1}, \dots, i_N]$  denote the skeleton and residual terms respectively. We also obtain a matrix  $X$  such that

$$Y = Y_{[:, \hat{I}]} X,$$

where  $X$  is a  $r \times N$  matrix that satisfies  $X_{[:, \hat{I}]} = I_r$ . We then compute the new coefficients as

$$\tilde{c}_{i_m} = c_{i_m} + \sum_{n=r+1}^N c_{i_n} X_{mn}, \quad m = 1, 2, \dots, r,$$

and use them to approximate

$$u(x) \approx \sum_{m=1}^r \tilde{c}_{i_m} g_{i_m}(x).$$

**Algorithm 3** Reduction using frequencies (dimension  $d = 1$ )

**Inputs:** Atoms  $g_l$  and coefficients  $c_l$  in the representation of  $u(x) = \sum_{l=1}^N c_l g_l(x)$ , number of frequencies samples  $r_p$ , and error tolerance  $\varepsilon > 0$ . We assume that a subroutine to compute the Fourier transform  $\hat{g}_l(\xi)$  is available.

**Outputs:** an index set  $I = [\hat{I}, \tilde{I}]$ , where  $\hat{I} = [i_1, \dots, i_r]$  contains indices of  $r$  skeleton terms and  $\tilde{I} = [i_{r+1}, \dots, i_N]$  contains indices of terms being removed from the final representation, and coefficients  $\tilde{c}_{i_m}, m = 1, \dots, r$ , such that  $u(x) \approx \sum_{m=1}^r \tilde{c}_{i_m} g_{i_m}(x)$  with accuracy of the matrix ID.

- (1) set  $\xi_{low} \in [10^{-2}, 10^{-3}]$  and compute  $\xi_{high} = \sqrt{\frac{-2 \log \frac{\pi^4 10^{-16}}{\sigma^2}}{\sigma^2}}$  where  $\sigma = \min_{l=1, \dots, N} \{\sigma_l\}$ .
- (2) initialize  $\xi_k = e^{\left(\log \xi_{low} + k \frac{\xi_{high} - \xi_{low}}{r_p}\right)}, k = 1, \dots, r_p$ .
- (3) construct matrix  $Y$  such that  $Y_{kl} = \hat{g}_l(\xi_k)$ .
- (4) compute a matrix ID of  $Y$  to obtain a index partition  $[\hat{I}, \tilde{I}]$  and a matrix  $X$  such that  $Y = Y_{[:, \hat{I}]} X$
- (5) compute new coefficients via  $\tilde{c}_{i_m} = c_{i_m} + \sum_{n=r+1}^N c_{i_n} X_{mn}$ , for  $m = 1, \dots, r$ .

The accuracy of this approximation appears to be the same as the accuracy of matrix ID. Unfortunately, for multivariate Gaussian atoms, the size of the matrix  $Y$  appears to grow too fast with the dimension  $d$  (except in the case of separated representations where such dependence is linear). While our approach via frequency vectors can be extended in a straightforward manner to dimensions  $d = 2, 3$ , it is of interest to construct an algorithm yielding high accuracy approximations in higher dimensions.

**2.4. Timings and comparisons.** We compare the performance of Algorithms 1, 2 and 3 in dimension  $d = 1$  by considering a univariate Gaussian mixture of the form (2.31)<sup>2</sup>. We choose  $N = 10000$ , and sample  $c_l, \sigma_l$  and  $\mu_l$  from uniform distributions  $\mathcal{U}(-1, 1)$ ,  $\mathcal{U}(0, 0.5)$ , and  $\mathcal{U}(-5, 5)$ , respectively. We apply Algorithms 1, 2 and 3 to reduce the number of terms in the Gaussian mixture and display the original function and the errors of the resulting approximations in Figure 2.1.

We run this experiment with different accuracy thresholds and display the resulting number of skeleton terms obtained by these algorithms in Table 1 (we use Algorithms 1 and 2 implemented in quadruple precision to make comparison possible for higher accuracies). We observe that the number of skeleton terms obtained by these algorithms differs only slightly (this difference is irrelevant for intended applications).

We also time Algorithm 1 (implemented in double precision) and examine its scaling as a function of the number of initial terms and the number of skeleton terms  $r$  in dimensions  $d = 1$  and  $d = 5$ . In this test we fix the initial number of terms  $N$  and vary the number of skeleton terms  $r$  and vice versa. The resulting representation may not achieve a particular accuracy, but our the goal here is to see how Algorithm 1 scales in  $N$  and  $r$ . We consider a Gaussian mixture with parameters  $c_l$  and  $\mu_l$  sampled from uniform distributions  $\mathcal{U}(-1, 1)$  and  $\mathcal{U}(-25, 25)^d$ , respectively, and set matrices  $\Sigma_l = U_l D_l U_l^T$ , where  $U_l$  is a  $d \times d$  random unitary matrix and  $D_l$  is a  $d \times d$  diagonal matrix with positive entries sampled from the uniform distribution  $\mathcal{U}(0, 0.01)^d$ . We report dimensions  $d$ , the number of initial terms  $N$ , the number of skeleton terms  $r$  and the running time  $T$  in seconds in Table 2. We implemented all algorithms in Fortran90 and compile them with Intel Fortran Compiler version 18.0.3. The computations are performed on a single core (without parallelization) using Intel i7-6700 CPU @3.4 GHz on a 64-bit Linux workstation with 64 GB of RAM.

**2.5. Applications of reduction algorithms.** In the following sections, we present several examples of application of reduction algorithms in both low and high dimensions. In Section 3 we represent

<sup>2</sup>A test code is available upon request.

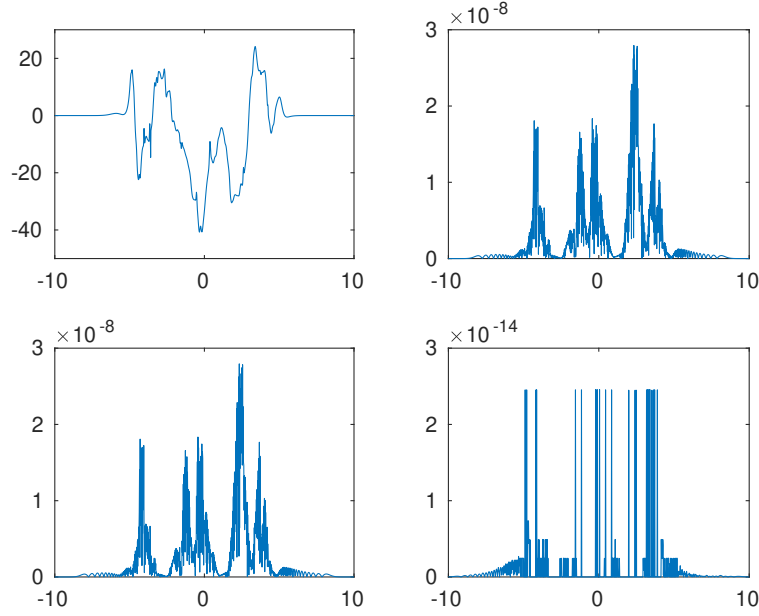


FIGURE 2.1. A function represented via a Gaussian mixture with  $10^4$  terms (top left). Relative errors obtained by using Algorithm 1 (requested accuracy  $10^{-7}$ ) yielding 300 terms (top right), by using Algorithm 2 (requested accuracy  $10^{-7}$ ) yielding 300 terms (bottom left) and by using Algorithm 3 (requested accuracy  $10^{-13}$ ) yielding 398 terms (bottom right).

Requested accuracy	Alg. 1	Actual accuracy	Alg. 2	Actual accuracy	Alg. 3	Actual accuracy
$10^{-3}$	216	$0.197e-3$	216	$0.197e-3$	217	$0.220e-3$
$10^{-5}$	261	$0.244e-5$	261	$0.244e-5$	259	$0.368e-5$
$10^{-7}$	300	$0.279e-7$	300	$0.279e-7$	297	$0.383e-7$
$10^{-9}$	330	$0.140e-9$	330	$0.140e-9$	330	$0.173e-9$
$10^{-11}$	361	$0.143e-11$	361	$0.143e-11$	361	$0.204e-11$
$10^{-13}$	390	$0.179e-13$	390	$0.179e-13$	398	$0.924e-14$

TABLE 1. Number of skeleton terms retained by Algorithms 1 and 2 (implemented in quadruple precision) and Algorithm 3 for different approximation accuracies. The requested accuracy is shown in the first column and (slightly different) resulting accuracies and the number of skeleton terms of the mixture are shown for each algorithm separately. In this example Algorithms 1 and 2 selected the same skeleton terms.

solutions of differential and integral equations in a functional form and adaptively solve these equations. We start with the free space Poisson's equation with non-separable right hand side and then present an example of solving an elliptic problem with variable coefficients; we consider both examples in dimensions  $d = 3$  through  $d = 7$ . In Section 4 we first use our algorithm in dimension  $d = 1$  to construct an efficient representation of the PDF of a cloud of points via kernel density estimation and compare it with results obtained via the usual approach. We then present an example of constructing PDFs in high

$d = 1$				$d = 5$			
Fixed $r = 10^2$		Fixed $N = 10^4$		Fixed $r = 10^2$		Fixed $N = 10^4$	
$N (\times 10^4)$	$T$	$r (\times 10^2)$	$T$	$N (\times 10^4)$	$T$	$r (\times 10^2)$	$T$
1	0.0960	1	0.0960	1	5.21	1	5.21
2	0.264	2	0.356	2	10.5	2	10.6
4	0.612	4	1.17	4	20.7	4	21.2
8	1.29	8	3.80	8	42.2	8	43.0
16	2.06	16	11.9	16	84.3	16	86.9

TABLE 2. Algorithm 1 timings  $T$  (in seconds) to be compared with the theoretical complexity estimate  $\mathcal{O}(r^2N + p(d)rN)$ . We observe the linear dependence on the number of initial terms  $N$ . Also, timings reveal that in dimension  $d = 5$  the cost of computing inner products dominates and, as a result, timing is practically linear as well. We note that the algorithm was implemented on a single core without any parallelization.

dimensions. Finally, we turn to kernel summation methods in high dimensions in Section 5, consider far-field evaluation in such computations and explore the problem of constructing equivalent sources in a similar setup. We also illustrate how a reduction algorithm can be used to partition points into groups (in a hierarchical fashion if desired).

### 3. REDUCTION ALGORITHMS FOR SOLVING DIFFERENTIAL AND INTEGRAL EQUATIONS

#### 3.1. Poisson equation in free space in high dimensions. The Poisson's equation

$$(3.1) \quad -\Delta u(\mathbf{x}) = f(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d,$$

arises in numerous applications in nearly all field of physics and computational chemistry (see e.g. [24]). The Reduction Algorithm 1 allows us to solve this equation in dimensions  $d \geq 3$  assuming that the charge distribution  $f(\mathbf{x})$  is given by a linear combination of multivariate Gaussian atoms. We obtain the solution via

$$(3.2) \quad u(\mathbf{x}) = \int_{\mathbb{R}^d} G(\mathbf{x} - \mathbf{y}) f(\mathbf{y}) d\mathbf{y},$$

where the free-space Green's function for (3.1) is given by the radial function

$$(3.3) \quad G(\mathbf{x}) = C_d \|\mathbf{x}\|^{2-d}, \quad C_d = \frac{\Gamma(\frac{d}{2} + 1)}{d(d-2)\pi^{\frac{d}{2}}},$$

where  $\|\cdot\| = \|\cdot\|_2$  is the standard  $l^2$ -norm. In order to evaluate the integral (3.2), we approximate the Green's function  $G$  via a linear combination of Gaussians (see e.g. [29, 11, 12]).

The error estimates in [12, Theorem 3] are based on discretizing the integral

$$\frac{1}{r^{d-2}} = \frac{1}{\Gamma(\frac{d-2}{2})} \int_{-\infty}^{\infty} e^{-r^2 e^t + \frac{d-2}{2} t} dt$$

as

$$(3.4) \quad G_{\infty}(r, h) = \frac{C_d h}{\Gamma(\frac{d-2}{2})} \sum_{l \in \mathbb{Z}} e^{hl(d-2)/2} e^{-e^{hl} r^2},$$

where the step size  $h$  satisfies

$$(3.5) \quad h \leq \frac{2\pi}{\log 3 + \frac{d-2}{2} \log(\cos 1)^{-1} + \log \varepsilon^{-1}}$$



and  $\varepsilon$  is any user-selected accuracy. Then [12, Theorem 3] implies

$$(3.6) \quad |G(r) - G_\infty(r, h)| \leq \varepsilon G(r), \quad \text{for all } r > 0.$$

To estimate the error of approximating the solution  $u$  in (3.2) using the series (3.4) instead of the Green's function (3.3), we first prove the following lemma.

**Lemma 5.** *For any  $d \geq 3$ ,  $e^{-1} \geq \varepsilon > 0$ , and  $f$  nonnegative in (3.1), there exist a step size  $h$  such that*

$$(3.7) \quad \left| u(\mathbf{x}) - \int_{\mathbb{R}^d} G_\infty(\|\mathbf{x} - \mathbf{y}\|, h) f(\mathbf{y}) d\mathbf{y} \right| \leq \varepsilon |u(\mathbf{x})|, \quad \text{for all } \mathbf{x} \neq 0.$$

*Proof.* From (3.2) and (3.6), we have

$$\begin{aligned} \left| u(\mathbf{x}) - \int_{\mathbb{R}^d} G_\infty(\|\mathbf{x} - \mathbf{y}\|, h) f(\mathbf{y}) d\mathbf{y} \right| &= \left| \int_{\mathbb{R}^d} [G(\mathbf{x} - \mathbf{y}) - G_\infty(\|\mathbf{x} - \mathbf{y}\|, h)] f(\mathbf{y}) d\mathbf{y} \right| \\ &\leq \int_{\mathbb{R}^d} |G(\mathbf{x} - \mathbf{y}) - G_\infty(\|\mathbf{x} - \mathbf{y}\|, h)| |f(\mathbf{y})| d\mathbf{y} \\ &\leq \varepsilon \int_{\mathbb{R}^d} G(\mathbf{x} - \mathbf{y}) |f(\mathbf{y})| d\mathbf{y} = \varepsilon \int_{\mathbb{R}^d} G(\mathbf{x} - \mathbf{y}) f(\mathbf{y}) d\mathbf{y} \\ &= \varepsilon u(\mathbf{x}) = \varepsilon |u(\mathbf{x})|. \end{aligned}$$

□

In our examples, we always consider functions  $f$  represented in the form

$$(3.8) \quad f(\mathbf{x}) = \sum_{l=1}^N c_l g_l(\mathbf{x}, \mu_l, \Sigma_l),$$

for some Gaussian atoms as in (1.3). In particular,

$$u(\mathbf{x}) = \sum_{l=1}^N c_l \int_{\mathbb{R}^d} G(\mathbf{x} - \mathbf{y}) g_l(\mathbf{x}, \mu_l, \Sigma_l) d\mathbf{y}$$

and

$$u^+(\mathbf{x}) = \sum_{l=1}^N |c_l| \int_{\mathbb{R}^d} G(\mathbf{x} - \mathbf{y}) g_l(\mathbf{x}, \mu_l, \Sigma_l) d\mathbf{y}$$

are both bounded. We assume that

$$(3.9) \quad \|u^+\|_{L^\infty} \leq c \|u\|_{L^\infty}$$

for a moderate size constant  $c$ . This assumption prevents representations of  $u$  that involve large coefficients  $c_l$  of opposite signs. We then have

**Lemma 6.** *Let  $d \geq 3$ ,  $e^{-1} \geq \varepsilon > 0$ , and  $f$  as in (3.8). If (3.9) holds, then there exist a step size  $h$  such that*

$$(3.10) \quad \left| \int_{\mathbb{R}^d} G(\mathbf{x} - \mathbf{y}, h) f(\mathbf{y}) d\mathbf{y} - \int_{\mathbb{R}^d} G_\infty(\|\mathbf{x} - \mathbf{y}\|, h) f(\mathbf{y}) d\mathbf{y} \right| \leq \varepsilon c |u(\mathbf{x})|, \quad \text{for all } \mathbf{x} \neq 0.$$

*Proof.* It follows from Lemma 5 that

$$|u(\mathbf{x}) - \tilde{u}(\mathbf{x})| \leq \varepsilon u^+(\mathbf{x})$$

where

$$\tilde{u}(\mathbf{x}) = \int_{\mathbb{R}^d} G_\infty(\|\mathbf{x} - \mathbf{y}\|, h) f(\mathbf{y}) d\mathbf{y}.$$

Therefore, using (3.9), we have

$$\|u(\mathbf{x}) - \tilde{u}\|_{L^\infty} \leq \varepsilon u^+(\mathbf{x}) \leq \varepsilon c \|u\|_{L^\infty}.$$

□

In practice, when using (3.4), we truncate the sum

$$(3.11) \quad G_{M,N}(r) = \frac{C_d h}{\Gamma\left(\frac{d-2}{2}\right)} \sum_{l=M}^N e^{hl(d-2)/2} e^{-e^{hl} r^2} = \frac{C_d h}{\Gamma\left(\frac{d-2}{2}\right)} \sum_{l=1}^{N_{terms}} e^{h(M+l-1)(d-2)/2} e^{-e^{h(M+l-1)} r^2}$$

so that the removed terms contribute less than  $\varepsilon$  and we limit the range of  $r$  to some interval of the form  $[\delta, R]$ ; the resulting approximation has  $N_{terms} = N - M + 1$ . In our computations we set the range of  $r$  to be  $[10^{-10}, 10^{10}]$  and the accuracy to be  $\varepsilon = 10^{-14}$ . As shown in Table 3, the number of terms in (3.11) depends on the dimension only weakly.

As an illustration, we first demonstrate our approach for a single Gaussian,

$$(3.12) \quad f(\mathbf{x}) = e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}.$$

Using (3.11), we approximate the solution  $u$  by

$$(3.13) \quad \begin{aligned} u_\varepsilon(\mathbf{x}) &= \frac{C_d h}{\Gamma\left(\frac{d-2}{2}\right)} \sum_{l=1}^{N_{terms}} e^{h(M+l-1)(d-2)/2} \int_{\mathbb{R}^d} e^{-e^{h(M+l-1)} \|\mathbf{x}-\mathbf{y}\|^2} f(\mathbf{y}) d\mathbf{y} \\ &= \frac{C_d h}{\Gamma\left(\frac{d-2}{2}\right)} \sum_{l=1}^{N_{terms}} e^{h(M+l-1)(d-2)/2} \int_{\mathbb{R}^d} e^{-e^{h(M+l-1)} \|\mathbf{x}-\mathbf{y}\|^2} e^{-\frac{1}{2}(\mathbf{y}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{y}-\boldsymbol{\mu})} d\mathbf{y}. \end{aligned}$$

Evaluating the integral explicitly (see Appendix A for details), we obtain

$$(3.14) \quad u_\varepsilon(\mathbf{x}) = \sum_{l=1}^{N_{terms}} c_l g_l(\mathbf{x}, \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l)$$

where

$$c_l = \frac{C_d h \pi^{\frac{3d}{4}}}{\Gamma\left(\frac{d-2}{2}\right)} e^{-h(M+l-1)(d+2)/2} \frac{(\det \boldsymbol{\Sigma})^{\frac{1}{2}}}{(\det \boldsymbol{\Sigma}_l)^{\frac{1}{4}}},$$

$$\boldsymbol{\Sigma}_l = \boldsymbol{\Sigma} + \begin{bmatrix} \frac{1}{2e^{h(M+l-1)}} & & & 0 \\ & \frac{1}{2e^{h(M+l-1)}} & & \\ & & \ddots & \\ 0 & & & \frac{1}{2e^{h(M+l-1)}} \end{bmatrix},$$

and

$$\boldsymbol{\mu}_l = \boldsymbol{\mu}.$$

The number of terms in the representation of  $u_\varepsilon$  is excessive and we reduce it using Algorithm 1 to obtain our final approximation as

$$(3.15) \quad \tilde{u}(\mathbf{x}) = \sum_{m=1}^{\tilde{N}} \tilde{c}_{i_m} g_{i_m}(\mathbf{x}, \boldsymbol{\mu}_{i_m}, \boldsymbol{\Sigma}_{i_m}),$$

where  $\tilde{c}_{i_m}$  are the new coefficients and  $i_m \in \hat{I}$  (see Algorithm 1 for details).

*Remark 7.* The representation of the kernel in (3.11) can be obtained for a large spatial range since the number of terms  $N_{terms}$  is proportional to the logarithm of the range. For this reason our approach is viable in high dimensions while employing the Fast Fourier Transform is not an option due to the size of the Fourier domain, c.f. [46].

In order to demonstrate the performance of our approach, we choose the right hand side  $f$  to be a Gaussian mixture with 100 terms,

$$f(\mathbf{x}) = \sum_{i=1}^{100} c_{f_i} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{f_i})^T \boldsymbol{\Sigma}_{f_i}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{f_i})}.$$

In the Gaussian mixture  $f$ , the coefficients  $c_{f_i}$  and means  $\mu_{f_i}$  are sampled from a one and a  $d$ -dimensional standard normal distributions respectively. The symmetric positive definite matrices  $\Sigma_{f_i}$  are constructed as

$$(3.16) \quad \Sigma_{f_i} = U_i^T U_i + \frac{1}{10} I_d,$$

where  $U_i$  is a  $d \times d$  matrix of standard normally distributed numbers and  $I_d$  is the  $d \times d$  identity matrix. We obtain  $u_\varepsilon$  in (3.14) and apply Algorithm 1 to reduce the number of terms to obtain  $\tilde{u}$  in (3.15). The results are displayed in Table 4, where we show the dimension of the problem,  $d$ , the number of terms,  $N_{terms}$ , in the approximation of the Green's function, the number of terms,  $N_{tot}$ , in the solution  $u_\varepsilon$  before reduction and its accuracy, the number of terms,  $\tilde{N}$ , in the solution  $\tilde{u}$  after reduction and its accuracy, and, finally, the relative error between  $u_\varepsilon$  and  $\tilde{u}$ .

In order to estimate the accuracy of the solution  $\tilde{u}$  of (3.1), we define the errors

$$h_\varepsilon(\mathbf{x}) = -\Delta u_\varepsilon(\mathbf{x}) - f(\mathbf{x}), \quad \tilde{h}(\mathbf{x}) = -\Delta \tilde{u}(\mathbf{x}) - f(\mathbf{x}), \quad h(\mathbf{x}) = u_\varepsilon(\mathbf{x}) - \tilde{u}(\mathbf{x}).$$

The usual approach to ascertain the size of  $h_\varepsilon$ ,  $\tilde{h}$ , and  $h$  by evaluating them on a lattice of grid points is impractical in high dimensions. Instead, we compute values of  $h_\varepsilon$ ,  $\tilde{h}$  and  $h$  at a collection of points in principle directions of the right hand side  $f(\mathbf{x})$ . To be precise, we first solve the eigenvalue problem for all matrices  $\Sigma_{f_i}$ ,  $i = 1, \dots, 100$ ,

$$\Sigma_{f_i} = \sum_{j=1}^d \lambda_j^{(i)} \mathbf{v}_j^{(i)} \left( \mathbf{v}_j^{(i)} \right)^T.$$

Here the eigenvectors  $\mathbf{v}_j^{(i)}$  identify principle directions for each Gaussian in  $f$  so that we can select an appropriate set of samples along those directions. We note that the actual range of the eigenvalues of matrices  $\Sigma_{f_i}$ ,  $\left\{ \lambda_j^{(i)} \right\}_{i=1, \dots, 100, j=1, \dots, d}$ , is  $\left[ \frac{1}{10}, 40 \right]$ . Next, for each pair of  $\left\{ \lambda_j^{(i)}, \mathbf{v}_j^{(i)} \right\}$ , we find an interval  $\left[ -s_j^{(i)}, s_j^{(i)} \right]$  by solving

$$e^{-\frac{\left( s_j^{(i)} \right)^2}{2\lambda_j^{(i)}}} = 10^{-10} \Leftrightarrow s_j^{(i)} = \left( -2\lambda_j^{(i)} \log 10^{-10} \right)^{1/2}$$

and generate an equally-spaced grid in  $\left[ -s_j^{(i)}, s_j^{(i)} \right]$  as

$$s_{jk}^{(i)} = -s_j^{(i)} + (k-1) \frac{2s_j^{(i)}}{N_s - 1}, \quad k = 1, \dots, N_s.$$

Finally, we select sample points

$$\mathbf{x}_{jk}^{(i)} = s_{jk}^{(i)} \mathbf{v}_j^{(i)} + \mu_{f_i}$$

and evaluate  $h_\varepsilon(\mathbf{x}_{jk}^{(i)})$ ,  $\tilde{h}(\mathbf{x}_{jk}^{(i)})$  and  $h(\mathbf{x}_{jk}^{(i)})$  for  $i = 1, \dots, 100$ ,  $j = 1, \dots, d$  and  $k = 1, \dots, N_s$ . In our experiment, we set  $N_s = 10$ , and report the resulting errors in Table 3 and 4. For these two tables, we use the notation  $\|f\|_\infty = \max_{i,j,k} \left| f(\mathbf{x}_{jk}^{(i)}) \right|$ .

**3.2. Second order elliptic equation with a variable coefficient.** In this example, we consider the second order linear elliptic equation,

$$(3.17) \quad -\nabla \cdot (a(\mathbf{x}) \nabla u(\mathbf{x})) + k^2 u(\mathbf{x}) = f(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d$$

where  $d \geq 3$  and  $k > 0$ . We assume that the variable coefficient  $a(x)$  is of the form

$$(3.18) \quad a(\mathbf{x}) = 1 + e^{-\frac{1}{2}(\mathbf{x} - \mu_a)^T \Sigma_a^{-1} (\mathbf{x} - \mu_a)},$$

$d$	$N_{terms}$	$N_{tot}$	$\ h_\varepsilon\ _\infty / \ f\ _\infty$	$\tilde{N}$	$\ \tilde{h}\ _\infty / \ f\ _\infty$	$\ h\ _\infty / \ u_\varepsilon\ _\infty$
3	345	247	$1.2e-9$	183	$1.8e-8$	$4.0e-8$
4	397	295	$1.3e-9$	227	$1.2e-7$	$9.8e-8$
5	386	282	$5.6e-10$	207	$7.8e-8$	$1.7e-9$
6	343	279	$1.9e-10$	197	$1.8e-7$	$2.6e-9$
7	354	240	$1.7e-10$	154	$2.6e-7$	$3.9e-9$

TABLE 3. Number of terms and relative errors of solving Poisson's equation in dimensions  $d = 3, \dots, 7$  where the forcing term is a single randomly generated multivariate Gaussian. The number of Gaussians to represent the Green's function in (3.11) is  $N_{terms}$ , the number of terms of  $u_\varepsilon$  in (3.14) after truncation of coefficients to  $10^{-10}$  is  $N_{tot}$  and, after applying Algorithm 1, the number of terms of  $\tilde{u}$  in (3.15) is  $\tilde{N}$ .

$d$	$N_{terms}$	$N_{tot}$	$\ h_\varepsilon\ _\infty / \ f\ _\infty$	$\tilde{N}$	$\ \tilde{h}\ _\infty / \ f\ _\infty$	$\ h\ _\infty / \ u_\varepsilon\ _\infty$
3	345	24694	$9.1e-10$	2978	$2.6e-4$	$1.4e-6$
4	397	29564	$5.1e-10$	3910	$4.4e-5$	$3.7e-7$
5	386	28103	$2.8e-10$	4602	$5.1e-5$	$4.3e-7$
6	343	27813	$1.7e-10$	5111	$9.3e-6$	$1.3e-7$
7	354	24153	$7.2e-11$	5591	$3.0e-6$	$8.3e-8$

TABLE 4. Number of terms and relative errors of solving Poisson's equation in dimensions  $d = 3, \dots, 7$  where the forcing term is a linear combination of 100 randomly generated multivariate Gaussians. The information displayed in each column is described in Table 3. The number of terms  $\tilde{N}$  is significantly larger than that in Table 3 since the principle directions of the matrices  $\Sigma_{f_i}$  are chosen at random causing the solution to have a larger number of terms.

such that  $\max_{\mathbf{x}} |a(\mathbf{x})| / \min_{\mathbf{x}} |a(\mathbf{x})| = 2$ , and choose the forcing function to be

$$(3.19) \quad f(\mathbf{x}) = e^{-\frac{1}{2}(\mathbf{x}-\mu_f)^T \Sigma_f^{-1}(\mathbf{x}-\mu_f)}, \quad \|f\|_{L^\infty} = 1.$$

The free space Green's function for the problem with a constant coefficient

$$-\Delta u(\mathbf{x}) + k^2 u(\mathbf{x}) = f(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^d$$

is given by

$$(3.20) \quad G(\mathbf{x}) = (2\pi)^{-\frac{d}{2}} \left( \frac{k}{\|\mathbf{x}\|} \right)^{\frac{d}{2}-1} K_{\frac{d}{2}-1}(k\|\mathbf{x}\|),$$

where  $K_{\frac{d}{2}-1}$  is a modified Bessel function of the second kind of order  $\frac{d}{2} - 1$ . We approximate the Green's function (3.20) by discretizing the integral

$$G(\mathbf{x}) = (4\pi)^{-\frac{d}{2}} \int_{-\infty}^{\infty} e^{-\frac{\|\mathbf{x}\|^2 e^t}{4} - k^2 e^{-t} + (\frac{d}{2}-1)t} dt$$

as

$$G_\infty(r) = (4\pi)^{-\frac{d}{2}} h \sum_{l \in \mathbb{Z}} e^{-\frac{r^2 e^{hl}}{4} - k^2 e^{-hl} + (\frac{d}{2}-1)hl},$$

where the step size  $h$  is selected to achieve the desired accuracy  $\varepsilon$  (see [12]). We then truncate the sum in the same manner as in (3.11) and obtain

$$\begin{aligned} G_{M,N}(r) &= (4\pi)^{-\frac{d}{2}} h \sum_{l=M}^N e^{-\frac{r^2 e^{hl}}{4} - k^2 e^{-hl} + (\frac{d}{2}-1)hl} \\ &= (4\pi)^{-\frac{d}{2}} h \sum_{l=1}^{N_{terms}} e^{-\frac{r^2 e^{h(l+M-1)}}{4} - k^2 e^{-h(l+M-1)} + (\frac{d}{2}-1)h(l+M-1)}, \end{aligned}$$

where the number of terms  $N_{terms} = M - N + 1$  in  $G_{M,N}$  weakly depends on the dimension (see Table 6). In our computation,  $k = 1$  and we select the accuracy range of  $G_{M,N}$  to be  $[10^{-7}, 10^2]$  with  $\varepsilon = 10^{-10}$ .

We rewrite (3.17) as an integral equation,

$$(3.21) \quad u(\mathbf{x}) - \int_{\mathbb{R}^d} G(\mathbf{x}-\mathbf{y}) \nabla \cdot ((a(\mathbf{y}) - 1) \nabla u(\mathbf{y})) d\mathbf{y} = \int_{\mathbb{R}^d} G(\mathbf{x}-\mathbf{y}) f(\mathbf{y}) d\mathbf{y}.$$

In order to solve (3.21), we first observe that, in a multiresolution basis, for a finite accuracy  $\varepsilon > 0$ , the non-standard form (see [5]) of the Green's function for (3.21) is banded on all scales. This implies that a set of basis functions that can represent the solution  $u$  is fully determined by the size of the bands of the multiresolution representation of the Green's function and of the right hand side  $f$ . This suggests that due to the interaction between the essential supports of the functions involved, we can identify a set of Gaussian atoms by performing one (or a few more) iterations of the integral equation (3.21) even if the fixed-point iteration does not converge. In this approach the accuracy is determined *a posteriori* and can be improved by additional iterations. From the so generated set of atoms, we obtain a basis of Gaussian atoms  $g_l(\mathbf{x}, \mu_l, \Sigma_l)$  by applying the reduction algorithm to identify the best linearly independent subset. Using this basis, we define the ansatz for  $u$  as

$$(3.22) \quad \tilde{u}(\mathbf{x}) = \sum_{l=1}^{\tilde{N}} c_l g_l(\mathbf{x}, \mu_l, \Sigma_l),$$

for some (unknown) coefficients  $c_l$ ,  $l = 1, \dots, \tilde{N}$ , to be determined; substituting  $\tilde{u}$  into either (3.21) or the differential equation (3.17) and computing appropriate inner products, we solve a system of linear algebraic equations for the coefficients  $c_l$ .

Specifically, we rewrite (3.21) as

$$u(\mathbf{x}) = \int_{\mathbb{R}^d} G(\mathbf{x}-\mathbf{y}) f(\mathbf{y}) d\mathbf{y} + \int_{\mathbb{R}^d} G(\mathbf{x}-\mathbf{y}) \nabla \cdot ((a(\mathbf{y}) - 1) \nabla u(\mathbf{y})) d\mathbf{y}$$

which leads to the iteration,

$$(3.23) \quad \begin{aligned} u_{n+1}(\mathbf{x}) &= u_0(\mathbf{x}) + \int_{\mathbb{R}^d} G(\mathbf{x}-\mathbf{y}) \nabla \cdot ((a(\mathbf{y}) - 1) \nabla u_n(\mathbf{y})) d\mathbf{y} \\ u_0(\mathbf{x}) &= \int_{\mathbb{R}^d} G(\mathbf{x}-\mathbf{y}) f(\mathbf{y}) d\mathbf{y}. \end{aligned}$$

To identify a set of Gaussian atoms, we perform one (or several) iteration(s), using as the initial  $u_0$  the collection of atoms in the representation of  $\int_{\mathbb{R}^d} G(\mathbf{x}-\mathbf{y}) f(\mathbf{y}) d\mathbf{y}$ . Using Algorithm 1 we then reduce the number of atoms by removing linearly dependent terms (we may repeat this step if we need to improve accuracy). As a result, we determine a basis of Gaussian atoms  $g_l(\mathbf{x}, \mu_l, \Sigma_l)$  to represent the solution of equation (3.17) as in (3.22). To find the coefficients  $c_l$ ,  $l = 1, \dots, \tilde{N}$ , we substitute (3.22) into the weak formulation of (3.17) to obtain the linear system

$$(3.24) \quad \sum_{l=1}^{\tilde{N}} c_l \langle -\nabla \cdot (a(\mathbf{x}) \nabla g_l(\mathbf{x}, \mu_l, \Sigma_l)), g_k(\mathbf{x}, \mu_k, \Sigma_k) \rangle = \langle f(\mathbf{x}), g_k(\mathbf{x}, \mu_k, \Sigma_k) \rangle, \quad k = 1, \dots, \tilde{N}.$$

The inner products  $\langle -\nabla \cdot (a(\mathbf{x}) \nabla g_l(\mathbf{x}, \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l)), g_k(\mathbf{x}, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \rangle$  are computed explicitly using integration by parts and the fact that

$$\nabla_{\mathbf{x}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})} = -\nabla_{\boldsymbol{\mu}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})},$$

leading to integrals involving only Gaussians (the result is then differentiated with respect to the shift parameter  $\boldsymbol{\mu}$ ).

Using the SVD, we solve the linear system (3.24) to obtain an approximate solution

$$\tilde{u}(\mathbf{x}) = \sum_{l=1}^{\tilde{N}} c_l g_l(\mathbf{x}, \boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l).$$

3.2.1. *Error estimates and results.* Since the exact solution  $u$  is not available, we verify that  $\tilde{u}$  is an approximate solution of (3.17) by evaluating the Fourier transform of the error on a particular set of vectors. Note that

$$h_{error}(\mathbf{x}) = -\nabla \cdot (a(\mathbf{x}) \nabla \tilde{u}(\mathbf{x})) + k^2 \tilde{u}(\mathbf{x}) - f(\mathbf{x}).$$

is a combination of Gaussians and products of Gaussians with low degree polynomials and, therefore, we can explicitly compute its Fourier transform,

$$(3.25) \quad \hat{h}_{error}(\boldsymbol{\xi}) = \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} (-\nabla \cdot (a(\mathbf{x}) \nabla \tilde{u}(\mathbf{x})) + k^2 \tilde{u}(\mathbf{x}) - f(\mathbf{x})) e^{-i\pi \mathbf{x} \cdot \boldsymbol{\xi}} d\mathbf{x}.$$

We then evaluate  $\hat{h}_{error}$  for selected vector arguments  $\boldsymbol{\xi}$ , which we call frequency vectors. To select these vectors, we use the principal directions of the matrices  $\boldsymbol{\Sigma}_l$  of the Gaussian atoms in the representation of  $\tilde{u}$ . To this end, we solve the eigenvalue problem

$$\boldsymbol{\Sigma}_l = \sum_{j=1}^d \lambda_j^{(l)} \mathbf{v}_j^{(l)} \left( \mathbf{v}_j^{(l)} \right)^T$$

and select the frequency vectors along the principal directions of  $\boldsymbol{\Sigma}_l$ . In our experiment, we choose  $s_{min} = 10^{-5}$  and  $s_{max} = (-2 \log(10^{-10}) / \lambda_{min})^{\frac{1}{2}}$ , where  $\lambda_{min} = \min_{l=1, \dots, \tilde{N}, j=1, \dots, d} \lambda_j^{(l)}$ , such that

$$e^{-\frac{\lambda_{min} s^2}{2}} \leq 10^{-10}$$

for  $s > s_{max}$ . We then sample  $s_k, k = 1, \dots, N_s$  using a logarithmic scale on the interval  $[s_{min}, s_{max}]$

$$s_k = e^{\log(s_{min} + (k-1) \frac{s_{max} - s_{min}}{N_s - 1})},$$

and select the  $d \cdot N_s \cdot \tilde{N}$  frequency vectors  $\boldsymbol{\xi}_{jk}^{(l)}$  to be

$$\boldsymbol{\xi}_{jk}^{(l)} = s_k \mathbf{v}_j^{(l)} \quad \text{for } j = 1, \dots, d, k = 1, \dots, N_s, l = 1, \dots, \tilde{N}.$$

In our experiment, we choose  $N_s = 10$ .

We notice that the number of terms in the solution  $\tilde{u}$  grows significantly with the dimension, if the matrices  $\boldsymbol{\Sigma}_a$  and  $\boldsymbol{\Sigma}_f$  are not related (they are effectively random) and/or the range of their eigenvalues is large. In our first experiment, we select matrices  $\boldsymbol{\Sigma}_a$  and  $\boldsymbol{\Sigma}_f$  in (3.18)-(3.19) to be

$$\boldsymbol{\Sigma}_a = U D_a U^T, \quad \text{and } \boldsymbol{\Sigma}_f = U D_f U^T,$$

where  $U$  is a  $d \times d$  random unitary matrix and  $D_a$  and  $D_f$  are  $d \times d$  diagonal matrices. We set the first two diagonal entries of  $D_a$  and  $D_f$  to be 0.1 and 20, and sample the other diagonal entry/entries from a uniform distribution  $\mathcal{U}(0.1, 20)$ . A random permutation is applied after all diagonal entries are generated. In the second experiment, we construct matrices  $\boldsymbol{\Sigma}_a$  and  $\boldsymbol{\Sigma}_f$  as

$$\boldsymbol{\Sigma}_a = U_a D_a U_a^T, \quad \boldsymbol{\Sigma}_f = U_f D_f U_f^T$$

$d$	$N_{terms}$	$N_{tot}$	$\tilde{N}$	$\ \hat{h}_{error}\ _{\infty} / \ \hat{f}\ _{\infty}$
3	104	5985	1184	$1.8e-6$
4	109	11624	2649	$3.5e-6$
5	117	16014	3640	$3.8e-6$
6	121	22466	4377	$8.0e-5$
7	127	31522	5573	$2.5e-5$

TABLE 5. Results of solving a second order elliptic equation with a variable coefficient where the principle directions of the matrices  $\Sigma_a$  and  $\Sigma_f$  in (3.18) and (3.19) are aligned and their eigenvalues are in the range  $(0.1, 20)$ .

$d$	$N_{terms}$	$N_{tot}$	$\tilde{N}$	$\ \hat{h}_{error}\ _{\infty} / \ \hat{f}\ _{\infty}$
3	104	15348	2586	$8.1e-7$
4	109	32688	6883	$1.9e-6$
5	117	40110	11793	$2.4e-6$
6	121	64664	18205	$6.7e-5$
7	127	75059	22966	

TABLE 6. Results of solving a second order elliptic equation with a variable coefficient where the principle directions of the matrices  $\Sigma_a$  and  $\Sigma_f$  in (3.18) and (3.19) are not aligned and their eigenvalues are in the range  $(0.1, 1)$ . We note that the accuracy estimation in dimension  $d = 7$  is computationally expensive and we skipped it.

where  $U_a$  and  $U_f$  are  $d \times d$  random unitary matrices,  $D_a$  and  $D_f$  are  $d \times d$  diagonal matrices. We set the first two diagonal entries of  $D_a$  and  $D_f$  to be 0.1 and 1, and sample the other diagonal entry/entries from a uniform distribution  $\mathcal{U}(0.1, 1)$ . Again we randomly permute the diagonals of  $D_a$  and  $D_f$ . We also notice that if the centers  $\mu_a$  and  $\mu_f$  are far away (no overlapping essential supports), then solving (3.17) is effectively the same as solving the Poisson's equation. In our tests, we select  $\mu_f = 0$  and  $\mu_a = (1, 0, \dots, 0)^T$  so that  $\|\mu_a - \mu_f\|_2 = 1$ .

In both experiments, we iterate (3.23) once to generate a set of Gaussian atoms. The results are displayed in Tables 5 and 6 where we show the dimension of the problem,  $d$ , the number of terms  $N_{terms}$  in the approximation of the Green's function, the number of terms  $N_{tot}$  obtained by performing one iteration in (3.23), the number of terms  $\tilde{N}$  in the solution  $\tilde{u}$  after reduction and the resulting accuracy. For these two tables, we use the notation  $\|\hat{g}\|_{\infty} = \max_{j,k,l} \left| \hat{g} \left( \xi_{jk}^{(l)} \right) \right|$ .

#### 4. KERNEL DENSITY ESTIMATION

We describe a new algorithmic approach to Kernel Density Estimation (KDE) based on Algorithm 1. KDE is a non-parametric method for constructing the PDF of data points used in cluster analysis, classification, and machine learning. The standard KDE construction is practical only in low dimensions,  $d = 1, 2, 3$  as it requires a Fourier transform of the data points, the cost of which grows exponentially with dimension (see e.g. [40, 4] for a technique based on the Fourier transform). Our approach avoids using the Fourier transform and is applicable in high dimensions. We note that a randomized approach that can be used for KDE estimation was recently suggested in [36]. In this paper we do not provide a comparison with other techniques that are applicable to KDE (e.g. reproducing kernel techniques which formulate the problem as a minimization of an objective function, see for example [21] and references

therein). We plan to develop our approach further and provide an appropriate comparison with other methods elsewhere.

The essence of KDE (see e.g. [43]) is to associate a smooth PDF  $f$  with data points  $\mathbf{x}_j \in \mathbb{R}^d$ ,  $j = 1, \dots, N$ ,

$$(4.1) \quad f(\mathbf{x}, h) = \frac{1}{N} \sum_{j=1}^N K_h(\mathbf{x} - \mathbf{x}_j),$$

where  $K_h(\mathbf{x}) = K(\mathbf{x}/h)/h$  and  $K$  is a nonnegative function with zero mean and  $\int_{\mathbb{R}^d} K(\mathbf{x}) d\mathbf{x} = 1$ . In what follows, we use a multivariate Gaussian as the kernel  $K$ . A naive implementation of (4.1) would require  $N$  evaluations of the kernel  $K_h$  for each point  $\mathbf{x}$  so that the computational cost of using this approach in a straightforward manner is prohibitive if  $N$  is large. The selection of the parameter  $h$ , the so-called bandwidth or scale parameter, is a well recognized delicate issue and, in our one dimensional example, we use  $h$  computed within Mathematica<sup>TM</sup> implementation of KDE.

In our approach, for a user selected target accuracy  $\varepsilon$ , we seek a *subset* of linear independent terms in (4.1) and express the remaining terms as their linear combinations. Thus, by removing redundant terms in the representation of  $f(\mathbf{x}, h)$ , we construct

$$(4.2) \quad F(\mathbf{x}, h) = \sum_{\ell=1}^r a_\ell K_h(\mathbf{x} - \mathbf{x}_{j_\ell}),$$

where  $r \ll N$  and

$$|f(\mathbf{x}, h) - F(\mathbf{x}, h)| \leq \varepsilon.$$

In other words, with accuracy  $\varepsilon$ , we obtain an approximation of the function  $f$  by a function  $F$  with a small number of terms. This reduction algorithm seeking a subset of linear independent terms in (4.1) can be used for any kernel in  $\mathbb{R}^d$  such that the cost of evaluating the multidimensional inner product

$$g_{ij} = \int_{\mathbb{R}^d} K_h(\mathbf{x} - \mathbf{x}_i) K_h(\mathbf{x} - \mathbf{x}_j) d\mathbf{x}$$

is reasonable i.e. depends mildly on the dimension  $d$ . For multivariate Gaussians, the values  $g_{ij}$  are available via an explicit expression, see Section 7.0.4.

The computational cost of our algorithm is  $\mathcal{O}(r^2N + p(d)rN)$ . Here  $N$  is the original number of data points and  $r$  is the final number of terms in the chosen linearly independent subset and  $p(d)$  is the cost of computing the inner product between two terms of the mixture. In typical KDE applications  $p(d) \sim d$  since usual kernels admit a separated representation.

**4.1. A comparison in dimension  $d = 1$ .** In low dimensions, the standard approach to KDE relies on using the Fast Fourier transform to both, assist in estimating the bandwidth parameter  $h$  and in constructing a more efficient representation of (4.1) on an equally spaced grid (see e.g. [43, Section 3.5]). Implementations of this approach can be found in many packages in dimensions  $d = 1, 2$ , e.g. Matlab, Mathematica, etc. While this approach is appropriate in low dimensions, an extension of this algorithm to high dimensions is prevented by the ‘‘curse of dimensionality’’. Thus, in high dimensions, only values at selected points can be computed (see [17] and Matlab implementation of KDE in high dimensions).

In order to illustrate our approach we provide a simple example with a bimodal distribution in dimension  $d = 1$ . Although this example is in one variable, it allows us to emphasize the differences between the existing KDE methods and our approach. We generate test data by using two normal distributions with means  $\mu_1 = 0$  and  $\mu_2 = 4$ . The exact PDF of this data is given by

$$(4.3) \quad g(x) = \frac{1}{2} \left( \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \frac{x^2}{\sigma^2}} + \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \frac{(x-4)^2}{\sigma^2}} \right),$$

where  $\sigma = 1$ . We then use KDE implemented in Mathematica<sup>TM</sup> with the Gaussian kernel  $K$ . Using  $N = 10^5$  data samples drawn from (4.3) (so that the initial sum (4.1) has  $N$  terms), the scaling parameter was set to  $h = 0.20121412622314902019$ . The true distribution (4.3) and the approximation error obtained



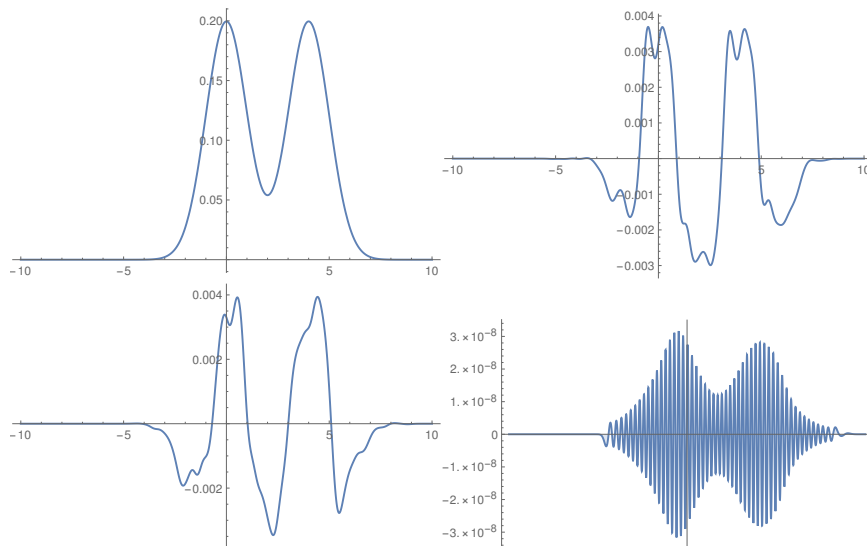


FIGURE 4.1. The distribution (4.3) (top left) and the error of its estimation with 241 term centered at an equally spaced grid obtain using KDE in Mathematica<sup>TM</sup> (top right). The error of estimating (4.3) using Algorithm 1 with 102 terms centered at a selected subset of the *original* data points (bottom left). We also show the difference between the definition of the PDF in (4.1) and its reduced version in (4.2) obtained using Algorithm 1 where accuracy was set to  $10^{-7}$  (bottom right).

by Mathematica<sup>TM</sup> by reducing (4.1) from  $N = 10^5$  terms to 241 Gaussian terms centered on an equally spaced grid are displayed in Figure 4.1. Using our algorithm with the same parameter  $h$ , we reduce (4.1) from  $N = 10^5$  terms to 102 terms centered at a selected *subset of the original data points*. The error of the resulting approximation is displayed in Figure 4.1, where we also show the difference between (4.1) and (4.2). The main point here is that while the standard approach in high dimensions becomes impractical (as it requires a multidimensional grid), our approach proceeds unchanged since the cost of the reduction algorithm depends on dimension only mildly.

*Remark 8.* We selected a much higher accuracy for reduction than the difference between the original and estimated PDFs in order to illustrate the fact that the accuracy limit of Algorithm 1 of about 7 to 8 digits is more than sufficient for this application. Note that to achieve a comparable accuracy for estimation of a PDF via KDE one needs  $\approx 10^{16}$  points since the accuracy improves as  $\mathcal{O}(1/\sqrt{N})$ .

*Remark 9.* Using KDE in high dimensions usually requires an additional assumption that points are located in a vicinity of a low dimensional manifold. If the intrinsic dimension of a collection of points is high, then the Gaussians assigned to the points will tend to be linearly independent since the expected distance between two points becomes increasingly large as the dimension  $d$  increases (see e.g. comments in e.g. [42, Section 1.5.3] and/or Figure 5 in [36]). Depending on dimension, one may need to use sufficiently “flat” Gaussians to address this issue (for example, in the problem of identifying a bimodal distribution in high dimensions). Such high-dimensional study exceeds the goal of this paper and will be reported elsewhere.

**4.2. An example in high ambient dimension.** We generate  $N = 10^5$  samples from a two dimensional Gaussian distribution with the PDF

$$g(\mathbf{y}) = \frac{1}{2} \left( \frac{1}{2\pi} e^{-\frac{1}{2}(\mathbf{y}-\mu_1)^T \Sigma_1^{-1}(\mathbf{y}-\mu_1)} + \frac{1}{2\pi} e^{-\frac{1}{2}(\mathbf{y}-\mu_2)^T \Sigma_2^{-1}(\mathbf{y}-\mu_2)} \right)$$

where  $\mathbf{y} = (y_1, y_2)$ ,  $\mu_1 = (0, 0)$ ,  $\mu_2 = (3, 3)$  and

$$\Sigma_1 = \begin{pmatrix} 2 & 0 \\ 0 & 0.5 \end{pmatrix}, \quad \Sigma_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We pad these samples with zeros so that they belong to a  $d$ -dimensional space and denote them by  $\{\mathbf{y}_i\}_{i=1}^N$ . We then apply a random rotation matrix  $U$  to obtain the test data  $\{\mathbf{x}_i\}_{i=1}^N$ . As the bandwidth parameter, we set

$$h = \left( \frac{4}{2d+1} \right)^{\frac{1}{d+4}} N^{-\frac{1}{d+4}},$$

a value that minimizes the mean integrated square error for an underlying standard normal distribution (see e.g. [43]). In our case, since the intrinsic dimension of the test data is 2, we set  $h = 0.14142135623730950488$ . The initial kernel density estimator using all test data is

$$(4.4) \quad f(\mathbf{x}) = \frac{1}{Nh^d} \sum_{i=1}^N \frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{x}_i)^T (h^2 I_d)^{-1}(\mathbf{x}-\mathbf{x}_i)},$$

where  $I_d$  is the  $d$ -by- $d$  identity matrix. We then reduce the number of terms in (4.4) using Algorithm 1 and obtain a sum of Gaussians with fewer terms,

$$\tilde{f}(\mathbf{x}) = \sum_{j=1}^{\tilde{N}} c_j \frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{x}_j)^T (h^2 I_d)^{-1}(\mathbf{x}-\mathbf{x}_j)}.$$

In our experiments, we choose dimensions  $d = 2, \dots, 16$  and error threshold  $\varepsilon = 10^{-1}$  in Algorithm 1. The number of terms after reduction is about 2000 for all dimensions and the approximation error is about  $1.2 \times 10^{-3}$  (see Figure 4.2).

In order to compare our approximation with the true PDF  $g(\mathbf{x})$ , we consider a point  $\mathbf{y} = (y_1, y_2, 0, \dots, 0) \in \mathbb{R}^d$  such that, under the rotation  $U$ , we have  $\mathbf{x} = U\mathbf{y}$ . Evaluating (4.4) at  $\mathbf{x}$ , we obtain

$$\begin{aligned} f(\mathbf{x}) &= f(U\mathbf{y}) \\ &= \frac{1}{Nh^d} \sum_{i=1}^N \frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{1}{2}(U(\mathbf{y}-\mathbf{y}_i))^T (h^2 I_d)^{-1} U(\mathbf{y}-\mathbf{y}_i)} \\ &= \frac{1}{Nh^d} \sum_{i=1}^N \frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{1}{2}(\mathbf{y}-\mathbf{y}_i)^T (h^2 I_d)^{-1}(\mathbf{y}-\mathbf{y}_i)} \\ &= c_d \frac{1}{Nh^2} \sum_{i=1}^N \frac{1}{2\pi} e^{-\frac{(y_1-y_1^{(i)})^2 + (y_2-y_2^{(i)})^2}{2h^2}}, \end{aligned}$$

where  $c_d = h^{2-d} (2\pi)^{\frac{2-d}{2}}$  and  $y_j^{(i)}$  denote the  $j$ -th component of  $\mathbf{y}_i$ . Notice that the kernel density estimator for the PDF of the original distribution in two dimensions is

$$g(y_1, y_2) \approx \frac{1}{Nh^2} \sum_{i=1}^N \frac{1}{2\pi} e^{-\frac{(y_1-y_1^{(i)})^2 + (y_2-y_2^{(i)})^2}{2h^2}} = \frac{1}{c_d} f(\mathbf{x})$$

Therefore, to estimate the size of the approximation error, we generate a set of  $64 \times 64$  equispaced grid points  $(\bar{y}_1^{(i)}, \bar{y}_2^{(j)})$ ,  $i, j = 1, \dots, 64$  in  $[-8, 8] \times [-8, 8]$ . We pad each  $(\bar{y}_1^{(i)}, \bar{y}_2^{(j)})$  with  $d-2$  zeros to embed

it into  $\mathbb{R}^d$ , which we denote as  $\bar{\mathbf{y}}_{ij} = (\bar{y}_1^{(i)}, \bar{y}_2^{(j)}, 0, \dots, 0) \in \mathbb{R}^d$ . We then apply the same rotation matrix  $U$  to obtain a set of points  $\bar{\mathbf{x}}_{ij} = U\bar{\mathbf{y}}_{ij} \in \mathbb{R}^d$ . In Figure 4.2 we illustrate the result for dimension  $d = 16$ ; we show the PDF  $g$ , the errors between  $g$  and the KDE estimate  $f$  before reduction and the estimate  $\tilde{f}$  after reduction, as well as the difference between  $f$  and  $\tilde{f}$ .

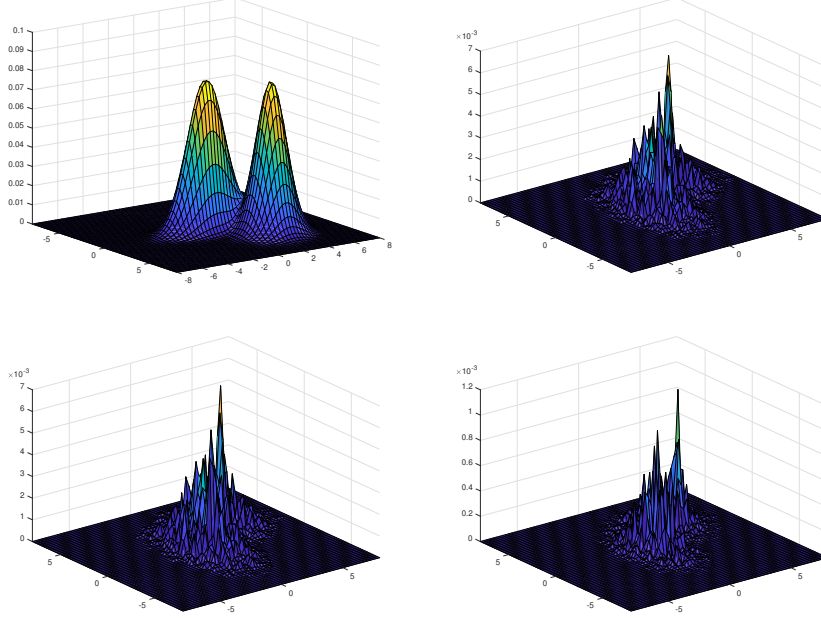


FIGURE 4.2. Using the grid  $(\bar{y}_1^{(i)}, \bar{y}_2^{(j)})$ ,  $i, j = 1, \dots, 64$ , and the rescaling constant  $c_d = h^{2-d} (2\pi)^{\frac{2-d}{2}}$ ,  $d = 16$ , we display the PDF  $g(\bar{y}_1^{(i)}, \bar{y}_2^{(j)})$  (top left) and the difference  $\left| g(\bar{y}_1^{(i)}, \bar{y}_2^{(j)}) - \frac{1}{c_d} f(U\bar{\mathbf{y}}_{ij}) \right|$  between  $g$  and constructed PDF  $f$  (top right). We also show the difference  $\left| g(\bar{y}_1^{(i)}, \bar{y}_2^{(j)}) - \frac{1}{c_d} \tilde{f}(U\bar{\mathbf{y}}_{ij}) \right|$  between  $g$  and  $\tilde{f}$  (bottom left) and the difference  $\frac{1}{c_d} \left| f(U\bar{\mathbf{y}}_{ij}) - \tilde{f}(U\bar{\mathbf{y}}_{ij}) \right|$  between  $f$  and  $\tilde{f}$  (bottom right).

## 5. FAR-FIELD SUMMATION IN HIGH DIMENSIONS

Given a large set of points in dimension  $d \gg 3$  with pairwise interaction via a non-oscillatory kernel, our approach provides a deterministic algorithm for fast summation in the far-field setup (i.e. where two groups of points are separated). Recently a randomized algebraic approach in a similar setup was suggested in [36]. Instead, we use Algorithm 1 as a tool to rapidly evaluate

$$g_m = \sum_{n=1}^N f_n K(\mathbf{x}_m, \mathbf{y}_n), \quad m = 1, \dots, M,$$

where  $M$  and  $N$  are large and  $K(\mathbf{x}, \mathbf{y})$  is a non-oscillatory kernel with a possible singularity at  $\mathbf{x} = \mathbf{y}$  (recall that kernels of mathematical physics typically have a singularity for coincident arguments  $\mathbf{x}$  and  $\mathbf{y}$ ).

Let us consider sources  $\{(\mathbf{y}_n, f_n)\}_{n=1}^N$  and targets  $\{(\mathbf{x}_m, g_m)\}_{m=1}^M$  occupying two distinct  $d$ -dimensional balls. Specifically, we assume that  $\|\mathbf{y}_n - \mathbf{y}_c\| \leq r_s$  and  $\|\mathbf{x}_m - \mathbf{x}_c\| \leq r_t$ , where  $\mathbf{y}_c, \mathbf{x}_c$  and  $r_s, r_t$  are the centers and the radii of the balls. We also assume that sources and targets are separated, i.e.,  $0 < r \leq \|\mathbf{x}_m - \mathbf{y}_n\| \leq R$ . The separation of sources and targets implies that in the evaluation of the kernel we are never close to a possible singularity of the kernel at  $\mathbf{x} = \mathbf{y}$ . Separation of sources and targets allows us to use a non-singular approximation of the kernel and reduce the problem to finding the best linearly independent subsets of sources (or targets) as defined by such approximate kernel. We also assume that the sources are located on a low-dimensional manifold embedded in high-dimensional space (see Remark 12 below).

In order to use Algorithm 1, we need to define an inner product that takes into account the assumption of separation of sources and targets. We illustrate this using the example of the Poisson kernel (3.3),  $K(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|^{-d+2}$  (without standard normalization). Approximating  $K(\mathbf{x}, \mathbf{y})$  as in Section 3.1 (c.f. (3.4)), we have

$$(5.1) \quad \left| K(\mathbf{x}, \mathbf{y}) - \sum_{l \in \mathbb{Z}} w_l e^{-\tau_l \|\mathbf{x} - \mathbf{y}\|^2} \right| \leq \varepsilon K(\mathbf{x}, \mathbf{y}).$$

Since sources and targets are separated, we drop terms in (5.1) with sufficiently large exponents (they produce a negligible contribution in the interval  $[r, R]$ ) as well as replace terms with small exponents using an algorithm described in [12]. As a result, we obtain

$$\tilde{K}(\mathbf{x}, \mathbf{y}) = \sum_{l=L_0}^{L_1} w_l e^{-\tau_l \|\mathbf{x} - \mathbf{y}\|^2},$$

such that

$$\left| K(\mathbf{x}, \mathbf{y}) - \tilde{K}(\mathbf{x}, \mathbf{y}) \right| < \tilde{\varepsilon}, \text{ for } r \leq \|\mathbf{x} - \mathbf{y}\| \leq R,$$

where  $\tilde{\varepsilon}$  is slightly larger than  $\varepsilon$ . It follows that

$$|g_m - \tilde{g}_m| \leq \tilde{\varepsilon},$$

where

$$\tilde{g}_m = \sum_{n=1}^N f_n \tilde{K}(\mathbf{x}_m, \mathbf{y}_n), \quad m = 1, \dots, M.$$

Since  $\|\mathbf{x} - \mathbf{x}_c\| \leq r_t$  implies  $r \leq \|\mathbf{x} - \mathbf{y}_n\| \leq R$ , we define the inner product as an integral over the ball  $\|\mathbf{x} - \mathbf{x}_c\| \leq r_t$ ,

$$(5.2) \quad \langle \tilde{K}(\cdot, \mathbf{y}_n), \tilde{K}(\cdot, \mathbf{y}_{n'}) \rangle_d = \int_{\|\mathbf{x} - \mathbf{x}_c\| \leq r_t} \tilde{K}(\mathbf{x}, \mathbf{y}_n) \tilde{K}(\mathbf{x}, \mathbf{y}_{n'}) d\mathbf{x}.$$

The inner product (5.2) can be reduced to a one dimensional integral, as we show next.

$$\begin{aligned} & \langle \tilde{K}(\cdot, \mathbf{y}_n), \tilde{K}(\cdot, \mathbf{y}_{n'}) \rangle_d \\ &= \sum_{l, l'=L_0}^{L_1} w_l w_{l'} \int_{\|\mathbf{x} - \mathbf{x}_c\| \leq r_t} e^{-\tau_l \|\mathbf{x} - \mathbf{y}_n\|^2} e^{-\tau_{l'} \|\mathbf{x} - \mathbf{y}_{n'}\|^2} d\mathbf{x} \\ &= \sum_{l, l'=L_0}^{L_1} w_l w_{l'} e^{-\frac{\tau_l \tau_{l'}}{\tau_l + \tau_{l'}} \|\mathbf{y}_n - \mathbf{y}_{n'}\|^2} \int_{\|\mathbf{x} - \mathbf{x}_c\| \leq r_t} e^{-(\tau_l + \tau_{l'}) \left\| \mathbf{x} - \frac{\tau_l \mathbf{y}_n + \tau_{l'} \mathbf{y}_{n'}}{\tau_l + \tau_{l'}} \right\|^2} d\mathbf{x} \\ &= \sum_{l, l'=L_0}^{L_1} \tilde{w}_{ll'}^{n, n'} I_{ll'}^{n, n'}, \end{aligned}$$

where

$$\tilde{w}_{ll'}^{n, n'} = w_l w_{l'} e^{-\frac{\tau_l \tau_{l'}}{\tau_l + \tau_{l'}} \|\mathbf{y}_n - \mathbf{y}_{n'}\|^2}$$

and

$$I_{l'l'}^{nm'} = \int_{\|\mathbf{x}-\mathbf{x}_c\| \leq r_t} e^{-\tilde{\tau}_{l'l'} \|\mathbf{x}-\tilde{\mathbf{y}}_{l'l'}^{nm'}\|^2} d\mathbf{x}, \quad \tilde{\tau}_{l'l'} = \tau_l + \tau_{l'}, \quad \tilde{\mathbf{y}}_{l'l'}^{nm'} = \frac{\tau_l \mathbf{y}_n + \tau_{l'} \mathbf{y}_{n'}}{\tau_l + \tau_{l'}}.$$

Using spherical coordinates in dimensions  $d \geq 3$ , we obtain

$$\begin{aligned} I_{l'l'}^{nm'} &= \int_{\|\mathbf{x}-\mathbf{x}_c\| \leq r_t} e^{-\tilde{\tau}_{l'l'} \|\mathbf{x}-\tilde{\mathbf{y}}_{l'l'}^{nm'}\|^2} d\mathbf{x} \\ &= e^{-\tilde{\tau}_{l'l'} \|\mathbf{x}_c-\tilde{\mathbf{y}}_{l'l'}^{nm'}\|^2} \int_{\|\mathbf{z}\| \leq r_t} e^{-\tilde{\tau}_{l'l'} \|\mathbf{z}\|^2} e^{-2\tilde{\tau}_{l'l'} \langle \mathbf{z}, \mathbf{x}_c-\tilde{\mathbf{y}}_{l'l'}^{nm'} \rangle} d\mathbf{z} \\ &= e^{-\tilde{\tau}_{l'l'} \|\mathbf{x}_c-\tilde{\mathbf{y}}_{l'l'}^{nm'}\|^2} \Omega_{d-2} \int_0^{r_t} \left( \int_0^\pi e^{-2\tilde{\tau}_{l'l'} r \|\mathbf{x}_c-\tilde{\mathbf{y}}_{l'l'}^{nm'}\| \cos \theta} \sin^{d-2} \theta d\theta \right) e^{-\tilde{\tau}_{l'l'} r^2} r^{d-1} dr \\ (5.3) \quad &= e^{-\tilde{\tau}_{l'l'} \|\mathbf{x}_c-\tilde{\mathbf{y}}_{l'l'}^{nm'}\|^2} \frac{2\pi^{\frac{d}{2}}}{(\tilde{\tau}_{l'l'} \|\mathbf{x}_c-\tilde{\mathbf{y}}_{l'l'}^{nm'}\|)^{\frac{d-2}{2}}} \int_0^{r_t} I_{\frac{d-2}{2}} \left( 2\tilde{\tau}_{l'l'} \|\mathbf{x}_c-\tilde{\mathbf{y}}_{l'l'}^{nm'}\| r \right) e^{-\tilde{\tau}_{l'l'} r^2} r^{\frac{d}{2}} dr, \end{aligned}$$

where  $\Omega_d$  is the surface area of the  $d$ -dimensional sphere embedded in  $(d+1)$ -dimensional space, i.e.,  $\Omega_d = \frac{2\pi^{\frac{d+1}{2}}}{\Gamma(\frac{d+1}{2})}$  and  $I_d$  is the  $d$ -th order modified Bessel function of the first kind (see [1, Eq. 9.6.18]). While in odd dimensions  $d$  the integrand in (5.3) can be extended from  $[0, r_t]$  to  $[-r_t, r_t]$  as a smooth function allowing us to use the trapezoidal rule, such extension is not available in even dimensions. Because of this, we choose to use quadratures on  $[0, r_t]$  developed in [10] (alternatively, one can use quadratures from [2]).

*Remark 10.* It is an important observation that the selection of the inner product for finding a linearly independent subset of functions is not limited to the standard one defined in (5.2). Observing that (5.2) approaches zero as  $d$  increases, in all of our experiments in dimensions  $d = 3, \dots, 128$ , we use (5.2) where we set  $d = 3$ . Thus, the inner product  $\langle \cdot, \cdot \rangle_3$  no longer corresponds to the integral between the functions  $\tilde{K}(\mathbf{x}, \mathbf{y}_n)$  and  $\tilde{K}(\mathbf{x}, \mathbf{y}_{n'})$ . However, since we use inner products only to identify the best linearly independent subset of sources (skeleton sources) and compute the coefficients to replace the remaining terms as linear combinations of these skeleton sources, there are many choices of inner products that will produce similar results.

Associating with each  $\mathbf{y}_n$  the source function  $K(\mathbf{x}, \mathbf{y}_n)$ , we use Algorithm 1 to find the skeleton terms (i.e. the skeleton sources) with indices  $\hat{I} = \{n_k\}_{k=1}^{r_s}$  allowing us to express the remaining source functions (for  $n \notin \hat{I}$ ) as

$$\left| \tilde{K}(\mathbf{x}, \mathbf{y}_n) - \sum_{k=1}^{r_s} \tilde{f}_{n_k} \tilde{K}(\mathbf{x}, \mathbf{y}_{n_k}) \right| \leq \varepsilon_1, \quad n \notin \hat{I}.$$

*Remark 11.* Using Algorithm 1 to find the skeleton sources requires  $\mathcal{O}(r_s^2 N + p(d)r_s N)$  operations and computing interactions between skeleton sources and targets requires additional  $\mathcal{O}(r_s M)$  operations. Clearly, instead of working with sources, we can work with targets. If targets are located on a low-dimensional manifold, we can associate the functions  $\{K(\mathbf{x}_m, \mathbf{y})\}_{m=1}^M$  with targets and use Algorithm 1 to find the skeleton targets. In such case, the computational cost becomes  $\mathcal{O}(r_t^2 M + p(d)r_t M + r_t N)$ , where  $r_t$  is the number of skeleton targets.

*Remark 12.* If sources are chosen from a random distribution in  $\mathbb{R}^d$  rather than located in a small neighborhood of a low-dimensional manifold, the expected distance between two sources  $\|\mathbf{y}_n - \mathbf{y}_{n'}\|$  becomes increasingly large as the dimension  $d$  increases (see e.g. comments in [42, Section 1.5.3] and examples in [36]). As a result, the functions of variable  $\mathbf{x}$ ,  $\tilde{K}(\mathbf{x}, \mathbf{y}_n)$  and  $\tilde{K}(\mathbf{x}, \mathbf{y}_{n'})$ , are effectively linearly independent

$d$	$\text{dist}_{near}$	$\text{dist}_{far}$	$r_s$	$error$
3	2.5310	5.6317	22	$2.6573e-07$
4	2.8922	5.3544	27	$1.5657e-07$
5	3.2019	5.0864	29	$3.0440e-08$
8	3.1261	5.1186	23	$3.2212e-08$
16	3.1881	5.1049	25	$2.5774e-08$
32	3.5261	5.0242	25	$1.1001e-08$
64	3.6577	4.7520	24	$1.4346e-09$
128	3.8410	4.4825	25	$1.7492e-09$

TABLE 7. Skeleton sources selected using the inner product in (5.2) in dimension  $d$ . We report the actual minimal and maximum distances between sources and targets ( $\text{dist}_{near}$  and  $\text{dist}_{far}$ ), the number of skeleton sources  $r_s$ , and the relative error in (5.4).

as  $d$  becomes large so that in order to have compressibility, the sources must have a low intrinsic dimension. Therefore, the assumption that sources are located in a small neighborhood of a low-dimensional manifold is not specific to our approach.

**5.1. Skeleton sources.** We illustrate our approach using sources located on a two-dimensional manifold embedded in a high-dimensional space. For our example, we generate points  $\{\mathbf{y}_n\}_{n=1}^N$ ,  $\mathbf{y}_n \in \mathbb{R}^d$  so that the first two coordinates are random variables drawn from the two-dimensional standard normal distribution and the remaining coordinates are set to zero. Next we apply a random rotation and rescale the points so that  $\|\mathbf{y}_n\| \leq 1$  for all  $n = 1, \dots, N$ . For targets, we draw points  $\{\mathbf{x}_m\}_{m=1}^M$  from the  $d$ -dimensional standard normal distribution and rescale them so that  $\|\mathbf{x}_m\| \leq 1$ . We then shift the first component of  $\{\mathbf{y}_n\}_{n=1}^N$  by 2 and that of  $\{\mathbf{x}_m\}_{m=1}^M$  by  $-2$  so that sources and targets are well separated. Finally, we select the coefficients of sources,  $\{f_n\}_{n=1}^N$ , from the uniform distribution  $\mathcal{U}(0, 1)$ . In all tests we set  $N = 10^4$  and  $M = 10^3$ . In Table 7 we report the actual minimal and maximum distances between sources and targets ( $\text{dist}_{near}$  and  $\text{dist}_{far}$ ), the number of skeleton sources  $r_s$ , and the relative error of the approximation,

$$(5.4) \quad error = \frac{\max_{m=1, \dots, M} |g_m - \tilde{g}_m|}{\max_{m=1, \dots, M} |g_m|},$$

for selected dimensions  $3 \leq d \leq 128$ . For dimensions  $d = 3, 4, 5$ , we use the Poisson kernel  $\|\mathbf{x} - \mathbf{y}\|^{-d+2}$  while for  $d \geq 8$ , we use the kernel  $\|\mathbf{x} - \mathbf{y}\|^{-1}$  (the fast decay of  $\|\mathbf{x} - \mathbf{y}\|^{-d+2}$  results in a negligible interaction between sources and targets in our setup).

**5.2. Equivalent sources.** In this example, we consider a similar setting as in Section 5.1 for  $d = 2, 3$ . We want to replace true sources  $\{\mathbf{y}_n\}_{n=1}^N$  located inside a ball by equivalent sources on its boundary as to reproduce their interaction with the targets within a selected accuracy  $\varepsilon$ . We expect the number of equivalent sources on the boundary to be significantly smaller than the number of original true sources so that pairwise interactions with targets can be computed rapidly. We note that such strategy is used in many numerical algorithms (see e.g. [49]) and here we demonstrate that our reduction algorithm can solve this problem.

We combine an initial set of candidate equivalent sources (note that their number will be reduced by the procedure) with the true sources and compute the Cholesky decomposition of their Gram matrix. We use Algorithm 1 with the inner product defined in (5.2) and modify the pivoting strategy to first pivot only among the candidate equivalent sources until we run out of significant pivots (i.e., pivots above the accuracy  $\varepsilon$ ); only then we switch to pivot among the true sources. Finally, we compute new

coefficients in the usual way (see Algorithm 1) noting that, initially, the candidate equivalent sources had zero coefficients. This approach allows us to (i) obtain the minimal number of equivalent sources and (ii) remove as many of the true sources as possible (we do not preclude the possibility of some of the true sources to remain).

To examine the performance of our approach, we draw source and target points from the  $d$ -dimensional standard normal distribution (where  $d = 2, 3$ ), rescale and translate these points so that targets are located in a ball of radius 1 centered at  $\mathbf{x}_c$  and sources are located in a ball of radius 0.9 centered at  $\mathbf{y}_c$ . We choose  $\mathbf{x}_c = (-2, 0)$ ,  $\mathbf{y}_c = (2, 0)$  for  $d = 2$  and  $\mathbf{x}_c = (-2, 0, 0)$ ,  $\mathbf{y}_c = (2, 0, 0)$  for  $d = 3$  to make sure sources and targets are well separated. Next we pick locations for the candidate equivalent sources on the surface of the ball of radius 1 centered at  $\mathbf{y}_c$ . In dimension  $d = 2$ , we pick

$$\mathbf{z}_k = \mathbf{y}_c + (\cos \theta_k, \sin \theta_k), \quad k = 1, \dots, K,$$

where the angles  $\theta_k$  are equally spaced on  $[0, 2\pi]$  with step size  $\frac{2\pi}{K}$ . In dimension  $d = 3$  we pick

$$\mathbf{z}_{kl} = \mathbf{y}_c + (\cos \theta_k \sin \phi_l, \sin \theta_k \sin \phi_l, \cos \phi_l), \quad k = 1, \dots, K, \quad l = 1, \dots, L,$$

where the angles  $\theta_k$  are equally spaced on  $[0, 2\pi]$  with step size  $\frac{2\pi}{K}$  and the angles  $\phi_l$  are the Gauss-Legendre nodes on  $[0, \pi]$ . In our experiments we choose a relatively small number of true sources and targets ( $N, M = 1000$ ) so that the result can be clearly visualized (see Figure 5.1 and 5.2). Note that the number of sources can be significantly higher since the algorithm is linear in this parameter. We demonstrate the results in Figure 5.1 and 5.2, where we display the original sources and targets and indicate both, candidate equivalent sources and selected equivalent sources obtained by Algorithm 1.

**5.3. Partitioning of points into groups.** Reduction Algorithm 1 can be used to subdivide scattered points into groups. Indeed, if a set of points (seeds) are specified beforehand then, like in Voronoi decomposition, all points can be split into groups by their proximity to the seeds, i.e. a point belongs to a group associated with a given seed if it is the closest to it among all seeds. There are several algorithms, e.g. Lloyd's algorithm [34], that use such seeds in an iterative procedure to optimize some properties of the sought subdivision. The question then becomes how to choose such seeds. In order to avoid poor clusterings the so-called k-means++ algorithm is often used [3]. We would like to point out that Algorithm 1 can be used to generate initial seeds using linear dependence (which is a proxy for distances between points). We only illustrate its potential use for selecting seeds and provide no comparison with k-means++ or spectral clustering algorithms (see e.g. [39] and references therein). We also do not provide a comparison with model-based clustering (see e.g. [33, 38]). Using Algorithm 1 to subdivide scattered points into groups requires further analysis and we plan to address it elsewhere.

For our experiment, we use the same set of points as in Section 4.2 and associate with each point a Gaussian centered at that point. The scale parameter of the Gaussians can be selected sufficiently large (so that the Gaussian is sufficiently flat) to cover the whole set of points. We can then use Algorithm 1 to select the seeds. In order to force Algorithm 1 to select a specific first point, we introduce an additional point as the mean

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i,$$

and associate it with an additional Gaussian which we place at the beginning of the list of Gaussian atoms.

The seeds are the first significant pivots produced by the algorithm and our choice of their number depends on the goals of the subdivision. By its nature, Algorithm 1 tends to push these seeds far away from each other. We observe that groups with a small number of points appear to contain outliers (see Figure 5.3), so that the resulting subdivision can be helpful in identifying them. Since the computational cost of Algorithm 1 is  $\mathcal{O}(r^2N + p(d)rN)$ , where  $N$  is the original number of points and  $r$  is the number of seeds, as long as the number of groups we are seeking is small, this algorithm is essentially linear. We note that we can subdivide the resulting groups further and, in a hierarchical fashion, build a tree

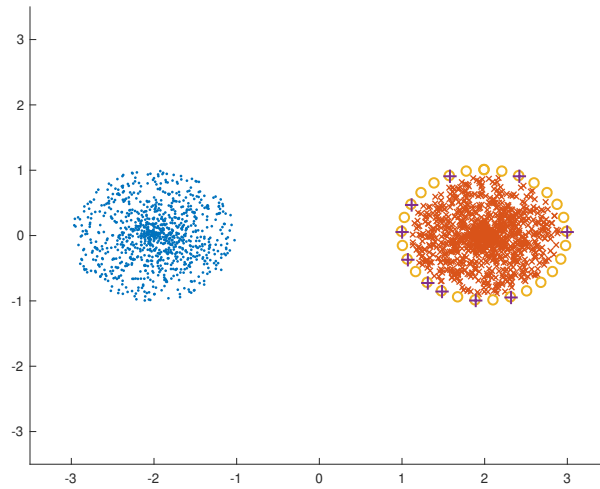


FIGURE 5.1. Example in dimension  $d = 2$ . We display  $M = 1000$  targets (marked with a dot on the left),  $N = 1000$  sources (marked with an “x” on the right), and  $K = 30$  candidate sources (marked with a circle on the right). Algorithm 1 selects 10 equivalent sources from the 30 candidate sources (marked with a +). The relative approximation error in (5.4) is  $1.3e - 07$ .

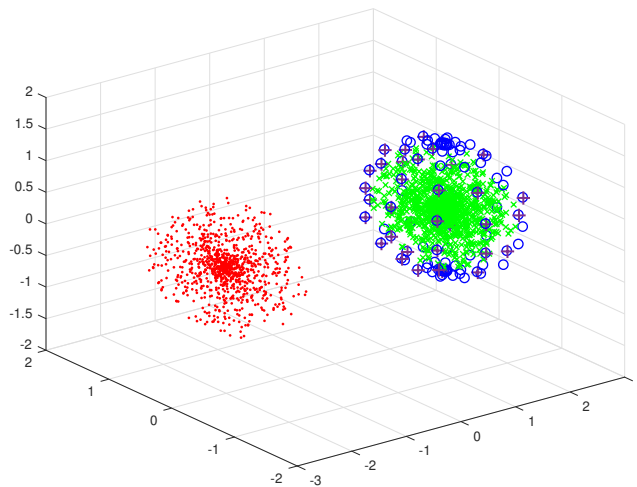


FIGURE 5.2. Example in dimension  $d = 3$ . We display  $M = 1000$  targets (marked with a dot on the left),  $N = 1000$  sources (marked with an “x” on the right), and  $K \times L = 10 \times 10$  candidate sources (marked with a circle on the right). Algorithm 1 selects 35 equivalent sources from the 100 candidate sources (marked with a +). The relative approximation error in (5.4) is  $8.7e - 08$ .

structure. In this paper we simply illustrate the use of Algorithm 1 for subdivision of points into groups and plan to develop applications of this approach elsewhere.



For the example in Figure 5.3 we use the two dimensional distribution of points described in Section 4.2. We choose the bandwidth parameter  $h = 200$  when selecting 4 seeds and  $h = 16$  when selecting 10 seeds in order to obtain the corresponding subdivisions of the set. Observe that outliers tend to be associated with linearly independent terms and, thus, form a subset with a small number of points.

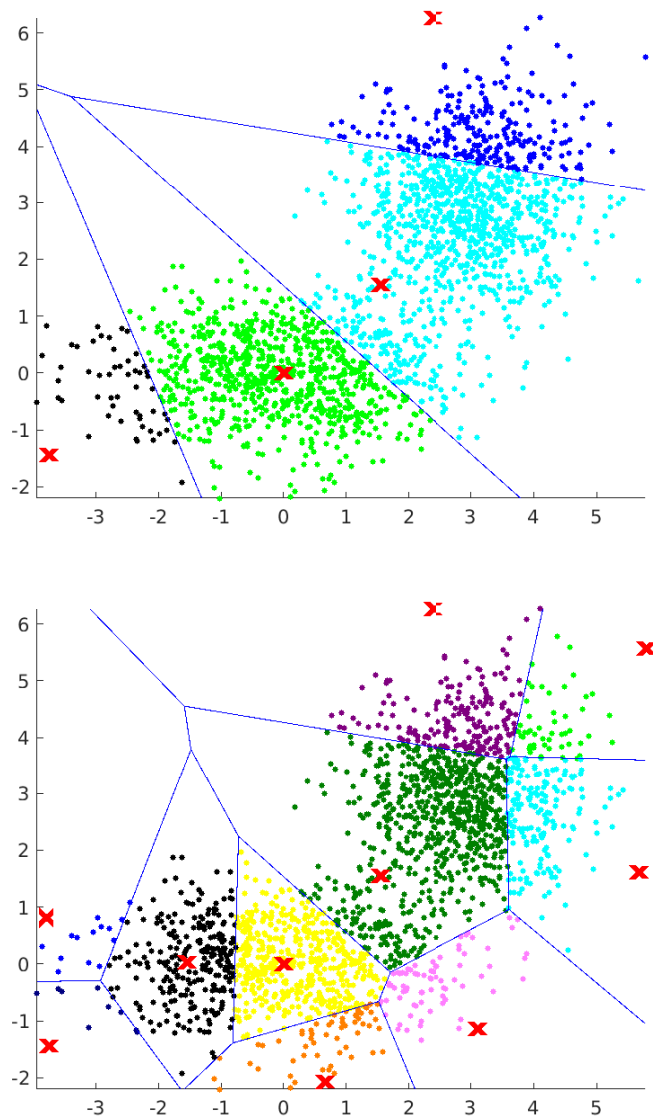


FIGURE 5.3. Subdivision of 2,000 points into groups using seeds (marked by “x”) produced by Algorithm 1. Illustrated are subdivisions into four groups (top) and into ten groups (bottom). Note that groups with a small number of points are likely to contain outliers.

## 6. CONCLUSIONS AND FURTHER WORK

We presented fast algorithms for reducing the number of terms in non-separated multivariate mixtures, analyzed them, and demonstrated their performance on a number of examples. These reduction algorithms allow us to work with non-separated multivariate mixtures which are a far reaching generalization of multivariate separated representations [8, 9, 7] and can be used as a tool for solving a variety of multidimensional problems. Further work is required to develop new numerical methods that use non-separated multivariate mixtures in applications, for example in quantum chemistry (see [14] for some initial results using separated Gaussian mixtures). We plan to pursue several multivariate problems with the techniques illustrated in this paper. Specifically, we plan to develop further our approach to KDE and compare it with existing techniques. We also plan to pursue the problem of hierarchical subdivision of point clouds into groups and its applications to clustering and detection of outliers.

## ACKNOWLEDGEMENTS

We would like to thank Zydrunas Gimbutas (NIST) and the anonymous reviewers for their many useful comments and suggestions.

## 7. APPENDIX A

As mentioned in the paper, computing with multivariate Gaussian mixtures is particularly convenient since all common operations result in explicit integrals. We present below the key identities for multivariate Gaussians using the standard  $L^1$  normalization,

$$N(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\det(2\pi\boldsymbol{\Sigma})^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right).$$

However, when computing integrals with Gaussians atoms, it is convenient to normalize them to have unit  $L^2$ -norm.

7.0.1. *Convolution of two normal distributions.*

$$(7.1) \quad \int_{\mathbb{R}^d} N(\mathbf{x} - \mathbf{y}, \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) N(\mathbf{y}, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2) d\mathbf{y} = N(\mathbf{x}, \boldsymbol{\mu}_1 + \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2).$$

7.0.2. *Sum of two quadratic forms.* Consider vectors  $\mathbf{x}$ ,  $\mathbf{a}$ , and  $\mathbf{b}$  and two symmetric positive definite matrices  $\mathbf{A}$  and  $\mathbf{B}$ . We have

$$(\mathbf{x} - \mathbf{a})^T \mathbf{A}(\mathbf{x} - \mathbf{a}) + (\mathbf{x} - \mathbf{b})^T \mathbf{B}(\mathbf{x} - \mathbf{b}) = (\mathbf{x} - \mathbf{c})^T (\mathbf{A} + \mathbf{B})(\mathbf{x} - \mathbf{c}) + (\mathbf{a} - \mathbf{b})^T \mathbf{C}(\mathbf{a} - \mathbf{b}),$$

where

$$\mathbf{c} = (\mathbf{A} + \mathbf{B})^{-1} (\mathbf{A}\mathbf{a} + \mathbf{B}\mathbf{b})$$

and

$$\mathbf{C} = \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1} \mathbf{B} = (\mathbf{A}^{-1} + \mathbf{B}^{-1})^{-1}.$$

7.0.3. *Product of two normal distributions.* We have

$$(7.2) \quad N(\mathbf{x}, \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) N(\mathbf{x}, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2) = N(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2) \cdot N\left(\mathbf{x}, \boldsymbol{\mu}_c, (\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1})^{-1}\right)$$

where

$$\boldsymbol{\mu}_c = (\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1})^{-1} (\boldsymbol{\Sigma}_1^{-1} \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_2^{-1} \boldsymbol{\mu}_2).$$

7.0.4. *Inner product of two normal distributions.* It follows that

$$(7.3) \quad \int_{\mathbb{R}^d} N(\mathbf{x}, \mu_1, \Sigma_1) N(\mathbf{x}, \mu_2, \Sigma_2) d\mathbf{x} = N(\mu_1, \mu_2, \Sigma_1 + \Sigma_2).$$

Indeed, from (7.1) we have

$$\begin{aligned} \int_{\mathbb{R}^d} N(\mathbf{x}, \mu_1, \Sigma_1) N(\mathbf{x}, \mu_2, \Sigma_2) d\mathbf{x} &= \int_{\mathbb{R}^d} N(2\mu_1 - \mathbf{x}, \mu_1, \Sigma_1) N(\mathbf{x}, \mu_2, \Sigma_2) d\mathbf{x} \\ &= N(2\mu_1, \mu_1 + \mu_2, \Sigma_1 + \Sigma_2) \\ &= N(\mu_1, \mu_2, \Sigma_1 + \Sigma_2). \end{aligned}$$

Alternatively, from (7.2) we have

$$\begin{aligned} \int_{\mathbb{R}^d} N(\mathbf{x}, \mu_1, \Sigma_1) N(\mathbf{x}, \mu_2, \Sigma_2) d\mathbf{x} &= N(\mu_1, \mu_2, \Sigma_1 + \Sigma_2) \int_{\mathbb{R}^d} N(\mathbf{x}, \mu_c, (\Sigma_1^{-1} + \Sigma_2^{-1})^{-1}) d\mathbf{x} \\ &= N(\mu_1, \mu_2, \Sigma_1 + \Sigma_2). \end{aligned}$$

#### REFERENCES

- [1] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions*, volume 55 of *Applied Math Series*. National Bureau of Standards, 1964.
- [2] Bradley K. Alpert. Hybrid Gauss-trapezoidal quadrature rules. *SIAM J. Sci. Comput.*, 20(5):1551–1584, 1999.
- [3] D. Arthur and S. Vassilvitskii. k-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1027–1035. Society for Industrial and Applied Mathematics, 2007.
- [4] A. Bernacchia and S. Pigolotti. Self-consistent method for density estimation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(3):407–422, 2011.
- [5] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms, I. *Comm. Pure Appl. Math.*, 44(2):141–183, 1991. Yale Univ. Technical Report YALEU/DCS/RR-696, August 1989.
- [6] G. Beylkin, G. Fann, R. J. Harrison, C. Kurcz, and L. Monzón. Multiresolution representation of operators with boundary conditions on simple domains. *Appl. Comput. Harmon. Anal.*, 33:109–139, 2012. <http://dx.doi.org/10.1016/j.acha.2011.10.001>.
- [7] G. Beylkin, J. Garcke, and M. J. Mohlenkamp. Multivariate regression and machine learning with sums of separable functions. *SIAM Journal on Scientific Computing*, 31(3):1840–1857, 2009.
- [8] G. Beylkin and M. J. Mohlenkamp. Numerical operator calculus in higher dimensions. *Proc. Natl. Acad. Sci. USA*, 99(16):10246–10251, August 2002.
- [9] G. Beylkin and M. J. Mohlenkamp. Algorithms for numerical analysis in high dimensions. *SIAM J. Sci. Comput.*, 26(6):2133–2159, July 2005.
- [10] G. Beylkin and L. Monzón. On generalized Gaussian quadratures for exponentials and their applications. *Appl. Comput. Harmon. Anal.*, 12(3):332–373, 2002.
- [11] G. Beylkin and L. Monzón. On approximation of functions by exponential sums. *Appl. Comput. Harmon. Anal.*, 19(1):17–48, 2005.
- [12] G. Beylkin and L. Monzón. Approximation of functions by exponential sums revisited. *Appl. Comput. Harmon. Anal.*, 28(2):131–149, 2010.
- [13] G. Beylkin, L. Monzón, and I. Satkauskas. On computing distributions of products of random variables via Gaussian multiresolution analysis. *Appl. Comput. Harmon. Anal.*, 2017. doi: 10.1016/j.acha.2017.08.008, see also arXiv:1611.08580.
- [14] G. Beylkin, L. Monzon, and X. Yang. Adaptive algorithm for electronic structure calculations using reduction of Gaussian mixtures. 2018. submitted to Proc. Royal Soc. A, <https://arxiv.org/abs/1812.09284>.
- [15] D. J. Biagioni, D. Beylkin, and G. Beylkin. Randomized interpolative decomposition of separated representations. *Journal of Computational Physics*, 281:116–134, 2015.
- [16] Åke Björck. Solving linear least squares problems by gram-schmidt orthogonalization. *BIT Numerical Mathematics*, 7(1):1–21, 1967.
- [17] Z. I. Botev, J. F. Grotowski, and D. P. Kroese. Kernel density estimation via diffusion. *The Annals of Statistics*, 38(5):2916–2957, 2010.
- [18] S. F. Boys. The integral formulae for the variational solution of the molecular many-electron wave equations in terms of gaussian functions with direct electronic correlation. *Proceedings of the Royal Society of London Series A- Mathematical and Physical Sciences*, 258(1294):402, 1960.
- [19] R. Bro. Parafac. Tutorial and Applications. *Chemometrics and Intelligent Laboratory Systems*, 38(2):149–171, 1997.
- [20] J. D. Carroll and J. J. Chang. Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition. *Psychometrika*, 35:283–320, 1970.
- [21] G. C. Cawley and N. L. C. Talbot. Reduced rank kernel ridge regression. *Neural Processing Letters*, 16(3):293–302, 2002.

- [22] S. Chandrasekaran and I. C. F. Ipsen. On rank-revealing factorisations. *SIAM Journal on Matrix Analysis and Applications*, 15(2):592–622, 1994.
- [23] Bengt Fornberg and Natasha Flyer. *A primer on radial basis functions with applications to the geosciences*. SIAM, 2015.
- [24] L. Genovese, T. Deutsch, A. Neelov, S. Goedecker, and G. Beylkin. Efficient solution of Poisson’s equation with free boundary conditions. *J. Chem. Phys.*, 125(7), 2006.
- [25] G. Golub and C. Van Loan. *Matrix Computations*. Johns Hopkins University Press, 3rd edition, 1996.
- [26] M. Gu and S. C. Eisenstat. Efficient algorithms for computing a strong rank-revealing qr factorization. *SIAM Journal on Scientific Computing*, 17(4):848–869, 1996.
- [27] N. Halko, P.-G. Martinsson, and J. A. Tropp. Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review*, 53(2):217–288, 2011.
- [28] R. J. Harrison, G. Beylkin, F. A. Bischoff, J. A. Calvin, G. I. Fann, J. Fosso-Tande, D. Galindo, J.R Hammond, R. Hartman-Baker, J. C Hill, et al. Madness: A multiresolution, adaptive numerical environment for scientific simulation. *SIAM J. Sci. Comput.*, 38(5):S123–S142, 2016. see also arXiv preprint arXiv:1507.01888.
- [29] R.J. Harrison, G.I. Fann, T. Yanai, Z. Gan, and G. Beylkin. Multiresolution quantum chemistry: basic theory and initial applications. *J. Chem. Phys.*, 121(23):11587–11598, 2004.
- [30] R. A. Harshman. Foundations of the Parafac procedure: model and conditions for an “explanatory” multi-mode factor analysis. Working Papers in Phonetics 16, UCLA, 1970.
- [31] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, second edition, 2013.
- [32] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51(3):455–500, 2009.
- [33] J. Li, S. Ray, and B. G. Lindsay. A nonparametric statistical approach to clustering via mode identification. *Journal of Machine Learning Research*, 8(Aug):1687–1723, 2007.
- [34] S. Lloyd. Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137, 1982.
- [35] J.V.L. Longstaff and K. Singer. The use of gaussian (exponential quadratic) wave functions in molecular problems. ii. wave functions for the ground state of the hydrogen atom and of hydrogen molecule. *Proc. R. Soc. London Ser. A-Math.*, 258(1294):421, 1960.
- [36] W. B. March and G. Biros. Far-field compression for fast kernel summation methods in high dimensions. *Applied and Computational Harmonic Analysis*, 43(1):39–75, 2017.
- [37] V. Maz’ya and G. Schmidt. *Approximate approximations*, volume 141 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2007.
- [38] P. D. McNicholas. Model-based clustering. *Journal of Classification*, 33(3):331–373, 2016.
- [39] A. Y. Ng, M. I. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in neural information processing systems*, pages 849–856, 2002.
- [40] T. A. O’Brien, K. Kashinath, N. R. Cavanaugh, W. D Collins, and J. P O’Brien. A fast and objective multidimensional kernel density estimation method: fastkde. *Computational Statistics & Data Analysis*, 101:148–160, 2016.
- [41] M. Reynolds, A. Doostan, and G. Beylkin. Randomized alternating least squares for canonical tensor decompositions: application to a PDE with random data. *SIAM J. Sci. Comput.*, pages A2634–A2664, 2016.
- [42] D. W. Scott. *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons, 2015.
- [43] B. W. Silverman. *Density estimation for statistics and data analysis*, volume 26. CRC press, 1986.
- [44] K. Singer. The use of gaussian (exponential quadratic) wave functions in molecular problems. i. General formulae for the evaluation of integrals. *Proc. R. Soc. London Ser. A-Math.*, 258(1294):412, 1960.
- [45] G. Tomasi and R. Bro. A comparison of algorithms for fitting the PARAFAC model. *Comput. Statist. Data Anal.*, 50(7):1700–1734, 2006.
- [46] F. Vico, L. Greengard, and M. Ferrando. Fast convolution with free-space green’s functions. *Journal of Computational Physics*, 323:191–203, 2016.
- [47] T. Yanai, G.I. Fann, Z. Gan, R.J. Harrison, and G. Beylkin. Multiresolution quantum chemistry: Analytic derivatives for Hartree-Fock and density functional theory. *J. Chem. Phys.*, 121(7):2866–2876, 2004.
- [48] T. Yanai, G.I. Fann, Z. Gan, R.J. Harrison, and G. Beylkin. Multiresolution quantum chemistry: Hartree-Fock exchange. *J. Chem. Phys.*, 121(14):6680–6688, 2004.
- [49] L. Ying, G. Biros, and D. Zorin. A kernel-independent adaptive fast multipole algorithm in two and three dimensions. *Journal of Computational Physics*, 196(2):591–626, 2004.