

APPROXIMATE SEPARABILITY OF GREEN'S FUNCTION FOR HIGH FREQUENCY HELMHOLTZ EQUATIONS

BJÖRN ENGQUIST AND HONGKAI ZHAO

ABSTRACT. Approximate separable representations of Green's functions for differential operators is a basic and an important aspect in the analysis of differential equations and in the development of efficient numerical algorithms for solving them. Being able to approximate a Green's function as a sum with few separable terms is equivalent to the existence of low rank approximation of corresponding discretized system. This property can be explored for matrix compression and efficient numerical algorithms. Green's functions for coercive elliptic differential operators in divergence form have been shown to be highly separable and low rank approximation for their discretized systems has been utilized to develop efficient numerical algorithms. The case of Helmholtz equation in the high frequency limit is more challenging both mathematically and numerically. In this work, we develop a new approach to study approximate separability for the Green's function of Helmholtz equation in the high frequency limit based on an explicit characterization of the relation between two Green's functions and a tight dimension estimate for the best linear subspace approximating a set of almost orthogonal vectors. We derive both lower bounds and upper bounds and show their sharpness for cases that are commonly used in practice.

1. INTRODUCTION

Given a linear differential operator, denoted by L , the Green's function, denoted by $G(\mathbf{x}, \mathbf{y})$, is defined as the fundamental solution in a domain $\Omega \subseteq R^n$ to the partial differential equation

$$(1) \quad \begin{cases} L_x G(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}), & \mathbf{x}, \mathbf{y} \in \Omega \subseteq R^n \\ \text{with boundary condition or condition at infinity,} \end{cases}$$

where $\delta(\mathbf{x} - \mathbf{y})$ is the Dirac delta function denoting an impulse source point at \mathbf{y} . In particular, general solutions of a partial differential equation can be obtained by superposition of fundamental solutions with source locations in Ω (and/or boundary of Ω).

Approximate separability of $G(\mathbf{x}, \mathbf{y})$ is defined as the following: given two sets $X, Y \subseteq \Omega \subseteq R^d$ (see Figure 1) and $\epsilon > 0$, there is a smallest N^ϵ such that there are $f_l(\mathbf{x}), g_l(\mathbf{y}), l =$

The work of Björn Engquist is partially supported by NSF grant DMS-1217203 and Texas Consortium for Computational Seismology.

The work of Hongkai Zhao is partially supported by NSF grant DMS-1115698 and ONR grant N00014-11-1-0602.

$1, 2, \dots, N^\epsilon$

$$(2) \quad \left\| G(\mathbf{x}, \mathbf{y}) - \sum_{l=1}^{N^\epsilon} f_l(\mathbf{x}) g_l(\mathbf{y}) \right\|_{X \times Y} \leq \epsilon, \quad \mathbf{x} \in X, \mathbf{y} \in Y,$$

where $\|\cdot\|_{X \times Y}$ is the norm of some function space which G, f_l, g_l belong to. If X and Y are compact and disjoint domains in R^d and $G(\mathbf{x}, \mathbf{y})$ is continuous in $X \times Y$, which is often the case of practical interest, there exists a polynomial approximation of $G(\mathbf{x}, \mathbf{y})$ in $X \times Y$ by Weierstrass approximation theorem which is separable. So there is a $N^\epsilon < \infty$ for any $\epsilon > 0$. The most interesting issue is how N^ϵ depends on ϵ , which manifests the intrinsic complexity of the PDE and its solution within ϵ -approximation. If one views $G(\mathbf{x}, \mathbf{y})$ as a family of functions on X parametrized by $\mathbf{y} \in Y$, this is equivalent to saying that the Kolmogorov n -width [17] for the family of functions $G(\mathbf{x}, \mathbf{y})$ in the $\|\cdot\|_X$ normed function space is ϵ when $n = N^\epsilon$. Kolmogorov n -width, which is used to characterize information content in information theory, is the best approximation of a set S in a normed space W by a n dimensional linear subspace L_n defined as

$$(3) \quad d_n(S, W) := \inf_{L_n} \sup_{f \in S} \inf_{g \in L_n} \|f - g\|_W,$$

Of course the role of \mathbf{x} and \mathbf{y} can be reversed.

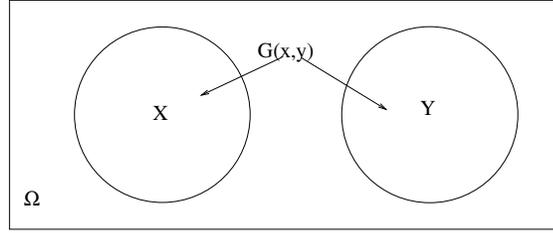


FIGURE 1. Green's function $G(\mathbf{x}, \mathbf{y})$ with dependence on $\mathbf{x} \in X$ and $\mathbf{y} \in Y$.

We introduce the following relations to simplify notations in later derivations. $x \gtrsim y$ means that there is a constant $\infty > c > 0$ such that $x \geq cy$, $x \lesssim y$ means that there is a constant $\infty > C > 0$ such that $x \leq Cy$, and $x \sim y$ means there are two constants $0 < c \leq C < \infty$ such that $cy \leq x \leq Cy$. For our results for Helmholtz equation (5), all constants are independent of the wave number k as $k \rightarrow \infty$.

In this study, we assume $X, Y \subseteq \Omega$ are two compact manifolds embedded in R^d with dimensions $\dim(X)$ and $\dim(Y)$ respectively, i.e., they may be compact domains in R^d , $\dim(X) = \dim(Y) = d = 1, 2, 3$, or compact two dimensional surfaces embedded in R^3 or one dimensional compact curves embedded in R^d , $d = 2, 3$. Without loss of generality, we assume $\dim(X) \geq \dim(Y) = s$.

One can get a sharper upper bound for N^ϵ based simply on regularity of the Green's function. For example, suppose X, Y are two disjoint compact domains in R^d and $G(\mathbf{x}, \mathbf{y})$ is $C^m(X \times Y)$, one can show that $N^\epsilon \lesssim \epsilon^{-\frac{d}{m}}$ using the following argument. Lay down a

uniform grid $\mathbf{y}_j, j = 1, 2, \dots, J \sim \epsilon^{-\frac{d}{m}}$ in Y with a grid size $h \sim \epsilon^{\frac{1}{m}}$. $\forall \mathbf{y} \in Y$, $G(\mathbf{x}, \mathbf{y})$ can be approximated by $(m - 1)$ -th order interpolation/extrapolation of $G(\mathbf{x}, \mathbf{y}_j)$, which is a linear combination of $G(\mathbf{x}, \mathbf{y}_j)$ and satisfies

$$(4) \quad |G(\cdot, \mathbf{y}) - \sum_{j: \mathbf{y}_j \in B_\delta(\mathbf{y})} a_j G(\cdot, \mathbf{y}_j)| \leq \|D_y^m G(\cdot, \cdot)\|_{L^\infty(X \times Y)} h^m \leq \epsilon,$$

where $B_\delta(\mathbf{y})$ is a ball centered at \mathbf{y} with a radius $\delta \sim h$. If Y is a rectangular domain and $G(\mathbf{x}, \mathbf{y})$, one can even use We call $G(\mathbf{x}, \mathbf{y})$ highly separable if N^ϵ depends on ϵ weakly, $N^\epsilon \leq O(|\log \epsilon|^p)$, for some $p > 0$.

When a linear PDE, such as (1), is discretized and numerically solved, high separability of its Green's function implies existence of low rank approximation of subsets of the discretized system, which provides a matrix compression and lies at the heart for many efficient algorithms. Typically low rank approximation has been used in two ways. One way is to utilize low rank approximation of the discretized Green's function, which is the kernel for boundary integral formulation, for fast matrix vector multiplication when solving boundary integral equations by iterative methods [8, 11, 12, 26, 27]. Similarly, it has been used to develop fast algorithms for evaluation of fast oscillatory scattering operator and Fourier integral operators [6, 24, 25]. The other way is to utilize low rank property to develop fast algorithms to solve a large linear system $Ax = b$ corresponding to a discretization of a PDE such as (1). Each column of the inverse matrix A^{-1} is a numerical approximation of the Green's function. Again low rank approximation for off-diagonal submatrices of A^{-1} , which is implied by high separability of the Green's function on two disjoint sets, is extensively explored in many fast algorithms to solve the linear system such as hierarchical matrix and structured inverse methods [2, 3, 7, 15, 23, 29, 31, 32]. Often the low rank approximation is computed or learned on the fly. Fast random algorithms or rank revealing type of methods [14, 22] can be used to find the leading singular values and corresponding singular vectors of a matrix. However, the computation cost increases dramatically if the rank is not sufficiently low. So both upper and lower bound estimates for approximate separability is of crucial importance in these applications.

In the literature, mostly upper bound estimates for highly separable cases were shown for Green's functions or kernel functions when developing fast numerical algorithms based on low rank approximation as mentioned above. These estimates are typically based on constructive approaches for Green's functions or kernel functions with explicit expression and using asymptotic expansion. Interesting study on spatial bandwidth and degree of freedom of scattered field have also been done in the engineering literature [4, 5], which shows that the scattered field is almost band limited and the degree of freedom is close to the Nyquist number in term of the effective (spatial) bandwidth of the scattered field and to the extension of the observation domain. A more general non-constructive mathematical approach was developed in [2] to show that the Green's function for a coercive elliptic operator in divergence form with L_∞ -coefficients is highly separable (Theorem 2.8) for two disjoint domains X, Y based on a key gradient estimate by Caccioppoli inequality. The method and result can be extended to Green's function of more general elliptic equations

with non-dominant lower order terms, such as convection-diffusion equations with small convection term or the Helmholtz equation (5) with small wave number k . Their method does not work when the lower order term is dominant which is the case for the Helmholtz equation with large k . It becomes a singularly perturbed problem and the gradient of the Green's function is unbounded almost everywhere as $k \rightarrow \infty$ due to fast oscillations. These issues are also reflected in numerical computation for these different PDEs. It is well known that there are many efficient numerical methods to solve the discretized system corresponding to differential operators that are elliptic dominant, such as iterative methods with various effective preconditioners and direct inverse methods as mentioned above. This is related to the intrinsic complexity manifested by the high separability of the corresponding Green's functions. On the other hand, it is well known that Helmholtz equation with large wave number is very difficult to solve numerically in practice. For example, all those well developed iterative methods for elliptic equations do not work effectively for this case [10].

Here we give another mathematics perspective by showing lower bounds for the approximate separability of the Green's function for Helmholtz equation in high frequency limit in terms of both ϵ and k . The lower bounds, which are sharp for many practical setups, show that the Green's function is not highly separable as $k \rightarrow \infty$ and manifests the intrinsic complexity of the solution space. In our study we give

- explicit characterization of the correlation or angle (in $L2$ normed space) between two Green's functions of Helmholtz equation (5) in the high frequency limit,

$$(\|G(\cdot, \mathbf{y}_1)\|_2 \|G(\cdot, \mathbf{y}_2)\|_2)^{-1} \int_X G(\mathbf{x}, \mathbf{y}_1) \overline{G(\mathbf{x}, \mathbf{y}_2)} d\mathbf{x} \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\alpha}, \quad k|\mathbf{y}_1 - \mathbf{y}_2| \rightarrow \infty$$

for some $\alpha > 0$ which depends on the dimension of X , its geometry and the locations of $\mathbf{y}_1, \mathbf{y}_2$ (see Theorem 2.1) based on generalized stationary phase analysis.

- lower bound estimate for the approximate separability for the Green's function of Helmholtz equation in the high frequency limit

$$N_k^\epsilon \gtrsim \begin{cases} k^{2\alpha}, & \alpha < \frac{s}{2}, \\ k^{s-\delta}, & \alpha \geq \frac{s}{2}, \end{cases}$$

and upper bound estimate

$$N_k^\epsilon \lesssim k^{s+\delta}$$

as $k \rightarrow \infty$ for two compact manifolds X and Y with $\dim(X) \geq \dim(Y) = s$ and any $\delta > 0$, where the constants in \gtrsim and \lesssim are independent of k for a fixed small ϵ (see Lemma 3.1 and Theorem 3.1 - 3.4). The lower bound is based on a tight dimension estimate improved from that for a set of nearly orthogonal random vectors by N. Alon [1] and Johnson-Lindenstrauss Lemma [19].

- explicit estimates and their sharpness for situations that are commonly used in practice (see Section 4). Our theory is also applied to show precise conditions if high separability (or low rank approximation after numerical discretization) can or can not be achieved for special set ups.

As far as we know, lower bound estimate for approximate separability of Green's functions of this type is the first in the literature. These bounds mathematically characterize intrinsic complexities for high frequency wave phenomena. We hope these studies and understandings can provide useful insights for developing fast numerical algorithms as well.

2. HELMHOLTZ EQUATION AND ITS GREEN'S FUNCTION

Let $G(\mathbf{x}, \mathbf{y})$ be the Green's function to the Helmholtz equation in free space,

$$(5) \quad \Delta_{\mathbf{x}}G(\mathbf{x}, \mathbf{y}) + k^2n^2(\mathbf{x})G(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in R^d,$$

where $k > 0$ is the wave number, $0 < c < n(\mathbf{x}) < C < \infty$ is the index of refraction and $\delta(\mathbf{x} - \mathbf{y})$ denotes a point source at \mathbf{y} . Suitable far field radiation condition has to be satisfied for uniqueness. The high frequency limit means the wave number $k \rightarrow \infty$, which poses challenge both mathematically and numerically due to faster and faster oscillations in the solution.

For completeness, we provide the general formula for the free space Green's function of Helmholtz equation (5) for any space dimension,

$$(6) \quad G_0(\mathbf{x}, \mathbf{y}) = c_d k^p \frac{H_p^{(1)}(k|\mathbf{x} - \mathbf{y}|)}{|\mathbf{x} - \mathbf{y}|^p}, \quad p = \frac{d-2}{2}, \quad c_d = \frac{1}{2i(2\pi)^p}, \quad \mathbf{x}, \mathbf{y} \in R^d, \mathbf{x} \neq \mathbf{y}.$$

$H_p^{(1)}(r)$ is the first kind Hankel function of order p which has the following asymptotic behavior: as $r \rightarrow 0$

$$(7) \quad H_p^{(1)}(r) = \begin{cases} -\frac{i}{\pi} \Gamma(p) \left(\frac{2}{r}\right)^p, & p \neq 0 \\ \frac{2i}{\pi} \log r, & p = 0 \end{cases}$$

where $\Gamma(p)$ is the Gamma function, and as $r \rightarrow \infty$

$$(8) \quad H_p^{(1)}(r) = \left(\frac{2}{\pi r}\right)^{\frac{1}{2}} e^{i(r - \frac{p\pi}{2} - \frac{\pi}{4})} + O(r^{-\frac{3}{2}}), \quad p \geq 0.$$

For $d = 3$, the Green's function takes the simplest form

$$(9) \quad G_0(\mathbf{x}, \mathbf{y}) = \frac{1}{4\pi} \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{|\mathbf{x} - \mathbf{y}|}, \quad \mathbf{x} \neq \mathbf{y}.$$

For $d = 2$,

$$(10) \quad G_0(\mathbf{x}, \mathbf{y}) = -\frac{i}{4} H_0^{(1)}(k|\mathbf{x} - \mathbf{y}|) = -\frac{1}{2\pi} \int_0^\infty e^{ik|\mathbf{x}-\mathbf{y}| \cosh \theta} d\theta, \quad \mathbf{x} \neq \mathbf{y},$$

and

$$(11) \quad \lim_{r \rightarrow 0^+} H_0^{(1)}(r) = \frac{2i}{\pi} \log r, \quad \lim_{r \rightarrow \infty} H_0^{(1)}(r) = \sqrt{\frac{2}{\pi r}} e^{i(r - \pi/4)} + O(r^{-3/2}).$$

Denote $B_\tau^d(\mathbf{y})$ and $S_\rho^d(\mathbf{y})$ to be a ball and a sphere in R^d centered at \mathbf{y} with radius τ and ρ respectively. We have

$$(12) \quad \int_{B_\tau^d(\mathbf{y})} |G_0(\mathbf{x}, \mathbf{y})|^2 d\mathbf{x} = c_d^2 k^{2p} \int_0^\tau d\rho \int_{S_\rho^d} \frac{[H_p^{(1)}(k\rho)]^2}{\rho^{2p}} ds = c_d^2 \omega_d k^{2p-2} \int_0^{k\tau} r [H_p^{(1)}(r)]^2 dr,$$

where $\omega_d = \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})}$ is the area of the unit sphere in R^d . From the asymptotic formula (7), we see that G_0 is square integrable at the singular source for $d = 2, 3$. Also from the asymptotic formula (8), we have $\|G_0(\cdot, \mathbf{y})\|_{2(B_\tau^d(\mathbf{y}))} \sim k^{\frac{d-3}{2}}$ as $k \rightarrow \infty$.

From the above explicit expressions for free space Green's function, we see that except for the case $d = 3$, there is a multiplication factor related to k for the magnitude of the Green's function. To characterize the angle or correlation between two Green's function and study the separability of the Green's function without the effect of this factor, we define the normalized Green's function as

$$(13) \quad \hat{G}(\mathbf{x}, \mathbf{y}) = \frac{G(\mathbf{x}, \mathbf{y})}{\|G(\cdot, \mathbf{y})\|_2}, \quad \mathbf{x} \in X \subset R^d, \quad \|G(\cdot, \mathbf{y})\|_2^2 = \int_X |G(\mathbf{x}, \mathbf{y})|^2 d\mathbf{x},$$

in our later study with the following understandings:

- $\|G(\cdot, \mathbf{y})\|_2$ is a smooth function of \mathbf{y} since fast oscillation due to rapid change of phase function is removed.
- When $d = 3$, all results for the normalized Green's function $\hat{G}(\mathbf{x}, \mathbf{y})$ can be extended to $G(\mathbf{x}, \mathbf{y})$ since there are constants $0 < c < C < \infty$ that are independent of k such that $c < |G(\mathbf{x}, \mathbf{y})| < C, \forall \mathbf{x} \in X, \mathbf{y} \in Y$, once two compact sets $X, Y \subset R^3$ are fixed.
- In this paper we prove results for $d = 2, 3$ for practical interest. Since the Green's function is square integrable at the source singularity, we allow overlaps between two compact domains X and Y when $\dim(X) = \dim(Y) = d = 2, 3$. All results for bounded and disjoint X and Y can be extended to $d > 3$.

Approximate separability of $G(\mathbf{x}, \mathbf{y})$ is defined as in (2) except that we now put a subscript k in N_k^ϵ to specifically indicate the dependence on k . The key issue is the dependence on k for a given $\epsilon > 0$ as $k \rightarrow \infty$. In practice, such as development of fast algorithms utilizing low rank approximation for the discretized system, X and Y are often disjoint and compact. Typical norms used are either $L_\infty(X \times Y)$ or $L_2(X \times Y)$. In our study we first show analysis and results in L_2 norm, which fits well with using SVD (singular value decomposition) for low rank approximation of a matrix, and then extend those results to L_∞ norm. Regarding $G(\mathbf{x}, \mathbf{y})$ as a family of functions in $L_2(X)$ parametrized by $\mathbf{y} \in Y$ (see Figure 1), the separability condition (2) in $L_2(X \times Y)$ is equivalent to the existence of a linear subspace $S_X \subset L_2(X)$ with dimension N_k^ϵ such that

$$(14) \quad \sqrt{\int_Y \|G(\mathbf{x}, \mathbf{y}) - P_{S_X} G(\mathbf{x}, \mathbf{y})\|_{L_2(X)}^2 d\mathbf{y}} \leq \epsilon,$$

where $P_{S_X} G(\mathbf{x}, \mathbf{y})$ is the projection of $G(\mathbf{x}, \mathbf{y})$ in S_X . This formulation is the same with the role of \mathbf{x} and \mathbf{y} exchanged.

We start with the study of approximate separability of the Green's function, $G_0(\mathbf{x}, \mathbf{y})$, for homogeneous medium, i.e., $n^2(x) \equiv 1$ in (5), in free space. When 2-norm is used as the metric, one important geometric characterization of relation between two vectors is the angle or correlation between them. Let $X \subset R^3$ be a compact domain and $\mathbf{y}_1, \mathbf{y}_2 \notin X$ be two points with $\delta = |\mathbf{y}_1 - \mathbf{y}_2| \ll 1$. It is easy to see that

$$\left| \langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle - 1 \right| = \left| \frac{1}{\|G_0(\cdot, \mathbf{y}_1)\|_2 \|G_0(\cdot, \mathbf{y}_2)\|_2} \int_X \frac{e^{ik(|\mathbf{x}-\mathbf{y}_1| - |\mathbf{x}-\mathbf{y}_2|)}}{|\mathbf{x}-\mathbf{y}_1| |\mathbf{x}-\mathbf{y}_2|} d\mathbf{x} - 1 \right| \lesssim k\delta$$

where the constant in \lesssim depends on domain X and the distance between X and $\mathbf{y}_1, \mathbf{y}_2$. In another word, two Green's functions become more and more correlated when the two source points become closer and closer. This is true in general for Green's function as long as $G(\mathbf{x}, \mathbf{y})$ is Lipschitz in \mathbf{y} . Actually for strictly elliptic operator of the following divergence form in $R^d, d \geq 3$,

$$(15) \quad L = - \sum_{i,j=1}^d \frac{\partial}{\partial x_j} (a_{ij}(\mathbf{x}) \frac{\partial}{\partial x_j}), \quad \lambda |\xi|^2 \leq \sum_{i,j=1}^d a_{ij}(\mathbf{x}) \xi_i \xi_j \leq \mu |\xi|^2,$$

where $a_{ij}(\mathbf{x})$ are bounded measurable functions and $0 < \lambda \leq \mu < \infty$ are two constants, there exists a unique Green's function [20, 13] $G(\mathbf{x}, \mathbf{y})$ satisfying

$$(16) \quad c(d, \lambda, \mu) |\mathbf{x} - \mathbf{y}|^{2-d} \leq G(\mathbf{x}, \mathbf{y}) \leq C(d, \lambda, \mu) |\mathbf{x} - \mathbf{y}|^{2-d},$$

where $0 < c(d, \lambda, \mu) < C(d, \lambda, \mu) < \infty$ are two constants. Given a compact domain $X \subset R^d$ and two points $\mathbf{y}_1, \mathbf{y}_2 \notin X$, define

$$\rho = \min[\min_{\mathbf{x} \in X} |\mathbf{x} - \mathbf{y}_1|, \min_{\mathbf{x} \in X} |\mathbf{x} - \mathbf{y}_2|], \quad K = \frac{C(d, \lambda, \mu)}{c(d, \lambda, \mu)} \left[1 + \frac{|\mathbf{y}_1 - \mathbf{y}_2|}{\rho} \right]^{d-2}.$$

Then we have

$$\frac{G(\mathbf{x}, \mathbf{y}_2)}{G(\mathbf{x}, \mathbf{y}_1)} \leq \frac{C(d, \lambda, \mu)}{c(d, \lambda, \mu)} \left[\frac{|\mathbf{x} - \mathbf{y}_1|}{|\mathbf{x} - \mathbf{y}_2|} \right]^{d-2} \leq \frac{C(d, \lambda, \mu)}{c(d, \lambda, \mu)} \left[\frac{|\mathbf{x} - \mathbf{y}_2| + |\mathbf{y}_1 - \mathbf{y}_2|}{|\mathbf{x} - \mathbf{y}_2|} \right]^{d-2} \leq K,$$

and vice versa. Given two disjoint compact domains $X, Y \subset R^d, d \geq 3$, with ρ being the distance between the two domains and r being the diameter of Y , the correlation between any two Green's function with sources at $\mathbf{y}_1, \mathbf{y}_2 \in Y$ is bounded by

$$(17) \quad 1 \geq \langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle_X \geq \tilde{K}^{-2}, \quad \tilde{K} = \frac{C(d, \lambda, \mu)}{c(d, \lambda, \mu)} \left[1 + \frac{r}{\rho} \right]^{d-2}.$$

Also Caccioppoli inequality gives a L_2 norm bound of the gradient of the Green's function away from the source singularity which is used in [2] to show that the Green's function for elliptic operator (and Helmholtz equation with small k) is highly separable. However, the picture is quite different in the more challenging regime of high frequency limit since the Green's function becomes more and more oscillatory as $k \rightarrow \infty$.

2.1. Decorrelation of Two Green's Function in High Frequency Limit in Homogeneous Medium. Here we study the angle between two Green's functions or how fast they decorrelate in term of the ratio of the separation distance of the two source points with respect to the wave length due to fast oscillations. Stationary phase theory will play an important role here. Define

$$(18) \quad \tilde{\phi}(\mathbf{x}) = |\mathbf{y}_1 - \mathbf{y}_2|^{-1} (|\mathbf{x} - \mathbf{y}_1| - |\mathbf{x} - \mathbf{y}_2|).$$

We have $|\tilde{\phi}(\mathbf{x})| \leq 1$ and

$$(19) \quad \begin{aligned} \nabla \tilde{\phi}(\mathbf{x}) &= |\mathbf{y}_1 - \mathbf{y}_2|^{-1} \left(\frac{\mathbf{x} - \mathbf{y}_1}{|\mathbf{x} - \mathbf{y}_1|} - \frac{\mathbf{x} - \mathbf{y}_2}{|\mathbf{x} - \mathbf{y}_2|} \right) \\ D^2 \tilde{\phi}(\mathbf{x}) &= |\mathbf{y}_1 - \mathbf{y}_2|^{-1} \left[\frac{I - \frac{(\mathbf{x} - \mathbf{y}_1)(\mathbf{x} - \mathbf{y}_1)^T}{|\mathbf{x} - \mathbf{y}_1|^2}}{|\mathbf{x} - \mathbf{y}_1|} - \frac{I - \frac{(\mathbf{x} - \mathbf{y}_2)(\mathbf{x} - \mathbf{y}_2)^T}{|\mathbf{x} - \mathbf{y}_2|^2}}{|\mathbf{x} - \mathbf{y}_2|} \right]. \end{aligned}$$

$|\nabla \tilde{\phi}(\mathbf{x})| \neq 0$ except for points on the line going through $\mathbf{y}_1, \mathbf{y}_2$ and outside the interval between $\mathbf{y}_1, \mathbf{y}_2$, where maximum value 1 or minimum value -1 of $\tilde{\phi}$ is attained (see Figure 2). Also $D^2 \tilde{\phi}(\mathbf{x})$ is degenerate in the direction of $\mathbf{y}_1 - \mathbf{y}_2$. However, for \mathbf{x} on the line and outside the interval between $\mathbf{y}_1, \mathbf{y}_2$, the Hessian in the plane perpendicular to the line, denoted by D_{\perp}^2 , is a multiple of identity matrix I_{\perp} in the plane,

$$(20) \quad D_{\perp}^2 \tilde{\phi}(\mathbf{x}) = \frac{\pm 1}{|\mathbf{x} - \mathbf{y}_1| |\mathbf{x} - \mathbf{y}_2|} I_{\perp},$$

where the sign depends whether maximum or minimum is attained at \mathbf{x} .

From the stationary phase result [16, 30] for $I(k) = \int e^{ik\phi(\mathbf{x})} u(\mathbf{x}) dx$, where $u \in C_c^{\infty}(R^d)$ and $\phi(\mathbf{x})$ has isolated stationary points $\mathbf{x}_m, m = 1, 2, \dots, M$, $|\nabla \phi(\mathbf{x}_m)| = 0$ and $D^2 \phi(\mathbf{x}_m)$ non-degenerate, one has

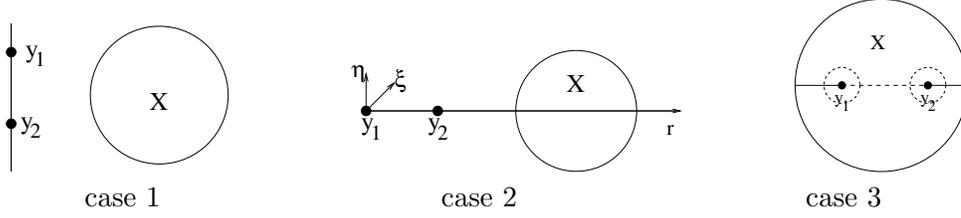
$$(21) \quad \begin{aligned} & \left| I(k) - \left(\frac{2\pi}{k}\right)^{d/2} \sum_{m=1}^M \frac{e^{ik\phi(\mathbf{x}_m)}}{|\det[D^2 \phi(\mathbf{x}_m)]|^{1/2}} e^{\frac{i\pi}{4} \text{sgn}(D^2 \phi(\mathbf{x}_m))} u(\mathbf{x}_m) \right| \\ & \leq C k^{-d/2-1} \|[D^2 \phi(\mathbf{x}_m)]^{-1}\|^{d/2+1} \sum_{\beta \leq s+2} \|D^{\beta} u\|_{L^2}, \end{aligned}$$

where $s > d/2$ and C is an universal constant independent of ϕ and u . However, we have to modify the standard stationary phase technique due to the following three complications in our case: (1) the stationary points may not be isolated, (2) the integration is on a compact domain X and the integrand u is not $C_0^{\infty}(X)$, and (3) there may be singularities in the integrand. Here is our result.

Theorem 2.1. *Assume $X \subset R^d, d = 2, 3$ is a compact domain with piecewise smooth boundary. $\hat{G}_0(\mathbf{x}, \mathbf{y})$ is the normalized free space Green's function. Depending on positions of $\mathbf{y}_1, \mathbf{y}_2$ relative to X and its boundary, there is some $\alpha > 0$ such that*

$$(22) \quad \left| \langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle \right| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\alpha}, \quad \min\{1, \frac{d-1}{2}\} \leq \alpha \leq \frac{d+1}{2}$$

as $k|\mathbf{y}_1 - \mathbf{y}_2| \rightarrow \infty$. The constant in \lesssim depends on X and the distances from $\mathbf{y}_1, \mathbf{y}_2$ to X .


 FIGURE 2. Different positions of $\mathbf{y}_1, \mathbf{y}_2$ relative to X

Proof. We prove for $d = 3$ first. Define

$$(23) \quad \tilde{k} = k|\mathbf{y}_1 - \mathbf{y}_2|, \quad \tilde{\phi}(\mathbf{x}) = |\mathbf{y}_1 - \mathbf{y}_2|^{-1}(|\mathbf{x} - \mathbf{y}_1| - |\mathbf{x} - \mathbf{y}_2|),$$

$$u(\mathbf{x}) = \frac{1}{\|G_0(\cdot, \mathbf{y}_1)\|_2 \|G_0(\cdot, \mathbf{y}_2)\|_2 |\mathbf{x} - \mathbf{y}_1| |\mathbf{x} - \mathbf{y}_2|},$$

and the operator

$$(24) \quad L = \frac{1}{|\nabla \tilde{\phi}(\mathbf{x})|^2} \sum_{j=1}^3 \frac{\partial \tilde{\phi}}{\partial x_j} \frac{\partial}{\partial x_j} \quad L^T = - \sum_{j=1}^3 \frac{\partial}{\partial x_j} \frac{1}{|\nabla \tilde{\phi}(\mathbf{x})|^2} \frac{\partial \tilde{\phi}}{\partial x_j}.$$

we have

$$(25) \quad \left| \langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle \right| = \left| \int_X e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u(\mathbf{x}) d\mathbf{x} \right|$$

Denote the line going through $\mathbf{y}_1, \mathbf{y}_2$ by $l_{\mathbf{y}_1}^{\mathbf{y}_2}$ and the part of $l_{\mathbf{y}_1}^{\mathbf{y}_2}$ outside the open interval between $\mathbf{y}_1, \mathbf{y}_2$ by $\tilde{l}_{\mathbf{y}_1}^{\mathbf{y}_2}$. Depending on the positions of $\mathbf{y}_1, \mathbf{y}_2$ relative to the domain X , we consider the three generic cases illustrated in Figure 2. All other cases can be deduced from these three cases.

Case 1. $\tilde{l}_{\mathbf{y}_1}^{\mathbf{y}_2} \cap X = \emptyset$, see Figure 2. Since there is no stationary point in X , i.e., $|\nabla \tilde{\phi}(\mathbf{x})| > c > 0, \forall \mathbf{x} \in X$, and $u(\mathbf{x})$ is smooth in X , from integration by part we have

$$(26) \quad \begin{aligned} \int_X e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u(\mathbf{x}) d\mathbf{x} &= \frac{1}{i\tilde{k}} \int_X (L e^{i\tilde{k}\tilde{\phi}(\mathbf{x})}) u(\mathbf{x}) d\mathbf{x} \\ &= \frac{1}{i\tilde{k}} \left[\int_X e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} (L^T u(\mathbf{x})) d\mathbf{x} + \int_{\partial X} |\nabla \tilde{\phi}(\mathbf{x})|^{-2} \left(\sum_{j=1}^3 \nu_j \frac{\partial \tilde{\phi}}{\partial x_j} \right) e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u(\mathbf{x}) dS(\mathbf{x}) \right] \\ &= -\frac{1}{\tilde{k}^2} \left[\int_X e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} ((L^T)^2 u(\mathbf{x})) d\mathbf{x} + \int_{\partial X} |\nabla \tilde{\phi}(\mathbf{x})|^{-2} \left(\sum_{j=1}^3 \nu_j \frac{\partial \tilde{\phi}}{\partial x_j} \right) e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} L^T u(\mathbf{x}) dS(\mathbf{x}) \right] \\ &\quad + \frac{1}{i\tilde{k}} \int_{\partial X} |\nabla \tilde{\phi}(\mathbf{x})|^{-2} \left(\sum_{j=1}^3 \nu_j \frac{\partial \tilde{\phi}}{\partial x_j} \right) e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u(\mathbf{x}) dS(\mathbf{x}). \end{aligned}$$

Integration by part can be continued. However, the leading term is the last term which is an oscillatory integral on the boundary ∂X . If the phase function $\tilde{\phi}(\mathbf{x})$ has isolated local minima and maxima on ∂X and $D^2 \tilde{\phi}(\mathbf{x})$ is not degenerate along ∂X at those extrema, the

boundary integral in the last term is $\lesssim \tilde{k}^{-\frac{d-1}{2}}$ by the stationary phase theory. Hence

$$(27) \quad \left| \int_X e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u(\mathbf{x}) d\mathbf{x} \right| \lesssim \tilde{k}^{-\frac{d+1}{2}}, \quad \text{as } \tilde{k} \rightarrow \infty.$$

If there is a piece of the boundary ∂X stays on a level set of $\tilde{\phi}(\mathbf{x})$, see Figure 3, the phase function $\tilde{\phi}(\mathbf{x})$ is constant on that piece, the boundary integral in the last term is $\lesssim 1$ and hence

$$(28) \quad \left| \int_X e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u(\mathbf{x}) d\mathbf{x} \right| \lesssim \tilde{k}^{-1}, \quad \text{as } \tilde{k} \rightarrow \infty.$$

All other scenarios are bounded in between.

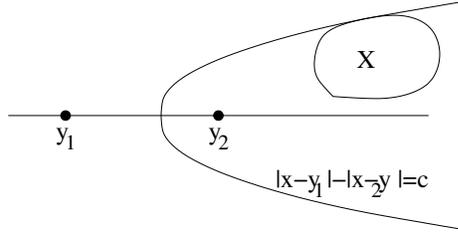


FIGURE 3. A piece of the boundary ∂X stays on a level set of $\tilde{\phi}(\mathbf{x})$

Case 2. $\mathbf{y}_1, \mathbf{y}_2$ are outside X but $\tilde{l}_{\mathbf{y}_1}^2 \cap X \neq \emptyset$, see Figure 2. Both $\tilde{\phi}(\mathbf{x})$ and $u(\mathbf{x})$ are smooth in X . However, all points on the line segment $\tilde{l}_{\mathbf{y}_1}^2 \cap X$ are stationary points with the same phase. Let's use a new coordinate system to evaluate the integral. The new orthogonal coordinate system is (r, ξ, η) , where the origin is at \mathbf{y}_1 and r -axis is in the direction $\mathbf{y}_1 - \mathbf{y}_2$, and (ξ, η) is an orthogonal system perpendicular to r , see Figure 2 (b). So

$$(29) \quad \int_X e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u(\mathbf{x}) d\mathbf{x} = \int_{r_1}^{r_2} \int_{X(r)} e^{i\tilde{k}\tilde{\phi}(r, \xi, \eta)} u(r, \xi, \eta) d\xi d\eta dr$$

where $X(r)$ denotes the intersection of X with the plane (r, ξ, η) at a fixed r and $r_1 = \min_{(r, \xi, \eta) \in X} r$, $r_2 = \max_{(r, \xi, \eta) \in X} r$.

For a fixed $r \in [r_1, r_2]$, if $\tilde{l}_{\mathbf{y}_1}^2 \cap X(r) = (r, 0, 0)$, it is the only stationary point in the plane (r, ξ, η) . Moreover, we have

$$(30) \quad \begin{aligned} \tilde{\phi}(r, 0, 0) &= 1, \\ D_{\xi\eta}^2 \tilde{\phi}(r, 0, 0) &= \frac{1}{r(|\mathbf{y}_1 - \mathbf{y}_2| + r)} I, \end{aligned}$$

$$u_1(r, 0, 0)|_{\psi=0} = \frac{1}{\|G_0(\cdot, \mathbf{y}_1)\|_2 \|G_0(\cdot, \mathbf{y}_2)\|_2 r (|\mathbf{y}_1 - \mathbf{y}_2| + r)},$$

where I is a 2×2 identity matrix. For each r , one can use a partition of unity for the domain $X(r)$ in (ξ, η) plane: $\chi_1(r, \xi, \eta) + \chi_2(r, \xi, \eta) \equiv 1, \forall (r, \xi, \eta) \in X(r)$. $0 \leq \chi_1(r, \xi, \eta) \leq 1$ is

smooth and is 1 in a small ball centered at $(r, 0, 0)$ and inside $X(r)$. $\chi_1(r, \xi, \eta)$ is zero near the boundary of $\partial X(r)$. Denote $u_i = \chi_i u, i = 1, 2$, then

$$(31) \quad \int_{X(r)} e^{i\tilde{k}\tilde{\phi}(r,\xi,\eta)} u(r, \xi, \eta) d\xi d\eta = \int_{X(r)} e^{i\tilde{k}\tilde{\phi}(r,\xi,\eta)} u_1(r, \xi, \eta) d\xi d\eta + \int_{X(r)} e^{i\tilde{k}\tilde{\phi}(r,\xi,\eta)} u_2(r, \xi, \eta) d\xi d\eta$$

Since there is no stationary point in the second integral, one can use integration by part argument as in case 1 to show that it is $\lesssim \tilde{k}^{-1}$. For the first integral, $(r, 0, 0)$ is the only stationary point. Apply the standard stationary phase result and from (30) we get

$$(32) \quad \left| \int_{X(r)} e^{i\tilde{k}\tilde{\phi}(r,\xi,\eta)} u_1(r, \xi, \eta) d\xi d\eta - \frac{2\pi i \tilde{k}^{-\frac{d-1}{2}} e^{i\tilde{k}\tilde{\phi}(r,0,0)}}{\|G_0(\cdot, \mathbf{y}_1)\|_2 \|G_0(\cdot, \mathbf{y}_2)\|_2} \right| \lesssim \tilde{k}^{-\frac{d+1}{2}}$$

It is important to note that the phase in the leading term after integration in (ξ, η) over $X(r)$ is independent of r , which means no fast oscillation when integrating in r .

For a fixed $r \in [r_1, r_2]$, if $\tilde{l}_{\mathbf{y}_1}^2 \cap X(r) = \emptyset$, there is no stationary point in $X(r)$. Hence $\int_{X(r)} e^{i\tilde{k}\tilde{\phi}(r,\xi,\eta)} u_2(r, \xi, \eta) d\xi d\eta$ will be less than the case when $\tilde{l}_{\mathbf{y}_1}^2 \cap X(r) \neq \emptyset$. So we have

$$(33) \quad \left| \int_X e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u(\mathbf{x}) d\mathbf{x} \right| = \left| \int_{r_1}^{r_2} \int_{X(r)} e^{i\tilde{k}\tilde{\phi}(r,\xi,\eta)} u(r, \xi, \eta) d\xi d\eta dr \right| \lesssim \tilde{k}^{\min\{-1, \frac{1-d}{2}\}} = \tilde{k}^{-1},$$

since $\frac{d-1}{2} = 1$ for $d = 3$.

Case 3. Let's consider the most general case where \mathbf{y}_1 and (or) \mathbf{y}_2 are in the interior of X , see Figure 2. The main contribution still comes from the line of stationary points $\tilde{l}_{\mathbf{y}_1}^2 \cap X$. However, singularities of u at \mathbf{y}_1 and \mathbf{y}_2 have to be taken care of. Assume that there is a ball with radius $r_0 < |\mathbf{y}_1 - \mathbf{y}_2|/4$ around each point $\mathbf{y}_1, \mathbf{y}_2$ contained in X . First, design a partition of unity functions, $\chi_0(\mathbf{x}), \chi_1(\mathbf{x}), \chi_2(\mathbf{x})$, each of which is smooth and non-negative and $\chi_0(\mathbf{x}) + \chi_1(\mathbf{x}) + \chi_2(\mathbf{x}) = 1, \forall \mathbf{x} \in X$. Here $\chi_1(\mathbf{x}), \chi_2(\mathbf{x})$ are 1 in a ball centered at $\mathbf{y}_1, \mathbf{y}_2$ respectively with radius $r_0/2$ and are zeros outside the ball with radius r_0 . $\chi_0(\mathbf{x}) = 1 - \chi_1(\mathbf{x}) - \chi_2(\mathbf{x})$. Denote

$$u(\mathbf{x}) = u(\mathbf{x})[\chi_0(\mathbf{x}) + \chi_1(\mathbf{x}) + \chi_2(\mathbf{x})] = u_0(\mathbf{x}) + u_1(\mathbf{x}) + u_2(\mathbf{x}).$$

We break the integral in (25) into three parts:

$$(34) \quad \int_X e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u(\mathbf{x}) d\mathbf{x} = \int_X e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} (u_0(\mathbf{x}) + u_1(\mathbf{x}) + u_2(\mathbf{x})) d\mathbf{x} = \text{I} + \text{II} + \text{III}.$$

The first term can be reduced to Case 2. Now let's look at the second term in (34). We change the integration to a spherical coordinate (r, θ, ψ) centered at \mathbf{y}_1 with $\theta \in [0, 2\pi]$ being the azimuthal angle, $\psi \in [0, \pi]$ being the polar angle and $\mathbf{y}_1 - \mathbf{y}_2$ being the polar axis. Then

$$(35) \quad \int_X e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u_1(\mathbf{x}) d\mathbf{x} = \int_0^{r_0} \int_{\partial B(\mathbf{y}_1, r)} e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u_1(\mathbf{x}) ds dr$$

It can be seen from (19) that $\nabla\tilde{\phi}(\mathbf{x}; \mathbf{y}_1, \mathbf{y}_2)$ is never aligned with the normal at \mathbf{x} of the sphere centered at \mathbf{y}_2 except at the intersections of $l_{\mathbf{y}_1}^{\mathbf{y}_2}$ with the sphere. So on any sphere $\partial B(\mathbf{y}_1, r)$ there are exactly two stationary points at $\psi = 0$ and $\psi = \pi$ which are non-degenerate. At the two stationary points we have

$$\begin{aligned}
(36) \quad & \tilde{\phi}(r, \theta, 0) = 1 \\
& D_{\perp}^2 \tilde{\phi}(r, \theta, 0) = \frac{-1}{r(r+|\mathbf{y}_1-\mathbf{y}_2|)} I_{\perp} \\
& u_1(r, \theta, 0) = \frac{\chi_1(r, \theta, 0)}{\|G_0(\cdot, \mathbf{y}_1)\|_2 \|G_0(\cdot, \mathbf{y}_2)\|_2 2r(|\mathbf{y}_1-\mathbf{y}_2|+r)} \\
& \tilde{\phi}(r, \theta, \pi) = |\mathbf{y}_1 - \mathbf{y}_2|^{-1} (2r - |\mathbf{y}_1 - \mathbf{y}_2|) \\
& D_{\perp}^2 \tilde{\phi}(r, \theta, \pi) = \frac{|\mathbf{y}_1-\mathbf{y}_2|^{-2r}}{r(|\mathbf{y}_1-\mathbf{y}_2|-r)} I_{\perp} \\
& u_1(r, \theta, \pi) = \frac{\chi_1(r, \theta, \pi)}{\|G_0(\cdot, \mathbf{y}_1)\|_2 \|G_0(\cdot, \mathbf{y}_2)\|_2 2r(|\mathbf{y}_1-\mathbf{y}_2|-r)}
\end{aligned}$$

where \perp denotes the tangent plane of the sphere. Note that modulo a scaling factor r^{-1} , u_1 and its derivatives, and $D_{\perp}^2 \tilde{\phi}$ as functions on $\partial B(\mathbf{y}_1, r)$ are all smooth and uniformly bounded, i.e., $|D_{\perp}^{\beta} u_1| = O(r^{-1})$ and $\|D_{\perp}^2 \tilde{\phi}\| = O(r^{-1})$. After scaling $\partial B(\mathbf{y}_1, r)$ to the unit sphere and apply the stationary phase result (21) to the two stationary phase points, one gets

$$\begin{aligned}
(37) \quad & \left| \int_{\partial B(\mathbf{y}_1, r)} e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u_1(\mathbf{x}) ds - \frac{2\pi i r^2 \tilde{k}^{-\frac{d-1}{2}}}{\|G_0(\cdot, \mathbf{y}_1)\|_2 \|G_0(\cdot, \mathbf{y}_2)\|_2} \left[-e^{i\tilde{k}} \chi_1(r, \theta, 0) + \frac{e^{i\tilde{k}|\mathbf{y}_1-\mathbf{y}_2|^{-1}(2r-|\mathbf{y}_1-\mathbf{y}_2|)} \chi_1(r, \theta, \pi)}{|\mathbf{y}_1-\mathbf{y}_2|^{-2r}} \right] \right| \\
& \lesssim \tilde{k}^{-\frac{d+1}{2}} r^2.
\end{aligned}$$

The righthand side in the above expression comes from an estimate of the righthand side term of the stationary phase formula (21) and the constant in \lesssim is uniformly bounded when $r \rightarrow 0$. The first term in the square bracket is the leading term from the stationary phase at $\psi = 0$ on the sphere $\partial B(\mathbf{y}_1, r)$ and the phase is constant in r . The second term in the bracket is the leading term from the stationary phase at $\psi = \pi$ on the sphere $\partial B(\mathbf{y}_1, r)$. However, it has a phase dependent on r which results in a higher order term after integration in r . Since all terms are integrable as $r \rightarrow 0$, we have

$$(38) \quad \left| \int_X e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u_1(\mathbf{x}) d\mathbf{x} \right| = \left| \int_0^{r_0} \int_{\partial B(\mathbf{y}_1, r)} e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u_1(\mathbf{x}) ds dr \right| \lesssim \tilde{k}^{-\frac{d-1}{2}}.$$

The third term in (34) can be shown in the same way.

From the above analysis we see that the main contribution for the integral of $\langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle_X$ may come from the stationary line $l_{\mathbf{y}_1}^{\mathbf{y}_2}$ and/or the boundary integral on $\partial\Omega$ after integration by part. All other cases can be reduced to the above three cases.

In 2D, the Green's function in free space is of the form (10) with the asymptotic formulas (11). For Case 1 and 2, the asymptotic formula (11) for $r \rightarrow \infty$ can be used as $k \rightarrow \infty$. Since the phase function in the exponential for $\langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle$ is also of the form $k(|\mathbf{x} - \mathbf{y}_1| - |\mathbf{x} - \mathbf{y}_2|)$, same arguments used above can be applied. So we have the following analogous results in 2D:

Case 1. Since the boundary ∂X is a one dimensional curve, there is some α , $1 \leq \alpha \leq \frac{d+1}{2} = \frac{3}{2}$, such that

$$(39) \quad \left| \langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle \right| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\alpha}, \quad \text{as } k|\mathbf{y}_1 - \mathbf{y}_2| \rightarrow \infty$$

Case 2. The leading contribution is due to the line of stationary phase $\tilde{l}_{\mathbf{y}_1}^{\mathbf{y}_2}$ except that the dimension orthogonal to the line is 1D, we have $\frac{d-1}{2} = \frac{1}{2}$ and

$$(40) \quad \left| \langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle \right| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\frac{1}{2}}, \quad \text{as } k|\mathbf{y}_1 - \mathbf{y}_2| \rightarrow \infty$$

Case 3. The singularity at the source is also integrable hence

$$(41) \quad \left| \langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle \right| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\frac{1}{2}}, \quad \text{as } k|\mathbf{y}_1 - \mathbf{y}_2| \rightarrow \infty$$

□

Here we give a few remarks related to the theorem above.

Remark 2.1. For $d = 3$, the same estimate also holds for two unnormalized Green's functions, i.e., $|\langle G_0(\cdot, \mathbf{y}_1), G_0(\cdot, \mathbf{y}_2) \rangle| \sim \left| \langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle \right|$, since $0 < c < \|G_0(\cdot, \mathbf{y})\|_{L_2(X)} < C < \infty$ as $k \rightarrow \infty$ for two constants c and C that only depend on X . However, this is not true for $d = 2$. If $\mathbf{y}_1, \mathbf{y}_2 \notin X$, $|\langle G_0(\cdot, \mathbf{y}_1), G_0(\cdot, \mathbf{y}_2) \rangle| \sim k^{-\frac{1}{2}} \left| \langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle \right|$ as $k \rightarrow \infty$ due to the asymptotic formula (11).

Remark 2.2. The estimate in Theorem 2.1 characterizes the correlation or angle between two normalized Green's function in term of the ratio of the separation distance between the two sources and the wavelength. One can also incorporate another scaling factor when the distance from the two points $\mathbf{y}_i, i = 1, 2$ to X is large compared to $|\mathbf{y}_1 - \mathbf{y}_2|$. Geometrically, this means that the difference between two distance functions, $|\mathbf{x} - \mathbf{y}_1| - |\mathbf{x} - \mathbf{y}_2|$, changes slowly with respect to $\mathbf{x} \in X$. Hence fast oscillation due to rapid change of the phase function, $ik(|\mathbf{x} - \mathbf{y}_1| - |\mathbf{x} - \mathbf{y}_2|)$ is discounted and the decorrelation rate of two Green's function is reduced.

Assume the size of X is $O(1)$ (otherwise one can first scale \mathbf{x} by the size of X for the Helmholtz equation (5)) and $\frac{|\mathbf{y}_1 - \mathbf{y}_2|}{\text{dist}(\mathbf{y}_1, X)} \sim \frac{|\mathbf{y}_1 - \mathbf{y}_2|}{\text{dist}(\mathbf{y}_2, X)} \sim \rho \ll 1$, which falls into either Case 1 or Case 2 in Theorem 2.1. From (19), we see that $\nabla \tilde{\phi}$ is scaled by ρ and $D^2 \tilde{\phi}, \det[D^2 \tilde{\phi}]$ are scaled by ρ, ρ^d respectively when they are not degenerate. The scaling for $u(\mathbf{x})$ is canceled due to the normalization according to the definition (23). When applying the stationary

phase result (21) at a point of stationary phase, one can see that \tilde{k} is rescaled to $\tilde{k}\rho$ if $\tilde{k}\rho \rightarrow \infty$. For Case 1, the main contribution for $\left| \langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle \right|$ comes from the last term in (26). There is a scaling factor of ρ^{-1} from $|\nabla \tilde{\phi}|^{-1}$ due to the integration by part and there are isolated stationary points on ∂X in general. So overall \tilde{k} is rescaled to $\tilde{k}\rho$ for Case 1. For Case 2 in Theorem 2.1, the main contribution for $\left| \langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle \right|$ comes from the line of stationary phase in general, where $D\tilde{\phi}$ and $D^2\tilde{\phi}$ are degenerate. Along the line of stationary phase, $D_1^2\tilde{\phi}$ and $\det[D_1^2\tilde{\phi}]$ is scaled by ρ^2 and $\rho^{2(d-1)}$ respectively from (20). Applying the stationary phase result (21) in the plane perpendicular to the line of stationary phase, \tilde{k} is rescaled to $\tilde{k}\rho^2$. From both cases we see that \tilde{k} is at least rescaled to $\tilde{k}\rho$ in the decorrelation estimate for two Green's function.

Remark 2.3. One can generalize the arguments in Theorem 2.1 to more general situations where X is a compact manifold embedded in \mathbb{R}^d with $\dim(X) = s < d$, such as a surface ($s = 2$) in \mathbb{R}^3 or a curve ($s = 1$) in \mathbb{R}^d , $d = 2, 3$. For example, if X is a compact manifold without boundary, e.g., a closed surface or curve, and two points $\mathbf{y}_1, \mathbf{y}_2 \notin X$, there is some α , $0 \leq \alpha \leq \frac{s}{2}$

$$(42) \quad \left| \langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle \right| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\alpha}, \quad \text{as } k|\mathbf{y}_1 - \mathbf{y}_2| \rightarrow \infty.$$

The two extreme cases are: (1) $\alpha = 0$ happens when there is a piece of X stays on a level set of the phase function $\tilde{\phi}(\mathbf{x}) = |\mathbf{y}_1 - \mathbf{y}_2|^{-1}(|\mathbf{x} - \mathbf{y}_1| - |\mathbf{x} - \mathbf{y}_2|)$; (2) $\alpha = \frac{s}{2}$ happens if the phase function $\tilde{\phi}(\mathbf{x})$, which has stationary phase points on a compact manifold, has isolated stationary phase points on X and $D^2\tilde{\phi}(\mathbf{x})$ is not degenerate along X at those points. The later case, $\alpha = \frac{s}{2}$, is more generic in practice.

If X is a compact manifold with boundary, there is some α , $0 \leq \alpha \leq \frac{s+1}{2}$ such that

$$(43) \quad \left| \langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle \right| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\alpha}, \quad \text{as } k|\mathbf{y}_1 - \mathbf{y}_2| \rightarrow \infty.$$

The two extreme cases are: (1) $\alpha = 0$ happens when there is a piece of X stays on a level set of the phase function $\tilde{\phi}(\mathbf{x})$; (2) $\alpha = \frac{s+1}{2}$ happens if the phase function $\tilde{\phi}(\mathbf{x})$ has no stationary phase in X and has isolated stationary phase points on ∂X and $D^2\tilde{\phi}(\mathbf{x})$ is not degenerate along ∂X at those points. If there are isolated stationary phase points in the interior of X , $\alpha = \frac{s}{2}$.

Remark 2.4. According to the Hessian estimate (20), there are two axisymmetric k dependent domains around the stationary line $\tilde{l}_{\mathbf{y}_1}^{\mathbf{y}_2}$ on each side of \mathbf{y}_1 and \mathbf{y}_2 , denoted by R_1 and R_2 respectively (see Figure 4), in which the phase function $\tilde{k}\tilde{\phi}$ does not change rapidly. For example, let's look at a point $\mathbf{x} \in \tilde{l}_{\mathbf{y}_1}^{\mathbf{y}_2}$ on the side of \mathbf{y}_2 and denote $r = \pm|\mathbf{x} - \mathbf{y}_2|$. Again we use the coordinate system $\mathbf{x} = (r, \xi, \eta)$ as shown in Figure 4. Since $\tilde{\phi}(r, 0, 0) = 1$, $|\nabla \tilde{\phi}(r, 0, 0)| = 0$, $r > 0$, for a point (r, ξ, η) with $r > 0$, $\sqrt{\xi^2 + \eta^2} \lesssim \frac{r(r+|\mathbf{y}_1-\mathbf{y}_2|)}{k|\mathbf{y}_1-\mathbf{y}_2|}$, from

(20) we have

$$(44) \quad \tilde{k}|\nabla\tilde{\phi}(r, \xi, \eta)| \lesssim \frac{k|\mathbf{y}_1 - \mathbf{y}_2|\sqrt{\xi^2 + \eta^2}}{r(r + |\mathbf{y}_1 - \mathbf{y}_2|)} \lesssim 1.$$

Hence $\langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle$ is not an oscillatory integral in R_1 or R_2 and the two Green's functions do not decorrelate fast in a subdomain contained in these two domains. We will also provide special k dependent setups of domains X, Y such that $G(\mathbf{x}, \mathbf{y}_1)$ and $G(\mathbf{x}, \mathbf{y}_2)$ do not decorrelate fast in X for two sources $\mathbf{y}_1, \mathbf{y}_2 \in Y$ in Section 4.2.

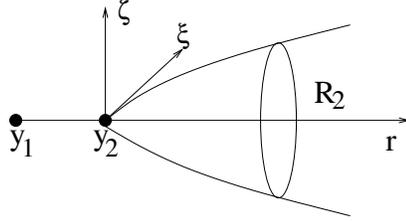


FIGURE 4. A domain where two Green's functions do not decorrelate fast.

Remark 2.5. The correlation between two Green's function can also be used for study imaging resolution using waves. Suppose X is a compact planar region in R^3 where the wave field is measured. $\mathbf{y}_1, \mathbf{y}_2 \in R^3$ are two point sources or scatterers. If the line connecting \mathbf{y}_1 and \mathbf{y}_2 is parallel to X , we have $|\langle G_0(\cdot, \mathbf{y}_1), G_0(\cdot, \mathbf{y}_2) \rangle| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\frac{3}{2}}$ in general as $k|\mathbf{y}_1 - \mathbf{y}_2| \rightarrow \infty$ since there is no stationary phase. While if the line connecting \mathbf{y}_1 and \mathbf{y}_2 intersects X perpendicularly, $|\langle G_0(\cdot, \mathbf{y}_1), G_0(\cdot, \mathbf{y}_2) \rangle| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\frac{1}{2}}$ as $k|\mathbf{y}_1 - \mathbf{y}_2| \rightarrow \infty$ since the intersection point is a stationary point. Hence it implies imaging resolution in the range direction is poorer than that in the plane parallel to X .

2.2. Decorrelation of Two Green's Function in Inhomogeneous Medium in High Frequency Limit. The situation is similar for inhomogeneous medium in the high frequency limit when the Green's function has a valid geometric optics approximation. The assumption for geometric optics ansatz is that the solution to the Helmholtz equation has the following expansion:

$$(45) \quad u(\mathbf{x}) = e^{ik\phi(\mathbf{x})} \sum_{m=0}^{\infty} a_m(\mathbf{x})(ik)^{-m},$$

where $\phi(\mathbf{x})$ is the phase function and $a_m(\mathbf{x})$ are the amplitude functions which satisfy

$$(46) \quad \begin{aligned} |\nabla\phi(\mathbf{x})| &= n(\mathbf{x}), & 2\nabla\phi(\mathbf{x}) \cdot \nabla a_0(\mathbf{x}) + \Delta\phi(\mathbf{x})a_0(\mathbf{x}) &= 0 \\ 2\nabla\phi(\mathbf{x}) \cdot \nabla a_m(\mathbf{x}) + \Delta\phi(\mathbf{x})a_m(\mathbf{x}) + \Delta a_{m-1}(\mathbf{x}) &= 0, & m &= 1, 2, \dots \end{aligned}$$

For the Green's function $G(\mathbf{x}, \mathbf{y})$ defined in (5), we have the following condition at the source point \mathbf{y} :

$$(47) \quad \lim_{\mathbf{x} \rightarrow \mathbf{y}} \left(\frac{\phi(\mathbf{x}, \mathbf{y})}{|\mathbf{x} - \mathbf{y}|} - n(\mathbf{y}) \right) = 0, \quad \lim_{\mathbf{x} \rightarrow \mathbf{y}} a_0(\mathbf{x}, \mathbf{y}) 4\pi |\mathbf{x} - \mathbf{y}| = 1.$$

The above geometric optics ansatz can be formulated as a Hamiltonian system, also called as Lagrangian formulation or ray tracing, which gives explicit ordinary differential equations (ODE) for bicharacteristics $(\mathbf{x}(t), \mathbf{p}(t))$ in phase space with Hamiltonian $H(\mathbf{x}, \mathbf{p}) = |\mathbf{p}| - n(\mathbf{x})$ and $\mathbf{p} = \nabla \phi$,

$$(48) \quad \begin{aligned} \frac{d\mathbf{x}(t, \mathbf{x}_0, \mathbf{p}_0)}{dt} &= \nabla_{\mathbf{p}} H(\mathbf{x}, \mathbf{p}) = \frac{\mathbf{p}}{n(\mathbf{x})}, & \mathbf{x}(0) &= \mathbf{x}_0, \\ \frac{d\mathbf{p}(t, \mathbf{x}_0, \mathbf{p}_0)}{dt} &= -\nabla_{\mathbf{x}} H(\mathbf{x}, \mathbf{p}) = \nabla n(\mathbf{x}), & \mathbf{p}(0) &= \mathbf{p}_0 = \nabla \phi(\mathbf{x}_0) \\ \frac{d\phi(\mathbf{x}(t, \mathbf{x}_0, \mathbf{p}_0))}{dt} &= \nabla \phi \cdot \frac{d\mathbf{x}}{dt} = n(\mathbf{x}), & \phi(0) &= \phi(\mathbf{x}_0) \end{aligned}$$

The projection of the bicharacteristics in the physical space, i.e., $\mathbf{x}(t, \mathbf{y}_0, \mathbf{p}_0)$, are called rays. If there is no caustics, i.e., two rays do not intersect in the physical space, each ray is a geodesic in the physical space with the slowness $n(x) = \frac{1}{c(\mathbf{x})}$ as the metric, where $c(\mathbf{x})$ is the wave speed. $|\phi(\mathbf{x}(t_2, \mathbf{x}_0, \mathbf{p}_0)) - \phi(\mathbf{x}(t_1, \mathbf{x}_0, \mathbf{p}_0))|$ is the shortest travel time between points $\mathbf{x}(t_1, \mathbf{x}_0, \mathbf{p}_0)$ and $\mathbf{x}(t_2, \mathbf{x}_0, \mathbf{p}_0)$. Moreover, the amplitude along each ray is given by

$$(49) \quad a_0^2(\mathbf{x}(t, \mathbf{x}_0)) = a_0^2(\mathbf{x}_0) \frac{n(\mathbf{x}_0)}{n(\mathbf{x}(t, \mathbf{x}_0))} \left| \frac{\partial \mathbf{x}(t, \mathbf{x}_0)}{\partial \mathbf{x}_0} \right|^{-1},$$

where $\left| \frac{\partial \mathbf{x}(t, \mathbf{x}_0)}{\partial \mathbf{x}_0} \right|$ is the determinant of the Jacobian $\frac{\partial \mathbf{x}(t, \mathbf{x}_0)}{\partial \mathbf{x}_0}$ meaning the geometric spreading of rays. Before caustics are formed, the determinant is always positive and bounded. Once $a_0(\mathbf{x})$ is known, $a_1(\mathbf{x}), a_2(\mathbf{x}), \dots$ can be solved consecutively from (46).

In particular, for the geometric optics ansatz for the Green's function with a point source at \mathbf{y}_0 , rays $\mathbf{x}(t, \mathbf{y}_0, \hat{\theta})$ are emanating from the source \mathbf{y}_0 in all directions and can be parametrized by the initial directions, i.e., the take-off angles $\hat{\theta} \in S^{d-1}$ on a unit sphere. The ODEs for the rays (48) have initial conditions $\mathbf{x}(0, \mathbf{y}_0, \hat{\theta}) = \mathbf{y}_0, \mathbf{p}(0, \mathbf{y}_0, \hat{\theta}) = n(\mathbf{y}_0)\hat{\theta}, \phi(0, \mathbf{y}_0, \hat{\theta}) = 0$, with $a_0(0, \hat{\theta})$ evenly distributed in all $\hat{\theta}$. If there is no caustics, every point \mathbf{x} has a unique ray passing through it, i.e., $\forall \mathbf{x}, \exists! t(\mathbf{x}), \hat{\theta}(\mathbf{x})$ such that $\mathbf{x}(t(\mathbf{x}), \mathbf{y}_0, \hat{\theta}(\mathbf{x})) = \mathbf{x}$. Moreover, the ray connecting \mathbf{y}_0 and \mathbf{x} is the geodesic, or the shortest travel time, between the two points with $n(x)$ being the slowness of the medium. Under no caustics assumptions, the Greens functions with sources at $\mathbf{y}_1, \mathbf{y}_2$ can be approximated by the following geometric optics ansatz

$$(50) \quad \left| G(\mathbf{x}, \mathbf{y}_j) - e^{ik\phi(\mathbf{x}, \mathbf{y}_j)} A_j(\mathbf{x}) \right| \lesssim k^{-(M+1)}, \quad j = 1, 2,$$

where $A_j(\mathbf{x}) = \sum_{m=0}^M a_m(\mathbf{x}, \mathbf{y}_j)(ik)^{-m}$, $j = 1, 2$. Hence

$$(51) \quad \left| \langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle - \int_X e^{i\tilde{k}\tilde{\phi}(\mathbf{x})} u(\mathbf{x}) d\mathbf{x} \right| \lesssim k^{-(M+1)},$$

where

$$\tilde{k} = k\phi(\mathbf{y}_1, \mathbf{y}_2), \quad \tilde{\phi}(\mathbf{x}) = \phi^{-1}(\mathbf{y}_1, \mathbf{y}_2)(\phi(\mathbf{x}, \mathbf{y}_1) - \phi(\mathbf{x}, \mathbf{y}_2)), \quad u(\mathbf{x}) = \frac{A_1(\mathbf{x})A_2(\mathbf{x})}{\|G(\cdot, \mathbf{y}_1)\|_2 \|G(\cdot, \mathbf{y}_2)\|_2}$$

Denote $\Gamma_{\mathbf{y}_1}^{\mathbf{y}_2}$ to be the unique ray that passes through \mathbf{y}_1 and \mathbf{y}_2 as illustrated in Figure 5 (a). If $n(\mathbf{x})$ is smooth and $0 < c \leq n(\mathbf{x}) \leq C < \infty$, one has (see Figure 5(b))

$$c|\mathbf{y}_1 - \mathbf{y}_2| \leq \phi(\mathbf{y}_1, \mathbf{y}_2) \leq C|\mathbf{y}_1 - \mathbf{y}_2|, \quad |\phi(\mathbf{x}, \mathbf{y}_1) - \phi(\mathbf{x}, \mathbf{y}_2)| \leq \phi(\mathbf{y}_2, \mathbf{y}_1)$$

So the phase function $\tilde{\phi}(\mathbf{x})$ attains the global maximum or minimum ± 1 on the part of the ray $\Gamma_{\mathbf{y}_1}^{\mathbf{y}_2}$ which is outside the interval between \mathbf{y}_1 and \mathbf{y}_2 , denoted by $\tilde{\Gamma}_{\mathbf{y}_1}^{\mathbf{y}_2}$. Moreover, $\nabla_{\mathbf{x}}\phi(\mathbf{x}, \mathbf{y}_1) - \nabla_{\mathbf{x}}\phi(\mathbf{x}, \mathbf{y}_2) \neq 0$ for any \mathbf{x} that is not on $\tilde{\Gamma}_{\mathbf{y}_1}^{\mathbf{y}_2}$ because the two different and unique geodesics connecting \mathbf{x}, \mathbf{y}_1 and \mathbf{x}, \mathbf{y}_2 respectively can not be tangent at \mathbf{x} . So $\tilde{\Gamma}_{\mathbf{y}_1}^{\mathbf{y}_2}$ is a stationary curve in the inhomogeneous case and plays the same role as the straight line $\tilde{l}_{\mathbf{y}_1}^{\mathbf{y}_2}$ in the homogeneous case. Depending on whether the ray passes through X or not, we get the same results as in Theorem 2.1 because $k\phi(\mathbf{y}_1, \mathbf{y}_2) \sim k|\mathbf{y}_1 - \mathbf{y}_2|$. This is true when \mathbf{y}_1 and/or \mathbf{y}_2 are in X as well since the amplitude a_0 satisfying (47) has exactly the same singularity as the homogeneous Green's function.

The main complication for geometric optics ansatz in heterogeneous medium is when rays cross each other, i.e., when caustics are formed. Although bicharacteristics in phase space are still well defined, the amplitude formula (49) breaks down. So in general, the above arguments can not be carried over to a general inhomogeneous medium. However, in the case that there are finite number of rays starting from \mathbf{y}_1 going through \mathbf{y}_2 and there is a partition of unity for the takeoff angle $\hat{\theta}$ on S^{d-1} such that there is a small cone around each ray where there is no caustics, then one can apply the above argument to each cone and get the same results.

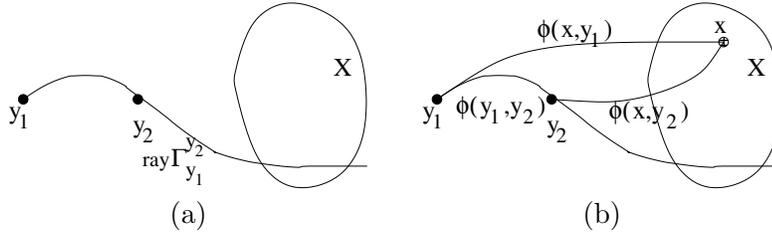


FIGURE 5. Rays in inhomogeneous medium

3. APPROXIMATE SEPARABILITY ESTIMATE FOR THE GREEN'S FUNCTION OF HELMHOLTZ EQUATION IN HIGH FREQUENCY LIMIT

In this section we present general estimates for the approximate separability of Green's function of Helmholtz equation in the high frequency limit. In the next section, we will apply these results to get explicit estimates for special setups that are of interest in practice. These estimates imply rank estimates for discretized operators which are important for developing fast algorithms for solving high frequency Helmholtz equation and its boundary integral equation counterpart.

3.1. Approximating a Set of Almost Orthogonal Vectors by a Linear Subspace.

First we present some background and introduce definitions for the approximation of a set of vectors using a linear subspace. It will be extended later to the approximation of Green's function in the infinite dimensional function space. Let $\mathbf{v}_m \in \mathbb{R}^d$, $m = 1, 2, \dots, N$ be a set of vectors. Define matrix $V = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N]$ and matrix $A = [a_{mn}]_{N \times N} = V^T V$. Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N \geq 0$ be the eigenvalues of A , then $\text{tr}(A) = \sum_{m=1}^N \lambda_m = \sum_{m=1}^N \|\mathbf{v}_m\|_2^2$. $\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \dots \geq \sqrt{\lambda_N} \geq 0$ are also called singular values for V . The best linear subspace \hat{S}_l of all linear subspace S_l of dimension l that approximates the set of vectors $\{\mathbf{v}_m\}_{m=1}^N$ in least square sense is the space spanned by the first l left singular vectors of V and satisfies

$$(52) \quad \sum_m^N \|\mathbf{v}_m - P_{\hat{S}_l} \mathbf{v}_m\|_2^2 = \min_{S_l, \dim(S)=l} \sum_m^N \|\mathbf{v}_m - P_{S_l} \mathbf{v}_m\|_2^2 = \sum_{m=l+1}^N \lambda_m,$$

where $P_{S_l} \mathbf{v}$ denotes projection of \mathbf{v} in S_l . In another word,

$$(53) \quad \lambda_l = \max_{\hat{\mathbf{e}} \in \mathbb{R}^d, \|\hat{\mathbf{e}}\|_2=1, \hat{\mathbf{e}} \perp \hat{S}_{l-1}} \sum_j^N |\mathbf{v}_j \cdot \hat{\mathbf{e}}|^2,$$

is the maximum reduction of approximation error in term of least square for the set of vectors $\{\mathbf{v}_m\}_{m=1}^N$ when adding one more dimension to the previous optimal $l-1$ dimensional linear subspace. Here we introduce two definitions for approximate rank estimate for a symmetric non-negative matrix A .

Definition 3.1. Given $\epsilon > 0$, $\bar{N}^\epsilon = \max_{1 \leq m \leq N} m$, s.t. $\sqrt{\lambda_m} \geq \epsilon$, i.e., \bar{N}^ϵ denotes the largest m such that $\sqrt{\lambda_m} \geq \epsilon$.

Definition 3.2. Given $1 \geq \epsilon > 0$, $\underline{N}^\epsilon = \min M$, s.t. $\sum_{m=M+1}^N \lambda_m \leq \epsilon^2 \sum_{m=1}^N \lambda_m$.

If $A = V^T V$, definition 3.2 implies that if a linear subspace S^ϵ satisfies

$$(54) \quad \frac{\sum_{m=1}^N \|\mathbf{v}_m - P_{S^\epsilon} \mathbf{v}_m\|_2^2}{\sum_{m=1}^N \|\mathbf{v}_m\|_2^2} \leq \epsilon^2$$

then $\dim(S^\epsilon) \geq \underline{N}^\epsilon$. Assume $0 < c < \|\mathbf{v}_m\|_2 < C < \infty, m = 1, 2, \dots, N$, we can conclude that if a linear subspace S^ϵ satisfies

$$(55) \quad \sqrt{\frac{\sum_{m=1}^N \|\mathbf{v}_m - P_{S^\epsilon} \mathbf{v}_m\|_2^2}{N}} \leq c\epsilon \quad \Rightarrow \quad \frac{\sum_{m=1}^N \|\mathbf{v}_m - P_{S^\epsilon} \mathbf{v}_m\|_2^2}{\sum_{m=1}^N \|\mathbf{v}_m\|_2^2} \leq \epsilon^2$$

then $\dim(S^\epsilon) \geq \underline{N}^\epsilon$. In another word, the least dimension of a linear subspace that can have an $c\epsilon$ -r.m.s. (root mean square) approximation of a set of vectors $\mathbf{v}_m, m = 1, 2, \dots, N$ is at least \underline{N}^ϵ . The root mean square approximation will lead to L_2 approximate separability estimate for Green's function in the continuous case.

In the previous section, we proved the rate of decorrelation of two Green's function: $|\langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\alpha}$ for some $\alpha > 0$ as $k|\mathbf{y}_1 - \mathbf{y}_2| \rightarrow \infty$. Geometrically it means that two Green's functions with sources separated a little more than one wavelength become almost orthogonal as $k \rightarrow \infty$. Intuitively, for two domains $X, Y \subset R^d$, if one views $G(\mathbf{x}, \mathbf{y})$ as a family of functions in $L_2(X)$ parametrized by $\mathbf{y} \in Y$ and lays down a uniform grid $\mathbf{y}_j \in Y$ with grid size $h = k^{-\beta}$ for any $0 < \beta < 1$, $G(\mathbf{x}, \mathbf{y}_j)$ is a set of almost orthogonal vectors in $L_2(X)$. A natural question is the least number of dimensions of a linear space that can contain a set of almost orthogonal vectors. This question has been studied in [1] by rank estimate for small off-diagonal perturbation of identity matrices, which is equivalent to the same question for a set of almost orthogonal unit vectors. In particular, the asymptotic estimate is optimal and is used to show the sharpness of Johnson-Lindenstraus Lemma [19]. However, this result can not address our problem adequately for the following two reasons. First, the assumption in [1] on almost orthogonality for a set vectors is only pairwise. In our problems, the set of vectors are Green's functions for a PDE which has spatial structure, i.e., the angle between two Green's functions depends on separation distance of the two sources. The spatial structure has to be taken into account to get sharp estimates. Second, approximate separability means that we need to estimate the least dimension of a linear subspace that can approximate a set of vectors to a certain tolerance instead of containing the whole set of vectors. Here we adopt the approach from [1] and develop more careful estimates in Lemma 3.1 for a set of Green's functions by taking into account both the spatial structure and approximation tolerance. The lemma is then used to prove lower bound estimates for approximate separability of the Green's function for Helmholtz equation in the high frequency limit.

For more generality, we assume X, Y are two compact manifolds embedded in R^d , i.e., they may be compact domains in R^d , or compact surfaces embedded in R^3 or compact curves embedded in $R^d, d = 2$ or 3 . Without loss of generality, we assume $d \geq \dim(X) \geq \dim(Y) = s$. For a smooth manifold Y with $\dim(Y) = s = 1, 2, 3$, it contains a local patch of size $O(1)$ that is diffeomorphism to a one dimensional line of unit length, a two dimensional unit square and a three dimensional unit cube respectively, which we will only consider in the following analysis. Let $G(\mathbf{x}, \mathbf{y})$ be the Green's function of the Helmholtz equation (5). The following Lemma 3.1 shows the dimension estimate for a linear subspace in $L_2(X)$ which approximate of a set of Green's function $G(\mathbf{x}, \mathbf{y}_m), \mathbf{y}_m \in Y$ with ϵ -r.m.s error.

Lemma 3.1. *Let X, Y be two compact manifolds embedded in $R^d, d = 2, 3$ and $d \geq \dim(X) \geq \dim(Y) = s$. If for any two points $\mathbf{y}_1, \mathbf{y}_2 \in Y$,*

$$|\langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\alpha} \quad \text{as } k|\mathbf{y}_1 - \mathbf{y}_2| \rightarrow \infty$$

for some $\alpha > 0$, then there are points $\mathbf{y}_m \in Y, m = 1, 2, \dots, N_\delta^s \sim k^{s-\delta}$, for any $0 < \delta < 1$ and arbitrary close to 0, such that for the set of Green's functions $\{G(\mathbf{x}, \mathbf{y}_m)\}_{m=1}^{N_\delta^s} \subset L_2(X)$ and matrix $A = \langle \hat{G}(\cdot, \mathbf{y}_m), \hat{G}(\cdot, \mathbf{y}_n) \rangle$

$$(56) \quad \underline{N}_k^\epsilon \gtrsim \begin{cases} (1 - \epsilon^2)^2 k^{2\alpha}, & \alpha < \frac{s}{2}, \\ (1 - \epsilon^2)^2 k^{s-\delta}, & \alpha \geq \frac{s}{2}, \end{cases}$$

and

$$(57) \quad \overline{N}_k^\epsilon \lesssim \begin{cases} \epsilon^{-4} k^{2(s-\alpha-\delta)}, & \alpha < \frac{s}{2}, \\ \epsilon^{-4} k^{s-\delta}, & \alpha \geq \frac{s}{2}, \end{cases}$$

as $k \rightarrow \infty$, where the constants in \lesssim and \gtrsim only depend on X, Y and $n(\mathbf{x})$.

Proof. We prove the statement for $X, Y \subset R^3$ and $\dim(Y) = s = 1, 2, 3$ respectively. The case for $X, Y \subset R^2$ can be proved in exactly the same way.

Case 1: $s = 1$. Y is a line of unit length in R^3 . Put down a uniform grid \mathbf{y}_m in Y with grid size $h = k^{\delta-1}$, $0 < \delta < 1$, such that $|\mathbf{y}_m - \mathbf{y}_n| \sim |m - n|h$, $m, n = 1, 2, \dots, n_k^h = k^{1-\delta}$. See Figure 6(a). Define the matrix

$$(58) \quad A = (a_{mn})_{n_k^h \times n_k^h}, \quad a_{mn} = \langle \hat{G}(\cdot, \mathbf{y}_m), \hat{G}(\cdot, \mathbf{y}_n) \rangle,$$

where

$$(59) \quad a_{mm} = 1, \quad |a_{mn}| \lesssim |m - n|^{-\alpha} k^{-\alpha\delta}, \quad m, n = 1, 2, \dots, n_k^h.$$

Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n_k^h} \geq 0$ be the eigenvalues of A . Then $\sum_{m=1}^{n_k^h} \lambda_m = n_k^h$. Since $\overline{N}_k^\epsilon = \max_{1 \leq m \leq n_k^h} m$, s.t. $\sqrt{\lambda_m} \geq \epsilon$ and $\underline{N}_k^\epsilon = \min M$, s.t. $\sum_{m=M+1}^{n_k^h} \lambda_m \leq \epsilon^2 \sum_{m=1}^{n_k^h} \lambda_m = \epsilon^2 n_k^h$, we have

$$\sum_{m=1}^{\underline{N}_k^\epsilon} \lambda_m \geq (1 - \epsilon^2) \sum_{m=1}^{n_k^h} \lambda_m = (1 - \epsilon^2) n_k^h$$

and

$$(60) \quad \sum_{m=1}^{n_k^h} \lambda_m^2 > \sum_{m=1}^{\underline{N}_k^\epsilon} \lambda_m^2 \geq \underline{N}_k^\epsilon \left[\frac{(1 - \epsilon^2) n_k^h}{\underline{N}_k^\epsilon} \right]^2 = \frac{[(1 - \epsilon^2) n_k^h]^2}{\underline{N}_k^\epsilon},$$

and

$$(61) \quad \sum_{m=1}^{n_k^h} \lambda_m^2 > \sum_{m=1}^{\overline{N}_k^\epsilon} \lambda_m^2 \geq \overline{N}_k^\epsilon \epsilon^4.$$

At the same time, for a fixed $\alpha > 0$ and take $0 < \delta < 1$ arbitrary close to 0,

$$\begin{aligned}
 \sum_{m=1}^{n_k^h} \lambda_m^2 &= \text{tr}(A^T A) = \sum_{m=1}^{n_k^h} \sum_{n=1}^{n_k^h} a_{mn}^2 \\
 &= \sum_{m=1}^{n_k^h} a_{mm}^2 + 2 \sum_{n=1}^{n_k^h-1} \sum_{m=n+1}^{n_k^h} a_{m,m-n}^2 \\
 (62) \quad &\lesssim n_k^h + 2 \sum_{n=1}^{n_k^h-1} (n_k^h - n) n^{-2\alpha} k^{-2\alpha\delta} \\
 &\lesssim \begin{cases} k^{1-\delta} + k^{2(1-\delta-\alpha)} \lesssim k^{2(1-\delta-\alpha)}, & \alpha < \frac{1}{2}, 2\alpha < 1 - \delta < 1 \\ k^{1-\delta} + k^{1-2\delta} \ln k \lesssim k^{1-\delta}, & \alpha = \frac{1}{2}, 0 < \delta < 1 \\ k^{1-\delta} + k^{1-\delta-2\alpha\delta} \lesssim k^{1-\delta}, & \alpha > \frac{1}{2}, 0 < \delta < 1 \end{cases}
 \end{aligned}$$

Hence for a fixed $\alpha > 0$ and any $0 < \delta < 1$ arbitrary close to 0, combining (60) and (62) we have

$$(63) \quad \underline{N}_k^\epsilon \gtrsim \begin{cases} (1 - \epsilon^2)^2 k^{2\alpha}, & \alpha < \frac{1}{2} \\ (1 - \epsilon^2)^2 k^{1-\delta}, & \alpha \geq \frac{1}{2} \end{cases},$$

and combining (61) and (62) we have

$$(64) \quad \overline{N}_k^\epsilon \lesssim \begin{cases} \epsilon^{-4} k^{2(1-\alpha-\delta)}, & \alpha < \frac{1}{2} \\ \epsilon^{-4} k^{1-\delta}, & \alpha \geq \frac{1}{2} \end{cases}.$$

Case 2: $s = 2$. Y is a unit square in R^3 . Again put down a uniform grid $\mathbf{y}_m, m = 1, 2, \dots, n_k^h = k^{2(1-\delta)}$ in Y with grid size $h = k^{\delta-1}, 0 < \delta < 1$ (see Figure 6(b)), and define matrix A as in (58). Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n_k^h} \geq 0$ be its eigenvalues, then $\sum_{m=1}^{n_k^h} \lambda_m = n_k^h$ and we have (60), (61). At the same time

$$(65) \quad \sum_{m=1}^{n_k^h} \lambda_m^2 = \text{tr}(A^T A) = \sum_{m=1}^{n_k^h} \sum_{n=1}^{n_k^h} a_{mn}^2.$$

Let's look at the sum of each row. Assume \mathbf{y}_m is the center of the square. We divide all other points into groups of 1st square neighbors, 2nd square neighbors, \dots , j -th square neighbors, denoted by $S_j, j = 1, 2, \dots, J \sim h^{-1} = k^{1-\delta}$. See Figure 6(b). S_j contains those $4(2j+1) - 4 = 8j$ grid points that are on the 4 sides of the square centered at \mathbf{y}_m with each side of length $2jh$. Then we have $jh \leq |\mathbf{y}_{n_j} - \mathbf{y}_m| \leq \sqrt{2}jh, \mathbf{y}_{n_j} \in S_j$, and hence $a_{m,n_j} \lesssim (kjh)^{-\alpha} = j^{-\alpha} k^{-\alpha\delta}$. For a fixed $\alpha > 0$ and take $0 < \delta < 1$ arbitrary close to 0, we

have

$$(66) \quad \begin{aligned} \sum_{n=1}^{n_k^h} a_{mn}^2 &= 1 + \sum_{j=1}^J \sum_{n_j \in S_j} a_{m,n_j}^2 \lesssim 1 + \sum_{j=1}^J j^{-2\alpha+1} k^{-2\alpha\delta} \\ &\lesssim \begin{cases} 1 + k^{2(1-\delta-\alpha)} \lesssim k^{2(1-\delta-\alpha)}, & \alpha < 1, \alpha < 1 - \delta < 1 \\ 1 + k^{-2\delta} \ln k \lesssim 1, & \alpha = 1, 0 < \delta < 1 \\ 1 + k^{-2\alpha\delta} \lesssim 1, & \alpha > 1, 0 < \delta < 1 \end{cases} \end{aligned}$$

Actually for any grid point \mathbf{y}_m , each j -th square neighbors of \mathbf{y}_m has at least $8j/4 = 2j$ points for $j = 1, 2, \dots, J \sim k^{1-\delta}$, e.g., if \mathbf{y}_m is a corner point of the square domain. Hence the above asymptotic formula is still true. For a fixed $\alpha > 0$ and any $0 < \delta < 1$ arbitrary close to 0, we have

$$(67) \quad \sum_{m=1}^{n_k^\epsilon} \lambda_m^2 \leq \sum_{m=1}^{n_k^h} \sum_{n=1}^{n_k^h} a_{mn}^2 \lesssim \begin{cases} k^{2(2(1-\delta)-\alpha)}, & \alpha < 1 \\ k^{2(1-\delta)}, & \alpha \geq 1 \end{cases}$$

Combining (67) with (60), we have

$$(68) \quad \underline{N}_k^\epsilon \gtrsim \begin{cases} (1 - \epsilon^2)^2 k^{2\alpha}, & \alpha < 1 \\ (1 - \epsilon^2)^2 k^{2(1-\delta)}, & \alpha \geq 1. \end{cases}$$

Combining (67) with (61), we have

$$(69) \quad \overline{N}_k^\epsilon \lesssim \begin{cases} \epsilon^{-4} k^{2(2(1-\delta)-\alpha)}, & \alpha < 1 \\ \epsilon^{-4} k^{2(1-\delta)}, & \alpha \geq 1 \end{cases}.$$

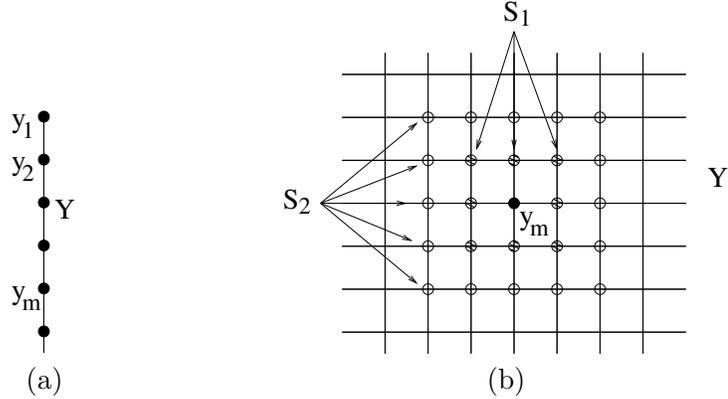


FIGURE 6. Green's functions with sources on a uniform grid.

Case 3: $s = 3$. Y is a unit cube in R^3 . Put down a uniform grid $\mathbf{y}_m, m = 1, 2, \dots, n_k^h = k^{3(1-\delta)}$ in Y with grid size $h = k^{\delta-1}, 0 < \delta < 1$ and define matrix A as in (58). Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{n_k^h} \geq 0$ be its eigenvalues, then $\sum_{m=1}^{n_k^h} \lambda_m = n_k^h$ and we have (60), (61). Similar to 2D case, assume \mathbf{y}_m is the center of the cube. We divide all other points into groups of 1st cube neighbors, 2nd cube neighbors, \dots , j -th cube neighbors, denoted by $C_j, j = 1, 2, \dots, J \sim h^{-1} = k^{1-\delta}$. C_j contains those $6(2j+1)^2 - 12(2j+1) + 8 = 24j^2 + 2$ grid points \mathbf{y}_{n_j} that are on the faces of the cube centered at \mathbf{y}_m with each face a square of length $2jh$. Then we have $jh \leq |\mathbf{y}_{n_j} - \mathbf{y}_m| \leq \sqrt{3}jh, \mathbf{y}_{n_j} \in C_j$, and $a_{m,n_j} \lesssim (kjh)^{-\alpha} = j^{-\alpha}k^{-\alpha\delta}$. For $0 < \delta < 1$ arbitrary close to 0, one has the row sum estimate

$$(70) \quad \begin{aligned} \sum_{n=1}^{n_k^h} a_{mn}^2 &= 1 + \sum_{j=1}^J \sum_{n_j \in C_j} a_{m,n_j}^2 \lesssim 1 + \sum_{j=1}^J (24j^2 + 2)j^{-2\alpha}k^{-2\alpha\delta} \\ &\lesssim \begin{cases} 1 + k^{3(1-\delta)-2\alpha} \lesssim k^{3(1-\delta)-2\alpha}, & \alpha < \frac{3}{2}, \frac{2}{3}\alpha < 1 - \delta < 1 \\ 1 + k^{-3\delta} \ln k \lesssim 1, & \alpha = \frac{3}{2}, 0 < \delta < 1 \\ 1 + k^{-2\alpha\delta} \lesssim 1, & \alpha > \frac{3}{2}, 0 < \delta < 1. \end{cases} \end{aligned}$$

Also this is true for any point \mathbf{y}_m which has at least $1/8$ of $24j^2 + 2$ points in its j -th cube neighbor C_j for $j = 1, 2, \dots, J \sim k^{1-\delta}$. Hence for a fixed $\alpha > 0$ and any $0 < \delta < 1$ arbitrary close to 0,

$$(71) \quad \sum_{m=1}^{n_k^\epsilon} \lambda_m^2 \leq \sum_{m=1}^{n_k^h} \sum_{n=1}^{n_k^h} a_{mn}^2 \lesssim \begin{cases} k^{2(3(1-\delta)-\alpha)}, & \alpha < \frac{3}{2} \\ k^{3(1-\delta)}, & \alpha \geq \frac{3}{2} \end{cases}$$

Combining (71) with (60), we have

$$(72) \quad \underline{N}_k^\epsilon \gtrsim \begin{cases} (1 - \epsilon^2)^2 k^{2\alpha}, & \alpha < \frac{3}{2} \\ (1 - \epsilon^2)^2 k^{3(1-\delta)}, & \alpha \geq \frac{3}{2} \end{cases}.$$

Combine (71) with (61), we have

$$(73) \quad \overline{N}_k^\epsilon \lesssim \begin{cases} \epsilon^{-4} k^{2(3(1-\delta)-\alpha)}, & \alpha < \frac{3}{2} \\ \epsilon^{-4} k^{3(1-\delta)}, & \alpha \geq \frac{3}{2} \end{cases}.$$

Replace $s\delta$ by δ in each of the above cases, we complete the proof. \square

Remark 3.1. When $d = 3$, if either $\dim(X) = \dim(Y) = 3$ or X and Y are separated, there are $0 < \underline{c} < \bar{c} < \infty$ independent of k such that $\underline{c} \leq \|G(\cdot, \mathbf{y})\|_2 \leq \bar{c}$ uniformly in $\mathbf{y} \in Y$. For the same set of Green's functions $G(\mathbf{x}, \mathbf{y}_j)$ as in the above proof and $A = \langle G(\cdot, \mathbf{y}_m), G(\cdot, \mathbf{y}_n) \rangle$, we have

$$(74) \quad \underline{c}n_k^h \leq \sum_{m=1}^{n_k^h} \lambda_m \leq \bar{c}n_k^h \quad \Rightarrow \quad \sum_{m=1}^{N_k^\epsilon} \lambda_m \geq (1 - \epsilon^2) \sum_{m=1}^{n_k^h} \lambda_m \geq (1 - \epsilon^2)\underline{c}n_k^h.$$

Hence inequalities (60) is replaced by the following:

$$(75) \quad \sum_{m=1}^{n_k^h} \lambda_m^2 > \sum_{m=1}^{\underline{N}_k^\epsilon} \lambda_m^2 \geq \underline{N}_k^\epsilon \left[\frac{(1-\epsilon^2)c n_k^h}{\underline{N}_k^\epsilon} \right]^2 = \frac{[(1-\epsilon^2)c n_k^h]^2}{\underline{N}_k^\epsilon},$$

and (61) is the same. The estimate of $\sum_{m=1}^{n_k^\epsilon} \lambda_m^2 \leq \sum_{m=1}^{n_k^h} \sum_{n=1}^{n_k^h} a_{mn}^2$ is amplified by \bar{c}^2 from the previous estimate at most. So the same results are true.

When $d = 2$, except for a scaling factor $k^{-\frac{1}{2}}$ for the Green's function, everything else is similar to the case $d = 3$. By the definition 3.2 of \underline{N}_k^ϵ , it is independent of a constant scaling of the whole set of vectors. So the results for \underline{N}_k^ϵ also holds for $A = \langle G(\cdot, \mathbf{y}_m), G(\cdot, \mathbf{y}_n) \rangle$.

Remark 3.2. The size of Y can be scaled for the result of Lemma 3.1. If $\alpha \geq \frac{s}{2}$, trace estimates (62), (67), and (71) gives $\sum_{m=1}^{n_k^h} \lambda_m^2 \lesssim n_k^h$. It can be seen easily from (60) and (61) that if Y is scaled to aY , k is scaled to ak in those estimates (56), (57) for \underline{N}_k^ϵ and \bar{N}_k^ϵ respectively. Otherwise, the only factor that can not be scaled with n_k^h in the trace estimates is $k^{-2\alpha\delta}$ in (62), (66) and (70). Hence, in addition to scale k to ak , there is an extra factor of $a^{2\alpha\delta}$ for the result of Lemma 3.1.

Remark 3.3. The key estimates in the proof of Lemma 3.1 are (60) (or (75)), (61), and the one for $\sum_{m=1}^{n_k^h} \lambda_m^2 = \text{tr}(A^T A)$. Although the estimate of $\text{tr}(A^T A)$ can be improved by more careful estimates of each row sum $\sum_{n=1}^{n_k^h} a_{mn}^2$ by taking into account different decorrelation rate according to Theorem 2.1, e.g., dividing those points in each j -th square (cube) neighbors of \mathbf{y}_m (see Figure 6) into directional cone sections according to whether the line connecting \mathbf{y}_m and its neighbors intersecting X or not, which gives different decorrelation rate of two Green's functions or different power α in $a_{m,n_j} \lesssim (kjh)^{-\alpha}$ for points in different cones, as long as there is a solid angle cone such that lines connecting those neighbors in that section and \mathbf{y}_m intersect X , the order of the estimate can not be improved. On the other hand, whether (60) and (61) are sharp or not is a more complicated issue. The answer depends on the variation of leading eigenvalues of the matrix $A = \langle \hat{G}(\cdot, \mathbf{y}_m), \hat{G}(\cdot, \mathbf{y}_n) \rangle$, which depends on the geometric setup of X and Y , and the choice of ϵ .

3.2. Lower Bound and Upper Bound Estimate for Approximate Separability of Green's Function. Now we use Lemma 3.1 to prove the following lower bound estimate for approximate separability (14) of Green's function for Helmholtz equation (5) in the high frequency limit.

Theorem 3.1. Let X, Y be two compact manifolds embedded in R^d , $d = 2, 3$, and $d \geq \dim(X) \geq \dim(Y) = s$. Assume that for any two points $\mathbf{y}_1, \mathbf{y}_2 \in Y$,

$$| \langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle | \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\alpha} \quad \text{as } k|\mathbf{y}_1 - \mathbf{y}_2| \rightarrow \infty$$

for some $\alpha > 0$. If there are $f_l(\mathbf{x}) \in L_2(X), g_l(\mathbf{y}) \in L_2(Y), l = 1, 2, \dots, N_k^\epsilon$ such that

$$(76) \quad \left\| \hat{G}(\mathbf{x}, \mathbf{y}) - \sum_{l=1}^{N_k^\epsilon} f_l(\mathbf{x})g_l(\mathbf{y}) \right\|_{L_2(X \times Y)} \leq \epsilon,$$

then

$$(77) \quad N_k^\epsilon \geq \begin{cases} c_\epsilon k^{2\alpha}, & \alpha < \frac{s}{2}, \\ c_\epsilon k^{s-\delta}, & \alpha \geq \frac{s}{2}, \end{cases}$$

for any $0 < \delta < 1$ and arbitrary close to 0 as $k \rightarrow \infty$, where $c_\epsilon \geq c(1 - (C\epsilon)^2)$ for some positive constants c and C that only depend on X, Y and $n(\mathbf{x})$.

Proof. Without loss of generality, we assume Y is a line of unit length, a unit square or a unit cube for $s = 1, 2, 3$ respectively. First, put down a uniform coarse grid in Y with grid size $\bar{h} = k^{\delta-1}$, $0 < \delta < 1$ and is arbitrary close to 0. The grid divides Y into cells $Y_m, m = 1, 2, \dots, N_k^{\bar{h}} = k^{s(1-\delta)}$. Divide each coarse cell Y_m further into uniform finer cells of size $\underline{h} < k^{-1}$, $Y_{m,n}, n = 1, 2, \dots, N_k^{\underline{h}} = (\bar{h}/\underline{h})^s$. See Figure 7. For a fixed n , $Y_{m,n}$ is in the same relative location in each coarse cell Y_m . The center and volume of each cell $Y_{m,n}$ is $\mathbf{y}_{m,n}$ and \underline{h}^s respectively. Define $\hat{G}_{\underline{h}}(\mathbf{x}, \mathbf{y}) = \hat{G}(\mathbf{x}, \mathbf{y}_{m,n}), \forall \mathbf{y} \in Y_{m,n}$ to be a piecewise constant function in \mathbf{y} . We show that by choosing $\underline{h} < k^{-1}$ small enough

$$(78) \quad \int_Y d\mathbf{y} \int_X |\hat{G}(\mathbf{x}, \mathbf{y}) - \hat{G}_{\underline{h}}(\mathbf{x}, \mathbf{y})|^2 d\mathbf{x} = \sum_{m=1}^{N_k^{\bar{h}}} \sum_{n=1}^{N_k^{\underline{h}}} \int_{Y_{m,n}} d\mathbf{y} \int_X |\hat{G}(\mathbf{x}, \mathbf{y}) - \hat{G}(\mathbf{x}, \mathbf{y}_{m,n})|^2 d\mathbf{x} \lesssim \epsilon^2$$

If X, Y are disjoint, assuming $n(\mathbf{x})$ is smooth in the Helmholtz equation (5) and there is no caustics, we have $|\nabla_{\mathbf{y}} \hat{G}(\mathbf{x}, \mathbf{y})| \lesssim k$ uniformly in $\mathbf{x} \in X$ and $\mathbf{y} \in Y$. One can pick $\underline{h} < k^{-1}$ small enough such that

$$|\hat{G}(\mathbf{x}, \mathbf{y}) - \hat{G}(\mathbf{x}, \mathbf{y}_{m,n})| \leq \epsilon, \quad \forall \mathbf{x} \in X, \mathbf{y} \in Y_{m,n}, \quad m = 1, 2, \dots, N_k^{\bar{h}}, n = 1, 2, \dots, N_k^{\underline{h}}$$

and get (78).

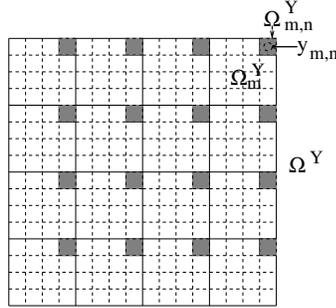


FIGURE 7. Two scale decomposition of the source domain.

If X and Y are not disjoint and $\dim(X) = \dim(Y) = d = 2$ or 3 , $G(\cdot, \mathbf{y}) \in L_2(X)$, $\forall \mathbf{y} \in Y$, we show that there is still $\underline{h} < k^{-1}$ small enough such that

$$(79) \quad \int_X |\hat{G}(\mathbf{x}, \mathbf{y}) - \hat{G}(\mathbf{x}, \mathbf{y}_{m,n})|^2 d\mathbf{x} \lesssim \epsilon^2, \quad \forall \mathbf{y} \in Y_{m,n},$$

which implies (78). From (12) and the asymptotic formulas (7), (8), we have

$$(80) \quad \int_{B_\tau(\mathbf{y})} |\hat{G}(\mathbf{x}, \mathbf{y})|^2 d\mathbf{x} \lesssim \tau$$

with the constant independent of k , where $B_\tau(\mathbf{y})$ denotes the ball with radius τ centered at \mathbf{y} . Hence there are balls $B_{\tau(\epsilon)}(\mathbf{y})$ and $B_{\tau(\epsilon)}(\mathbf{y}_{m,n})$ with radius $\tau(\epsilon) \sim \epsilon^2$ such that

$$(81) \quad \int_{X \cap B_{\tau(\epsilon)}(\mathbf{y})} |\hat{G}(\mathbf{x}, \mathbf{y})|^2 d\mathbf{x} \leq \epsilon^2, \quad \int_{X \cap B_{\tau(\epsilon)}(\mathbf{y}_{m,n})} |\hat{G}(\mathbf{x}, \mathbf{y}_{m,n})|^2 d\mathbf{x} \leq \epsilon^2$$

for a given ϵ , Denote $X_\epsilon = X \cap (B_{\tau(\epsilon)}(\mathbf{y}) \cup B_{\tau(\epsilon)}(\mathbf{y}_{m,n}))$ and $X_\epsilon^C = X \setminus X_\epsilon$. For $\mathbf{y} \in Y_{m,n}$ we have

$$\begin{aligned} & \int_X |\hat{G}(\mathbf{x}, \mathbf{y}) - \hat{G}(\mathbf{x}, \mathbf{y}_{m,n})|^2 d\mathbf{x} \\ &= \int_{X_\epsilon^C} |\hat{G}(\mathbf{x}, \mathbf{y}) - \hat{G}(\mathbf{x}, \mathbf{y}_{m,n})|^2 d\mathbf{x} + \int_{X_\epsilon} |\hat{G}(\mathbf{x}, \mathbf{y}) - \hat{G}(\mathbf{x}, \mathbf{y}_{m,n})|^2 d\mathbf{x} = I + II \end{aligned}$$

Since $|\nabla_{\mathbf{y}} \hat{G}(\mathbf{x}, \mathbf{y})| \lesssim \max(k\tau^{-1}(\epsilon), \tau^{-2}(\epsilon))$, $\forall \mathbf{x} \in X_\epsilon^C$, by choosing $\underline{h} < k^{-1}$ small enough we get $I \lesssim \epsilon^2$. From (81) we get $II \lesssim \epsilon^2$. Hence we prove (79).

Let the linear subspace $S_X = \text{span}\{f_p(\mathbf{x})\}_{p=1}^{N_k^\epsilon} \subset L_2(X)$, then

$$\int_Y \|\hat{G}(\mathbf{x}, \mathbf{y}) - P_{S_X} \hat{G}(\mathbf{x}, \mathbf{y})\|_{L_2(X)}^2 d\mathbf{y} \leq \epsilon^2,$$

where P_{S_X} is the projection onto S_X . From (78), we have

$$\int_Y \|(I - P_{S_X})[\hat{G}(\mathbf{x}, \mathbf{y}) - \hat{G}_{\underline{h}}(\mathbf{x}, \mathbf{y})]\|_{L_2(X)}^2 d\mathbf{y} \lesssim O(\epsilon^2),$$

where I is the identity map, from (78) we have

$$\begin{aligned} & \epsilon^2 \gtrsim \int_Y \|(I - P_{S_X})\hat{G}_{\underline{h}}(\mathbf{x}, \mathbf{y})\|_{L_2(X)}^2 d\mathbf{y} \\ &= \sum_{m=1}^{N_k^{\bar{h}}} \sum_{n=1}^{N_k^{\underline{h}}} \int_{Y_{m,n}} \|(I - P_{S_X})\hat{G}(\mathbf{x}, \mathbf{y}_{m,n})\|_{L_2(X)}^2 d\mathbf{y} \\ (82) \quad &= \underline{h}^s \sum_{m=1}^{N_k^{\bar{h}}} \sum_{n=1}^{N_k^{\underline{h}}} \|\hat{G}(\mathbf{x}, \mathbf{y}_{m,n}) - P_{S_X} \hat{G}(\mathbf{x}, \mathbf{y}_{m,n})\|_{L_2(X)}^2 \\ &= (\underline{h}/\bar{h})^s \bar{h}^s \sum_{m=1}^{N_k^{\bar{h}}} \sum_{n=1}^{N_k^{\underline{h}}} \|\hat{G}(\mathbf{x}, \mathbf{y}_{m,n}) - P_{S_X} \hat{G}(\mathbf{x}, \mathbf{y}_{m,n})\|_{L_2(X)}^2 \\ &\gtrsim \frac{1}{N_k^{\underline{h}}} \sum_{n=1}^{N_k^{\underline{h}}} \frac{1}{N_k^{\bar{h}}} \sum_{m=1}^{N_k^{\bar{h}}} \|\hat{G}(\mathbf{x}, \mathbf{y}_{m,n}) - P_{S_X} \hat{G}(\mathbf{x}, \mathbf{y}_{m,n})\|_{L_2(X)}^2 \end{aligned}$$

Assume

$$(83) \quad \sum_{m=1}^{N_k^{\bar{h}}} \|\hat{G}(\mathbf{x}, \mathbf{y}_{m,\bar{n}}) - P_{S_X} \hat{G}(\mathbf{x}, \mathbf{y}_{m,\bar{n}})\|_{L_2(X)}^2 = \min_n \sum_{m=1}^{N_k^{\bar{h}}} \|\hat{G}(\mathbf{x}, \mathbf{y}_{m,n}) - P_{S_X} \hat{G}(\mathbf{x}, \mathbf{y}_{m,n})\|_{L_2(X)}^2$$

Then there is a constant $C > 0$ such that

$$(84) \quad \frac{1}{N_k^{\bar{h}}} \sum_{m=1}^{N_k^{\bar{h}}} \|\hat{G}(\mathbf{x}, \mathbf{y}_{m,\bar{n}}) - P_{S_X} \hat{G}(\mathbf{x}, \mathbf{y}_{m,\bar{n}})\|_{L_2(X)}^2 \leq C^2 \epsilon^2$$

Since $\mathbf{y}_{m,\bar{n}} \in Y, m = 1, 2, \dots, N_k^{\bar{h}} = k^{s(1-\delta)}$ forms a uniform grid with grid size $\bar{h} = k^{\delta-1}$, we can apply Lemma 3.1 to get

$$\dim(S_X) \geq \begin{cases} c(1 - (C\epsilon)^2)^2 k^{2\alpha} & \alpha < \frac{s}{2} \\ c(1 - (C\epsilon)^2)^2 k^{s-\delta} & \alpha \geq \frac{s}{2} \end{cases}$$

for any $0 < \delta < 1$ and arbitrary close to 0 as $k \rightarrow \infty$, where $C > 0, c > 0$ are some constants that depend only on X and Y and $n(\mathbf{x})$. □

Remark 3.4. *Theorem 3.1 presents the intrinsic mathematical difficulty for numerical computation of Helmholtz equation with large wave number k . When $n = 3$, the same lower bound estimate for approximate separability holds for unnormalized Green's function too. Hence, a discretized system, such as the discretized kernel in boundary integral formulation or inverse of the matrix corresponding to direct discretization of the PDE and its off-diagonal sub-matrices corresponding to a mesh that resolves the wavelength, does not have low rank approximation for large k due to the decorrelation of Green's functions caused by fast oscillations. However, when $n = 2$, for separated X and Y , $|G(\mathbf{x}, \mathbf{y})| \lesssim k^{-\frac{1}{2}}, \forall (\mathbf{x}, \mathbf{y}) \in X \times Y$ which approximate separability trivial when k is large enough.*

Remark 3.5. *Although it seems that the lower bound estimate (77) in Theorem 3.1 shows a weak dependence of N_k^ϵ on ϵ as $k \rightarrow \infty$, the choice of ϵ can affect the sharpness of the lower bound as commented in Remark 3.3.*

Here we also give an upper bound estimate for the approximate separability of Green's function in the high frequency limit. The intuition is that the Green's functions with sources located within a wavelength are correlated. So use the linear subspace spanned by Green's functions sampled uniformly in the source domain with a grid size smaller than wavelength should approximate the whole family of Green's functions good enough.

Theorem 3.2. *Let X, Y be two compact manifolds embedded in $R^d, d = 2, 3$ and $d \geq \dim(X) \geq \dim(Y) = s$. For any $\epsilon > 0$ and $\delta > 0$, there are $f_l(\mathbf{x}) \in L_2(X), g_l(\mathbf{y}) \in$*

$L_2(Y), l = 1, 2, \dots, N_k^\epsilon \leq Ck^{s+\delta}$ such that

$$(85) \quad \left\| \hat{G}(\mathbf{x}, \mathbf{y}) - \sum_{l=1}^{N_k^\epsilon} f_l(\mathbf{x})g_l(\mathbf{y}) \right\|_{L_2(X \times Y)} \leq \epsilon$$

as $k \rightarrow \infty$, where $C > 0$ is some constant that depends on X, Y and $n(\mathbf{x})$.

Proof. Put down a uniform grid in Y , which is assumed to be a line of unit length, a unit square or a unit cube for $s = 1, 2, 3$ respectively, with grid size $h = k^{-1-\delta/2}$. Denote the grid point as $\mathbf{y}_m, m = 1, 2, \dots, N_k^h = k^{s(1+\delta/2)}$. Denote the linear subspace $S_X = \text{span}\{\hat{G}(\mathbf{x}, \mathbf{y}_m)\}_{m=1}^{N_k^h} \subset L_2(X)$. We claim

$$(86) \quad \|\hat{G}(\mathbf{x}, \mathbf{y}) - P_{S_X} \hat{G}(\mathbf{x}, \mathbf{y})\|_{L_2(X)} \leq |Y|^{-\frac{1}{2}} \epsilon$$

for k large enough, where $|Y|$ denotes the volume of Y .

If X and Y are disjoint, assuming $n(\mathbf{x})$ is smooth in the Helmholtz equation (5) and there is no caustics, we have $|\nabla_{\mathbf{y}} \hat{G}(\mathbf{x}, \mathbf{y})| \lesssim k, \|D_{\mathbf{y}}^2 \hat{G}(\mathbf{x}, \mathbf{y})\| \lesssim k^2$, where the bound is uniform in X and Y . Given a non-grid point $\mathbf{y} \in Y$, $\hat{G}(\mathbf{x}, \mathbf{y})$ can be approximated by a linear interpolation, which is convex combination of Green's function at its neighboring grid points. To be precise, suppose $\mathbf{y} \in Y$ lies in the s dimensional simplex with vertices $\mathbf{y}_{m_1}, \dots, \mathbf{y}_{m_{s+1}}$. Let $r_{\mathbf{y}}^1, \dots, r_{\mathbf{y}}^{s+1}$ be the barycentric coordinates for \mathbf{y} , i.e.,

$$\mathbf{y} = \sum_{j=1}^{s+1} r_{\mathbf{y}}^j \mathbf{y}_{m_j}, \quad 1 \geq r_{\mathbf{y}}^j \geq 0, \quad \sum_{j=1}^{s+1} r_{\mathbf{y}}^j = 1.$$

Then we have the following linear interpolation

$$(87) \quad |\hat{G}(\mathbf{x}, \mathbf{y}) - \sum_{j=1}^{s+1} r_{\mathbf{y}}^j \hat{G}(\mathbf{x}, \mathbf{y}_{m_j})| \lesssim \|D_{\mathbf{y}}^2 \hat{G}(\mathbf{x}, \mathbf{y})\| h^2 \lesssim k^{-\delta},$$

and hence (86) is true when k is large enough.

If X and Y are not disjoint and $\dim(X) = \dim(Y) = d = 2, 3$, $\hat{G}(\cdot, \mathbf{y}) \in L_2(X)$. For a given ϵ , there are balls $B_{\tau(\epsilon)}(\mathbf{y})$ and $B_{\tau(\epsilon)}(\mathbf{y}_{m_j})$ with radius $0 < \tau(\epsilon) \sim \epsilon^2$ centered at \mathbf{y} and $\mathbf{y}_{m_j}, j = 1, 2, \dots, d+1$ such that

$$(88) \quad \int_{B_{\tau(\epsilon)}(\mathbf{y})} |\hat{G}(\mathbf{x}, \mathbf{y})|^2 d\mathbf{x} \leq \frac{\epsilon^2}{8|Y|(d+2)}, \quad \int_{B_{\tau(\epsilon)}(\mathbf{y}_{m_j})} |\hat{G}(\mathbf{x}, \mathbf{y}_{m_j})|^2 d\mathbf{x} \leq \frac{\epsilon^2}{8|Y|(d+2)}.$$

Also we have $|\nabla_{\mathbf{y}} \hat{G}(\mathbf{x}, \mathbf{y})| \lesssim \max(k\tau^{-1}(\epsilon), \tau^{-2}(\epsilon)), \|D_{\mathbf{y}}^2 \hat{G}(\mathbf{x}, \mathbf{y})\| \lesssim \max(k^2\tau^{-1}(\epsilon), k\tau^{-2}(\epsilon), \tau^{-3}(\epsilon))$ and hence

$$(89) \quad |\hat{G}(\mathbf{x}, \mathbf{y}) - \sum_{j=1}^{s+1} r_{\mathbf{y}}^j \hat{G}(\mathbf{x}, \mathbf{y}_{m_j})| \lesssim \|D_{\mathbf{y}}^2 \hat{G}(\mathbf{x}, \mathbf{y})\| h^2 \lesssim \max(k^{2(1-\gamma)}, k^{1-2\gamma}\tau^{-2}(\epsilon), k^{-2\gamma}\tau^{-3}(\epsilon)).$$

for any $\mathbf{x} \in X \setminus B_{\tau(\epsilon)}(\mathbf{y}) \cup (\cup_{j=1}^{d+1} B_{\tau(\epsilon)}(\mathbf{y}_{m_j}))$. By decomposing the integration in (86) on X into two parts, one on $X_\epsilon = X \cap (B_{\tau(\epsilon)}(\mathbf{y}) \cup (\cup_{j=1}^{d+1} B_{\tau(\epsilon)}(\mathbf{y}_{m_j})))$ and the other on $X_\epsilon^C = X \setminus (B_{\tau(\epsilon)}(\mathbf{y}) \cup (\cup_{j=1}^{d+1} B_{\tau(\epsilon)}(\mathbf{y}_{m_j})))$, we have

$$\begin{aligned}
 & \int_{X_\epsilon} |\hat{G}(\mathbf{x}, \mathbf{y}) - \sum_{j=1}^{d+1} r_{\mathbf{y}}^j \hat{G}(\mathbf{x}, \mathbf{y}_{m_j})|^2 d\mathbf{x} \\
 & \leq 2 \int_{X_\epsilon} |\hat{G}(\mathbf{x}, \mathbf{y})|^2 + \sum_{j=1}^{d+1} r_{\mathbf{y}}^j |\hat{G}(\mathbf{x}, \mathbf{y}_{m_j})|^2 d\mathbf{x} \\
 (90) \quad & \leq 2 \int_{B_{\tau(\epsilon)}(\mathbf{y}) \cup (\cup_{j=1}^{d+1} B_{\tau(\epsilon)}(\mathbf{y}_{m_j}))} |\hat{G}(\mathbf{x}, \mathbf{y})|^2 + \sum_{j=1}^{d+1} r_{\mathbf{y}}^j |\hat{G}(\mathbf{x}, \mathbf{y}_{m_j})|^2 d\mathbf{x} \\
 & \leq \frac{\epsilon^2}{2|Y|},
 \end{aligned}$$

and from (89)

$$(91) \quad \int_{X_\epsilon^C} |\hat{G}(\mathbf{x}, \mathbf{y}) - \sum_{j=1}^{d+1} r_{\mathbf{y}}^j \hat{G}(\mathbf{x}, \mathbf{y}_{m_j})|^2 d\mathbf{x} \leq \frac{\epsilon^2}{2|Y|}$$

when k is large enough. Combining the above two parts we get (86) when k is large enough, which implies

$$(92) \quad \sqrt{\int_Y \|\hat{G}(\mathbf{x}, \mathbf{y}) - P_{S_X} \hat{G}(\mathbf{x}, \mathbf{y})\|_{L_2(X)}^2 d\mathbf{y}} \leq \epsilon.$$

□

Remark 3.6. *The above upper bound holds for unnormalized Green's function when $d = 3$.*

If $\dim(X) = \dim(Y) = d = 2, 3$, since the Green's function belongs to $L_2(X \times Y)$, our approximate separability estimates in L_2 norm is valid for general compact domains X and Y , disjoint or not. However, if X and Y are disjoint, which is the case for most applications, our results in L_2 norm are also true in $L_\infty(X \times Y)$ norm. Since L_∞ norm is stronger than L_2 norm in a compact domain, the lower bound for approximate separability in Theorem 3.1 immediately extends to L_∞ norm. Also, first part of the proof in Theorem 3.2 directly extends to L_∞ norm. We summarize these two results below.

Theorem 3.3. *Let X, Y be two disjoint compact manifolds embedded in \mathbb{R}^d , $d = 2, 3$, and $d \geq \dim(X) \geq \dim(Y) = s$. Assume that for any two points $\mathbf{y}_1, \mathbf{y}_2 \in Y$,*

$$|\langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\alpha} \quad \text{as } k|\mathbf{y}_1 - \mathbf{y}_2| \rightarrow \infty$$

for some $\alpha > 0$. If there are $f_l(\mathbf{x}) \in L_\infty(X)$, $g_l(\mathbf{y}) \in L_\infty(Y)$, $l = 1, 2, \dots, N_k^\epsilon$ such that

$$(93) \quad \left| \hat{G}(\mathbf{x}, \mathbf{y}) - \sum_{l=1}^{N_k^\epsilon} f_l(\mathbf{x}) g_l(\mathbf{y}) \right| \leq \epsilon, \quad \forall \mathbf{x} \in X, \forall \mathbf{y} \in Y,$$

then

$$(94) \quad N_k^\epsilon \geq \begin{cases} c_\epsilon k^{2\alpha}, & \alpha < \frac{s}{2}, \\ c_\epsilon k^{d-\delta}, & \alpha \geq \frac{s}{2}, \end{cases}$$

for any $0 < \delta < 1$ and arbitrary close to 0 as $k \rightarrow \infty$, where $c_\epsilon \geq c(1 - (C\epsilon)^2)^2$ for some positive constants c and C that only depend on X, Y and $n(\mathbf{x})$.

Theorem 3.4. *Let X, Y be two separated compact manifolds embedded in $R^d, d = 2, 3$, and $d \geq \dim(X) \geq \dim(Y) = s$. For any $\epsilon > 0$ and $\delta > 0$, there are $f_l(\mathbf{x}) \in L_\infty(X), g_l(\mathbf{y}) \in L_\infty(Y), l = 1, 2, \dots, N_k^\epsilon \leq Ck^{d+\delta}$ such that*

$$(95) \quad \left| \hat{G}(\mathbf{x}, \mathbf{y}) - \sum_{l=1}^{N_k^\epsilon} f_l(\mathbf{x})g_l(\mathbf{y}) \right| \leq \epsilon, \quad \forall \mathbf{x} \in X, \forall \mathbf{y} \in Y,$$

as $k \rightarrow \infty$, where $C > 0$ is some constant that depends on X, Y and $n(\mathbf{x})$.

Remark 3.7. *In Theorem 3.2, upper bound estimate for the approximate separability of Green's function for the Helmholtz equation in the high frequency limit in L_2 norm is derived based on separable approximation using linear combination (interpolation) of a set of Green's functions with sources located on a uniform grid. It is also possible to obtain an upper bound estimate in the L_2 norm for the approximate separability of Green's function for the Helmholtz equation in the high frequency limit in homogeneous medium for an arbitrary bounded domain based on separable approximation using the eigenfunctions of Laplace operator and Weyl's asymptotic formula [28] for the eigenvalues.*

Suppose $G(\mathbf{x}, \mathbf{y})$ is the Green's function in a bounded domain $\Omega \subset R^d$ satisfying

$$(96) \quad \begin{cases} \Delta_{\mathbf{x}}G(\mathbf{x}, \mathbf{y}) + k^2G(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y}), & \mathbf{x}, \mathbf{y} \in \Omega, \\ G(\mathbf{x}, \mathbf{y}) = 0, & \mathbf{x} \in \partial\Omega. \end{cases}$$

Let $u_m(\mathbf{x}), \|u_m\|_{L_2(\Omega)} = 1, m = 1, 2, \dots$ be the normalized eigenfunctions for the Laplace operator

$$(97) \quad \Delta u_m(\mathbf{x}) = \lambda u_m(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad u_m(\mathbf{x}) = 0, \mathbf{x} \in \partial\Omega$$

with eigenvalues $0 > \lambda_1 \geq \lambda_2 \geq \dots$. Hence $u_m(\mathbf{x}), m = 1, 2, \dots$ are also the normalized eigenfunctions for the homogeneous Helmholtz operator with eigenvalues $\lambda_m + k^2, m = 1, 2, \dots$. Here we assume the domain Ω is not resonant, i.e., $\lambda_m + k^2 \neq 0, \forall m$. Since $u_m(\mathbf{x})$ forms an orthonormal basis for $L_2(\Omega)$ and $G(\mathbf{x}, \mathbf{y}) \in L_2(\Omega)$ for $d = 2, 3$, one has the following expansion

$$(98) \quad G(\mathbf{x}, \mathbf{y}) = \sum_{m=1}^{\infty} (\lambda_m + k^2)^{-1} u_m(\mathbf{y})u_m(\mathbf{x}).$$

The Weyl's asymptotic formula gives

$$(99) \quad |\lambda_m| \sim \frac{4\pi^2 m^{2/d}}{(C_d |\Omega|)^{2/d}}$$

for large m , where $|\Omega|$ is the volume of Ω . Choose a large enough integer $M \gtrsim k^{d+\delta}$, $\delta > 0$, then $|\lambda_m + k^2|^{-1} \lesssim |\lambda_m|^{-1} \lesssim m^{-\frac{2}{d}}$ for $m > M$. For any $\epsilon > 0$, we have

$$(100) \quad \int_{\Omega} \int_{\Omega} |G(\mathbf{x}, \mathbf{y}) - \sum_{m=1}^M (\lambda_m + k^2)^{-1} u_m(\mathbf{y}) u_m(\mathbf{x})|^2 d\mathbf{x} d\mathbf{y} \lesssim \sum_{m=M+1}^{\infty} m^{-\frac{4}{d}} \lesssim M^{1-\frac{4}{d}} < \epsilon^2$$

when k is large enough for $d = 2, 3$. Hence one can use eigenfunctions of Laplace operators to construct a separable approximation for the Green's function of Helmholtz equation. In particular for any two subdomains X, Y of Ω , (100) implies

$$(101) \quad \begin{aligned} & \left\| G(\mathbf{x}, \mathbf{y}) - \sum_{m=1}^M (\lambda_m + k^2)^{-1} u_m(\mathbf{y}) u_m(\mathbf{x}) \right\|_{L_2(X \times Y)} \\ & \leq \left\| G(\mathbf{x}, \mathbf{y}) - \sum_{m=1}^M (\lambda_m + k^2)^{-1} u_m(\mathbf{y}) u_m(\mathbf{x}) \right\|_{L_2(\Omega \times \Omega)} < \epsilon \end{aligned}$$

which shows that $N_k^\epsilon \lesssim k^{d+\delta}$ for any $\delta > 0$ in the high frequency limit.

Remark 3.8. It can be seen from Theorem 3.1 and 3.2 as well as Remark 3.7 that if two Green's functions at different sources decorrelate fast, $\left| \langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle \right| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\alpha}$ with $\alpha \geq \frac{s}{2}$, $s = \dim(Y) \leq \dim(X)$, the upper and lower estimate of the approximate separability of Green's function for Helmholtz equation is sharp in the high frequency limit. In this scenario, a set of Green's functions with sources uniformly distributed in a domain or a set of leading eigenfunctions for Laplace operator form a good basis to represent an arbitrary Green's function or solution. However, the representation is not much compressible and hence low rank approximation in numerical computation is not feasible. Moreover, to compute a set of densely distributed Green's functions or a large number of eigenfunctions of Laplace operator of order at least $O(k^s)$ for a s -dimensional manifold for a general setup in practice is computationally challenging. In certain setups, explicitly known special functions/basis can be explored to develop fast algorithms even for dense matrices that do not have low rank approximation, such as fast Fourier transform (FFT).

4. EXAMPLES

In this section we apply our previous general results to a few setups that are of practical interests in 3D. Again we assume that X, Y are two compact manifolds embedded in R^3 and $\dim(X) \geq \dim(Y) = s$. First, in Section 4.1, we discuss several examples for which the Green's function for Helmholtz equation is not highly separable in the high frequency limit. In particular, if the decorrelation of two Green's functions is fast enough, $\left| \langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle \right| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\alpha}$, $\alpha \geq \frac{s}{2}$, the lower bound and upper bound for the approximate separability of Green's function for Helmholtz equations in the high frequency limit are sharp. As discussed in the proof of Theorem 2.1, Remark 2.3 and Section 2.2, if a generic ray going through two points, $\mathbf{y}_1, \mathbf{y}_2 \in Y$, only intersect X at isolated points, then $\alpha = \frac{s}{2}$ and hence both lower bound and upper bound are sharp. For all these examples,

where X and Y are given and fixed, the lower bound in Theorem 3.1, sharp or not, implies that there is no low rank approximation in the corresponding discretized system due to the fast decorrelation of Green's function of Helmholtz equation in the high frequency limit. However, there are special k dependent setups for X and Y , which are discussed in Section 4.2, where the Green's function for Helmholtz equation is highly separable even in the high frequency limit. High separability in these special setups, which implies availability of low rank approximation in the corresponding discretized system, can be explored to develop fast algorithms. In the following study, all constants are positive and only depend on X, Y and $n(\mathbf{x})$ by assuming ϵ is small so that the weak dependence of those estimates in Theorem 3.1, 3.3 on it is neglected.

4.1. Examples of not highly separable Green's function.

1) X and Y are two disjoint compact domains in R^3 , $\dim(X) = \dim(Y) = s = 3$. For two points $\mathbf{y}_1, \mathbf{y}_2 \in Y$, we can only claim $|\langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle_X| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-1}$ in general from Theorem 2.1 since for any point $\mathbf{y} \in Y$ there is a cone with \mathbf{y} as the vertex and a solid angle. A segment of the ray (a straight line in homogeneous medium) connecting \mathbf{y} and a point in the cone stays in X which gives a 1D curve of stationary phase. Since $\alpha = 1 < \frac{s}{2}$ our lower bound and upper bound estimates are not sharp. Theorem 3.1, 3.3 give the lower bound $N_k^\epsilon \gtrsim k^2$ while Theorem 3.2, 3.4 give the upper bound $N_k^\epsilon \lesssim k^{3+\delta}$ for any $\delta > 0$.

2) X and Y are two disjoint compact surfaces in 3D, $\dim(X) = \dim(Y) = s = 2$. This is a typical scenario for boundary integral methods when X and Y are two pieces of the boundary of scatterers. In general, the ray (straight line in homogeneous medium) going through two points $\mathbf{y}_1, \mathbf{y}_2 \in Y$ intersects X at most finite number of times, i.e., there are only isolated stationary points for the oscillatory surface integral of $|\langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle_X|$, so $\alpha = \frac{s}{2}$ if there are isolated stationary phase points in the interior of X or $\alpha = \frac{s+1}{2}$ if there are only isolated stationary phase points on the boundary ∂X (see Remark 2.3). For both cases, we have the following sharp estimate from Theorem 3.1- 3.4

$$k^{2-\delta} \lesssim N_k^\epsilon \lesssim k^{2+\delta}, \quad \forall \delta > 0,$$

as $k \rightarrow \infty$.

Another interesting scenario is when people compute the direct inverse of the discretized linear system for Helmholtz equation using multi-frontal method in which the full linear system is reduced to smaller but dense linear systems corresponding to unknowns living on planar domains such as those depicted in Figure 8. Low rank approximation or further skeletonization for these smaller but dense linear systems is crucial for a fast numerical solver. We use our theory to give the approximate separability estimates for the typical setups in the high frequency limit. Since $s = 2$, the upper bound is $N_k^\epsilon \lesssim k^{2+\delta}$ for any $\delta > 0$. Now let's look at the lower bound. We first look at three typical configurations in homogeneous medium where rays are (piecewise) straight lines: (a) if X, Y are two disjoint coplanar regions as shown in Figure 8(a), the least rate of decorrelation between two Greens's function is $\alpha = \frac{1}{2}$, i.e., $|\langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle_X| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\frac{1}{2}}$ by Theorem 2.1

when a segment of the ray going through $\mathbf{y}_1, \mathbf{y}_2$ stay in X . So one has $N_k^\epsilon \gtrsim k$ as $k \rightarrow \infty$; (b) if X, Y are two disjoint planar regions that are not coplanar nor parallel to each other, e.g. perpendicular to each other as shown in Figure 8(b). The least rate of decorrelation between two Greens's function is $\alpha = 1$ by applying the standard stationary phase result since any ray going through two points $\mathbf{y}_1, \mathbf{y}_2 \in Y$ intersect X at most finite times (0 or 1). So one has $N_k^\epsilon \gtrsim k^{2-\delta}, \forall \delta > 0$ as $k \rightarrow \infty$; (c) if X, Y are two planar regions in parallel as shown in Figure 8(c), the least rate of decorrelation between two Greens's function is $\alpha = \frac{3}{2}$ (in general since any ray going through two points $\mathbf{y}_1, \mathbf{y}_2 \in Y$ does not intersect X) or 1 (if part of the boundary ∂X stays on a level set of the phase function $|\mathbf{x} - \mathbf{y}_1| - |\mathbf{x} - \mathbf{y}_2|$), see Remark 2.3. So one has $N_k^\epsilon \gtrsim k^{2-\delta}, \forall \delta > 0$ as $k \rightarrow \infty$. If the medium is inhomogeneous, a ray going through two points $\mathbf{y}_1, \mathbf{y}_2$ is not straight line and intersect a planar region X only a finite number of times. So $\alpha \geq 1$ and we have the sharp low bound estimate $N_k^\epsilon \gtrsim k^{2-\delta}, \forall \delta > 0$ in general for the above three configurations. For a discretization with a fixed ratio of grid size and wavelength, the sharp lower bound means that the discretized matrix and its sub matrices are full rank modulo a constant.

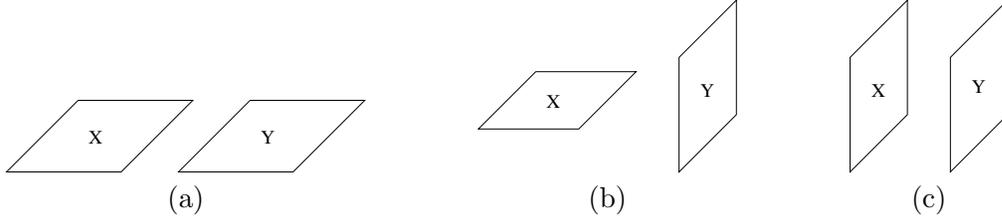


FIGURE 8. Two planar surfaces in 3D.

3) $X \in R^3$ is a compact domain, Y is a compact smooth curve ($\dim(Y) = s = 1$) or surface ($\dim(Y) = s = 2$) and X, Y are disjoint. Given any two different points $\mathbf{y}_1, \mathbf{y}_2 \in Y$, we have $|\langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle_X| \lesssim (k|\mathbf{y}_1 - \mathbf{y}_2|)^{-\alpha}$ for some $1 \leq \alpha \leq 2$ from Theorem 2.1. Since $\alpha \geq \frac{s}{2}$, from Theorem 3.1- 3.4 one has the following sharp estimates

$$k^{s-\delta} \lesssim N_k^\epsilon \lesssim k^{s+\delta}, \quad \forall \delta > 0,$$

as $k \rightarrow \infty$.

4) X is a 2D surface or 1D curve and Y is a 1D curve. One has $\alpha = 1$ if X is a surface or $\alpha = \frac{1}{2}$ if X is a curve using standard stationary phase theory for a general X and two points $\mathbf{y}_1, \mathbf{y}_2 \in Y$. So for both scenarios we have sharp bounds

$$k^{1-\delta} \lesssim N_k^\epsilon \gtrsim k^{1+\delta} \quad \forall \delta > 0,$$

as $k \rightarrow \infty$.

4.2. Highly Separable Cases. Although in this study we have shown that the lower bound for the number of terms, N_k^ϵ , in the approximate separability (2) of the Green's function $G(\mathbf{x}, \mathbf{y})$ for Helmholtz equation in the high frequency limit grows with certain

power of k (Theorem 3.1, 3.3) in general, there are special k dependent setups of X and Y where $G(\mathbf{x}, \mathbf{y})$ is highly separable, i.e., N_k^ϵ does not depend on k and depends on ϵ logarithmically. These special setups can be explored to develop fast algorithms in practice by utilizing low rank approximation.

Let's assume the Green's function is or can be approximated by the form $G(\mathbf{x}, \mathbf{y}) = A(\mathbf{x}, \mathbf{y})e^{ik\phi(\mathbf{x}, \mathbf{y})}$, where $A(\mathbf{x}, \mathbf{y})$ is the amplitude function and $\phi(\mathbf{x}, \mathbf{y})$ is the phase function. The key point is that fast decorrelation or almost orthogonality between Green's functions with different source locations due to rapid change in the phase, as we have shown for general domains X and Y , are not present in those special setups. Since the amplitude function $A(\mathbf{x}, \mathbf{y})$ is independent of k , typically one needs to find $\phi_1(\mathbf{x})$ and $\phi_2(\mathbf{y})$ such that $k(\phi(\mathbf{x}, \mathbf{y}) - \phi_1(\mathbf{x}) - \phi_2(\mathbf{y}))$ is uniformly bounded with respect to $\mathbf{x} \in X, \mathbf{y} \in Y$ and k and has a highly separable approximation, e.g., Taylor expansion in \mathbf{x}, \mathbf{y} , and hence $e^{ik(\phi(\mathbf{x}, \mathbf{y}) - \phi_1(\mathbf{x}) - \phi_2(\mathbf{y}))}$ has a high separable approximation due to the fast convergence of Taylor expansion for e^{ikz} for $|z| \leq C < \infty$. Also it is easy to see that the phase difference between two Green's function at different sources $\mathbf{y}_1, \mathbf{y}_2$ can be written as

$$k(\phi(\mathbf{x}, \mathbf{y}_1) - \phi(\mathbf{x}, \mathbf{y}_2)) = k(\phi_2(\mathbf{y}_1) - \phi_2(\mathbf{y}_2)) + k[(\phi(\mathbf{x}, \mathbf{y}_1) - \phi_1(\mathbf{x}) - \phi_2(\mathbf{y}_1)) - (\phi(\mathbf{x}, \mathbf{y}_2) - \phi_1(\mathbf{x}) - \phi_2(\mathbf{y}_2))]$$

which is a constant phase shift $k(\phi_2(\mathbf{y}_1) - \phi_2(\mathbf{y}_2))$ plus a term that is bounded uniformly with respect to $\mathbf{x} \in X, \mathbf{y} \in Y$ and k . Hence no fast oscillatory integral is present to decorrelate two Green's functions. For simplicity 3D free space Green's function (5) are used for illustration in the following examples.

1) X, Y are two line segments that are collinear as shown in Figure 9(a) in homogeneous medium

$$(102) \quad \langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle = \frac{1}{\|G_0(\cdot, \mathbf{y}_1)\|_2 \|G_0(\cdot, \mathbf{y}_2)\|_2} e^{ik(\mathbf{y}_2 - \mathbf{y}_1)} \int_X \frac{1}{|\mathbf{x} - \mathbf{y}_1| |\mathbf{x} - \mathbf{y}_2|} d\mathbf{x}.$$

There is no fast oscillatory integral to decorrelate two Green's functions. Denote the axis going through these two line segments as r , we have

$$(103) \quad G_0(\mathbf{x}, \mathbf{y}) = \frac{1}{4\pi} \frac{e^{-ik(r_{\mathbf{x}} - r_{\mathbf{y}})}}{r_{\mathbf{y}} - r_{\mathbf{x}}} = \frac{1}{4\pi} e^{-ikr_{\mathbf{x}}} e^{ikr_{\mathbf{y}}} r_{\mathbf{y}}^{-1} \sum_{m=0}^{\infty} \left(\frac{r_{\mathbf{x}}}{r_{\mathbf{y}}}\right)^m.$$

In this trivial case, $\phi(\mathbf{x}, \mathbf{y}) = \phi_1(\mathbf{x}) + \phi_2(\mathbf{y})$, where $\phi_1(\mathbf{x}) = -r_{\mathbf{x}}, \phi_2(\mathbf{y}) = r_{\mathbf{y}}$. It is easy to see that the geometric series can be truncated at $N_k^\epsilon = (\log \frac{l_X}{l_X + \rho})^{-1} \log(4\pi\rho\epsilon)$ to get an a separable approximation for any $\epsilon > 0$ independent of k , where l_X is the length of X and ρ is the distance between X and Y . The same argument can be applied to two disjoint curve segments X and Y that lie on the same ray in heterogeneous medium. In this case

$$(104) \quad \langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle = \frac{1}{\|G(\cdot, \mathbf{y}_1)\|_2 \|G(\cdot, \mathbf{y}_2)\|_2} e^{ik\phi(\mathbf{y}_1, \mathbf{y}_2)} \int_X A(\mathbf{x}, \mathbf{y}_1) A(\mathbf{x}, \mathbf{y}_2) d\mathbf{x},$$

where $\phi(\mathbf{y}_1, \mathbf{y}_2)$ is the travel time between \mathbf{y}_1 and \mathbf{y}_2 and $A(\mathbf{x}, \mathbf{y}_1), A(\mathbf{x}, \mathbf{y}_2)$ are the corresponding amplitudes in geometric optics ansatz as discussed in Section 2.2.

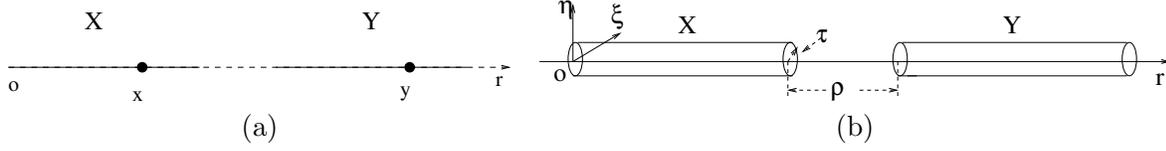


FIGURE 9. Two special setups of X and Y that allow highly separable approximation of the Green's function.

2) X and Y are two disjoint thin cylinders around a line as shown in Figure 9(b). This 3D setup is analogous to the 2D setup in [21]. Numerical test will be presented in Section 5.

Let the longitudinal axis be r and the other two orthogonal coordinates in the plane perpendicular to r be ξ, η . Let $\mathbf{x} = (r_x, \xi_x, \eta_x) \in X$ and $\mathbf{y} = (r_y, \xi_y, \eta_y) \in Y$. Denote $\rho = \inf_{\mathbf{x} \in X, \mathbf{y} \in Y} (r_x - r_y)$ and $\tau = \sup_{\mathbf{x} \in X, \mathbf{y} \in Y} \sqrt{\xi^2 + \eta^2}$. Assume $k\tau < \frac{1}{2}$, $\mu = \frac{\tau}{\rho} < \frac{1}{2}$. Again in this case $\phi_1(\mathbf{x}) = -r_x$, $\phi_2(\mathbf{y}) = r_y$ as in the previous case. One has

$$k|\phi(\mathbf{x}, \mathbf{y}) - \phi_1(\mathbf{x}) - \phi_2(\mathbf{y})| = k(|\mathbf{x} - \mathbf{y}| - (r_y - r_x)) < 2k\tau = 1$$

Next, we give an explicit separable construction using asymptotic expansions.

(105)

$$\begin{aligned} k|\mathbf{x} - \mathbf{y}| &= k\sqrt{(r_x - r_y)^2 + (\xi_x - \xi_y)^2 + (\eta_x - \eta_y)^2} \\ &= k(r_y - r_x) + k\sqrt{(\xi_x - \xi_y)^2 + (\eta_x - \eta_y)^2} \sum_{m=1}^{\infty} \frac{(-1)^m (2m)!}{(1-2m)(m!)^2 4^m} \frac{((\xi_x - \xi_y)^2 + (\eta_x - \eta_y)^2)^{m-1/2}}{(r_y - r_x)^{2m-1}}. \end{aligned}$$

Note that $k\sqrt{(\xi_x - \xi_y)^2 + (\eta_x - \eta_y)^2} \leq 2k\tau < 1$ and $\frac{(-1)^m (2m)!}{(1-2m)(m!)^2 4^m} < \frac{1}{2m}$. So the second term in the above expression can be bounded by a geometric series $\sum_{m=1}^{\infty} (\frac{2\tau}{\rho})^{2m-1}$. So

(106)

$$\langle \hat{G}_0(\cdot, \mathbf{y}_1), \hat{G}_0(\cdot, \mathbf{y}_2) \rangle = \frac{1}{\|G_0(\cdot, \mathbf{y}_1)\|_2 \|G_0(\cdot, \mathbf{y}_2)\|_2} e^{ik(r_{y_1} - r_{y_2})} \int_X \frac{e^{i\tilde{\phi}(\mathbf{x})}}{|\mathbf{x} - \mathbf{y}_1| |\mathbf{x} - \mathbf{y}_2|} d\mathbf{x},$$

where $\tilde{\phi}(\mathbf{x}) = k[(|\mathbf{x} - \mathbf{y}_1| - (r_{y_1} - r_x)) - (|\mathbf{x} - \mathbf{y}_2| - (r_{y_2} - r_x))]$ and $|\tilde{\phi}(\mathbf{x})| = O(1), \forall \mathbf{x} \in X$. Again no fast oscillation due to rapid change of phase is present in the integral. Now we show explicitly the approximate separation based on the expansion (105).

For any tolerance $\epsilon > 0$, take q such that $(\frac{2\tau}{\rho})^{2q+1} (1 - \frac{2\tau}{\rho})^{-1} < \epsilon$. So only the first $q = O(|\log \epsilon|)$ terms are needed in the summation in (105). Denote $Q(\mathbf{x}, \mathbf{y})$ be the first q term expansion,

$$Q(\mathbf{x}, \mathbf{y}) = k\sqrt{(\xi_x - \xi_y)^2 + (\eta_x - \eta_y)^2} \sum_{m=1}^q \frac{(-1)^m (2m)!}{(1-2m)(m!)^2 4^m} \frac{((\xi_x - \xi_y)^2 + (\eta_x - \eta_y)^2)^{m-1/2}}{(r_y - r_x)^{2m-1}}.$$

$Q(\mathbf{x}, \mathbf{y})$ is bounded independent of k . So

$$e^{ik|\mathbf{x} - \mathbf{y}|} = e^{ik(r_y - r_x)} e^{i(Q(\mathbf{x}, \mathbf{y}) + \epsilon)} = e^{ik(r_y - r_x)} e^{iQ(\mathbf{x}, \mathbf{y})} + O(\epsilon).$$

Since e^x can again have a $p = O(|\log \epsilon|)$ term polynomial expansion for any tolerance ϵ for a bounded x , we have

$$e^{iQ(\mathbf{x}, \mathbf{y})} = \sum_{l=0}^p \frac{[iQ(\mathbf{x}, \mathbf{y})]^l}{l!} + O(\epsilon).$$

Let's look at each term in the summation. Since each term in the expansion of $[Q(\mathbf{x}, \mathbf{y})]^l$ is like

$$\left[k \sqrt{(\xi_x - \xi_y)^2 + (\eta_x - \eta_y)^2} \right]^l \frac{\left(\sqrt{(\xi_x - \xi_y)^2 + (\eta_x - \eta_y)^2} \right)^m}{(r_y - r_x)^m}$$

where the integer m ranges from l to $(2q-1)l$ and $m+l$ is even. One only needs to keep those $m \leq 2q$ terms in the expansion because $\left(\frac{2r}{\rho}\right)^{2q+1} = O(\epsilon)$ and $k \sqrt{(\xi_x - \xi_y)^2 + (\eta_x - \eta_y)^2} < 1$. So altogether we have $O(pq) = O(|\log \epsilon|^2)$ terms of the form

$$(107) \quad \frac{((\xi_x - \xi_y)^2 + (\eta_x - \eta_y)^2)^m}{(r_y - r_x)^l},$$

where $0 < l \leq m \lesssim |\log \epsilon|$ are integers. Each term

$$((\xi_x - \xi_y)^2 + (\eta_x - \eta_y)^2)^m = [(\xi_x^2 + \eta_x^2) - 2\xi_x \xi_y - 2\eta_x \eta_y + (\xi_y^2 + \eta_y^2)]^m$$

can be expanded into $\frac{(m+3)!}{m!3!} = O(|\log \epsilon|^3)$ separable terms. Further more, since $r_x > r_y \geq 0$, $O(|\log \epsilon|)$ leading terms are needed in the following expansion to have an ϵ approximation,

$$(108) \quad (r_y - r_x)^{-l} = r_x^{-l} \left[1 - \sum_{m=1}^{\infty} \left(\frac{r_y}{r_x} \right)^m \right]^l.$$

Hence a separable ϵ -approximation of $e^{ik|\mathbf{x}-\mathbf{y}|}$ requires $O(|\log \epsilon|^6)$ terms. The last term we need to make a separable approximation is $\frac{1}{|\mathbf{x}-\mathbf{y}|}$.

$$\begin{aligned} \frac{1}{|\mathbf{x}-\mathbf{y}|} &= (r_y - r_x)^{-1} \left[1 + \frac{(\xi_x - \xi_y)^2 + (\eta_x - \eta_y)^2}{(r_x - r_y)^2} \right]^{-\frac{1}{2}} \\ &= r_x^{-l} \left[1 - \sum_{m=1}^{\infty} \left(\frac{r_y}{r_x} \right)^m \right]^l \left[1 + \sum_{m=1}^{\infty} \frac{(-1)^m (2(m+1))!}{2(2m+1)((m+1)!)^2 4^m} \frac{((\xi_x - \xi_y)^2 + (\eta_x - \eta_y)^2)^m}{(r_y - r_x)^{2m}} \right] \end{aligned}$$

Both summations in the above the formula can be truncated at $O(|\log \epsilon|)$ terms with ϵ error. Each term in the summation in the second bracket is similar to the term in (107) which can be approximated by at most $O(|\log \epsilon|^4)$ separable terms. So $\frac{1}{|\mathbf{x}-\mathbf{y}|}$ can also be approximated by $O(|\log \epsilon|^6)$ separable terms. Combine all these terms together we have $N_k^\epsilon \leq O(|\log \epsilon|^{12})$ for this setup.

3) Here we include two setups of X and Y that has have been proposed and used for developing fast algorithms.

(a) In [24, 6, 18] fast butterfly algorithms for computing highly oscillatory Fourier integral

operators and boundary integrals for Helmholtz equation were developed. The key idea is a dyadic decomposition of two domains A, B into tree structures T_A, T_B , from root to leaf, and a recursive pairing of $X \in T_A$ and $Y \in T_B$ such that the phase function or free space Green's function restricted on $X \times Y$ has good separability property or low rank approximation in discrete setting. The condition for pairing X and Y is that the product of the radius of X and Y is less than or equal to $\frac{1}{k}$. For this setup, the key observation is that one can construct a simple separable phase function, $\phi(\mathbf{x}_0, \mathbf{y}) + \phi(\mathbf{x}, \mathbf{y}_0) - \phi(\mathbf{x}_0, \mathbf{y}_0)$ that can approximate the original phase function uniformly well, i.e., $k|\phi(\mathbf{x}, \mathbf{y}) - \phi(\mathbf{x}_0, \mathbf{y}) - \phi(\mathbf{x}, \mathbf{y}_0) + \phi(\mathbf{x}_0, \mathbf{y}_0)|$ is uniformly bounded for all $k, \mathbf{x} \in X, \mathbf{y} \in Y$, where $\mathbf{x}_0, \mathbf{y}_0$ are the centers of X, Y respectively. Under an analytic function assumption for $\phi(\mathbf{x}, \mathbf{y})$, $k(\phi(\mathbf{x}, \mathbf{y}) - \phi(\mathbf{x}_0, \mathbf{y}) - \phi(\mathbf{x}, \mathbf{y}_0) + \phi(\mathbf{x}_0, \mathbf{y}_0))$ can be approximated by a Taylor expansion with $O(|\log \epsilon|)$ terms (see the proof in [6]).

(b) In [8] fast directional multilevel algorithms for oscillatory kernels were developed based on directional low rank property of free space Green's function of Helmholtz equation on two domains X_r and Y_r that satisfies directional parabolic separation condition: Y_r is a ball of radius r centered at a point c and X_r is the domain containing all points which are at a distance r^2 or greater from c and belong to a cone centered at c with spanning angle $\frac{1}{r}$. Here r can be thought in unit of wavelength $\lambda = \frac{2\pi}{k}$. For this setup, the phase function $|\mathbf{x} - \mathbf{y}| = \frac{k}{2\pi} |\lambda \mathbf{x} - \lambda \mathbf{y}|$ can be uniformly approximated by $\hat{\mathbf{x}} \cdot (\mathbf{x} - \mathbf{y}) = |\mathbf{x}| - \hat{\mathbf{x}} \cdot \mathbf{y} = |\mathbf{x}| - (\hat{\mathbf{x}} - \hat{\mathbf{l}}) \cdot \mathbf{y} - \hat{\mathbf{l}} \cdot \mathbf{y}$ which can be further approximated by $|\mathbf{x}| - \hat{\mathbf{l}} \cdot \mathbf{y}$ uniformly, where $\hat{\mathbf{x}} = \frac{\mathbf{x}}{|\mathbf{x}|}$ and $\hat{\mathbf{l}}$ is the direction the cone centered at. In another word, let $\phi(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$ and $\phi_1(\mathbf{x}) = |\mathbf{x}|, \phi_2(\mathbf{y}) = -\hat{\mathbf{l}} \cdot \mathbf{y}$, then $|\phi(\mathbf{x}, \mathbf{y}) - \phi_1(\mathbf{x}) - \phi_2(\mathbf{y})|$ is uniformly bounded and has an $O(\log |\epsilon|)$ term separable approximation with any $\epsilon > 0$ error.

4) Here we present a scaling argument to show a high separable asymptotic regime for two general disjoint domains, X and Y , as $k \rightarrow \infty$. Then we use this asymptotic argument to show that the condition for the setup of butterfly algorithm is tight. Let $r(X), r(Y)$ denote the diameters of X, Y respectively. Denote the distance between X and Y to be $dist(X, Y)$. Without loss of generality, let $r(Y) \leq r(X)$ and $r(X) = O(1)$. Actually the size of X is not restricted since it can always be scaled to $O(1)$ by scaling \mathbf{x} to $\frac{\mathbf{x}}{r(X)}$ for the Helmholtz equation (5). Assume $r(Y) \ll dist(X, Y)$. From the scaling argument in Remark 2.2, the rate of decorrelation of two Green's functions is rescaled to $|\langle \hat{G}(\cdot, \mathbf{y}_1), \hat{G}(\cdot, \mathbf{y}_2) \rangle| \lesssim \left(\frac{|\mathbf{y}_1 - \mathbf{y}_2|^2 k}{dist(X, Y)} \right)^{-\alpha}, \forall \mathbf{y}_1, \mathbf{y}_2 \in Y$ for some $\alpha > 0$. So if $r(Y) \lesssim \sqrt{\frac{dist(X, Y)}{k}}$, there is no fast decorrelation of two Green's function. On the other hand, if $r(Y) \gtrsim \sqrt{\frac{dist(X, Y)}{k^{1-\delta}}}, \forall \delta > 0$, one can put a grid with a grid size a little larger than $\sqrt{\frac{dist(X, Y)}{k}}$ in Y so that they become more and more decorrelated as $k \rightarrow \infty$. At the same time, the number of grid points in Y becomes larger and larger as $k \rightarrow \infty$. Use a similar argument as in Lemma 3.1 one can show a lower bound estimate for approximate separability which grows with some power of k . The above discussion can be further scaled with $dist(X, Y)$. For example, if $dist(X, Y) = O(k)$, the size of Y can be $O(1)$.

In a typical setup for butterfly algorithm [24, 6, 18], there are two domains A, B whose sizes are of $O(1)$ and $dist(A, B) = O(1)$. For example, A, B may be disjoint boundaries of scatterers in boundary integral methods. Dyadic decomposition of A, B gives tree structures T_A, T_B with $L = O(\log k)$ levels, where the roots are A, B and the leaf nodes at level L are of size $O(k^{-1})$. Then the interaction between A and B through a highly oscillatory kernel function, e.g., the Green's function for Helmholtz equation or a Fourier integral operator, is computed based on a recursive pairing of nodes $X, A \supseteq X \in T_A$, and $Y, B \supseteq Y \in T_B$, such that the level of $X, l(X)$, and the level of $Y, l(Y)$, satisfy $l(X) + l(Y) = L$. The key observation is that the interaction between X and Y has a low rank approximation. It was shown in [6] that the low rank approximation is guaranteed if $r(X)r(Y) = O(k^{-1})$ which is implied by condition $l(X) + l(Y) = L$ (plus some analyticity condition on the phase function). It can be easily seen that high separability for the setup for butterfly algorithms falls into the asymptotic regime discussed above. The requirement of analyticity of the phase function is equivalent to requiring $dist(X, Y) > c > 0$ for some fixed c and the condition $r(X)r(Y) = O(k^{-1})$ implies the smaller domain Y (or X) satisfies $r(Y)$ (or $r(X)$) $\leq O(k^{-\frac{1}{2}})$. In particular, this condition is barely satisfied when $r(X) \sim r(Y)$ or $l(X) = l(Y)$. If $r(A)r(B) = O(k^\delta)$ for any $\delta > 0$ or $dist(X, Y) \rightarrow 0$ as $k \rightarrow \infty$, then the condition is violated and low rank approximation is not valid as discussed above.

4.3. Approximate separability with boundary condition. So far we have shown approximate separability estimates for Green's function of Helmholtz equation in high frequency limit either in the whole space. Here we present an example when boundary and reflection are present. First we study approximate separability of Green's function for Helmholtz equation in half space with homogeneous Dirichlet boundary condition as illustrated in Figure 10(a). The Green's function for the upper space, denoted by $G_1(\mathbf{x}, \mathbf{y})$, can be explicitly constructed from free space Green's function, denoted by $G_0(\mathbf{x}, \mathbf{y})$,

$$G_1(\mathbf{x}, \mathbf{y}) = G_0(\mathbf{x}, \mathbf{y}) - G_0(\mathbf{x}, \bar{\mathbf{y}}),$$

where $\bar{\mathbf{y}}$ is the mirror image of \mathbf{y} with respect to the boundary. Decorrelation of G_1 can be deduced from that of G_0 . Given two disjoint compact domains X and Y in the upper half space and two points $\mathbf{y}_1, \mathbf{y}_2 \in Y$, if the line $l_{\mathbf{y}_1}^{\mathbf{y}_2}$ connecting these two points intersects with X , or none of the the lines $l_{\mathbf{y}_1}^{\bar{\mathbf{y}}_2}, l_{\bar{\mathbf{y}}_1}^{\mathbf{y}_2}, l_{\bar{\mathbf{y}}_1}^{\bar{\mathbf{y}}_2}$ intersects with X , we have

(109)

$$|\langle G_1(\cdot, \mathbf{y}_1), G(\cdot, \mathbf{y}_2) \rangle|$$

$$= |\langle G_0(\mathbf{x}, \mathbf{y}_1), G_0(\mathbf{x}, \mathbf{y}_2) \rangle - \langle G_0(\mathbf{x}, \mathbf{y}_1), G_0(\mathbf{x}, \bar{\mathbf{y}}_2) \rangle - \langle G_0(\mathbf{x}, \bar{\mathbf{y}}_1), G_0(\mathbf{x}, \mathbf{y}_2) \rangle + \langle G_0(\mathbf{x}, \bar{\mathbf{y}}_1), G_0(\mathbf{x}, \bar{\mathbf{y}}_2) \rangle|$$

$$\lesssim |\langle G_0(\cdot, \mathbf{y}_1), G_0(\cdot, \mathbf{y}_2) \rangle|$$

since $|\mathbf{y}_1 - \mathbf{y}_2| \leq \min(|\mathbf{y}_1 - \bar{\mathbf{y}}_2|, |\bar{\mathbf{y}}_1 - \mathbf{y}_2|, |\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2|)$ and $dist(\mathbf{y}_j, X) < dist(\bar{\mathbf{y}}_j, X), j = 1, 2$. Otherwise, decorrelation of $G_1(\mathbf{x}, \mathbf{y}_1)$ and $G_1(\mathbf{x}, \mathbf{y}_2)$ may be slower than that of $G_0(\mathbf{x}, \mathbf{y}_1)$ and $G_0(\mathbf{x}, \mathbf{y}_2)$. In general, for two disjoint compact domains X and Y , there is a cone with solid angle at each point $\mathbf{y} \in Y$ with \mathbf{y} as the vertex such that a line connecting \mathbf{y}

and any other point in the cone will intersect X . Correlations between Green's functions of these points are the leading terms in the estimate of $tr(A^T A)$ as discussed in Remark 3.3. Hence the relation (109) implies that the proof and result in Lemma 3.1 and the lower bound estimates in Theorem 3.1 and Theorem 3.3 for X and Y without boundary hold for this case too. It can also be easily seen that separable approximation of Green's function $G_1(\mathbf{x}, \mathbf{y})$ for $\mathbf{x} \in X, \mathbf{y} \in Y$ can be obtained by combining separable approximation of $G_0(\mathbf{x}, \mathbf{y})$ for $\mathbf{x} \in X, \mathbf{y} \in Y$ and that of $G_0(\mathbf{x}, \bar{\mathbf{y}})$ for $\mathbf{x} \in X, \bar{\mathbf{y}} \in \bar{Y}$, where \bar{Y} is the mirror image of Y with respect to the boundary. So the upper bound estimates in Theorem 3.2 and Theorem 3.4 hold too here. Although the asymptotic formulas of the lower bound and upper bound estimate for approximate separability of Green's function in this case is similar to those in free space, the number of terms, N^ϵ , needed for a separable approximation for the Green's function with a given error tolerance $\epsilon > 0$ defined in (2) will increase due to reflection at the boundary in general. It can be argued by the following. The Green's function $G_1(\mathbf{x}, \mathbf{y})$ viewed as a family function in X parametrized by $\mathbf{y} \in Y$ is composed of a family of free space Green's function $G_0(\mathbf{x}, \mathbf{y})$ with \mathbf{y} in a larger domain $Y \cup \bar{Y}$. Or geometrically, instead of there is one ray connecting any two points $\mathbf{x} \in X$ and $\mathbf{y} \in Y$ in free space, now there are two rays with the presence of a reflection boundary. The phase function becomes more complicated and needs more terms in a separable approximation. As the distance from the boundary to X and Y becomes larger and larger, the effect from the boundary becomes less and less. In general, one can expect that presence of reflection and scattering can cause the increase of N^ϵ , the number of terms needed for a separable approximation for the Green's function defined in (2).

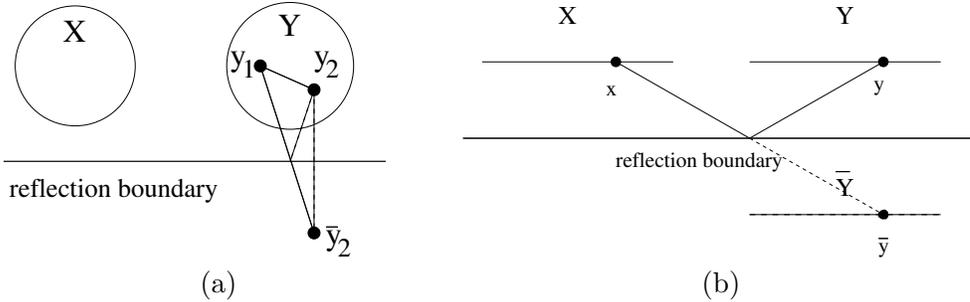


FIGURE 10. Half space setups with a reflection boundary.

Here, we give a more explicit study of the boundary effect on a particular setup as depicted in Figure 10(b), where X and Y are two disjoint collinear line segments parallel to the boundary. Without the reflection boundary, $G_0(\mathbf{x}, \mathbf{y}), \mathbf{x} \in X, \mathbf{y} \in Y$ is highly separable which is discussed as the first special case in Section 4.2. However, if the distance from the boundary to X, Y is small compared to the wavelength and the separation distance between X and Y , the special setup of two thin cylinders in Section 4.2 or study in [21, 9] shows that $G_1(\mathbf{x}, \mathbf{y}), \mathbf{x} \in X, \mathbf{y} \in Y$ is still highly separable. This is the key observation for the low rank approximation used in the sweeping preconditioner for Helmholtz equation in 2D

in [9]. However, this highly separable property does not hold if the boundary is not close enough. Also high separability does not hold in an analogous setup in 3D. As discussed in case 2) in Section 4.1, even for two general disjoint compact co-planar 2D regions X, Y without the boundary the number of terms needed for a separable approximation is at least $O(k)$ as $k \rightarrow \infty$.

We can extend the above study further to Helmholtz equation in a waveguide, see Figure 11. The Green's function in the waveguide with a source at \mathbf{y} is an infinite sum of homogeneous Green's functions with sources that are mirror images of each other with respect the top and bottom boundary, i.e.,

$$(110) \quad G_2(\mathbf{x}, \mathbf{y}) = G_0(\mathbf{x}, \mathbf{y}) + \sum_{m=1}^{\infty} (-1)^m G_0(\mathbf{x}, \mathbf{y}_m^+) + \sum_{m=1}^{\infty} (-1)^m G_0(\mathbf{x}, \mathbf{y}_m^-),$$

where \mathbf{y}_1^+ is the mirror image of \mathbf{y} with respect to the top boundary, $\mathbf{y}_m^+, m = 2, 3, \dots$ is the mirror image of \mathbf{y}_{m-1}^- with respect to the top boundary. Similarly, \mathbf{y}_1^- is the mirror image of \mathbf{y} with respect to the bottom boundary and $\mathbf{y}_m^-, m = 2, 3, \dots$ is the mirror image of \mathbf{y}_{m-1}^+ with respect to the bottom boundary. Due to the two reflection boundaries, the Green's function $G_2(\mathbf{x}, \mathbf{y}), \mathbf{y} \in Y$ in the waveguide involves free space Green's function $G_0(\mathbf{x}, \mathbf{y})$ with \mathbf{y} belonging an infinite union of Y and its images and their images with respect to the top and bottom boundary respectively. In another word, there is one direct ray in free space and infinite many broken rays due to the waveguide setup connecting any two points $\mathbf{x} \in X$ and $\mathbf{y} \in Y$. Of course the decay factor $\frac{1}{|\mathbf{x}-\mathbf{y}|}$ will play a role here too. For the setup of two collinear line segments X and Y shown in Figure 11(b), highly separability does not hold even if both boundaries are very close.

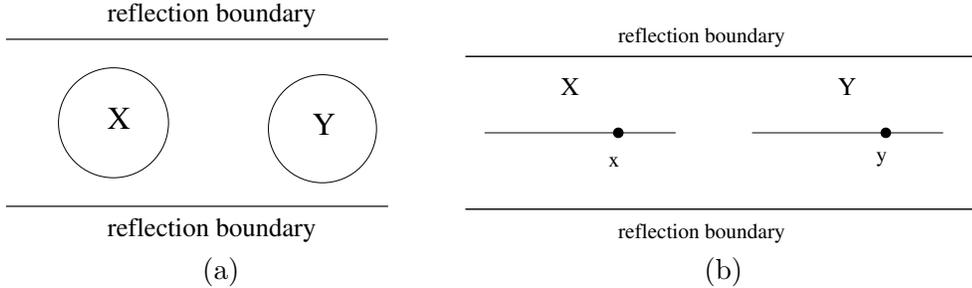


FIGURE 11. Helmholtz equation in a waveguide.

5. NUMERICAL TESTS

Here we show a few numerical tests related to our analysis in previous sections. In all our numerical tests free space Green's functions are used. Our computational grid size h resolves the wavelength $\lambda = 2\pi/k$, $h = \lambda/15$ in 2D and $h = \lambda/13$ in 3D.

Example 1 We test the decorrelation of two Green's functions, $\left| \langle \hat{G}_1(\cdot, \mathbf{y}_1), \hat{G}_1(\cdot, \mathbf{y}_2) \rangle_X \right|$, where X is a compact domain. In this test, we show results for k starting from 50 with an increment of 5.

Figure 12 shows the decorrelation of two Green's functions in 2D. Here the domain X is a disc centered at $(0, 0)$ with radius 0.4. In Figure 12(a) the two points $\mathbf{y}_1, \mathbf{y}_2$ lie on x-axis. So the line through $\mathbf{y}_1, \mathbf{y}_2$ intersects X and hence there is a line of stationary points for $\langle \hat{G}_1(\cdot, \mathbf{y}_1), \hat{G}_1(\cdot, \mathbf{y}_2) \rangle_X$. In Figure 12(b), the line through $\mathbf{y}_1, \mathbf{y}_2$ is parallel to y-axis which does not intersect X . Hence there is no stationary point for $\langle \hat{G}_1(\cdot, \mathbf{y}_1), \hat{G}_1(\cdot, \mathbf{y}_2) \rangle_X$. As shown in the proof of Theorem 2.1, two Green's functions in this case decorrelate much faster than the afore mentioned case.

Also two Green's functions decorrelate faster when $|\mathbf{y}_1 - \mathbf{y}_2|$ becomes larger in general as shown in Theorem 2.1,. It can be seen that as $\mathbf{y}_1, \mathbf{y}_2$ are further away from X , two Green's functions become more correlated due to scaling of the gradient and the Hessian of the phase function in term of the distance from $\mathbf{y}_1, \mathbf{y}_2$ to X as explained in Remark 2.2.

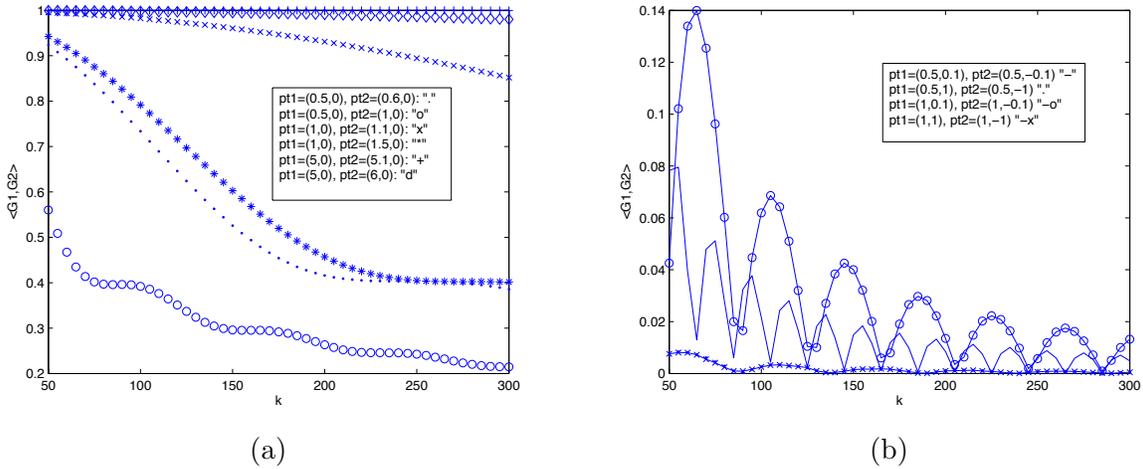


FIGURE 12. Decorrelation of two Green's functions in 2D.

Figure 13 shows a test with homogeneous Dirichlet boundary condition. The boundary is an infinite line that is parallel to x-axis located at $y = -d$. In Figure 13(a), X is again a disc as before. One can see that the boundary condition does affect the correlation between two Green's functions. However, the asymptotic behavior is similar to the one without boundary condition as $k \rightarrow \infty$. In Figure 13(b), X is a line segment on x-axis between $[-0.5, 0.5]$, which is co-linear with points $\mathbf{y}_1, \mathbf{y}_2$. With no boundary conditions, the two Green's functions are just different by a constant phase and hence fully correlated. When the boundary is present and is not too close or too far away from the two points and X , we do see more decorrelations due to the boundary boundary. When the boundary

is either very close or the boundary is very far, the two Green's functions become highly correlated as explained in Section 4.3.

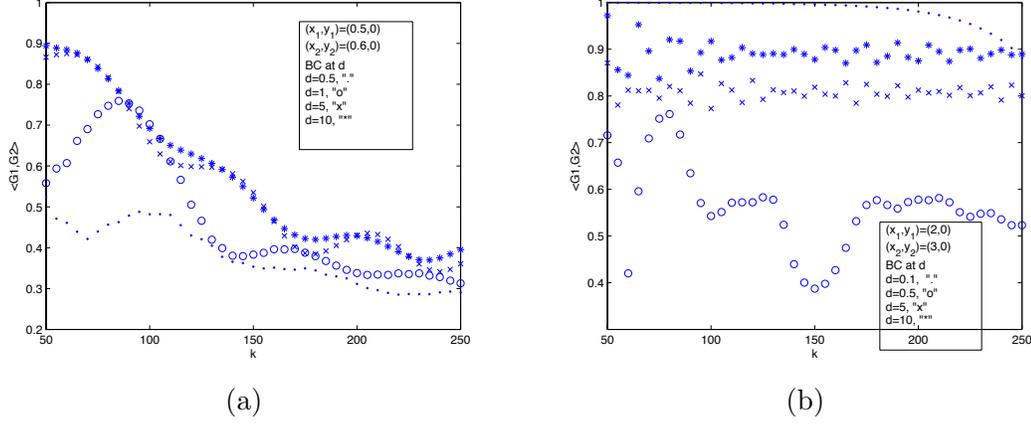


FIGURE 13. Decorrelation of two Green's functions in 2D with a boundary.

Figure 14 shows corresponding tests in 3D. The behavior in 3D is similar to those in 2D. The domain X is a ball centered at the origin with radius 0.4. Figure 14(a) shows a test where the two points $\mathbf{y}_1, \mathbf{y}_2$ lie on x -axis, hence the line through $\mathbf{y}_1, \mathbf{y}_2$ intersects X and two Green's functions decorrelate relatively slow. Figure 14(b) shows the effect of a Dirichlet boundary condition.

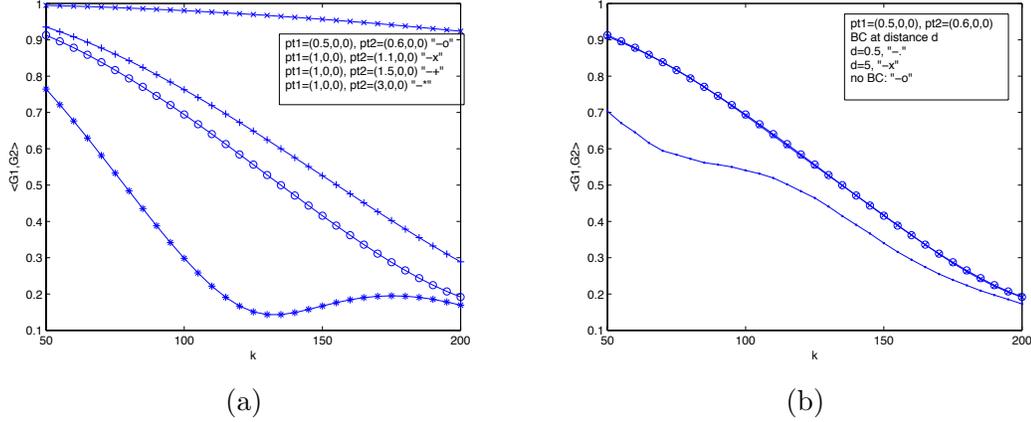


FIGURE 14. Decorrelation of two Green's functions in 3D.

Example 2 Here we present singular value decomposition (SVD) pattern for matrix $G(\mathbf{x}_i, \mathbf{y}_j)$, where $\mathbf{x}_i, \mathbf{y}_j$ are uniformly distributed points in X and Y respectively with a

grid size that resolves the wavelength. 3D free space Green's function are used in the test. In these tests, we show results for k ranging from 10 to 150 with an increment of 5. In the following figures, solid lines show the number of leading singular values that is larger than ϵ vs. the wave number k while dotted lines shows N_k^ϵ defined by Definition 3.2. In all tests we use three threshold values $\epsilon = 10^{-2}, 10^{-4}, 10^{-6}$.

Figure 15(a) is the result for $X = \{(x, y, z) | x \in [-0.5, 0.5], y = 0, z = 0\}$ and $Y = \{(x, y, z) | x = 0, y = 0.2, z \in [-0.5, 0.5]\}$. Figure 15(b) is the result for X , a sphere centered at origin with radius 0.5, and $Y = \{(x, y, z) | x \in [0.6, 1.6], y = 0.6, z = 0\}$. Linear growth as predicted by our analysis is clearly seen for both cases.

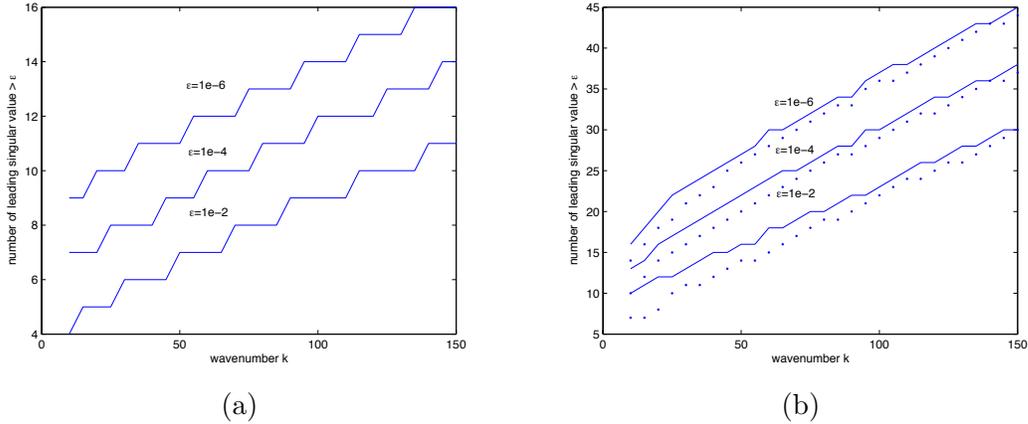


FIGURE 15. Leading singular values vs. wave number for $\dim(Y) = 1$.

Figure 16 gives an example of two square domains X, Y of length 0.4 for each side with different relative position corresponding to the three setups as discussed in Section 4.1 case 2) and demonstrated in Figure 8. In Figure 16(a), X, Y are side by side and coplanar with minimum distance 0.1 between them. In Figure 16(b), X, Y are side by side and orthogonal to each other with minimum distance 0.1 between them. In Figure 16(c), X, Y are parallel with distance 0.3 between them. As analyzed in Section 4.1, case (a) has the slowest rate of decorrelation between two Green's function and hence also the slowest growth of the number of leading singular values and N_k^ϵ while case (c) has the fastest rate of decorrelation between two Green's function and hence also the fastest growth of the number of leading singular values and N_k^ϵ among the three cases as $k \rightarrow \infty$. At least linear growth is observed in case (a). In both case (b) and (c) a quadratic growth, as predicted by sharp lower bound and upper bound estimates and analyzed in Section 4.1 case 2, is observed.

Figure 17(a) shows an example of two spheres of radius 0.2 with a separation distance of 0.2 between the two spheres. At the maximum wave number $k = 150$, 20,000 points are randomly distributed on the surface of each sphere with the number of points being proportional to wave number for wave numbers in between. Again one sees that the number

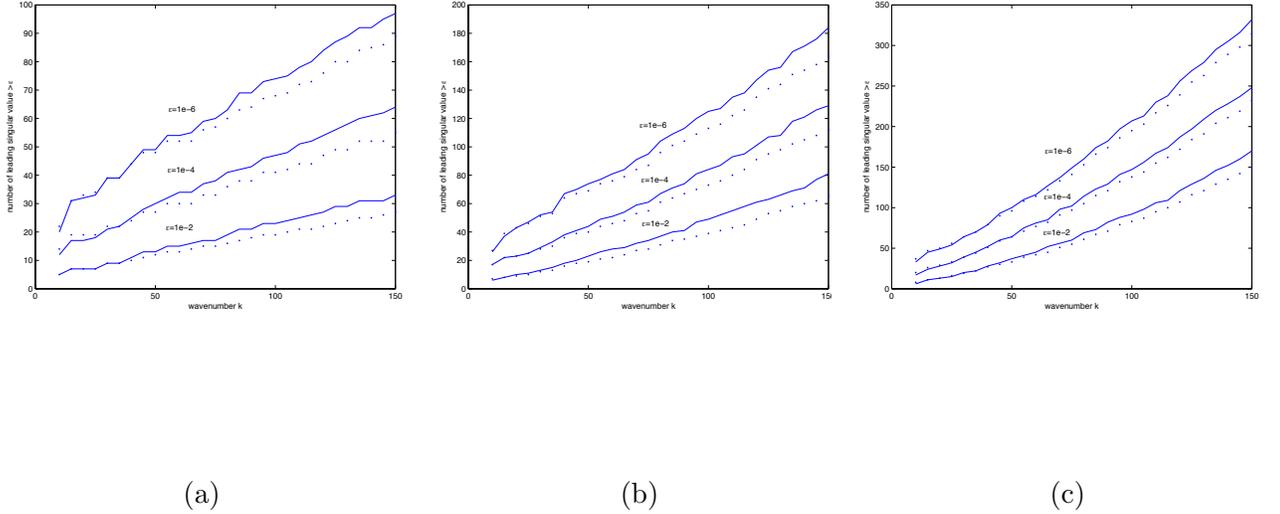


FIGURE 16. Leading singular values vs. wave number for two squares in 3D.

of leading singular values grows quadratically as analyzed in Section 4.1 case 2). Figure 17(b) shows an example of two thin cylinders as illustrated in Figure 9(b). The radius of each cylinder is half wavelength and the length of each cylinder is 0.4. The separation distance between the two cylinders is one wavelength. The SVD pattern agrees with our high separability result in Section 4.2 nicely.

6. CONCLUSION

In this work, approximate separability of Green's functions of Helmholtz equations in the high frequency limit, which has direct implication for low rank approximation for the corresponding discretized linear system, is studied in details. By characterizing the decorrelation rate of two Green's function due to fast oscillations in various situations and showing a tight dimension estimate for the approximation of a set of almost orthogonal vectors, we prove an explicit and sharp asymptotic formula for the lower bound for the number of terms needed for a separable approximation of Green's function in the high frequency limit. It gives a rigorous mathematical argument for the complexity and difficulty for solving high frequency Helmholtz equation numerically. Application to setups that are commonly used in practice is presented. Numerical tests show full agreement with the analysis. Development of efficient numerical algorithms based on this study will be investigated in the future.

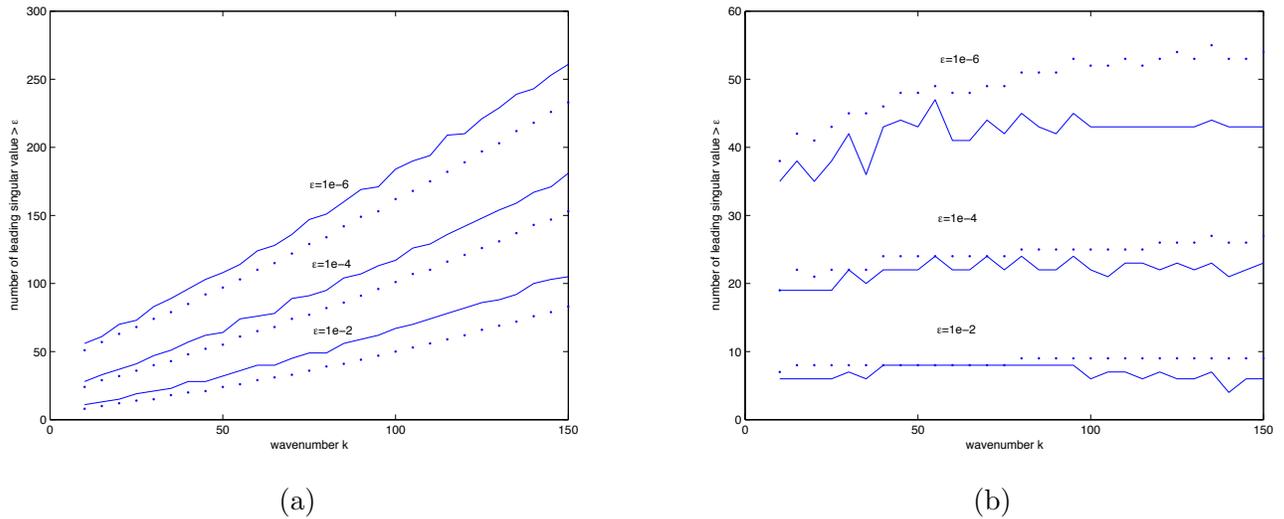


FIGURE 17. SVD pattern for two spheres and two thin cylinders.

Acknowledgement The authors would like to thank Lawrence C. Evans and Terence Tao for helpful discussions.

REFERENCES

- [1] Noga Alon, *Problems and results in extremal combinatorics, I*, Discrete Math. 273, 31-53, 2003.
- [2] M. Bebendorf and W. Hackbusch, *Existence of \mathcal{H} -matrix approximants to the inverse FE-matrix of elliptic operator with L^∞ -coefficients*, Numer. Math. 95 (2003), no. 1, 1–28.
- [3] S. Börm, *Approximation of solution operators of elliptic partial differential equations by \mathcal{H} - and \mathcal{H}^2 -matrices*, Numer. Math. 115 (2010), no. 2, 165-193.
- [4] O. M. Bucci and G. Franceschetti, *On the spatial bandwidth of scattered fields*, IEEE Transactions on Antennas and Propagation, Vol 35(12), pp1445-1455, 1987.
- [5] O. M. Bucci and G. Franceschetti, *On the degrees of freedom of scattered fields*, IEEE Transactions on Antennas and Propagation, Vol 37(7), pp 918-926, 1989.
- [6] E. Candes, L. Demanet, and L. Ying, *A fast butterfly algorithm for the computation of Fourier integral operators*, Multiscale Model. Simul. 7 (2009), no. 4, 1727–1750.
- [7] S. Chandrasekaran, P. Dewilde, M. Gu and N. Somasunderam, *On the numerical rank of the off-diagonal blocks of Schur complements of discretized elliptic PDEs*, SIAM J. Matrix Anal. Appl. 31 (2010), no. 5, 2261–2290.
- [8] B. Engquist and L. Ying, *Fast directional multilevel algorithms for oscillatory kernels*, SIAM Journal on Scientific Computing 29 (2007).
- [9] B. Engquist and L. Ying, *Sweeping preconditioner for the Helmholtz equation: Hierarchical matrix representation*, Communications in Pure and Applied Mathematics 64, 2011.
- [10] O. Ernst and M.J. Gander, *Why it is Difficult to Solve Helmholtz Problems with Classical Iterative Methods*, Numerical Analysis of Multiscale Problems, I. Graham, T. Hou, O. Lakkis and R. Scheichl, Editors, pp. 325-363, Springer Verlag, 2012.
- [11] L. Greengard and V. Rokhlin, *A fast algorithm for particle simulations*, J. Comput. Phys., vol. 73, pp. 325–348, 1987.

- [12] L. Greengard, *The rapid evaluation of potential fields in particle systems*, ACM Distinguished Dissertations, MIT Press, Cambridge, MA, 1988.
- [13] M. Grüter and K. -O Wildman, *The Green function for uniformly elliptic equations*, Manuscripta Math. 37 (1982), no. 3, 303–342.
- [14] M. Gu and S. Eisenstat *Efficient algorithms for computing a strong rank-revealing QR factorization*, SIAM J. Sci. Comput., 17 (1996), pp. 848–869.
- [15] K. Ho and L. Ying, *Hierarchical interpolative factorization for elliptic operators: differential equations*, preprint.
- [16] Lars Hörmander, *The Analysis of Linear Partial Differential Operators, I*, Springer-Verlag, 1983.
- [17] A. N. Kolmogorov, *Über die beste Annäherung von Funktionen einer Funktionklasse*, Ann. Math., 37 (1936), 107–111.
- [18] S. Luo, J. Qian and R. Burridge, *Fast Huygens sweeping methods for Helmholtz equations in inhomogeneous media in the high frequency regime*, J. Comput. Phys. 270, 378–401, 2014.
- [19] W. B. Johnson and J. Lindenstrauss, *Extensions of Lipschitz mappings into a Hilbert space*, Contemp. Math. 26, AMS, Providence, RI, 189–206, 1984.
- [20] W. Littman, G. Stampacchia and H. F. Weinberger, *Regular points for elliptic equations with discontinuous coefficients*, Annali della Scuola Normale Superiore di Pisa - Classe di Scienze, Sér. 3, 17 no. 1-2, p. 43–77, 1963 .
- [21] P.-G. Martinsson and V. A. Rokhlin, *A fast direct solver for scattering problems involving elongated structures*, J. Comput. Phys. 221, no. 1, 288–302, 2007.
- [22] P.G. Martinsson, V. Rokhlin, and M. Tygert, *A randomized algorithm for the approximation of matrices*, Applied and Computational Harmonic Analysis, 30(1), pp. 47–68, 2011.
- [23] P.-G. Martinsson, *A fast direct solver for a class of elliptic partial differential equations*, J. Sci. Comput. 38 (2009), no. 3, 316–330.
- [24] E. Michielssen and A. Boag, *A multilevel matrix decomposition algorithm for analyzing scattering from large structures*, IEEE Trans. Antennas and Propagation, 44 (1996), no. 8, pp. 1086–1093.
- [25] M. O’Neil, F. Woolfe, and V. Rokhlin, *An algorithm for the rapid evaluation of special function transforms*, Applied and Computational Harmonic Analysis, 28 (2010), no. 2, pp. 203–226.
- [26] V. Rokhlin, *Rapid solution of integral equations of scattering theory in two dimensions*, J. Comput. Phys., vol. 86 (1990), pp. 414–439.
- [27] V. Rokhlin, *Diagonal forms of translation operators for the Helmholtz equation in three dimensions*, Appl. Comput. Harmon. Anal. 1 (1993), no. 1, 82–93.
- [28] J. Roe, *Elliptic operators, topology, and asymptotic methods*, Chapman & Hall/CRC Research Notes in Mathematics Series, 2nd edition, 1999.
- [29] P. G. Schmitz and L. Ying, *A fast direct solver for elliptic problems on general meshes in 2D*, J. Comput. Phys. 231 (2012), no. 4, 1314–1338.
- [30] András Vasy, *Lecture notes at Stanford University*.
- [31] J. Xia, S. Chandrasekaran, M. Gu and X. S. Li, *Superfast multifrontal method for large structured linear systems of equations*, SIAM J. Matrix Anal. Appl. 31(2009), no. 3, 1382–1411.
- [32] J. Xia, *Efficient structured multifrontal factorization for general large sparse matrices*, SIAM J. Sci. Comput. 35 (2013), no. 2, A832–A860.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF TEXAS AND ICES, UNIVERSITY OF TEXAS AT AUSTIN,
1 UNIVERSITY STATION C1200, AUSTIN, TX, 78712.

E-mail address: `engquist@ices.utexas.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, IRVINE, CA 92697.

E-mail address: `zhao@math.uci.edu`