Supplementary Material of ICCV2013 Paper #1965

Xu Wang, Stefan Atev, John Wright and Gilad Lerman

1. Alternative Formulation: Hashing with points

An alternative formulation. The probabilities p_1 and p_2 are two important parameters. The algorithm performs better if the gap between them is large. In this section, we propose a modified hashing family for which we can write an analytic expression for p_1 and p_2.

Definition 1.1. *For each point $\boldsymbol{x} \in \mathbb{R}^D$ and $\eta \in \mathbb{R}^+$, we associate a function $h_{\boldsymbol{x},\eta}$ on $\mathrm{G}(D,d)$:*

$$h_{\boldsymbol{x},\eta} : \mathrm{G}(D,d) \longrightarrow \{0,1\}$$

such that for any $L \in \mathrm{G}(D,d)$

$$\begin{aligned} h_{\boldsymbol{x},\eta}(L) &= 0, \qquad \textit{if} \quad \mathrm{dist}_{\mathrm{G}}(\boldsymbol{x}, L) > \eta; \\ h_{\boldsymbol{x},\eta}(L) &= 1, \qquad \textit{if} \quad \mathrm{dist}_{\mathrm{G}}(\boldsymbol{x}, L) \leq \eta, \end{aligned} \tag{1}$$

where $\mathrm{dist}(\boldsymbol{x}, L)$ is the Euclidean distance between $\boldsymbol{x}$ and L.

Let $\mathcal{H}$ be this family of functions. Let μ be the normal distribution on $\mathbb{R}^D$ with mean $\mathbf{0}$ and variance $\mathbf{1}$ in each direction. Its density function is $e^{-||\boldsymbol{x}||_2^2/2}/(2\pi)^{D/2}$. By identifying $\mathcal{H}$ with R^D, we get a measure μ on this hashing family. Following similar arguments, it is easy to see that $(\mathcal{H}, \mu)$ is *locality sensitive hashing* family on $\mathrm{G}(D,d)$.

The maximal value of $\mu(\boldsymbol{x} \in \mathbb{R}^D | h_{\boldsymbol{x}}(L_1) = h_{\boldsymbol{x}}(L_2), \ \mathrm{dist}_{\mathrm{G}}(L_1, L_2) = R)$ as a function of $L_1, L_2 \in \mathrm{G}(D,d)$ is achieved when $\theta_1(L_1, L_2) = R$ and $\theta_i(L_1, L_2) = 0$ for $i = 2, ..., d$. The minimal value of $\mu(\boldsymbol{x} \in \mathbb{R}^D | h_{\boldsymbol{x}}(L_1) = h_{\boldsymbol{x}}(L_2), \ \mathrm{dist}_{\mathrm{G}}(L_1, L_2) = R)$ is achieved when $\theta_i(L_1, L_2) = R/\sqrt{d}$ for $i = 1, ..., d$.
Therefore,

$$\begin{aligned} p_1 = &\min_{\mathrm{dist}_{\mathrm{G}}(L_1,L_2) \leq R} \mu(\boldsymbol{x} \in \mathbb{R}^D | h_{\boldsymbol{x}}(L_1) = h_{\boldsymbol{x}}(L_2)) \\ &= 1 - 2\mu(\boldsymbol{x} \in \mathbb{R}^D | h_{\boldsymbol{x}}(L_1) = 1, \theta_i(L_1, L_2) = R/\sqrt{d}, \\ &\quad \forall 1 \leq i \leq d) + 2\mu(\boldsymbol{x} \in \mathbb{R}^D | h_{\boldsymbol{x}}(L_1) = h_{\boldsymbol{x}}(L_2) = 1, \\ &\quad \theta_i(L_1, L_2) = R/\sqrt{d}, \forall 1 \leq i \leq d) \end{aligned} \tag{2}$$

To compute the probability in the RHS of (3), we note that for the underlying Gaussian measure $\mathrm{dist}^2(\boldsymbol{x}, L)$ has chi-squared distribution $\mathcal{X}^2_{D-d}$ with $D-d$ degree of freedom. Therefore, the first probability in the RHS of (3) is $F(\eta^2; D-d)$, that is the cdf function at η^2 of the $\mathcal{X}^2_{D-d}$. To compute the second probability, we distinguish between the projection of X onto the orthogonal complement of $L_1 \oplus L_2$, whose distance from L_1 and L_2 distributes like $\mathcal{X}^2_{D-2d}$ (we denote the pdf of this distribution by $f(t; D-2d)$) and the projection onto $L_1 \oplus L_2$. For elements in the latter projection, we assign coordinates $(x_1, y_1, ..., x_d, y_d)$ so that the projection onto L_1 is $(x_1, 0, x_2, 0, ...)$ and its distance from L_2 (obtained by dot product with the normals $\{(\sin\theta_i, -\cos\theta_i)\}_{i=1}^d$ of L_2 in $L_1 \oplus L_2$) is $(\sum_{i=1}^d (x_i \sin\theta_i - y_i \cos\theta_i)^2)^{1/2}$. Using this observation, we obtain that

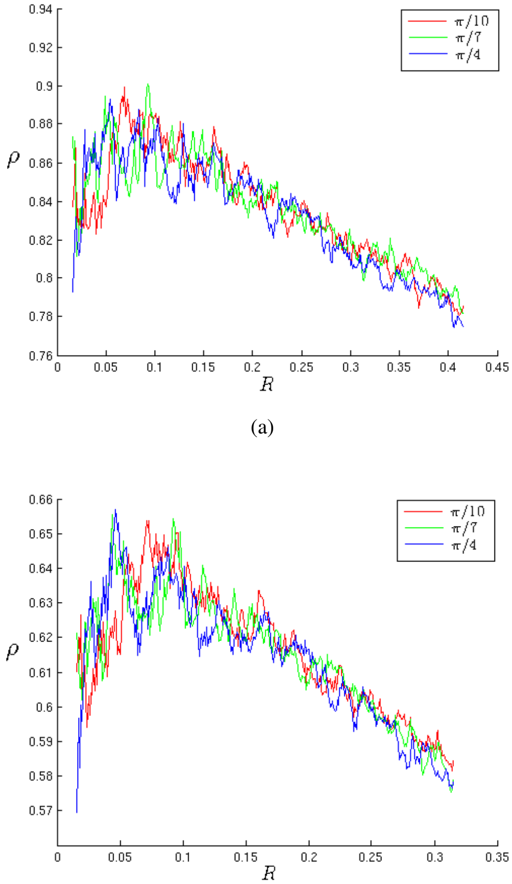
$$\begin{aligned} p_1 = &\ 1 - 2F(\eta^2; D-d) \\ &+ 2\int_0^{\eta^2} f(t; D-2d)dt \int_{\sum_{i=1}^d y_i^2 \leq \eta^2 - t} \Pi_{i=1}^d \frac{1}{\sqrt{2\pi}} e^{-\frac{y_i^2}{2}} dy_i \\ &\times \int_{\sum_{i=1}^d (x_i \sin\frac{R}{\sqrt{d}} - y_i \cos\frac{R}{\sqrt{d}})^2 \leq \eta^2 - t} \Pi_{i=1}^d \frac{1}{\sqrt{2\pi}} e^{-\frac{x_i^2}{2}} dx_i \end{aligned} \tag{3}$$

Similarly,

$$\begin{aligned} p_2 = &\max_{\mathrm{dist}_{\mathrm{G}}(L_1,L_2) \geq cR} \mu(\boldsymbol{x} \in \mathbb{R}^D | h_{\boldsymbol{x}}(L_1) = h_{\boldsymbol{x}}(L_2)) \\ &= 1 - 2F(\eta^2; D-d) + \frac{1}{\pi}\int_0^{\eta^2} f(t; D-d-1)dt \times \\ &\int_{y^2 \leq \eta^2 - t} e^{-\frac{y^2}{2}} dy \int_{(x\sin(cR) - y\cos(cR))^2 \leq \eta^2 - t} e^{-\frac{x^2}{2}} dx \end{aligned} \tag{4}$$

2. Numerical Investigation of parameters

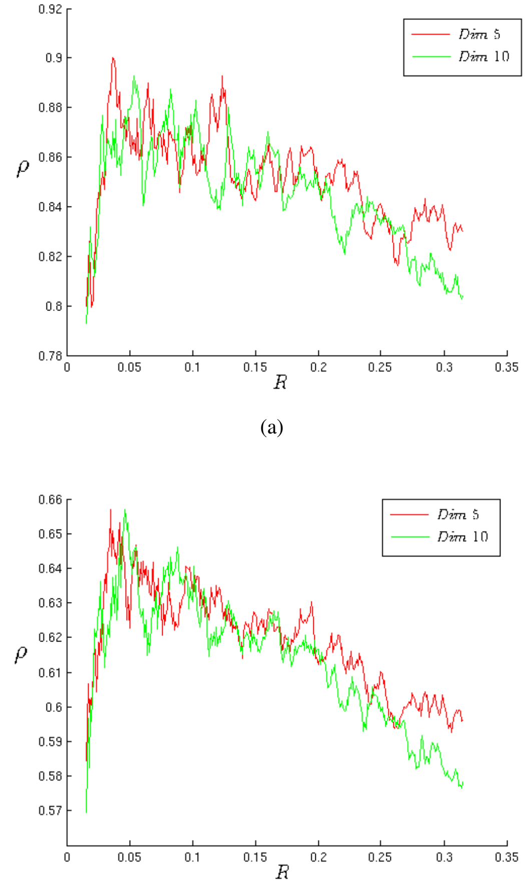
The sublinearity exponent ρ of GLH algorithm (with N^ρ sublinear time) depends on the probability p_1 and p_2. It is desirable to know the dependence of p_1 and p_2 (or alternatively ρ) on the underlying parameters c, R, θ_0, D and d (also η if using alternative formulation). While it is hard to determine this in theory for the general case, we apply Monte-Carlo integration to estimate p_1 and p_2 in various instances and thus try to infer their dependence on the underlying parameters in these cases. In two paragraphs below, we consider the original formulation and the alternative formulation of GLH respectively.



(a)

(b)

Figure 1. Comparing different sizes of neighborhoods



(a)

(b)

Figure 2. Comparing different dimensions of ambient spaces

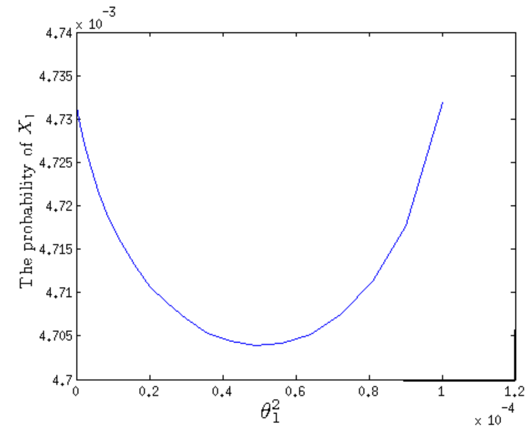
Original GLH: Hashing with lines In this paragraph, the goal is to show how p_1 and p_2 (and ρ) depends on various parameters. By definition of p_1 and p_2, they are volumns of particular areas on the sphere with uniform measure. Monte-Carlo integration is chosen to estimate these volumns. In the following experiments, we pick 100,000 random points uniformly from the sphere and check the percentages of points which are in the areas corresponding to p_1 and p_2. These percentages are taken as an estimation of p_1 and p_2.

Here, both the query points and the database are from $\mathrm{G}(10, 1)$. We demonstrate the dependence of ρ on R for some fixed values of c and θ_0, In Figure 1, ρ is plotted against R when $\theta_0 = \pi/10, \pi/7, \pi/4$ respectively, where Figure 1(a) $c = 1.1$ and Figure 1(b) $c = 1.5$. In this case, different values of θ_0 result in similar exponents.

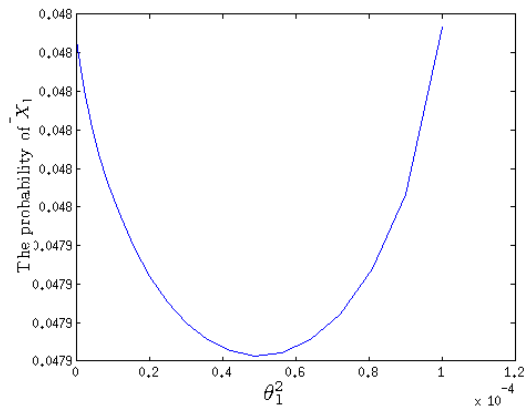
Next, we demonstrate the dependence of ρ on D by observing both $D = 5$ and $D = 10$ and maintaining $d = 1$. We also fix $\theta_0 = \pi/4$. The results are shown in Figure 2 (in Figure 2(a) $c = 1.5$ and in Figure 2(b) $c = 1.1$). The red plot is for $D = 5$ and the blue plot for $D = 10$. The sublinearity exponent ρ is fairly stable as the ambient dimension increases from 5 to 10.

Alternative GLH: Hashing with points We shall use the alternative formulation of GLH defined in 1.1. The formulae to calculate p_1 and p_2 are given by 3 and 4. Denote $\mathrm{Pmin}_R = \min_{\mathrm{dist}_{\mathrm{G}}(L_1, L_2) = R} \mu(\boldsymbol{x} \in \mathbb{R}^D | h_{\boldsymbol{x}}(L_1) = h_{\boldsymbol{x}}(L_2))$ and $\mathrm{Pmax}_R = \max_{\mathrm{dist}_{\mathrm{G}}(L_1, L_2) = R} \mu(\boldsymbol{x} \in \mathbb{R}^D | h_{\boldsymbol{x}}(L_1) = h_{\boldsymbol{x}}(L_2))$.

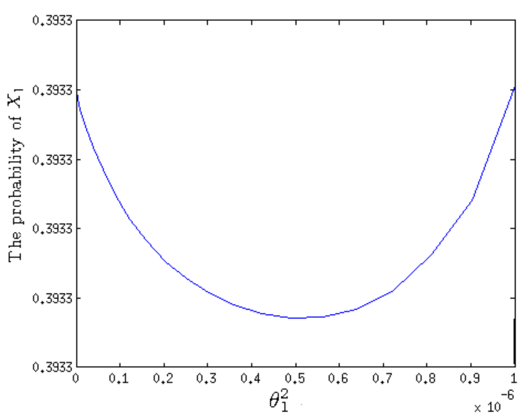
In this paragraph, we demonstrate that the condition $c > \sqrt{d}$ in main theorems is necessary to establish LSH property. The observation is that there exists a probability spread (a gap between Pmax and Pmin for a given R) when subspaces are of dimension d larger than one. In other words, $\mathrm{Pmin}_R \neq \mathrm{Pmax}_R$. Because of this spread, to ensure $\mathrm{Pmax}_R > p_1 = \mathrm{Pmin}_R > p_2 = \mathrm{Pmax}_{cR}$ means c can't be close to 1. The theoretical analysis leads to the condition $c > \sqrt{d}$. Numerical results below also support



(a)



(b)



(c)

Figure 3. The probability depends on principal angles

this. Since $\mu(h_{l,\theta_0} \in \mathcal{H} | h_{l,\theta_0}(L_1) = h_{l,\theta_0}(L_2))$ has a linear relation with $\mu(\mathcal{L}_\cap)$, we compute $\mu(\mathcal{L}_\cap)$ instead (In figure 3, $\mathcal{L}_\cap$ is shown as X_1).

In figure 3, we work with the space $\mathrm{G}(4, 2)$. For each subfigure in 3, we fix $R = \sqrt{\theta_1^2 + \theta_2^2}$ (the distance) and

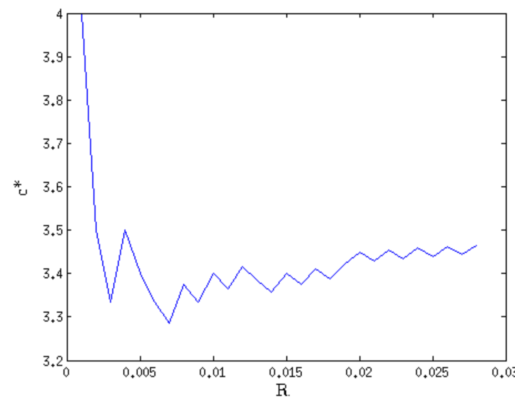


Figure 4. The constant c and the dimension d

η. Then, $\mu(X_1)$ (or $\mu(\mathcal{L}_\cap)$ for consistency) is computed for different pairs of principal angles (θ_1, θ_2) and plotted against θ_1^2. In figure 3(a), $R^2 = 0.01$ and $\eta = 0.01$; In figure 3(b), $R^2 = 0.01$ and $\eta = 0.1$; In figure 3(c), $R^2 = 0.001$ and $\eta = 1$.

From figure 3, it is easy to see that the probability is minimized when both of principal angles are equal given distance R is fixed. This verifies the general theory that states the probability reaches its maximum Pmax_R if there is only one nonzero principal angle and reaches its minimum Pmin_R if all angles are equal.

Now we show how the parameter c is related to the dimension of subspaces. Particularly, we are interested in the minimal value of c that ensures $p_1 > p_2$ for a given R. In figure 4, the minimal value of c for each R is plotted against R in the case of $G(20, 10)$. The figures show that the lower bound of constant c is approximately $\sqrt{d}$ where $d = 10$ is the dimension of subspaces.

3. Proof of Main Theorems

Since each hash function in $\mathcal{H}_{\theta_0}(d_1, d_2, D)$ corresponds to a line in $\mathbb{R}^D$, we can identify $\mathcal{H}_{\theta_0}(d_1, d_2, D)$ with the unit sphere $\mathbb{S}^{D-1}$ and assign to it a probability measure which is induced by the uniform probability measure on $\mathbb{S}^{D-1}$. We denote throughout this section the measure by μ, that is, $\mathbb{P} := \mu$.

Proof of Theorem 3.2. We fix $0 < \theta_0 < \pi/6$. For $L_1 \in \mathrm{G}(D, d)$ and $L_2 \in \mathrm{G}(D, 1)$, $\mu(h_{l,\theta_0} \in \mathcal{H}_{\theta_0}(1, d, D) | h_{l,\theta_0}(L_1) = h_{l,\theta_0}(L_2))$ depends only on the principal angle, which is the elevation angle $\theta(L_1, L_2)$ of the line L_2 with respect to the d-dimensional subspace L_1. We denote this probability by $g(\theta(L_1, L_2))$. To prove the theorem, we need only to show that $g(\theta)$ is a decreasing function of θ. Indeed, then $p_1 = g(\theta) > g(c\theta) = p_2$.

Let

$$B_{\mathrm{G}(D,1)}(L,\theta_0)=\{l\in \mathrm{G}(D,1)|\mathrm{dist}_{\mathrm{G}}(l,L)<\theta_0\},$$
$$\mathcal{L}_{\cap}(L_1,L_2)=\{l\in \mathrm{G}(D,1)|h_{l,\theta_0}(L_1)=h_{l,\theta_0}(L_2)=1\},$$

and

$$\mathcal{L}_{\cup^c}(L_1,L_2)=\{l\in \mathrm{G}(D,1)|h_{l,\theta_0}(L_1)=h_{l,\theta_0}(L_2)=0\}. \tag{5}$$

We note that,

$$\begin{aligned}\mathcal{L}_{\cap}(L_1,L_2)&=B_{\mathrm{G}(D,1)}(L_1,\theta_0)\cap B_{\mathrm{G}(D,1)}(L_2,\theta_0),\\ \mathcal{L}_{\cup^c}(L_1,L_2)&=(B_{\mathrm{G}(D,1)}(L_1,\theta_0)\cup B_{\mathrm{G}(D,1)}(L_2,\theta_0))^c,\end{aligned} \tag{6}$$

and

$$\begin{aligned}g(\theta(L_1,L_2))=&1-\mu(B_{\mathrm{G}(D,1)}(L_1,\theta_0))\\ &-\mu(B_{\mathrm{G}(D,1)}(L_2,\theta_0))+2\mu(\mathcal{L}_{\cap}(L_1,L_2)).\end{aligned} \tag{7}$$

Since $\mu(B_{\mathrm{G}(D,1)}(L_1,\theta_0))$ and $\mu(B_{\mathrm{G}(D,1)}(L_2,\theta_0))$ in (7) are independent of $\theta(L_1,L_2)$, it is enough to show that $\mu(\mathcal{L}_{\cap}(L_1,L_2))$ decrease as $\theta(L_1,L_2)$ increases.

Let L_1 be a d-dimensional subspace in $\mathbb{R}^D$ and L_2, L_3 be two lines in $\mathbb{R}^D$ such that $\theta(L_1,L_2)=\alpha$ and $\theta(L_1,L_3)=\alpha+\beta$ ($0<\beta,\alpha$ and $\alpha+\beta<\pi/6$ and $\alpha+\theta_0<\pi/4$). Let $\{e_i\}_{i=1}^D$ be a basis of $\mathbb{R}^D$ such that

$$\begin{aligned}L_1&=\mathrm{span}\{e_1,...,e_d\},\\ L_2&=\mathrm{span}\{\cos\alpha e_1+\sin\alpha e_{d+1}\},\end{aligned}$$

We may rotate L_3 in a direction orthogonal to L_1 and maintain the elevation angle $\theta(L_1,L_3)$ so that L_3 is modified as follows

$$L_3=\mathrm{span}\{\cos(\alpha+\beta)e_1+\sin(\alpha+\beta)e_{d+1}\}.$$

Throughout the rest of the proof we express coordinates and operators w.r.t. the basis $\{e_i\}_{i=1}^D$.

Let A be the rotation of R^D which rotates L_2 to L_3 within the subspace $\mathrm{span}\{e_1,e_{d+1}\}$. We denote the image of a line l under the rotation A by $A(l)$ and note that $A(L_2)=L_3$.

Let l be the line passing through the point $(a_1,...,a_D)\in\mathbb{S}^{D-1}$ and such that $l\in(B_{\mathrm{G}(D,1)}(L_1,\theta_0))^c\cap B_{\mathrm{G}(D,1)}(L_2,\theta_0)$. Since $l\in(B_{\mathrm{G}(D,1)}(L_1,\theta_0))^c$ and $\alpha+\theta_0<\pi/4$

$$\sum_{i=d+1}^{D}a_i^2>\sin\theta_0 \text{ and } a_1>a_{d+1}. \tag{8}$$

The image $A(l)$ is the line passing through $(a_1\cos\beta-a_{d+1}\sin\beta,a_2,...,a_d,a_1\sin\beta+a_{d+1}\cos\beta,...,a_D)$. The elevation angle $\theta(L_1,A(l))$ of $A(l)$ with respect to L_1 is

$$\begin{aligned}&\sin^{-1}(\sqrt{(a_1\sin\beta+a_{d+1}\cos\beta)^2+a_{d+2}^2...+a_D^2})\\ &>\sin^{-1}(\sqrt{a_{d+1}^2+...+a_D^2})>\theta_0.\end{aligned}$$

. Therefore, $A(l)\in(B_{\mathrm{G}(D,1)}(L_1,\theta_0))^c\cap B_{\mathrm{G}(D,1)}(L_3,\theta_0)$. That is, $A((B_{\mathrm{G}(D,1)}(L_1,\theta_0))^c\cap B_{\mathrm{G}(D,1)}(L_2,\theta_0))\subset(B_{\mathrm{G}(D,1)}(L_1,\theta_0))^c\cap B_{\mathrm{G}(D,1)}(L_3,\theta_0)$. Consequently,

$$\begin{aligned}&\mu((B_{\mathrm{G}(D,1)}(L_1,\theta_0))^c\cap B_{\mathrm{G}(D,1)}(L_2,\theta_0))\\ &=\mu(A((B_{\mathrm{G}(D,1)}(L_1,\theta_0))^c\cap B_{\mathrm{G}(D,1)}(L_2,\theta_0)))\\ &\le\mu((B_{\mathrm{G}(D,1)}(L_1,\theta_0))^c\cap B_{\mathrm{G}(D,1)}(L_3,\theta_0)).\end{aligned} \tag{9}$$

In view of (6), we can rewrite (9) as

$$\begin{aligned}&\mu(B_{\mathrm{G}(D,1)}(L_2,\theta_0)/\mathcal{L}_{\cap}(L_1,L_2))\\ &\qquad\le\mu(B_{\mathrm{G}(D,1)}(L_3,\theta_0)/\mathcal{L}_{\cap}(L_1,L_3))\end{aligned} \tag{10}$$

. Since $\mu(B_{\mathrm{G}(D,1)}(L_2,\theta_0))=\mu(B_{\mathrm{G}(D,1)}(L_3,\theta_0))$, (10) implies that

$$\mu(\mathcal{L}_{\cap}(L_1,L_2))\ge\mu(\mathcal{L}_{\cap}(L_1,L_3))$$

. That is, $\mu(\mathcal{L}_{\cap}(L_1,l))$ is a decreasing function of $\theta(L_1,l)$ for any l satisfying $\theta(L_1,l)<\min\{\pi/6,\pi/4-\theta_0\}$. Combining this observation with 7 we conclude that $g(\theta(L_1,L_2))$ is a decreasing function of $\theta(L_1,L_2)$ and thus conclude the proof. □

Proof of Theorem 3.3. We use similar notations as in the proof of Theorem 3.2. We fix $0<\theta_0<\pi/6$. For $L_1,L_2\in\mathrm{G}(D,d)$, the probability

$$\mu(h_{l,\theta_0}\in\mathcal{H}_{\theta_0}(d,D)|h_{l,\theta_0}(L_1)=h_{l,\theta_0}(L_2))$$

depends only on the principal angles between L_1 and L_2. Indeed, this probability equals the RHS of (7) (here, L_1, $L_2\in\mathrm{G}(D,d)$) and $\mu(\mathcal{L}_{\cap}(L_1,L_2))$ depends only on the relative position between the two subspaces. We denote this probability by $g(\theta_1,...,\theta_d)$ where $\{\theta_i\}_{i=1}^d$ are the principal angles.

It is obvious that if $L_1=L_2$, then the corresponding probability is $g(0,...,0)=1$ and it obtains the maximal value among all principal angles $(\theta_1,...,\theta_d)$. We will show that the directional derivatives of $g(\theta_1,...,\theta_d)$ w.r.t. any direction at the origin is strictly negative. We use our estimates to obtain a lower bound on g in a ball of radius R around the origin when R is sufficiently small and an upper bound on g in the ball of radius cR and use these bounds to conclude that $p_1>p_2$.

For convenience, we drop the requirements that $\theta_1\ge...\ge\theta_d$, but only assume that $0\le\theta_1,...,\theta_d\le\pi/2$. More precisely, for any $\theta_1,...,\theta_d\in[0,\pi/2]^d$ we can parametrize L_1 (in the right coordinate system) as $L_1=\{(x_1,...,x_d,0,...,0)|x_i\in\mathbb{R}\}$ and then $L_2=\{(x_1\cos\theta_1,...,x_d\cos\theta_d,x_1\sin\theta_1,...,x_d\sin\theta_d,0,...,0)|x_i\in\mathbb{R}\}$. This $\theta_1,...,\theta_d$ parametrize the relative position between L_1 and L_2 even though they don't satisfy $\theta_1\ge...\ge\theta_d$. We note that with this convention, $g(\theta_1,...,\theta_d)$ is invariant to permutations of its arguments.

We will verify the following two lemmas. Lemma 3.1 asserts that the probability $g(\theta_1, ..., \theta_d)$ indeed decreases around the origin in the coordinate directions. Lemma 3.2 reestablishes the connection between directional derivatives and coordinate derivatives (to use the chain rule we verify continuity of the partial derivatives). Moreover, it shows that the ratio of change in the fastest descent direction (the direction where θ_i change at the same rate) and in the slowest descent direction (the coordinate direction) is bounded by a factor of $\sqrt{d}$.

Lemma 3.1. *For each i, the coordinate directional derivative* $\frac{\partial g}{\partial \theta_i}|_{(\theta_1,...,\theta_d)=(0,...,0)}$ *is negative.*

Lemma 3.2. *For* $s = s_1\frac{\partial}{\partial \theta_1} + ... + s_d\frac{\partial}{\partial \theta_d}$, $\sqrt{d}||\frac{\partial g}{\partial \theta_1}||_2 \leq ||\frac{\partial g}{\partial s}||_2 \leq ||\frac{\partial g}{\partial \theta_1}||_2$.

The proofs of these two lemmas are given in the Appendix. Now, we give the proof of Theorem 3.3.

Let $\frac{\partial g}{\partial \theta_i} = -a$ $(a > 0)$ and $S(R) = \{(\theta_1, ..., \theta_d) | \sum_{i=1}^d \theta_i^2 = R^2\}$. Applying Taylor expansion to $g(\theta_1, ..., \theta_d)$ at the origin and Lemmas 3.1 and 3.2, we obtain that

$$\begin{aligned} &\max_{S(cR)} g(\theta_1, ..., \theta_d) \leq g(0, ..., 0) - acR + O(c^2R^2), \\ &\text{and } \min_{S(R)} g(\theta_1, ..., \theta_d) \geq g(0, ..., 0) - \sqrt{d}aR + O(R^2). \end{aligned} \tag{11}$$

Therefore, if $c > \sqrt{d}$ and R is sufficiently small, then,

$$\min_{S(R)} g(\theta_1, ..., \theta_d) > \max_{S(cR)} g(\theta_1, ..., \theta_d).$$

If we choose $p_1 = \min_{S(R)} g(\theta_1, ..., \theta_d)$ and $p_2 = \max_{S(cR)} g(\theta_1, ..., \theta_d)$, then $\mathcal{H}_{\theta_0}(d, D)$ is an (R, c, p_1, p_2)-sensitive hashing family by the definition of g and the LSH family.

□

Proof of Theorem 3.4. Fixing $0 < \theta_0 < \pi/6$, the neighborhood $B_{\mathrm{G}(D,1)}(L, \theta_0) = \{l \in \mathrm{G}(D, 1) | \mathrm{dist}_{\mathrm{G}}(l, L) < \theta_0\}$ of a line L is a hyperspherical cap (on the unit sphere).

Let L_1, L_2 and L_2' be three lines in $\mathbb{R}^D$ with $\mathrm{dist}(L_1, L_2) = \theta$ and $\mathrm{dist}(L_1, L_2') = c\theta$ for some $c, \theta > 0$. Moreover, let

$$\begin{aligned} &X_1 = B_{\mathrm{G}(D,1)}(L_1, \theta_0) \backslash B_{\mathrm{G}(D,1)}(L_2, \theta_0), \\ &X_2 = B_{\mathrm{G}(D,1)}(L_1, \theta_0) \cap B_{\mathrm{G}(D,1)}(L_2, \theta_0), \\ &\text{and } X_3 = X_2 \backslash B_{\mathrm{G}(D,1)}(L_2', \theta_0). \end{aligned} \tag{12}$$

Using this notation, we formulate the following lemma, which we later prove in the appendix.

Lemma 3.3. *Assume that* $R = \alpha/\sqrt{D}$ *for a fixed* $\alpha > 0$ *and that* $cR = O(1)$. *The measures* $x_1 := \mu(X_1)$ *and* $x_3 := \mu(X_3)$ *satisfy the following properties: when* $D \to \infty$, $x_1, x_3 \to 0$ *and* $\lim_{D\to\infty} x_3/x_1 > e^{\alpha^2/2}$.

We conclude Theorem 3.4 as follows. We denote $p_1 = \mu(h_{l,\theta_0} | h_{l,\theta_0}(L_1) = h_{l,\theta_0}(L_2)) = 1 - 2x_1$ and $p_2 = \mu(h_{l,\theta_0} | h_{l,\theta_0}(L_1) = h_{l,\theta_0}(L_2')) = 1 - 2(x_1 + x_3)$.

$$\begin{aligned} \lim_{D\to\infty} \rho &= \lim_{D\to\infty} \ln(p_1)/\ln(p_2) \\ &= \lim_{D\to\infty} \ln(1 - 2x_1)/\ln(1 - 2x_1 - 2x_3) \\ &= \lim_{D\to\infty} x_1/(x_1 + x_3) \qquad \text{(since } x_1, x_3 \to 0\text{)} \\ &= \lim_{D\to\infty} 1/(1 + x_3/x_1) \\ &< 1/(1 + e^{\alpha^2/2}). \qquad \text{(by Lemma 3.3)} \end{aligned} \tag{13}$$

□

Proof of Theorem 3.5. The GLH algorithm returns a point that is within cR distance of the query if there is a point in the dataset that is within R distance of the query. When the query is in the dataset, it is guaranteed to have a point within R distance of the query for any $R > 0$. Therefore, we can pick R arbitrarily small, and we note that p_1 approaches 1 as R approaches zero ($\ln(p_1)$ approaches zero). Moreover, we can pick c such that $cR = O(1)$. This can keep p_2 to be a fixed constant less than 1. That is, if the query is in the dataset, we are able to adjust c and R such that $\rho = \ln(p_1)/\ln(p_2) = \epsilon$ for any $\epsilon > 0$. □

A. Appendix

A.1. Local Behavior of $g(\theta_1, ..., \theta_d)$

Throughout this section, we use the following coordinate representation.

$$\begin{aligned} &\mathbb{R}^D: && (x_1, ..., x_d, y_1, ..., y_d, z_{2d+1}, ..., z_D), \\ &L_1: && \{(x_1, x_2, ..., x_d, 0, ..., 0) | x_i \in \mathbb{R}\}, \\ &\text{and} && \\ &L_2: && \{(x_1 \cos\theta_1, ..., x_d \cos\theta_d, x_1 \sin\theta_1, ..., \\ & && x_d \sin\theta_d, 0, ..., 0) | x_i \in \mathbb{R}, \theta_i > 0\}. \end{aligned} \tag{14}$$

Proof of Lemma 3.1 First, we study the derivative along the coordinate direction $\frac{\partial}{\partial \theta_1}$ at the origin. This is the case where $\theta_1 = \epsilon$ and $\theta_i = 0$ for $i \neq 1$. Suppose L_1 is given in (14) and L_2^ϵ is given by

$$\{(x_1 \cos\epsilon, x_2, ..., x_d, x_1 \sin\epsilon, 0, ..., 0) | x_i \in \mathbb{R}, \epsilon > 0\}.$$

Denote $g(\epsilon, 0..., 0) = g(\epsilon)$ for short.

Let $(a_1, ..., a_{2d-2}) \in \mathbb{R}^{2d-2}$. We define the following quantities:

$A(a_1, ..., a_{2d-2})$ is subset of S^{D-1} with coordinates:

$$(x_2, ..., x_d, y_2, ..., y_d) = (a_1, ..., a_{2d-2}),$$

$$h(\epsilon, a_1, ..., a_{2d-2}) = \mu(h_{l,\theta_0} \in \mathcal{H}_{\theta_0}(d, D) \\ |l \in A(a_1, ..., a_{2d-2}), h_{l,\theta_0}(L_1) \neq h_{l,\theta_0}(L_2^\epsilon)),$$

and

$\mathrm{Vol}_d(r)$ is volume of $(d-1)$-dim. sphere of radius r.

Then, we can write $g(\epsilon)$ as an integral of $h(\epsilon, a_1, ..., a_{2d-2})$ as follows:

$$g(\epsilon) = 1 - \int_{(a_1,..,a_{2d-2}) \in D^{2d-2}} h(\epsilon, a_1, .., a_{2d-2}) d\nu. \quad (15)$$

We observe that $\frac{\partial h(\epsilon,0,...,0)}{\partial \epsilon}$ where $a_i = 0, \forall i$. Using polar coordinates (r, θ) on (x_1, y_1)-plane. $h(\epsilon, 0, ..., 0)$ can be written in this way:

$$\frac{2}{\mathrm{Vol}_D(1)} \int_{-\epsilon/2}^{\theta_0} d\theta \int_{\frac{\cos\theta_0}{\cos\theta}}^{\frac{\cos\theta_0}{\cos(\theta+\epsilon)}} \mathrm{Vol}_{D-2d}(\sqrt{1-r^2}) r dr.$$

Therefore, the derivative $\frac{\partial h(\epsilon,0,...,0)}{\partial \epsilon}|_{\epsilon=0}$ is equal to

$$\frac{2}{\mathrm{Vol}_D(1)} \int_0^{\theta_0} \mathrm{Vol}_{D-2d}(\sqrt{1 - \frac{\cos^2\theta_0}{\cos^2\theta}}) \frac{\cos^2\theta_0 \sin\theta}{\cos^3\theta} d\theta.$$

This is bigger than zero. Since $\frac{\partial h}{\partial \epsilon}|_{\epsilon=0}$ is continuous and non-negative on $(a_1, ..., a_{2d})$, and when $(a_1, ..., a_{2d}) = 0$, $\frac{\partial h}{\partial \epsilon}|_{\epsilon=0}$ is strictly positive. We conclude,

$$\frac{\partial g(\theta_1, .., \theta_d)}{\partial \theta_1}|_{(\theta_1,...,\theta_d)=0} = \frac{\partial g(\epsilon)}{\partial \theta_1}|_{\epsilon=0} \\ = -\int_{(a_1,..,a_{2d}) \in D^{2d-2}} \frac{\partial h(\epsilon, a_1, .., a_{2d})}{\partial \epsilon}|_{\epsilon=0} d\nu < 0.$$

By symmetry,

$$\frac{\partial g}{\partial \theta_i} = \frac{\partial g}{\partial \theta_1} = -a < 0$$

for some fixed a and all i. □

Proof of Lemma 3.2 By symmetry of the function $g(\theta_1, ..., \theta_d)$, we need show the usual chain rule $\frac{\partial g}{\partial s} = s_1 \frac{\partial g}{\partial \theta_1} + ... + s_d \frac{\partial g}{\partial \theta_d}$ holds for the region $\{\theta_i \geq 0, \forall i\}$. Firstly, we show that the partial derivatives of $g(\theta_1, ..., \theta_d)$ are continuous up to the origin in the region $\{\theta_i \geq 0, \forall i\}$. Then, the chain rule follows from this fact.

Notice that $g(\theta_1, ..., \theta_d) = 1 - 2\mu(B_{\mathrm{G}(D,1)}(L, \theta_0)) + 2\mu(\mathcal{L}_\cap(L_1, L_2))$. Since the first two terms are constants, it is enough to show that the derivatives of $\mu(\mathcal{L}_\cap(L_1, L_2))$ is continuous up to the origin in the region $s_i \geq 0$. We shall prove this for a general class of functions.

Firstly, we introduce a class of rotations.

Definition A.1. *Given angles $\{\theta_i\}_{i=1}^d$ ($d > 0$), a rotation $A(\theta_1, ..., \theta_d)$ on $\mathbb{R}^D$ is defined as follows.*

For a point $\boldsymbol{x} = (x_1, ..., x_D) \in \mathbb{R}^D$, the i-th coordinate of the image $A(\theta_1, ..., \theta_d)(\boldsymbol{x})$ is equal to

$$\begin{cases} x_i \cos\theta_i, + x_{d+i} \sin\theta_i, & 1 \leq i \leq d, \\ -x_{i-d} \sin\theta_{i-d} + x_i \cos\theta_{i-d}, & d+1 \leq i \leq 2d, \\ x_i, & i \geq 2d. \end{cases}$$

In other words, $A(\theta_1, ..., \theta_d)$ rotates the first $2d$ coordinates. For a set $X \subset \mathbb{R}^D$, we denote its image under this rotation by $A(\theta_1, ..., \theta_d)(X)$.

Then, we define two set of functions.

Definition A.2. *Let $\mu_{\mathbb{S}^{D-1}}$ be the uniform measure on the unit sphere $\mathbb{S}^{D-1}$ and $\mu_{\mathbb{R}^D}$ be the Borel measure on $\mathbb{R}^D$. Given two smooth-boundary regions U and V on $\mathbb{S}^{D-1}$, a function of $(\theta_1, ..., \theta_d)$ is defined by*

$$G_{UV}(\theta_1, ..., \theta_d) = \mu_{\mathbb{S}^{D-1}}(U \cap A(\theta_1, ..., \theta_d)(V))$$

Moreover, denote by $C[U]$ and $C[V]$ the cones generated by connecting U and V with the origin respectively. We define

$$CG_{UV}(\theta_1, ..., \theta_d) = \mu_{\mathbb{R}^D}(C[U] \cap A(\theta_1, ..., \theta_d)(C[V]))$$

In the following, a convex polytope cone is a convex cone with vertex at the origin such that sides are hyperplanes and the base is enclosed by the unit sphere. Denote $\Theta = (\theta_1, ..., \theta_d)$ and $\boldsymbol{e}_i$ and $\boldsymbol{v}_i$ be the i-th coordinate direction of Θ and $\mathbb{R}^D$ respectively. Let X be a convex polytope cone with sides $\{F_i\}_{i=1}^S$. Let X_Θ be the cone with sides $A(\Theta)(F_1)$ and $\{F_i\}_{i=2}^S$.

Lemma A.3. *$\mu_{\mathbb{R}^D}(X_\Theta)$ is contiuously differentiable w.r.t. $\{\frac{\partial}{\partial \theta_i}\}_{i=1}^d$ in $[0, \alpha_1] \times ... \times [0, \alpha_d]$ for some positive numbers $\alpha_1, ..., \alpha_d > 0$.*

Let n^Θ be the unit normal direction of $A(\Theta)(F_1)$ and α^Θ be the elevation angle between n^Θ and the subspace spanned by $\{\boldsymbol{v}_1, \boldsymbol{v}_{d+1}\}$. Let Ω^Θ be the region on $A(\Theta)(F_1)$ enclosed by the other sides $\{F_i\}_{i=2}^S$ and the base $\mathbb{S}^{D-1}$. Specifically, $\Omega^{\mathbf{0}} = X$ when Θ is the origin $\mathbf{0}$.

We show that $\frac{\partial \mu_{\mathbb{R}^D}(X_\Theta)}{\partial \theta_1}$ is continuous.

Let $\Delta\Theta = \epsilon \boldsymbol{e}_1$. The angle $\mathrm{Ang}(\Theta, \Delta\Theta)$ between n^Θ and $n^{\Theta+\Delta\Theta}$ is $\cos^{-1}[\cos^2\alpha^\Theta \cos\epsilon + \sin^2\alpha^\Theta]$. Let $\mathrm{ProjNorm}(\boldsymbol{x}, \Theta, \Delta\Theta)$ be the norm of the projection of a point $\boldsymbol{x} \in A(\Theta)(F_1)$ to the plane spanned by n^Θ and

$n^{\Theta+\Delta\Theta}$. By direct computation, $\mathrm{ProjNorm}(\boldsymbol{x}, \Theta, \Delta\Theta)$ is equal to

$$\frac{(\cos\epsilon - 1)[n_1^\Theta x_1 + n_{d+1}^\Theta x_{d+1}] + \sin\epsilon[n_{d+1}^\Theta x_1 - n_1^\Theta x_{d+1}]}{(\sin^2\alpha^\Theta \cos^2\alpha^\Theta (1-\cos\epsilon)^2 + \cos^2\alpha^\Theta \sin^2\epsilon)^{1/2}}$$

Then, we can express the partial derivative as follows.

$$\frac{\partial \mu_{\mathbb{R}^D}(X_\Theta)}{\partial\theta_1} = \lim_{\epsilon\to 0}\int_{r=0}^1 r^2 dr \int_{\boldsymbol{x}\in\Omega^\Theta} \mathrm{ProjNorm}(\boldsymbol{x}, \Theta, \Delta\Theta) \times \mathrm{Ang}(\Theta, \Delta\Theta) d\boldsymbol{x}/\epsilon.$$

By applying Tyler's expansion, we have

$$\frac{\partial \mu_{\mathbb{R}^D}(X_\Theta)}{\partial\theta_1} = \int_{\boldsymbol{x}\in\Omega^\Theta} \frac{2(n_{d+1}^\Theta x_1 - n_1^\Theta x_{d+1})}{3\cos\alpha^\Theta(\sin^2\alpha^\Theta\cos^2\alpha^\Theta + 2\cos^2\alpha^\Theta)} d\boldsymbol{x}$$

Since n^Θ and α^Θ are continous as Θ approaches the origin. Moreover, the domain Ω^Θ will approach $\Omega^{\mathbf{0}}$. Therefore,

$$\lim_{\Theta\to\mathbf{0}} \frac{\partial \mu_{\mathbb{R}^D}(X_\Theta)}{\partial\theta_1} = \frac{\partial \mu_{\mathbb{R}^D}(X_\Theta)}{\partial\theta_1}\Big|_{\Theta=\mathbf{0}}.$$

This means $\mu_{\mathbb{R}^D}(X_\Theta)$ is continuously differentiable. $\square$

Lemma A.4. *if $C[U]$ and $C[V]$ are two convex polytope cones, then there exists some positive numbers $\alpha_1, ..., \alpha_d > 0$ such that $CG_{UV}(\theta_1, ..., \theta_d)$ has continuous partial derivatives in $[0, \alpha_1] \times ... \times [0, \alpha_d]$.*

The intersection $C[U] \cap A(\theta_1, ..., \theta_d)(C[V])$ is also a convex polytope cone. Its sides from U are fixed and sides from $A(\theta_1, ..., \theta_d)(C[V])$ are moving as $(\theta_1, ..., \theta_d)$ change. We can decompose the rotation of its moving sides into individual rotations of each moving side. Following Lemma A.3, The intersection has continuous partial derivatives if one side is moving. After combining individual rotations, we have partial derivatives of $CG_{UV}(\theta_1, ..., \theta_d)$ are continuous in $[0, \alpha_1] \times ... \times [0, \alpha_d]$. $\square$

For general smooth-boundary region U on $\mathbb{S}^{D-1}$, we approximate $C[U]$ by unions of polytope cones. Let $\{\mathcal{P}_i = \{X_{ij}\}_{j=1}^{N_i}\}_{i=1}^\infty$ be a sequence of partitions of $\mathbb{S}^{D-1}$ satisfying:

- $\forall i, \cup_{j=1}^{N_i} X_{ij} = \mathbb{S}^{D-1}$
- $\forall i, j$, $C[X_{ij}]$ is a polytope cone.
- if $i < k$, each piece X_{ij} of $\mathcal{P}_i$ is a union of pieces in $\mathcal{P}_k$. That is, $\mathcal{P}_k$ is a refinement of $\mathcal{P}_i$.
- $\max_{1\le j\le N_i} diam(X_{ij}) \le 1/n$ for $i = n, \forall n$ ($diam(X_{ij})$ is the diameter of X_{ij})

For each n, let $U^n = \bigcup_{X_{nj}\subset U;\, X_{nj}\in\mathcal{P}_n} X_{nj}$. Then we have an increasing sequence $U^1 \subset ... \subset U^n \subset ... \subset U$ and $\bigcup_{n=1}^\infty U^n = U$. Notice that $C[U^n]$ can be expressed as a finite collection of polyhedra cones.

From now on, we fix the sequence of partitions $\{\mathcal{P}_i\}_{i=1}^\infty$ and denote $G_{U^nV^n}$ and $CG_{U^nV^n}$ by G_{UV}^n and CG_{UV}^n respectively.

Lemma A.5. *Given U, V and $1 \le i \le d$, $\frac{\partial CG_{UV}}{\partial\theta_i}$ is continuous at the origin.*

By Lemma A.4, $\frac{\partial CG_{UV}^n}{\partial\theta_i}$ is continuous on $[0, \alpha_1^n] \times ... \times [0, \alpha_d^n]$ for $1 \le n \le \infty$ where α_i^j is positive $\forall i, j$.

$$\begin{aligned}\frac{\partial CG_{UV}(\Theta)}{\partial\theta_i} &= \lim_{\epsilon\to 0}\frac{CG_{UV}(\Theta+\epsilon\boldsymbol{e}_i) - CG_{UV}(\Theta)}{\epsilon} \\ &= \lim_{\epsilon\to 0}\frac{CG_{UV}^n(\Theta+\epsilon\boldsymbol{e}_i) - CG_{UV}^n(\Theta) - c\epsilon/n^{D-2}}{\epsilon} \\ &= \frac{\partial CG_{UV}^n(\Theta)}{\partial\theta_i} - c/n^{D-2}\end{aligned} \tag{16}$$

In the second equality, c is a bounded constant. If $\frac{\partial CG_{UV}}{\partial\theta_i}$ is not continuous at the origin, then there exists $\epsilon > 0$ such that $\forall\delta > 0$, there exist points $|\boldsymbol{a} - \boldsymbol{b}| < \delta$ s.t. $|\frac{\partial CG_{UV}(b)}{\partial\theta_i} - \frac{\partial CG_{UV}(a)}{\partial\theta_i}| > \epsilon$.

On the other hand, we can pick $N > 0$ so that $c/N^{D-2} \le \epsilon/2$. Moreover, we can pick $\delta > 0$ so that if $|\boldsymbol{a} - \boldsymbol{b}| < \delta$, then $|\frac{\partial CG_{UV}^N(b)}{\partial\theta_i} - \frac{\partial CG_{UV}^N(a)}{\partial\theta_i}| \le \epsilon/2$ since $\frac{\partial CG_{UV}^N}{\partial\theta_i}$ is continuous. By 16, we have $|\frac{\partial CG_{UV}(b)}{\partial\theta_i} - \frac{\partial CG_{UV}(a)}{\partial\theta_i}| < \epsilon$ for $|\boldsymbol{a} - \boldsymbol{b}| < \delta$. This contradiction asserts that $\frac{\partial CG_{UV}}{\partial\theta_i}$ is continuous at the origin. $\square$

Lemma A.6. *if CG_{UV} has continuous partial derivatives, then G_{UV} also has continuous partial derivatives.*

We have the following relation:

$$CG_{UV}(\theta_1, ..., \theta_d) = \int_{r=0}^1 r^{D-1} G_{UV}(\theta_1, ..., \theta_d) dr.$$

It is easy to see the lemma holds. $\square$

By above lemmas, it is easy to see that G_{UV} has continuous partial derivatives. And $g(\theta_1, ..., \theta_d) = G_{UU}(\theta_1, ..., \theta_d)$ with U and $V = A(\theta_1, ..., \theta_d)(U)$ be neighborhoods of L_1 and $L_2 = A(\theta_1, ..., \theta_d)(L_1)$ respectively. Thus, $g(\theta_1, ..., \theta_d)$ has continuous partial derivatives and the chain rule holds for it.

Using the above result, we can now prove Lemma 3.2.

By chain rule, $\frac{\partial g}{\partial s} = |s_1|\frac{\partial g}{\partial\theta_1} + ... + |s_d|\frac{\partial g}{\partial\theta_d}$. Since $\frac{\partial g}{\partial\theta_i} = \frac{\partial g}{\partial\theta_j}$, $\frac{\partial g}{\partial s} = (|s_1| + ... + |s_d|)\frac{\partial g}{\partial\theta_1}$. Note $s_1^2 + ... + s_d^2 = 1$. By Cauchy-Schwarz inequality, $1 \le |s_1| + ... + |s_d| \le \sqrt{d}$. Thus, the inequality for directional derivatives follows. $\square$

A.2. Hyperspherical Area As D Approaches Infinity

In this section, we use probability measure on $\mathbb{S}^{D-1}$ such that $\mu(\mathbb{S}^{D-1}) = 1$.

Lemma A.7. *For any fixed* $0 < \theta_1 < \theta_0 < \pi/6$, *both* $\mu(B_{\mathrm{G}(D,1)}(L, \theta_0))$ *and the ratio of* $\mu(B_{\mathrm{G}(D,1)}(L, \theta_1))/\mu(B_{\mathrm{G}(D,1)}(L, \theta_0))$ *approaches zero, as the dimension* D *approaches infinity.*

The usual volumn of hyperspherical cap is $\pi^{(D-2)/2}/\Gamma(D/2)\int_0^\theta \sin^{D-1}(t)dt$. So,

$$\begin{aligned}
&\mu(B_{\mathrm{G}(D,1)}(L, \theta_0)) \\
&\qquad = \mathrm{Vol}(B_{\mathrm{G}(D,1)}(L, \theta_0))/\,\mathrm{Vol}(\mathbb{S}^{D-1}) \\
&\qquad = \int_0^{\theta_1} \sin^{D-1}(t)dt \Big/ (2\pi) \\
&\qquad \to 0, \text{ as } D \text{ approaches infinity.}
\end{aligned} \tag{17}$$

$$\begin{aligned}
&\mu(B_{\mathrm{G}(D,1)}(L, \theta_1))/\mu(B_{\mathrm{G}(D,1)}(L, \theta_0)) \\
&\quad = \mathrm{Vol}(B_{\mathrm{G}(D,1)}(L, \theta_1))/\,\mathrm{Vol}(B_{\mathrm{G}(D,1)}(L, \theta_0)) \\
&\quad = \int_0^{\theta_1} \sin^{D-1}(t)dt \Big/ \int_0^{\theta_0} \sin^{D-1}(t)dt \\
&\quad < \int_0^{\theta_1} \sin^{D-1}(t)dt \Big/ (\sin^{D-1}(\frac{\theta_0+\theta_1}{2}) * \frac{\theta_0-\theta_1}{2}) \\
&\quad < \frac{2}{\theta_0-\theta_1}\int_0^{\theta_1} [\sin(t)/\sin(\frac{\theta_0+\theta_1}{2})]^{D-1}dt \\
&\quad \to 0, \text{ as } D \text{ approaches infinity.}
\end{aligned} \tag{18}$$

Integrals in (17),(18) approach zero because θ_0, θ_1 are fixed and the integrant approaches zero as D approaches infinity. □

Proof of Lemma 3.3. We use the same notation as in Proof of Theorem 3.4. In addition, we denote $B_{\mathrm{G}(D,1)}(L_1, \theta_0) \cap B_{\mathrm{G}(D,1)}(L_2', \theta_0)$ by X_4 and $\mu(X_4)$ by x_4 and $\mu(B_{\mathrm{G}(D,1)}(L_1, \theta_0))$ by x. It is easy to see $B_{\mathrm{G}(D,1)}(L_1, \theta_0) = X_1 \cup X_2$ and $X_2 = X_3 \cup X_4$.

Since X_1 and X_3 are subsets of the hyperspherical cap $\mu(B_{\mathrm{G}(D,1)}(L, \theta_0))$. By Lemma A.7 above, $\mu(B_{\mathrm{G}(D,1)}(L, \theta_0))$ approaches zero. So, x_1 and x_3 also approach zero as D approaches infinity.

Now, we show $\lim\limits_{D\to\infty} x_3/x_1 > e^{\alpha^2/2}$.

Let L_3 be the line passing through the middle point of the great circle connecting L_1 and L_2. Then, X_2 contains the hyperspherical cap $B_{\mathrm{G}(D,1)}(L_3, \theta_0 - \frac{\theta}{2})$. this implies $x_3 + x_4 > \mu(B_{\mathrm{G}(D,1)}(L_3, \theta_0 - \frac{\theta}{2}))$.

Moreover, since X_4 is contained in a hyperspherical cap with smaller angle than θ_0, $\lim\limits_{D\to\infty} x_4/x = 0$ by Lemma A.7.

First, we compute

$$\begin{aligned}
&\lim_{D\to\infty}\int_0^{\theta_0} \sin^{D-1}(t)dt \Big/ \int_0^{\theta_0-\frac{\theta}{2}} \sin^{D-1}(t)dt \\
&= 1 + \lim_{D\to\infty}\int_{\theta_0-\frac{\theta}{2}}^{\theta_0} \sin^{D-1}(t)dt \Big/ \int_0^{\theta_0-\frac{\theta}{2}} \sin^{D-1}(t)dt \\
&< 1 + \lim_{D\to\infty} (\sin^D(\theta_0) * \frac{\theta}{2})/(\sin^D(\theta_0-\theta) * \frac{\theta}{2}) \\
&= 1 + \lim_{D\to\infty} (\sin(\theta_0)/\sin(\theta_0-\theta))^D \\
&= 1 + \lim_{D\to\infty} \cos^D(\theta) \\
&= 1 + \lim_{D\to\infty} (1-\theta^2/2 + O(\theta^4))^D \\
&= 1 + \lim_{D\to\infty} (1-(\alpha^2/2)/D + O(D^2))^D \\
&= 1 + e^{-\alpha^2/2}.
\end{aligned} \tag{19}$$

Then,

$$\begin{aligned}
&\lim_{D\to\infty} x_3/x_1 > \lim_{D\to\infty} x_3/(x-x_3) \\
&= \lim_{D\to\infty} (x/x_3 - 1)^{-1} \\
&= \lim_{D\to\infty} (x/(x_3+x_4) - 1)^{-1} \\
&\geq \lim_{D\to\infty} \left(x \Big/ \mu(B_{\mathrm{G}(D,1)}(L_3, \theta_0 - \frac{\theta}{2})) - 1 \right)^{-1} \\
&= \lim_{D\to\infty} \left(\int_0^{\theta_0} \sin^{D-1}(t)dt \Big/ \int_0^{\theta_0-\frac{\theta}{2}} \sin^{D-1}(t)dt - 1 \right)^{-1} \\
&> e^{\alpha^2/2}. \qquad \text{(by (19))}
\end{aligned}$$

□