A Proofs

A.1 Proof of Theorem 2

Theorem 2. *For any fixed γ, let $z \sim N(0,\gamma^2)$, define $D_s = \mathbb{E}[(\sin(Q(z)) - \sin(z))^2]$, $\zeta_s = Cov\big(\sin(Q(z)) - \sin(z), \sin(z)\big)$, and D_c and ζ_c analogously for cosine function. Further denote $\triangle_c = \mathbb{E}[\cos(Q(z))] - e^{-\frac{\gamma^2}{2}}$ and $\tilde{D}_c = D_c - \triangle_c^2$. Denote $V_s^* = \frac{1}{2}\left[1 - e^{-2\rho^2\gamma^2}\right]$ and $V_c^* = \frac{1}{2}\left[1 + e^{-2\rho^2\gamma^2}\right] - e^{-\rho^2\gamma^2}$. Assume x, y are two normalized samples with correlation ρ. Then at γ, $K_Q(x,y)$ is lower and upper bounded respectively by*

$$K(x,y) - D_s - D_c + 2e^{-\frac{\gamma^2(1-\rho^2)}{2}}\big(C_{1-} + C_{2-}\big),$$
$$K(x,y) + D_s + D_c + 2e^{-\frac{\gamma^2(1-\rho^2)}{2}}\big(C_{1+} + C_{2+}\big),$$

where $C_{1\pm} = (C_1C_2 \pm \sqrt{(1-C_1^2)(1-C_2^2)})\sqrt{D_sV_s^*}$, $C_{2\pm} = \big[(C_3C_4 \pm \sqrt{(1-C_3^2)(1-C_4^2)})\sqrt{D_cV_c^*} + e^{-\frac{\gamma^2}{2}}\triangle_c\big]$, with

$$C_1 = \sqrt{\frac{2\zeta_s}{D_s(1-e^{-2\gamma^2})}}, \qquad C_2 = \frac{e^{-\frac{\gamma^2(1-\rho)^2}{2}} - e^{-\frac{\gamma^2(1+\rho)^2}{2}}}{\sqrt{2(1-e^{-2\gamma^2})V_s^*}},$$
$$C_3 = \sqrt{\frac{\zeta_c}{\tilde{D}_c(\frac{1}{2}\left[1+e^{-2\gamma^2}\right] - e^{-\gamma^2})}}, \qquad C_4 = \frac{\frac{1}{2}\left[e^{-\frac{\gamma^2(1-\rho)^2}{2}} + e^{-\frac{\gamma^2(1+\rho)^2}{2}}\right] - e^{-\frac{\gamma^2(1+\rho^2)}{2}}}{\sqrt{(\frac{1}{2}\left[1+e^{-2\gamma^2}\right] - e^{-\gamma^2})V_c^*}}.$$

Proof. First we look at the sine function. In this proof, we will use the notation $(u,v) = (w^Tx, w^Ty) \sim N\left(0, \gamma^2\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right)$ to denote the projected data. Q is a general quantizer applied to the linear random projections. We have

$$\begin{aligned}
\mathbb{E}[\sin(Q(u)\sin(Q(v))] &= \mathbb{E}\Big[\big(\sin(Q(u)) - \sin(u) + \sin(u)\big) \\
&\qquad \big(\sin(Q(v)) - \sin(v) + \sin(v)\big)\Big] \\
&= \mathbb{E}[(\sin(Q(u)) - \sin(u))(\sin(Q(v)) - \sin(v))] \\
&\quad + 2\mathbb{E}[(\sin(Q(u)) - \sin(u))\sin(v)] + E[\sin(u)\sin(v)] \\
&\triangleq T_1 + 2T_2 + E[\sin(u)\sin(v)]. \qquad (8)
\end{aligned}$$

By Young's inequality, we have

$$-D_s \le T_1 \le D_s. \qquad (9)$$

To bound the second term, the following identities would be useful. For $u \sim N(0,\gamma^2)$,

$$\begin{aligned}
\mathbb{E}[\cos(u)] &= e^{-\frac{\gamma^2}{2}}, \\
\mathbb{E}[\sin(au)\sin(bu)] &= \frac{1}{2}\left[e^{-\frac{\gamma^2(a-b)^2}{2}} - e^{-\frac{\gamma^2(a+b)^2}{2}}\right], \\
\mathbb{E}[\cos(au)\cos(bu)] &= \frac{1}{2}\left[e^{-\frac{\gamma^2(a-b)^2}{2}} + e^{-\frac{\gamma^2(a+b)^2}{2}}\right].
\end{aligned}$$

Since we can write $v = \rho u + \sqrt{1-\rho^2}Z$ with $Z \sim N(0, \gamma^2)$ independent of u, we have

$$\begin{aligned} T_2 &= \mathbb{E}\Big[(\sin(Q(u)) - \sin(u))\sin(\rho u + \sqrt{1-\rho^2}Z)\Big] \\ &= \mathbb{E}\Big[(\sin(Q(u)) - \sin(u))\sin(\rho u)\cos(\sqrt{1-\rho^2}Z)\Big] \\ &\quad + \mathbb{E}\Big[(\sin(Q(u)) - \sin(u))\cos(\rho u)\sin(\sqrt{1-\rho^2}Z)\Big] \\ &= e^{-\frac{\gamma^2(1-\rho^2)}{2}}\mathbb{E}\Big[(\sin(Q(u)) - \sin(u))\sin(\rho u)\Big]. \end{aligned} \tag{10}$$

By assumption,

$$\begin{aligned} Cov[\sin(Q(u)) - \sin(u), \sin(u)] &= \zeta_s, \\ Var[\sin(Q(u)) - \sin(u)] &= D_s, \end{aligned}$$

Now we can compute

$$\begin{aligned} Var[\sin(u)] &= \frac{1}{2}\left[1 - e^{-2\gamma^2}\right], \\ Cov[\sin(\rho u), \sin(u)] &= \frac{1}{2}\left[e^{-\frac{\gamma^2(1-\rho)^2}{2}} - e^{-\frac{\gamma^2(1+\rho)^2}{2}}\right], \\ Var[\sin(\rho u)] &= \frac{1}{2}\left[1 - e^{-2\rho^2\gamma^2}\right] \triangleq V_s^*. \end{aligned}$$

The correlation coefficients are

$$\begin{aligned} C_1 &\triangleq Corr[\sin(Q(u)) - \sin(u), \sin(u)] = \frac{\sqrt{2}\zeta_s}{\sqrt{D_s(1-e^{-2\gamma^2})}}, \\ C_2 &\triangleq Corr[\sin(\rho u), \sin(u)] = \frac{e^{-\frac{\gamma^2(1-\rho)^2}{2}} - e^{-\frac{\gamma^2(1+\rho)^2}{2}}}{\sqrt{(1-e^{-2\gamma^2})(1-e^{-2\rho^2\gamma^2})}}. \end{aligned}$$

By Cauchy-Schwartz inequality, we know that $Corr[\sin(Q(u)) - \sin(u), \sin(\rho u)]$ is bounded between

$$C_{l1} \triangleq C_1C_2 - \sqrt{(1-C_1^2)(1-C_2^2)},$$

$$C_{u1} \triangleq C_1C_2 + \sqrt{(1-C_1^2)(1-C_2^2)}.$$

Therefore, we have that

$$\begin{aligned} C_{l1}\sqrt{D_sV_s^*} \le E\Big[(\sin(Q(u)) - \sin(u))\sin(\rho u)\Big] \\ \le C_{u1}\sqrt{D_sV_s^*}. \end{aligned}$$

Combining with (8), (9) and (10) gives the expression for the sine part. For the cosine, we can use similar approach. From now on, denote $Q = Q_{c,\gamma}$. In particular, we can have

$$\mathbb{E}[\cos(Q(u)\cos(Q(v))] = T_3 + 2T_4 + E[\cos(u)\cos(v)],$$

where

$$\begin{aligned} T_3 &= \mathbb{E}[(\cos(Q(u)) - \cos(u))(\cos(Q(v)) - \cos(v))], \\ T_4 &= \mathbb{E}[(\cos(Q(u)) - \cos(u))\cos(v)]. \end{aligned}$$

Similarly, we have

$$-D_c \le T_3 \le D_c,$$

and

$$T_4 = e^{-\frac{\gamma^2(1-\rho^2)}{2}}\mathbb{E}\Big[(\cos(Q(u)) - \cos(u))\cos(\rho u)\Big].$$

Similarly, we can obtain

$$\begin{aligned}
&Cov[\cos(Q(u)) - \cos(u), \cos(u)] = \zeta_c, \\
&Var[\cos(Q(u)) - \cos(u)] = \tilde{D}_c, \\
&Var[\cos(u)] = \frac{1}{2}\left[1 + e^{-2\gamma^2}\right] - e^{-\gamma^2}, \\
&Cov[\cos(\rho u), \cos(u)] \\
&= \frac{1}{2}\left[e^{-\frac{\gamma^2(1-\rho)^2}{2}} + e^{-\frac{\gamma^2(1+\rho)^2}{2}}\right] - e^{-\frac{\gamma^2(1+\rho^2)}{2}}, \\
&Var[\cos(\rho u)] = \frac{1}{2}\left[1 + e^{-2\rho^2\gamma^2}\right] - e^{-\rho^2\gamma^2} \triangleq V_c^*.
\end{aligned}$$

The remaining part is similar, where we use Cauchy-Schwartz to bound the correlation of $\cos(Q(u)) - \cos(u)$ and $\cos(\rho u)$. We omit it for conciseness. The desired result is obtained by combining two parts and noticing that $\mathbb{E}[\sin(u)\sin(v) + \cos(u)\cos(v)] = e^{-\gamma^2(1-\rho)} = K(u, v)$.

□

A.2 Proof of Theorem 3

Theorem 3. ***(Uniform Approximation Error)*** *Assume the sample space $\mathcal{S}$ is the unit sphere (normalized data). Let QRP-RFF estimators be defined as (6). Let $\Gamma \sim N(0, \gamma^2)$ in Definition 2. Suppose a quantizer Q is mean smooth w.r.t.* sin *and* cos *functions with Lipschitz constant L_Q^s and L_Q^c, respectively. Then for $\forall\ \epsilon > 0$, with probability at least $1 - 4e^{-k\epsilon^2/256}$,*

$$|\hat{K}_Q(x, y) - K_Q(x, y)| \le \epsilon, \quad \text{for } \forall x, y \in \mathcal{S},$$

when $k \ge \frac{512d}{\epsilon^2}\log(\frac{64\max\{L_Q^s, L_Q^c\}\gamma}{\epsilon} + 1)$.

Proof. We denote the sample space (unit sphere) as $\mathcal{S} = \mathbb{S}^{d-1}$. Let $\tilde{\mathcal{S}}_\triangle$ be a $\triangle$-net placed on $\mathcal{S}$. We then can express any $x \in \mathcal{S}$ as $x = \tilde{x} + r_x$, for the center $\tilde{x} \in \tilde{\mathcal{S}}_\triangle$ and $\|r_x\| \le \triangle$.

Define

$$\begin{aligned}
K_Q^s(x, y) &= \mathbb{E}[\sin(Q(w^T x))\sin(Q(w^T y))], \\
K_Q^c(x, y) &= \mathbb{E}[\cos(Q(w^T x))\cos(Q(w^T y))].
\end{aligned}$$

We have

$$\begin{aligned}
|\hat{K}_Q(x, y) - K_Q(x, y)| &= |\hat{K}_Q^s(x, y) + \hat{K}_Q^c(x, y) - K_Q^s(x, y) - K_Q^c(x, y)| \\
&\le |\hat{K}_Q^s(x, y) - K_Q^s(x, y)| + |\hat{K}_Q^c(x, y) - K_Q^c(x, y)|. \qquad (11)
\end{aligned}$$

As before, we mainly provide details on the sine part, and the reasoning applies to the cosine part similarly. For any $x, y \in \mathcal{S}$, firstly we assume that the following two events hold:

$$\begin{aligned}
\Omega_1: &\quad \sup_{\tilde{x} \in \tilde{\mathcal{S}}_\triangle} \frac{1}{k}\sum_{i=1}^k \sup_{\|r\| \le \triangle} |\sin(Q(w_i^T \tilde{x} + w_i^T r)) - \sin(Q(w_i^T \tilde{x}))| \le L_Q^s \gamma \triangle + \epsilon_1, \\
\Omega_2: &\quad \sup_{\tilde{x}, \tilde{y} \in \tilde{\mathcal{S}}_\triangle} |\hat{K}_Q^s(\tilde{x}, \tilde{y}) - K_Q^s(\tilde{x}, \tilde{y})| \le \epsilon_2.
\end{aligned}$$

For any $x, y \in \mathcal{S}$, we have the following bound by triangle inequality,

$$\begin{aligned}
|\hat{K}_Q^s(x, y) - K_Q^s(x, y)| \le\ & |\hat{K}_Q^s(x, y) - \hat{K}_Q^s(\tilde{x}, y)| + |\hat{K}_Q^s(\tilde{x}, y) - \hat{K}_Q^s(\tilde{x}, \tilde{y})| \\
&+ |\hat{K}_Q^s(\tilde{x}, \tilde{y}) - K_Q^s(\tilde{x}, \tilde{y})| + |K_Q^s(x, y) - K_Q^s(\tilde{x}, \tilde{y})| \\
\triangleq\ & T_1 + T_2 + T_3 + T_4. \qquad (12)
\end{aligned}$$

We now bound these terms separately. We have

$$
\begin{aligned}
T_1 &= \frac{1}{k}\Big|\sum_{i=1}^{k}\sin(Q(w_i^T\tilde{x}+w_i^Tr_x))\sin(Q(w_i^T\tilde{y}+w_i^Tr_y))-\sin(Q(w_i^T\tilde{x}))\sin(Q(w_i^T\tilde{y}+w_i^Tr_y))\Big| \\
&= \frac{1}{k}\Big|\sum_{i=1}^{k}\big[\sin(Q(w_i^T\tilde{x}+w_i^Tr_x))-\sin(Q(w_i^T\tilde{x}))\big]\sin(Q(w_i^T\tilde{y}+w_i^Tr_y))\Big| \\
&\le L_Q^s\gamma\triangle+\epsilon_1,
\end{aligned}
$$

where the last line is due to event Ω_1 and boundedness of sine function. Similarly,

$$
\begin{aligned}
T_2 &= \frac{1}{k}\Big|\sum_{i=1}^{k}\sin(Q(w_i^T\tilde{x}))\big[\sin(Q(w_i^T\tilde{y}+w_i^Tr_y))-\sin(Q(w_i^T\tilde{y}))\big]\Big| \\
&\le L_Q^s\gamma\triangle+\epsilon_1.
\end{aligned}
$$

The event Ω_2 directly implies that

$$T_3 \le \epsilon_2.$$

For T_4, by mean smoothness assumption we have

$$
\begin{aligned}
T_4 &= \Big|\mathbb{E}\big[\sin(Q(w_i^T\tilde{x}+w_i^Tr_x))\sin(Q(w_i^T\tilde{y}+w_i^Tr_y))-\sin(Q(w_i^T\tilde{x}))\sin(Q(w_i^T\tilde{y}))\big]\Big| \\
&= \Big|\mathbb{E}\big[\sin(Q(w_i^T\tilde{x}+w_i^Tr_x))\sin(Q(w_i^T\tilde{y}+w_i^Tr_y))-\sin(Q(w_i^T\tilde{x}))\sin(Q(w_i^T\tilde{y})) \\
&\qquad +\sin(Q(w_i^T\tilde{x}+w_i^Tr_x))\sin(Q(w_i^T\tilde{y}))-\sin(Q(w_i^T\tilde{x}+w_i^Tr_x))\sin(Q(w_i^T\tilde{y}))\big]\Big| \\
&\le \mathbb{E}\Big|\sin(Q(w_i^T\tilde{x}+w_i^Tr_x))[\sin(Q(w_i^T\tilde{y}+w_i^Tr_y))-\sin(Q(w_i^T\tilde{y}))] \\
&\qquad -[\sin(Q(w_i^T\tilde{x}+w_i^Tr_x))-\sin(Q(w_i^T\tilde{x}))]\sin(Q(w_i^T\tilde{y}))\Big| \\
&\le L_Q^s\mathbb{E}\big[\|w_i^Tr_x\|+\|w_i^Tr_y\|\big] \\
&\le 2L_Q^s\gamma\triangle.
\end{aligned}
$$

Summing up ingredients together in (12) we get that in event Ω_1 and Ω_2, we have

$$|\hat{K}_Q^s(x,y)-K_Q^s(x,y)| \le 2\epsilon_1+\epsilon_2+4L_Q^s\gamma\triangle.$$

To derive a high probability bound, we now investigate the two events. First, we have the complement

$$P\Big[\Omega_1^c\Big] = P\Big[\sup_{\tilde{x}\in\tilde{\mathcal{S}}_\triangle}\frac{1}{k}\sum_{i=1}^{k}\sup_{\|r\|\le\triangle}|\sin(Q(w_i^T\tilde{x}+w_i^Tr))-\sin(Q(w_i^T\tilde{x}))| \ge L_Q^s\gamma\triangle+\epsilon_1\Big].$$

Since the terms in the summation, $\sup_{\|r\|\le\triangle}|\sin(Q(w_i^T\tilde{x}+w_i^Tr))-\sin(Q(w_i^T\tilde{x}))|$, are i.i.d. random variables, for any $\tilde{x}\in\tilde{\mathcal{S}}_\triangle$ the expectation admits

$$\mathbb{E}\Big[\sup_{\|r\|\le\triangle}|\sin(Q(w_i^T\tilde{x}+w_i^Tr))-\sin(Q(w_i^T\tilde{x}))|\Big] \le L_Q^s\gamma\triangle,$$

due to mean smoothness of Q. By Hoeffding's inequality on bounded variables, we get $\forall\tilde{x}\in\tilde{\mathcal{S}}_\triangle$

$$P\Big[\frac{1}{k}\sum_{i=1}^{k}\sup_{\|r\|\le\triangle}|\sin(Q(w_i^T\tilde{x}+w_i^Tr))-\sin(Q(w_i^T\tilde{x}))| \ge L_Q^s\gamma\triangle+\epsilon_1\Big] \le e^{-2k^2\epsilon_1^2/4k} = e^{-\frac{k\epsilon_1^2}{2}}.$$

Applying union bound over all $\tilde{x}\in\tilde{\mathcal{S}}_\triangle$, we obtain

$$P\Big[\Omega_1^c\Big] \le |\tilde{\mathcal{S}}_\triangle|e^{-k\epsilon_1^2/2} \le (\frac{2}{\triangle}+1)^d e^{-k\epsilon_1^2/2},$$

where the last inequality is due to the bound on covering number of the unit sphere (Corollary 4.2.13 in Vershynin (2018)). When $k \geq \frac{4d}{\epsilon_1^2}\log(\frac{2}{\triangle}+1)$, we have $P[\Omega_1^c] \leq e^{-k\epsilon_1^2/4}$. For Ω_2, applying Hoeffding's inequality yields a point-wise bound, where for $\forall \tilde{x}, \tilde{y} \in \tilde{\mathcal{S}}_\triangle$,

$$P\Big[|\hat{K}_Q^s(\tilde{x},\tilde{y}) - K_Q^s(\tilde{x},\tilde{y})| \geq \epsilon_2\Big] = P\Big[\frac{1}{k}|\sum_{i=1}^k \sin(Q(\tilde{x}))\sin(Q(\tilde{y})) - K_Q^s(\tilde{x},\tilde{y})| \geq \epsilon_2\Big]$$
$$\leq 2e^{-k\epsilon_2^2/2}.$$

Casting an union bound over $(\tilde{x},\tilde{y}) \in \tilde{\mathcal{S}}_\triangle \times \tilde{\mathcal{S}}_\triangle$ yields

$$\begin{aligned} P\Big[\Omega_2^c\Big] &= P\Big[\sup_{\tilde{x},\tilde{y}\in\tilde{\mathcal{S}}_\triangle} |\hat{K}_Q^s(\tilde{x},\tilde{y}) - K_Q^s(\tilde{x},\tilde{y})| \geq \epsilon_2\Big] \\ &\leq 2\binom{|\tilde{\mathcal{S}}_\triangle|}{2} e^{-k\epsilon_2^2/2} \\ &\leq (\frac{2}{\triangle}+1)^{2d} e^{-k\epsilon_2^2/2}. \end{aligned}$$

Consequently, $P[\Omega_2^c] \leq e^{-k\epsilon_2^2/4}$ when $k \geq \frac{8d}{\epsilon_1^2}\log(\frac{2}{\triangle}+1)$. Therefore, we obtain that when $k \geq 4d\log(\frac{2}{\triangle}+1)\max\{\epsilon_1^{-2}, 2\epsilon_2^{-2}\}$,

$$P\Big[\Omega_1^c \cup \Omega_2^c\Big] \leq e^{-k\epsilon_1^2/4} + e^{-k\epsilon_2^2/4}.$$

Now by letting $\epsilon_1 = \epsilon_2 = \epsilon/8$, and choosing $\triangle = \frac{\epsilon}{32L_Q^s\gamma}$, we have proved that when $k \geq \frac{512d}{\epsilon^2}\log(\frac{64L_Q^s\gamma}{\epsilon}+1)$, the error of sine part is bounded as

$$|\hat{K}_Q^s(x,y) - K_Q^s(x,y)| \leq \epsilon/2,$$

with probability at least $1-2e^{-k\epsilon^2/256}$. Similarly analysis can be used to bound the cosine part. For conciseness we omit the detailed proof. It is true that when $k \geq \frac{512d}{\epsilon^2}\log(\frac{64L_Q^c\gamma}{\epsilon}+1)$, with probability $1-2e^{-k\epsilon^2/256}$ we have

$$|\hat{K}_Q^c(x,y) - K_Q^c(x,y)| \leq \epsilon/2.$$

Therefore, by (11) and union bound we know that when $k \geq \frac{512d}{\epsilon^2}\log(\frac{64\max\{L_Q^s,L_Q^c\}\gamma}{\epsilon}+1)$, the kernel approximation error is uniformly bounded by

$$|\hat{K}_Q(x,y) - K_Q(x,y)| \leq \epsilon,$$

for $\forall x, y \in \mathcal{S}$, with probability $1-4e^{-k\epsilon^2/256}$. This completes the proof. $\square$