A The Power of CQ for high SNR

Let us first briefly describe what we believe is an important mistake in Chen & Qin (2010) - all notations, equation numbers and theorems in this paragraph refer to those in Chen & Qin (2010). Using the test statistic $T_n/\hat{\sigma}_{n1}$ defined below Theorem 2, we can derive the power under assumption (3.5) as

$$\begin{aligned} & P_1\left(\frac{T_n}{\hat{\sigma}_{n1}} > \xi_\alpha\right) = \\ = \quad & P_1\left(\frac{T_n - \|\mu_1-\mu_2\|^2}{\hat{\sigma}_{n2}} > \frac{\hat{\sigma}_{n1}}{\hat{\sigma}_{n2}}\xi_\alpha - \frac{\|\mu_1-\mu_2\|^2}{\hat{\sigma}_{n2}}\right) \\ \to \quad & \Phi\left(\frac{\|\mu_1-\mu_2\|^2}{\hat{\sigma}_{n2}}\right) \text{ (the denominator is } \textit{not } \hat{\sigma}_{n1}) \\ = \quad & \Phi\left(\frac{\sqrt{n}\|\mu_1-\mu_2\|^2}{\sqrt{(\mu_1-\mu_2)^T\Sigma(\mu_1-\mu_2)}}\right) \end{aligned}$$

which should be the expression for power that they derive in Eq.(3.12), the most important differnce being the presence of $\sqrt{n}$ instead of n in the numerator. They also do not have an explicit Berry-Esseen bound dealing with the deviation from normality.

B Remarks for this Appendix

B.1 Taylor Expansion

In all our calculations, we use the Taylor expansion for the function e^{-x} around 0. More specifically, we have

$$\int_u\int_v e^{-\frac{(u-v)^2}{\gamma^2}} p_i(u)q_i(v)dudv = \int_u\int_v \left[1 - \frac{(u-v)^2}{\gamma^2} + \frac{e^{-\frac{\lambda_{uv}(u-v)^2}{\gamma^2}}(u-v)^4}{2\gamma^4}\right] p_i(u)q_i(v)dudv$$

where $\lambda_{uv} \in [0,1]$. The above equality follows from the exact formula for Taylor expansions having exact residuals. Note that

$$e^{-\frac{\lambda_{uv}(u-v)^2}{\gamma^2}} \leq 1.$$

When $\gamma = \Omega(\sqrt{d})$ and fourth moments of the distributions p_i and q_i exist, the above integral becomes

$$\int_u\int_v e^{-\frac{(u-v)^2}{\gamma^2}} p_i(u)q_i(v)dudv = \int_u\int_v \left[1 - \frac{(u-v)^2}{\gamma^2}\right] p_i(u)q_i(v)dudv + o\left(\frac{1}{\gamma^2}\right)$$

Similarly, an higher order expansion can also be obtained by assuming existence of sixth order moments. For ease of exposition, we drop $o(1)$ throughout our calculations. To emphasize this issue, we use $\approx$ symbol in our calculations to indicate that the $o(1)$ term is ignored.

B.2 Independent Coordinates

In our calculations, we assume that the coordinates of x, y are independent and that their central moments are σ^2, μ_3, μ_4. In other words, we use $U = I$ in Assumption 1 to derive expressions in this Appendix. However, this is only for ease of exposition and all our proofs hold even when $U \neq I$. This can be seen from the following argument.

$$\begin{aligned} \|x-y\|^2 &= \|Us + \mu_P - Ut - \mu_Q\|^2 \\ &= \left\|U\left(s - t + U^\top(\mu_P - \mu_Q)\right)\right\|^2 \\ &= \|(s + U^\top\mu_P) - (t + U^\top\mu_Q)\|^2 \\ &= \|s' - t'\|^2. \end{aligned}$$

where $s' = s + U^\top \mu_P$ and $t' = t + U^\top \mu_Q$. Since $U^T \mu_P$ and $U^T \mu_Q$ are just rotated mean vectors, the coordinates of s' and t' are independent (since the coordinates of s, t are independent in assumption A1) and s', t' still have the same central moments as s, t.

Using the above relation, we can rewrite our calculations involving $e^{-\|x-y\|^2/\gamma^2}$ in terms of $e^{-\|s'-t'\|^2/\gamma^2}$. Note that the difference between the means of (the distributions on) x, y is $\|\mu_P - \mu_Q\|^2$ and the that the difference between the means of (the distributions on) s', t' is also $\|U^\top \mu_P - U^\top \mu_Q\|^2 = \|\mu_P - \mu_Q\|^2$ since U is orthogonal. So all the problem parameters remain the same, except we shift from non-independent coordinates for x, y to independent coordinates for s, t.

C Proof of Lemma 1

First note that we can rewrite the population MMD^2 as

$$\begin{aligned}
\mathrm{MMD}^2 &= E_{x,x' \sim P}[k(x,x')] + E_{y,y' \sim Q}[k(y,y')] - 2E_{x \sim P, y \sim Q}[k(x,y)] \\
&= \int_u \int_v e^{-\frac{\|u-v\|^2}{\gamma^2}} p(u)p(v) dv du + \int_u \int_v e^{-\frac{\|u-v\|^2}{\gamma^2}} q(u)q(v) dv du - 2 \int_u \int_v e^{-\frac{\|u-v\|^2}{\gamma^2}} p(u)q(v) dv du
\end{aligned}$$

We calculate each of these integrals in the following manner. Since the coordinates of the P and Q are independent, we have

$$\begin{aligned}
\int_u \int_v e^{-\frac{\|u-v\|^2}{\gamma^2}} p(u)p(v) dv du &= \prod_i \left(\int_u \int_v e^{-\frac{(u-v)^2}{\gamma^2}} p_i(u) p_i(v) dv du \right) \\
&\approx \prod_i \int_u \int_v \left[1 - \frac{(u-v)^2}{\gamma^2} \right] p_i(u) p_i(v) du dv \\
&= \left(1 - \frac{2\sigma^2}{\gamma^2} \right)^d
\end{aligned}$$

The last two steps follow from the fact that the coordinates are independent and definition of the second moments of the distributions p_i and q_i (see Section F.1 of the Appendix). Similarly the corresponding term for distribution Q is

$$\int_u \int_v e^{-\frac{\|u-v\|^2}{\gamma^2}} q(u)q(v) dv du \approx \left(1 - \frac{2\sigma^2}{\gamma^2} \right)^d$$

For the final term, we have

$$\begin{aligned}
\int_u \int_v e^{-\frac{\|u-v\|^2}{\gamma^2}} p(u)q(v) dv du &\approx \prod_i \left(\int_u \int_v \left[1 - \frac{(u-v)^2}{\gamma^2} \right] p_i(u)\, q_i(v)\, du\, dv \right) \\
&= \prod_i \left(\int_u \int_v \left(1 - \frac{(u-\mu_{Pi})^2}{\gamma^2} - \frac{(v-\mu_{Pi})^2}{\gamma^2} + 2\frac{(u-\mu_{Pi})(v-\mu_{Pi})}{\gamma^2} \right) p_i(u) q_i(v) du dv \right) \\
&= \prod_i \left(\int_v \left(1 - \frac{\sigma^2}{\gamma^2} - \frac{(v - \mu_{Qi} + \mu_{Qi} - \mu_{Pi})^2}{\gamma^2} \right) q_i(v) dv \right) \\
&= \prod_i \left(1 - \frac{\sigma^2}{\gamma^2} - \frac{\sigma^2}{\gamma^2} - \frac{\delta_i^2}{\gamma^2} \right)
\end{aligned}$$

The second step follows from since integral. The third step follows from independence of the coordinates. The fourth step follows from taylor expansion. The final few steps follow from the definition of second moment of the distributions (see Section F.1 of the Appendix). Combining the above terms, we have

$$\begin{aligned}
\text{MMD}^2 &\approx \prod_i \left(1 - \frac{2\sigma^2}{\gamma^2}\right) + \prod_i \left(1 - \frac{2\sigma^2}{\gamma^2}\right) - 2\prod_i \left(1 - \frac{\sigma^2}{\gamma^2} - \frac{\sigma^2}{\gamma^2} - \frac{\delta_i^2}{\gamma^2}\right) \\
&\approx 1 - \sum_i \frac{2\sigma^2}{\gamma^2} + 1 - \sum_i \frac{2\sigma^2}{\gamma^2} - 2\left(1 - \sum_i \frac{\sigma^2}{\gamma^2} - \sum_i \frac{\sigma^2}{\gamma^2} - \sum_i \frac{\delta_i^2}{\gamma^2}\right) \\
&= \frac{2\|\delta\|^2}{\gamma^2}
\end{aligned}$$

D Proof of Lemma 2

The variance for the linear time MMD is given by

$$var_{z,z,'}(h(z,z')) = \mathbb{E}_{z,z'}[h^2(z,z')] - (\mathbb{E}_{z,z'}h(z,z'))^2$$

where $h(z,z') = k(x,x') + k(y,y') - k(x,y') - k(x',y)$ where $x, x' \sim P$ and $y, y' \sim Q$ and $\mathbb{E}_{z,z'}[h(z,z')] = MMD^2$. Hence the second term is just $(\mathbb{E}_{z,z'}h(z,z'))^2 = (MMD^2)^2$. Let us concentrate on the first term:

$$\begin{aligned}
\mathbb{E}_{z,z'}[h^2(z,z')] = &\ E_{x,x'\sim P}k^2(x,x') + E_{y,y'\sim Q}k^2(y,y') + 2E_{x\sim P, y\sim Q}k^2(x,y) \\
&+ 2E_{x,x'\sim P, y,y'\sim Q}k(x,x')k(y,y') + 2E_{x,x'\sim P, y,y'\sim Q}k(x,y')k(x',y) \\
&- 4E_{x,x'\sim P, y\sim Q}k(x,x')k(x,y) - 4E_{x\sim P, y,y'\sim Q}k(x,y)k(y,y')
\end{aligned}$$

Hence, there are five kinds of terms to calculate

1. $E_{x,x'\sim P}k^2(x,x')$ (from which $E_{y,y'\sim Q}k^2(y,y')$ can follow)
2. $E_{x\sim P, y\sim Q}k^2(x,y)$
3. $E_{x,x'\sim P, y,y'\sim Q}k(x,x')k(y,y')$
4. $E_{x,x'\sim P, y,y'\sim Q}k(x,y')k(x',y)$
5. $E_{x,x'\sim P, y\sim Q}k(x,x')k(x,y)$ (from which $E_{x\sim P, y,y'\sim Q}k(x,y)k(y,y')$ can follow)

Let us calculate these five terms in order.

D.1 Term 1: $E_{x,x'\sim P}k^2(x,x')$

$$\begin{aligned}
&= \int_{x,x'\sim P} e^{-2\frac{\|x-x'\|^2}{\gamma^2}} p(x)p(x')dxdx' \\
&= \prod_i \left(\int_{x,x'} e^{-2\frac{(x_i - x_i')^2}{\gamma^2}} p_i(x_i)p_i(x_i')dx_i dx_i'\right) \\
&\approx \prod_i \left(1 - \frac{4\sigma^2}{\gamma^2} + \frac{4\mu_4}{\gamma^4} + \frac{12\sigma^4}{\gamma^4}\right) \\
&\approx 1 - \frac{4d\sigma^2}{\gamma^2} + \frac{4d\mu_4}{\gamma^4} + \frac{12d\sigma^4}{\gamma^4} + \frac{8d(d-1)\sigma^4}{\gamma^4}
\end{aligned}$$

The third step follows from our calculations in Section F.1 of the Appendix. Note that the extra terms arise from considering all cross terms with denominator γ^4.

D.2 Term 2: $E_{x\sim P, y\sim Q} k^2(x,y)$

$$
\begin{aligned}
&= \int_{x\sim P}\int_{y'\sim Q} e^{-2\frac{\|x-y'\|^2}{\gamma^2}} p(x)q(y')dx dy' \\
&= \prod_i \int\int e^{-2\frac{(x_i-y'_i)^2}{\gamma^2}} p_i(x_i)q_i(y'_i)dx_i dy'_i \\
&\approx \prod_i \left(1 - \frac{4\sigma^2}{\gamma^2} - \frac{2\delta_i^2}{\gamma^2} + \frac{4\mu_4}{\gamma^4} + \frac{24\sigma^2\delta_i^2}{\gamma^4} + \frac{12\sigma^4}{\gamma^4} + \frac{2\delta_i^4}{\gamma^4}\right) \\
&\approx 1 - \frac{4d\sigma^2}{\gamma^2} - \frac{2\|\delta\|^2}{\gamma^2} + \frac{4d\mu_4}{\gamma^4} + \frac{24\sigma^2\|\delta\|^2}{\gamma^4} + \frac{12d\sigma^4}{\gamma^4} + \frac{2\|\delta\|_4^4}{\gamma^4} + \frac{8d(d-1)\sigma^4}{\gamma^4} + \frac{8(d-1)\sigma^2\|\delta\|^2}{\gamma^4} + \frac{2\|\delta\|^4 - 2\|\delta\|_4^4}{\gamma^4}
\end{aligned}
$$

The third step follows from our calculations in Section F.1 of the Appendix.

D.3 Term 3: $E_{x,x'\sim P, y,y'\sim Q} k(x,x')k(y,y')$

$$
\begin{aligned}
&= \int_{x,x'\sim P}\int_{y,y'\sim Q} e^{-\frac{\|x-x'\|^2}{\gamma^2}} e^{-\frac{\|y-y'\|^2}{\gamma^2}} p(x)p(x')q(y)q(y')dx dx' dy dy' \\
&= \prod_i \int\int e^{-\frac{(x_i-x'_i)^2}{\gamma^2}} p_i(x_i)p_i(x'_i)dx_i dx'_i \prod_i \int\int e^{-\frac{(y_i-y'_i)^2}{\gamma^2}} q_i(y_i)q_i(y'_i)dy_i dy'_i \\
&\approx \prod_i \left(1 - \frac{2\sigma^2}{\gamma^2} + \frac{\mu_4}{\gamma^4} + \frac{3\sigma^4}{\gamma^4}\right)\left(1 - \frac{2\sigma^2}{\gamma^2} + \frac{\mu_4}{\gamma^4} + \frac{3\sigma^4}{\gamma^4}\right) \\
&\approx \prod_i \left(1 - \frac{4\sigma^2}{\gamma^2} + \frac{2\mu_4}{\gamma^4} + \frac{10\sigma^4}{\gamma^4}\right) \\
&\approx 1 - \frac{4d\sigma^2}{\gamma^2} + \frac{2d\mu_4}{\gamma^4} + \frac{10d\sigma^4}{\gamma^4} + \frac{8d(d-1)\sigma^4}{\gamma^4}
\end{aligned}
$$

The third step follows from our calculations in Section F.1 of the Appendix.

D.4 Term 4: $E_{x,x'\sim P, y,y'\sim Q} k(x,y)k(x',y')$

$$
\begin{aligned}
&= \int_{x,x'\sim P}\int_{y,y'\sim Q} e^{-\frac{\|x-y\|^2}{\gamma^2}} e^{-\frac{\|x'-y'\|^2}{\gamma^2}} p(x)p(x')q(y)q(y')dx dx' dy dy' \\
&\approx \prod_i \left(1 - \frac{2\sigma^2}{\gamma^2} - \frac{\delta_i^2}{\gamma^2} + \frac{\mu_4}{\gamma^4} + \frac{6\sigma^2\delta_i^2}{\gamma^4} + \frac{3\sigma^4}{\gamma^4} + \frac{\delta_i^4}{2\gamma^4}\right)^2 \\
&= \prod_i \left(1 - \frac{4\sigma^2}{\gamma^2} - \frac{2\delta_i^2}{\gamma^2} + \frac{2\mu_4}{\gamma^4} + \frac{16\sigma^2\delta_i^2}{\gamma^4} + \frac{10\sigma^4}{\gamma^4} + \frac{2\delta_i^4}{\gamma^4}\right) \\
&\approx 1 - \frac{4d\sigma^2}{\gamma^2} - \frac{2\|\delta\|^2}{\gamma^2} + \frac{2d\mu_4}{\gamma^4} + \frac{16\sigma^2\|\delta\|^2}{\gamma^4} + \frac{10d\sigma^4}{\gamma^4} + \frac{2\|\delta\|_4^4}{\gamma^4} + \frac{8d(d-1)\sigma^4}{\gamma^4} + \frac{2\|\delta\|^4 - 2\|\delta\|_4^4}{\gamma^4} + \frac{8(d-1)\sigma^2\|\delta\|^2}{\gamma^4}
\end{aligned}
$$

The third step follows from our calculations in Section F.1 of the Appendix.

D.5 Term 5: $E_{x,x'\sim P,y\sim Q}k(x,x')k(x,y)$

$$
\begin{aligned}
&= \int_{x,x'\sim P}\int_{y\sim Q} e^{-\frac{\|x-x'\|^2}{\gamma^2}} e^{-\frac{\|x'-y\|^2}{\gamma^2}} p(x)p(x')q(y)dxdx'dy \\
&= \prod_i \left(\int\int e^{-\frac{\|x_i-x_i'\|^2}{\gamma^2}} e^{-\frac{\|x_i'-y_i\|^2}{\gamma^2}} p_i(x_i)p_i(x_i')q(y_i)\right) \\
&\approx \prod_i \left(1 - \frac{4\sigma^2}{\gamma^2} - \frac{\delta_i^2}{\gamma^2} + \frac{3\mu_4}{\gamma^4} + \frac{8\sigma^2\delta_i^2}{\gamma^4} + \frac{9\sigma^4}{\gamma^4} + \frac{\delta^4}{2\gamma^4} + \frac{2\mu_3\delta_i}{\gamma^4}\right) \\
&\approx 1 - \frac{4d\sigma^2}{\gamma^2} - \frac{\|\delta\|^2}{\gamma^2} + \frac{3d\mu_4}{\gamma^4} + \frac{8\sigma^2\|\delta\|^2}{\gamma^4} + \frac{9d\sigma^4}{\gamma^4} + \cancel{\frac{\|\delta\|_4^4}{2\gamma^4}} + \frac{2\mu_3\sum_i\delta_i}{\gamma^4} + \frac{8d(d-1)\sigma^4}{\gamma^4} + \frac{4(d-1)\sigma^2\|\delta\|^2}{\gamma^4} + \frac{\|\delta\|^4 \cancel{-\|\delta\|_4^4}}{2\gamma^4}
\end{aligned}
$$

The second step follows from our calculations in Section F.2 of the Appendix. Combining the all the terms above, we get the following bound on the variance.

D.6 The bound on $\mathbb{E}_{z,z'}[h^2(z,z')]$

$$
\begin{aligned}
&\approx E_{x,x'\sim P}k^2(x,x') + E_{y,y'\sim Q}k^2(y,y') + 2E_{x\sim P,y\sim Q}k^2(x,y) \\
&\quad + 2E_{x,x'\sim P,y,y'\sim Q}k(x,x')k(y,y') + 2E_{x,x'\sim P,y,y'\sim Q}k(x,y')k(x',y) \\
&\quad - 4E_{x,x'\sim P,y\sim Q}k(x,x')k(x,y) - 4E_{x\sim P,y,y'\sim Q}k(x,y)k(y,y') \\
&= \left(\cancel{1 - \frac{4d\sigma^2}{\gamma^2}} + \cancel{\frac{4d\mu_4}{\gamma^4}} + \boxed{\frac{12d\sigma^4}{\gamma^4}} + \cancel{\frac{8d(d-1)\sigma^4}{\gamma^4}}\right) \\
&\quad + \left(\cancel{1 - \frac{4d\sigma^2}{\gamma^2}} + \cancel{\frac{4d\mu_4}{\gamma^4}} + \boxed{\frac{12d\sigma^4}{\gamma^4}} + \cancel{\frac{8d(d-1)\sigma^4}{\gamma^4}}\right) \\
&\quad + 2\left(\cancel{1 - \frac{4d\sigma^2}{\gamma^2} - \frac{2\|\delta\|^2}{\gamma^2}} + \cancel{\frac{4d\mu_4}{\gamma^4}} + \boxed{\frac{24\sigma^2\|\delta\|^2}{\gamma^4}} + \boxed{\frac{12d\sigma^4}{\gamma^4}} + \cancel{\frac{8d(d-1)\sigma^4}{\gamma^4}} + \cancel{\frac{8(d-1)\sigma^2\|\delta\|^2}{\gamma^4}} + \boxed{\frac{2\|\delta\|^4}{\gamma^4}}\right) \\
&\quad + 2\left(\cancel{1 - \frac{4d\sigma^2}{\gamma^2}} + \cancel{\frac{2d\mu_4}{\gamma^4}} + \boxed{\frac{10d\sigma^4}{\gamma^4}} + \cancel{\frac{8d(d-1)\sigma^4}{\gamma^4}}\right) \\
&\quad + 2\left(\cancel{1 - \frac{4d\mu_2}{\gamma^2} - \frac{2\|\delta\|^2}{\gamma^2}} + \cancel{\frac{2d\mu_4}{\gamma^4}} + \boxed{\frac{16\sigma^2\|\delta\|^2}{\gamma^4}} + \boxed{\frac{10d\sigma^4}{\gamma^4}} + \cancel{\frac{8d(d-1)\sigma^4}{\gamma^4}} + \boxed{\frac{2\|\delta\|^4}{\gamma^4}} + \cancel{\frac{8(d-1)\sigma^2\|\delta\|^2}{\gamma^4}}\right) \\
&\quad - 4\left(\cancel{1 - \frac{4d\sigma^2}{\gamma^2} - \frac{\|\delta\|^2}{\gamma^2}} + \cancel{\frac{3d\mu_4}{\gamma^4}} + \boxed{\frac{8\sigma^2\|\delta\|^2}{\gamma^4}} + \boxed{\frac{9d\sigma^4}{\gamma^4}} + \cancel{\frac{2\mu_3\sum_i\delta_i}{\gamma^4}} + \cancel{\frac{8d(d-1)\sigma^4}{\gamma^4}} + \cancel{\frac{4(d-1)\sigma^2\|\delta\|^2}{\gamma^4}} + \boxed{\frac{\|\delta\|^4}{2\gamma^4}}\right) \\
&\quad - 4\left(\cancel{1 - \frac{4d\sigma^2}{\gamma^2} - \frac{\|\delta\|^2}{\gamma^2}} + \cancel{\frac{3d\mu_4}{\gamma^4}} + \boxed{\frac{8\sigma^2\|\delta\|^2}{\gamma^4}} + \boxed{\frac{9d\sigma^4}{\gamma^4}} - \cancel{\frac{2d\mu_3\sum_i\delta_i}{\gamma^4}} + \cancel{\frac{8d(d-1)\sigma^4}{\gamma^4}} + \cancel{\frac{4(d-1)\sigma^2\|\delta\|^2}{\gamma^4}} + \boxed{\frac{\|\delta\|^4}{2\gamma^4}}\right) \\
&= \boxed{\frac{4\|\delta\|^4}{\gamma^4} + \frac{16d\sigma^4}{\gamma^4} + \frac{16\sigma^2\|\delta\|^2}{\gamma^4}}
\end{aligned}
$$

Finally, using the bound derived above on $\mathbb{E}_{z,z'}[h^2(z,z')]$, the bound on variance is

$$\mathrm{var}_{z,z,'}(h(z,z')) = \mathbb{E}_{z,z'}[h^2(z,z')] - (\mathbb{E}_{z,z'}h(z,z'))^2 = \frac{16d\sigma^4}{\gamma^4} + \frac{16\sigma^2\|\delta\|^2}{\gamma^4}.$$

E Proof of Lemma 3

E.1 Upper bound on τ_4

We derive the upper bound on τ_4 in this section. An upper bound on $E_{z,z'}[(h(z,z') - E_{z,z'}[h(z,z')])^4]$ can be obtain in the following manner. First note that

$$E_{z,z'}[(h(z,z') - E_{z,z'}[h(z,z')])^4] = E[h^4(z,z')] - 3(\mathrm{MMD}^2)^4 - 4E[h^3(z,z')]\mathrm{MMD}^2 + 6E[h^2(z,z')](\mathrm{MMD}^2)^2$$
$$= 16\kappa_4 - \frac{48\|\delta\|^8}{\gamma^8} - 64\kappa_3\frac{\|\delta\|^2}{\gamma^2} + \frac{96\|\delta\|^8}{\gamma^8} + \frac{384d\sigma^4\|\delta\|^4}{\gamma^8} + \frac{384\sigma^2\|\delta\|^6}{\gamma^8}$$

where $\kappa_4 = E[h^4(z,z')]$ and $\kappa_3 = E[h^3(z,z')]$.

Calculations for κ_4

We now calculate an upper bound to $E_{z,z'}[h^4(z,z')]$ in the following manner. With slight abuse of notation, we use x_i to denote the i^{th} coordinate of x. We first note that

$$\begin{aligned}
E_{z,z'}[h^4(z,z')] &= E_{z,z'}[k(x,x') + k(y,y') - k(x,y') - k(x',y)]^4 \\
&\approx E_{z,z'}\left[1 - \frac{\|x-x'\|^2}{\gamma^2} + 1 - \frac{\|y-y'\|^2}{\gamma^2} - 1 + \frac{\|x-y'\|^2}{\gamma^2} - 1 + \frac{\|x'-y'\|^2}{\gamma^2}\right]^4 \\
&= 16E_{z,z'}\left[\frac{(x^\top x' + y^\top y' - x^\top y' - x'^\top y)}{\gamma^2}\right]^4 \\
&= 16E_{z,z'}\left[\frac{\sum_{j=1}^d (x_j - y_j)(x'_j - y'_j)}{\gamma^2}\right]^4 \\
&= \frac{16}{\gamma^8}E_{z,z'}\left[\sum_{k_1+\cdots+k_d=4}\binom{4}{k_1\cdots k_d}\prod_{1\le i\le d}(x_i-y_i)^{k_i}(x'_i-y'_i)^{k_i}\right] \\
&= \frac{16}{\gamma^8}\sum_{k_1+\cdots+k_d=4}\binom{4}{k_1\cdots k_d}\prod_{1\le i\le d}(E_z[(x_i-y_i)^{k_i}])^2
\end{aligned}$$

The above summation splits into five different sums, based on the different ways to write $k_1 + \cdots + k_d = 4$ - we derive these terms using the calculations in Section F.1 and Section F.2, as well as some terms from the Variance calculations in Section D, and explain in brackets which way to sum the k_is to 4 was used.

$$\begin{aligned}
\kappa \quad = \quad & \frac{1}{\gamma^8}\sum_i [2\mu_4 + 12\sigma^2\delta_i^2 + 6\sigma^4 + \delta_i^4]^2 \quad \text{(using (4,0,0...))} \\
+ \quad & \frac{4}{\gamma^8}\sum_{i\ne j}(\delta_i^3 + 6\sigma^2\delta_i)^2\delta_j^2 \quad \text{(using (3,1,0,0...))} \\
+ \quad & \frac{3}{\gamma^8}\sum_{i\ne j}(4\sigma^4 + \delta_i^4 + 4\sigma^2\delta_i^2)(4\sigma^4 + \delta_j^4 + 4\sigma^2\delta_j^2) \quad \text{(using (2,2,0,0...))} \\
+ \quad & \frac{6}{\gamma^8}\sum_{i\ne j\ne k}(4\sigma^4 + \delta_i^4 + 4\sigma^2\delta_i^2)\delta_j^2\delta_k^2 \quad \text{(using (2,1,1,0,0...))} \\
+ \quad & \frac{1}{\gamma^8}\sum_{i\ne j\ne k\ne l}\delta_i^2\delta_j^2\delta_k^2\delta_l^2 \quad \text{(using (1,1,1,1,0,0...))}
\end{aligned}$$

Expanding the each of the above terms further, we get

Term 1: $\frac{1}{\gamma^8}\Bigg[4d\mu_4^2 + (144+12)\sigma^4\sum_i \delta_i^4 + 36\sigma^8 d + \sum_i \delta_i^8$

$+ 48\mu_4\sigma^2\|\delta\|^2 + 24d\mu_4\sigma^4 + 4\mu_4\sum_i \delta_i^4 + 144\sigma^6\|\delta\|^2 + 24\sigma^2\sum_i \delta_i^6\Bigg]$

Term 2: $\frac{4}{\gamma^8}\Bigg[\sum_{i\neq j}\delta_i^6\delta_j^2 + 36\sigma^4\sum_{i\neq j}\delta_i^2\delta_j^2 + 12\sigma^2\sum_{i\neq j}\delta_i^4\delta_j^2\Bigg]$

Term 3: $\frac{3}{\gamma^8}\Bigg[8d(d-1)\sigma^8 + \sum_{i\neq j}\delta_i^4\delta_j^4 + 8\sigma^4(d-1)\sum_i \delta_i^4 + 32\sigma^6\|\delta\|^2(d-1) + 8\sigma^2\sum_{i\neq j}\delta_i^4\delta_j^2 + 16\sigma^4\sum_{i\neq j}\delta_i^2\delta_j^2\Bigg]$

Term 4: $\frac{6}{\gamma^8}\Bigg[4\sigma^4(d-2)\sum_{i\neq j}\delta_i^2\delta_j^2 + \sum_{i\neq j\neq k}\delta_i^4\delta_j^2\delta_k^2 + 4\sigma^2\sum_{i\neq j\neq k}\delta_i^2\delta_j^2\delta_k^2\Bigg]$

Term 5: $\frac{1}{\gamma^8}\Bigg[\sum_{i\neq j\neq k\neq l}\delta_i^2\delta_j^2\delta_k^2\delta_l^2\Bigg]$

Calculations for κ_3

Similar to the multinomial expansion for κ_4, we have

$$\begin{aligned}
\kappa_3 &= \frac{1}{\gamma^6}\sum_i(\delta_i^6 + 36\sigma^4\delta_i^2 + 12\sigma^2\delta_i^4) \quad \text{(using (3,0,0,0...))}\\
&+ \frac{3}{\gamma^6}\sum_{i\neq j}(4\sigma^4 + \delta_i^4 + 4\sigma^2\delta_i^2)\delta_j^2 \quad \text{(using (2,1,0,0...))}\\
&+ \frac{1}{\gamma^6}\sum_{i\neq j\neq k}\delta_i^2\delta_j^2\delta_k^2 \quad \text{(using (1,1,1,0,0...))}
\end{aligned}$$

Using the above expansion, we get

$$\begin{aligned}
\kappa_3\frac{\|\delta\|^2}{\gamma^2} &= \frac{1}{\gamma^8}\sum_{i\neq j}(\delta_i^6\delta_j^2 + 36\sigma^4\delta_i^2\delta_j^2 + 12\sigma^2\delta_i^4\delta_j^2)\\
&+ \frac{1}{\gamma^8}\sum_i(\delta_i^8 + 36\sigma^4\delta_i^4 + 12\sigma^2\delta_i^6)\\
&+ \frac{3}{\gamma^8}\sum_{i\neq j\neq k}(4\sigma^4\delta_j^2\delta_k^2 + \delta_i^4\delta_j^2\delta_k^2 + 4\sigma^2\delta_i^2\delta_j^2\delta_k^2)\\
&+ \frac{3}{\gamma^8}\sum_{i\neq j}(4\sigma^4\delta_i^2\delta_j^2 + \delta_i^6\delta_j^2 + 4\sigma^2\delta_i^4\delta_j^2)\\
&+ \frac{3}{\gamma^8}\sum_{i\neq j}(4\sigma^4\delta_j^4 + \delta_i^4\delta_j^4 + 4\sigma^2\delta_i^2\delta_j^4)\\
&+ \frac{1}{\gamma^8}\sum_{i\neq j\neq k\neq l}\delta_i^2\delta_j^2\delta_k^2\delta_l^2 + \frac{3}{\gamma^8}\sum_{i\neq j\neq k}\delta_i^4\delta_j^2\delta_k^2
\end{aligned}$$

Also note the following expansions of $\|\delta\|^8$ and $\|\delta\|^6$.

$$\frac{\|\delta\|^8}{\gamma^8} = \sum_i \delta_i^8 + 4\sum_{i\neq j} \delta_i^6\delta_j^2 + 3\sum_{i\neq j}\delta_i^4\delta_j^4 + 6\sum_{i\neq j\neq k}\delta_i^4\delta_j^2\delta_k^2 + \sum_{i\neq j\neq k\neq l}\delta_i^2\delta_j^2\delta_k^2\delta_l^2$$
$$\frac{\|\delta\|^6}{\gamma^6} = \sum_i \delta_i^6 + 3\sum_{i\neq j}\delta_i^4\delta_j^2 + \sum_{i\neq j\neq k}\delta_i^2\delta_j^2\delta_k^2$$

Putting all terms together

Using the above calculations for κ_3 and κ_4, we obtain the following bound on $E_{z,z'}[(h(z,z') - E_{z,z'}[h(z,z')])^4]$.

$$\begin{aligned}
E_{z,z'}[(h(z,z') - E_{z,z'}[h(z,z')])^4] &= E[h^4(z,z')] - 3(\text{MMD}^2)^4 - 4E[h^3(z,z')]\text{MMD}^2 + 6E[h^2(z,z')](\text{MMD}^2)^2 \\
&= 16\kappa_4 - \frac{48\|\delta\|^8}{\gamma^8} - 64\kappa_3\frac{\|\delta\|^2}{\gamma^2} + \frac{96\|\delta\|^8}{\gamma^8} + \frac{384d\sigma^4\|\delta\|^4}{\gamma^8} + \frac{384\sigma^2\|\delta\|^6}{\gamma^8} \\
&= \frac{16}{\gamma^2}\Bigg(4d\mu_4^2 + 36\sigma^8 d + 24d\mu_4\sigma^4 + 24d(d-1)\sigma^8 + (96d\sigma^6 + 48\sigma^6 + 48\mu_4\sigma^2)\sum_i \delta_i^2 \\
&\qquad + (132\sigma^4 + 4\mu_4)\sum_i \delta_i^4 + 144\sigma^4\sum_{i\neq j}\delta_i^2\delta_j^2\Bigg) - 64\left(24\sigma^4\sum_i\delta_i^4 + 24\sigma^4\sum_{i\neq j}\delta_i^2\delta_j^2\right) \\
&= \frac{1}{\gamma^8}\Bigg(64d\mu_4^2 + 576\sigma^8 d + 384d\mu_4\sigma^4 + 384d(d-1)\sigma^8 + (1536d\sigma^6 + 768\sigma^6 + 768\mu_4\sigma^2)\sum_i\delta_i^2 \\
&\qquad + (576\sigma^4 + 64\mu_4)\sum_i\delta_i^4 + 768\sigma^4\sum_{i\neq j}\delta_i^2\delta_j^2\Bigg) \\
&= (3 + o(1))V^2
\end{aligned}$$

where we substituted κ_4, κ_3 in the third equationand the $\|\delta\|^6$ and $\|\delta\|^8$ terms perfectly cancel out.

F Helpful Calculations for Lemma 1, 2, 3, 4

F.1 Double Integrals

$$\begin{aligned}
\int_u\int_v e^{-\frac{(u-v)^2}{\gamma^2}} f(u)g(v)dudv &\approx \int_u\int_v \left[1 - \frac{(u-v)^2}{\gamma^2} + \frac{(u-v)^4}{2\gamma^4}\right] f(u)g(v)dudv \\
&= 1 - \frac{2\sigma^2}{\gamma^2} - \frac{\delta^2}{\gamma^2} + \frac{\mu_4}{\gamma^4} + \frac{6\sigma^2\delta^2}{\gamma^4} + \frac{3\sigma^4}{\gamma^4} + \frac{\delta^4}{2\gamma^4}
\end{aligned}$$

because $\int\int \frac{(u-v)^2}{\gamma^2} f(u)g(v)dudv$

$$\begin{aligned}
&= \int\int \frac{((u-\mu_f) - (v-\mu_f))^2}{\gamma^2} f(u)g(v)dudv \\
&= \int\left(\frac{\sigma^2}{\gamma^2} + \frac{(v-\mu_f)^2}{\gamma^2}\right) g(v)dv = \frac{2\sigma^2}{\gamma^2} + \frac{\delta^2}{\gamma^2}
\end{aligned}$$

and $\int_u \int_v \frac{(u-v)^4}{\gamma^4} f(u)g(v)dvdu$

$$
\begin{aligned}
&= \int_u \int_v \frac{((u-\mu_f)-(v-\mu_f))^4}{\gamma^4} f(u)g(v)dvdu \\
&= \int_u \int_v \left(\frac{(u-\mu_f)^4}{\gamma^4} + \frac{(v-\mu_f)^4}{\gamma^4} - \frac{4(u-\mu_f)^3(v-\mu_f)}{\gamma^4} - \frac{4(u-\mu_f)(v-\mu_f)^3}{\gamma^4} + \frac{6(u-\mu_f)^2(v-\mu_f)^2}{\gamma^4} \right) f(u)g(v)dudv \\
&= \int_v \left(\frac{\mu_4}{\gamma^4} + \frac{(v-\mu_f)^4}{\gamma^4} - \frac{4\mu_3(v-\mu_f)}{\gamma^4} + \frac{6\sigma^2(v-\mu_f)^2}{\gamma^4} \right) g(v)dv \\
&= \frac{\mu_4}{\gamma^4} + \left[\frac{\mu_4}{\gamma^4} - \frac{4\mu_3\delta}{\gamma^4} + \frac{6\sigma^2\delta^2}{\gamma^4} + \frac{\delta^4}{\gamma^4} \right]_1 + \left[\frac{4\mu_3\delta}{\gamma^4} \right]_2 + \left[\frac{6\sigma^2(\sigma^2+\delta^2)}{\gamma^4} \right]_3 \\
&= \frac{2\mu_4}{\gamma^4} + \frac{12\sigma^2\delta^2}{\gamma^4} + \frac{6\sigma^4}{\gamma^4} + \frac{\delta^4}{\gamma^4}
\end{aligned}
$$

Finally, we have

$$
\begin{aligned}
\int_u \int_v (u-v)^3 f(u)g(v)dvdu &= \int_u \int_v ((u-\mu_f)-(v-\mu_f))^3 f(u)g(v)dudv \\
&= \int_v \mu_3 - 3\sigma^2(v-\mu_f) - (v-\mu_f)^3 g(v)dv \\
&= 3\sigma^2\delta + \delta^3 + 3\sigma^2\delta = \delta^3 + 6\sigma^2\delta.
\end{aligned}
$$

F.2 Triple Integral

$$
\begin{aligned}
&\int_u \int_v \int_y e^{-\frac{(u-v)^2}{\gamma^2}} e^{-\frac{(v-y)^2}{\gamma^2}} f(u)g(v)g(y)dudvdy \\
&= \int_u \int_v \int_y \left[1 - \frac{(u-v)^2}{\gamma^2} + \frac{(u-v)^4}{2\gamma^4}\right] \left[1 - \frac{(v-y)^2}{\gamma^2} + \frac{(v-y)^4}{2\gamma^4}\right] f(u)g(v)g(y)dudvdy \\
&= \int_u \int_v \int_y \left[1 - \frac{(u-v)^2}{\gamma^2} - \frac{(v-y)^2}{\gamma^2} + \frac{(u-v)^4}{2\gamma^4} + \frac{(v-y)^4}{2\gamma^4} + \frac{(u-v)^2(v-y)^2}{\gamma^4}\right] f(u)g(v)g(y)dudvdy \\
&= 1 - \frac{2\sigma^2}{\gamma^2} - \frac{\delta^2}{\gamma^2} - \frac{2\sigma^2}{\gamma^2} + \frac{1}{2}\left[\frac{2\mu_4}{\gamma^4} + \frac{12\sigma^2\delta^2}{\gamma^4} + \frac{6\sigma^4}{\gamma^4} + \frac{\delta^4}{\gamma^4}\right] + \frac{1}{2}\left[\frac{2\mu_4}{\gamma^4} + \frac{6\sigma^4}{\gamma^4}\right] \\
&\quad + \frac{1}{\gamma^4}\int_v \left[\sigma^2 + (v-\mu_f)^2\right]\left[\sigma^2 + (v-\mu_g)^2\right] g(v)dv \\
&= 1 - \frac{2\sigma^2}{\gamma^2} - \frac{\delta^2}{\gamma^2} - \frac{2\sigma^2}{\gamma^2} + \frac{1}{2}\left[\frac{2\mu_4}{\gamma^4} + \frac{12\sigma^2\delta^2}{\gamma^4} + \frac{6\sigma^4}{\gamma^4} + \frac{\delta^4}{\gamma^4}\right] + \frac{1}{2}\left[\frac{2\mu_4}{\gamma^4} + \frac{6\sigma^4}{\gamma^4}\right] \\
&\quad + \frac{3\sigma^4}{\gamma^4} + \frac{2\sigma^2\delta^2}{\gamma^4} + \frac{\mu_4}{\gamma^4} - \frac{2\mu_3\delta}{\gamma^4} \\
&= 1 - \frac{4\sigma^2}{\gamma^2} - \frac{\delta^2}{\gamma^2} + \frac{3\mu_4}{\gamma^4} + \frac{8\sigma^2\delta^2}{\gamma^4} + \frac{9\sigma^4}{\gamma^4} + \frac{\delta^4}{2\gamma^4} - \frac{2\mu_3\delta}{\gamma^4}
\end{aligned}
$$

The last equality is obtained from the following:

$$
\int_v \left[\sigma^2 + (v-\mu_f)^2\right]\left[\sigma^2 + (v-\mu_g)^2\right] g(v)dv = \sigma^4 + \sigma^2(\sigma^2 + (\mu_g - \mu_f)^2) + \sigma^4 + \int_v (v-\mu_f)^2(v-\mu_g)^2 g(v)dv
$$

G Additional Experiments

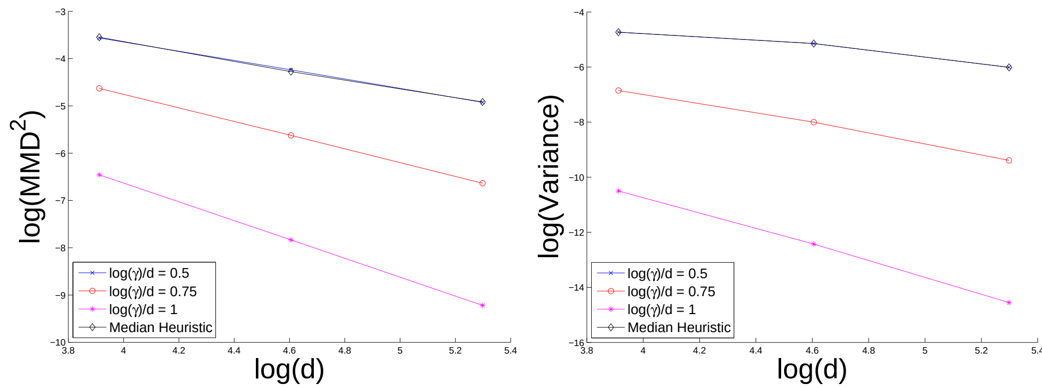

Figure 5: A plot for MMD^2 and Variance of linear statistic, when $n = 1000$ for Normal distribution with identity covariance and $\Psi = 1$, for bandwidths $d^{0.5}, d, d^{0.75}$. Note that these plots provide empirical verification for Lemma 1 and Lemma 2

.