APPENDIX A
PROOF OF LEMMA 1

Case 1. f is a Lipschitz function. For any $\epsilon > 0$, there is a constant $r > 0$ such that $2lr < \epsilon$, for any $x \in \mathbb{R}^d$, where l is the Lipschitz constant. Hence

$$\frac{1}{\mu(B(x,r_x))}\int_{B(x,r_x)} |f(y)-f(x)|d\mu(y) < 2lr_x < \epsilon,$$

which means that Equation (2) is obtained. For any $x \in A$, $r_x \leq r$,

$$\frac{\int_{\mathbb{R}^d} |f(x')-f(x)|k(x',x;r_x,m)d\mu(x')}{\int_{\mathbb{R}^d} k(x',x;r_x,m)d\mu(x')} \leq 2lCr_x < C\epsilon,$$

where

$$C = \frac{\int_{\mathbb{R}^d} |x|e^{-|x|^m}d\mu(x)}{\int_{\mathbb{R}^d} e^{-|x|^m}d\mu(x)}.$$

We obtain that

$$\lim_{r\to 0}\sup_{x\in A}|\frac{1}{\mu(B(x,r_x))}\int_{B(x,r_x)} fd\mu - f(x)| = 0,$$
$$\lim_{r\to 0}\sup_{x\in A}|\frac{\int_{\mathbb{R}^d}(f(x')-f(x))k(x',x;r_x,m)d\mu(x')}{\int_{\mathbb{R}^d} k(x',x;r_x,m)d\mu(x')}| = 0.$$

Therefore,

$$\lim_{r\to 0}\sup_{x\in A}|\frac{1}{\mu(B(x,r_x))}\int_{B(x,r_x)} fd\mu|$$
$$= \sup_{x\in A}|f(x)| = \sup_{x\in A}\lim_{r\to 0}|\frac{1}{\mu(B(x,r_x))}\int_{B(x,r_x)} fd\mu|,$$

and

$$\lim_{r\to 0}\sup_{x\in A}|\frac{\int_{\mathbb{R}^d} f(x')k(x',x;r_x,m)d\mu(x')}{\int_{\mathbb{R}^d} k(x',x;r_x,m)d\mu(x')}|$$
$$= \sup_{x\in A}|f(x)| = \lim_{r\to 0}\sup_{x\in A}|\frac{1}{\mu(B(x,r_x))}\int_{B(x,r_x)} fd\mu|,$$

which implies that

$$\sup_{x\in A}\lim_{r\to 0}\frac{1}{\mu(B(x,r_x))}\int_{B(x,r_x)} fd\mu$$
$$= \lim_{r\to 0}\sup_{x\in A}\frac{1}{\mu(B(x,r_x))}\int_{B(x,r_x)} fd\mu$$
$$= \lim_{r\to 0}\sup_{x\in A}\frac{\int_{\mathbb{R}^d} f(x')k(x',x;r_x,m)d\mu(x')}{\int_{\mathbb{R}^d} k(x',x;r_x,m)d\mu(x')}.$$

Case 2. f is a continuous function with a compact support set. We can construct sequence $\{f_n\}_{n=1}^{\infty}$, where f_n is a Lipschitz function, such that $\sup_{x\in\mathbb{R}^d}|f_n(x)-f(x)| \to 0$, when $n \to \infty$. Based on Case 1, f_n satisfies Equation (3).

$$\begin{aligned}&\sup_{x\in A}\lim_{r\to 0}\left|\frac{1}{\mu(B(x,r_x))}\int_{B(x,r_x)} f_n d\mu\right|\\ &= \lim_{r\to 0}\sup_{x\in A}\left|\frac{\int_{\mathbb{R}^d} f_n(x')k(x',x;r_x,m)d\mu(x')}{\int_{\mathbb{R}^d} k(x',x;r_x,m)d\mu(x')}\right|\\ &= \sup_{x\in A}|f_n(x)|,\end{aligned} \tag{24}$$

If we take $n \to \infty$ for above equation, we obtain the result. This lemma is proved.

Appendix B
Proof of Theorem 2

Based on Lemma 1 and the definition of L_d, we have

$$\begin{aligned} &\sup_x |f_p(x) - f_q(x)| \\ =& \sup_x \lim_{r\to 0} \left| \frac{\int_{B(x,r_x)} f_p(x') - f_q(x')dx'}{\int_{B(x,r_x)} dx'} \right| \\ =& \lim_{r\to 0} \sup_x |L_d(x)|. \end{aligned} \tag{25}$$

Equation (25) is obtained by using Equation (3). It directly points out the relation between $L_d(x)$ and the maximum density discrepancy between p and q. Thus, the last term in Equation (25) can be used to measure the discrepancy between p and q. However, we cannot calculate $\sup_x |L_d(x)|$ because the maximum value of $|L_d(x)|$ cannot be estimated using observations. Because $\sup_x |\cdot|$ can be replaced with $\|\cdot\|_{L^\infty}$, based on Lemma 1, we arrive at the following equations.

$$\begin{aligned} &\sup_x |f_p(x) - f_q(x)| \\ =& \sup_{x\in \mathrm{supp}(p+q)} |f_p(x) - f_q(x)| \\ =& \lim_{r\to 0} \sup_{x\in \mathrm{supp}(p+q)} |L_d(x)| \\ =& \lim_{r\to 0} \|L_d(x)\|_{L^\infty(\mu^+)}, \end{aligned}$$

where μ^+ is $p+q$.

Then, we prove the following lemma to demonstrate the relation between $L^{\tilde{p}}$-norm and L^∞-norm.

Lemma 4. *Let μ be a finite measure in $\mathbb{R}^d$, Then we have the following equality.*

$$\lim_{\tilde{p}\to\infty} \|f\|_{L^{\tilde{p}}(\mu)} = \|f\|_{L^\infty(\mu)}, \tag{26}$$

where f is a μ-measurable function. Moreover, if $\mu(\mathbb{R}^d) = 2$, then for $\tilde{p} > \log(2)/\log(c+1)$,

$$|\|f\|_{L^{\tilde{p}}(\mu)} - \|f\|_{L^\infty(\mu)}| \le \|f\|_{L^\infty(\mu)} - M'\mu(A)^{1/\tilde{p}},$$

where c is a positive constant $(c \le 1)$ to make sure $\mu(A) \le 1$, here $M' \le \|f\|_{L^\infty(\mu)}(1-c)$, $A = \{x \in \mathbb{R}^d : f \ge M'\}$.

Proof. **Case 1.** We assume $M = \|f\|_{L^\infty}$ and $A = \{x \in \mathbb{R}^d : f \ge M'\}$, where $M' < M$. Then

$$\|f\|_{L^{\tilde{p}}(\mu)} = \Big(\int_{\mathbb{R}^d} |f|^{\tilde{p}} d\mu\Big)^{1/\tilde{p}} \ge M'(\mu(A))^{1/\tilde{p}},$$

which means that

$$\liminf_{\tilde{p}\to\infty} \|f\|_{L^{\tilde{p}}(\mu)} \ge M'.$$

Let $M' \to M$, then

$$\liminf_{\tilde{p}\to\infty} \|f\|_{L^{\tilde{p}}(\mu)} \ge M.$$

Because $\|f\|_{L^{\tilde{p}}(\mu)} \le (\int_{\mathbb{R}^d} M^{\tilde{p}} d\mu)^{1/\tilde{p}} = M(\mu(\mathbb{R}^d))^{1/\tilde{p}}$, we obtain that

$$\limsup_{\tilde{p}\to\infty} \|f\|_{L^{\tilde{p}}(\mu)} \le M.$$

Therefore, we get that

$$\lim_{\tilde{p}\to\infty} \|f\|_{L^{\tilde{p}}(\mu)} = \|f\|_{L^\infty(\mu)}.$$

Case 2. When $\mu(\mathbb{R}^d) = 2$ and $\tilde{p} > \log(2)/\log(c+1)$, we obtain that $\|f\|_{L^\infty}(2^{1/\tilde{p}} - 1) < c\|f\|_{L^\infty} \le \|f\|_{L^\infty(\mu)} - M'\mu(A)^{1/\tilde{p}}$, where c is a positive constant $(c \le 1)$, $M' \le \|f\|_{L^\infty(\mu)}(1-c)$ and $A = \{x \in \mathbb{R}^d : f \ge M'\}$.

Because

$$|\|f\|_{L^{\tilde{p}}(\mu)} - \|f\|_{L^\infty(\mu)}| \le \max\{\|f\|_{L^\infty}(2^{1/\tilde{p}} - 1), \|f\|_{L^\infty(\mu)} - M'\mu(A)^{1/\tilde{p}}\}.$$

Therefore,

$$|\|f\|_{L^{\tilde{p}}(\mu)} - \|f\|_{L^\infty(\mu)}| \le \|f\|_{L^\infty(\mu)} - M'\mu(A)^{1/\tilde{p}}.$$

This lemma is proved. □

Based on Lemma 4, we have the following equation, which is an equivalent form for Equation (6),

$$\sup_x |f_p(x) - f_q(x)| = \lim_{r\to 0} \lim_{\tilde{p}\to+\infty} \|L_d(x)\|_{L^{\tilde{p}}(\mu^+)},$$

where μ^+ is $p+q$. This theorem is proved.

APPENDIX C
PROOF OF THEOREM 3

We begin by considering the simple situation that $d=1$ and $m=1$. Under this situation, the Formula (20) has an equivalent form:

$$T \circ h = 0 \Rightarrow \int_{-\infty}^{+\infty} (x')^n dh(x') = 0, \forall n \in \mathbb{Z}_{\geq 0}, \tag{27}$$

where $\mathbb{Z}_{\geq 0} = \{0, 1, 2, ...\}$. If r_x is a constant for any $x \in \mathbb{R}^d$, then the Formula (27) is proved using:

$$\lim_{x\to\infty} \frac{\partial^n (T \circ h)(x)}{\partial x^n} = 0 \Rightarrow \int_{-\infty}^{+\infty} (x')^n dh(x') = 0.$$

Thus, we are motivated to calculate the n^{th} derivative of the function $(T \circ h)(x)$ with respect to x to prove the Formula (27) under the simple situation. In the rest of this subsection, we denote $k(x', x; r_x, m)$ by $exp(b_m(x', x))$, and $K_m^{(n)}$ by n^{th} derivatives $k(x', x; r_x, m)$ with respect to x, and $b_m^{(n)}$ by n^{th} derivatives $b_m(x', x)$ with respect to x. Based on the definitions of $k(x', x; r_x, m)$ and r_x, we have

$$\begin{aligned} b_1(x', x) &= -\frac{(x'-x)^2}{r_x^2} \\ &= -\frac{(x'-x)^2}{\beta \int_{-\infty}^{+\infty} (x'-x)^2 d_{p_{x'}+q_{x'}}}. \end{aligned}$$

Denote m_p^1 and m_p^2 by the first moment and the second moment of p, and denote m_q^1 and m_q^2 by the first moment and the second moment of q, $b_1(x', x)$ is expressed as follows:

$$\begin{aligned} b_1(x', x) &= \frac{-(x'-x)^2}{\beta(2x^2 - 2(m_p^1 + m_q^1)x + m_p^2 + m_q^2)} \\ &= \frac{-(x'-x)^2}{c_2x^2 + c_1x + c_0}. \end{aligned} \tag{28}$$

Then, the following lemma is proved to show that, $\forall n \in \mathbb{Z}_{\geq 0}$, $b_1^{(n)}$ is a polynomial function with respect to x' and the degree of $b_1^{(n)}$ is always 2.

Lemma 5. *Given p, q defined in Problem 1 and $b_1(x', x)$ in Equation (28), we have*

$$b_1^{(n)} = \frac{\partial^n b_1(x', x)}{\partial x^n} = a_{0n}(x) + a_{1n}(x)x' + a_{2n}(x)(x')^2, \tag{29}$$

where $a_{0n}(x), a_{1n}(x)$ and $a_{2n}(x)$ are rational functions with respect to x and $a_{0n}(x) = \mathcal{O}(x^{-(1+n)})$, $a_{1n}(x) = \mathcal{O}(x^{-(1+n)})$ and $a_{2n}(x) = \mathcal{O}(x^{-(2+n)})$.

Proof. According to Equation (28), we have

$$b_1(x', x) = -\frac{1}{c_2x^2 + c_1x + c_0}(x')^2 + 2\frac{x}{c_2x^2 + c_1x + c_0}x' - \frac{x^2}{c_2x^2 + c_1x + c_0}.$$

Since $\beta > 0$, c_2 is thus over 0, which indicates that $b_1^{(n)}$ is a polynomial function with respect to x' and the degree of $b_1^{(n)}$ is always 2. Let $C_2(x) = -(c_2x^2 + c_1x + c_0)^{-1}$, we have

$$b_1(x', x) = C_2(x)(x')^2 - C_2(x)xx' + C_2(x)x^2.$$

Then, we have

$$\frac{\partial^n b_1(x', x)}{\partial x^n} = \frac{\partial^n C_2(x)}{\partial x^n}(x')^2 + \frac{\partial^n xC_2(x)}{\partial x^n}x' + \frac{\partial^n x^2C_2(x)}{\partial x^n}$$

Because of the definition of $C_2(x)$, we know

$$\frac{\partial^n C_2(x)}{\partial x^n} = \mathcal{O}(x^{-(2+n)}), \ \frac{\partial^n xC_2(x)}{\partial x^n} = \mathcal{O}(x^{-(1+n)}), \ \frac{\partial^n x^2C_2(x)}{\partial x^n} = \mathcal{O}(x^{-(1+n)}).$$

□

Based on Lemma 5, Faà di Bruno's formula and Bell polynomial, the n^{th} derivative of the function $k(x', x; r_x, 1)$ with respect to x is given in the following lemma.

Lemma 6. *Given p, q defined in Problem 1, $d = 1$ and $k(x', x; r_x, 1)$ in Equation (17), we have*

$$K_1^{(n)} = \frac{\partial^n k(x', x; r_x, 1)}{\partial x^n} = K_1^{(0)} \sum_{i=0}^{2n} c_{in}(x)(x')^i,$$

where $c_{in}(x)$ is a rational function with respect to x, $c_{in}(x) = \mathcal{O}(x^{-i-n})$, $i = 1, ..., 2n$, and $c_{0n}(x) = \mathcal{O}(x^{-1-n})$.

Proof. To prove this lemma, the basic idea is to use Faà di Bruno's formula and Bell polynomial. Based on the Faà di Bruno's formula and Bell polynomial, we have

$$\begin{aligned} K_1^{(n)} &= \frac{\partial^n exp(b_1(x', x))}{\partial x^n} \\ &= K_1^{(0)} B_n(b_1^{(1)}, b_1^{(2)}, ..., b_1^{(n)}) \\ &= K_1^{(0)} \sum_{(m_1,...,m_n)} \frac{n!}{m_1! m_2! \cdots m_n!} \prod_{j=1}^{n} \left(\frac{b_1^{(j)}}{j!} \right)^{m_j}, \end{aligned} \tag{30}$$

where B_n is called the n^{th} complete exponential Bell polynomial, the sum is over all $n-$tuples of nonnegative integers $(m_1, ..., m_n)$ satisfying the constraint $1m_1 + 2m_2 + 3m_3 + ... + nm_n = n$. Substituting Equation (29) into Equation (30), we have

$$K_1^{(n)} = K_1^{(0)} \sum_{(m_1,...,m_n)} \frac{n!}{1m_1! 2m_2! \cdots nm_n!} \prod_{j=1}^{n} \left(a_{0j}(x) + a_{1j}(x)x' + a_{2j}(x)(x')^2 \right)^{m_j}.$$

Then, using the trinomial expansion, we have

$$K_1^{(n)} = K_1^{(0)} \sum_{(m_1,...,m_n)} \frac{n!}{1m_1! 2m_2! \cdots nm_n!} \prod_{j=1}^{n} \sum_{j_0, j_1, j_2} \frac{m_j!}{j_0! j_1! j_2!} a_{0j}^{j_0}(x)(a_{1j}(x)x')^{j_1}(a_{2j}(x)(x')^2)^{j_2},$$

where $j_0 + j_1 + j_2 = m_j$. Without loss of the generality, we assume that $x > 0$. Then, based on two constraints: 1) $1m_1 + 2m_2 + 3m_3 + ... + nm_n = n$ and 2) $j_0 + j_1 + j_2 = m_j$, we know the maximum coefficients of x' and $(x')^2$ appears in the tuple $(0, 0, ..., 1)$. Based on Lemma 5, the coefficients have the same order with x^{-1-n} and x^{-2-n}, respectively. Similarly, we know the maximum coefficients of $(x')^3$ and $(x')^4$ appear in the tuple $(1, 0, ..., 1, 0)$ and the coefficients have the same order with x^{-3-n} and x^{-4-n}, respectively. Finally, the maximum coefficients of $(x')^{2n-1}$ and $(x')^{2n}$ appears in the tuple $(n, 0, ..., 0)$ and the coefficients have the same order with $x^{-2n-n+1}$ and x^{-3n}, respectively. This lemma is proved. ☐

Based on Lemma 6, the Formula (27) is proved in the following theorem.

Theorem 8. *Given p, q defined in Problem 1, $d = 1$ and $m = 1$, we have*

$$\frac{\partial^n (T \circ h)(x)}{\partial x^n} = \int_{-\infty}^{+\infty} K_1^{(0)} \sum_{i=0}^{2n} c_{in}(x)(x')^i dh(x').$$

Then, because $\frac{\partial^n (T \circ h)(x)}{\partial x^n} = 0$ for any $x \in \mathbb{R}^d$ and h has a compact support set, we have

$$\int_{-\infty}^{+\infty} (x')^l dh(x') = 0, \ \forall l \in \mathbb{Z}_{\geq 0}.$$

Proof. We begin by considering $n = 1$. When $n = 1$, we have

$$\frac{\partial (T \circ h)(x)}{\partial x} = \int_{-\infty}^{+\infty} K_1^{(0)} \sum_{i=0}^{2} c_{in}(x)(x')^i dh(x') = 0.$$

Because $(T \circ h)(x) = 0$, we know

$$\begin{aligned} &\int_{-\infty}^{+\infty} K_1^{(0)} (c_{11}(x)x' + c_{21}(x)(x')^2) dh(x') = 0 \\ \Leftrightarrow &\int_{-\infty}^{+\infty} K_1^{(0)} c_{11}(x)x' dh(x') = -\int_{-\infty}^{+\infty} K_1^{(0)} c_{21}(x)(x')^2 dh(x') \\ \Leftrightarrow &\int_{-\infty}^{+\infty} K_1^{(0)} \frac{c_{11}(x)}{x^{-2}} x' dh(x') = -\int_{-\infty}^{+\infty} K_1^{(0)} \frac{c_{21}(x)}{x^{-2}} (x')^2 dh(x') \end{aligned}$$

Because $\frac{\partial^n (T\circ h)(x)}{\partial x^n}=0$ for any $x\in\mathbb{R}^d$, taking $x\to+\infty$, we know $x^2c_{11}(x)\to$ a nonzero constant, $x^2c_{21}(x)\to 0$ and $K_1^{(0)}$ is a positive constant. This means

$$\int_{-\infty}^{+\infty}(x')dh(x')=0.$$

We also obtain that

$$\int_{-\infty}^{+\infty}K_1^{(0)}\frac{c_{22}'(x)}{x^{-4}}(x')^2dh(x')=-\int_{-\infty}^{+\infty}K_1^{(0)}\frac{c_{32}(x)}{x^{-4}}(x')^3dh(x').$$

Taking $x\to+\infty$, we know $x^4c_{22}'(x)\to$ a nonzero constant, $x^4c_{32}(x)\to 0$ and $K_1^{(0)}\to$ a positive constant. This means

$$\int_{-\infty}^{+\infty}(x')^2dh(x')=0.$$

Using the same idea, we prove that

$$\int_{-\infty}^{+\infty}(x')^l dh(x')=0,\ \forall l\in\mathbb{Z}_{\geq 0}.$$

This theorem is proved. □

Then, similar to Lemma 5, we prove that $\forall n\in\mathbb{Z}_{\geq 0}$, $b_m^{(n)}$ is a polynomial function with respect to x' and the degree of $b_m^{(n)}$ is always $2m$ ($m>1$), and a following corollary is directly obtained via Lemma 5 and Theorem 8.

Corollary 1. *Given p,q defined in Problem 1, if $d=1$, we will have*

$$\frac{\partial^n (T\circ h)(x)}{\partial x^n}=\int_{-\infty}^{+\infty}K_m^{(0)}\sum_{i=0}^{2mn}c_i(x)(x')^i dh(x').$$

Then, because $\frac{\partial^n (T\circ h)(x)}{\partial x^n}=0$ for any $x\in\mathbb{R}^d$ and h has a compact support set, we have

$$\int_{-\infty}^{+\infty}(x')^l dh(x')=0,\ \forall l\in\mathbb{Z}_{\geq 0}.$$

Proof of Theorem 3: Next, we consider the situation that $d>1$. When $d>1$, $b_m(x',x)$ has the following expression:

$$\begin{aligned}b_m(x',x)&=\frac{-\|x'-x\|^{2m}}{\beta^m\int_{\mathbb{R}^d}\|x'-x\|^{2m}d_{p_{x'}+q_{x'}}}\\&=\frac{-(\sum_{j=1}^d((x')_j-x_j)^2)^m}{\beta^m\int_{\mathbb{R}^d}(\sum_{j=1}^d((x')_j-x_j)^2)^m d_{p_{x'}+q_{x'}}}\\&=-\sum_{i_0,i_1,i_2}\frac{m!}{i_0!i_1!i_2!}\frac{(\sum_{j=1}^d x_j^2)^{i_0}(\sum_{j=1}^d -2x_j(x')_j)^{i_1}(\sum_{j=1}^d(x')_j^2)^{i_2}}{\beta^m\int_{\mathbb{R}^d}(\sum_{j=1}^d((x')_j-x_j)^2)^m d_{p_{x'}+q_{x'}}},\end{aligned}\tag{31}$$

where $i_0+i_1+i_2=m$. We use r_x to denote $\beta^m\int_{\mathbb{R}^d}(\sum_{j=1}^d((x')_j-x_j)^2)^m d_{p_{x'}+q_{x'}}$. Then, based on the trinomial expansion, $b_m(x',x)$ is rewritten as follows:

$$b_m(x',x)=-\sum_{i_0,i_1,i_2}\frac{(-2)^{i_1}m!}{i_0!i_1!i_2!}\left(\sum_{\{v_{ji_0}\}_{j=1}^d}\prod_{j=1}^d x_j^{2v_{ji_0}}\cdot\sum_{\{v_{ji_1}\}_{j=1}^d}\prod_{j=1}^d(x_jx_j')^{v_{ji_1}}\cdot\sum_{\{v_{ji_2}\}_{j=1}^d}\prod_{j=1}^d(x_j')^{2v_{ji_2}}\right)\frac{1}{r_x},\tag{32}$$

where $v_{ji_0},v_{ji_1},v_{ji_2}\in\mathbb{Z}_{\geq 0}$, $\sum_{j=1}^d v_{ji_0}=i_0$, $\sum_{j=1}^d v_{ji_1}=i_1$ and $\sum_{j=1}^d v_{ji_2}=i_2$. In Equation (32), we can re-arrange it according to $\prod_{j=1}^d(x_j')^{v_{jl}}$ as follows, where $\sum_{j=1}^d v_{jl}=l$.

$$b_m(x',x)=\sum_{l=0}^{2m}\sum_{\{v_{jl}\}_{j=1}^d}A_{\{v_{jl}\}_{j=1}^d}(x)\prod_{j=1}^d(x_j')^{v_{jl}}.\tag{33}$$

Based on Equation (32) and $i_0+i_1+i_2=m$, we have

$$\begin{aligned}A_{\{v_{jl}\}_{j=1}^d}(x)&=\mathcal{O}\left(\frac{\sum_{\{v_{j,2m-l}\}_{j=1}^d}\prod_{j=1}^d x_j^{v_{j,2m-l}}}{r_x}\right),\\&=\mathcal{O}\left(\frac{\sum_{\{v_{j,2m-l}\}_{j=1}^d}\prod_{j=1}^d x_j^{v_{j,2m-l}}}{\sum_{\{v_{j,2m}\}_{j=1}^d}\prod_{j=1}^d x_j^{v_{j,2m}}}\right).\end{aligned}\tag{34}$$

where $\sum_{j=1}^{d} v_{j,2m-l} = 2m - l$ and $\sum_{j=1}^{d} v_{j,2m} = 2m$. So, we have

$$\frac{\partial^n A_{\{v_{jl}\}_{j=1}^d}(x)}{\partial x_j^n} = \mathcal{O}\left(\frac{1}{\sum_{\{v_{j,l+n}\}_{j=1}^d} \prod_{j=1}^{d} x_j^{v_{j,l+n}}}\right), \tag{35}$$

where $l > 0$ and $\sum_{j=1}^{d} v_{j,l+1} = l + 1$. Then, we have

$$\frac{\partial^n (T \circ h)(x)}{\partial x_j^n} = \int_{\mathbb{R}^d} K_m^{(0)} \sum_{l=0}^{2mn} \sum_{\{v_{jl}\}_{j=1}^d} \frac{\partial^n A_{\{v_{jl}\}_{j=1}^d}(x)}{\partial x_j^n} \prod_{j=1}^{d} (x'_j)^{v_{jl}} dh(x').$$

Since $\frac{\partial^n (T\circ h)(x)}{\partial x_j^n} = 0$, based on the same idea used to prove Theorem 8, we obtain

$$\begin{aligned}
&\int_{\mathbb{R}^d} \sum_{\{v_{jl}\}_{j=1}^d} \lim_{x_j \to +\infty} x_j^{l+n} \frac{\partial^n A_{\{v_{jl}\}_{j=1}^d}(x)}{\partial x_j^n} \prod_{j=1}^{d} (x'_j)^{v_{jl}} dh(x') \\
=& \sum_{\{v_{jl}\}_{j=1}^d} \lim_{x_j \to +\infty} x_j^{l+n} \frac{\partial^n A_{\{v_{jl}\}_{j=1}^d}(x)}{\partial x_j^n} \int_{\mathbb{R}^d} \prod_{j=1}^{d} (x'_j)^{v_{jl}} dh(x') \\
=& \ 0,
\end{aligned} \tag{36}$$

where $\lim_{x_j \to +\infty} x_j^{l+n} \frac{\partial^n A_{\{v_{jl}\}_{j=1}^d}(x)}{\partial x_j^n}$ is a function related to $x_1, ..., x_{j-1}, x_{j+1}, ..., x_d$ according to Equation (35). Since Equation (36) is correct for any $x \in \mathbb{R}^d$, we know

$$\int_{\mathbb{R}^d} \prod_{j=1}^{d} (x'_j)^{v_{jl}} dh(x') = 0.$$

Theorem 3 is proved.

Appendix D
Proof of Theorem 4

We first write a *multi-index* μ is an element of $(\mathbb{Z}_{\geq 0})^d$ as follows:

$$\Lambda(d) = (\mathbb{Z}_{\geq 0})^d.$$

Then, we recall a notation for this proof. For $\mu = (\mu_1, ..., \mu_d) \in \Lambda(d)$ and $x = (x_1, ..., x_d) \in \mathbb{R}^d$, set

$$x^\mu = x_1^{\mu_1} x_2^{\mu_2} \cdots x_d^{\mu_d}.$$

If we want to prove $(T \circ (p - q))(x)$ is a real analytic function defined on $\mathbb{R}^d$, we need to prove that, $\forall x_0 \in \mathbb{R}^d$,

$$(T \circ (p - q))(x) = \sum_{\mu \in \Lambda(d)} a_\mu (x - x_0)^\mu,$$

where $x \in U_{x_0}$ and U_{x_0} is an open set containing x_0. In the following, we use $h = p - q$ for a short expression. Because e^{-x^m}, $\|x - x'\|^2$ and r_x are real analytic functions and $r_x > 0$, $k(x', x; r_x, m)$ is also a real analytic function (based on Proposition 2.2.2 in [28]). Then $\forall x_0, x'_0 \in A$, there exists an open set $B(x_0, r) \times B(x'_0, r)$, we can write $k(x', x; r_x, m)$ as power series:

$$k(x', x; r_x, m) = \sum_{\mu \in \Lambda(d), \mu' \in \Lambda(d)} a_{\mu,\mu'} (x - x_0)^\mu (x' - x'_0)^{\mu'}.$$

Let $A = \mathrm{supp}(h)$ and any $x_0 \in \mathbb{R}^d$, because h has a compact support set A, there are L balls $B(x'_i, \epsilon_{x'_i})$ to cover A, where $i = 1, ..., L$, such that $k(x', x; r_x, m)$ can be written as power series in every ball $B(x_0, \epsilon_{x'_i}) \times B(x'_i, \epsilon_{x'_i})$, where $i = 1, ..., L$. Through these balls, we can construct L subsets A_i such that every A_i is contained in the ball $B(x'_i, \epsilon_{x'_i})$, and

$$A = \cup_{i=1}^{L} A_i, \ \ A_i \cap A_j = \emptyset \ \forall i \neq j, i, j \in \{1, 2, ..., L\}.$$

For every $x \in B(x_0, \epsilon)$, where $\epsilon = \min_{i=1,...,L} \epsilon_{x'_i}$, we have

$$\begin{aligned}(T \circ h)(x) &= \sum_{i=1}^{L} \int_{A_i} k(x', x; r_x, m) dh(x') \\ &= \sum_{i=1}^{L} \int_{A_i} \sum_{\mu \in \Lambda(d), \mu' \in \Lambda(d)} a_{\mu,\mu'} (x - x_0)^{\mu} (x' - x'_i)^{\mu'} dh(x') \\ &= \sum_{i=1}^{L} \sum_{\mu \in \Lambda(d)} \sum_{\mu' \in \Lambda(d)} \int_{A_i} a_{\mu,\mu'} (x' - x'_i)^{\mu'} dh(x') (x - x_0)^{\mu} \\ &= \sum_{\mu \in \Lambda(d)} \sum_{i=1}^{L} \sum_{\mu' \in \Lambda(d)} \int_{A_i} a_{\mu,\mu'} (x' - x'_i)^{\mu'} dh(x') (x - x_0)^{\mu}.\end{aligned}$$

Taking

$$a_{\mu} = \sum_{i=1}^{L} \sum_{\mu' \in \Lambda(d)} \int_{A_i} a_{\mu,\mu'} (x' - x'_i)^{\mu'} dh(x'),$$

we have

$$(T \circ h)(x) = \sum_{\mu \in \Lambda(d)} a_{\mu} (x - x_0)^{\mu}.$$

This theorem is proved.

APPENDIX E
PROOF OF THEOREM 5

To prove Theorem 5, we first prove the following lemma:

Lemma 7. *For any real analytic function f in $\mathbb{R}^d$, if Lebesgue measure $\mu(\{x : f = 0\}) > 0$, then $f = 0$.*

Proof. Because $\mu(\{x : f = 0\}) > 0$, there exists a closed ball $\overline{B(0, r)}$ such that $\mu(\{x : f = 0\} \cap \overline{B(0, r)}) > 0$.

Step 1
Let $B = \{x : f = 0\} \cap \overline{B(0, r)}$, $A_0 = \{x \in B : |Df| \neq 0\}$ and $B_0 = \{x \in B : |Df| = 0\}$.

Claim 1: $\mu(B_0) > 0$.
According to Implicit Function Theorem, we obtain that for every $x \in A_0$, there exists an open set $U_x \subset \mathbb{R}^d$ such that $U_x \cap A_0$ is a $d - 1$ dimensional manifold M_x^{d-1}. Moreover, M_x^{d-1} can be presented as a graph:

$$M_x^{d-1} = \{(x.g(x)) : x \in U\}, \tag{37}$$

where U is an open set in $\mathbb{R}^{d-1}$ and U is bounded. Therefore, $\mu(M_x^{d-1}) = 0$.
According to Vitali's Covering Theorem, there exist countable manifolds $\{M_i^{d-1}\}_{i=1}^{\infty}$ and zero measurable sets C, such that

$$A_0 \subset \cup_{i=1}^{\infty} M_i^{d-1} \cup C,$$

which implies $\mu(A_0) \leq \sum_{i=1}^{\infty} \mu(M_i^{d-1}) + \mu(C) = 0$.
But $\mu(A_0) + \mu(B_0) = \mu(B) > 0$, we obtain that

$$\mu(B_0) > 0.$$

Step 2
Assume for $i \in \mathbb{Z}^+$

$$A_i = \{x \in B_{i-1} : |D^{(i)} f| \neq 0\},$$

$$B_i = \{x \in B_{i-1} : |D^{(i)} f| = 0\},$$

and $k_1, k_2, ..., k_i \in \{1, 2, ..., d\}$,

$$A_i(k_1, k_2, ..., k_i) = \{x \in B_i : \frac{\partial^{k_1 k_2 ... k_i} f}{\partial x_{k_1} \partial x_{k_2} ... \partial x_{k_i}} \neq 0\}.$$

Then we know that

$$B_{i-1} = A_i \cup B_i,$$

$$A_i = \cup_{k_1,k_2...,k_i} A_i(k_1, k_2, ..., k_i).$$

Claim 2: $\mu(B_i) > 0, i \in \mathbb{Z}_{\geq 0}$.
We use Mathematical induction to prove **Claim 2**.
1) $\mu(B_0) > 0$ has been proved in **Claim 1**.

2) Assume $\mu(B_i) > 0$, where $i \geq 0$, we need to prove $\mu(B_{i+1}) > 0$.
If $\mu(B_{i+1}) = 0$, then $\mu(A_{i+1}) > 0$, because $\mu(B_{i+1}) + \mu(A_{i+1}) = \mu(B_i) > 0$.

Therefore, there exists $k_1, k_2, ..., k_i, k_{i+1}$ such that

$$\mu(A_{i+1}(k_1, k_2, ..., k_{i+1})) > 0.$$

Consider the function

$$g = \frac{\partial^{k_2...k_{i+1}} f}{\partial x_{k_2}...\partial x_{k_{i+1}}},$$

then

$$A_{i+1}(k_1, k_2, ..., k_{i+1}) \subset \{x \in B_i : g \neq 0\}.$$

According to the idea of **Claim 1**, we can prove that there exist countable manifolds $\{M_i^{d-1}\}_{i=1}^{\infty}$ and zero measurable set C, such that

$$\{x \in B_i : g \neq 0\} \subset \cup_{i=1}^{\infty} M_i^{d-1} \cup C,$$

which implies that

$$\mu(A_{i+1}(k_1, k_2, ..., k_{i+1})) \leq \mu(\{x \in B_i : g \neq 0\}) \leq \sum_{i=1}^{\infty} \mu(M_i^{d-1}) + \mu(C) = 0.$$

This conflicts with $\mu(A_{i+1}) > 0$. Hence, $\mu(B_{i+1}) > 0$.

Step 3
We obtain a sequence $\{B_i\}_{i=1}^{\infty}$, which satisfy

$$B_{i+1} \subset B_i,$$

$$\mu(B_i) > 0,$$

and B_i is compact, $i \in \mathbb{Z}_{\geq 0}$.

Claim 3: there exists $p \in \mathbb{R}^d$ such that $p \in B_i, \quad i \in \mathbb{Z}_{\geq 0}$.
Because $\mu(B_i) > 0$, we can find a point $x_i \in B_i$. Then we obtain a sequence

$$\{x_i\}_{i=0}^{\infty}.$$

Because $B_{i+1} \subset B_i$, we know that

$$x_i \in B_j, \quad i \leq j.$$

Hence,

$$\{x_i\}_{i=0}^{\infty} \subset B_0.$$

Therefore, there exists a subsequence $\{x_{ik}\}_{k=0}^{\infty}$ of $\{x_i\}_{i=0}^{\infty}$ such that there exists a point p satisfying

$$x_{ik} \to p, \quad k \to +\infty, \tag{38}$$

because B_0 is compact.

Because B_i is closed, we obtain that $p \in B_i$.

Step 4 According to the definition of B_i, we know that

$$D^{(i)} f(p) = 0,$$

for $i \in \mathbb{Z}_{\geq 0}$.
According to the property of analytic function, we obtain that $f(x) = 0$, for any $x \in \mathbb{R}^d$. $\square$

Proof for Theorem 5:
According to Theorem 4, $T \circ (p - q)$ is a real analytic function. Moreover, $\mathrm{supp}\{p + q\} \subset \{x \in \mathbb{R}^d : T \circ (p - q) = 0\}$.

According to the assumption $\mu(\mathrm{supp}\{p + q\}) > 0$, we get

$$\mu(\{x \in \mathbb{R}^d : T \circ (p - q) = 0\}) > 0,$$

which implies

$$T \circ (p - q)(x) = 0, \forall x \in \mathbb{R}^d,$$

according to Lemma 7. This theorem is proved.

Appendix F
Proof of Lemma 2

Using the definition of $V(r_x)$ and $L_d(x)$, we have

$$|L_d(x)| = \left| \int \frac{k(x', x)}{V(r_x)} dp(x') - dq(x') \right|.$$

Without loss of generality, we assume that $\int \frac{k(x',x)}{V(r_x)} dp(x') \geq \int \frac{k(x',x)}{V(r_x)} dq(x')$. Thus, $|L_d(x)|$ is less than $\int \frac{k(x',x)}{V(r_x)} dp(x')$. Based on the Equation (12) and $0 \leq k(x', x_1) \leq 1$, we have

$$\begin{aligned} |L_d(x)| &= \left| \int \frac{m\Gamma(\frac{d}{2}) k(x', x)}{\Gamma(\frac{d}{2m})(\sqrt{\pi} r_x)^d} dp(x') - dq(x') \right| \\ &< \left| \int \frac{m\Gamma(\frac{d}{2}) k(x', x)}{\Gamma(\frac{d}{2m})(\sqrt{\pi} r_x)^d} dp(x') \right| \\ &\leq \frac{m\Gamma(d/2)}{\Gamma(d/2m)(\sqrt{\pi} r_x)^d}. \end{aligned}$$

Because $(d/2m) \cdot \Gamma(d/2m) = \Gamma(1 + d/2m)$, we have

$$|L_d(x)| < \frac{d\Gamma(d/2)}{2\Gamma(1 + d/2m)(\sqrt{\pi} r_x)^d}.$$

This lemma is proved.

Appendix G
Proof of Theorem 6

We first prove the case for $|\mathrm{MLD}^{\tilde{p}}(p, q) - \widehat{\mathrm{MLD}}^{\tilde{p}}(X_1, X_2)|$ and then prove the case for $|\mathrm{MLD}(p, q) - \widehat{\mathrm{MLD}}(X_1, X_2)|$.

Case 1. We denote $\Delta(p, q, X_1, X_2) = \widehat{\mathrm{MLD}}^{\tilde{p}}(X_1, X_2) - \mathrm{MLD}^{\tilde{p}}(p, q)$. Then, in terms of Lemma 1, changing either x_{1i} or x_{2i} in $\Delta(p, q, X_1, X_2)$ results in changes in magnitude of at most $d^{\tilde{p}}\Gamma^{\tilde{p}}(d/2)/(\Gamma^{\tilde{p}}(1 + d/2m)(\sqrt{\pi} C_r)^{d\tilde{p}} n_1)$ or $d^{\tilde{p}}\Gamma^{\tilde{p}}(d/2)/(\Gamma^{\tilde{p}}(1 + d/2m)(\sqrt{\pi} C_r)^{d\tilde{p}} n_2)$, respectively. We can apply McDiarmid's inequality [10], given a denominator in the exponent of

$$n_1 \left(\frac{d^{\tilde{p}}\Gamma^{\tilde{p}}(d/2)}{\Gamma^{\tilde{p}}(1 + d/2m)(\sqrt{\pi} C_r)^{d\tilde{p}} n_1} \right)^2 + n_2 \left(\frac{d^{\tilde{p}}\Gamma^{\tilde{p}}(d/2)}{\Gamma^{\tilde{p}}(1 + d/2m)(\sqrt{\pi} C_r)^{d\tilde{p}} n_2} \right)^2 = \frac{2d^{2\tilde{p}}\Gamma^{2\tilde{p}}(d/2)(n_1 + n_2)}{\Gamma^{2\tilde{p}}(1 + d/2m) n_1 n_2 \pi^{d\tilde{p}} C_r^{2d\tilde{p}}}.$$

to obtain

$$Pr_{X_1,X_2}(\Delta(p, q, X_1, X_2) - \mathbb{E}_{X_1,X_2}(\Delta(p, q, X_1, X_2)) > \epsilon) \leq exp\left(\frac{-\epsilon^2 \Gamma^{2\tilde{p}}(1 + d/2m) n_1 n_2 \pi^{d\tilde{p}} C_r^{2d\tilde{p}}}{2d^{2\tilde{p}}\Gamma^{2\tilde{p}}(d/2)(n_1 + n_2)} \right). \tag{39}$$

So, we need to calculate $\mathbb{E}_{X_1,X_2}(\Delta(p, q, X_1, X_2))$. First, we need to prove the following formula.

$$\sqrt{n_1}\Big(\frac{1}{n_1} \sum_{j=1}^{n_1} \frac{k(x_{1j}, x_{1i})}{V(r_{x_{1i}})} - \int_{\mathbb{R}^d} \frac{k(x', x_{1i})}{V(r_{x_{1i}})} dp_{x'} \Big) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_{1i}^2), \tag{40}$$

where $\sigma_{1,1i}^2 = Var(k(x_{1j}, x_{1i})/V(r_{x_{1i}}))$. Because

$$\begin{aligned}\mathbb{E}\Big(\frac{1}{n_1}\sum_{j=1}^{n_1}\frac{k(x_{1j},x_{1i})}{V(r_{x_{1i}})}\Big) &= \frac{1}{n_1}\sum_{j=1}^{n_1}\mathbb{E}\Big(\frac{k(x_{1j},x_{1i})}{V(r_{x_{1i}})}\Big)\\ &= \frac{n_1}{n_1}\int_{\mathbb{R}^d}\frac{k(x_{1j},x_{1i})}{V(r_{x_{1i}})}dp_{x_{1j}}\\ &= \int_{\mathbb{R}^d}\frac{k(x',x_{1i})}{V(r_{x_{1i}})}dp_{x'},\end{aligned}$$

$k(x', x; r_x, m) \leq 1$ and $r_x \geq C_r > 0$ (meaning that $\sigma_{1i}^2 < +\infty$), using the central limit theorem, we prove Formula (40). Thus, we have

$$\frac{1}{n_1}\sum_{j=1}^{n_1}\frac{k(x_{1j},x_{1i})}{V(r_{x_{1i}})} = \int_{\mathbb{R}^d}\frac{k(x',x_{1i})}{V(r_{x_{1i}})}dp_{x'} + O_{p_1}(\frac{1}{\sqrt{n_1}}), \tag{41}$$

where $O_{p_1}(1/\sqrt{n_1}) = 1/\sqrt{n_1}\mathcal{N}(0, \sigma_{1,1i}^2)$. Similarly, we know

$$\frac{1}{n_2}\sum_{j=1}^{n_1}\frac{k(x_{2j},x_{1i})}{V(r_{x_{1i}})} = \int_{\mathbb{R}^d}\frac{k(x',x_{1i})}{V(r_{x_{1i}})}dq_{x'} + O_{p_2}(\frac{1}{\sqrt{n_2}}), \tag{42}$$

where $O_{p_2}(1/\sqrt{n_2}) = 1/\sqrt{n_2}\mathcal{N}(0, \sigma_{2,1i}^2)$. Substituting Equations (41) and (42) into $\widehat{\text{MLD}}_l$, we have

$$\begin{aligned}\widehat{\text{MLD}}_l &= \frac{1}{n_l}\sum_{i=1}^{n_l}\left|\frac{1}{n_1}\sum_{j=1}^{n_1}\frac{k(x_{1j},x_{li})}{V(r_{x_{li}})} - \frac{1}{n_2}\sum_{j=1}^{n_2}\frac{k(x_{2j},x_{li})}{V(r_{x_{li}})}\right|^{\tilde{p}},\\ &= \frac{1}{n_l}\sum_{i=1}^{n_l}\left|\int_{\mathbb{R}^d}\frac{k(x',x_{li})}{V(r_{x_{li}})}d_{p_{x'}-q_{x'}} + O_p(\frac{1}{\sqrt{n_1}}+\frac{1}{\sqrt{n_2}})\right|^{\tilde{p}},\end{aligned} \tag{43}$$

where $O_p(1/\sqrt{n_1} + 1/\sqrt{n_2}) = \mathcal{N}(0, \sigma_{1,1i}^2/n_1 + \sigma_{2,1i}^2/n_2)$. Using Taylor series, expand Equation (43) at the value $\int \frac{k(x',x_{li})}{V(r_x)} d_{p_{x'}-q_{x'}}$, we have

$$\widehat{\text{MLD}}_l = \frac{1}{n_l}\sum_{i=1}^{n_l}\left(\left|\int_{\mathbb{R}^d}\frac{k(x',x_{li})}{V(r_{x_{li}})}d_{p_{x'}-q_{x'}}\right|^{\tilde{p}} + \left|O_p'(\frac{1}{\sqrt{n_1}}+\frac{1}{\sqrt{n_2}})\right|\right),$$

where $O_p'(1/\sqrt{n_1}+1/\sqrt{n_2}) = \mathcal{N}(0, M\sigma_{1,li}^2/n_1 + M\sigma_{2,li}^2/n_2)$ and M is a positive finite number. Because $\int_{\mathbb{R}^d}\frac{k(x',x_{li})}{V(r_{x_{li}})}d_{p_{x'}-q_{x'}}$ is a function related to x_{li}, So, we have

$$\begin{aligned}\mathbb{E}_{X_1,X_2}(\widehat{\text{MLD}}_l - \text{MLD}_l) &= \mathbb{E}_{X_1,X_2}\left(O_{p_1}(1/\sqrt{n_l}) + \frac{1}{n_l}\sum_{i=1}^{n_l}\left|O_p'(\frac{1}{\sqrt{n_1}}+\frac{1}{\sqrt{n_2}})\right|\right)\\ &= \frac{1}{n_l}\sum_{i=1}^{n_l}\mathbb{E}\left(\left|O_p'(\frac{1}{\sqrt{n_1}}+\frac{1}{\sqrt{n_2}})\right|\right).\end{aligned}$$

Because $|O_p'|$ is a random variable that obeys the half-normal distribution, we know

$$\mathbb{E}_{X_1,X_2}(\widehat{\text{MLD}}_l - \text{MLD}_l) = \frac{1}{n_l}\sum_{i=1}^{n_l}\frac{\sqrt{2}}{\sqrt{\pi}}\sqrt{M\Big(\frac{\sigma_{1,li}^2}{n_1}+\frac{\sigma_{2,li}^2}{n_2}\Big)}.$$

Thus, we know

$$\mathbb{E}_{X_1,X_2}(\Delta(p,q,X_1,X_2)) = \sum_{l=1}^{2}\frac{1}{n_l}\sum_{i=1}^{n_l}\frac{\sqrt{2}}{\sqrt{\pi}}\sqrt{M\Big(\frac{\sigma_{1,li}^2}{n_1}+\frac{\sigma_{2,li}^2}{n_2}\Big)} \tag{44}$$

Combining Inequality (39) and Equation (44), we have

$$\begin{aligned}&\Pr_{X_1,X_2}(\widehat{\text{MLD}}^{\tilde{p}}(X_1,X_2) - \text{MLD}^{\tilde{p}}(p,q) - \sum_{l=1}^{2}\frac{1}{n_l}\sum_{i=1}^{n_l}\frac{\sqrt{2}}{\sqrt{\pi}}\sqrt{M\Big(\frac{\sigma_{1,li}^2}{n_1}+\frac{\sigma_{2,li}^2}{n_2}\Big)} > \epsilon)\\ &\leq \exp\left(\frac{-\epsilon^2\Gamma^{2\tilde{p}}(1+d/2m)n_1n_2\pi^{d\tilde{p}}C_r^{2d\tilde{p}}}{2d^{2\tilde{p}}\Gamma^{2\tilde{p}}(d/2)(n_1+n_2)}\right).\end{aligned} \tag{45}$$

Case 2. Because of Inequality (45), we have

$$\begin{aligned}
&\Pr_{X_1,X_2}(\mathrm{MLD}(p,q)-\widehat{\mathrm{MLD}}(X_1,X_2)-U^{1/\tilde{p}}>\epsilon)\\
=\ &\Pr_{X_1,X_2}(\mathrm{MLD}^{\tilde{p}}(p,q)>(\widehat{\mathrm{MLD}}(X_1,X_2)+U^{1/\tilde{p}}+\epsilon)^{\tilde{p}})\\
\le\ &\Pr_{X_1,X_2}(\mathrm{MLD}^{\tilde{p}}(p,q)>\widehat{\mathrm{MLD}}^{\tilde{p}}(X_1,X_2)+U+\epsilon^{\tilde{p}})\\
\le\ &\exp\left(\frac{-\epsilon^{2\tilde{p}}\Gamma^{2\tilde{p}}(1+d/2m)n_1n_2\pi^{d\tilde{p}}C_r^{2d\tilde{p}}}{2d^{2\tilde{p}}\Gamma^{2\tilde{p}}(d/2)(n_1+n_2)}\right),
\end{aligned}$$

where

$$\begin{aligned}
U&=\sum_{l=1}^{2}\frac{1}{n_l}\sum_{i=1}^{n_l}\frac{\sqrt{2}}{\sqrt{\pi}}\sqrt{M\Big(\frac{\sigma_{1,li}^2}{n_1}+\frac{\sigma_{2,li}^2}{n_2}\Big)}\\
&=\frac{1}{\sqrt{n}}\sum_{l=1}^{2}\frac{1}{n_l}\sum_{i=1}^{n_l}\frac{\sqrt{2}}{\sqrt{\pi}}\sqrt{M\Big(\frac{\sigma_{1,li}^2}{\lambda_1}+\frac{\sigma_{2,li}^2}{\lambda_2}\Big)},
\end{aligned}$$

$n=n_1+n_2$ and $\lambda_l=n_l/n$. Hence, the following inequality is obtained.

$$\Pr_{X_1,X_2}(|\mathrm{MLD}(p,q)-\widehat{\mathrm{MLD}}(X_1,X_2)|-\mathcal{O}(n^{-\frac{1}{2\tilde{p}}})>\epsilon)\le 2\exp\left(\frac{-\epsilon^{2\tilde{p}}\Gamma^{2\tilde{p}}(1+d/2m)n_1n_2\pi^{d\tilde{p}}C_r^{2d\tilde{p}}}{2d^{2\tilde{p}}\Gamma^{2\tilde{p}}(d/2)(n_1+n_2)}\right).$$

The theorem is proved.

Appendix H
Proof of Lemma 3

It is clear that $k(z_j,z_i)/V(z_i)$, as a n-by-n matrix, will not change further if X_1 and X_2 are obtained. The one-time permutation of $\widehat{\mathrm{MLD}}(X_1,X_2)$ uniformly chooses n_1 from z_i as the new X_1 and selects the rest of z_i as the new X_2, and then calculates the new $\widehat{\mathrm{MLD}}(X_1,X_2)$. We do not need to calculate this new $\widehat{\mathrm{MLD}}(X_1,X_2)$ from the very beginning (*e.g.* re-calculating $k(x_{1j},x_{1i})$) but rather only from $k(z_j,z_i)/V(z_i)$. This allows us to find the asymptotic null distribution of $\widehat{\mathrm{MLD}}(X_1,X_2)$.

Consider the following random variable b_j,

$$P\left(b_j=n_1^{-1}\frac{k(z_j,x_{1i})}{V(r_{x_{1i}})}\right)=\lambda_1, P\left(b_j=-n_2^{-1}\frac{k(z_j,x_{1i})}{V(r_{x_{1i}})}\right)=\lambda_2,$$

where P means the probability and $\lambda_1+\lambda_2=1$. Under the permutation null ($p=q$), we know

$$n_1^{-1}\sum_{j=1}^{n_1}\frac{k(x_{1j},x_{1i})}{V(r_{x_{1i}})}-n_2^{-1}\sum_{j=1}^{n_2}\frac{k(x_{2j},x_{1i})}{V(r_{x_{1i}})}=\sum_{j=1}^{n}b_j.$$

Because $L(z_i,z_j)$ is fixed after obtaining X_1 and X_2, each b_j is independent and $n_1^{-1}\sum_{j=1}^{n_1}\frac{k(x_{1j},x_{1i})}{V(r_{x_{1i}})}-n_2^{-1}\sum_{j=1}^{n_2}\frac{k(x_{2j},x_{1i})}{V(r_{x_{1i}})}$ is the sum of n independent random variables. To apply the Lyapunov central limit theorem to $\sum b_j$, we need to verify the following condition.

$$\lim_{n\to\infty}\frac{\sum_{j=1}^{n}\mathbb{E}\Big(\big(b_j-\mathbb{E}(b_j)\big)^3\Big)}{s_n^3}=0,\ s_n^2=\sum_{j=1}^{n}\mathbb{E}\Big(\big(b_j-\mathbb{E}(b_j)\big)^2\Big) \tag{46}$$

In terms of the definition of b_j, we know $\mathbb{E}(b_j)=0$ and

$$\begin{aligned}
\mathbb{E}\Big(\big(b_j-\mathbb{E}(b_j)\big)^3\Big)&=\lambda_1n_1^{-3}\Big(\frac{k(z_j,x_{1i})}{V(r_{x_{1i}})}\Big)^3-\lambda_2n_2^{-3}\Big(\frac{k(z_j,x_{1i})}{V(r_{x_{1i}})}\Big)^3,\\
\mathbb{E}\Big(\big(b_j-\mathbb{E}(b_j)\big)^2\Big)&=\lambda_1n_1^{-2}\Big(\frac{k(z_j,x_{1i})}{V(r_{x_{1i}})}\Big)^2+\lambda_2n_2^{-2}\Big(\frac{k(z_j,x_{1i})}{V(r_{x_{1i}})}\Big)^2
\end{aligned}$$

It is clear that any non-zero $n_1^{-1}\frac{k(x_{1j},x_{1i})}{V(r_{x_{1i}})}$ or $n_2^{-1}\frac{k(x_{2j},x_{1i})}{V(r_{x_{1i}})}$ has a minimum value, denoted by $L_{\min}>0$, and a maximum value, $L_{max}<+\infty$. If $\mathbb{E}\big(\big(b_j-\mathbb{E}(b_j)\big)^3\big)>0$, we have

$$\frac{\sum_{j=1}^{n}\mathbb{E}\Big(\big(b_j-\mathbb{E}(b_j)\big)^3\Big)}{s_n^3}<\frac{n\lambda_1L_{\max}^3}{n^{\frac{3}{2}}(\lambda_1+\lambda_2)L_{\min}^2}=\mathcal{O}\left(\frac{1}{\sqrt{n}}\right).$$

If $\mathbb{E}\big((b_j-\mathbb{E}(b_j))^3\big)<0$, we have

$$\frac{\sum_{j=1}^{n}\mathbb{E}\Big(\big(b_j-\mathbb{E}(b_j)\big)^3\Big)}{s_n^3}>\frac{-n\lambda_2 L_{max}^3}{n^{\frac{3}{2}}(\lambda_1+\lambda_2)L_{min}^2}=\mathcal{O}\left(\frac{-1}{\sqrt{n}}\right).$$

Hence, the condition in Formula (46) is verified, which means, due to the Lyapunov central limit theorem, we have

$$\frac{1}{s_n}\sum_{j=1}^{n}(b_j-\mathbb{E}(b_j))\xrightarrow{\mathcal{D}}\mathcal{N}(0,1).$$

Using the definitions of b_j and s_n, this lemma is proved.

APPENDIX I
PROOF OF THEOREM 7

The main idea to prove this theorem is to construct a new random variable $a_{x_{1i}}$ as per the following.

$$a_{x_{1i}}=\frac{1}{n_1^{1/\tilde{p}}}L_d(x_{1i}).$$

So, we know $\widehat{\text{MLD}}_1=\sum_i|a_{x_{1i}}^{\tilde{p}}|$. Similarly, we can construct $a_{x_{2i}}$ and $\widehat{\text{MLD}}_2=\sum_i|a_{x_{2i}}^{\tilde{p}}|$. Since $k(z_j,x_{1i})$ are fixed after obtaining X_1 and X_2 and all of x_{1i} are independent, all of $a_{x_{1i}}$ are independent. Thus, applying Lyapunov's central limit theorem to $\sum_i|a_{x_{1i}}^{\tilde{p}}|$, this theorem can be proved.

To use Lyapunov's central limit theorem, we first need to know whether the first, second and third moments of the random variable $|a_{x_{1i}}^{\tilde{p}}|$ are finite. If these three moments are finite, we can easily verify the Lyapunov condition:

$$\lim_{n\to\infty}\frac{\sum_{i=1}^{n_1}\mathbb{E}\Big(\big(|a_{x_{1i}}^{\tilde{p}}|-\mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|)\big)^3\Big)}{s_{n_1}^3}=0,\ \ s_{n_1}^2=\sum_{i=1}^{n_1}\mathbb{E}\Big(\big(|a_{x_{1i}}^{\tilde{p}}|-\mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|)\big)^2\Big). \tag{47}$$

From Lemma 3, it is clear that $L_d(x_{1i})$ is finite. Thus, the value of $|a_{x_{1i}}^{\tilde{p}}|$ is always finite. Assume $|a_{x_{1i}}^{\tilde{p}}|$ is bounded by a finite number M, we have $\mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|)<M$, $\mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|^2)<M^2$ and $\mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|^3)<M^3$. So, we have

$$\frac{\sum_{i=1}^{n_1}\mathbb{E}\Big(\big(|a_{x_{1i}}^{\tilde{p}}|-\mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|)\big)^3\Big)}{s_{n_1}^3}<\frac{8n_1M^3}{2n_1^{\frac{3}{2}}M^2}=\mathcal{O}(\frac{1}{\sqrt{n_1}}),$$

which means that the condition in Formula (47) is verified. Based on the Lyapunov's central limit theorem, we have

$$\frac{1}{s_{n_1}}\sum_{i=1}^{n_1}(|a_{x_{1i}}^{\tilde{p}}|-\mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|))\xrightarrow{\mathcal{D}}\mathcal{N}(0,1).$$

Hence,

$$\frac{1}{s_{n_1}}\widehat{\text{MLD}}_1-\frac{1}{s_{n_1}}\sum_{i=1}^{n_1}\mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|)\xrightarrow{\mathcal{D}}\mathcal{N}(0,1).$$

Similarly, we will obtain

$$\frac{1}{s_{n_1}+s_{n_2}}\big(\widehat{\text{MLD}}_1+\widehat{\text{MLD}}_2\big)-\frac{1}{s_{n_1}+s_{n_2}}\left(\sum_{i=1}^{n_1}\mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|)+\sum_{i=1}^{n_2}\mathbb{E}(|a_{x_{2i}}^{\tilde{p}}|)\right)\xrightarrow{\mathcal{D}}\mathcal{N}(0,1),$$

where

$$s_{n_2}^2=\sum_{i=1}^{n_2}\mathbb{E}\Big(\big(|a_{x_{2i}}^{\tilde{p}}|-\mathbb{E}(|a_{x_{2i}}^{\tilde{p}}|)\big)^2\Big).$$

Then, we will show how to calculate $\mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|)$, $\mathbb{E}(|a_{x_{2i}}^{\tilde{p}}|)$, s_{n_1} and s_{n_2} using the fact

$$\frac{1}{\sigma_{x_{1i}}}L_d(x_{1i})\xrightarrow{\mathcal{D}}\mathcal{N}(0,1),$$

where $\lambda_1=\lim_{n\to\infty}n_1/n$, $\lambda_2=\lim_{n\to\infty}n_2/n$, $\sigma_{x_{1i}}^2=(\lambda_1 n_1^{-2}+\lambda_2 n_2^{-2})\sum_{j=1}^{n}k(z_j,x_{1i})^2/V^2(r_{x_{1i}})$ and $\mathcal{D}$ is "converges in distribution". We assume that Y is a random variable that obeys standard normal distribution. Then, based on the continuous mapping theorem, we have

$$\frac{1}{\sigma_{x_{1i}}^{\tilde{p}}}|L_d^{\tilde{p}}(x_{1i})|\xrightarrow{\mathcal{D}}|Y^{\tilde{p}}|.$$

Because $|L_d^{\tilde{p}}(x_{1i})|$ is finite, $|L_d^{\tilde{p}}(x_{1i})|$ is uniformly integrable, which means that $\lim_{n\to\infty}\mathbb{E}(|L_d^{\tilde{p}}(x_{1i})|^k/\sigma_{x_{1i}}^{k\tilde{p}}) = \mathbb{E}(|Y^{\tilde{p}}|^k)$, $k=1,2,3$. Because Y obeys a standard normal distribution, we can calculate the probability density function related to $|Y^{\tilde{p}}|$ as follows.

$$f_{|Y^{\tilde{p}}|}(x) = \frac{2}{\sqrt{2\pi}\tilde{p}} x^{\frac{1}{\tilde{p}}-1} exp(-\frac{x^{\frac{2}{\tilde{p}}}}{2}). \tag{48}$$

Then, using the definition of the gamma function and the integral of the probability density function of the gamma distribution equals 1, we have

$$\begin{aligned}\mathbb{E}(|L_d^{\tilde{p}}(x_{1i})|^k/\sigma_{x_{1i}}^{k\tilde{p}}) &= \frac{2}{\sqrt{2\pi}\tilde{p}}\int_0^{+\infty} x^{\frac{1}{\tilde{p}}-1+k} exp(-\frac{x^{\frac{2}{\tilde{p}}}}{2})dx \\ &= \frac{2}{\sqrt{2\pi}\tilde{p}}\int_0^{+\infty} t^{\frac{\tilde{p}}{2}(\frac{1}{\tilde{p}}-1+k)} exp(-\frac{t}{2})\frac{\tilde{p}}{2} t^{\frac{\tilde{p}}{2}-1}dt \\ &= \frac{1}{\sqrt{2\pi}}\int_0^{+\infty} t^{\frac{1}{2}(1+k\tilde{p})-1} exp(-\frac{t}{2})dt \\ &= \frac{1}{\sqrt{2\pi}} 2^{\frac{1}{2}(1+k\tilde{p})}\Gamma\Big(\frac{1}{2}(1+k\tilde{p})\Big).\end{aligned} \tag{49}$$

In Equation (49), we use $t = x^{2/\tilde{p}}$. So, we have

$$\mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|) = \mathbb{E}(|L_d^{\tilde{p}}(x_{1i})|/n_1) = \frac{2^{\frac{1}{2}(1+\tilde{p})}\Gamma\Big(\frac{1}{2}(1+\tilde{p})\Big)}{n_1\sqrt{2\pi}}\sigma_{x_{1i}}^{\tilde{p}}, \tag{50}$$

$$\begin{aligned}\mathbb{E}\Big(\big(|a_{x_{1i}}^{\tilde{p}}| - \mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|)\big)^2\Big) &= \mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|^2) - (\mathbb{E}(|a_{x_{1i}}^{\tilde{p}}|))^2 \\ &= \frac{\sigma_{x_{1i}}^{2\tilde{p}}}{n_1^2\sqrt{2\pi}} 2^{\frac{1}{2}(1+2\tilde{p})}\Gamma\Big(\frac{1}{2}(1+2\tilde{p})\Big) - \frac{\sigma_{x_{1i}}^{2\tilde{p}}}{2\pi n_1^2} 2^{1+\tilde{p}}\Big(\Gamma\Big(\frac{1}{2}(1+\tilde{p})\Big)\Big)^2 \\ &= \frac{2^{\frac{1}{2}+\tilde{p}}\Gamma\Big(\frac{1}{2}+\tilde{p}\Big) - \frac{2^{1+\tilde{p}}}{\sqrt{2\pi}}\Big(\Gamma\Big(\frac{1}{2}(1+\tilde{p})\Big)\Big)^2}{n_1^2\sqrt{2\pi}}\sigma_{x_{1i}}^{2\tilde{p}}.\end{aligned} \tag{51}$$

According to definition of s_{n_1}, we know

$$s_{n_1} = \frac{2^{\frac{1}{2}+\tilde{p}}\Gamma\Big(\frac{1}{2}+\tilde{p}\Big) - \frac{2^{1+\tilde{p}}}{\sqrt{2\pi}}\Big(\Gamma\Big(\frac{1}{2}(1+\tilde{p})\Big)\Big)^2}{n_1^2\sqrt{2\pi}}\sum_{i=1}^{n_1}\sigma_{x_{1i}}^{2\tilde{p}}. \tag{52}$$

Similarly, we can give s_{n_2} and we have

$$\frac{\widehat{\mathrm{MLD}}^{\tilde{p}}(X_1,X_2) - \mu_{\mathrm{MLD}}}{\sigma_{\mathrm{MLD}}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \tag{53}$$

where $\mathcal{D}$ means “converges in distribution”, and both μ_{MLD} and σ_{MLD} are related to $\sigma_{x_{li}}^2$ in Lemma 3, i.e.,

$$\mu_{\mathrm{MLD}} = C_\mu\left(\sum_{i=1}^{n_1}\frac{\sigma_{x_{1i}}^{\tilde{p}}}{n_1} + \sum_{i=1}^{n_2}\frac{\sigma_{x_{2i}}^{\tilde{p}}}{n_2}\right)$$

and

$$\sigma_{\mathrm{MLD}}^2 = C_\sigma\left(\sum_{i=1}^{n_1}\frac{\sigma_{x_{1i}}^{2\tilde{p}}}{n_1^2} + \sum_{i=1}^{n_2}\frac{\sigma_{x_{2i}}^{2\tilde{p}}}{n_2^2}\right),$$

where

$$C_\mu = \frac{2^{\frac{1}{2}(1+\tilde{p})}\Gamma\Big(\frac{1}{2}(1+\tilde{p})\Big)}{\sqrt{2\pi}},$$

with $\Gamma(\cdot)$ referring to the gamma function,

$$C_\sigma = \frac{2^{\frac{1}{2}+\tilde{p}}\Gamma\Big(\frac{1}{2}+\tilde{p}\Big) - \frac{2^{1+\tilde{p}}}{\sqrt{2\pi}}\Big(\Gamma\Big(\frac{1}{2}(1+\tilde{p})\Big)\Big)^2}{\sqrt{2\pi}},$$

$\sigma_{x_{1i}}^2 = (\lambda_1 n_1^{-2} + \lambda_2 n_2^{-2})\sum_{j=1}^n k(z_j,x_{1i})^2/V^2(r_{x_{1i}})$, and $\sigma_{x_{2i}}^2 = (\lambda_1 n_1^{-2} + \lambda_2 n_2^{-2})\sum_{j=1}^n k(z_j,x_{2i})^2/V^2(r_{x_{2i}})$.