Supplementary Material for "Private Minimum Hellinger Distance Estimation via Hellinger Distance Differential Privacy"

Fengnan Deng fdeng2@gmu.edu
Department of Statistics
George Mason University
Fairfax, VA 22030, USA

Anand N. Vidyashankar avidyash@gmu.edu
Department of Statistics
George Mason University
Fairfax, VA 22030, USA

Editor: Action Editor name

6 Proofs

In this section we provide the proofs of the main results of the paper.

6.1 Proof of Theorem 6 and 8

We start with the proof of Theorem 6. We begin with the case $\lambda(\lambda+1) \neq 0$. Since $M_1(w, D)$ is (λ, ϵ)-PDP, the power divergence between $M_1(w, D)$ and $M_1(w, D')$ is atmost ϵ_1. For brevity, we denote the random variables $M_1(w, D)$ and $M_1(w, D')$ by X_1 and X_2 respectively. Let $p_1(\cdot)$ and $p_2(\cdot)$ denote their densities. Thus, by the $(\lambda, \epsilon)-$PDP property it follows that,

$$\frac{1}{\lambda(\lambda+1)} \mathbf{E}_{p_2}\left[\left(\frac{p_1(X_2)}{p_2(X_2)}\right)^{\lambda+1}\right] \leq \epsilon_1 + \frac{1}{\lambda(\lambda+1)} \tag{1}$$

Next, let the random variables $Y_1|X_1$ and $Y_2|X_2$ represent the compositions of mechanisms $M_2(M_1(w, D), D)$ given $M_1(w, D)$ and $M_2(M_1(w, D'), D')$ given $M_1(w, D')$ with conditional densities $q_{Y_1|X_1}(\cdot)$ and $q_{Y_2|X_2}(\cdot)$ respectively. Again using the PDP property, it follows that for a generic random variable $V \sim q_{Y_2|X_2}(\cdot)$,

$$\frac{1}{\lambda(\lambda+1)} \mathbf{E}_{q_{Y_2|X_2}}\left[\left(\frac{q_{Y_1|X_1}(V)}{q_{Y_2|X_2}(V)}\right)^{\lambda+1}\right] \leq \epsilon_2 + \frac{1}{\lambda(\lambda+1)} \tag{2}$$

Now, to calculate power divergence of $M^{(2)}(w, D)$ and $M^{(2)}(w, D')$, we need to calculate the joint power divergence of (X_1, Y_1) and (X_2, Y_2). Now, using that the joint density is the product of conditional density and the marginal density, it follows that

$$D_\lambda(M^{(2)}(w, D), M^{(2)}(w, D')) = \frac{1}{\lambda(\lambda+1)} \mathbf{E}_{p_2}\left\{\left(\frac{p_1(X_2)}{p_2(X_2)}\right)^{\lambda+1} \mathbf{E}_{q_{Y_2|X_2}}\left[\left(\frac{q_{Y_1|X_1}(V)}{q_{Y_2|X_2}(V)}\right)^{\lambda+1}\right] - 1\right\}.$$

Now, suppose $\lambda(\lambda+1) > 0$. Then, using (1) and (2) it follows that

$$\begin{aligned} D_\lambda(M^{(2)}(w,D), M^{(2)}(w,D')) &\leq \frac{1}{\lambda(\lambda+1)}[\epsilon_1(\lambda(\lambda+1))+1][\epsilon_2(\lambda(\lambda+1))+1] - \frac{1}{\lambda(\lambda+1)} \\ &= \epsilon_1+\epsilon_2+\lambda(\lambda+1)\epsilon_1\epsilon_2. \end{aligned}$$

Next, if $\lambda(\lambda+1) < 0$, then using the condition $0 < \epsilon < -[\lambda(\lambda+1)]^{-1}$ the above inequality continues to hold. Finally, consider the case $\lambda = 0$. In this case, the power divergence reduces to the Kullback Leibler (KL) divergence between the densities. Hence, using the PDP property, it follows that

$$D_0(p_1,p_2) = KL(p_1,p_2) \leq \epsilon_1, \quad \text{and} \quad D_0(q_{Y_1|X_1}, q_{Y_2|X_2}) = KL(q_{Y_1|X_1}, q_{Y_2|X_2}) \leq \epsilon_2.$$

Hence, with $V \sim q_{Y_1|X_1}(\cdot)$, we have

$$D_\lambda(M^{(2)}(w,D), M^{(2)}(w,D')) = \mathbf{E}_{p_1}\left[\log\frac{p_1(X_2)}{p_2(X_2)}\right] + \mathbf{E}_{q_{Y_1|X_1}}\left[\log\frac{q_{Y_1|X_1}(V)}{q_{Y_2|X_2}(V)}\right] \leq \epsilon_1+\epsilon_2.$$

The proof of the case $\lambda = -1$ is similar. This completes the proof of (1). Next, the proof of (2) follows exactly as in (1) except that $M_2(X_1,D)|X_1$ and $M_2(X_2,D')|X_2$ are now replaced by $M_2(w,D)$ and $M_2(w,D')$. That is, we replace $\frac{q_{Y_1|X_1}(V)}{q_{Y_2|X_2}(V)}$ in (2) by $\frac{q_{Y_1}(V)}{q_{Y_2}(V)}$, where q_{Y_1} and q_{Y_2} are the unconditional distributions of $M_2(w,D)$ and $M_2(w,D')$ respectively and V has the density q_{Y_2}.

To start the proof of part (3), we first notice that adjacent $\mathbf{D}^{(2)}$ and $\mathbf{D}^{(2)'}$ can be decomposed into two distinct cases. By definition,

$$||\mathbf{D}^{(2)} - \mathbf{D}^{(2)'}||_H = \sum_{i=1}^{2} ||D_i - D_i'||_H = 1.$$

Since the Hamming distance is a non-negative integer, the above equation holds if either: Case (1): $||D_1 - D_1'||_H = 1$ and $||D_2 - D_2'||_H = 0$; or Case (2): $||D_1 - D_1'||_H = 0$ and $||D_2 - D_2'||_H = 1$. Let, as before, $p_1(\cdot)$, $p_2(\cdot)$ denote the distributions of $M_1(w_1,D_1)$, $M_1(w_1,D_1')$. Also, let $q_1(\cdot)$, and $q_2(\cdot)$ are density functions of $M_2(w_2,D_2)$ and $M_2(w_2,D_2')$ respectively. The joint density of $M^{(2)}(\mathbf{W}^{(2)},\mathbf{D}^{(2)})$ and $M^{(2)}(\mathbf{W}^{(2)},\mathbf{D}^{(2)'})$ are therefore given by $h_1(x,y)$ and $h_2(x,y)$ respectively, where $h_1(x,y) = p_1(x)q_1(y)$ and $h_2(x,y) = p_2(x)q_2(y)$.

If case (1) happens, $q_1(\cdot) = q_2(\cdot)$ holds since $D_2 = D_2'$. This leads to $\frac{h_1(x,y)}{h_2(x,y)} = \frac{p_1(x)}{p_2(x)}$. The PD between $M^{(2)}(\mathbf{W}^{(2)},\mathbf{D}^{(2)})$ and $M^{(2)}(\mathbf{W}^{(2)},\mathbf{D}^{(2)'})$ is reduced to the PD between $M_1(w_1,D_1)$ and $M_1(w_1,D_1')$, since

$$\begin{aligned} D_\lambda(M^{(2)}(\mathbf{W}^{(2)},\mathbf{D}^{(2)}), M^{(2)}(\mathbf{W}^{(2)},\mathbf{D}^{(2)'})) =& \frac{1}{\lambda(\lambda+1)}\mathbf{E}_{h_2}\left[\left(\frac{h_1(X,Y)}{h_2(X,Y)}\right)^{\lambda+1} - 1\right] \\ =& \frac{1}{\lambda(\lambda+1)}\mathbf{E}_{p_2}\left[\left(\frac{p_1(X)}{p_2(X)}\right)^{\lambda+1} - 1\right] \\ =& D_\lambda(M_1(w_1,D_1), M_1(w_1,D_1')) \leq \epsilon_1. \end{aligned}$$

The last inequality follows from $M_1(w_1, D_1)$ is $(\lambda, \epsilon_1)-$PDP. Similarly in Case (2),

$$D_\lambda(M^{(2)}(\mathbf{W}^{(2)}, \mathbf{D}^{(2)}), M^{(2)}(\mathbf{W}^{(2)}, \mathbf{D}^{(2)'})) = D_\lambda(M_2(w_2, D_2), M_2(w_2, D_2')) \leq \epsilon_2.$$

Combining Case (1) and Case (2), we get $D_\lambda(M^{(2)}(\mathbf{W}^{(2)}, \mathbf{D}^{(2)}), M^{(2)}(\mathbf{W}^{(2)}, \mathbf{D}^{(2)'})) \leq \max\{\epsilon_1, \epsilon_2\}$, which implies that the parallel composition is $(\lambda, \max\{\epsilon_1, \epsilon_2\})-$PDP. This completes the proof of Theorem 6. The Proof of Theorem 8 follows by taking $\lambda = -\frac{1}{2}$ and noticing that the objective function is $D_{-\frac{1}{2}}(\cdot, \cdot) = 2D^2_{HD}(\cdot, \cdot)$. ■

6.2 Proof of Theorem 7 and Proposition 12

We begin with the case when Y_i's are $N(0, \sigma^2)$. In this case, using Lemma B.1 we have that

$$D_\lambda(M(w, D), M(w, D')) \quad = \quad \frac{1}{\lambda(\lambda+1)}\left[e^{\frac{\lambda(\lambda+1)||\mathbf{v}||_2^2}{2\sigma^2}} - 1\right],$$

where $\mathbf{v}$ is the difference between the mean of $M(w, D)$ and the mean of $M(w, D')$, and for $r = 1, 2$, $||\mathbf{v}||_r = \Delta_{L_r} W$. Thus, $D_\lambda(M(w, D), M(w, D')) \leq \epsilon$ is equivalent to

$$\sigma \geq ||\mathbf{v}||_2 \sqrt{\frac{\lambda(\lambda+1)}{2\log(1+\lambda(\lambda+1)\epsilon)}},$$

which is well-defined for all values of $0 < \epsilon < [-\lambda(\lambda+1)]^{-1}$. Thus, we choose

$$\sigma^2_{\lambda,\epsilon} = (\Delta_{L_2} W)^2 \frac{\lambda(\lambda+1)}{2\log(1+\lambda(\lambda+1)\epsilon)}.$$

Next, when $\lambda = 0$,

$$D_0(M(w, D), M(w, D')) = \sum_{i=1}^{m} \frac{||w_1 - w_2||_2^2}{2\sigma^2}$$

Again, $D_0 M(w, D), M(w, D') \leq \epsilon$ is equivalent to

$$\sigma^2_{0,\epsilon} = \frac{(\Delta_{L_2} W)^2}{2\epsilon}.$$

The case for $\lambda = -1$ is similar. Next, turning to the Laplace case, using Lemma B.2, notice that

$$\begin{aligned} D_\lambda(M(w, D), M(w, D')) &= \frac{1}{\lambda(\lambda+1)}\left[\left(\prod_{i=1}^{m} \int_{\mathbb{R}} \frac{1}{2b} e^{-\frac{(\lambda+1)|y_i - v_i| - \lambda|y_i|}{b}} \mathrm{d}y_i\right) - 1\right] \\ &\leq \frac{1}{\lambda(\lambda+1)}\left[e^{\frac{sign(\lambda)(\lambda+1)||v||_1}{b}} - 1\right]. \end{aligned}$$

Similarly,

$$D_\lambda(M(w, D'), M(w, D)) \leq \frac{1}{\lambda(\lambda+1)}\left[e^{\frac{sign(\lambda+1)(\lambda)||v||_1}{b}} - 1\right].$$

Thus, $D_\lambda(M(w,D), M(w,D')) \le \epsilon$ is equivalent to

$$b \ge \max\left\{\frac{sign(\lambda)(\lambda+1)||v||_1}{\log(\lambda(\lambda+1)\epsilon+1)}, \frac{sign(\lambda+1)(\lambda)||v||_1}{\log(\lambda(\lambda+1)\epsilon+1)}\right\}.$$

Thus, we choose $b_{\lambda,\epsilon}$ to be

$$b_{\lambda,\epsilon} = \max\left\{\frac{sign(\lambda)(\lambda+1)\Delta_{L_1}W}{\log(\lambda(\lambda+1)\epsilon+1)}, \frac{sign(\lambda+1)(\lambda)\Delta_{L_1}W}{\log(\lambda(\lambda+1)\epsilon+1)}\right\}.$$

Turning to the case $\lambda(\lambda+1) = 0$, we note that

$$\begin{aligned} D_0(M(w,D), M(w,D')) &= \sum_{i=1}^{m} \int_{\mathbb{R}} \frac{1}{2b} e^{-\frac{|y_i|}{b}} \cdot \frac{|y_i - v_i| - |y_i|}{b} \mathrm{d}y_i \\ &\le \frac{||w_1 - w_2||_1}{b} \end{aligned}$$

Hence, $D_0(M(w,D), M(w,D')) \le \epsilon$ implies $b \ge \frac{\Delta w}{\epsilon}$. Thus,

$$b_{0,\epsilon} = \frac{\Delta_{L_1}W}{\epsilon}.$$

The proof for the case $\lambda = -1$ is similar. Finally, the proof for the HDP case follows by taking $\lambda = -\frac{1}{2}$ and replacing ϵ by 2ϵ to obtain $\epsilon-$HDP. ■

6.3 Proof of Corollary 9

The proof is based on the following iterative argument for adaptive and sequential compositions. Setting $\epsilon_1 = \epsilon$ and $\epsilon_2 = h_1(\epsilon)$ it follows from Theorem 8 part 1. and part 2., that $h_2(\epsilon) = \epsilon + h_1(\epsilon) - \frac{1}{2}\epsilon h_1(\epsilon)$. Now iterating, we obtain $h_{j+1}(\epsilon) = \epsilon + h_j(\epsilon) - \frac{1}{2}\epsilon h_j(\epsilon)$. The proof for the parallel compositions follows from part 3. of Theorem 8. ■

6.4 Proof of Proposition 10

The proof of the Proposition follows using a comparison argument. Recall that the total variation distance between two densities can be expressed as one-half the L_1-norm, which is bounded above by the Hellinger distance between the densities. That is,

$$TV(p_1,p_2) = \frac{1}{2}||p_1 - p_2||_1 \le HD(p_1,p_2).$$

Now, if $HD^2(p_1,p_2) \le \epsilon$, then $TV(p_1,p_2) \le \sqrt{\epsilon}$ which implies that $M(\cdot,\cdot)$ satisfies $\sqrt{\epsilon}-$TV privacy. Hence, using Ghazi and Issa (2024) page 209, it follows that M also satisfies $(0,\sqrt{\epsilon})$ differential privacy. Turning to $\mu-$GDP, we now use the Corollary 1 in Dong et al. (2022) to get $\mu = 2\Phi^{-1}(\frac{\sqrt{\epsilon}+1}{2})$. ■

6.5 Proof of Theorem 11

First notice that by the definition of group privacy, we need to calculate $D_{HD}(M(w,D), M(w,D'))$ for $k-$neighbor datasets D and D'. Now, by definition of $k-$neighbor datasets, there exist

$D = B_0, B_1, B_2, \cdots B_k = D'$, such that $||B_i - B_{i+1}||_H = 1$ for all $i = 0, 1, \cdots, (k-1)$. Also, since $M(w, B_i)$ is $\epsilon-$HDP for all $0 \le i \le k$, we get that

$$D_{HD}(M(w, B_i), M(w, B_{i+1})) \le \epsilon.$$

Now, using that $D_{HD}^{\frac{1}{2}}(\cdot, \cdot) = HD(\cdot, \cdot)$ is a metric, using the triangle inequality

$$D_{HD}^{\frac{1}{2}}(M(w, D), M(w, D')) \le \sum_{i=1}^{k} D_{HD}^{\frac{1}{2}}(M(w, B_{i-1}), M(w, B_i)) \le k\sqrt{\epsilon}. \tag{3}$$

The result follows by squaring both sides of (3). ■

6.6 Proof of Proposition 19

To show $H_n(\boldsymbol{\theta})$ is $\alpha-$Lipschitz continuous, it is enough to show that the $(i, j)^{th}$ component of $H_n(\boldsymbol{\theta})$ is Lipschitz continuous for all (i, j), where the Lipschitz constant depends only on m and the upper bounds in assumptions **(U1)** and **(U2)**. That is, we will show that for any $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)} \in \Theta$,

$$|H_{n,i,j}(\boldsymbol{\theta}^{(1)}) - H_{n,i,j}(\boldsymbol{\theta}^{(2)})| \le \alpha_{i,j} ||\boldsymbol{\theta}^{(1)} - \boldsymbol{\theta}^{(2)}||_2,$$

where $H_{n,i,j}(\boldsymbol{\theta})$ is the $(i, j)^{th}$ component of $H_n(\boldsymbol{\theta})$. Recall that

$$H_{n,i,j}(\boldsymbol{\theta}) = -\int_{\mathbb{R}} g_n^{\frac{1}{2}}(x) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},j}(x) u_{\theta,i}(x) \mathrm{d}x - 2\int_{\mathbb{R}} g_n^{\frac{1}{2}}(x) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},i,j}(x) \mathrm{d}x.$$

Then for any $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)} \in \Theta$,

$$\begin{aligned} |H_{n,i,j}(\boldsymbol{\theta}^{(1)}) - H_{n,i,j}(\boldsymbol{\theta}^{(2)})| \quad \le \quad & \int_{\mathbb{R}} g_n^{\frac{1}{2}}(x) |T_{1,i,j}(\boldsymbol{\theta}^{(1)}, x) - T_{1,i,j}(\boldsymbol{\theta}^{(2)}, x)| \mathrm{d}x \\ & +2\int_{\mathbb{R}} g_n^{\frac{1}{2}}(x) |T_{2,i,j}(\boldsymbol{\theta}^{(1)}, x) - T_{2,i,j}(\boldsymbol{\theta}^{(2)}, x)| \mathrm{d}x, \end{aligned} \tag{4}$$

where

$$T_{1,i,j}(\boldsymbol{\theta}, x) = f_{\theta}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},j}(x) u_{\boldsymbol{\theta},i}(x), \quad T_{2,i,j}(\boldsymbol{\theta}, x) = f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},i,j}(x).$$

Notice that for $\boldsymbol{\theta} \in \Theta$, $T_{1,i,j}(\boldsymbol{\theta}, x)$ and $T_{2,i,j}(\boldsymbol{\theta}, x)$ are differentiable in $\boldsymbol{\theta}$ by Assumption **(U2)**. Using Cauchy-Schwarz inequality and the upper bounds in Assumption **(U2)**, for $\boldsymbol{\theta} \in \Theta$ we obtain that $g_n^{1/2}(x) ||\nabla T_{1,i,j}(\boldsymbol{\theta}, x)||_2$ and $g_n^{1/2}(x) ||\nabla T_{2,i,j}(\boldsymbol{\theta}, x)||_2$ are integrable with respect to x, where the gradient is taken with respect to $\boldsymbol{\theta}$. By the mean value theorem and Cauchy-Schwarz inequality, there exists $\boldsymbol{\theta}^{(1)*}$ and $\boldsymbol{\theta}^{(2)*}$ on the line between $\boldsymbol{\theta}^{(1)}$ and $\boldsymbol{\theta}^{(2)}$, such that

$$|T_{1,i,j}(\boldsymbol{\theta}^{(1)}, x) - T_{1,i,j}(\boldsymbol{\theta}^{(2)}, x)| \le ||\nabla T_{1,i,j}(\boldsymbol{\theta}^{(1)*}, x)||_2 \cdot ||\boldsymbol{\theta}^{(1)} - \boldsymbol{\theta}^{(2)}||_2 \tag{5}$$

$$|T_{2,i,j}(\boldsymbol{\theta}^{(1)}, x) - T_{2,i,j}(\boldsymbol{\theta}^{(2)}, x)| \le ||\nabla T_{2,i,j}(\boldsymbol{\theta}^{(2)*}, x)||_2 \cdot ||\boldsymbol{\theta}^{(1)} - \boldsymbol{\theta}^{(2)}||_2. \tag{6}$$

By the convexity of Θ (see Assumption **(A1)**), we obtain $\boldsymbol{\theta}^{(1)*}, \boldsymbol{\theta}^{(2)*} \in \Theta$. Now, multiplying both sides of (5) and (6) by $g_n^{\frac{1}{2}}(\cdot)$ and using the integrability described above, it follows that

$$\int g_n^{1/2}(x)||\nabla T_{1,i,j}(\boldsymbol{\theta}, x)||_2 \mathrm{d}x \leq \left(\sup_{\boldsymbol{\theta}\in\Theta} \int g_n^{1/2}(x)||\nabla T_{1,i,j}(\boldsymbol{\theta}, x)||_2 \mathrm{d}x \right) ||\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2||_2.$$

Using similar arguments for $T_{2,i,j}$ and setting

$$0 < \alpha_{i,j} = \sup_{\boldsymbol{\theta}\in\Theta} \left\{ \int g_n^{1/2}(x)||\nabla T_{1,i,j}(\boldsymbol{\theta}, x)||_2 \mathrm{d}x + 2 \int g_n^{1/2}(x)||\nabla T_{2,i,j}(\boldsymbol{\theta}, x)||_2 \mathrm{d}x \right\} < \infty.$$

It follows that

$$|H_{n,i,j}(\boldsymbol{\theta}^{(1)}) - H_{n,i,j}(\boldsymbol{\theta}^{(2)})| \leq \alpha_{i,j}||\boldsymbol{\theta}^{(1)} - \boldsymbol{\theta}^{(2)}||_2.$$

This completes the proof. ∎

Before we prove the theorem, we recall that

$$g_n(x) = \frac{1}{nc_n} \sum_{i=1}^{n} K\left(\frac{x - X_i}{c_n}\right).$$

and for the neighboring i.i.d. observations $\{X_1', X_2, \cdots, X_n\}$, the corresponding density estimator is

$$\tilde{g}_n(x) = \frac{1}{nc_n} \sum_{i=2}^{n} K\left(\frac{x - X_i}{c_n}\right) + \frac{1}{nc_n} K\left(\frac{x - X_1'}{c_n}\right).$$

The corresponding loss functions are given by

$$L_n(\boldsymbol{\theta}) = 2HD^2(g_n, f_{\boldsymbol{\theta}}) \;\text{ and }\; \tilde{L}_n(\boldsymbol{\theta}) = 2HD^2(\tilde{g}_n, f_{\boldsymbol{\theta}}),$$

and the Hessian of the loss functions are given by $H_n(\boldsymbol{\theta})$ and $\tilde{H}_n(\boldsymbol{\theta})$.

6.7 Proof of Proposition 20 and Theorem 24

We begin with the proof of Proposition 20. First, notice that for all $1 \leq i \leq m$

$$\frac{\partial}{\partial\theta_i} L_n(\boldsymbol{\theta}) = -2 \int_{\mathbb{R}} g_n^{\frac{1}{2}}(x) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},i}(x) \mathrm{d}x$$

Hence,

$$\bar{\Delta}^{(i)} := \frac{\partial}{\partial\theta_i}(L_n(\boldsymbol{\theta}) - \tilde{L}_n(\boldsymbol{\theta})) = 2 \int_{\mathbb{R}} (g_n^{\frac{1}{2}}(x) - \tilde{g}_n^{\frac{1}{2}}(x)) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},i}(x) \mathrm{d}x$$

where we have suppressed n in the notation $\bar{\Delta}^{(i)}$. Using Cauchy-Schwarz inequality, the $HD^2(g_n, \tilde{g}_n)$ is bounded above by $||g_n - \tilde{g}_n||_1$ and, assumption **(A2)** it follows that

$$\bar{\Delta}^{(i)} \leq 2HD(g_n, \tilde{g}_n) \cdot \left[\int_{\mathbb{R}} f_{\boldsymbol{\theta}}(x) u_{\boldsymbol{\theta},i}^2(x) \mathrm{d}x \right]^{1/2} \leq C_i ||g_n - \tilde{g}_n||_1^{\frac{1}{2}}.$$

Now,

$$||g_n - \tilde{g}_n||_1^{1/2} = \left[\frac{1}{n \cdot c_n} \int_{\mathbb{R}} |K\left(\frac{x - X_1}{c_n}\right) - K\left(\frac{x - X_1'}{c_n}\right)|\mathrm{d}x\right]^{1/2} \leq \left(\frac{2}{n}\right)^{1/2}. \tag{7}$$

Hence $\bar{\Delta}^{(i)} \leq 2C_i\left(\frac{2}{n}\right)^{1/2}$. Let $\bar{\boldsymbol{\Delta}} = [\bar{\Delta}^{(1)}, \cdots, \bar{\Delta}^{(m)}]$. Then $\Delta_{L_1}(\nabla L_n(\boldsymbol{\theta})) \leq ||\bar{\boldsymbol{\Delta}}||_1 \leq C \cdot m \cdot n^{-\frac{1}{2}}$ and $\Delta_{L_2}(\nabla L_n(\boldsymbol{\theta})) \leq ||\bar{\boldsymbol{\Delta}}||_2 \leq C \cdot \sqrt{m} \cdot n^{-\frac{1}{2}}$, where $C = \max_i\{2\sqrt{2}C_i\}$.

We now turn to the Hessian. Recall the definition $u_{\boldsymbol{\theta},i}(x) = \frac{1}{f_{\boldsymbol{\theta}}(x)} \cdot \frac{\partial}{\partial\theta_i} f_{\boldsymbol{\theta}}(x)$, $u_{\boldsymbol{\theta},i,j} = \frac{\partial}{\partial\theta_j} u_{\boldsymbol{\theta},i}$, and

$$H_{n,i,j}(\boldsymbol{\theta}) = -\int_{\mathbb{R}} g_n^{\frac{1}{2}}(x) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x)[u_{\boldsymbol{\theta},j}(x)u_{\boldsymbol{\theta},i}(x) + 2u_{\boldsymbol{\theta},i,j}(x)]\mathrm{d}x.$$

Hence,

$$\bar{\Delta}^{(i,j)} := H_{n,i,j}(\theta) - \tilde{H}_{n,i,j}(\theta) = \int_{\mathbb{R}} (g_n^{\frac{1}{2}}(x) - \tilde{g}_n^{\frac{1}{2}}(x)) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x)[u_{\boldsymbol{\theta},j}(x)u_{\boldsymbol{\theta},i}(x) + 2u_{\boldsymbol{\theta},i,j}(x)]\mathrm{d}x$$

Using Cauchy-Schwarz inequality, the $HD^2(g_n, \tilde{g}_n)$ is bounded above by $||g_n - \tilde{g}_n||_1$ and, assumptions **(U1)-(U2)** it follows that

$$\begin{aligned}
|\bar{\Delta}^{(i,j)}| \leq& 2HD(g_n, \tilde{g}_n) \cdot \left[\int_{\mathbb{R}} f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x)[u_{\boldsymbol{\theta},j}(x)u_{\boldsymbol{\theta},i}(x) + 2u_{\boldsymbol{\theta},i,j}(x)]\right]^{1/2} \\
\leq& (C_{i,1} + C_{i,2})||g_n - \tilde{g}_n||_1^{\frac{1}{2}},
\end{aligned}$$

where $C_{i,1}$ and $C_{i,2}$ are upper bounds given in assumptions **(U1)** and **(U2)**. Using (7), it follows that $\bar{\Delta}^{(i,j)} \leq C_{i,j} \cdot n^{-\frac{1}{2}}$. Let $\bar{\boldsymbol{\Delta}}$ be a $m \times m$ matrix with $(i,j)^{\text{th}}$ element $\bar{\Delta}^{(i,j)}$. Then $\Delta_{L_1}(H_n(\boldsymbol{\theta})) \leq ||\bar{\boldsymbol{\Delta}}||_1 \leq C \cdot m \cdot n^{-\frac{1}{2}}$ and $\Delta_{L_2}(\nabla L_n(\boldsymbol{\theta})) \leq ||\bar{\boldsymbol{\Delta}}||_2 \leq C \cdot m \cdot n^{-\frac{1}{2}}$ for some $0 < C < \infty$. This completes the proof of Proposition 20.

We next turn to the Proof of Theorem 24. In this case, we require the sensitivity is taken on a compact set A_n as in assumption **(U3)**. To reduce notational complexity, *redefine* $\tilde{g}_n(x)$ as follows:

$$\tilde{g}_n(x) = \frac{1}{nc_n}\sum_{i=2}^{n} K\left(\frac{x - X_i}{c_n}\right)\mathbf{1}_{(X_i \in B_n)} + \frac{1}{nc_n} K\left(\frac{x - X_1'}{c_n}\right)\mathbf{1}_{(X_1' \in B_n)},$$

where $X_1, X_1' \in B_n$. Now the $\bar{\Delta}^{(i)}$ is given by

$$\bar{\Delta}^{(i)} := \bar{\Delta}^{(i,1)} + \bar{\Delta}^{(i,2)},$$

where

$$\begin{aligned}
\bar{\Delta}^{(i,1)} =& 2\int_{\mathbb{R}} (\bar{g}_n^{\frac{1}{2}}(x) - \tilde{g}_n^{\frac{1}{2}}(x)) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},i}(x) \cdot \mathbf{1}_{(x \in A_n)}\mathrm{d}x, \quad \text{and} \\
\bar{\Delta}^{(i,2)} =& 2\int_{\mathbb{R}} (\bar{g}_n^{\frac{1}{2}}(x) - \tilde{g}_n^{\frac{1}{2}}(x)) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},i}(x) \cdot \mathbf{1}_{(x \notin A_n)}\mathrm{d}x.
\end{aligned}$$

Using the equation $a^{\frac{1}{2}} - b^{\frac{1}{2}} = \frac{a-b}{2b^{1/2}} - \frac{(a^{1/2}-b^{1/2})^2}{2b^{1/2}}$, and denoting $R_{\boldsymbol{\theta}}(x) = |f_{\boldsymbol{\theta}}^{1/2}(x)u_{\boldsymbol{\theta},i}(x)|$, we obtain

$$
\begin{aligned}
|\bar{\Delta}^{(i,1)}| &\leq T_1 + T_2, \quad \text{where} \\
T_1 &= \int_{\mathbb{R}} \left| \frac{\bar{g}_n(x) - \tilde{g}_n(x)}{2\sqrt{\tilde{g}_n(x)}} \right| \cdot R_{\boldsymbol{\theta}}(x) \cdot \mathbf{1}_{(x\in A_n)} \mathrm{d}x \text{ and} \\
T_2 &= \int_{\mathbb{R}} \left| \frac{(\sqrt{\bar{g}_n(x)} - \sqrt{\tilde{g}_n(x)})^2}{2\sqrt{\tilde{g}_n(x)}} \right| \cdot R_{\boldsymbol{\theta}}(x) \cdot \mathbf{1}_{(x\in A_n)} \mathrm{d}x.
\end{aligned}
$$

We first develop the upper bound of T_1 and use the fact that $T_2 \leq T_1$ almost surely to get the final answer. Using the Hölder's inequality with $p \in (1,2)$ and integrability of $|R_{\boldsymbol{\theta}}(x)|^q$ in assumption **(U3)** and the boundedness of the kernel function $K(\cdot)$, it follows that

$$
\begin{aligned}
T_1 \leq & C_1 \cdot \left[\int_{\mathbb{R}} \left| \frac{g_n(x) - \tilde{g}_n(x)}{2\sqrt{\tilde{g}_n(x)}} \right|^p \cdot \mathbf{1}_{(x\in A_n)} \mathrm{d}x \right]^{\frac{1}{p}} \\
\leq & C_1 \cdot \left[\sup_{x\in A_n} \{ |g_n(x) - \tilde{g}_n(x)| \} \right]^{\frac{1}{p}} \cdot \left[\int_{\mathbb{R}} \frac{|g_n(x) - \tilde{g}_n(x)|^{p-1}}{(2\sqrt{\tilde{g}_n(x)})^p} \cdot \mathbf{1}_{(x\in A_n)} \mathrm{d}x \right]^{\frac{1}{p}} \\
\leq & C_1 \cdot C_2 \left(\frac{1}{n} \right)^{\frac{1}{p}} \cdot \left[\int_{\mathbb{R}} \frac{|g_n(x) - \tilde{g}_n(x)|^{p-1}}{(2\sqrt{\tilde{g}_n(x)})^p} \cdot \mathbf{1}_{(x\in A_n)} \mathrm{d}x \right]^{\frac{1}{p}}
\end{aligned}
$$

where $0 < C_1 < \infty$ is a constant (independent of $\boldsymbol{\theta}$) obtained from assumption **(U3)**. We turn to the last term on the RHS and show it converges to 0 almost surely under assumption **(U3)**. Notice that $\inf_{x\in A_n} \tilde{g}_n(x) \geq \delta_n$, we obtain

$$
\int_{\mathbb{R}} \frac{|g_n(x) - \tilde{g}_n(x)|^{p-1}}{(2\sqrt{\tilde{g}_n(x)})^p} \cdot \mathbf{1}_{(x\in A_n)} \mathrm{d}x \leq \int_{\mathbb{R}} \frac{|g_n(x) - \tilde{g}_n(x)|^{p-1}}{(2\sqrt{\delta_n})^p} \cdot \mathbf{1}_{(x\in A_n)} \mathrm{d}x.
$$

The RHS of the above equation is bounded above by

$$
\frac{1}{(2\sqrt{\delta_n})^p \cdot (n \cdot c_n)^{p-1}} \left[\int_{\mathbb{R}} \left| K(\frac{x - X_1'}{c_n}) - K(\frac{x - X_1}{c_n}) \right|^{p-1} \mathrm{d}x \right].
$$

Next, we establish the upper bound of the term above. By assumption **(A2)**, $K(\cdot)$ has compact support, say $[-\beta, \beta]$. Next, fix any $X_1 = x_1$, $X_1' = x_1'$ and $x_1, x_1' \in B_n$. Then write

$$
\begin{aligned}
S =& supp_x \left(K(\frac{x - x_1}{c_n}) \right) \cup supp_x \left(K(\frac{x - x_1'}{c_n}) \right) \\
=& [x_1 - \beta c_n, x_1 + \beta c_n] \cup [x_1' - \beta c_n, x_1' + \beta c_n]
\end{aligned}
$$

and note that $\lambda(S) \leq 4\beta c_n$, where $\lambda(S)$ is the Lebesgue measure of S. Notice that $h(x) = x^{p-1}$ is concave for $p \in (1,2)$ on $x \in (0,\infty)$, using Jensen's inequality, it follows that

$$\begin{aligned}
\int_{\mathbb{R}} \left| K(\frac{x-x_1'}{c_n}) - K(\frac{x-x_1}{c_n}) \right|^{p-1} \mathrm{d}x =& \lambda(S) \int_S \left| K(\frac{x-x_1'}{c_n}) - K(\frac{x-x_1}{c_n}) \right|^{p-1} \cdot \frac{1}{\lambda(S)} \mathrm{d}x \\
=& \lambda(S) \mathbf{E} \left[\left| K(\frac{X-x_1'}{c_n}) - K(\frac{X-x_1}{c_n}) \right|^{p-1} \right] \\
\leq& \lambda(S) \mathbf{E} \left[\left| K(\frac{X-x_1'}{c_n}) - K(\frac{X-x_1}{c_n}) \right| \right]^{p-1} \\
=& \lambda(S) \left[\int_S \left| K(\frac{x-x_1'}{c_n}) - K(\frac{x-x_1}{c_n}) \right| \cdot \frac{1}{\lambda(S)} \mathrm{d}x \right]^{p-1} \\
=& [\lambda(S)]^{2-p} \left[\int_S \left| K(\frac{x-x_1'}{c_n}) - K(\frac{x-x_1}{c_n}) \right| \mathrm{d}x \right]^{p-1} \\
\leq& C_3 \cdot c_n.
\end{aligned}$$

Hence by assumption **(U3)**,

$$\int_{\mathbb{R}} \left[\frac{|g_n(x) - \tilde{g}_n(x)|^{p-1}}{(2\sqrt{\tilde{g}_n(x)})^p} \cdot \mathbf{1}_{(x \in A_n)} \right] \mathrm{d}x \leq C_4 \cdot \frac{c_n^{2-p}}{\delta_n^{p/2} \cdot n^{p-1}} \to 0.$$

Therefore, we proved $T_1 \leq C_5 \left(\frac{1}{n}\right)^{1/p}$ for some $C_5 \in (0,\infty)$. Next, we show $T_2 \leq T_1$ almost surely. To this end,

$$T_2 \leq \int \frac{|\sqrt{g_n(x)} - \sqrt{g_n'(x)}| \cdot |\sqrt{g_n(x)} + \sqrt{g_n'(x)}|}{2\sqrt{g_n'(x)}} \cdot R_\theta(x) \mathrm{d}x = T_1$$

We turn to $\bar{\Delta}^{(i,2)}$. Using Cauchy-Schwarz inequality and assumption **(U3)**, it follows that

$$\bar{\Delta}^{(i,2)} \leq C_6 \int_{\mathbb{R}} \left| K(\frac{x-x_1'}{c_n}) - K(\frac{x-x_1}{c_n}) \right| \cdot \mathbf{1}_{(x \notin A_n)} \mathrm{d}x,$$

where $x_1, x_1' \in B_n$. Notice that if $x \notin A_n$, then $x \notin S$. This implies the RHS of above inequality is zero. Now combining the upper bounds of $\bar{\Delta}^{(i,1)}$ and $\bar{\Delta}^{(i,2)}$, we have proved that under assumption **(U3)**, for $i = 1, \cdots, m$, $\bar{\Delta}^{(i)} \leq C_6 \left(\frac{1}{n}\right)^{1/p_n}$. Now setting $\bar{\boldsymbol{\Delta}} = [\bar{\Delta}^{(1)}, \cdots, \bar{\Delta}^{(m)}]$, we obtain $\Delta_{L_1}(\nabla L_n(\boldsymbol{\theta})) \leq ||\bar{\boldsymbol{\Delta}}||_1 \leq C \cdot m \cdot n^{-\frac{1}{p}}$ and $\Delta_{L_2}(\nabla L_n(\boldsymbol{\theta})) \leq ||\bar{\boldsymbol{\Delta}}||_2 \leq C \cdot \sqrt{m} \cdot n^{-\frac{1}{p}}$. Turning to the sharp sensitivity for Hessian, the proof follows a similar method and we obtain $\Delta_{L_1}(H_n(\boldsymbol{\theta})) \leq C \cdot m \cdot n^{-\frac{1}{p}}$ and $\Delta_{L_2}(\nabla L_n(\boldsymbol{\theta})) \leq C \cdot m \cdot n^{-\frac{1}{p}}$. This completes the proof of Theorem 24. ∎

6.8 Proof of Proposition 23

For PGD, recalling that the mechanism $M_k(w, D) = w - \eta\left(\nabla L_n(w) + \Delta_n \cdot c_{\epsilon'} Z_k\right)$. M_k satisfies $\epsilon'-$HDP by Proposition 12 and the post processing property. Now starting with the initial estimate $w = \hat{\boldsymbol{\theta}}_n^{(0)}$, we obtain $\hat{\boldsymbol{\theta}}_n^{(k)}$, for $k \geq 1$ using the iteration

$$\hat{\boldsymbol{\theta}}_n^{(k)} = M_k(\hat{\boldsymbol{\theta}}_n^{(k-1)}, D) = \hat{\boldsymbol{\theta}}_n^{(k-1)} - \eta\left(\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k-1)}) + \Delta_n \cdot c_{\epsilon'} Z_k\right).$$

Hence, by Corollary 9, $\hat{\boldsymbol{\theta}}_n^{(K)}$ satisfies $h_K(\epsilon')-$HDP. Finally, by the choice of ϵ' satisfying $h_K(\epsilon') = \epsilon$, it follows that $\hat{\boldsymbol{\theta}}_n^{(K)}$ satisfies $\epsilon-$HDP. Next, turning to PNR, the mechanism is $M_k(w, D) = w - (H_n(w) + W_{n,k})^{-1} (\nabla L_n(w) + N_{n,k})$, where $W_{n,k} \in \mathbb{R}^{m\times m}$ and $N_{n,k} \in \mathbb{R}^{m\times 1}$ are the independent random variables added to satisfy the HDP property and

$$N_{n,k} = \Delta_n \cdot c_{\epsilon'/2} \cdot Z_k, \quad W_{n,k} = \Delta_n^{(H)} \cdot c_{\epsilon'/2} \cdot \tilde{Z}_k.$$

Let $M_{k,1}(w, D) = (H_n(w) + W_{n,k})^{-1}$, $M_{k,2}(w, D) = (\nabla L_n(w) + N_{n,k})$. Then $M_{k,1}$ and $M_{k,2}$ satisfies $\frac{\epsilon'}{2}-$HDP by Proposition 12, Proposition 13, and the post processing property. Hence, by Corollary 9, it follows that $M_k(w, D) = w - M_{k,1}(w, D) \cdot M_{k,2}(w, D)$ satisfies $\epsilon'-$HDP. Finally, starting with the initial estimate $w = \hat{\boldsymbol{\theta}}_n^{(0)}$, we obtain $\hat{\boldsymbol{\theta}}_n^{(k)}$ for $k \geq 1$ by iterating

$$\hat{\boldsymbol{\theta}}_n^{(k)} = M_k(\hat{\boldsymbol{\theta}}_n^{(k-1)}, D) = \hat{\boldsymbol{\theta}}_n^{(k-1)} - M_{k,1}(\hat{\boldsymbol{\theta}}_n^{(k-1)}, D) \cdot M_{k,1}(\hat{\boldsymbol{\theta}}_n^{(k-1)}, D)$$

Hence, by Corollary 9, $\hat{\boldsymbol{\theta}}_n^{(K)}$ satisfies $h_K(\epsilon')-$HDP. Also, by the choice of ϵ' satisfying $h_K(\epsilon') = \epsilon$, it follows that $\hat{\boldsymbol{\theta}}_n^{(K)}$ satisfies $\epsilon-$HDP. Finally, to obtain the bounds for $h_K(\epsilon K^{-1})$, first notice that $h_2(\epsilon K^{-1}) = h_1(\epsilon K^{-1}) + \epsilon' - \frac{1}{2} h_1(\epsilon K^{-1})\epsilon' \leq h_1(\epsilon K^{-1}) + \epsilon' = 2\epsilon K^{-1}$. Iterating, it follows that $h_K(\epsilon K^{-1}) \leq \epsilon$. Now, turning to the lower bound, by iterating Corollary 9 we obtain

$$h_K(\epsilon K^{-1}) = \epsilon' + \epsilon' \left[(K-1) - \frac{1}{2} \sum_{j=1}^{K-1} h_j(\epsilon K^{-1}) \right] = K\epsilon' - \epsilon' \frac{1}{2} \sum_{j=1}^{K-1} h_j(\epsilon K^{-1}).$$

Next, using the upper bound, $h_j(\epsilon K^{-1}) \leq \epsilon \cdot jK^{-1}$, we obtain

$$h_K(\epsilon K^{-1}) \geq \epsilon[1 - \epsilon(K-1)(4K)^{-1}]. \quad \blacksquare$$

6.9 Proof of Theorem 25

The proof of the theorem relies on the behavior of the Hellinger loss function at *private* estimates. Intuitively, we show that under ASLSC and τ_2-smoothness, the closeness of the loss functions implies the closeness of the parameter estimates and vice-versa. This is achieved via Lemma 6.1-Lemma 6.3. We recall that N is defined in Proposition 17. In this proof, for the ease of exposition, we set τ_1 and τ_2 to be $2\tau_1$ and $2\tau_2$.

Lemma 6.1 *Assume that assumptions* **(A1)-(A8)** *in Appendix A and* **(U1)-(U2)** *hold and that* $||\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_g||_2 \leq \frac{1}{2}r$. *Also, assume that for all* $k = 1, \cdots, K$, *and* $n \geq N$, $||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}||_2 \leq \frac{1}{2}r$ *with probability* $1 - \frac{k\xi}{K}$ *and* $||\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)}||_2 \leq r$ *with probability* $1 - \frac{(k+1)\xi}{K}$, *where* r *is as defined in Proposition 17. Let* $N_{n,k} = \Delta_n c_{\epsilon'} Z_k$, *where* $Z_k \sim N(\mathbf{0}, \mathbf{I})$. *Then, with probability* $1 - \frac{k\xi}{K}$,

$$L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq (1-\gamma)^k (L_n(\hat{\boldsymbol{\theta}}_n^{(0)}) - L_n(\hat{\boldsymbol{\theta}}_n)) + \frac{3r \cdot ||N_{n,k}||_2}{2\gamma}, \tag{8}$$

where η *and* γ *are chosen such that* $0 < \gamma \leq 2\eta\tau_1 \leq 2\eta\tau_2 < 1$.

Proof: Recall that

$$Q_k(\boldsymbol{\theta}) = L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + \langle \nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + N_k, \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_n^{(k)} \rangle + \frac{1}{2\eta} ||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_n^{(k)}||_2^2.$$

Since $\hat{\boldsymbol{\theta}}_n^{(k+1)}$ minimizes $Q_k(\boldsymbol{\theta})$, it follows that by setting $\boldsymbol{\theta}_\gamma = \gamma \hat{\boldsymbol{\theta}}_n + (1-\gamma)\hat{\boldsymbol{\theta}}_n^{(k)}$, that

$$\begin{aligned} Q_k(\hat{\boldsymbol{\theta}}_n^{(k+1)}) \leq Q_k(\boldsymbol{\theta}_\gamma) =& L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + \gamma \langle \nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}), \hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)} \rangle \\ &+ \frac{\gamma^2}{2\eta} ||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}||_2^2 + \gamma \langle N_{n,k}, \hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)} \rangle. \end{aligned} \tag{9}$$

Now using Proposition 18 part (1) we obtain

$$\langle \nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}), \hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)} \rangle \leq L_n(\hat{\boldsymbol{\theta}}_n) - L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - \tau_1 ||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}||_2^2.$$

Now, using this bound in the inequality (9), we obtain

$$\begin{aligned} Q_k(\hat{\boldsymbol{\theta}}_n^{(k+1)}) \leq & L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - \gamma [L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n)] + \left(\frac{\gamma^2}{2\eta} - \gamma \tau_1 \right) ||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}||_2^2 \\ &+ \gamma \langle N_{n,k}, \hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)} \rangle, \end{aligned} \tag{10}$$

yielding the upper bound of $Q_k(\hat{\boldsymbol{\theta}}_n^{(k+1)})$. We next obtain a lower bound for $Q_k(\hat{\boldsymbol{\theta}}_n^{(k+1)})$. To this end, we use part (3) of Proposition 18. Specifically, using $L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) \geq L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - \langle \nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}), \hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)} \rangle - \tau_2 ||\hat{\boldsymbol{\theta}}_n^{(k)} - \hat{\boldsymbol{\theta}}_n^{(k+1)}||_2^2$, we obtain that

$$Q_k(\hat{\boldsymbol{\theta}}_n^{(k+1)}) \geq L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) + (\frac{1}{2\eta} - \tau_2) ||\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)}||_2^2 + \langle N_{n,k}, \hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)} \rangle.$$

Now since $2\eta\tau_2 \leq 1$, it follows that

$$Q_k(\hat{\boldsymbol{\theta}}_n^{(k+1)}) \geq L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) + \langle N_{n,k}, \hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)} \rangle. \tag{11}$$

Now using (10) and (11) it follows that

$$\begin{aligned} L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq & (1-\gamma)(L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n)) + \left(\frac{\gamma^2}{2\eta} - \gamma \tau_1 \right) ||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}||_2^2 \\ &+ \gamma \langle N_{n,k}, \hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)} \rangle - \langle N_{n,k}, \hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)} \rangle. \end{aligned}$$

Now choosing γ so that $0 < \gamma \leq 2\eta\tau_1 < 1$ and applying Cauchy-Schwarz inequality, it follows that

$$\begin{aligned} L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq & (1-\gamma)(L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n)) + ||N_{n,k}||_2 (||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}||_2 \\ &+ ||\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)}||_2) \\ \leq & (1-\gamma)(L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n)) + \frac{3r||N_{n,k}||_2}{2}, \end{aligned} \tag{12}$$

where the last inequality follows from the assumptions $||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}||_2 \leq \frac{1}{2}r$ and $||\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)}||_2 \leq r$. Now iterating the above inequality, it follows that

$$\begin{aligned} L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq & (1-\gamma)^k (L_n(\hat{\boldsymbol{\theta}}_n^{(0)}) - L_n(\hat{\boldsymbol{\theta}}_n)) + \frac{3r \cdot ||N_{n,k}||_2}{2} \cdot \frac{(1-(1-\gamma)^k)}{\gamma} \\ \leq & (1-\gamma)^k (L_n(\hat{\boldsymbol{\theta}}_n^{(0)}) - L_n(\hat{\boldsymbol{\theta}}_n)) + \frac{3r \cdot ||N_{n,k}||_2}{2\gamma}. \quad \blacksquare \end{aligned}$$

Our next key result is Lemma 6.3 below, which verifies that under the assumptions of Lemma 6.3 the private and non-private estimators are close for large n and for every iteration $k = 0, 1, \cdots K$. The proof of this lemma relies on the notion that, under the assumptions in the Appendix A and **(U1)-(U3)**, if the loss functions are "close", then arguments of the loss functions are also "close". This is the content of our next lemma and the proof is based on almost sure local strong convexity and is provided in Appendix D.

Lemma 6.2 *Let assumptions* **(A1)-(A8)** *in Appendix A and* **(U1)-(U2)** *hold. Then for* $\boldsymbol{\theta} \in B_r(\boldsymbol{\theta}_g)$ *and* $n \geq N$*, if* $L_n(\boldsymbol{\theta}) - L_n(\hat{\boldsymbol{\theta}}) \leq \frac{r^2}{4}\tau_1$ *then* $||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}||_2 \leq \frac{r}{2}$*. Furthermore, if* $||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}||_2 \leq \frac{r}{2}$ *for* $\boldsymbol{\theta} \in B_r(\boldsymbol{\theta}_g)$*, then for* $n \geq N$*,* $L_n(\boldsymbol{\theta}) - L_n(\hat{\boldsymbol{\theta}}) \leq \frac{r^2}{4}\tau_2$*.*

We next turn to the key result verifying the validity of the conditions in Lemma 6.1 above.

Lemma 6.3 *Under assumptions* **(A1)-(A8)** *and* **(U1)-(U2)***, for* $\eta \leq \frac{1}{\tau_2}$*, assume that for* $n \geq N$*,* $\hat{\boldsymbol{\theta}}_n \in B_{r/c}(\boldsymbol{\theta}_g) \subset B_{r/2}(\boldsymbol{\theta}_g)$*, where* $c > 2\left(\frac{\tau_2}{\tau_1}\right)^{\frac{1}{2}}$*, then there exists* $\hat{\boldsymbol{\theta}}_n^{(0)}$*, such that* $L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq \tau_1 \frac{r^2}{4}$ *and* $||\hat{\boldsymbol{\theta}}_n^{(k)} - \hat{\boldsymbol{\theta}}_n||_2 \leq \frac{r}{2}$ *hold with probability* $1 - \frac{k\xi}{K}$ *for all* $k = 0, \cdots, K$*.*

The proof of this lemma is similar to the proof of Lemma 18 in Avella-Medina et al. (2023). A mildly different proof is given in the Appendix D.

We now turn to the proof of Theorem 25.

Proof of Theorem 25 : Using Proposition 23 with K replaced by K_n, it follows that $\hat{\boldsymbol{\theta}}_n^{(K_n)}$ satisfies $\epsilon-$HDP. We next turn to verification of the utility. The key idea is to use Proposition 18 (1) and Lemma 6.1 and iterate until the required bound is reached. Towards this, using $\nabla L_n(\hat{\boldsymbol{\theta}}_n) = (0, \cdots, 0)$ and taking $\gamma \in (0, 2\eta\tau_1)$, it follows that

$$||\hat{\boldsymbol{\theta}}_n^{(k)} - \hat{\boldsymbol{\theta}}_n||_2^2 \leq \frac{(1-\gamma)^k (L_n(\hat{\boldsymbol{\theta}}_n^{(0)}) - L_n(\hat{\boldsymbol{\theta}}_n))}{\tau_1} + \frac{3r \cdot ||N_{n,k}||_2}{2\gamma\tau_1}.$$

Using concentration inequality for L_2-norm of the Gaussian vector (see Rigollet and Hütter (2023)), namely,

$$P\left(||Z_k||_2 \geq 4\sqrt{m} + 2\sqrt{2[\log K - \log \xi]}\right) \leq \frac{\xi}{K},$$

it follows by setting $\epsilon' = \frac{\epsilon}{K}$ that

$$P(||N_{n,k}||_2 \leq \Delta_n c_{\epsilon'}\left(4\sqrt{m} + 2\sqrt{2[\log K - \log \xi]}\right) := r_{noi}) > 1 - \frac{\xi}{K}. \tag{13}$$

We emphasize here that r_{noi} depends on n, K_n and ξ. Now, first consider the case $L_n(\hat{\boldsymbol{\theta}}_n^{(0)}) \neq L_n(\hat{\boldsymbol{\theta}}_n)$. By choosing $k > k_0 := \max\{0, \frac{-\log(L_n(\hat{\boldsymbol{\theta}}_n^{(0)})-L_n(\hat{\boldsymbol{\theta}}_n))+\log(3r\cdot r_{noi})-\log(2\gamma)}{\log(1-\gamma)}\}$, it follows that with probability $(1-\frac{k\xi}{K})$, and for all $k \geq k_0$

$$||\hat{\boldsymbol{\theta}}_n^{(k)} - \hat{\boldsymbol{\theta}}_n||_2^2 \leq \frac{3r \cdot r_{noi}}{\gamma\tau_1} := C_0^2 r_{noi}. \tag{14}$$

We notice here that this bound is of order $\Delta_n^{\frac{1}{2}}(K_n \log K_n)^{\frac{1}{4}}$. However, this will not yield efficiency. Our goal is to remove the square root from Δ_n. This suggests one needs larger k in the above bound. This is accomplished by additional iterations (see Theorem 2 in Avella-Medina et al. (2023)). To this end, we need the following claim, whose proof is given below.

Claim: For $k > k_0$, choose n, K such that $r_{noi}^{\frac{1}{2}} < \frac{1}{2\eta}C_0$. Then

$$L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq (1-\gamma)(L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n)) + (2\eta\tau_2 + \frac{3}{2})C_0 r_{noi}^{\frac{3}{2}}.$$

Using the claim with $k = k_0 + j - 1$ and iterating we get

$$\begin{aligned} L_n(\hat{\boldsymbol{\theta}}_n^{(k_0+j)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq& (1-\gamma)^j (L_n(\hat{\boldsymbol{\theta}}_n^{(k_0)}) - L_n(\hat{\boldsymbol{\theta}}_n)) + \left[\sum_{i=0}^{j}(1-\gamma)^i\right](2\eta\tau_2 + \frac{3}{2})C_0 r_{noi}^{\frac{3}{2}} \\ \leq& (1-\gamma)^j (L_n(\hat{\boldsymbol{\theta}}_n^{(k_0)}) - L_n(\hat{\boldsymbol{\theta}}_n)) + \frac{1}{\gamma}(2\eta\tau_2 + \frac{3}{2})C_0 r_{noi}^{\frac{3}{2}}. \end{aligned}$$

Now, using Proposition 18 (1) and utilizing $\nabla L_n(\hat{\boldsymbol{\theta}}_n) = (0, \cdots, 0)$, it follows that $||\hat{\boldsymbol{\theta}}_n^{(k_0+j)} - \hat{\boldsymbol{\theta}}_n||_2^2 \leq \frac{L_n(\hat{\boldsymbol{\theta}}_n^{(k_0+j)})-L_n(\hat{\boldsymbol{\theta}}_n)}{\tau_1}$ and hence

$$||\hat{\boldsymbol{\theta}}_n^{(k_0+j)} - \hat{\boldsymbol{\theta}}_n||_2^2 \leq \frac{(1-\gamma)^j(L_n(\hat{\boldsymbol{\theta}}_n^{(k_0)}) - L_n(\hat{\boldsymbol{\theta}}_n))}{\tau_1} + (2\eta\tau_2 + \frac{3}{2})\frac{C_0}{\gamma\tau_1} r_{noi}^{\frac{3}{2}}.$$

Next, we choose

$$j \geq k_1 := \max\left\{0, \frac{-\log(L_n(\hat{\boldsymbol{\theta}}_n^{(k_0)}) - L_n(\hat{\boldsymbol{\theta}}_n)) + \log\left(\frac{1}{2}\frac{C_0}{\gamma\tau_1} r_{noi}^{\frac{3}{2}}\right)}{\log(1-\gamma)}\right\},$$

and setting $C_1 = (\frac{2\eta\tau_2+2}{\gamma\tau_1}C_0)^{\frac{1}{2}}$, we obtain for $k > k_0 + k_1$,

$$||\hat{\boldsymbol{\theta}}_n^{(k)} - \hat{\boldsymbol{\theta}}_n||_2^2 \leq C_1^2 r_{noi}^{\frac{3}{2}}.$$

We notice that the power of r_{noi} is now 3/2 and is still below the required power of 2. Hence, continuing the iterations and using the Claim with starting value $k_0 + k_1 + \cdots k_i$, we obtain for $k \geq k_0 + k_1 + \cdots k_i$,

$$||\hat{\boldsymbol{\theta}}_n^{(k)} - \hat{\boldsymbol{\theta}}||_2^2 \leq C_i^2 r_{noi}^{2-\frac{1}{2^i}} \quad \text{if} \quad r_{noi}^{\frac{1}{2^i}} \leq \frac{C_{i-1}}{2\eta},$$

where $C_i = \left(\frac{2\eta\tau_2+2}{\gamma\tau_1}C_{i-1}\right)^{\frac{1}{2}} = \left(\frac{2\eta\tau_2+2}{\gamma\tau_1}\right)^{1-\frac{1}{2^i}} \cdot C_0^{\frac{1}{2^i}}$. Finally, taking $i = \log_2(n)$ we get $k > k_0 + \cdots + k_{\log_2(n)}$

$$||\hat{\boldsymbol{\theta}}_n^{(k)} - \hat{\boldsymbol{\theta}}_n||_2^2 \leq C^2_{\log_2(n)} r_{noi}^{2-\frac{1}{n}}, \quad \text{if} \quad r_{noi}^{\frac{1}{n}} \leq \frac{C_{\log_2(n)-1}}{2\eta},$$

where $C_{\log_2(n)} = \left(\frac{2\eta\tau_2+2}{\gamma\tau_1}\right)^{1-\frac{1}{n}} \cdot C_0^{\frac{1}{n}}$. Now, letting $n \to \infty$, notice that $C_{\log_2(n)}$ converges to $C_\infty(\gamma) := 2(\eta\tau_2 + 1)(\gamma\tau_1)^{-1}$. Also, notice that $r_{noi}^{\frac{1}{n}}$ converges to 1. Now choosing $\gamma \in (0, 2\eta\tau_1)$ and $C_\infty > 1$ (such a γ exists) it follows that

$$\limsup_{n\to\infty} r_{noi}^{-2}||\hat{\boldsymbol{\theta}}_n^{(k_n)} - \hat{\boldsymbol{\theta}}_n||_2 = C_\infty. \tag{15}$$

This requires $K \geq k_0 + k_1 + \cdots + k_{\log_2(n)} \sim (\log n) \cdot (\log r_{noi})$ which implies $K \geq c \log n$, since r_{noi} is bounded by a constant by choice of n and K. Next, we notice that $r_{noi} = \Delta_n c_{\epsilon'} \left(4\sqrt{m} + 2\sqrt{2[\log K - \log \xi]}\right)$ by Theorem 24 and $\Delta_n \sim c \cdot n^{-\frac{1}{p}}$ for $p \in (1, 2]$. Hence, (15) becomes

$$||\hat{\boldsymbol{\theta}}_n^{(K_n)} - \hat{\boldsymbol{\theta}}_n||_2 \leq c \cdot n^{-\frac{1}{p}} (K_n \log(K_n/\xi))^{\frac{1}{2}},$$

for large n with high probability. Thus, to complete the proof of the Theorem, we now establish the claim.
Proof of the Claim: Notice that by Proposition 18 inequality 4 that

$$||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2 = ||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - \nabla L_n(\hat{\boldsymbol{\theta}}_n)||_2 \leq 2\tau_2||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}||_2. \tag{16}$$

Now, first using equation (8) in Definition 21 and the expression above and applying (14) it follows that

$$||\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)}||_2 \leq 2\eta\tau_2 C_0 r_{noi}^{\frac{1}{2}} + \eta r_{noi}.$$

From the inequality (12) in the proof of Lemma 6.1, using (14) and from (16) it follows that

$$L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq (1-\gamma)(L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n)) + (2\eta\tau_2 + 1)C_0 r_{noi}^{\frac{3}{2}} + \eta r_{noi}^2.$$

Next, choosing n, K such that $r_{noi}^{\frac{1}{2}} < \frac{1}{2\eta}C_0$ it follows that

$$L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq (1-\gamma)(L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n)) + (2\eta\tau_2 + \frac{3}{2})C_0 r_{noi}^{\frac{3}{2}}.$$

This completes the proof of the claim and the Theorem. ∎

6.10 Proof of Theorem 27

The proof of the Theorem relies on the Lemma 6.4-Lemma 6.6 whose proofs use matrix concentration inequality and is similar to the idea of proof of Theorem 25. We recall that the concentration inequality for L_2-norm of the Gaussian vector and matrix, (see Rigollet and Hütter (2023); Tropp et al. (2015)), is given by

$$P\left(||N_{n,k}||_2 \leq \Delta_n \cdot c_{\epsilon'/2} \cdot [4\sqrt{m} + 2\sqrt{2(\log 2K - \log \xi)}]\right) \geq 1 - \frac{\xi}{2K}, \quad \text{and}$$
$$P\left(||W_{n,k}||_2 \leq \Delta_n^{(H)} \cdot c_{\epsilon'/2} \cdot \sqrt{2m \log(4Km/\xi)}\right) \geq 1 - \frac{\xi}{2K}.$$

We use these upper bounds on the norms with probability $1 - \frac{\xi}{K}$ in the following lemmas and proofs. In this proof, for the ease of exposition, we set τ_1 and τ_2 to be $2\tau_1$ and $2\tau_2$, and choose $\eta = 1$. Our first lemma provides a useful alternative expression for $\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)}$.

Lemma 6.4

$$\left(H_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + W_{n,k}\right)^{-1} \left(\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + N_{n,k}\right) = H_n^{-1}(\hat{\boldsymbol{\theta}}_n^{(k)}) \nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + \tilde{N}_{n,k}.$$

Under assumptions **(A1)**-**(A8)** *of Appendix A and* **(U1)**-**(U2)**, *if* $\hat{\boldsymbol{\theta}}_n^{(k)} \in B_r(\boldsymbol{\theta}_g)$, *then for large* n, $||\tilde{N}_{n,k}||_2 \leq \frac{\|N_{n,k}\|_2}{2\tau_1} + \frac{B_1 \cdot ||W_{n,k}||_2}{2\tau_1^2} + \frac{||N_{n,k}||_2 \cdot ||W_{n,k}||_2}{2\tau_1^2}$ *holds with probability* $1 - \frac{\xi}{2K}$. *Additionally,* $\kappa \sim n^{-\frac{1}{p}}(K \log(K/\xi))^{\frac{1}{2}}$, *there exists* N_κ *such that for all* $n > N_\kappa$ *and* $k = 1, 2, \cdots K$ $P(||\tilde{N}_{n,k}||_2 \leq \kappa) > 1 - \frac{\xi}{K}$.

Proof: Using Neumann series formula, note that

$$\begin{aligned}
&\left(H_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + W_{n,k}\right)^{-1} \left(\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + N_{n,k}\right) \\
=& H_n^{-1}(\hat{\boldsymbol{\theta}}_n^{(k)}) \left[\mathbf{I} + \sum_{j=1}^{\infty} (-W_{n,k} H_n^{-1}(\hat{\boldsymbol{\theta}}_n^{(k)}))^j\right] \left(\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + N_{n,k}\right) \\
:=& H_n^{-1}(\hat{\boldsymbol{\theta}}_n^{(k)}) \nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + \tilde{N}_{n,k},
\end{aligned}$$

where

$$\tilde{N}_{n,k} = H_n^{-1}(\hat{\boldsymbol{\theta}}_n^{(k)}) \left\{ N_{n,k} + \left[\sum_{j=1}^{\infty} (-W_{n,k} H_n^{-1}(\hat{\boldsymbol{\theta}}_n^{(k)}))^j\right] \left(\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + N_{n,k}\right) \right\}.$$

Now, applying the properties of matrix norms, Proposition 17, and Proposition 14, we obtain

$$||\tilde{N}_{n,k}||_2 \leq \frac{1}{2\tau_1} \cdot \left[\|N_{n,k}\|_2 + \left[\sum_{j=1}^{\infty} \left(\frac{||W_{n,k}||_2}{2\tau_1}\right)^j\right] (B_1 + ||N_{n,k}||_2)\right].$$

Let n be large enough such that $||W_{n,k}||_2 \leq \tau_1$ with probability $1 - \frac{\xi}{2K}$. Then it follows that

$$||\tilde{N}_{n,k}||_2 \leq \frac{\|N_{n,k}\|_2}{2\tau_1} + \frac{B_1 \cdot ||W_{n,k}||_2}{2\tau_1^2} + \frac{||N_{n,k}||_2 \cdot ||W_{n,k}||_2}{2\tau_1^2}.$$

Notice that as $n \to \infty$, both $||N_{n,k}||_2$ and $||W_{n,k}||_2$ converge to 0 in probability at rate $n^{-\frac{1}{p}}(K \log(K/\xi))^{\frac{1}{2}}$. ∎

Lemma 6.5 *Under assumptions* **(A1)-(A8)** *and* **(U1)-(U2)**, *if* $\hat{\boldsymbol{\theta}}_n^{(k)} \in B_r(\boldsymbol{\theta}_g)$, *then* $||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)})||_2 \leq \frac{\alpha}{2\tau_1^2}||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2^2 + C||\tilde{N}_{n,k}||_2$ *holds with probability* $1-\frac{\xi}{K}$.

Proof: Recall that from the PNR iteration, namely, $\hat{\boldsymbol{\theta}}_n^{(k+1)} = \hat{\boldsymbol{\theta}}_n^{(k)} - H_n^{-1}(\hat{\boldsymbol{\theta}}_n^{(k)})\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + \tilde{N}_{n,k}$, that $\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + H_n(\hat{\boldsymbol{\theta}}_n^{(k)}) \cdot [\hat{\boldsymbol{\theta}}_n^{(k)} - \hat{\boldsymbol{\theta}}_n^{(k+1)} - \tilde{N}_{n,k}] = 0$. We now rewrite $||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)})||_2$ as

$$||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)})||_2 = ||T_1 - T_2 + H_n(\hat{\boldsymbol{\theta}}_n^{(k)})\tilde{N}_{n,K}||_2, \quad \text{where}$$
$$T_1 = \nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - \nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) \ \text{ and } \ T_2 = H_n(\hat{\boldsymbol{\theta}}_n^{(k)})(\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)}).$$

Notice that $T_1 - T_2$ can be written as

$$\begin{aligned} T_1 - T_2 =& \int_0^1 H_n(\hat{\boldsymbol{\theta}}_n^{(k)} + t(\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)})) \cdot (\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)})\mathrm{d}t - \int_0^1 H_n(\hat{\boldsymbol{\theta}}_n^{(k)})(\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)})\mathrm{d}t \\ =&(\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)}) \int_0^1 \left[H_n(\hat{\boldsymbol{\theta}}_n^{(k)} + t(\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)})) - H_n(\hat{\boldsymbol{\theta}}_n^{(k)})\right]\mathrm{d}t. \end{aligned}$$

Using Proposition 19 (namely the Lipschitz property of the Hessian), it follows that with probability 1,

$$||T_1 - T_2||_2 \leq ||\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)}||_2 \cdot \int_0^1 \alpha \cdot t \cdot ||\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)}||_2 \mathrm{d}t = \frac{\alpha}{2}||\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)}||_2^2.$$

Using the upper bound of Proposition 14, $\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)} = -H_n^{-1}(\hat{\boldsymbol{\theta}}_n^{(k)})\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + \tilde{N}_{n,k}$, Proposition 17, and for large n that $||\tilde{N}_{n,k}||_2 \leq 1$ with probability $1-\frac{\xi}{K}$, we obtain

$$||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)})||_2 \leq \frac{\alpha}{2}||\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n^{(k)}||_2^2 + B_2||\tilde{N}_{n,k}||_2 \leq \frac{\alpha}{2\tau_1^2}||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2^2 + C||\tilde{N}_{n,k}||_2,$$

where the constant $C \in (0,\infty)$ only depends on α, τ_1, B_1, B_2. ■

The next lemma concerns the "distance" between the private and non-private estimators at every iteration, and the proof is based on induction. The choice of $\hat{\boldsymbol{\theta}}_n^{(0)}$, verifies the assumption that the assumptions in Lemma 6.5 hold; that's is, for all $k = 1, 2\cdots K$, $\hat{\boldsymbol{\theta}}_n^{(k)} \in B_r(\boldsymbol{\theta}_g)$.

Lemma 6.6 *Under assumption* **(A1)-(A8)** *and* **(U1)-(U2)**, *if* $\hat{\boldsymbol{\theta}}_n \in B_{r/2}(\boldsymbol{\theta}_g)$, *and* $||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(0)})||_2 \leq \min\{\frac{\tau_1 r}{2}, \frac{\tau_1^2}{\alpha}\}$, *then for* $k = 0, 1, \cdots, K$, $||\hat{\boldsymbol{\theta}}_n^{(k)} - \hat{\boldsymbol{\theta}}_n||_2 \leq \frac{r}{2}$ *holds with probability* $1-\frac{k\xi}{K}$.

Proof: We prove the lemma using the following claim:
Claim: If $\hat{\boldsymbol{\theta}}_n \in B_{r/2}(\boldsymbol{\theta}_g)$ and $||\nabla L_n(\boldsymbol{\theta})||_2 \leq 2\tau_1 r$, then $||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_n||_2 \leq r$.

First, we finish the proof of the lemma using the Claim and then prove the Claim. We prove the lemma by induction. First notice that by assumption $||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(0)})||_2 \leq \min\{\frac{\tau_1 r}{2}, \frac{\tau_1^2}{\alpha}\} \leq \tau_1 r$ and hence from the claim it follows, with $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_n^{(0)}$ and r replaced by $\frac{r}{2}$, that $||\hat{\boldsymbol{\theta}}_n^{(0)} - \hat{\boldsymbol{\theta}}_n||_2 \leq \frac{r}{2}$. We start the inductive hypothesis with $k = k_0$. That is, assume for $k = k_0$, $||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k_0)})||_2 \leq \min\{\tau_1 r, \frac{\tau_1^2}{\alpha}\}$ and $||\hat{\boldsymbol{\theta}}_n^{(k_0)} - \hat{\boldsymbol{\theta}}_n||_2 \leq \frac{r}{2}$, and $\hat{\boldsymbol{\theta}}_n^{(k_0)} \in B_r(\boldsymbol{\theta}_g)$.

Also from Lemma 6.5, and for large n such that $||\tilde{N}_{n,k}||_2 \leq \min\{\tau_1 r, \frac{\tau_1^2}{\alpha}\}$ with probability $1 - \frac{\xi}{K}$, we obtain

$$\begin{aligned}||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k_0+1)})||_2 \leq& \frac{\alpha}{2\tau_1^2}||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k_0)})||_2^2 + C||\tilde{N}_{n,k}||_2 \\ \leq& \frac{\alpha}{2\tau_1^2} \cdot \left(\frac{\tau_1^2}{\alpha}\right)^2 + C||\tilde{N}_{n,k}||_2 \leq \min\{\frac{\tau_1^2}{\alpha}, \tau_1 r\}.\end{aligned}$$

Now, applying the claim with $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_n^{(k_0+1)}$ and replacing r by $\frac{r}{2}$, it follows that $||\hat{\boldsymbol{\theta}}_n^{(k_0+1)} - \hat{\boldsymbol{\theta}}_n||_2 \leq \frac{r}{2}$. This completes the induction. Now, we turn to the proof of the claim.

Proof of the claim: The proof uses the ASLSC property and is similar to the one used in Avella-Medina et al. (2023). Specifically, we establish the proof using contradiction. To this end, suppose $||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_n||_2 > r$; let $\tilde{\boldsymbol{\theta}}$ denote the point on the boundary of $\mathcal{B}_r(\hat{\boldsymbol{\theta}})$. By Proposition 18,

$$\nabla L_n(\tilde{\boldsymbol{\theta}})^T \cdot (\tilde{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}) \geq 2\tau_1||\tilde{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}||_2^2.$$

Define $\boldsymbol{v} = \frac{\tilde{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}}{||\tilde{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}||_2}$; then we have

$$\nabla L_n(\tilde{\boldsymbol{\theta}})^T \cdot \boldsymbol{v} \geq 2\tau_1||\tilde{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}||_2 = 2\tau_1 r.$$

Set $f(t) = \nabla L_n(\hat{\boldsymbol{\theta}} + t \cdot v)^T \cdot v$ for $t \geq 0$, then $f'(t) = \boldsymbol{v}^T H_n(\hat{\boldsymbol{\theta}} + t \cdot \boldsymbol{v}) \cdot \boldsymbol{v} \geq 0$, since Hessian matrix is positive definite by Proposition 17. Hence $f(t)$ is increasing in t and this implies that

$$||\nabla L_n(\boldsymbol{\theta})||_2 \geq \nabla L_n(\boldsymbol{\theta})^T \cdot \boldsymbol{v} \geq \nabla L_n(\tilde{\boldsymbol{\theta}})^T \cdot \boldsymbol{v} \geq 2\tau_1 r$$

which is a contradiction since $||\nabla L_n(\boldsymbol{\theta})||_2 \leq 2\tau_1 r$. Therefore, it follows that

$$||\nabla L_n(\boldsymbol{\theta})||_2 \leq 2\tau_1 r \Longrightarrow ||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}||_2 \leq r.$$

This completes the proof of the claim and the lemma. ∎

We now turn to the proof of the Theorem.

Proof of Theorem 27 : Using Proposition 23 with K replaced by K_n, it follows that $\hat{\boldsymbol{\theta}}_n^{(K_n)}$ satisfies $\epsilon-$HDP. We next turn to verification of the utility. We assume N is large enough to satisfy the conditions in Lemma 6.4. That is for $n > N$ such that $P(||\tilde{N}_{n,k}||_2 \leq r_{noi}) \geq 1 - \frac{\xi}{K}$ for $r_{noi} \sim n^{-\frac{1}{p}}(K\log(K/\xi))^{\frac{1}{2}}$. We will use Lemma 6.5 and Lemma 6.6 to obtain the following claim:

Claim: For $\hat{\boldsymbol{\theta}}_n \in B_{r/2}(\boldsymbol{\theta}_g)$ and $||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(0)})||_2 \leq \min\{\frac{\tau_1 r}{2}, \frac{\tau_1^2}{\alpha}\}$, the inequality

$$\frac{\alpha}{2\tau_1^2}||\nabla L_n(\hat{\theta}_n^{(K)})||_2 \leq \left(\frac{\alpha}{2\tau_1^2}||\nabla L_n(\hat{\theta}_n^{(0)})||_2\right)^{2^K} + 3C \cdot r_{noi}$$

holds for some constant $C \in (0, \infty)$ with probability $1 - \xi$.

Using Proposition 18 (2), and multiplying both side by $\frac{\alpha}{\tau_1}$, we obtain

$$\frac{\alpha}{\tau_1}||\hat{\boldsymbol{\theta}}_n^{(K)} - \hat{\boldsymbol{\theta}}_n||_2 \leq \frac{\alpha}{2\tau_1^2}||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(K)}) - \nabla L_n(\hat{\boldsymbol{\theta}}_n)||_2.$$

Now using the fact that $\nabla L_n(\hat{\boldsymbol{\theta}}_n) = 0$, the claim, we obtain (since $||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(0)})||_2 \leq \min\{\frac{\tau_1 r}{2}, \frac{\tau_1^2}{\alpha}\}$)

$$\frac{\alpha}{\tau_1}||\hat{\boldsymbol{\theta}}_n^{(K)} - \hat{\boldsymbol{\theta}}_n||_2 \leq \left(\frac{1}{2}\right)^{2^K} + 3Cr_{noi}.$$

Choose K large such that $\left(\frac{1}{2}\right)^{2^K} \leq Cr_{noi}$, that is $K \geq \frac{1}{\log 2}\log\frac{\log Cr_{noi}}{\log(1/2)}$, then

$$||\hat{\boldsymbol{\theta}}_n^{(K)} - \hat{\boldsymbol{\theta}}_n||_2 \leq 4Cr_{noi}.$$

By Lemma 6.4, $r_{noi} \sim n^{-\frac{1}{p}}(K\log(K/\xi))^{\frac{1}{2}}$. Using the sharp bound of $\Delta_n^{(H)}$ in Theorem 24, we obtain the utility bound. This also implies that $K \geq C' \log\log n$ for some $C' \in (0,\infty)$. We complete the proof by establishing the claim.

Proof the the claim: We prove the claim by induction. Notice that for $k=1$, the claim is true by Lemma 6.5 and Lemma 6.6. Assume that the claim holds for $k = k_0$. Then for $k = k_0+1$, using Lemma 6.5 and the choice of $\hat{\boldsymbol{\theta}}_n^{(0)}$ such that $||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(0)})||_2 \leq \min\{\frac{\tau_1 r}{2}, \frac{\tau_1^2}{\alpha}\}$, it follows that

$$\frac{\alpha}{2\tau_1^2}||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k_0+1)})||_2 \leq \left(\frac{\alpha}{2\tau_1^2}||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k_0)})||_2\right)^2 + Cr_{noi}.$$

Now by inductive hypothesis, it follows that

$$\begin{aligned}\frac{\alpha}{2\tau_1^2}||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k_0+1)})||_2 &\leq \left[\left(\frac{\alpha}{2\tau_1^2}||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(0)})||_2\right)^{2^{k_0}} + 3Cr_{noi}\right]^2 + Cr_{noi} \\ &\leq \left(\frac{\alpha}{2\tau_1^2}||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(0)})||_2\right)^{2^{k_0+1}} + \frac{3}{2}Cr_{noi} + 9C^2r_{noi}^2 + Cr_{noi}.\end{aligned}$$

Let n be large such that $9C^2r_{noi}^2 \leq \frac{1}{2}Cr_{noi}$. It then follows that

$$\frac{\alpha}{2\tau_1^2}||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k_0+1)})||_2 \leq \left(\frac{\alpha}{2\tau_1^2}||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(0)})||_2\right)^{2^{k_0+1}} + 3Cr_{noi}.$$

This completes the induction and the proof of the Claim and the Theorem. ∎

We next turn to the proof of Theorem 29. First, we recall that Q is the distribution associated with the mechanism, representing the noise distribution.

6.11 Proof of Theorem 29

We begin with part (1). Suppose $\boldsymbol{\theta}_n^{(K_n)}$ is obtained using the PGD or PNR algorithm. Then, using Theorem 25 or Theorem 27 with $p \in (1,2)$ it follows that $n^{\frac{1}{2}}||\hat{\boldsymbol{\theta}}_n^{(K_n)} - \hat{\boldsymbol{\theta}}_n||_2$ converges to zero in probability (with respect to the joint distribution of $P_g \times Q$) since $K_n \sim \log n$ for PGD algorithm and $K_n \sim \log\log n$ for the PNR algorithm. Turning to part (2), observe that

$$(\hat{\boldsymbol{\theta}}_n^{(K_n)} - \boldsymbol{\theta}_g) = (\hat{\boldsymbol{\theta}}_n^{(K_n)} - \hat{\boldsymbol{\theta}}_n) + (\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_g). \tag{17}$$

Now, taking the norm, the first term on the RHS of the above equation converges to 0 in probability by part (1), and the second term converges to zero almost surely under the assumptions **(A1)-(A8)** in appendix A. Finally, turning to part (3), by multiplying both sides of (17) by $\sqrt{n}$, the first term converges to zero in $P_g \times Q$ probability by part (1). The second term converges to a normal distribution under the assumptions **(A1)-(A8)** in Appendix A under P_g, by Theorem A.1 in Appendix A. Hence, $n^{\frac{1}{2}}(\hat{\boldsymbol{\theta}}_n^{(K_n)} - \boldsymbol{\theta}_g)$ converges in distribution (under $P_g \times Q$) to a multivariate normal distribution; that is,

$$\lim_{n\to\infty} P_g \times Q \left(n^{\frac{1}{2}}(\hat{\boldsymbol{\theta}}_n^{(K_n)} - \boldsymbol{\theta}_g) \leq \boldsymbol{x} \right) = P(\mathbf{Z} \leq \boldsymbol{x}),$$

where $\mathbf{Z} \sim N(0, \Sigma_g)$.

Appendix A.

Let $f(x)$ and $g(x)$ be any two probability density functions. The Hellinger distance between $f(x)$ and $g(x)$ is defined as the L_2-norm of the difference between the square root of density functions, that is,

$$HD^2(f,g) = ||f^{\frac{1}{2}}(x) - g^{\frac{1}{2}}(x)||_2^2 = \int \left[f^{\frac{1}{2}}(x) - g^{\frac{1}{2}}(x)\right]^2 \mathrm{d}x.$$

Let $\{X_1, X_2, \cdots, X_n\}$ be i.i.d. real-valued random variables with density $g(\cdot)$, and postulated to belong to a parametric family $\{f_{\boldsymbol{\theta}} : \boldsymbol{\theta} \in \Theta \subset \mathbb{R}^m\}$. The minimum Hellinger distance estimator in the population, $\boldsymbol{\theta}_g$, if it exists, is the minimizer of the $||f_{\boldsymbol{\theta}}^{\frac{1}{2}} - g^{\frac{1}{2}}||_2$; that is,

$$\boldsymbol{\theta}_g = \underset{\boldsymbol{\theta}\in\Theta}{\operatorname{argmin}} ||f_{\boldsymbol{\theta}}^{\frac{1}{2}} - g^{\frac{1}{2}}||_2 = \underset{\boldsymbol{\theta}\in\Theta}{\operatorname{argmin}} HD(f_{\boldsymbol{\theta}}, g).$$

When $g(\cdot) = f_{\boldsymbol{\theta}_0}(\cdot)$, $\boldsymbol{\theta}_g = \boldsymbol{\theta}_0$. We also assume that $\boldsymbol{\theta}_g$ and $\boldsymbol{\theta}_0$ belong to the interior of Θ. Beran (1977) and Cheng and Vidyashankar (2006) establish that under the assumption,

(A1) *$\Theta \subset \mathbb{R}^m$ is compact and convex and the family $\{f_{\boldsymbol{\theta}}(\cdot) : \boldsymbol{\theta} \in \Theta\}$ is identifiable; that is, if $\boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_2$ then $f_{\boldsymbol{\theta}_1}(\cdot) \neq f_{\boldsymbol{\theta}_2}(\cdot)$ on a set of positive Lebesgue measure.*

that $\boldsymbol{\theta}_g$ exists and is unique. We will assume this condition holds. In practice, one replaces $g(\cdot)$ by $g_n(\cdot)$, where $g_n(\cdot)$ is a nonparametric estimate of of $g(\cdot)$; specifically, a kernel density estimator, defined below.

$$g_n(x) = \frac{1}{n \cdot c_n} \sum_{i=1}^{n} K\left(\frac{x - X_i}{c_n}\right).$$

The MHDE is obtained by minimizing the loss function

$$\hat{\boldsymbol{\theta}}_n = \arg\min_{\boldsymbol{\theta}\in\Theta} L_n(\boldsymbol{\theta}), \quad \text{where } L_n(\boldsymbol{\theta}) = \int_{\mathbb{R}} (\sqrt{f_{\boldsymbol{\theta}}(x)} - \sqrt{g_n(x)})^2 \mathrm{d}x.$$

Asymptotic properties of $\hat{\boldsymbol{\theta}}_n$ rely on the bandwidth c_n and additional regularity assumptions on the parametric family. We provide the assumptions below:

(A2) *The kernel function $K(\cdot)$ is symmetric (about 0) density with compact support. The bandwidth c_n satisfies $c_n \to 0$, $n^{\frac{1}{2}} c_n^2 \to 0$, $n^{\frac{1}{2}} c_n \to \infty$.*

(A3) *$f_{\boldsymbol{\theta}}(x)$ is twice continuously differentiable in $\boldsymbol{\theta}$. Also, the Fisher information matrix $I(\boldsymbol{\theta})$ is positive definite and continuous in $\boldsymbol{\theta}$ with finite maximum eigenvalue.*

(A4) $||\mathbf{u}_{\boldsymbol{\theta}}(\cdot) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(\cdot)||_2$, $||\dot{\mathbf{u}}_{\boldsymbol{\theta}}(\cdot) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(\cdot)||_2$, $||\mathbf{u}_{\boldsymbol{\theta}}(\cdot)\mathbf{u}_{\boldsymbol{\theta}}^T(\cdot) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(\cdot)||_2$ *exist and are continuous in $\boldsymbol{\theta}$.*

(A5) *Let $\{a_n, n \geq 1\}$ be a sequence diverging to infinity. Assume $\lim_{n\to\infty} n \sup_{t\in supp(K)} \mathbf{P}(|X - c_n t| > a_n) = 0$, where $supp(K)$ is the support of the kernel density $K(\cdot)$ and X is a generic random variable with density $f_{\boldsymbol{\theta}_g}(\cdot)$.*

(A6) *Let* $M(n) = \sup_{|x| \leq a_n} \sup_{t \in supp(K)} |f_{\boldsymbol{\theta}_g}^{-1}(x) f_{\boldsymbol{\theta}_g}(x + tc_n)|$. *Assume* $\sup_{n \geq 1} M(n) < \infty$.

(A7) *The score function has a regular central behavior,*

$$\lim_{n\to\infty} (n^{\frac{1}{2}} c_n)^{-1} \int_{-a_n}^{a_n} \mathbf{u}_{\boldsymbol{\theta}_g}(x) \mathrm{d}x = \mathbf{0}; \text{ also, assume that } \lim_{n\to\infty} (n^{\frac{1}{2}} c_n^4) \int_{-a_n}^{a_n} \mathbf{u}_{\boldsymbol{\theta}_g}(x) \mathrm{d}x = \mathbf{0}.$$

(A8) *The score function is smooth in an* L_2 *sense; i.e.*

$$\lim_{n\to\infty} \sup_{t \in supp(K)} \int_{\mathbb{R}} [\mathbf{u}_{\boldsymbol{\theta}_g}(x + tc_n) - \mathbf{u}_{\boldsymbol{\theta}_g}(x)]^2 f_{\boldsymbol{\theta}_g}(x) \mathrm{d}x = \mathbf{0}.$$

It is known that, under the above conditions, $\hat{\boldsymbol{\theta}}_n$ is known to be unique, consistent, and asymptotically efficient (see Beran (1977), Cheng and Vidyashankar (2006)). Write $g(x) = \lim_{n\to\infty} g_n(x)$ (which exists by **(A2)**) and set

$$\rho_{\boldsymbol{\theta}}(x) = -4 \left[\int g^{\frac{1}{2}}(x) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) [\mathbf{u}_{\boldsymbol{\theta}}(x) \mathbf{u}_{\boldsymbol{\theta}}^T(x) + 2\dot{\mathbf{u}}_{\boldsymbol{\theta}}(x)] \mathrm{d}x \right]^{-1} \cdot \nabla f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) \quad \text{and}$$

$$\Sigma_g = 4^{-1} \int \rho_{\boldsymbol{\theta}_g}(x) \rho_{\boldsymbol{\theta}_g}^T(x) \mathrm{d}x.$$

The next theorem is concerned with the limit distribution of MHDE and is similar to the proof in Cheng and Vidyashankar (2006) when the true model is $g(\cdot)$.

Theorem A.1 *Under the assumptions* **(A1)**-**(A8)**, $\rho_{\boldsymbol{\theta}}(\cdot)$ *is continuous at* $\boldsymbol{\theta}_g$. *Furthermore,*

1. $||\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_g||_2 \xrightarrow{P} 0$,
2. $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_g) \xrightarrow{d} N(0, \Sigma_g)$.
3. *In particular, if* $g(\cdot) = f_{\boldsymbol{\theta}_0}(\cdot)$, *then* $\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) \xrightarrow{d} N(0, I^{-1}(\boldsymbol{\theta}_0))$.

Appendix B.

B.1 Gaussian mechanism

Lemma B.1 *For two m dimensional random variable $\mathbf{X} \sim N(\mathbf{w}_1, \sigma^2 \cdot \mathbf{I})$ and $\mathbf{Y} \sim N(\mathbf{w}_2, \sigma^2 \cdot \mathbf{I})$, the power divergence with parameter λ is given by*

$$D_\lambda(\mathbf{X},\mathbf{Y}) = \begin{cases} \frac{1}{\lambda(\lambda+1)}\left[e^{\frac{\lambda(\lambda+1)||\mathbf{v}||_2^2}{2\sigma^2}} - 1\right], & \lambda(\lambda+1) \neq 0 \\ \frac{||\mathbf{v}||_2^2}{2\sigma^2}, & \lambda(\lambda+1) = 0, \end{cases}$$

where $\mathbf{v} = \mathbf{w}_1 - \mathbf{w}_2$. In particular, if $\lambda = -\frac{1}{2}$ then $D_\lambda(\mathbf{X},\mathbf{Y}) = -4\left[e^{-\frac{||\mathbf{v}||_2^2}{8}} - 1\right]$.

Proof: Denote the density function for $\mathbf{X}$ and $\mathbf{Y}$ by $p(\cdot)$ and $q(\cdot)$ correspondingly, that is

$$p(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^m} e^{-\frac{||\mathbf{x}-\mathbf{w}_1||_2^2}{2\sigma^2}} \quad \text{and} \quad q(\mathbf{x}) = \frac{1}{(\sqrt{2\pi}\sigma)^m} e^{-\frac{||\mathbf{x}-\mathbf{w}_2||_2^2}{2\sigma^2}}.$$

Let $\mathbf{y} = \mathbf{x} - \mathbf{w}_2$, $\mathbf{v} = \mathbf{w}_1 - \mathbf{w}_2$, and denote by y_i, v_i the i^{th} element of $\mathbf{y}$ and $\mathbf{v}$. For the case $\lambda(\lambda+1) \neq 0$, the power divergence with parameter λ between $\mathbf{X}$ and $\mathbf{Y}$ is given by

$$\begin{aligned} D_\lambda(\mathbf{X},\mathbf{Y}) =& \frac{1}{\lambda(\lambda+1)} \int_{\mathbb{R}^m} \left[\frac{p^{\lambda+1}(\mathbf{x})}{q^{\lambda+1}(\mathbf{x})} \cdot q(\mathbf{x}) - q(\mathbf{x})\right] \mathrm{d}\mathbf{x} \\ =& \frac{1}{\lambda(\lambda+1)} \left[\int_{\mathbb{R}^m} \frac{1}{(\sqrt{2\pi}\sigma)^m} e^{-\frac{(\lambda+1)||\mathbf{x}-\mathbf{w}_1||_2^2 - \lambda||\mathbf{x}-\mathbf{w}_2||_2^2}{2\sigma^2}} \mathrm{d}\mathbf{x} - 1\right] \\ =& \frac{1}{\lambda(\lambda+1)} \left[\int_{\mathbb{R}^m} \frac{1}{(\sqrt{2\pi}\sigma)^m} e^{-\frac{(\lambda+1)||\mathbf{y}-\mathbf{v}||_2^2 - \lambda||\mathbf{y}||_2^2}{2\sigma^2}} \mathrm{d}\mathbf{y} - 1\right] \\ =& \frac{1}{\lambda(\lambda+1)} \left[\left(\prod_{i=1}^m \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\lambda+1)(y_i-v_i)^2 - \lambda y_i^2}{2\sigma^2}} \mathrm{d}y_i\right) - 1\right] \\ =& \frac{1}{\lambda(\lambda+1)} \left[\left(\prod_{i=1}^m e^{\frac{\lambda(\lambda+1)v_i^2}{2\sigma^2}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_i-(1+\lambda)v_i)^2}{2\sigma^2}} \mathrm{d}y_i\right) - 1\right] \\ =& \frac{1}{\lambda(\lambda+1)} \left[e^{\frac{\lambda(\lambda+1)||\mathbf{v}||_2^2}{2\sigma^2}} - 1\right]. \end{aligned}$$

Next consider the case $\lambda(\lambda+1) = 0$. Denote the i^{th} element of $\mathbf{x}$, $\mathbf{w}_1$, and $\mathbf{w}_2$ by $x_i, w_{1,i}, w_{2,i}$ correspondingly. First we study the case $\lambda = 0$.

$$\begin{aligned}
D_0(\mathbf{X},\mathbf{Y}) &= \int_{\mathbb{R}^m} p(\mathbf{x}) \log(\frac{p(\mathbf{x})}{q(\mathbf{x})})\mathrm{d}x \\
&= \int_{\mathbb{R}^m} p(\mathbf{x}) \frac{-(||\mathbf{x}-\mathbf{w}_1||_2^2 - ||\mathbf{x}-\mathbf{w}_2||_2^2)}{2\sigma^2}\mathrm{d}\mathbf{x} \\
&= \sum_{i=1}^m \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-w_{1,i})^2}{2\sigma^2}} \cdot \frac{(w_{1,i}-w_{2,i})(2x-w_{1,i}-w_{2,i})}{2\sigma^2}\mathrm{d}x_i \\
&= \sum_{i=1}^m \left(\frac{w_{1,i}-w_{2,i}}{2\sigma^2} \mathbf{E}_{X\sim N(w_{1,i},\sigma^2)}[2X - w_{1,i} - w_{2,i}] \right) \\
&= \frac{||\mathbf{w}_1-\mathbf{w}_2||_2^2}{2\sigma^2} = \frac{||\mathbf{v}||_2^2}{2\sigma^2}
\end{aligned}$$

The case $\lambda = -1$ is similar and this completes the proof. ∎

B.2 Laplace mechanism

Lemma B.2 *For two m dimensional random variable $\mathbf{X}$ and $\mathbf{Y}$, where $X_i \sim Lap(w_{1,i}, b)$ and $Y_i \sim Lap(w_{2,i}, b)$, the power divergence between them, with parameter λ is given by*

$$D_\lambda(\mathbf{X},\mathbf{Y}) = \begin{cases} \frac{1}{\lambda(\lambda+1)} \left[\left(\prod_{i=1}^m \int_{\mathbb{R}} \frac{1}{2b} e^{-\frac{(\lambda+1)|y_i-v_i|-\lambda|y_i|}{b}} \mathrm{d}y_i \right) - 1 \right], & \lambda(\lambda+1) \neq 0 \\ \sum_{i=1}^m \int_{\mathbb{R}} \frac{1}{2b} e^{-\frac{|y_j|}{b}} \cdot \frac{|y_i-v_i|-|y_i|}{b} \mathrm{d}y_i, & \lambda(\lambda+1) = 0, \end{cases}$$

where $v_i = w_{1,i} - w_{2,i}$. Furthermore

$$D_\lambda(\mathbf{X},\mathbf{Y}) \leq \begin{cases} \frac{1}{\lambda(\lambda+1)} \left[e^{\frac{sign(\lambda)(\lambda+1)||\mathbf{v}||_1}{b}} - 1 \right], & \lambda(\lambda+1) \neq 0 \\ \frac{||\mathbf{v}||_1}{b}, & \lambda(\lambda+1) = 0, \end{cases}$$

In particular, if $\lambda = -\frac{1}{2}$, then $D_\lambda(\mathbf{X},\mathbf{Y}) = -4 \left[\left(\prod_{i=1}^m \int_{\mathbb{R}} \frac{1}{2b} e^{-\frac{|y_i-v_i|+|y_i|}{2b}} \mathrm{d}y_i \right) - 1 \right] \leq -4 \left[e^{\frac{-||\mathbf{v}||_1}{2b}} - 1 \right]$.

Proof: Denote the density function for $\mathbf{X}$ and $\mathbf{Y}$ by $p(\cdot)$ and $q(\cdot)$ correspondingly, that is

$$p(\mathbf{x}) = \frac{1}{(2b)^m} e^{-\frac{||\mathbf{x}-\mathbf{w}_1||_1}{b}} \quad \text{and} \quad q(\mathbf{x}) = \frac{1}{(2b)^m} e^{-\frac{||\mathbf{x}-\mathbf{w}_2||_1}{b}}.$$

Let $\mathbf{y} = \mathbf{x} - \mathbf{w}_2$, $\mathbf{v} = \mathbf{w}_1 - \mathbf{w}_2$, and denote y_i, v_i the i^{th} element of $\mathbf{y}$ and $\mathbf{v}$. For the case $\lambda(\lambda+1) \neq 0$, the power divergence with parameter λ between $\mathbf{X}$ and $\mathbf{Y}$ is given by

$$\begin{aligned}
D_\lambda(\mathbf{X},\mathbf{Y}) &= \frac{1}{\lambda(\lambda+1)} \int_{\mathbb{R}^m} \frac{p^{\lambda+1}(\mathbf{x})}{q^{\lambda+1}(\mathbf{x})} \cdot q(\mathbf{x}) - q(\mathbf{x})\mathrm{d}\mathbf{x} \\
&= \frac{1}{\lambda(\lambda+1)} \left[\int_{\mathbb{R}^m} \frac{1}{(2b)^m} e^{-\frac{(\lambda+1)||\mathbf{y}-\mathbf{v}||_1 - \lambda||\mathbf{y}||_1}{b}} \mathrm{d}\mathbf{y} - 1 \right] \\
&= \frac{1}{\lambda(\lambda+1)} \left[\left(\prod_{i=1}^m \int_{\mathbb{R}} \frac{1}{2b} e^{-\frac{(\lambda+1)|y_i-v_i|-\lambda|y_i|}{b}} \mathrm{d}y_i \right) - 1 \right].
\end{aligned}$$

Furthermore,

$$
\begin{aligned}
D_\lambda(\mathbf{X},\mathbf{Y}) \leq & \frac{1}{\lambda(\lambda+1)}\left[\left(\prod_{i=1}^m \int_{\mathbb{R}} \frac{1}{2b} e^{-\frac{(\lambda+1)|y_i| - sign(\lambda)(\lambda+1)|v_i| - \lambda|y_i|}{b}} \mathrm{d}y_i\right) - 1\right] \\
= & \frac{1}{\lambda(\lambda+1)}\left[\left(\prod_{i=1}^m e^{\frac{sign(\lambda)(\lambda+1)|v_i|}{b}} \int_{\mathbb{R}} \frac{1}{2b} e^{\frac{|y_i|}{b}} \mathrm{d}y_i\right) - 1\right] \\
= & \frac{1}{\lambda(\lambda+1)}\left[e^{\frac{sign(\lambda)(\lambda+1)||\mathbf{v}||_1}{b}} - 1\right].
\end{aligned}
$$

Next we consider the case $\lambda(\lambda+1) = 0$. Denote the i^{th} element of $\mathbf{x}$, $\mathbf{w}_1$, and $\mathbf{w}_2$ by $x_i, w_{1,i}, w_{2,i}$ correspondingly. We first study the case $\lambda = 0$.To this end,

$$
\begin{aligned}
D_0(\mathbf{X},\mathbf{Y}) &= \int_{\mathbb{R}^m} p(\mathbf{x}) \log\left(\frac{p(\mathbf{x})}{q(\mathbf{x})}\right) \mathrm{d}\mathbf{x} \\
&= \int_{\mathbb{R}^m} p(\mathbf{x}) \frac{-(||\mathbf{x}-\mathbf{w}_1||_1 - ||\mathbf{x}-\mathbf{w}_2||_1)}{b} \mathrm{d}\mathbf{x} \\
&= \sum_{i=1}^m \int_{\mathbb{R}} \frac{1}{2b} e^{-\frac{|x_i - w_{1,i}|}{b}} \cdot \frac{|x_i - w_{2,i}| - |x_i - w_{1,i}|}{b} \mathrm{d}x_i \\
&= \sum_{i=1}^m \int_{\mathbb{R}} \frac{1}{2b} e^{-\frac{|y_i|}{b}} \cdot \frac{|y_i - v_i| - |y_i|}{b} \mathrm{d}y_i.
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
D_0(\mathbf{X},\mathbf{Y}) \leq & \sum_{i=1}^m \int_{\mathbb{R}} \frac{1}{2b} e^{-\frac{|x_i - w_{1,i}|}{b}} \cdot \left|\frac{|x_i - w_{2,i}| - |x_i - w_{1,i}|}{b}\right| \mathrm{d}x_i \\
\leq & \sum_{i=1}^m \int_{\mathbb{R}} \frac{1}{2b} e^{-\frac{|x_i - w_{1,i}|}{b}} \cdot \frac{|w_{1,i} - w_{2,i}|}{b} \mathrm{d}x_i = \frac{||\mathbf{w}_1 - \mathbf{w}_2||_1}{b}.
\end{aligned}
$$

The case $\lambda = -1$ is similar, and this completes the proof. ∎

B.3 Exact Laplace mechanism

Lemma B.3 *For two m dimensional random variable $\mathbf{X}$ and $\mathbf{Y}$, where $X_i \sim Lap(w_{1,i}, b)$ and $Y_i \sim Lap(w_{2,i}, b)$, the $\lambda-$power divergence between, $D_\lambda(\mathbf{X},\mathbf{Y})$, for $\lambda(\lambda+1) \neq 0$ or $\lambda(\lambda+1) \neq -\frac{1}{2}$ given by*

$$
\frac{1}{\lambda(\lambda+1)}\left[\left(\prod_{i=1}^m \frac{1}{2b}\left[e^{\frac{\lambda|v_i|}{b}}\left(b + \frac{b}{2\lambda+1}\right) + e^{-\frac{(\lambda+1)|v_i|}{b}}\left(b - \frac{b}{2\lambda+1}\right)\right]\right) - 1\right].
$$

If $\lambda(\lambda+1) = 0$ or $\lambda(\lambda+1) = -\frac{1}{2}$, then

$$
D_\lambda(\mathbf{X},\mathbf{Y}) = \begin{cases} -4\left[\left(\prod_{i=1}^m e^{\frac{-|v_i|}{2b}} + \frac{|v_i|}{2b} e^{-\frac{|v_i|}{2b}}\right) - 1\right], & \lambda = -\frac{1}{2} \\ \frac{1}{2b}\sum_{i=1}^m \left[2|v_i| - 2b + 2be^{-\frac{|v_i|}{b}}\right], & \lambda(\lambda+1) = 0. \end{cases}
$$

Proof: In case $\lambda(\lambda+1)\neq 0, \lambda\neq -\frac{1}{2}$, using Lemma B.2, it follows that

$$D_\lambda(\mathbf{X},\mathbf{Y}) =\frac{1}{\lambda(\lambda+1)}\left[\left(\prod_{i=1}^m\int_{\mathbb{R}}\frac{1}{2b}e^{-\frac{(\lambda+1)|y_i-v_i|-\lambda|y_i|}{b}}\mathrm{d}y_i\right)-1\right].$$

For each i, we remove the absolute sign by studying case $v_i<0$ and $v_i\geq 0$. If $v_i<0$,

$$\begin{aligned}
&\int_{\mathbb{R}}\frac{1}{2b}e^{-\frac{(\lambda+1)|y_i-v_i|-\lambda|y_i|}{b}}\mathrm{d}y_i\\
=&\frac{1}{2b}\left[\int_{-\infty}^{v_i}e^{-\frac{-(\lambda+1)(y_i-v_i)+\lambda y_i}{b}}\mathrm{d}y_i+\int_{v_i}^{0}e^{-\frac{(\lambda+1)(y_i-v_i)+\lambda y_i}{b}}\mathrm{d}y_i+\int_{0}^{\infty}e^{-\frac{(\lambda+1)(y_i-v_i)-\lambda y_i}{b}}\mathrm{d}y_i\right]\\
=&\frac{1}{2b}\left[e^{\frac{-\lambda v_i}{b}}\left(b+\frac{b}{2\lambda+1}\right)+e^{\frac{(\lambda+1)v_i}{b}}\left(b-\frac{b}{2\lambda+1}\right)\right].
\end{aligned}$$

If $v_i\geq 0$,

$$\begin{aligned}
&\int_{\mathbb{R}}\frac{1}{2b}e^{-\frac{(\lambda+1)|y_i-v_i|-\lambda|y_i|}{b}}\mathrm{d}y_i\\
=&\frac{1}{2b}\left[\int_{-\infty}^{0}e^{-\frac{-(\lambda+1)(y_i-v_i)+\lambda y_i}{b}}\mathrm{d}y_i+\int_{0}^{v_i}e^{-\frac{-(\lambda+1)(y_i-v_i)-\lambda y_i}{b}}\mathrm{d}y_i+\int_{v_i}^{\infty}e^{-\frac{(\lambda+1)(y_i-v_i)-\lambda y_i}{b}}\mathrm{d}y_i\right]\\
=&\frac{1}{2b}\left[e^{\frac{\lambda v_i}{b}}\left(b+\frac{b}{2\lambda+1}\right)+e^{\frac{-(\lambda+1)v_i}{b}}\left(b-\frac{b}{2\lambda+1}\right)\right].
\end{aligned}$$

Combining the cases $v_i<0$ and $v_i\geq 0$, we get

$$\int_{\mathbb{R}}\frac{1}{2b}e^{-\frac{(\lambda+1)|y_i-v_i|-\lambda|y_i|}{b}}\mathrm{d}y_i=\frac{1}{2b}\left[e^{\frac{\lambda|v_i|}{b}}\left(b+\frac{b}{2\lambda+1}\right)+e^{-\frac{(\lambda+1)|v_i|}{b}}\left(b-\frac{b}{2\lambda+1}\right)\right].$$

Therefore,

$$D_\lambda(\mathbf{X},\mathbf{Y})=\frac{1}{\lambda(\lambda+1)}\left[\left(\prod_{i=1}^m\frac{1}{2b}\left[e^{\frac{\lambda|v_i|}{b}}\left(b+\frac{b}{2\lambda+1}\right)+e^{-\frac{(\lambda+1)|v_i|}{b}}\left(b-\frac{b}{2\lambda+1}\right)\right]\right)-1\right].$$

We now turn to the case $\lambda=-\frac{1}{2}$. If $v_i<0$, by the same calculation of the integral, it follows that

$$\int_{\mathbb{R}}\frac{1}{2b}e^{-\frac{(\lambda+1)|y_i-v_i|-\lambda|y_i|}{b}}\mathrm{d}y_i=e^{\frac{v_i}{2b}}-\frac{v_i}{2b}e^{\frac{v_i}{2b}}.$$

If $v_i\geq 0$,

$$\int_{\mathbb{R}}\frac{1}{2b}e^{-\frac{(\lambda+1)|y_i-v_i|-\lambda|y_i|}{b}}\mathrm{d}y_i=e^{\frac{-v_i}{2b}}+\frac{v_i}{2b}e^{-\frac{v_i}{2b}}.$$

Combining $v_i<0$ and $v_i\geq 0$, we get

$$\int_{\mathbb{R}}\frac{1}{2b}e^{-\frac{(\lambda+1)|y_i-v_i|-\lambda|y_i|}{b}}\mathrm{d}y_i=e^{\frac{-|v_i|}{2b}}+\frac{|v_i|}{2b}e^{-\frac{|v_i|}{2b}}.$$

Therefore

$$D_\lambda(\mathbf{X},\mathbf{Y}) = -4\left[\left(\prod_{i=1}^{m} e^{\frac{-|v_i|}{2b}} + \frac{|v_i|}{2b}e^{-\frac{|v_i|}{2b}}\right) - 1\right].$$

Finally, we turn to the case $\lambda(\lambda+1)=0$. We study the case $\lambda=0$. Using Lemma B.2, it follows that

$$D_0(\mathbf{X},\mathbf{Y}) = \sum_{i=1}^{m}\int_{\mathbb{R}} \frac{1}{2b}e^{-\frac{|y_j|}{b}} \cdot \frac{|y_i - v_i| - |y_i|}{b}\mathrm{d}y_i.$$

If $v_i \geq 0$,

$$\begin{aligned}
&\int_{\mathbb{R}} e^{-\frac{|y_i|}{b}} \cdot \frac{|y_i - v_i| - |y_i|}{b}\mathrm{d}y_i \\
=&\int_{-\infty}^{0} e^{-\frac{-y_i}{b}}\frac{-(y_i - v_i) + y_i}{b}\mathrm{d}y_i + \int_{0}^{v_i} e^{-\frac{y_i}{b}}\frac{-(y_i - v_i) - y_i}{b}\mathrm{d}y_i + \int_{v_i}^{\infty} e^{-\frac{y_i}{b}}\frac{(y_i - v_i) - y_i}{b}\mathrm{d}y_i \\
=&2v_i - 2b + 2be^{-\frac{v_i}{b}}.
\end{aligned}$$

If $v_i < 0$,

$$\begin{aligned}
&\int_{\mathbb{R}} e^{-\frac{|y_i|}{b}} \cdot \frac{|y_i - v_i| - |y_i|}{b}\mathrm{d}y_i \\
=&\int_{-\infty}^{v_i} e^{-\frac{-y_i}{b}}\frac{-(y_i - v_i) + y_i}{b}\mathrm{d}y_i + \int_{v_i}^{0} e^{-\frac{-y_i}{b}}\frac{(y_i - v_i) + y_i}{b}\mathrm{d}y_i + \int_{0}^{\infty} e^{-\frac{y_i}{b}}\frac{(y_i - v_i) - y_i}{b}\mathrm{d}y_i \\
=&-2v_i - 2b + 2be^{\frac{v_i}{b}}.
\end{aligned}$$

Combining the cases $v_i < 0$ and $v_i \geq 0$, we get

$$\int_{\mathbb{R}} e^{-\frac{|y_i|}{b}} \cdot \frac{|y_i - v_i| - |y_i|}{b}\mathrm{d}y_i = 2|v_i| - 2b + 2be^{-\frac{|v_i|}{b}}.$$

Therefore,

$$D_0(\mathbf{X},\mathbf{Y}) = \frac{1}{2b}\sum_{i=1}^{m}\left[2|v_i| - 2b + 2be^{-\frac{|v_i|}{b}}\right].$$

The case $\lambda = -1$ is similar, and this completes the proof. ■

Appendix C.

C.1 Proof of Remark 5

Link to $\rho-$zCDP: Suppose a mechanism M satisfies $(\lambda, \epsilon)-$PDP for some $\lambda > 0$; this is equivalent to the statement that M satisfies $(\lambda+1, \frac{1}{\lambda}\log(\epsilon\lambda(\lambda+1)+1)-$RDP. Since $\frac{1}{\lambda}\log(\epsilon\lambda(\lambda+1)+1) \leq (\lambda+1)\epsilon$, it follows that M satisfies $(\lambda+1, (\lambda+1)\epsilon)-$RDP. Hence, by the definition of $\rho-$zCDP, it follows that M satisfies $\epsilon-$zCDP.
Link to $(\epsilon, \delta)-$differential privacy: Suppose a mechanism M satisfies $(\lambda, \epsilon)-$PDP, then by definition, $D_\lambda(f_1, f_2) \leq \epsilon$, where f_1 is the density of $M(w, D)$ and f_2 is the density of $M(w, D')$. We now determine the relationship to $(\epsilon, \delta)-$DP.

If $\lambda > 0$, then

$$\begin{aligned} D_\lambda(f_1, f_2) = \frac{1}{\lambda(\lambda+1)}\left[\int_{\mathbb{R}^m} \frac{f_1^{\lambda+1}(x)}{f_2^{\lambda}(x)}\mathrm{d}x - 1\right] \leq \epsilon \\ \iff \int_{\mathbb{R}^m} \frac{f_1^{\lambda+1}(x)}{f_2^{\lambda}(x)}\mathrm{d}x \leq \lambda(\lambda+1)\epsilon + 1 = e^{\log(\lambda(\lambda+1)\epsilon+1)}. \end{aligned}$$

For any set $A \subset \mathbb{R}^m$, applying Holder inequality for $p = \lambda+1$ and $q = \frac{\lambda+1}{\lambda}$, it follows that

$$\begin{aligned} \mathbf{P}_{X\sim f_1}(X \in A) = \int_A f_1(x)\mathrm{d}x = \int_A \frac{f_1(x)}{[f_2(x)]^{\frac{\lambda}{\lambda+1}}} \cdot [f_2(x)]^{\frac{\lambda}{\lambda+1}}\mathrm{d}x \\ \leq \left(\int_A \left(\frac{f_1(x)}{[f_2(x)]^{\frac{\lambda}{\lambda+1}}}\right)^p \mathrm{d}x\right)^{\frac{1}{p}} \cdot \left(\int_A \left([f_2(x)]^{\frac{\lambda}{\lambda+1}}\right)^q \mathrm{d}x\right)^{\frac{1}{q}} \\ = \left(\int_A \frac{[f_1(x)]^{\lambda+1}}{[f_2(x)]^{\lambda}}\mathrm{d}x\right)^{\frac{1}{\lambda+1}} \cdot \left(\int_A f_2(x)\mathrm{d}x\right)^{\frac{\lambda}{\lambda+1}} \\ \leq \left(\int_{\mathbb{R}^m} \frac{[f_1(x)]^{\lambda+1}}{[f_2(x)]^{\lambda}}\mathrm{d}x\right)^{\frac{1}{\lambda+1}} \cdot \left(\int_A f_2(x)\mathrm{d}x\right)^{\frac{\lambda}{\lambda+1}} \\ \leq e^{\frac{1}{\lambda+1}\log(\lambda(\lambda+1)\epsilon+1)} \cdot [\mathbf{P}_{X\sim f_2}(X \in A)]^{\frac{\lambda}{\lambda+1}} \\ = \left[e^{\frac{1}{\lambda}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \mathbf{P}_{X\sim f_2}(X \in A)\right]^{\frac{\lambda}{\lambda+1}}. \end{aligned}$$

If $e^{\frac{1}{\lambda}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \mathbf{P}_{X\sim f_2}(X \in A) > \delta^{\frac{\lambda+1}{\lambda}}$, then

$$\begin{aligned} \mathbf{P}_{X\sim f_1}(X \in A) \leq \left[e^{\frac{1}{\lambda}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \mathbf{P}_{X\sim f_2}(X \in A)\right]^{\frac{\lambda}{\lambda+1}} \\ = e^{\frac{1}{\lambda}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \mathbf{P}_{X\sim f_2}(X \in A) \cdot \left[e^{\frac{1}{\lambda}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \mathbf{P}_{X\sim f_2}(X \in A)\right]^{\frac{-1}{\lambda+1}} \\ \leq e^{\frac{1}{\lambda}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \mathbf{P}_{X\sim f_2}(X \in A) \cdot \delta^{\frac{-1}{\lambda}} \\ = e^{\frac{1}{\lambda}\log(\frac{\lambda(\lambda+1)\epsilon+1}{\delta})} \cdot \mathbf{P}_{X\sim f_2}(X \in A) \end{aligned} \tag{18}$$

If $e^{\frac{1}{\lambda}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \mathbf{P}_{X\sim f_2}(X \in A) \leq \delta^{\frac{\lambda+1}{\lambda}}$, then

$$\mathbf{P}_{X\sim f_1}(X \in A) \leq \left[e^{\frac{1}{\lambda}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \mathbf{P}_{X\sim f_2}(X \in A)\right]^{\frac{\lambda}{\lambda+1}} = \delta.$$

Therefore

$$\mathbf{P}_{X\sim f_1}(X \in A) \leq e^{\frac{1}{\lambda}\log(\frac{\lambda(\lambda+1)\epsilon+1}{\delta})} \cdot \mathbf{P}_{X\sim f_2}(X \in A) + \delta.$$

This implies that M satisfies $(\frac{1}{\lambda}\log(\frac{\lambda(\lambda+1)\epsilon+1}{\delta}), \delta)-$DP.

If $\lambda < -1$, write $\lambda' = -\lambda - 1 > 0$,

$$D_\lambda(f_2, f_1) = \frac{1}{\lambda(\lambda+1)}\left[\int_{\mathbb{R}^m} \frac{f_2^{\lambda+1}(x)}{f_1^{\lambda}(x)}\mathrm{d}x - 1\right] \leq \epsilon$$

$$\Longleftrightarrow \int_{\mathbb{R}^m} \frac{f_2^{\lambda+1}(x)}{f_1^{\lambda}(x)}\mathrm{d}x \leq \lambda(\lambda+1)\epsilon + 1 = e^{\log(\lambda(\lambda+1)\epsilon+1)}$$

$$\Longleftrightarrow \int_{\mathbb{R}^m} \frac{f_1^{\lambda'+1}(x)}{f_2^{\lambda'}(x)}\mathrm{d}x \leq \lambda'(\lambda'+1)\epsilon + 1 = e^{\log(\lambda'(\lambda'+1)\epsilon+1)}.$$

Applying Holder inequality for $p = \lambda' + 1 > 1$, $q = \frac{\lambda'+1}{\lambda'} > 1$, by the same method, we obtain

$$\mathbf{P}_{X\sim f_1}(X \in A) \leq \left[e^{\frac{1}{\lambda'}\log(\lambda'(\lambda'+1)\epsilon+1)} \cdot \mathbf{P}_{X\sim f_2}(X \in A)\right]^{\frac{\lambda'}{\lambda'+1}}. \tag{19}$$

Using the same δ, it follows that

$$\mathbf{P}_{X\sim f_1}(X \in A) \leq e^{\frac{-1}{\lambda+1}\log(\frac{\lambda(\lambda+1)\epsilon+1}{\delta})} \cdot \mathbf{P}_{X\sim f_2}(X \in A) + \delta.$$

This implies that M satisfies $(\frac{-1}{\lambda+1}\log(\frac{\lambda(\lambda+1)\epsilon+1}{\delta}), \delta)-$DP.

Link to $\mu-$GDP: Suppose a mechanism M satisfies $(\lambda, \epsilon)-$PDP, then $D_\lambda(f_1, f_2) \leq \epsilon$, where f_1 is the density of $M(w, D)$ and f_2 is the density of $M(w, D')$. Consider the one observation hypothesis test:

$$H : X \sim f_1 \quad vs \quad K : X \sim f_2.$$

Using the Neyman–Pearson lemma, the most powerful test function is given by

$$\tau(x) = \begin{cases} 1, & x \in A_\alpha \\ 0, & \text{otherwise,} \end{cases}$$

and A_α is determined by $\mathbf{P}_{X\sim f_1}(X \in A_\alpha) = \alpha$.

For $\lambda > 0$, by (18), it follows that

$$\begin{aligned}\mathbf{P}_{X\sim f_2}(X \in A_\alpha) \leq & e^{\frac{1}{\lambda+1}\log(\lambda(\lambda+1)\epsilon+1)} \cdot [\mathbf{P}_{X\sim f_1}(X \in A_\alpha)]^{\frac{\lambda}{\lambda+1}} \\ = & e^{\frac{1}{\lambda+1}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \alpha^{\frac{\lambda}{\lambda+1}}.\end{aligned}$$

To get μ such that M satisfies $\mu - GDP$, from the definition of $\mu - GDP$ in Dong et al. (2022), we need

$$1 - \mathbf{P}_{X\sim f_2}(X \in A_\alpha) \geq \Phi(\Phi^{-1}(1-\alpha) - \mu).$$

We only need to show for any $\alpha \in [0,1]$,

$$
\begin{aligned}
1 - e^{\frac{1}{\lambda+1}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \alpha^{\frac{\lambda}{\lambda+1}} &\geq \Phi(\Phi^{-1}(1-\alpha) - \mu) \\
\Longleftrightarrow \mu &\geq \Phi^{-1}(1-\alpha) - \Phi^{-1}(1 - e^{\frac{1}{\lambda+1}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \alpha^{\frac{\lambda}{\lambda+1}}).
\end{aligned}
$$

μ can be chosen such that

$$
\mu = \sup_{\alpha\in[0,1]} \{\Phi^{-1}(1-\alpha) - \Phi^{-1}(1 - e^{\frac{1}{\lambda+1}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \alpha^{\frac{\lambda}{\lambda+1}})\}.
$$

For $\lambda < -1$ by (19) and $\lambda' = -\lambda - 1$, we obtain

$$
\mathbf{P}_{X\sim f_2}(X \in A_\alpha) \leq e^{\frac{-1}{\lambda}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \alpha^{\frac{\lambda+1}{\lambda}}.
$$

To get μ such that M satisfies $\mu - GDP$, we need

$$
1 - \mathbf{P}_{X\sim f_2}(X \in A_\alpha) \geq \Phi(\Phi^{-1}(1-\alpha) - \mu).
$$

We only need to show for any $\alpha \in [0,1]$,

$$
\begin{aligned}
1 - e^{\frac{-1}{\lambda}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \alpha^{\frac{\lambda+1}{\lambda}} &\geq \Phi(\Phi^{-1}(1-\alpha) - \mu) \\
\Longleftrightarrow \mu &\geq \Phi^{-1}(1-\alpha) - \Phi^{-1}(1 - e^{\frac{-1}{\lambda}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \alpha^{\frac{\lambda+1}{\lambda}}).
\end{aligned}
$$

μ can be chosen such that

$$
\mu = \sup_{\alpha\in[0,1]} \{\Phi^{-1}(1-\alpha) - \Phi^{-1}(1 - e^{\frac{-1}{\lambda}\log(\lambda(\lambda+1)\epsilon+1)} \cdot \alpha^{\frac{\lambda+1}{\lambda}})\}.
$$

This completes the proof. ∎

Appendix D.

D.1 Details on the convergence of $H_n(\cdot)$ to $H_\infty(\cdot)$ in Proposition 15

We write $H_{n,i,j}(\boldsymbol{\theta})$ as the $i-th$ row and $j-th$ column element of $H_n(\boldsymbol{\theta})$, and $I_{i,j}(\boldsymbol{\theta})$ as the $i-th$ row and $j-th$ column element of $I(\boldsymbol{\theta})$. Then we only need to show $H_{n,i,j}(\boldsymbol{\theta}_0) \to I_{i,j}(\boldsymbol{\theta}_0)$ for any $i, j = 1, \cdots, m$. Recall that

$$H_{n,i,j}(\boldsymbol{\theta}) = \frac{\partial^2}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j} L_n(\boldsymbol{\theta}) = -T_{1,n,i,j}(\boldsymbol{\theta}) - 2T_{2,n,i,j}(\boldsymbol{\theta}),$$

where

$$T_{1,n,i,j}(\boldsymbol{\theta}) = \int_{\mathbb{R}} g_n^{\frac{1}{2}}(x) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},j}(x) u_{\boldsymbol{\theta},i}(x) \mathrm{d}x, \quad T_{2,n,i,j}(\boldsymbol{\theta}) = \int_{\mathbb{R}} g_n^{\frac{1}{2}}(x) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},i,j}(x) \mathrm{d}x.$$

We decompose $T_{1,n,i,j}(\boldsymbol{\theta})$ and $T_{2,n,i,j}(\boldsymbol{\theta})$ as follows:

$$T_{1,n,i,j}(\boldsymbol{\theta}) = T^{(1)}_{1,n,i,j}(\boldsymbol{\theta}) + I_{i,j}(\boldsymbol{\theta}), \quad \text{and} \quad T_{2,n,i,j}(\boldsymbol{\theta}) = T^{(1)}_{2,n,i,j}(\boldsymbol{\theta}) - I_{i,j}(\boldsymbol{\theta}),$$

where

$$\begin{aligned} T^{(1)}_{1,n,i,j}(\boldsymbol{\theta}) &= \int_{\mathbb{R}} \left(g_n^{\frac{1}{2}}(x) - f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) \right) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},j}(x) u_{\boldsymbol{\theta},i}(x) \mathrm{d}x, \quad \text{and} \\ T^{(1)}_{2,n,i,j}(\boldsymbol{\theta}) &= \int_{\mathbb{R}} \left(g_n^{\frac{1}{2}}(x) - f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) \right) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},i,j}(x) \mathrm{d}x. \end{aligned}$$

Then $H_{n,i,j}(\boldsymbol{\theta})$ can be written as follows:

$$H_{n,i,j}(\boldsymbol{\theta}) = I_{i,j}(\boldsymbol{\theta}) - D_{n,i,j}(\boldsymbol{\theta}), \tag{20}$$

where $D_{n,i,j}(\boldsymbol{\theta}) = T^{(1)}_{1,n,i,j}(\boldsymbol{\theta}) + 2T^{(1)}_{2,n,i,j}(\boldsymbol{\theta})$. We are going to show $D_{n,i,j}(\boldsymbol{\theta}) \to D_{i,j}(\boldsymbol{\theta})$ almost surely as $n \to \infty$, where $D_{i,j}(\boldsymbol{\theta}) = \int_{\mathbb{R}} \left(g^{\frac{1}{2}}(x) - f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) \right) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) [u_{\boldsymbol{\theta},i}(x) u_{\boldsymbol{\theta},j}(x) + 2u_{\boldsymbol{\theta},i,j}(x)] \mathrm{d}x$. Notice that

$$D_{n,i,j}(\boldsymbol{\theta}) = D_{i,j}(\boldsymbol{\theta}) - \int_{\mathbb{R}} \left(g_n^{\frac{1}{2}}(x) - g^{\frac{1}{2}}(x) \right) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) [u_{\boldsymbol{\theta},i}(x) u_{\boldsymbol{\theta},j}(x) + 2u_{\boldsymbol{\theta},i,j}(x)] \mathrm{d}x.$$

We are going to show $\int_{\mathbb{R}} \left(g_n^{\frac{1}{2}}(x) - g^{\frac{1}{2}}(x) \right) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) [u_{\boldsymbol{\theta},i}(x) u_{\boldsymbol{\theta},j}(x) + 2u_{\boldsymbol{\theta},i,j}(x)] \mathrm{d}x \to 0$ almost surely as $n \to \infty$. Using Cauchy–Schwarz inequality and the upper bounds in assumption **(U1)-(U2)**, it follows that as $n \to \infty$,

$$\begin{aligned} &\left| \int_{\mathbb{R}} \left(g_n^{\frac{1}{2}}(x) - g^{\frac{1}{2}}(x) \right) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) [u_{\boldsymbol{\theta},i}(x) u_{\boldsymbol{\theta},j}(x) + 2u_{\boldsymbol{\theta},i,j}(x)] \mathrm{d}x \right| \\ \le & HD(g_n, g) \cdot \mathbf{E}_{\boldsymbol{\theta}} \left[u^2_{\boldsymbol{\theta},i}(X) u^2_{\boldsymbol{\theta},j}(X) \right] + 2HD(g_n, g) \cdot \mathbf{E}_{\boldsymbol{\theta}} \left[u^2_{\boldsymbol{\theta},i,j}(X) \right] \\ \le & c \cdot HD(g_n, g) \xrightarrow{a.s.} 0. \end{aligned}$$

The convergence follows from $HD(g_n, g) \xrightarrow{a.s.} 0$ when $n \to \infty$, since by assumption **(A2)**, $||g_n - g||_1$ converges to zero almost surely. Thus, $H_\infty(\boldsymbol{\theta}) = I(\boldsymbol{\theta}) - D(\boldsymbol{\theta})$. This completes the proof.

D.2 Establish the upper bound for $D_{i,j}(\cdot)$ in Proposition 16

Using Cauchy- Schwarz inequality and **(U1)-(U2)** the result follows. To see this, notice that

$$|D_{i,j}(\boldsymbol{\theta})| = \left| \int_{\mathbb{R}} \left(g^{\frac{1}{2}}(x) - f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) \right) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) [u_{\boldsymbol{\theta},i}(x) u_{\boldsymbol{\theta},j}(x) + 2u_{\boldsymbol{\theta},i,j}(x)] \mathrm{d}x \right|.$$

Now, splitting the RHS of the above equation, we see that it is bounded above by

$$\left| \int_{\mathbb{R}} \left(g^{\frac{1}{2}}(x) - f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) \right) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},i}(x) u_{\boldsymbol{\theta},j}(x) \mathrm{d}x \right| + 2 \left| \int_{\mathbb{R}} \left(g^{\frac{1}{2}}(x) - f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) \right) f_{\boldsymbol{\theta}}^{\frac{1}{2}}(x) u_{\boldsymbol{\theta},i,j}(x) \mathrm{d}x \right|.$$

Now, applying Cauchy-Schwarz inequality, we get

$$\begin{aligned} |D_{i,j}(\boldsymbol{\theta})| \leq & HD(g, f_{\boldsymbol{\theta}}) \cdot \mathbf{E}_{\boldsymbol{\theta}} \left[u^2_{\boldsymbol{\theta},i}(X) u^2_{\boldsymbol{\theta},j}(X) \right] + 2HD(g, f_{\boldsymbol{\theta}}) \cdot \mathbf{E}_{\boldsymbol{\theta}} \left[u^2_{\boldsymbol{\theta},i,j}(X) \right] \\ \leq & c \cdot HD(g, f_{\boldsymbol{\theta}}), \end{aligned} \tag{21}$$

where $0 < c = \sup_{\boldsymbol{\theta} \in \Theta} \max \left\{ \mathbf{E}_{\boldsymbol{\theta}} \left[u^2_{\boldsymbol{\theta},i}(X) u^2_{\boldsymbol{\theta},j}(X) \right], 2\mathbf{E}_{\boldsymbol{\theta}} \left[u^2_{\boldsymbol{\theta},i,j}(X) \right] \right\} < \infty$.

D.3 Proof of Lemma 6.2

Statement: Let assumptions **(A1)-(A8)** and **(U1)-(U2)** hold. Then for $\boldsymbol{\theta} \in B_r(\boldsymbol{\theta}_g)$ and $n \geq N$, if $L_n(\boldsymbol{\theta}) - L_n(\hat{\boldsymbol{\theta}}) \leq \frac{r^2}{4}\tau_1$ then $||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}||_2 \leq \frac{r}{2}$. Furthermore, if $||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}||_2 \leq \frac{r}{2}$ for $\boldsymbol{\theta} \in B_r(\boldsymbol{\theta}_g)$, then for $n \geq N$, $L_n(\boldsymbol{\theta}) - L_n(\hat{\boldsymbol{\theta}}) \leq \frac{r^2}{4}\tau_2$.
Proof: Let $n \geq N$ and $\boldsymbol{\theta}, \hat{\boldsymbol{\theta}} \in B_r(\boldsymbol{\theta}_g)$. Suppose $L_n(\boldsymbol{\theta}) - L_n(\hat{\boldsymbol{\theta}}) \leq \frac{r^2}{4}\tau_1$. Then using Proposition 18 (1), it follows that $L_n(\boldsymbol{\theta}) \geq L_n(\hat{\boldsymbol{\theta}}) + \langle \nabla L_n(\hat{\boldsymbol{\theta}}), \boldsymbol{\theta} - \hat{\boldsymbol{\theta}} \rangle + \tau_1 ||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}||_2^2$. Since $\nabla L_n(\hat{\boldsymbol{\theta}}) = (0, \cdots, 0)$ it follows that $\frac{r^2\tau_1}{4} \geq L_n(\boldsymbol{\theta}) - L_n(\hat{\boldsymbol{\theta}}) \geq \tau_1 ||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}||_2^2$, the result follows. The rest of the proof follows similarly, using Proposition 18 (3); that is, if $||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}||_2 \leq \frac{r}{2}$, then $L_n(\boldsymbol{\theta}) - L_n(\hat{\boldsymbol{\theta}}) \leq \langle \nabla L_n(\hat{\boldsymbol{\theta}}), \boldsymbol{\theta} - \hat{\boldsymbol{\theta}} \rangle + \tau_{2,n} ||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}||_2^2 = \tau_2 ||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}||_2^2 \leq \frac{r^2}{4}\tau_2$. ■

D.4 Proof of Lemma 6.3

Statement: Under assumptions **(A1)-(A8)** and **(U1)-(U2)**, for $\eta \leq \frac{1}{\tau_2}$, assume that for $n \geq N$, $\hat{\boldsymbol{\theta}}_n \in B_{r/c}(\boldsymbol{\theta}_g) \subset B_{r/2}(\boldsymbol{\theta}_g)$, where $c > 2\left(\frac{\tau 2}{\tau_1}\right)^{\frac{1}{2}}$, then there exists $\hat{\boldsymbol{\theta}}_n^{(0)}$, such that $L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq \tau_1 \frac{r^2}{4}$ and $||\hat{\boldsymbol{\theta}}_n^{(k)} - \hat{\boldsymbol{\theta}}_n||_2 \leq \frac{r}{2}$ hold with probability $1 - \frac{k\xi}{K}$ for all $k = 0, \cdots, K$.
Proof: The Lemma states that under the stated conditions, the estimators from each iteration, $\hat{\boldsymbol{\theta}}_n^{(k)}$, belong to the ball $B_{r/2}(\hat{\boldsymbol{\theta}}_n) \subset B_r(\boldsymbol{\theta}_g)$. We prove the result by induction. First, for $k = 0$, we choose the initial estimator to be a consistent estimator of $\boldsymbol{\theta}_g$. Hence for large n, $||\hat{\boldsymbol{\theta}}_n^{(0)} - \boldsymbol{\theta}_g|| \leq \left(\frac{\tau_1}{\tau_2}\right)^{\frac{1}{2}} \frac{r}{2} - \frac{r}{c}$. Hence, for large n, $||\hat{\boldsymbol{\theta}}_n^{(0)} - \hat{\boldsymbol{\theta}}_n||_2 \leq \left(\frac{\tau_1}{\tau_2}\right)^{\frac{1}{2}} \frac{r}{2}$. By Lemma 6.2, $L_n(\hat{\boldsymbol{\theta}}_n^{(0)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq \tau_1 \frac{r^2}{4}$ and $||\hat{\boldsymbol{\theta}}_n^{(0)} - \hat{\boldsymbol{\theta}}_n||_2 \leq \frac{r}{2}$ hold. Hence, by induction hypothesis, let $L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq \tau_1 \frac{r^2}{4}$ and $||\hat{\boldsymbol{\theta}}_n^{(k)} - \hat{\boldsymbol{\theta}}_n||_2 \leq \frac{r}{2}$ hold. We will establish that $L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq \tau_1 \frac{r^2}{4}$ and $||\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n||_2 \leq \frac{r}{2}$. The proof of this relies on the behavior

of $||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2$, $||N_{n,k}||_2$, and their relationships which is described in the following claim whose proof is relegated to the end.
Claim: If $||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2 \geq \sqrt{\frac{(1+2\eta\tau_2)B||N_{n,k}||_2+\eta\tau_2||N_{n,k}||_2^2}{1-\eta\tau_2}}$, where B is the upper bound of $||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2$ from Proposition 14, then the following inequality holds.

$$L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n).$$

Using the claim for k, and the assumption $L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq \tau_1\frac{r^2}{4}$, we obtain that $L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq \tau_1\frac{r^2}{4}$. Next, applying Lemma 6.2 we get $||\hat{\boldsymbol{\theta}}_n^{(k+1)} - \hat{\boldsymbol{\theta}}_n||_2 \leq \frac{r}{2}$. This completes the proof under the condition of the claim.

Next, we turn to the case $||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2 < \sqrt{\frac{(1+2\eta\tau_2)B||N_{n,k}||_2+\eta\tau_2||N_{n,k}||_2^2}{1-\eta\tau_2}}$. By (13), it follows that with probability $1-\frac{k\xi}{K}$ (since we have k iterations here)

$$||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2 < \sqrt{\frac{(1+2\eta\tau_2)Br_{noi}+\eta\tau_2 r_{noi}^2}{1-\eta\tau_2}} := \bar{r}_{noi}.$$

Using Proposition 18 (1), $\tau_1||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}||_2^2 \leq L_n(\hat{\boldsymbol{\theta}}_n) - L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - \langle \nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}), \hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}\rangle$. Since $\hat{\boldsymbol{\theta}}_n$ is the minimizer of $L_n(\boldsymbol{\theta})$, it follows that $\tau_1||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}||_2^2 \leq |\langle \nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}), \hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}\rangle|$. Now, applying the Cauchy-Schwarz inequality, it follows that $\tau_1||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}||_2^2 \leq ||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2 \cdot ||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}||_2$. Hence, we obtain

$$||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k)}||_2 \leq \frac{||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2}{\tau_1} \leq \frac{\bar{r}_{noi}}{\tau_1}.$$

Now using $\hat{\boldsymbol{\theta}}_n^{(k)} - \hat{\boldsymbol{\theta}}_n^{(k+1)} = \eta(\nabla L_n(\hat{\boldsymbol{\theta}}_n^{k}) + N_k)$, it follows that

$$\begin{aligned}||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k+1)}||_2 \leq& ||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{k}||_2 + \eta||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2 + \eta||N_k||_2 \leq \frac{\bar{r}_{noi}}{\tau_1} + \eta\bar{r}_{noi} + \eta r_{noi}\\ \leq& \left(\frac{\tau_1}{\tau_2}\right)^{\frac{1}{2}}\frac{r}{2} \leq \frac{r}{2},\end{aligned}$$

where the last inequality follows by taking n large. This is equivalent to choosing n such that

$$\Delta_n\frac{\left(4\sqrt{m}+2\sqrt{2\log(\frac{K}{\xi})}\right)}{\sqrt{8\log(1-0.5\frac{\epsilon}{K})}} \leq r_u.$$

Finally, the inequality $L_n(\hat{\boldsymbol{\theta}}^{(k+1)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq \tau_1\frac{r^2}{4}$ follows using Lemma 6.2 and $||\hat{\boldsymbol{\theta}}_n - \hat{\boldsymbol{\theta}}_n^{(k+1)}||_2 \leq \left(\frac{\tau_1}{\tau_2}\right)^{\frac{1}{2}}\frac{r}{2}$. This completes the induction. To complete the proof of the Lemma, we now establish the claim.
Proof of the claim: Using equation (8) in Definition 21, and let $\boldsymbol{\theta}^* = \gamma\hat{\boldsymbol{\theta}}_n^{(k+1)} + (1-\gamma)\hat{\boldsymbol{\theta}}_n^{(k)}$ for some $\gamma \in [0,1]$ in the Taylor expansion of $L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)})$ up-to second order, and apply

Cauchy–Schwarz inequality to get

$$\begin{aligned}
L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - L_n(\hat{\boldsymbol{\theta}}_n) =& L_n(\hat{\boldsymbol{\theta}}_n^{(k)} - \eta(\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + N_k)) - L_n(\hat{\boldsymbol{\theta}}_n) \\
\leq& L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n) - \eta||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2^2 + \eta||N_{n,k}||_2 \cdot ||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2 \\
&+ \frac{\eta^2}{2}\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})^T H_n(\boldsymbol{\theta}^*)\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) + \frac{\eta^2}{2}N_{n,k}^T H_n(\boldsymbol{\theta}^*)N_{n,k} \\
&+ \eta^2 N_{n,k}^T H_n(\boldsymbol{\theta}^*)\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}).
\end{aligned}$$

Furthermore, use Proposition 18 (3) and Cauchy–Schwarz inequality to get

$$\begin{aligned}
&\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})^T H_n(\theta^*)\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) \leq 2\tau_2||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2^2, \\
&N_{n,k}^T H_n(\boldsymbol{\theta}^*)N_k \leq 2\tau_2||N_{n,k}||_2^2, \quad \text{and} \\
&|N_{n,k}^T H_n(\boldsymbol{\theta}^*)\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})| \leq ||N_{n,k}^T||_2 \cdot ||H_n(\boldsymbol{\theta}^*)\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2 \leq 2\tau_2||N_{n,k}^T||_2 \cdot ||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2.
\end{aligned}$$

These give the upper bound of $L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - L_n(\hat{\boldsymbol{\theta}}_n)$ as follows:

$$L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n) + Term_{n,k},$$

where $Term_{n,k}$ is given by

$$\begin{aligned}
&-\eta||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2^2 + \eta||N_{n,k}||_2 \cdot ||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2 + \eta^2\tau_2||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2^2 + \eta^2\tau_2||N_{n,k}||_2^2 \\
&+ 2\eta^2\tau_2||N_{n,k}^T||_2 \cdot ||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2.
\end{aligned}$$

Hence the condition $||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2 \geq \sqrt{\frac{(1+2\eta\tau_2)B||N_{n,k}||_2 + \eta\tau_2||N_{n,k}||_2^2}{1-\eta\tau_2}}$ implies that $Term_{n,k} \leq 0$, that is

$$\begin{aligned}
&-\eta||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2^2 + \eta||N_{n,k}||_2 \cdot ||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2 + \eta^2\tau_2||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2^2 \\
&+ \eta^2\tau_2||N_{n,k}||_2^2 + 2\eta^2\tau_2||N_{n,k}^T||_2 \cdot ||\nabla L_n(\hat{\boldsymbol{\theta}}_n^{(k)})||_2 \leq 0,
\end{aligned}$$

and furthermore

$$L_n(\hat{\boldsymbol{\theta}}_n^{(k+1)}) - L_n(\hat{\boldsymbol{\theta}}_n) \leq L_n(\hat{\boldsymbol{\theta}}_n^{(k)}) - L_n(\hat{\boldsymbol{\theta}}_n).$$

This completes the proof. ■

Appendix E.

In this appendix, we provide a Monte-Carlo approximation to the loss function and give calculation details for the Normal distribution used in numerical experiments.

$$\begin{aligned}\tilde{L}_n(\boldsymbol{\theta}) =&2\int_{\mathbb{R}}(\sqrt{f_{\boldsymbol{\theta}}(x)}-\sqrt{g_n(x)})^2\mathrm{d}x = 2\left[2-2\int_{\mathbb{R}} g_n(x)\left(\frac{f_{\boldsymbol{\theta}}(x)}{g_n(x)}\right)^{\frac{1}{2}}\mathrm{d}x\right]\\ \approx&2\left[2-\frac{2}{r_n}\sum_{i=1}^{r_n}\left(\frac{f_{\boldsymbol{\theta}}(X_{n,i})}{g_n(X_{n,i})}\right)^{\frac{1}{2}}\right],\end{aligned}$$

where $\{X_{n,1},\cdots X_{n,r_n}\}|(X_1,\cdots X_n)$ are i.i.d. $g_n(\cdot)$. Next, the gradient is given by

$$\nabla\tilde{L}_n(\boldsymbol{\theta}) = -\frac{2}{r_n}\sum_{i=1}^{r_n}\left(\frac{f_{\boldsymbol{\theta}}(X_{n,i})}{g_n(X_{n,i})}\right)^{\frac{1}{2}} u_{\boldsymbol{\theta}}(X_{n,i}),$$

where $u_{\boldsymbol{\theta}}(x)=\frac{\nabla f_{\boldsymbol{\theta}}(x)}{f_{\boldsymbol{\theta}}(x)}$, while the Hessian is given by

$$\begin{aligned}\tilde{H}_n(\boldsymbol{\theta}) =&\frac{1}{r_n}\sum_{i=1}^{r_n}\left(\frac{f_{\boldsymbol{\theta}}(X_{n,i})}{g_n(X_{n,i})}\right)^{\frac{1}{2}} u_{\boldsymbol{\theta}}(X_{n,i})\cdot u_{\boldsymbol{\theta}}^T(X_{n,i})\\ &-\frac{2}{r_n}\sum_{i=1}^{r_n}\left(\frac{f_{\boldsymbol{\theta}}(X_{n,i})}{g_n(X_{n,i})}\right)^{\frac{1}{2}}\frac{1}{f_{\boldsymbol{\theta}}(X_{n,i})}H_f(X_{n,i}),\end{aligned}$$

where H_f is Hessian of $f_{\boldsymbol{\theta}}(\cdot)$. For the normal distribution, the gradient is given by

$$\nabla f_{\boldsymbol{\theta}}(x) = f_{\boldsymbol{\theta}}(x)\cdot\begin{bmatrix}\frac{x-\mu}{\sigma^2}\\ \frac{(x-\mu)^2-\sigma^2}{\sigma^3}\end{bmatrix},\quad u_{\boldsymbol{\theta}}(x)=\begin{bmatrix}\frac{x-\mu}{\sigma^2}\\ \frac{(x-\mu)^2-\sigma^2}{\sigma^3}\end{bmatrix},$$

and the Hessian of $f_{\boldsymbol{\theta}}(\cdot)$ is given by

$$H_f(x) = f_{\boldsymbol{\theta}}(x)\cdot\begin{bmatrix}\frac{(x-\mu)^2-\sigma^2}{\sigma^4} & \frac{(x-\mu)((x-\mu)^2-3\sigma^2)}{\sigma^5}\\ \frac{(x-\mu)((x-\mu)^2-3\sigma^2)}{\sigma^5} & \frac{(x-\mu)^4-5\sigma^2(x-\mu)^2+2\sigma^4}{\sigma^6}\end{bmatrix}.$$

E.1 Supplemental pictures and tables for PNR

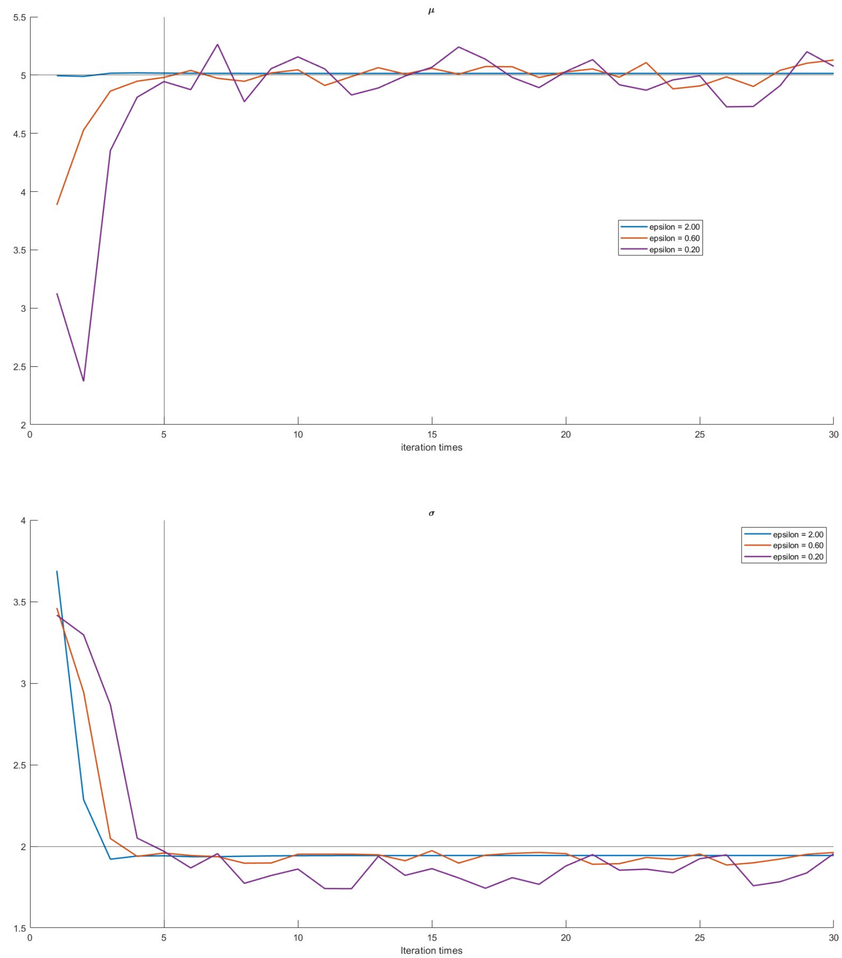

Figure E.1: Private Newton's method path

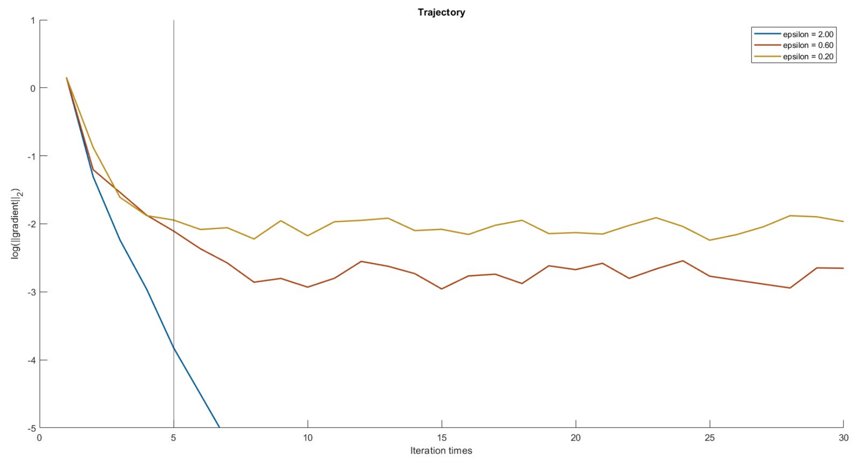

Figure E.2: Private Newton's method trajectory

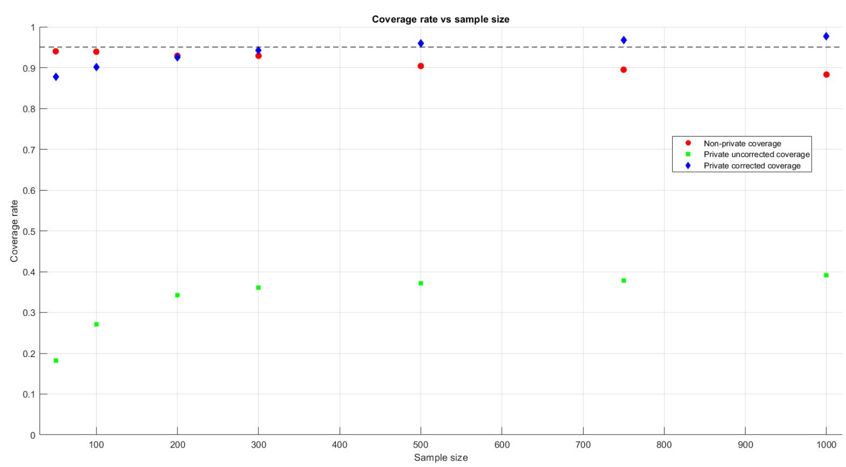

Figure E.3: Private and non-private Newton's method 95% confidence interval coverage

		ϵ		
		2.00	0.60	0.20
Estimator	μ: Mean (Std. Error)	5 (0.08)	4.948 (0.332)	4.868 (1.756)
	σ: Mean (Std. Error)	1.975 (0.076)	1.987 (0.349)	2.196 (1.99)
CI coverage for μ	Corrected	0.883	0.977	0.95
	Uncorrected	0.883	0.391	0.247
CI coverage for σ	Corrected	0.739	0.913	0.904
	Uncorrected	0.739	0.442	0.264

Table E.1: Results for different values of ϵ (Newton-Raphson). Sample size is 1000, $K = 5$.

		$\lambda = 1$, ϵ		
		Non-private	1.20	0.40
Estimator	μ: Mean (Std. Error)	4.991 (0.083)	4.986 (0.295)	4.96 (2.009)
	σ: Mean (Std. Error)	1.984 (0.058)	2.023 (0.21)	2.038 (1.809)
CI coverage for μ	Corrected	0.861	0.823	0.817
	Uncorrected	0.861	0.371	0.292
CI coverage for σ	Corrected	0.819	0.93	0.913
	Uncorrected	0.819	0.332	0.266

Table E.2: Results for different values of ϵ (Gradient descent). Sample size is 1000, $K = 50$.

		$\lambda = 1$, ϵ		
		Non-private	1.20	0.40
Estimator	μ: Mean (Std. Error)	5 (0.08)	4.927 (0.727)	4.76 (5.962)
	σ: Mean (Std. Error)	1.975 (0.076)	2.107 (1.265)	2.513 (7.652)
CI coverage for μ	Corrected	0.883	0.965	0.934
	Uncorrected	0.883	0.288	0.208
CI coverage for σ	Corrected	0.739	0.918	0.894
	Uncorrected	0.739	0.309	0.235

Table E.3: Results for different values of epsilon (Newton-Raphson). Sample size is 1000, $K = 5$.

	Contamination percentage α				
	0%	5%	10%	20%	30%
MLE (Std. Error)	5.001 (0.063)	5.241 (0.061)	5.476 (0.059)	5.952 (0.056)	6.422 (0.052)
PMHDE $\epsilon = 2$ (Std. Error)	5 (0.08)	5.174 (0.085)	5.309 (0.087)	5.555 (0.091)	5.778 (0.096)
PMHDE $\epsilon = 0.6$ (Std. Error)	4.952 (0.326)	5.119 (0.333)	5.252 (0.341)	5.472 (0.38)	5.646 (0.42)
PMHDE $\epsilon = 0.2$ (Std. Error)	4.942 (13.391)	4.905 (5.895)	5.054 (2.78)	5.349 (4.023)	5.169 (15.035)

Table E.4: Contamination results, Newton-Raphson, sample size is 1000, $\mu = 5$

E.2 Estimation and coverage rate for $(\lambda, \epsilon)-$PDP

In this section, we provide PMHDE and coverage rates for PGD and PNR algorithms for different values of λ.

		$\lambda = -0.1$, ϵ		
		Non-private	1.20	0.40
Estimator	μ: Mean (Std. Error)	4.991 (0.083)	4.992 (0.212)	4.979 (0.454)
	σ: Mean (Std. Error)	1.984 (0.058)	2.001 (0.152)	2.045 (0.291)
CI coverage for μ	Corrected	0.861	0.836	0.82
	Uncorrected	0.861	0.468	0.33
CI coverage for σ	Corrected	0.819	0.931	0.927
	Uncorrected	0.819	0.428	0.296

Table E.5: Results for different values of ϵ (Gradient descent). Sample size is 1000, $K = 50$.

		$\lambda = -0.1$, ϵ		
		Non-private	1.20	0.40
Estimator	μ: Mean (Std. Error)	5 (0.08)	4.955 (0.348)	4.827 (3.731)
	σ: Mean (Std. Error)	1.975 (0.076)	1.992 (0.353)	2.223 (1.554)
CI coverage for μ	Corrected	0.883	0.977	0.948
	Uncorrected	0.883	0.382	0.234
CI coverage for σ	Corrected	0.739	0.917	0.902
	Uncorrected	0.739	0.41	0.264

Table E.6: Results for different values of ϵ (Newton). Sample size is 1000, $K = 5$.

		$\lambda = 0.5$, ϵ		
		Non-private	1.20	0.40
Estimator	μ: Mean (Std. Error)	4.991 (0.083)	4.991 (0.256)	4.976 (0.549)
	σ: Mean (Std. Error)	1.984 (0.058)	2.013 (0.18)	2.056 (0.322)
CI coverage for μ	Corrected	0.861	0.826	0.817
	Uncorrected	0.861	0.396	0.306
CI coverage for σ	Corrected	0.819	0.931	0.922
	Uncorrected	0.819	0.379	0.277

Table E.7: Results for different values of ϵ (Gradient descent). Sample size is 1000, $K = 50$.

		$\lambda = 0.5$, ϵ		
		Non-private	1.20	0.40
Estimator	μ: Mean (Std. Error)	5 (0.08)	4.942 (0.483)	4.808 (2.064)
	σ: Mean (Std. Error)	1.975 (0.076)	2.03 (0.523)	2.297 (2.313)
CI coverage for μ	Corrected	0.883	0.972	0.94
	Uncorrected	0.883	0.322	0.222
CI coverage for σ	Corrected	0.739	0.915	0.899
	Uncorrected	0.739	0.355	0.254

Table E.8: Results for different values of ϵ (Newton). Sample size is 1000, $K = 5$.

E.3 Additional results for HDP and robustness evaluation

In this section, we provide additional results for PMHDE for sample sizes 200, 300, and 500 for both PGD and PNR algorithms. As explained in Section 4, when n and ϵ are both small, the algorithms can produce aberrant values, reducing their usefulness. For this reason, we use only estimates within the lower 0.7% and upper 99.5% percentiles of a Gaussian distribution with non-private $\hat{\mu}_n$ and $\hat{\sigma}_n$. All the Tables in this section are based on such a thresholding strategy. Since the confidence intervals are unaffected by thresholding, we retain all the simulation experiments for constructing the confidence intervals.

Tables E.9 and E.10 provide the estimators and the coverage rates for sample size 200, while Tables E.11 and E.12 provide the behavior of PMHDE under contamination for the sample size 200. The corresponding Tables for sample size 300 and 500 are given in Tables E.13, E.14, E.15, E.16, E.17, E.18, E.19, E.20 respectively.

Sample size 200:

		ϵ		
		2.00	0.60	0.20
Estimator	μ: Mean (Std. Error)	4.992 (0.153)	4.978 (0.538)	4.844 (1.22)
	σ: Mean (Std. Error)	1.952 (0.104)	2.036 (0.588)	1.744 (6.164)
CI coverage for μ	Corrected	0.921	0.839	0.626
	Uncorrected	0.921	0.51	0.27
CI coverage for σ	Corrected	0.846	0.921	0.6
	Uncorrected	0.846	0.484	0.216

Table E.9: Results for different values of ϵ (Gradient descent). Sample size is 200, $K = 50$.

		ϵ		
		2.00	0.60	0.20
Estimator	μ: Mean (Std. Error)	4.993 (0.148)	4.819 (1.242)	4.703 (1.972)
	σ: Mean (Std. Error)	1.93 (0.139)	2.465 (2.452)	3.118 (3.792)
CI coverage for μ	Corrected	0.929	0.925	0.888
	Uncorrected	0.929	0.343	0.157
CI coverage for σ	Corrected	0.771	0.879	0.829
	Uncorrected	0.771	0.359	0.169

Table E.10: Results for different values of ϵ (Newton). Sample size is 200, $K = 5$.

	Contamination percentage α				
	0%	5%	10%	20%	30%
MLE (Std. Error)	5 (0.142)	5.238 (0.138)	5.483 (0.134)	5.948 (0.127)	6.418 (0.119)
PMHDE $\epsilon = 2$ (Std. Error)	4.99 (0.153)	5.169 (0.155)	5.301 (0.158)	5.562 (0.165)	5.783 (0.172)
PMHDE $\epsilon = 0.6$ (Std. Error)	4.977 (0.521)	5.141 (0.534)	5.283 (0.54)	5.508 (0.548)	5.721 (0.576)
PMHDE $\epsilon = 0.2$ (Std. Error)	4.856 (1.228)	5.024 (1.16)	5.13 (1.22)	5.313 (1.215)	5.463 (1.306)

Table E.11: Contamination results, gradient descent, sample size is 200, $\mu = 5$.

	Contamination percentage α				
	0%	5%	10%	20%	30%
MLE (Std. Error)	5 (0.142)	5.238 (0.138)	5.483 (0.134)	5.948 (0.127)	6.418 (0.119)
PMHDE $\epsilon = 2$ (Std. Error)	4.991 (0.148)	5.174 (0.15)	5.308 (0.151)	5.578 (0.155)	5.823 (0.156)
PMHDE $\epsilon = 0.6$ (Std. Error)	4.809 (1.275)	4.974 (1.221)	5.021 (1.216)	5.17 (1.308)	5.304 (1.375)
PMHDE $\epsilon = 0.2$ (Std. Error)	4.744 (1.964)	4.815 (2.071)	4.866 (2.116)	4.954 (2.21)	5.002 (2.206)

Table E.12: Contamination results, Newton, sample size is 200, $\mu = 5$.

Sample size 300:

		ϵ		
		2.00	0.60	0.20
Estimator	μ: Mean (Std. Error)	4.989 (0.128)	4.979 (0.405)	4.926 (0.866)
	σ: Mean (Std. Error)	1.962 (0.086)	2.023 (0.338)	1.952 (1.971)
CI coverage for μ	Corrected	0.918	0.837	0.733
	Uncorrected	0.918	0.493	0.317
CI coverage for σ	Corrected	0.85	0.944	0.763
	Uncorrected	0.85	0.477	0.29

Table E.13: Results for different values of ϵ (Gradient descent). Sample size is 300, $K = 50$.

		ϵ		
		2.00	0.60	0.20
Estimator	μ: Mean (Std. Error)	4.993 (0.123)	4.863 (0.934)	4.697 (1.722)
	σ: Mean (Std. Error)	1.946 (0.113)	2.233 (1.442)	2.785 (6.688)
CI coverage for μ	Corrected	0.929	0.943	0.897
	Uncorrected	0.929	0.361	0.182
CI coverage for σ	Corrected	0.777	0.9	0.836
	Uncorrected	0.777	0.377	0.204

Table E.14: Results for different values of ϵ (Newton). The Sample size is 300, $K = 5$.

	Contamination percentage α				
	0%	5%	10%	20%	30%
MLE (std.error)	4.999 (0.115)	5.238 (0.113)	5.474 (0.11)	5.938 (0.103)	6.433 (0.096)
PMHDE $\epsilon = 2$ (Std. Error)	4.989 (0.128)	5.164 (0.132)	5.309 (0.135)	5.542 (0.14)	5.756 (0.149)
PMHDE $\epsilon = 0.6$ (Std. Error)	4.992 (0.408)	5.161 (0.398)	5.291 (0.408)	5.514 (0.416)	5.717 (0.434)
PMHDE $\epsilon = 0.2$ (Std. Error)	4.925 (0.858)	5.095 (0.868)	5.241 (0.844)	5.396 (0.905)	5.562 (0.932)

Table E.15: Contamination results, gradient descent, sample size is 300, $\mu = 5$.

	Contamination percentage α				
	0%	5%	10%	20%	30%
MLE (Std. Error)	4.999 (0.115)	5.238 (0.113)	5.474 (0.11)	5.938 (0.103)	6.433 (0.096)
PMHDE $\epsilon = 2$ (Std. Error)	4.993 (0.124)	5.173 (0.127)	5.32 (0.129)	5.562 (0.132)	5.806 (0.135)
PMHDE $\epsilon = 0.6$ (Std. Error)	4.856 (0.947)	5.009 (0.922)	5.114 (0.882)	5.296 (0.991)	5.42 (1.055)
PMHDE $\epsilon = 0.2$ (Std. Error)	4.724 (1.724)	4.863 (1.729)	4.904 (1.782)	5.023 (1.849)	5.063 (1.894)

Table E.16: Contamination results, Newton, sample size is 300, $\mu = 5$.

Sample size 500:

		ϵ		
		2.00	0.60	0.20
Estimator	μ: Mean (Std. Error)	4.989 (0.106)	4.986 (0.296)	4.951 (0.559)
	σ: Mean (Std. Error)	1.973 (0.072)	2.016 (0.213)	2.058 (0.799)
CI coverage for μ	Corrected	0.892	0.842	0.798
	Uncorrected	0.892	0.489	0.326
CI coverage for σ	Corrected	0.848	0.941	0.888
	Uncorrected	0.848	0.44	0.301

Table E.17: Results for different values of ϵ (gradient descent). Sample size is 500, $K = 50$.

		ϵ		
		2.00	0.60	0.20
Estimator	μ: Mean (Std. Error)	4.995 (0.102)	4.929 (0.631)	4.785 (1.322)
	σ: Mean (Std. Error)	1.959 (0.094)	2.089 (1.514)	2.554 (2.009)
CI coverage for μ	Corrected	0.904	0.96	0.916
	Uncorrected	0.904	0.371	0.213
CI coverage for σ	Corrected	0.767	0.908	0.874
	Uncorrected	0.767	0.391	0.243

Table E.18: Results for different values of ϵ. Sample size is 500 (Newton), $K = 5$.

	Contamination percentage α				
	0%	5%	10%	20%	30%
MLE (Std. Error)	4.998 (0.09)	5.24 (0.087)	5.473 (0.086)	5.946 (0.08)	6.424 (0.076)
PMHDE $\epsilon = 2$ (Std. Error)	4.989 (0.107)	5.157 (0.109)	5.293 (0.115)	5.526 (0.119)	5.738 (0.126)
PMHDE $\epsilon = 0.6$ (Std. Error)	4.986 (0.293)	5.157 (0.294)	5.291 (0.299)	5.514 (0.302)	5.719 (0.312)
PMHDE $\epsilon = 0.2$ (Std. Error)	4.987 (0.556)	5.146 (0.553)	5.256 (0.571)	5.464 (0.583)	5.65 (0.636)

Table E.19: Contamination results, gradient descent, sample size is 500, $\mu = 5$.

	Contamination percentage α				
	0%	5%	10%	20%	30%
MLE (Std. Error)	4.998 (0.09)	5.24 (0.087)	5.473 (0.086)	5.946 (0.08)	6.424 (0.076)
PMHDE $\epsilon = 2$ (Std. Error)	4.995 (0.103)	5.169 (0.105)	5.309 (0.11)	5.553 (0.113)	5.793 (0.117)
PMHDE $\epsilon = 0.6$ (Std. Error)	4.915 (0.631)	5.069 (0.606)	5.2 (0.629)	5.383 (0.659)	5.546 (0.716)
PMHDE $\epsilon = 0.2$ (Std. Error)	4.777 (1.316)	4.923 (1.306)	4.995 (1.384)	5.151 (1.386)	5.259 (1.419)

Table E.20: Contamination results, Newton, sample size is 500, $\mu = 5$.

References

Marco Avella-Medina, Casey Bradshaw, and Po-Ling Loh. Differentially private inference via noisy optimization. *The Annals of Statistics*, 51(5):2067–2092, 2023.

Rudolf Beran. Minimum hellinger distance estimates for parametric models. *The annals of Statistics*, pages 445–463, 1977.

An-lin Cheng and Anand N Vidyashankar. Minimum hellinger distance estimation for randomized play the winner design. *Journal of statistical planning and inference*, 136(6): 1875–1910, 2006.

Jinshuo Dong, Aaron Roth, and Weijie J Su. Gaussian differential privacy. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(1):3–37, 2022.

Elena Ghazi and Ibrahim Issa. Total variation meets differential privacy. *IEEE Journal on Selected Areas in Information Theory*, 2024.

Philippe Rigollet and Jan-Christian Hütter. High-dimensional statistics. *arXiv preprint arXiv:2310.19244*, 2023.

Joel A Tropp et al. An introduction to matrix concentration inequalities. *Foundations and Trends® in Machine Learning*, 8(1-2):1–230, 2015.