Noname manuscript No.
(will be inserted by the editor)

Supplementary Material for "Robust Statistical Inference Based on the C-Divergence Family"

Avijit Maji · Abhik Ghosh ·
Ayanendranath Basu · Leandro Pardo

S1 Proofs

S1.1 Proof of Theorem 3

We will first state and prove one lemma before going into the proof of the main theorem.

Lemma 1. *If $N'_\alpha(\infty)$ is bounded, then we have*

$$C(\epsilon g, 0) = \epsilon^{1+\alpha} N'_\alpha(\infty) \int g^{1+\alpha},$$

which is bounded and is also increasing in ϵ for all $\alpha \geq 0$.

Proof. Note that $C(\epsilon g, 0) = \int c(\epsilon g, 0)$. But

$$\begin{aligned} c(\epsilon g, 0) &= \lim_{f \to 0} c(\epsilon g, f) = \lim_{f \to 0} \left\{ \left[N\left(\frac{\epsilon g}{f} - 1 \right) \right] f^{1+\alpha} \right\} \\ &= \lim_{f \to 0} \left\{ \frac{\left[N\left(\frac{\epsilon g}{f} - 1 \right) \right]}{\left(\frac{\epsilon g}{f} \right)^{1+\alpha}} (\epsilon g)^{1+\alpha} \right\} = N'_\alpha(\infty)(\epsilon g)^{1+\alpha}. \end{aligned}$$

The rest of it is straightforward given that $N'_\alpha(\infty)$ is bounded and $\alpha \geq 0$. □

Now we will prove the main theorem.

Avijit Maji · Abhik Ghosh · Ayanendranath Basu
Indian Statistical Institute, 203, B.T. Road, Kolkata-700108, India.
Tel.: +91 33 2575 2806, Fax: +91 33 2577 3104.
Leandro Pardo
Department of Statistics and O.R., Complutense University of Madrid, 28040 Madrid, Spain.
E-mail: avijit.maji@hotmail.com, abhianik@gmail.com, ayanbasu@isical.ac.in, lpardo@mat.ucm.es.

Proof of Theorem 3. Let θ_n denote the minimizer of $C(h_{\epsilon,n}, f_\theta)$ where ϵ denotes the level of contamination. That is, $\theta_n = T(H_{\epsilon,n})$, the minimum C-divergence functional at $H_{\epsilon,n}$. If breakdown occurs, there exists a sequence $\{v_n\}$ such that $|\theta_n| \to \infty$ as $n \to \infty$. Now, consider

$$C(h_{\epsilon,n}, f_{\theta_n}) = \int_{L_n} c(h_{\epsilon,n}, f_{\theta_n}) + \int_{L_n^c} c(h_{\epsilon,n}, f_{\theta_n}), \tag{S1}$$

where $L_n = \{x : g(x) > \max(v_n(x), f_{\theta_n}(x))\}$ and $c(g, f)$ is the integrand of $C(g, f)$.

Now from Assumption (BP1), $\int_{L_n} v_n(x) \to 0$, and from Assumption (BP2), $\int_{L_n} f_{\theta_n}(x) \to 0$, so under $v_n(\cdot)$ and $f_{\theta_n}(\cdot)$, the set L_n converges to a set of zero probability as $n \to \infty$. Thus, on L_n,

$$c(h_{\epsilon,n}, f_{\theta_n}) \to c((1-\epsilon)g, 0) \quad \text{as} \quad n \to \infty.$$

and so by using Dominated Convergence Theorem (DCT)

$$\left| \int_{L_n} c(h_{\epsilon,n}, f_{\theta_n}) - \int_{L_n} c((1-\epsilon)g, 0) \right| \to 0, \tag{S2}$$

and further by Assumption (BP1) we have

$$\left| \int_{L_n} c((1-\epsilon)g, 0) - \int_{g>0} c((1-\epsilon)g, 0) \right| \to 0. \tag{S3}$$

Thus by Equations (S2) and (S3) we have

$$\left| \int_{L_n} c(h_{\epsilon,n}, f_{\theta_n}) - \int_{g>0} c((1-\epsilon)g, 0) \right| \to 0. \tag{S4}$$

So, by Lemma 1 and Assumption (C), we get

$$\int_{L_n} c(h_{\epsilon,n}, f_{\theta_n}) \to C((1-\epsilon)g, 0). \tag{S5}$$

Next by Assumption (BP1), $\int_{L_n^c} g(x) \to 0$ as $n \to \infty$, so under $g(\cdot)$, the set L_n^c converges to a set of zero probability. Hence similarly, we get

$$\left| \int_{L_n^c} c(h_{\epsilon,n}, f_{\theta_n}) - \int_{L_n^c} c(\epsilon v_n, f_{\theta_n}) \right| \to 0. \tag{S6}$$

Now by Assumption (BP3) we have,

$$\int c(\epsilon v_n, f_{\theta_n}) \geq \int c(\epsilon f_{\theta_n}, f_{\theta_n}) = N(\epsilon - 1) M_f^\alpha.$$

Using Equations (S5) and (S6), we get, when breakdown occurs,

$$\liminf_{n\to\infty} C(h_{\epsilon,n}, f_{\theta_n}) \geq N(\epsilon - 1) M_f^\alpha + C((1-\epsilon)g, 0) = a_1(\epsilon). \tag{S7}$$

We will have a contradiction to our assumption that $\{v_n\}$ is a sequence for which breakdown occurs if we can show that there exists a constant value θ^* in the parameter space such that for the same sequence $\{v_n\}$,

$$\limsup_{n\to\infty} C(h_{\epsilon,n}, f_{\theta_n}) < a_1(\epsilon) \tag{S8}$$

as then the $\{\theta_n\}$ sequence above could not minimize $C(h_{\epsilon,n}, f_{\theta_n})$ for every n.

We will now show that Equation (S8) is true for all $\epsilon < 1/2$ under the model when we choose θ^* to be the minimizer of $\int c((1-\epsilon)g, f_\theta)$. For any fixed θ, let $B_n = \{x : v_n(x) > max(g(x), f_\theta(x))\}$. From Assumption (BP1), $\int_{B_n} g(x) \to 0$, $\int_{B_n} f_\theta(x) \to 0$ and $\int_{B_n^c} v_n \to 0$. Thus, under v_n, the set B_n^c converges to a set of zero probability, while under g and f_θ, the set B_n converges to a set of zero probability. Thus on B_n, $c(h_{\epsilon,n}, f_\theta) \to c(\epsilon v_n, 0)$ as $n \to \infty$. So, by using DCT

$$\left| \int_{B_n} c(h_{\epsilon,n}, f_\theta) - \int_{v_n>0} c(\epsilon v_n, 0) \right| \to 0.$$

Similarly, we have $\left| \int_{B_n^c} c(h_{\epsilon,n}, f_\theta) - \int c((1-\epsilon)g, f_\theta) \right| \to 0$. Therefore, by Lemma 1 and Assumption (BP3),

$$\limsup_{n\to\infty} C(h_{\epsilon,n}, f_\theta) \le C((1-\epsilon)g, f_\theta) + C(\epsilon g, 0).$$

Now if $g = f_{\theta^g}$ then substituting $\theta = \theta^g$ in the above equation, and using the fact that $C((1-\epsilon)f_{\theta^g}, f_{\theta^g}) = N(-\epsilon)M_f^\alpha$, we get that $\theta^* = \theta^g$ satisfies

$$\lim_{n\to\infty} C(h_{\epsilon,n}, f_{\theta^*}) = N(-\epsilon)M_f^\alpha + C(\epsilon g, 0) = a_3(\epsilon), \qquad \text{say.}$$

Consequently, asymptotically there is no breakdown for ϵ level contamination when $a_3(\epsilon) < a_1(\epsilon)$. Notice that $a_1(\epsilon)$ and $a_3(\epsilon)$ are strictly decreasing and increasing respectively in ϵ by Lemma 1 and $a_1(1/2) = a_3(1/2)$. Hence asymptotically there is no breakdown and $\limsup_{n\to\infty} |T(H_{\epsilon,n})| < \infty$ for $\epsilon < 1/2$. □

S1.2 Proof of Theorem 4

To prove the consistency and asymptotic normality of the minimum C-divergence estimator, we will, from now on, assume that Conditions (A1)–(A7), presented in Section 4 of the main paper, hold. We will first prove some preliminary lemmas. Define $\eta_n(x) = \sqrt{n}(\sqrt{\delta_n} - \sqrt{\delta_g})^2$.

Lemma 2. *For any $k \in [0, 2]$, we have*

1. $E[\eta_n^k(x)] \le n^{\frac{k}{2}} E[|\delta_n(X) - \delta_g(X)|]^k \le \left[\frac{g(x)(1-g(x))}{f_\theta^2(x)}\right]^{\frac{k}{2}}$.
2. $E[|\delta_n(X) - \delta_g(X)|] \le \frac{2g(x)(1-g(x))}{f_\theta(x)}$.

Proof. For $a, b \geq 0$, we have the inequality $(\sqrt{a}-\sqrt{b})^2 \leq |a-b|$. So we get

$$E[\eta_n^k(x)] = n^{\frac{k}{2}} E[(\sqrt{\delta_n}-\sqrt{\delta_g})^2]^k \leq n^{\frac{k}{2}} E[|\delta_n - \delta_g|]^k.$$

For the next part, see that, $nd_n(x) \sim \text{binomial}(n, g(x))$ for all x. Now, for any $k \in [0,2]$, we get by the Lyapounov's inequality that

$$\begin{aligned} E[|\delta_n(X)-\delta_g(X)|]^k &\leq \left[E(\delta_n(X)-\delta_g(X))^2\right]^{\frac{k}{2}} \\ &= \frac{1}{f_\theta^k(x)}\left[E(d_n(X)-g(X))^2\right]^{\frac{k}{2}} = \frac{1}{f_\theta^k(x)}\left[\frac{g(x)(1-g(x))}{n}\right]^{\frac{k}{2}}. \end{aligned}$$

For the second part, note that

$$E[|\delta_n(X)-\delta_g(X)|] = \frac{1}{f_\theta(x)}\left[E|d_n(X)-g(X)|\right] \leq \frac{2g(x)(1-g(x))}{f_\theta(x)},$$

where the last inequality follows from the result about the mean-deviation of a binomial random variable. □

Lemma 3. *$E[\eta_n^k(x)] \to 0$, as $n \to \infty$, for $k \in [0,2)$.*

Proof. This follows from Theorem 4.5.2 of Chung (1974) by noting that $n^{1/4}(d_n^{1/2}(x)-g^{1/2}(x)) \to 0$ with probability one for each $x \in \mathcal{X}$ and by Lemma 2, $\sup_n E[\eta_n^k(x)]$ is bounded. □

Let us now define, $a_n(x) = K(\delta_n(x)) - K(\delta_g(x))$ and $b_n(x) = (\delta_n(x) - \delta_g(x))K'(\delta_g(x))$. We will need the limiting distributions of

$$S_{1n} = \sqrt{n}\sum_x a_n(x) f_\theta^{1+\alpha}(x) u_\theta(x) \text{ and } S_{2n} = \sqrt{n}\sum_x b_n(x) f_\theta^{1+\alpha}(x) u_\theta(x).$$

Define $\tau_n(x) = \sqrt{n}|a_n(x) - b_n(x)|$.

Lemma 4. *Suppose Assumption (A5) holds. Then $E|S_{1n}-S_{2n}| \to 0$ as $n \to \infty$, and hence $S_{1n}-S_{2n} \xrightarrow{P} 0$ as $n \to \infty$.*

Proof. By Lemma 2.15 of Basu et al. (2011) [or, Lindsay (1994), Lemma 25], there exists some positive constant β such that

$$\tau_n(x) \leq \beta\sqrt{n}(\sqrt{\delta}_n - \sqrt{\delta}_g)^2 = \beta\eta_n(x).$$

Also, by Lemma 2, $E[\tau_n(x)] \leq \beta\frac{g^{1/2}(x)}{f_\theta(x)}$. And by Lemma 2, $E[\tau_n(x)] = \beta E[\eta_n(x)] \to 0$ as $n \to \infty$. Thus we get,

$$\begin{aligned} E|S_{1n}-S_{2n}| &\leq \sum_x E[\tau_n(x)] f_\theta^{1+\alpha}(x)|u_\theta(x)| \\ &\leq \beta\sum_x g^{1/2}(x) f_\theta^\alpha(x)|u_\theta(x)| < \infty \qquad \text{(by Assumption A5)}. \end{aligned}$$

So, by using DCT, $E|S_{1n}-S_{2n}| \to 0$ as $n \to \infty$. Hence, by Markov's inequality, $S_{1n}-S_{2n} \xrightarrow{P} 0$ as $n \to \infty$. □

Lemma 5. *Suppose* $V_g = V_g\left[K'(\delta_g(X))f_\theta^\alpha(X)u_\theta(X)\right]$ *is finite. Then*

$$S_{1n} \to N(0, V_g).$$

Proof. Note that, by the Lemma 4, the asymptotic distribution of S_{1n} and S_{2n} are the same. Now, we have

$$\begin{aligned} S_{2n} &= \sqrt{n}\sum_x(\delta_n(x) - \delta_g(x))K'(\delta_g(x))f_\theta^{1+\alpha}(x)u_\theta(x) \\ &= \sqrt{n}\sum_x(d_n(x) - g(x))K'(\delta_g(x))f_\theta^{\alpha}(x)u_\theta(x) \\ &= \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n\left[K'(\delta_g(X_i))f_\theta^\alpha(X_i)u_\theta(X_i) - E_g\{K'(\delta_g(X))f_\theta^\alpha(X)u_\theta(X)\}\right]\right) \\ &\to N(0, V_g), \end{aligned}$$

where the last relation follows by the central limit theorem. □

Proof of Theorem 4.
Consistency: Consider the behavior of $C(d_n, f_\theta)$ on a sphere Q_a which has radius a and center at θ^g. We will show, for sufficiently small a, that with probability tending to one

$$C(d_n, f_\theta) > C(d_n, f_{\theta^g}) \text{ for all } \theta \text{ on the surface of } Q_a,$$

so that the C-divergence has a local minimum with respect to θ in the interior of Q_a. At a local minimum, the estimating equations must be satisfied. Therefore, for any $a > 0$ sufficiently small, the minimum C-divergence estimating equations have a solution θ_n within Q_a with probability tending to one as $n \to \infty$. Now taking a Taylor series expansion of $C(d_n, f_\theta)$ about $\theta = \theta^g$, we get

$$\begin{aligned} &C(d_n, f_{\theta^g}) - C(d_n, f_\theta) \\ &= -\sum_j(\theta_j - \theta_j^g)\nabla_j C(d_n, f_\theta)|_{\theta=\theta^g} - \frac{1}{2}\sum_{j,k}(\theta_j - \theta_j^g)(\theta_k - \theta_k^g)\nabla_{jk}C(d_n, f_\theta)|_{\theta=\theta^g} \\ &-\frac{1}{6}\sum_{j,k,l}(\theta_j - \theta_j^g)(\theta_k - \theta_k^g)(\theta_l - \theta_l^g)\nabla_{jkl}C(d_n, f_\theta)|_{\theta=\theta^{**}} \\ &= S_1 + S_2 + S_3, \qquad (say) \end{aligned} \tag{S9}$$

where θ^{**} lies between θ^g and θ. We will now consider each term one-by-one. For the linear term S_1, we consider

$$\nabla_j C(d_n, f_\theta)|_{\theta=\theta^g} = -\sum_x K(\delta_n^g(x))f_{\theta^g}^{1+\alpha}(x)u_{j\theta^g}(x) \tag{S10}$$

where $\delta_n^g(x)$ is $\delta_n(x)$ evaluated at $\theta = \theta^g$. We will now show that

$$\sum_x K(\delta_n^g(x))f_{\theta^g}^{1+\alpha}(x)u_{j\theta^g}(x) \xrightarrow{P} \sum_x K(\delta_g^g(x))f_{\theta^g}^{1+\alpha}(x)u_{j\theta^g}(x), \tag{S11}$$

as $n \to \infty$ and note that the right hand side of above is zero by definition of the minimum C-divergence estimator. By Assumption (A7) and the fact that $d_n(x) \to g(x)$ almost surely by the strong law of large numbers (SLLN), it follows that

$$\begin{aligned}|K'(\delta)| &= |N''(\delta)(\delta+1) - \alpha N'(\delta)| \\ &\leq |N''(\delta)(\delta+1)| + \alpha|N'(\delta)| \leq C_1, \qquad (\text{ say }) \end{aligned} \tag{S12}$$

for any δ in between $\delta_n^g(x)$ and $\delta_g^g(x)$ (uniformly in x). So, by using the one-term Taylor series expansion,

$$\begin{aligned}&|\sum_x K(\delta_n^g(x)) f_{\theta^g}^{1+\alpha}(x) u_{j\theta^g}(x) - \sum_x K(\delta_g^g(x)) f_{\theta^g}^{1+\alpha}(x) u_{j\theta^g}(x)| \\ \leq\ & C_1 \sum_x |\delta_n^g(x) - \delta_g^g(x)| f_{\theta^g}^{1+\alpha}(x) |u_{j\theta^g}(x)|.\end{aligned}$$

However, by Lemma 2,

$$E\left[|\delta_n^g(x) - \delta_g^g(x)|\right] \leq \frac{(g(x)(1-g(x))^{1/2}}{f_{\theta^g}(x)\sqrt{n}} \to 0 \quad \text{as} \quad n \to \infty. \tag{S13}$$

and, by Lemma 2, we have

$$\begin{aligned}&E\left[C_1 \sum_x |\delta_n^g(x) - \delta_g^g(x)| f_{\theta^g}^{1+\alpha}(x) |u_{j\theta^g}(x)|\right] \\ \leq\ & 2C_1 \sum_x g^{1/2}(x) f_{\theta^g}^{\alpha}(x) |u_{j\theta^g}(x)| \ < \ \infty. \qquad (\text{by Assumption A5})\end{aligned} \tag{S14}$$

Hence, by DCT, we get ,

$$E\left|\sum_x K(\delta_n^g(x)) f_{\theta^g}^{1+\alpha}(x) u_{j\theta^g}(x) - \sum_x K(\delta_g^g(x)) f_{\theta^g}^{1+\alpha}(x) u_{j\theta^g}(x)\right| \to 0, \tag{S15}$$

as $n \to \infty$, so that by Markov's inequality we have the desired claim. Therefore, we have

$$\nabla_j C(d_n, f_\theta)|_{\theta=\theta^g} \xrightarrow{P} 0. \tag{S16}$$

Thus, with probability tending to one, $|S_1| < pa^3$, where p is the dimension of θ and a is the radius of Q_a.

Next we consider the quadratic term S_2. We have,

$$\begin{aligned}&\nabla_{jk} C(d_n, f_\theta)|_{\theta=\theta^g} = \nabla_k \left(- \sum_x K(\delta_n(x)) f_\theta^{1+\alpha}(x) u_{j\theta}(x)|_{\theta=\theta^g} \right) \\ =\ & \sum_x K'(\delta_n^g(x))(\delta_n^g(x)+1) f_{\theta^g}^{1+\alpha}(x) u_{j\theta^g}(x) u_{k\theta^g}(x) \\ & - \sum_x K(\delta_n^g(x)) f_{\theta^g}^{1+\alpha}(x) u_{jk\theta^g}(x) \\ & -(1+\alpha) \sum_x K(\delta_n^g(x)) f_{\theta^g}^{1+\alpha}(x) u_{j\theta^g}(x) u_{k\theta^g}(x).\end{aligned} \tag{S17}$$

We will now show that

$$\sum_x K'(\delta_n^g(x))(\delta_n^g(x)+1)f_{\theta^g}^{1+\alpha}(x)u_{j\theta^g}(x)u_{k\theta^g}(x)$$
$$\xrightarrow{P} \sum_x K'(\delta_g^g(x))(\delta_g^g(x)+1)f_{\theta^g}^{1+\alpha}(x)u_{j\theta^g}(x)u_{k\theta^g}(x). \tag{S18}$$

Note that as in Equation (S12), we have

$$|K''(\delta)(\delta+1)| \leq |N'''(\delta)(\delta+1)^2| + (1-\alpha)|N''(\delta)(\delta+1)| \leq C_2, \text{ say}, \tag{S19}$$

for every δ lying in between $\delta_n^g(x)$ and $\delta_g^g(x)$ (uniformly in x). So, by using the one-term Taylor series expansion,

$$\begin{aligned}&|K'(\delta_n^g(x))(\delta_n^g(x)+1) - K'(\delta_g^g(x))(\delta_g^g(x)+1)|\\ &\leq |\delta_n^g(x)-\delta_g^g(x)||K''(\delta_n^g(x))(\delta_n^g(x)+1)+K'(\delta_n^g(x))|\\ &\leq |\delta_n^g(x)-\delta_g^g(x)|(C_2+C_1).\end{aligned}$$

Thus, we get

$$\begin{aligned}&\left|\sum_x K'(\delta_n^g(x))(\delta_n^g(x)+1)f_{\theta^g}^{1+\alpha}(x)u_{j\theta^g}(x)u_{k\theta^g}(x)\right.\\ &\qquad\left.-\sum_x K'(\delta_g^g(x))(\delta_g^g(x)+1)f_{\theta^g}^{1+\alpha}(x)u_{j\theta^g}(x)u_{k\theta^g}(x)\right|\\ &\leq (C_1+C_2)\sum_x |\delta_n^g(x)-\delta_g^g(x)|f_{\theta^g}^{1+\alpha}(x)|u_{j\theta^g}(x)u_{k\theta^g}(x)|.\end{aligned}$$

Since by Assumption (A5), we have $\sum_x g^{1/2}(x)f_{\theta^g}^{1+\alpha}(x)|u_{j\theta^g}(x)u_{k\theta^g}(x)| < \infty$, the desired result of Equation (S18) follows by an approach similar to the proof of Equation (S11). Similarly, we also get that

$$\sum_x K(\delta_n^g(x))f_{\theta^g}^{1+\alpha}(x)u_{jk\theta^g}(x) \xrightarrow{P} \sum_x K(\delta_g^g(x))f_{\theta^g}^{1+\alpha}u_{jk\theta^g}(x),$$

and

$$\sum_x K(\delta_n^g(x))f_{\theta^g}^{1+\alpha}(x)u_{j\theta^g}(x)u_{k\theta^g}(x) \xrightarrow{P} \sum_x K(\delta_g^g(x))f_{\theta^g}^{1+\alpha}(x)u_{j\theta^g}(x)u_{k\theta^g}(x).$$

Thus, combining Equation (S18) with the above two, we get that

$$\nabla_k\left(\sum_x K(\delta_n(x))f_\theta^{1+\alpha}(x)u_{j\theta}(x)|_{\theta=\theta^g}\right) \xrightarrow{P} -J_g^{j,k}. \tag{S20}$$

But

$$\begin{aligned}2S_2 = &\sum_{j,k}\left\{\nabla_k\left(\sum_x K(\delta_n(x))f_\theta^{1+\alpha}(x)u_{j\theta}(x)|_{\theta=\theta^g}\right) - (-J_g^{j,k})\right\}(\theta_j-\theta_j^g)(\theta_k-\theta_k^g)\\ &+\sum_{j,k}\left\{-\left(J_g^{j,k}\right)(\theta_j-\theta_j^g)(\theta_k-\theta_k^g)\right\}.\end{aligned} \tag{S21}$$

Now the absolute value of the first term in the above Equation (S21) is $< p^2a^3$ with probability tending to one. And, the second term in Equation (S21) is a negative definite quadratic form in the variables $(\theta_j - \theta_j^g)$. Letting λ_1 be the largest eigenvalue of J_g, the quadratic form is $< \lambda_1 a^2$. Combining the two terms, we see that there exists $c > 0$ and $a_0 > 0$ such that for $a < a_0$, we have $S_2 < -ca^2$ with probability tending to one.

Finally, considering the cubic term S_3, we have

$$
\begin{aligned}
&-\nabla_{jkl}C(d_n, f_\theta)|_{\theta=\theta^{**}}\\
&= \sum_x K''(\delta_n^{**}(x))(\delta_n^{**}(x)+1)^2 f_{\theta^{**}}^{1+\alpha}(x)u_{j\theta^{**}}(x)u_{k\theta^{**}}(x)u_{l\theta^{**}}(x)\\
&+ \sum_x K'(\delta_n^{**}(x))(\delta_n^{**}(x)+1) f_{\theta^{**}}^{1+\alpha}(x)u_{j\theta^{**}}(x)u_{k\theta^{**}}(x)u_{l\theta^{**}}(x)\\
&- (1+\alpha)\sum_x K'(\delta_n^{**}(x))(\delta_n^{**}(x)+1) f_{\theta^{**}}^{1+\alpha}(x)u_{j\theta^{**}}(x)u_{k\theta^{**}}(x)u_{l\theta^{**}}(x)\\
&- \sum_x K'(\delta_n^{**}(x))(\delta_n^{**}(x)+1) f_{\theta^{**}}^{1+\alpha}(x)u_{j\theta^{**}}(x)u_{kl\theta^{**}}(x)\\
&- \sum_x K'(\delta_n^{**}(x))(\delta_n^{**}(x)+1) f_{\theta^{**}}^{1+\alpha}(x)u_{k\theta^{**}}(x)u_{jl\theta^{**}}(x)\\
&- \sum_x K'(\delta_n^{**}(x))(\delta_n^{**}(x)+1) f_{\theta^{**}}^{1+\alpha}(x)u_{l\theta^{**}}(x)u_{jk\theta^{**}}(x)\\
&+ (1+\alpha)\sum_x K(\delta_n^{**}(x)) f_{\theta^{**}}^{1+\alpha}(x)u_{l\theta^{**}}(x)u_{jk\theta^{**}}(x)\\
&+ \sum_x K(\delta_n^{**}(x)) f_{\theta^{**}}^{1+\alpha}(x)u_{jkl\theta^{**}}(x) - \sum_x K(\delta_n^{**}(x)) f_{\theta^{**}}^{1+\alpha}(x)u_{j\theta^{**}}(x)u_{kl\theta^{**}}(x)\\
&- (1+\alpha)\sum_x K'(\delta_n^{**}(x))(\delta_n^{**}(x)+1) f_{\theta^{**}}^{1+\alpha}(x)u_{j\theta^{**}}(x)u_{k\theta^{**}}(x)u_{l\theta^{**}}(x)\\
&+ (1+\alpha)^2\sum_x K(\delta_n^{**}(x)) f_{\theta^{**}}^{1+\alpha}(x)u_{j\theta^{**}}(x)u_{k\theta^{**}}(x)u_{l\theta^{**}}(x)\\
&- \sum_x K(\delta_n^{**}(x)) f_{\theta^{**}}^{1+\alpha}(x)u_{k\theta^{**}}(x)u_{jl\theta^{**}}(x), \qquad \text{(S22)}
\end{aligned}
$$

where $\delta_n^{**}(x) = \frac{d_n(x)}{f_{\theta^{**}}(x)} - 1$. We will now show that all the terms in the RHS of the above Equation (S22) are bounded. Let us denote the terms in Equation (S22), in order, by (i), (ii),..., (xii), respectively.

For the first term (i), we use Equation (S19) to get

$$
\begin{aligned}
&\left|\sum_x K''(\delta_n^{**}(x))(\delta_n^{**}(x)+1)^2 f_{\theta^{**}}^{1+\alpha}(x)u_{j\theta^{**}}(x)u_{k\theta^{**}}(x)u_{l\theta^{**}}(x)\right|\\
&\leq C_2 \sum_x |\delta_n^{**}(x)+1| M_{j,k,l}(x) f_{\theta^{**}}(x) = C_2 \sum_x d_n(x) M_{j,k,l}(x) \quad \text{(by CLT)}\\
&\to C_2\ E_g[M_{j,k,l}(X)] < \infty. \quad \text{(by Assumption (A6))} \qquad \text{(S23)}
\end{aligned}
$$

Thus term (i) is bounded. Now for the terms (ii), (iii), (ix) we again use Equation (S12) to get

$$\begin{aligned}
&\left|\sum_x K'(\delta_n^{**}(x))(\delta_n^{**}(x)+1) f_{\theta^{**}}^{1+\alpha}(x) u_{j\theta^{**}}(x) u_{k\theta^{**}}(x) u_{l\theta^{**}}(x)\right| \\
&\leq C_1 \sum_x |\delta_n^{**}(x)+1| M_{j,k,l}(x) f_{\theta^{**}}(x) = C_1 \sum_x d_n(x) M_{j,k,l}(x) \quad \text{(by CLT)} \\
&\to C_1 \ E_g[M_{j,k,l}(X)] < \infty. \quad \text{(by Assumption (A6))} \qquad \text{(S24)}
\end{aligned}$$

so that the terms (ii), (iii) and (ix) are also bounded. Similarly, the terms (iv), (v) and (vi) are bounded as in case of term (ii) and using Equation (S12) and Assumption (A6). Next for the terms (vii), (viii), (x), (xi) and (xii) we will consider the relation.

$$|K(\delta)| = |\int_0^\delta K'(\delta) d\delta| \leq C_1 |\delta| < C_1 |\delta + 1|. \qquad \text{(S25)}$$

Also, the terms (vii), (viii), (x), (xi) and (xii) are individually bounded by

$$\begin{aligned}
&C_1 \sum_x |\delta_n^{**}(x)+1| M(x) f_{\theta^{**}}(x) \ \text{(or some suitable multiple of } C_1) \\
&= C_1 \sum_x d_n(x) M(x) \to C_1 \ E_g[M(X)] < \infty, \qquad \text{(S26)}
\end{aligned}$$

by the CLT and Assumption (A6), where $M(x) = M_{jkl}(x) + M_{jk,l}(x) + M_{jl,k}(x) + M_{j,kl}(x) + M_{j,k,l}(x)$. Hence, we have $|S_3| < ba^3$ on the sphere Q_a with probability tending to one.

Combining the three inequalities we get that

$$\max(S_1 + S_2 + S_3) < -ca^2 + (b+p)a^3 \quad < 0 \quad \text{for} \quad \left[a < \frac{c}{b+p}\right].$$

Thus, for any sufficiently small a, there exists a sequence of roots $\theta_n = \theta_n(a)$ to the minimum C-divergence estimating equation such that $P(||\theta_n - \theta^g||_2 < a)$ converges to one, where $||.||_2$ denotes the L_2-norm.

It remains to show that we can determine such a sequence independent of a. For let θ_n^* be the root which is closest to θ^g. This exists because the limit of a sequence of roots is again a root by the continuity of the C-divergence. This completes the proof of the consistency part.

Asymptotic Normality: We expand $\sum_x K(\delta_n(x)) f_\theta^{1+\alpha}(x) u_\theta(x)$ in Taylor series about $\theta = \theta^g$ to get a θ' lying in-between θ and θ^g such that

$$\begin{aligned}
&\sum_x K(\delta_n(x)) f_\theta^{1+\alpha}(x) u_\theta(x) = \sum_x K(\delta_n^g(x)) f_{\theta^g}^{1+\alpha}(x) u_{\theta^g}(x) \\
&+ \sum_k (\theta_k - \theta_k^g) \nabla_k \left(\sum_x K(\delta_n(x)) f_\theta^{1+\alpha}(x) u_\theta(x) \right) \Bigg|_{\theta=\theta^g} \\
&+ \frac{1}{2} \sum_{k,l} (\theta_k - \theta_k^g)(\theta_l - \theta_l^g) \nabla_{kl} \left(\sum_x K(\delta_n(x)) f_\theta^{1+\alpha}(x) u_\theta(x) \right) \Bigg|_{\theta=\theta'}. \qquad \text{(S27)}
\end{aligned}$$

Now, let θ_n be the solution of the minimum C-divergence estimating equation, which can be assumed to be consistent by the previous part. Replace θ by θ_n in the above Equation (S27) so that the LHS of the equation becomes zero and hence we get

$$\begin{aligned}
&-\sqrt{n}\sum_x K(\delta_n^g(x))f_{\theta^g}^{1+\alpha}(x)u_{\theta^g}(x)\\
&=\sqrt{n}\sum_k(\theta_{nk}-\theta_k^g)\times\left\{\nabla_k\left(\sum_x K(\delta_n(x))f_\theta^{1+\alpha}(x)u_\theta(x)\right)|_{\theta=\theta^g}\right.\\
&\qquad\left.+\frac{1}{2}\sum_l(\theta_{nl}-\theta_l^g)\nabla_{kl}\left(\sum_x K(\delta_n(x))f_\theta^{1+\alpha}(x)u_\theta(x)\right)|_{\theta=\theta'}\right\}. \quad \text{(S28)}
\end{aligned}$$

Note that, the first term within the bracketed quantity in the RHS of the above Equation (S28) converges to J_g with probability tending to one, while the second bracketed term is an $o_p(1)$ term (as proved in the proof of consistency part). Also, by using the Lemma 5, we get that

$$\begin{aligned}
\sqrt{n}\sum_x K(\delta_n^g(x))f_{\theta^g}^{1+\alpha}(x)u_{\theta^g}(x) &= \sqrt{n}\sum_x\left[K(\delta_n^g(x))-K(\delta_g^g(x))\right]f_{\theta^g}^{1+\alpha}(x)u_{\theta^g}(x)\\
&= S_{1n}|_{\theta=\theta^g}\xrightarrow{\mathcal{D}}N_p(0,V_g). \quad \text{(S29)}
\end{aligned}$$

Therefore, by Lehmann (1983, Lemma 4.1), it follows that $\sqrt{n}(\theta_n-\theta^g)$ has an asymptotic $N_p(0,J_g^{-1}V_gJ_g^{-1})$ distribution. □

S1.3 Proof of Theorem 10

We consider the second order Taylor series expansion of $C_\gamma(f_\theta,f_{\theta_0})$ around $\theta=\theta_0$ at $\theta=\hat\theta_\alpha$ as,

$$\begin{aligned}
C_\gamma(f_{\hat\theta_\alpha},f_{\theta_0}) &= C_\gamma(f_{\theta_0},f_{\theta_0})+\sum_{i=1}^p\nabla_iC_\gamma(f_\theta,f_{\theta_0})|_{\theta=\theta_0}(\hat\theta_\alpha^i-\theta_0^i)\\
&\quad+\frac{1}{2}\sum_{i,j}\nabla_{ij}C_\gamma(f_\theta,f_{\theta_0})|_{\theta=\theta_0}(\hat\theta_\alpha^i-\theta_0^i)(\hat\theta_\alpha^j-\theta_0^j)+o(||\hat\theta_\alpha-\theta_0||^2),
\end{aligned}$$

where the superscripts denotes the corresponding components. Now we have $C_\gamma(f_{\theta_0},f_{\theta_0})=0$ and $\nabla_iC_\gamma(f_\theta,f_{\theta_0})|_{\theta=\theta_0}=0$. Note that the above second order partial derivative of $C_\gamma(f_\theta,f_{\theta_0})$ at $\theta=\theta_0$ is independent of λ and so we will denote them by $a_{ij}^\gamma(\theta_0)$. Then note that $A_\gamma(\theta_0)=\left(a_{ij}^\gamma(\theta_0)\right)_{i,j=1,\ldots,p}=\nabla^2C_\gamma(f_\theta,f_{\theta_0})\Big|_{\theta=\theta_0}$. Now from the above Taylor series expansion it is clear that $T_{C_\gamma}(\hat\theta_\alpha,\theta_0)=2nC_\gamma(f_{\hat\theta_\alpha},f_{\theta_0})$ and $\sqrt{n}(\hat\theta_\alpha-\theta_0)^TA_\gamma(\theta_0)\sqrt{n}(\hat\theta_\alpha-\theta_0)$ have the same asymptotic distribution. Now we know from Basu et al. (1998) that the asymptotic distribution of $\sqrt{n}(\hat\theta_\alpha-\theta_0)$ is normal with mean zero and variance $J_\alpha^{-1}(\theta_0)V_\alpha(\theta_0)J_\alpha^{-1}(\theta_0)$.

Further we know that for $X \sim N_q(0, \Sigma)$, and a $q-$dimensional real symmetric matrix A, the distribution of the quadratic form $X^T AX$ is the same as that of $\sum_{i=1}^{r} \zeta_i Z_i^2$, where $Z_1, \ldots, Z_r$ are independent standard normal variables, $r = rank(\Sigma A \Sigma)$, $r \geq 1$ and $\zeta_1, \ldots, \zeta_r$ are the nonzero eigenvalues of $A\Sigma$ (Dik and Gunst, 1985, Corollary 2.1). Applying this result with $X = \sqrt{n}(\hat{\theta}_\alpha - \theta_0)$ we get the theorem.

S1.4 Proof of Theorem 11

Fix some $\theta^* \neq \theta_0$. Consider the first order Taylor series expansion of $C_\gamma(f_{\hat{\theta}_\alpha}, f_{\theta_0})$ under f_{θ^*} as

$$C_\gamma(f_{\hat{\theta}_\alpha}, f_{\theta_0}) = C_\gamma(f_{\theta^*}, f_{\theta_0}) + M_{C_\gamma}(\theta^*)^T(\hat{\theta}_\alpha - \theta^*) + o(||\hat{\theta}_\alpha - \theta^*||)$$

where M_{C_γ} is as defined in the theorem. Now we know that, under θ^*,

$$\sqrt{n}(\hat{\theta}_\alpha - \theta^*) \to N(0, J_\alpha^{-1}(\theta^*)V_\alpha(\theta^*)J_\alpha^{-1}(\theta^*)) \quad \text{as } n \to \infty$$

and $\sqrt{n} \times o(||\hat{\theta}_\alpha - \theta^*||) = o_p(1)$. Thus we get that the random variables $\sqrt{n}\left[C_\gamma(f_{\hat{\theta}_\alpha}, f_{\theta_0}) - C_\gamma(f_{\theta^*}, f_{\theta_0})\right]$ and $M_{C_\gamma}(\theta^*)^T\sqrt{n}(\hat{\theta}_\alpha - \theta^*)$ have the same asymptotic distribution . Therefore, we have

$$\sqrt{n}\left[C_\gamma(f_{\hat{\theta}_\alpha}, f_{\theta_0}) - C_\gamma(f_{\theta^*}, f_{\theta_0})\right] \to N(0, \sigma(\theta^*))$$

where $\sigma(\theta^*)$ is as given in Equation (35) of the main text. Then the desired approximation to the power function follows from the above asymptotic distribution.

S1.5 Proof of Theorem 12

From the asymptotic distribution of the MDPDE, we have

$$\sqrt{n_i}\left({}^{(i)}\hat{\theta}_\alpha - \theta_i\right) \to N(0, J_\alpha^{-1}(\theta_i)V_\alpha(\theta_i)J_\alpha^{-1}(\theta_i))$$

for $i = 1, 2$, with $n_1 = n, n_2 = m$. Let $\frac{m}{m+n} \to \omega$ as $m, n \to \infty$. Then we have

$$\sqrt{\frac{mn}{m+n}}\left({}^{(1)}\hat{\theta}_\alpha - \theta_1\right) \to N(0, \omega J_\alpha^{-1}(\theta_1)V_\alpha(\theta_1)J_\alpha^{-1}(\theta_1))$$

and

$$\sqrt{\frac{mn}{m+n}}\left({}^{(2)}\hat{\theta}_\alpha - \theta_2\right) \to N(0, (1-\omega)J_\alpha^{-1}(\theta_2)V_\alpha(\theta_2).J_\alpha^{-1}(\theta_2)).$$

Now, under $H_0 : \theta_1 = \theta_2$, we get that

$$\sqrt{\frac{mn}{m+n}}\left({}^{(1)}\hat{\theta}_\alpha -{}^{(2)} \hat{\theta}_\alpha\right) \to N(0, J_\alpha^{-1}(\theta_1)V_\alpha(\theta_1)J_\alpha^{-1}(\theta_1)).$$

Next consider the second order Taylor series expansion of $C_\gamma(f_{\theta_1}, f_{\theta_2})$ around $\theta_1 = \theta_2$ at $\left({}^{(1)}\hat{\theta}_\alpha, {}^{(1)}\hat{\theta}_\alpha\right)$ as follows.

$$\begin{aligned}
C_\gamma(f_{{}^{(1)}\hat{\theta}_\alpha}, f_{{}^{(2)}\hat{\theta}_\alpha}) = & \frac{1}{2}\sum_{i,j=1}^{p}\left(\frac{\partial^2 C_\gamma(f_{\theta_1}, f_{\theta_2})}{\partial\theta_{1i}\partial\theta_{1j}}\right)_{\theta_1=\theta_2}(\hat{\theta}_\alpha^{1i}-\theta_{1i})(\hat{\theta}_\alpha^{1j}-\theta_{1j}) \\
& + \sum_{i,j=1}^{p}\left(\frac{\partial^2 C_\gamma(f_{\theta_1}, f_{\theta_2})}{\partial\theta_{1i}\partial\theta_{2j}}\right)_{\theta_1=\theta_2}(\hat{\theta}_\alpha^{1i}-\theta_{1i})(\hat{\theta}_\alpha^{2j}-\theta_{2j}) \\
& + \frac{1}{2}\sum_{i,j=1}^{p}\left(\frac{\partial^2 C_\gamma(f_{\theta_1}, f_{\theta_2})}{\partial\theta_{2i}\partial\theta_{2j}}\right)_{\theta_1=\theta_2}(\hat{\theta}_\alpha^{2i}-\theta_{2i})(\hat{\theta}_\alpha^{2j}-\theta_{2j}) \\
& + o\left(||{}^{(1)}\hat{\theta}_\alpha - \theta_1||^2\right) + o\left(||{}^{(2)}\hat{\theta}_\alpha - \theta_2||^2\right).
\end{aligned}$$

But for $i = 1, \ldots, p$, we have

$$\frac{\partial C_\gamma(f_{\theta_1}, f_{\theta_2})}{\partial\theta_{1i}} = \int N'\left(\frac{f_{\theta_1}}{f_{\theta_2}} - 1\right)\ f_{\theta_2}^\alpha\ \frac{\partial f_{\theta_1}}{\partial\theta_{1i}}.$$

We also have,

$$\begin{aligned}
\left(\frac{\partial^2 C_\gamma(f_{\theta_1}, f_{\theta_2})}{\partial\theta_{1i}\partial\theta_{1j}}\right)_{\theta_1=\theta_2} &= a_{ij}^\gamma(\theta_1), \\
\left(\frac{\partial^2 C_\gamma(f_{\theta_1}, f_{\theta_2})}{\partial\theta_{1i}\partial\theta_{2j}}\right)_{\theta_1=\theta_2} &= -\left(\frac{\partial^2 C_\gamma(f_{\theta_1}, f_{\theta_2})}{\partial\theta_{1i}\partial\theta_{1j}}\right)_{\theta_1=\theta_2} = -a_{ij}^\gamma(\theta_1), \\
\left(\frac{\partial^2 C_\gamma(f_{\theta_1}, f_{\theta_2})}{\partial\theta_{2i}\partial\theta_{2j}}\right)_{\theta_1=\theta_2} &= \left(\frac{\partial^2 C_\gamma(f_{\theta_1}, f_{\theta_2})}{\partial\theta_{1i}\partial\theta_{1j}}\right)_{\theta_1=\theta_2} = a_{ij}^\gamma(\theta_1).
\end{aligned}$$

Therefore, we get

$$\begin{aligned}
2C_\gamma(f_{{}^{(1)}\hat{\theta}_\alpha}, f_{{}^{(2)}\hat{\theta}_\alpha}) = & ({}^{(1)}\hat{\theta}_\alpha - \theta_1)^T A_\gamma(\theta_1)({}^{(1)}\hat{\theta}_\alpha - \theta_1) \\
& -2({}^{(1)}\hat{\theta}_\alpha - \theta_1)^T A_\gamma(\theta_1)({}^{(2)}\hat{\theta}_\alpha - \theta_1) \\
& +({}^{(2)}\hat{\theta}_\alpha - \theta_1)^T A_\gamma(\theta_1)({}^{(2)}\hat{\theta}_\alpha - \theta_1) \\
& + o\left(||{}^{(1)}\hat{\theta}_\alpha - \theta_1||^2\right) + o\left(||{}^{(2)}\hat{\theta}_\alpha - \theta_2||^2\right) \\
= & ({}^{(1)}\hat{\theta}_\alpha - {}^{(2)}\hat{\theta}_\alpha)^T A_\gamma(\theta_1)({}^{(1)}\hat{\theta}_\alpha - {}^{(2)}\hat{\theta}_\alpha) \\
& + o\left(||{}^{(1)}\hat{\theta}_\alpha - \theta_1||^2\right) + o\left(||{}^{(2)}\hat{\theta}_\alpha - \theta_2||^2\right),
\end{aligned}$$

with

$$o\left(||{}^{(1)}\hat{\theta}_\alpha - \theta_1||^2\right) = o_p\left(\frac{1}{n}\right) \quad \text{and} \quad o\left(||{}^{(2)}\hat{\theta}_\alpha - \theta_2||^2\right) = o_p\left(\frac{1}{m}\right).$$

Thus the asymptotic distribution of

$$S_{\gamma,\lambda}({}^{(1)}\hat{\theta}_\alpha, {}^{(2)}\hat{\theta}_\alpha) = \frac{2nm}{n+m}\ C_\gamma(f_{{}^{(1)}\hat{\theta}_\alpha}, f_{{}^{(2)}\hat{\theta}_\alpha})$$

coincides with the distribution of the random variable $\sum_{i=1}^{r}\ \zeta_i^{\gamma,\alpha}(\theta_1)Z_i^2$.

S2 Lehmann Conditions

The following four conditions from Lehmann (1983, page 429) are referred to as the Lehman Conditions in the main paper; we report them here for the shake of completeness.

(A) There exists an open subset ω of Θ containing the true parameter point θ such that for almost all x the density $f_\theta(x)$ admits all third derivatives

$$(\partial^3/\partial\theta_j\partial\theta_k\partial\theta_l)f_\theta(x) \text{ for all } \theta \in \omega.$$

(B) The first and second logarithmic derivatives of f satisfy the equations $E_\theta\left[\frac{\partial}{\partial\theta_j}\log f_\theta(x)\right] = 0$, for all j and the information matrix $I(\theta)$ with $(j,k)^{th}$ element is defined as

$$I_{jk}(\theta) = E_\theta\left[\frac{\partial}{\partial\theta_j}\log f(X,\theta).\frac{\partial}{\partial\theta_k}\log f(X,\theta)\right] = E_\theta\left[-\frac{\partial^2}{\partial\theta_j\partial\theta_k}\log f(X,\theta)\right].$$

(C) The matrix $I(\theta)$ is a positive definite matrix.

(D) There exist functions M_{jkl} such that

$$\left|\frac{\partial^3}{\partial\theta_j\partial\theta_k\partial\theta_l}\log f_\theta(x)\right| \leq M_{jkl}(x) \text{ for all } \theta \in \omega$$

where $m_{jkl} = E_\theta\left[M_{jkl}(X)\right] < \infty$ for all j, k, l.

S3 Additional Numerical Illustrations: The MGPDEs under Poisson Model

Here we report the bias and MSE of the MGPDEs for several more values of (α, λ) combination near the optimum region under the simulation set-up (with sample size $n = 50$) as described in Section 3.3 of the main paper. Tables S1-S2 present the bias and MSE values, respectively, under the pure data scenarios whereas the same under contamination scenario are presented in Tables S3-S4.

References

1. Basu, A., I. R. Harris, N. L. Hjort, and M. C. Jones (1998). Robust and efficient estimation by minimising a density power divergence. *Biometrika*, **85**, 549–559.
2. Basu, A., H. Shioya, and C. Park (2011). *Statistical Inference: The Minimum Distance Approach.* Chapman & Hall/CRC, Boca Raton, FL.
3. Chung, K. L. (1974). *A Course in Probability Theory.* Academic Press, New York.
4. Dik, J. J. and de Gunst, M. C. M. (1985). The distribution of general quadratic forms in normal variables. *Statistica Neerlandica*, **39**, 14–26.
5. Lehmann, E. L. (1983). *Theory of Point Estimation.* John Wiley & Sons.
6. Lindsay, B. G. (1994). Efficiency versus robustness: The case for minimum Hellinger distance and related methods. *Annals of Statistics*, **22**, 1081–1114.

Table S1 Bias of the MGPDEs of the Poisson mean under pure data ($n = 50$) for various values of λ and α

$\lambda\downarrow\alpha\rightarrow$	0	0.1	0.2	0.3	0.4	0.5
−0.9	−0.347	−0.3082	−0.2739	−0.2436	−0.2168	−0.1928
−0.8	−0.234	−0.2097	−0.1872	−0.1666	−0.1477	−0.1303
−0.7	−0.17	−0.1544	−0.139	−0.1243	−0.1103	−0.097
−0.6	−0.1266	−0.117	−0.1067	−0.0962	−0.0857	−0.0753
−0.5	−0.094	−0.089	−0.0826	−0.0754	−0.0676	−0.0596
−0.4	−0.068	−0.0667	−0.0635	−0.059	−0.0535	−0.0474
−0.3	−0.0462	−0.0481	−0.0476	−0.0454	−0.042	−0.0375
−0.2	−0.0271	−0.0319	−0.0339	−0.0338	−0.0321	−0.0292
−0.1	−0.0099	−0.0173	−0.0216	−0.0234	−0.0233	−0.0218
0	0.006	−0.0039	−0.0103	−0.0139	−0.0154	−0.0153
0.1	0.0211	0.0089	0.0004	−0.005	−0.008	−0.0092
0.2	0.0356	0.0213	0.0107	0.0035	−0.001	−0.0034
0.3	0.0493	0.0336	0.0209	0.0118	0.0058	0.0021

Table S2 MSE of the MGPDEs of the Poisson mean under pure data ($n = 50$) for various values of λ and α

$\lambda\downarrow\alpha\rightarrow$	0	0.1	0.2	0.3	0.4	0.5
−0.9	0.3269	0.3037	0.284	0.2677	0.2544	0.2438
−0.8	0.1994	0.1936	0.1888	0.1853	0.1829	0.1819
−0.7	0.1493	0.1499	0.151	0.1526	0.1548	0.1578
−0.6	0.1244	0.1278	0.1316	0.1359	0.1405	0.1456
−0.5	0.1104	0.115	0.1204	0.1261	0.1321	0.1385
−0.4	0.1019	0.1071	0.1131	0.1197	0.1266	0.1337
−0.3	0.0966	0.1017	0.1081	0.1151	0.1225	0.1303
−0.2	0.0932	0.098	0.1044	0.1116	0.1194	0.1275
−0.1	0.0912	0.0954	0.1015	0.1087	0.1167	0.125
0	0.0902	0.0935	0.0992	0.1063	0.1143	0.1228
0.1	0.0903	0.0923	0.0974	0.1042	0.1121	0.1207
0.2	0.0916	0.0916	0.0959	0.1023	0.11	0.1186
0.3	0.0943	0.0918	0.0947	0.1006	0.1081	0.1165

Table S3 Bias of the MGPDEs of the Poisson mean under contaminated data ($n = 50$) for various values of λ and α

$\lambda\downarrow\alpha\rightarrow$	0	0.1	0.2	0.3	0.4	0.5
−0.9	−0.3273	−0.3004	−0.2558	−0.2201	−0.1896	−0.1627
−0.8	−0.2169	−0.2131	−0.1742	−0.1448	−0.1205	−0.099
−0.7	−0.1521	−0.1663	−0.1296	−0.1039	−0.0833	−0.0651
−0.6	−0.1059	−0.1375	−0.1009	−0.0771	−0.059	−0.0431
−0.5	−0.0682	−0.1196	−0.0811	−0.058	−0.0415	−0.0275
−0.4	−0.0319	−0.11	−0.0673	−0.0436	−0.0282	−0.0155
−0.3	0.0144	−0.1088	−0.0586	−0.0329	−0.0177	−0.006
−0.2	0.1091	−0.1178	−0.0553	−0.0252	−0.0094	0.0018
−0.1	0.4646	−0.1394	−0.0595	−0.0209	−0.003	0.0081
0	1.9957	−0.1476	−0.0752	−0.0216	0.0012	0.0132
0.1	4.0455	0.4136	−0.0981	−0.0304	0.002	0.0168
0.2	5.3173	2.9705	0.1214	−0.0475	−0.0029	0.0182
0.3	6.0811	4.7843	2.2009	0.0681	−0.0145	0.0155

Table S4 MSE of the MGPDEs of the Poisson mean under contaminated data ($n = 50$) for various values of λ and α

$\lambda\downarrow\alpha\rightarrow$	0	0.1	0.2	0.3	0.4	0.5
−0.9	0.338	0.3192	0.2924	0.2736	0.2592	0.2481
−0.8	0.2156	0.2142	0.2017	0.1953	0.1919	0.1905
−0.7	0.1674	0.1725	0.1659	0.1651	0.1667	0.1697
−0.6	0.1441	0.1518	0.148	0.1502	0.1546	0.1601
−0.5	0.132	0.1404	0.1378	0.1418	0.1478	0.1548
−0.4	0.1267	0.1342	0.1314	0.1364	0.1435	0.1516
−0.3	0.1283	0.1317	0.1272	0.1326	0.1405	0.1492
−0.2	0.1526	0.1335	0.1243	0.1297	0.138	0.1473
−0.1	0.4363	0.143	0.123	0.1272	0.1359	0.1456
0	4.8308	0.1613	0.1249	0.1252	0.1338	0.1438
0.1	18.0435	0.3685	0.1338	0.1241	0.1317	0.142
0.2	30.4329	10.1105	0.1507	0.1265	0.1297	0.1399
0.3	39.3892	25.0289	5.772	0.1348	0.1294	0.1376