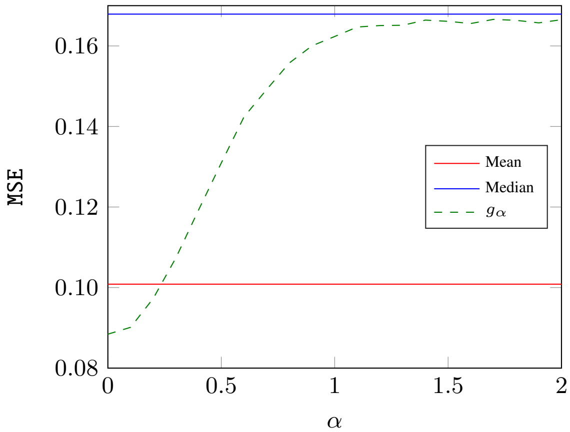
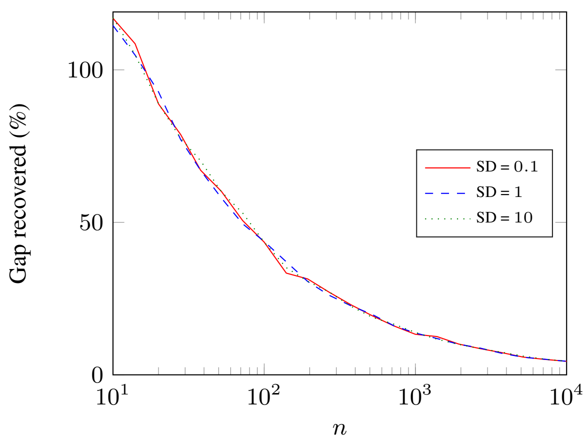
Appendix



(a) The MSEs of the sample mean, the sample median, and g_α as a function of $|\alpha|$ (the distance of the phantom from the true mean) for the standard Gaussian and $n = 8$.



(b) The difference between the MSEs of the sample median and g_α as a percentage of the difference between the MSEs of the sample median and the sample mean, when the underlying distribution is the standard Gaussian. To measure the best-case improvement, we set $\alpha = 0$ (i.e., at the true mean).

Figure 3. Adding a single real-valued phantom when the number of samples n is even.

A. Proof of Lemma 3.3

Fix a symmetric distribution $D \in \mathcal{D}^{\text{sym}}$ with mean μ, PDF f, and CDF F. Let $g_t(x_1, \ldots, x_n) = x_{(t)}$ denote the t^{th} order statistic. We want to prove $\texttt{MSE}(g_t, D) \geq \texttt{MSE}(g_{t+1}, D)$ for $t \leq (n-1)/2$. The comparison between g_{n-t} and g_{n-t+1} follows from symmetry of D.

We prove this by considering each "negative" example where the squared error of g_t is less than the squared error of g_{t+1} by an amount d, and map it to a unique "positive" example where the squared error of g_{t+1} is less than the squared error of g_t by d. The result follows by ensuring that the negative example has at most as much probability density as its corresponding positive example. There are two cases of negative examples.

Case 1: $x_{(t)} = \mu + a$ and $x_{(t+1)} = \mu + b$, where $0 \leq a < b$. The squared error of g_t is $d = b^2 - a^2$ less than that of g_{t+1}. Let us map it to the positive example where $x_{(t)} = \mu - b$ and $x_{(t+1)} = \mu - a$. In this (unique) positive example, the squared error of g_{t+1} is exactly $b^2 - a^2$ less than that of g_t. Let f_N and f_P denote the probability densities of the negative and the positive examples. We need to show $f_N \leq f_P$. Now, f_P/f_N is

$$\frac{f(\mu-a)f(\mu-b)\,[F(\mu-b)]^{t-1}\,[1-F(\mu-a)]^{n-t-1}}{f(\mu+a)f(\mu+b)\,[F(\mu+a)]^{t-1}\,[1-F(\mu+b)]^{n-t-1}} = \left(\frac{F(\mu+a)}{F(\mu-b)}\right)^{n-2t} \geq 1,$$

where the first transition holds due to symmetry of D around μ, and the final transition holds because $F(\mu-b) \leq 1/2 \leq F(\mu+a)$ and $n - 2t > 0$.

Case 2: $x_{(t)} = \mu - a$ and $x_{(t+1)} = \mu + b$, where $0 \leq a < b$. Here, $x_{(t)}$ and $x_{(t+1)}$ are on different sides of μ. We map it to the (unique) positive example where $x_{(t)} = \mu - b$ and $x_{(t+1)} = \mu + a$, thus maintaining them on different sides. Both examples admit an identical difference of $b^2 - a^2$ between the squared errors, and the ratio f_P/f_N is

$$\frac{f(\mu+a)f(\mu-b)\,[F(\mu-b)]^{t-1}\,[1-F(\mu+a)]^{n-t-1}}{f(\mu-a)f(\mu+b)\,[F(\mu-a)]^{t-1}\,[1-F(\mu+b)]^{n-t-1}} = \left(\frac{F(\mu-a)}{F(\mu-b)}\right)^{n-2t} \geq 1.$$

For the final inequality, note that we still have $F(\mu-b) \leq F(\mu-a)$ because $b > a \geq 0$. ■

B. Proof of Proposition 3.4

Fix $\alpha \in \mathbb{R}$. First, observe that g_α is the generalized median obtained by placing one phantom on α, and an equal number of phantoms on $-\infty$ and ∞. The following alternative formulation of g_α provides further intuition.

$$g_\alpha(\mathbf{x}) = \begin{cases} x_{(n/2)} & \text{if } \alpha \leq x_{(n/2)}, \\ \alpha & \text{if } x_{(n/2)} \leq \alpha \leq x_{(n/2+1)}, \\ x_{(n/2+1)} & \text{if } x_{(n/2+1)} \leq \alpha. \end{cases}$$

Thus, g_α always chooses among the left median, the right median, and α. Fix a distribution $D \in \mathcal{D}^{\text{sym}}$ with mean μ.

Let $\texttt{med}_\ell(\mathbf{x}) = x_{(n/2)}$ and $\texttt{med}_r(\mathbf{x}) = x_{(n/2+1)}$ denote the left and the right medians. We show that $\texttt{MSE}(g_\alpha, D) \leq \texttt{MSE}(\texttt{med}_\ell, D) = \texttt{MSE}(\texttt{med}_r, D)$. Comparison with other order statistics then follows immediately from Lemma 3.3.

Note that $\texttt{MSE}(\texttt{med}_\ell, D) = \texttt{MSE}(\texttt{med}_r, D)$ holds due to the symmetry of D.

Suppose $\mu \geq \alpha$. Observe that $g_\alpha(\mathbf{x}) \neq \texttt{med}_\ell(\mathbf{x})$ implies $\texttt{med}_\ell(\mathbf{x}) \leq g_\alpha(\mathbf{x}) < \alpha \leq \mu$, and in that case $g_\alpha(\mathbf{x})$ is closer to μ than $\texttt{med}_\ell(\mathbf{x})$. This yields $\texttt{MSE}(g_\alpha, D) \leq \texttt{MSE}(\texttt{med}_\ell, D)$. For $\mu \leq \alpha$, a similar argument establishes $\texttt{MSE}(g_\alpha, D) \leq \texttt{MSE}(\texttt{med}_r, D)$. The proof now follows from the fact that $\texttt{MSE}(\texttt{med}_\ell, D) = \texttt{MSE}(\texttt{med}_r, D)$ for any symmetric distribution D. ■