Supplemental: Active Diagnosis via AUC Maximization

Gowtham Bellala, Jason Stanley, Clayton Scott
Electrical Engineering and Computer Science
University of Michigan, Ann Arbor

Suresh K. Bhavnani
Institute for Translational Sciences
University of Texas Medical Branch, Galveston

1 AUC Estimation

The area under the ROC curve can be approximated using lower rectangles, upper rectangles or by using a linear approximation, as shown in Figure 1. The expressions related to each of these approximations are

$$\underline{\mathbf{A}}_l(\mathbf{z}_\mathcal{A}) = \sum_{t=0}^{M-1}(1-\widehat{\mathrm{MR}}_t)(\widehat{\mathrm{FAR}}_{t+1}-\widehat{\mathrm{FAR}}_t)$$

$$\underline{\mathbf{A}}_u(\mathbf{z}_\mathcal{A}) = \sum_{t=0}^{M-1}(1-\widehat{\mathrm{MR}}_{t+1})(\widehat{\mathrm{FAR}}_{t+1}-\widehat{\mathrm{FAR}}_t)$$

$$\underline{\mathbf{A}}_m(\mathbf{z}_\mathcal{A}) = \sum_{t=0}^{M-1}(1-\frac{\widehat{\mathrm{MR}}_t+\widehat{\mathrm{MR}}_{t+1}}{2})(\widehat{\mathrm{FAR}}_{t+1}-\widehat{\mathrm{FAR}}_t).$$

Substituting the estimates for miss rate and false alarm rate, the corresponding approximations for the area above the ROC curve are given by

$$\overline{\mathbf{A}}_l(\mathbf{z}_\mathcal{A}) = \frac{\sum_{i=1}^{M}\sum_{j=i}^{M}\Pr(X_{r(i)}=0|\mathbf{z}_\mathcal{A})\Pr(X_{r(j)}=1|\mathbf{z}_\mathcal{A})}{\sum_{i=1}^{M}\Pr(X_i=1|\mathbf{z}_\mathcal{A})\sum_{i=1}^{M}\Pr(X_i=0|\mathbf{z}_\mathcal{A})} \tag{1a}$$

$$\overline{\mathbf{A}}_u(\mathbf{z}_\mathcal{A}) = \frac{\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\Pr(X_{r(i)}=0|\mathbf{z}_\mathcal{A})\Pr(X_{r(j)}=1|\mathbf{z}_\mathcal{A})}{\sum_{i=1}^{M}\Pr(X_i=1|\mathbf{z}_\mathcal{A})\sum_{i=1}^{M}\Pr(X_i=0|\mathbf{z}_\mathcal{A})} \tag{1b}$$

$$\overline{\mathbf{A}}_m(\mathbf{z}_\mathcal{A}) = \frac{\sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\Pr(X_{r(i)}=0|\mathbf{z}_\mathcal{A})\Pr(X_{r(j)}=1|\mathbf{z}_\mathcal{A})}{\sum_{i=1}^{M}\Pr(X_i=1|\mathbf{z}_\mathcal{A})\sum_{i=1}^{M}\Pr(X_i=0|\mathbf{z}_\mathcal{A})}$$

$$+\frac{\sum_{i=1}^{M}\Pr(X_i=0|\mathbf{z}_\mathcal{A})\Pr(X_i=1|\mathbf{z}_\mathcal{A})}{2\sum_{i=1}^{M}\Pr(X_i=1|\mathbf{z}_\mathcal{A})\sum_{i=1}^{M}\Pr(X_i=0|\mathbf{z}_\mathcal{A})}. \tag{1c}$$

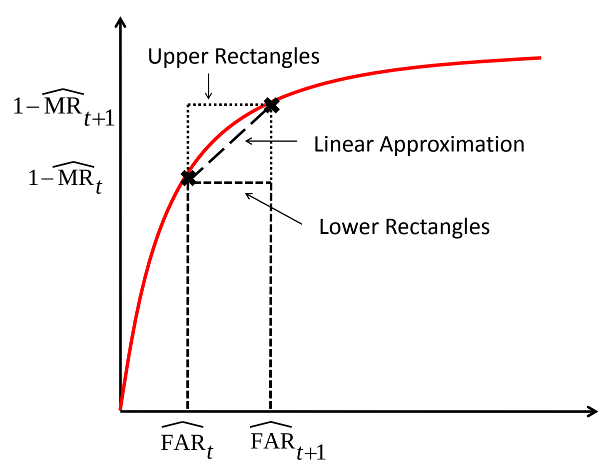

Figure 1: Demonstrates the different approximations for area under the ROC curve

1.1 Choice of Upper rectangles

As mentioned in the paper, query selection based on AUC approximated using the upper rectangles performs better than the other two. We will now provide an intuitive explanation for this phenomenon.

Using the result in Proposition 1 below, note that under a single fault assumption, the approximations to the area above the ROC curve in (1) reduce to

$$\overline{\mathbf{A}}_l(\mathbf{z}_{\mathcal{A}}) = \frac{\sum_{i=1}^{M} 2i\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) - \Pr^2(X_i = 1|\mathbf{z}_{\mathcal{A}})}{2(M-1)} + c_l,$$

$$\overline{\mathbf{A}}_m(\mathbf{z}_{\mathcal{A}}) = \frac{\sum_{i=1}^{M} 2i\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}})}{2(M-1)} + c_m,$$

$$\overline{\mathbf{A}}_u(\mathbf{z}_{\mathcal{A}}) = \frac{\sum_{i=1}^{M} 2i\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) + \Pr^2(X_i = 1|\mathbf{z}_{\mathcal{A}})}{2(M-1)} + c_u,$$

where c_l, c_m and c_u are constants that do not contribute to query selection.

Now note that all the three approximations have the same first term, which corresponds to the expected rank of the faults in the ranked list. However, they differ with respect to the second term, which makes the crucial difference in terms of the query selected. More specifically, given two or more queries with the the same expected rank value (i.e., same value for the first term), query selected using $\overline{\mathbf{A}}_u(\mathbf{z}_{\mathcal{A}})$ chooses the one that most evenly distributes the posterior probability mass of 1 among all the objects, while query selected using $\overline{\mathbf{A}}_l(\mathbf{z}_{\mathcal{A}})$ chooses the one that assigns most of the probability mass to one object, and the query selected using $\overline{\mathbf{A}}_m(\mathbf{z}_{\mathcal{A}})$ just picks one at random. Hence, the queries selected using $\overline{\mathbf{A}}_l(\mathbf{z}_{\mathcal{A}})$ and $\overline{\mathbf{A}}_m(\mathbf{z}_{\mathcal{A}})$ are more prone to increase the posterior fault probability of one (or few) object(s), thereby creating a bias towards those objects in the queries selected there after. However, this is overcome by the AUC-based query selection criterion approximated using the upper rectangles.

2 Proof of Proposition 1

Proposition 1. (Extended) *The estimates for the area above the ROC curve, $\overline{\mathbf{A}}_l(\mathbf{z}_{\mathcal{A}})$, $\overline{\mathbf{A}}_m(\mathbf{z}_{\mathcal{A}})$ and $\overline{\mathbf{A}}_u(\mathbf{z}_{\mathcal{A}})$ in* (1) *can be equivalently expressed as*

$$\overline{\mathbf{A}}_l(\mathbf{z}_{\mathcal{A}}) = \frac{1}{2} + \frac{\mathbf{U}(\mathbf{z}_{\mathcal{A}}) + \mathbf{V}(\mathbf{z}_{\mathcal{A}})}{2\mathbf{W}(\mathbf{z}_{\mathcal{A}})}$$
$$\overline{\mathbf{A}}_m(\mathbf{z}_{\mathcal{A}}) = \frac{1}{2} + \frac{\mathbf{U}(\mathbf{z}_{\mathcal{A}})}{2\mathbf{W}(\mathbf{z}_{\mathcal{A}})}$$
$$\overline{\mathbf{A}}_u(\mathbf{z}_{\mathcal{A}}) = \frac{1}{2} + \frac{\mathbf{U}(\mathbf{z}_{\mathcal{A}}) - \mathbf{V}(\mathbf{z}_{\mathcal{A}})}{2\mathbf{W}(\mathbf{z}_{\mathcal{A}})}$$

where

$$\mathbf{U}(\mathbf{z}_{\mathcal{A}}) = \sum_{i=1}^{M}(2i - M - 1)\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) \quad \text{(2a)}$$
$$\mathbf{V}(\mathbf{z}_{\mathcal{A}}) = \sum_{i=1}^{M}\Pr(X_i = 1|\mathbf{z}_{\mathcal{A}})\Pr(X_i = 0|\mathbf{z}_{\mathcal{A}}) \quad \text{(2b)}$$
$$\mathbf{W}(\mathbf{z}_{\mathcal{A}}) = \sum_{i=1}^{M}\Pr(X_i = 1|\mathbf{z}_{\mathcal{A}})\sum_{i=1}^{M}\Pr(X_i = 0|\mathbf{z}_{\mathcal{A}}).$$

Proof. We will show the equivalence result for $\overline{\mathbf{A}}_u(\mathbf{z}_{\mathcal{A}})$, and the other two results follow by observing that

$$\overline{\mathbf{A}}_l(\mathbf{z}_{\mathcal{A}}) = \overline{\mathbf{A}}_u(\mathbf{z}_{\mathcal{A}}) + \frac{\mathbf{V}(\mathbf{z}_{\mathcal{A}})}{\mathbf{W}(\mathbf{z}_{\mathcal{A}})}$$
$$\overline{\mathbf{A}}_m(\mathbf{z}_{\mathcal{A}}) = \overline{\mathbf{A}}_u(\mathbf{z}_{\mathcal{A}}) + \frac{\mathbf{V}(\mathbf{z}_{\mathcal{A}})}{2\mathbf{W}(\mathbf{z}_{\mathcal{A}})}.$$

We will now show the equivalence result for $\overline{\mathbf{A}}_u(\mathbf{z}_{\mathcal{A}})$. Let $\mathbf{N}(\mathbf{z}_{\mathcal{A}}) := \sum_{i=1}^{M-1}\sum_{j=i+1}^{M}\Pr(X_{r(i)} = 0|\mathbf{z}_{\mathcal{A}})\Pr(X_{r(j)} = 1|\mathbf{z}_{\mathcal{A}})$ denote its numerator. Then, the result follows by observing that

$$\begin{aligned}
&\sum_{i=1}^{M}\Pr(X_i = 0|\mathbf{z}_{\mathcal{A}})\sum_{i=1}^{M}\Pr(X_i = 1|\mathbf{z}_{\mathcal{A}}) \\
&= \sum_{i=1}^{M}\Pr(X_{r(i)} = 0|\mathbf{z}_{\mathcal{A}})\sum_{i=1}^{M}\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) \\
&= \mathbf{N}(\mathbf{z}_{\mathcal{A}}) + \sum_{i=1}^{M}\Pr(X_{r(i)} = 0|\mathbf{z}_{\mathcal{A}})\sum_{j=1}^{i}\Pr(X_{r(j)} = 1|\mathbf{z}_{\mathcal{A}}) \\
&= \mathbf{N}(\mathbf{z}_{\mathcal{A}}) + \sum_{i=1}^{M}\Pr(X_{r(i)} = 0|\mathbf{z}_{\mathcal{A}})\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) \\
&\quad + \sum_{i=2}^{M}\sum_{j=1}^{i-1}\Pr(X_{r(i)} = 0|\mathbf{z}_{\mathcal{A}})\Pr(X_{r(j)} = 1|\mathbf{z}_{\mathcal{A}}), \qquad (3)
\end{aligned}$$

where the last term in the above expression can be expressed in terms of $\mathbf{N}(\mathbf{z}_{\mathcal{A}})$ using the relation $\Pr(X_{r(i)} = 0|\mathbf{z}_{\mathcal{A}}) = 1 - \Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}})$,

$$\begin{aligned}
&\sum_{i=2}^{M}\sum_{j=1}^{i-1}\Pr(X_{r(i)} = 0|\mathbf{z}_{\mathcal{A}})\Pr(X_{r(j)} = 1|\mathbf{z}_{\mathcal{A}}) \\
&= \sum_{i=2}^{M}\sum_{j=1}^{i-1}\Big[1 - \Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) - \Pr(X_{r(j)} = 0|\mathbf{z}_{\mathcal{A}}) \\
&\qquad + \Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}})\Pr(X_{r(j)} = 0|\mathbf{z}_{\mathcal{A}})\Big] \\
&= \sum_{i=2}^{M}\sum_{j=1}^{i-1}\Big[-\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) + \Pr(X_{r(j)} = 1|\mathbf{z}_{\mathcal{A}}) \\
&\qquad + \Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}})\Pr(X_{r(j)} = 0|\mathbf{z}_{\mathcal{A}})\Big]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=2}^{M} -(i-1)\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) \\
&\quad + \sum_{i=1}^{M-1} (M-i)\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) \\
&\quad + \sum_{i=2}^{M}\sum_{j=1}^{i-1} \Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}})\Pr(X_{r(j)} = 0|\mathbf{z}_{\mathcal{A}}) \\
&= \sum_{i=1}^{M}(M-2i+1)\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) \\
&\quad + \sum_{j=1}^{M-1}\sum_{i=j+1}^{M} \Pr(X_{r(j)} = 0|\mathbf{z}_{\mathcal{A}})\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) \\
&= \sum_{i=1}^{M}(M-2i+1)\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) + \mathbf{N}(\mathbf{z}_{\mathcal{A}}).
\end{aligned}
$$

Finally, substituting the above relation in (3), we get

$$
\begin{aligned}
&\sum_{i=1}^{M} \Pr(X_i = 0|\mathbf{z}_{\mathcal{A}}) \sum_{i=1}^{M} \Pr(X_i = 1|\mathbf{z}_{\mathcal{A}}) \\
&= 2\mathbf{N}(\mathbf{z}_{\mathcal{A}}) + \sum_{i=1}^{M} \Pr(X_{r(i)} = 0|\mathbf{z}_{\mathcal{A}})\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) \\
&\qquad + \sum_{i=1}^{M}(M-2i+1)\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) \\
&= 2\mathbf{N}(\mathbf{z}_{\mathcal{A}}) + \sum_{i=1}^{M} \Pr(X_i = 0|\mathbf{z}_{\mathcal{A}})\Pr(X_i = 1|\mathbf{z}_{\mathcal{A}}) \\
&\qquad + \sum_{i=1}^{M}(M-2i+1)\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}})
\end{aligned}
$$

from which, the result follows. □

3 Proof of Theorem 1

Since $\underline{\mathbf{A}}(\mathbf{z}_{\mathcal{A}}) = 1 - \overline{\mathbf{A}}(\mathbf{z}_{\mathcal{A}})$, Theorem 1 corresponding to the adaptive monotonicity of $\underline{\mathbf{A}}(\mathbf{z}_{\mathcal{A}})$ can be equivalently stated as shown below.

Theorem 1. *Under the single fault assumption of Section* 4.1*, the quality functions* $\overline{\mathbf{A}}_l(\mathbf{Z}_{\mathcal{A}})$ *and* $\overline{\mathbf{A}}_m(\mathbf{Z}_{\mathcal{A}})$ *as defined in* (1a), (1c) *are adaptive monotone, i.e.,* $\forall \mathcal{A}' \subseteq \mathcal{A}$

$$\overline{\mathbf{A}}_l(\mathbf{Z}_{\mathcal{A}}) \leq \overline{\mathbf{A}}_l(\mathbf{Z}_{\mathcal{A}'}) \text{ and } \overline{\mathbf{A}}_m(\mathbf{Z}_{\mathcal{A}}) \leq \overline{\mathbf{A}}_m(\mathbf{Z}_{\mathcal{A}'})$$

Proof. Let $\mathbf{z}_{\mathcal{A}}$ denote the responses to queries in the set $\mathcal{A}$. To prove adaptive monotonicity for $\overline{\mathbf{A}}_l(\mathbf{Z}_{\mathcal{A}})$, it suffices to show that for any query $j \notin \mathcal{A}$, $\overline{\mathbf{A}}_l(\mathbf{z}_{\mathcal{A}}) - \mathbb{E}_{Z_j}[\overline{\mathbf{A}}_l(\mathbf{z}_{\mathcal{A}} \cup Z_j)] \geq 0$ (Golovin and Krause, 2010). Similarly, for $\overline{\mathbf{A}}_m(\mathbf{Z}_{\mathcal{A}})$, we need to show that $\overline{\mathbf{A}}_m(\mathbf{z}_{\mathcal{A}}) - \mathbb{E}_{Z_j}[\overline{\mathbf{A}}_m(\mathbf{z}_{\mathcal{A}} \cup Z_j)] \geq 0$.

Under single fault assumption,

$$
\begin{aligned}
\overline{\mathbf{A}}_l(\mathbf{z}_{\mathcal{A}}) &= \frac{1}{2} + \frac{\mathbf{U}(\mathbf{z}_{\mathcal{A}}) + \mathbf{V}(\mathbf{z}_{\mathcal{A}})}{2(M-1)}, \\
\overline{\mathbf{A}}_m(\mathbf{z}_{\mathcal{A}}) &= \frac{1}{2} + \frac{\mathbf{U}(\mathbf{z}_{\mathcal{A}})}{2(M-1)},
\end{aligned}
$$

where $\mathbf{U}(\mathbf{z}_{\mathcal{A}})$ and $\mathbf{V}(\mathbf{z}_{\mathcal{A}})$ are as defined in (2a) and (2b), respectively. Hence, the adaptive monotonicity of $\overline{\mathbf{A}}_l(\mathbf{z}_{\mathcal{A}})$ and $\overline{\mathbf{A}}_m(\mathbf{z}_{\mathcal{A}})$ follows by showing that $\forall j \notin \mathcal{A}$

$$
\begin{aligned}
&\mathbf{U}(\mathbf{z}_{\mathcal{A}}) - \mathbb{E}_{Z_j}[\mathbf{U}(\mathbf{z}_{\mathcal{A}} \cup Z_j)] \geq 0, \text{ and} \\
&\mathbf{V}(\mathbf{z}_{\mathcal{A}}) - \mathbb{E}_{Z_j}[\mathbf{V}(\mathbf{z}_{\mathcal{A}} \cup Z_j)] \geq 0,
\end{aligned}
$$

which follow from Lemma 1 and 2, below. □

Lemma 1. *Let* $\mathbf{z}_{\mathcal{A}}$ *denote the observed responses to queries in the set* $\mathcal{A}$*. Then, for any query* $j \notin \mathcal{A}$*,*

$$\mathbf{U}(\mathbf{z}_{\mathcal{A}}) - \mathbb{E}_{Z_j}[\mathbf{U}(\mathbf{z}_{\mathcal{A}} \cup Z_j)] \geq 0$$

Proof. Under single fault assumption, $\mathbf{U}(\mathbf{z}_{\mathcal{A}}) = -(M+1) + \sum_{i=1}^{M} 2i\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}})$. Hence, the result follows by showing that $\forall\ j \notin \mathcal{A}$,

$$
\begin{aligned}
\sum_{i=1}^{M} i \Bigg\{ &\Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}}) \\
&- \Big[\Pr(Z_j = 0|\mathbf{z}_{\mathcal{A}})\Pr(X_{r_0(i)} = 1|\mathbf{z}_{\mathcal{A}}, 0) \\
&+ \Pr(Z_j = 1|\mathbf{z}_{\mathcal{A}})\Pr(X_{r_1(i)} = 1|\mathbf{z}_{\mathcal{A}}, 1)\Big]\Bigg\} \geq 0. \quad (4)
\end{aligned}
$$

As mentioned earlier, the rank order depends on the queries chosen $\mathcal{A}$ and their observed responses $\mathbf{z}_{\mathcal{A}}$. Hence, to differentiate the rank orders in the above expression, we use $r(i)$ to denote the rank order of the objects based on the observed responses $\mathbf{z}_{\mathcal{A}}$, and $r_0(i)$, $r_1(i)$ to denote the rank order of the objects based on the observed responses $\mathbf{z}_{\mathcal{A}} \cup 0$ and $\mathbf{z}_{\mathcal{A}} \cup 1$ to queries in $\mathcal{A} \cup \{j\}$.

Note that (4) is equivalent to showing

$$
\begin{aligned}
\sum_{i=1}^{M}(M-i+1)\Bigg\{\Big[&\Pr(Z_j = 0|\mathbf{z}_{\mathcal{A}})\Pr(X_{r_0(i)} = 1|\mathbf{z}_{\mathcal{A}}, 0) \\
&+ \Pr(Z_j = 1|\mathbf{z}_{\mathcal{A}})\Pr(X_{r_1(i)} = 1|\mathbf{z}_{\mathcal{A}}, 1)\Big] \\
&- \Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}})\Bigg\} \geq 0. \quad (5)
\end{aligned}
$$

Let $\mathbf{f_t}(\mathbf{r}, \mathbf{z}_{\mathcal{A}}) := \sum_{i=1}^{t} \Pr(X_{r(i)} = 1|\mathbf{z}_{\mathcal{A}})$, i.e., the probability mass of the top t objects in the ranked list given by $\mathbf{r}$. Then,

$$\sum_{i=1}^{M}(M-i+1)\Pr(X_{r(i)}|\mathbf{z}_{\mathcal{A}}) = \sum_{t=1}^{M}\mathbf{f_t}(\mathbf{r},\mathbf{z}_{\mathcal{A}}),$$

and hence (5) is equivalent to showing

$$\sum_{t=1}^{M}\Bigg[\Pr(Z_j=0|\mathbf{z}_{\mathcal{A}})\mathbf{f_t}(\mathbf{r}_0,\mathbf{z}_{\mathcal{A}}\cup 0) + \Pr(Z_j=1|\mathbf{z}_{\mathcal{A}})\mathbf{f_t}(\mathbf{r}_1,\mathbf{z}_{\mathcal{A}}\cup 1)\Bigg] - \mathbf{f_t}(\mathbf{r},\mathbf{z}_{\mathcal{A}}) \geq 0.$$

Now, note that

$$\mathbf{f_t}(\mathbf{r}_0,\mathbf{z}_{\mathcal{A}}\cup 0) \geq \mathbf{f_t}(\mathbf{r},\mathbf{z}_{\mathcal{A}}\cup 0) = \sum_{i=1}^{t}\Pr(X_{r(i)}=1|\mathbf{z}_{\mathcal{A}},0).$$

Since the rank order $\mathbf{r}_0$ corresponds to the decreasing order of the posterior probabilities in $\{\Pr(X_i = 1|\mathbf{z}_{\mathcal{A}},0)\}_{i=1}^{M}$, the probability mass of the top t objects in this ranked list is greater than any other t objects. Similarly, $\mathbf{f_t}(\mathbf{r}_1,\mathbf{z}_{\mathcal{A}}\cup 1) \geq \mathbf{f_t}(\mathbf{r},\mathbf{z}_{\mathcal{A}}\cup 1)$. Hence,

$$\begin{aligned}
&\Pr(Z_j=0|\mathbf{z}_{\mathcal{A}})\mathbf{f_t}(\mathbf{r}_0,\mathbf{z}_{\mathcal{A}}\cup 0)\\
&\qquad + \Pr(Z_j=1|\mathbf{z}_{\mathcal{A}})\mathbf{f_t}(\mathbf{r}_1,\mathbf{z}_{\mathcal{A}}\cup 1) && \text{(6a)}\\
&\geq \Pr(Z_j=0|\mathbf{z}_{\mathcal{A}})\mathbf{f_t}(\mathbf{r},\mathbf{z}_{\mathcal{A}}\cup 0)\\
&\qquad + \Pr(Z_j=1|\mathbf{z}_{\mathcal{A}})\mathbf{f_t}(\mathbf{r},\mathbf{z}_{\mathcal{A}}\cup 1) && \text{(6b)}\\
&= \sum_{i=1}^{t}\Bigg[\Pr(Z_j=0|\mathbf{z}_{\mathcal{A}})\Pr(X_{r(i)}=1|\mathbf{z}_{\mathcal{A}},0)\\
&\qquad + \Pr(Z_j=1|\mathbf{z}_{\mathcal{A}})\Pr(X_{r(i)}=1|\mathbf{z}_{\mathcal{A}},1)\Bigg] && \text{(6c)}\\
&= \sum_{i=1}^{t}\Bigg[\Pr(Z_j=0|X_{r(i)}=1)\Pr(X_{r(i)}=1|\mathbf{z}_{\mathcal{A}})\\
&\qquad + \Pr(Z_j=1|X_{r(i)}=1)\Pr(X_{r(i)}=1|\mathbf{z}_{\mathcal{A}})\Bigg] && \text{(6d)}\\
&= \sum_{i=1}^{t}\Pr(X_{r(i)}=1|\mathbf{z}_{\mathcal{A}}) = \mathbf{f_t}(\mathbf{r},\mathbf{z}_{\mathcal{A}}). && \text{(6e)}
\end{aligned}$$

Thus proving the inequality.

Note that in the above equation, (6d) follows from (6c) by observing that under a single fault assumption, $X_i = 1 \iff \mathbf{X} = \mathbb{I}_i$, and hence, using the conditional independence assumption of Section 2, the posterior probability can be expressed as

$$\begin{aligned}
\Pr(X_i=1|\mathbf{z}_{\mathcal{A}},z) &= \Pr(\mathbf{X}=\mathbb{I}_i|\mathbf{z}_{\mathcal{A}},z)\\
&= \frac{\Pr(\mathbf{X}=\mathbb{I}_i)\Pr(\mathbf{z}_{\mathcal{A}}|\mathbf{X}=\mathbb{I}_i)\Pr(Z_j=z|\mathbf{X}=\mathbb{I}_i)}{\Pr(Z_j=z|\mathbf{z}_{\mathcal{A}})\Pr(\mathbf{Z}_{\mathcal{A}}=\mathbf{z}_{\mathcal{A}})}\\
&= \frac{\Pr(\mathbf{X}=\mathbb{I}_i|\mathbf{z}_{\mathcal{A}})\Pr(Z_j=z|\mathbf{X}=\mathbb{I}_i)}{\Pr(Z_j=z|\mathbf{z}_{\mathcal{A}})}\\
&= \frac{\Pr(X_i=1|\mathbf{z}_{\mathcal{A}})\Pr(Z_j=z|X_i=1)}{\Pr(Z_j=z|\mathbf{z}_{\mathcal{A}})}. \qquad (7)
\end{aligned}$$

□

Lemma 2. *Let $\mathbf{z}_{\mathcal{A}}$ denote the observed responses to queries in the set $\mathcal{A}$. Then, for any query $j \notin \mathcal{A}$,*

$$\mathbf{V}(\mathbf{z}_{\mathcal{A}}) - \mathbb{E}_{Z_j}[\mathbf{V}(\mathbf{z}_{\mathcal{A}}\cup Z_j)] \geq 0$$

Proof. Note that under single fault assumption, $\mathbf{V}(\mathbf{z}_{\mathcal{A}}) = 1 - \sum_{i=1}^{M}\Pr^2(X_i = 1|\mathbf{z}_{\mathcal{A}})$. Hence, we need to show that $\forall\ j \notin \mathcal{A}$,

$$\sum_{i=1}^{M}\Bigg\{\Bigg[\Pr(Z_j=0|\mathbf{z}_{\mathcal{A}})\Pr^2(X_i=1|\mathbf{z}_{\mathcal{A}},0) + \Pr(Z_j=1|\mathbf{z}_{\mathcal{A}})\Pr^2(X_i=1|\mathbf{z}_{\mathcal{A}},1)\Bigg] - \Pr^2(X_i=1|\mathbf{z}_{\mathcal{A}})\Bigg\} \geq 0. \qquad (8)$$

Substituting the expression for posterior probability from (7) in the LHS of (8), we get

$$\begin{aligned}
&\sum_{i=1}^{M}\Bigg\{\Pr^2(X_i=1|\mathbf{z}_{\mathcal{A}})\Bigg[\frac{\Pr^2(Z_j=0|X_i=1)}{\Pr(Z_j=0|\mathbf{z}_{\mathcal{A}})} + \frac{\Pr^2(Z_j=1|X_i=1)}{\Pr(Z_j=1|\mathbf{z}_{\mathcal{A}})} - 1\Bigg]\Bigg\}\\
&= \sum_{i=1}^{M}\Bigg\{\Pr^2(X_i=1|\mathbf{z}_{\mathcal{A}})\Bigg[\frac{\Big(1-\Pr(Z_j=1|X_i=1)\Big)^2}{\Pr(Z_j=0|\mathbf{z}_{\mathcal{A}})} + \frac{\Pr^2(Z_j=1|X_i=1)}{\Pr(Z_j=1|\mathbf{z}_{\mathcal{A}})} - 1\Bigg]\Bigg\},\\
&= \sum_{i=1}^{M}\Bigg\{\Pr^2(X_i=1|\mathbf{z}_{\mathcal{A}})\Bigg[\frac{\Big(\Pr(Z_j=1|X_i=1)-\Pr(Z_j=1|\mathbf{z}_{\mathcal{A}})\Big)^2}{\Pr(Z_j=1|\mathbf{z}_{\mathcal{A}})\Pr(Z_j=0|\mathbf{z}_{\mathcal{A}})}\Bigg]\Bigg\}\\
&\geq 0
\end{aligned}$$

where the last equality follows by using the relation $\Pr(Z_j = 0|\mathbf{z}_{\mathcal{A}}) = 1 - \Pr(Z_j = 1|\mathbf{z}_{\mathcal{A}})$, and completing the square. Thus, proving the result.

□

4 Proof of Proposition 2

Proposition 2. *Under the single fault assumption along with the conditional independence assumption of Section 2, the entropy-based query selection criterion in (2) reduces to*

$$j^* := \operatorname*{argmin}_{j \notin \mathcal{A}} \sum_{i=1}^{M} \Pr(X_i = 1|\mathbf{z}_\mathcal{A}) H\Big(\Pr(Z_j = 0|X_i = 1)\Big) - H\Big(\Pr(Z_j = 0|\mathbf{z}_\mathcal{A})\Big)$$

where $H(p) := -p \log_2 p - (1-p)\log_2(1-p)$ *denotes the binary entropy function.*

Proof. The entropy-based query selection criterion is given by

$$j^* = \operatorname*{argmin}_{j \notin \mathcal{A}} \sum_{z=0,1} \Pr(Z_j = z|\mathbf{z}_\mathcal{A}) H(\mathbf{X}|\mathbf{Z}_\mathcal{A}, z). \quad (9)$$

Since, under single fault assumption, $X_i = 1 \Longleftrightarrow \mathbf{X} = \mathbb{I}_i$, we need to show that the above query selection criterion reduces to

$$j^* := \operatorname*{argmin}_{j \notin \mathcal{A}} \sum_{i=1}^{M} \Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}) H\Big(\Pr(Z_j = 0|\mathbf{X} = \mathbb{I}_i)\Big) - H\Big(\Pr(Z_j = 0|\mathbf{z}_\mathcal{A})\Big).$$

We show this by first noting that under the single fault assumption, the conditional entropy reduces to

$$-\sum_{i=1}^{M} \Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}, z) \log \Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}, z).$$

In addition, as noted in (7), under the conditional independence assumption of Section 2, the posterior probability can be expressed as

$$\Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}, z) = \frac{\Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}) \Pr(Z_j = z|\mathbf{X} = \mathbb{I}_i)}{\Pr(Z_j = z|\mathbf{z}_\mathcal{A})}. \quad (10)$$

Substituting the above expression in (9), we get

$$\sum_{z=0,1} \Pr(Z_j = z|\mathbf{z}_\mathcal{A}) H(\mathbf{X}|\mathbf{Z}_\mathcal{A}, z) = -\sum_{z=0,1} \sum_{i=1}^{M} \Bigg[\Pr(Z_j = z|\mathbf{z}_\mathcal{A}) \Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}, z) \log \frac{\Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}) \Pr(Z_j = z|\mathbf{X} = \mathbb{I}_i)}{\Pr(Z_j = z|\mathbf{z}_\mathcal{A})}\Bigg]. \quad (11)$$

This expression can be broken down into 3 different terms. The first term is given by

$$\begin{aligned} &-\sum_{z=0,1} \sum_{i=1}^{M} \Bigg[\Pr(Z_j = z|\mathbf{z}_\mathcal{A}) \Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}, z) \log \Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A})\Bigg] \\ &= -\sum_{i=1}^{M} \Bigg[\Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}) \log \Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}) \sum_{z=0,1} \Pr(Z_j = z|\mathbf{X} = \mathbb{I}_i)\Bigg] \\ &= H(\mathbf{X}|\mathbf{z}_\mathcal{A}), \end{aligned}$$

where the second equality follows from (10) and the last equality follows since $\sum_z \Pr(Z_j = z|\mathbf{X} = \mathbb{I}_i) = 1$.

The second term is given by

$$\begin{aligned} &-\sum_{z=0,1} \sum_{i=1}^{M} \Bigg[\Pr(Z_j = z|\mathbf{z}_\mathcal{A}) \Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}, z) \log \frac{1}{\Pr(Z_j = z|\mathbf{z}_\mathcal{A})}\Bigg] \\ &= -\sum_{z=0,1} \Bigg[\Pr(Z_j = z|\mathbf{z}_\mathcal{A}) \log \frac{1}{\Pr(Z_j = z|\mathbf{z}_\mathcal{A})} \sum_{i=1}^{M} \Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}, z)\Bigg] \\ &= -H\Big(\Pr(Z_j = 0|\mathbf{z}_\mathcal{A})\Big), \end{aligned}$$

where the last equality follows since $\sum_{i=1}^{M} \Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}, z) = 1$.

The last term is given by

$$\begin{aligned} &-\sum_{z=0,1} \sum_{i=1}^{M} \Bigg[\Pr(Z_j = z|\mathbf{z}_\mathcal{A}) \Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}, z) \log \Pr(Z_j = z|\mathbf{X} = \mathbb{I}_i)\Bigg] \\ &= -\sum_{i=1}^{M} \Bigg[\Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}) \Bigg(\sum_{z=0,1} \Pr(Z_j = z|\mathbf{X} = \mathbb{I}_i) \log \Pr(Z_j = z|\mathbf{X} = \mathbb{I}_i)\Bigg)\Bigg] \\ &= \sum_{i=1}^{M} \Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}) H\Big(\Pr(Z_j = 0|\mathbf{X} = \mathbb{I}_i)\Big). \end{aligned}$$

Substituting these 3 terms back into (11), we get

$$\sum_{z=0,1} \Pr(Z_j = z|\mathbf{z}_\mathcal{A}) H(\mathbf{X}|\mathbf{Z}_\mathcal{A}, z)$$
$$= H(\mathbf{X}|\mathbf{z}_\mathcal{A}) - H\Big(\Pr(Z_j = 0|\mathbf{z}_\mathcal{A})\Big)$$
$$+ \sum_{i=1}^{M} \Pr(\mathbf{X} = \mathbb{I}_i|\mathbf{z}_\mathcal{A}) H\Big(\Pr(Z_j = 0|\mathbf{X} = \mathbb{I}_i)\Big),$$

and the result follows since $H(\mathbf{X}|\mathbf{z}_\mathcal{A})$ does not depend on the query j.

$\square$

5 Experiments

In this section, we provide more experimental evidence to support our argument that AUC-based query selection under single-fault assumption (AUC+SF) is a reliable, practical alternative to BPEA in large scale diagnosis problems.

We compare the performance of the three query selection criteria, i.e., BPEA, AUC-based query selection under single fault assumption (AUC+SF), and entropy-based query selection under single fault assumption (Entropy+SF), on two different datasets. The first dataset is a random bipartite diagnosis graph generated using the standard Preferential Attachment (PA) random network model. The second dataset is a network topology built using the BRITE generator, which simulates an Internet-like topology at the Autonomous Systems level.

Figures 2 and 3 compare the performance of the three query selection criteria on the two datasets, for different values of prior probability α, leak and inhibition probabilities ρ_l and ρ_i. In these figures, the area under the ROC curve (AUC) is obtained by ranking the objects based on their posterior probabilities, which in turn are computed using a single-fault assumption. Alternatively, note that these posterior probabilities could be estimated using belief propagation on these networks (as the networks are small in size), and the ranking obtained there after could be used to compute the AUC. Figures 4 and 5 compare the three query selection criteria using AUC computed through BP based ranking. Finally, the information gain is computed using BPEA as described in (Zheng et al., 2005).

Note from these figures that AUC+SF invariably performs better than Entropy+SF, and often comparable to BPEA, while having a computational complexity that is orders less than that of BPEA, thereby making it a robust, practical alternative to BPEA in large scale diagnosis problems.

References

D. Golovin and A. Krause. Adaptive Submodularity: A new approach to active learning and stochastic optimization. *In COLT*, 2010.

A. X. Zheng, I. Rish, and A. Beygelzimer. Efficient test selection in active diagnosis via entropy approximation. *In UAI*, 2005.

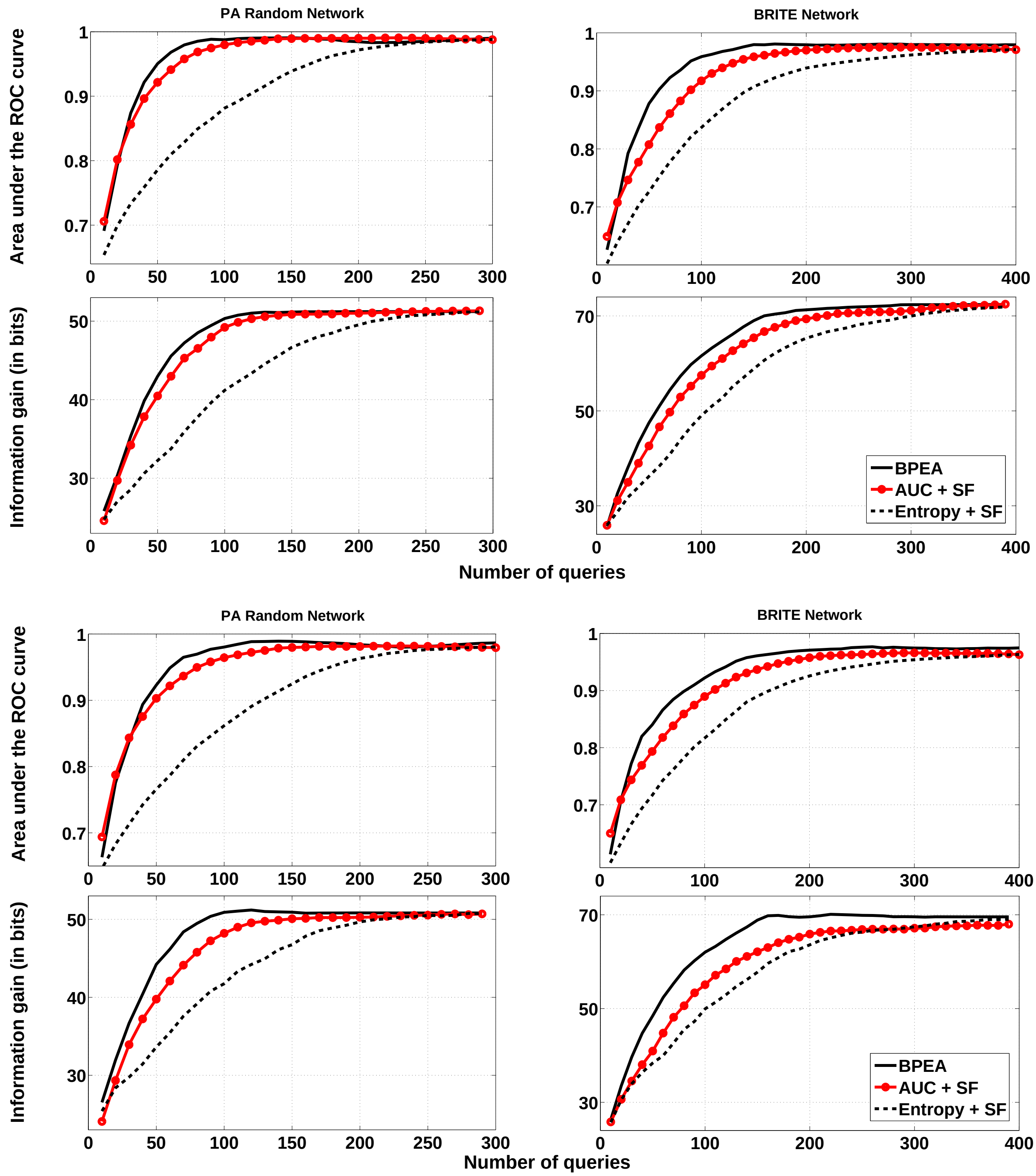

Figure 2: The plots in the first column correspond to a dataset generated using the PA model, and the second column corresponds to a BRITE network. The figure in the top corresponds to $(\alpha, \rho_i, \rho_l) = (0.03, 0.05, 0.05)$, and the figure in the bottom corresponds to $(\alpha, \rho_i, \rho_l) = (0.03, 0.1, 0.1)$.

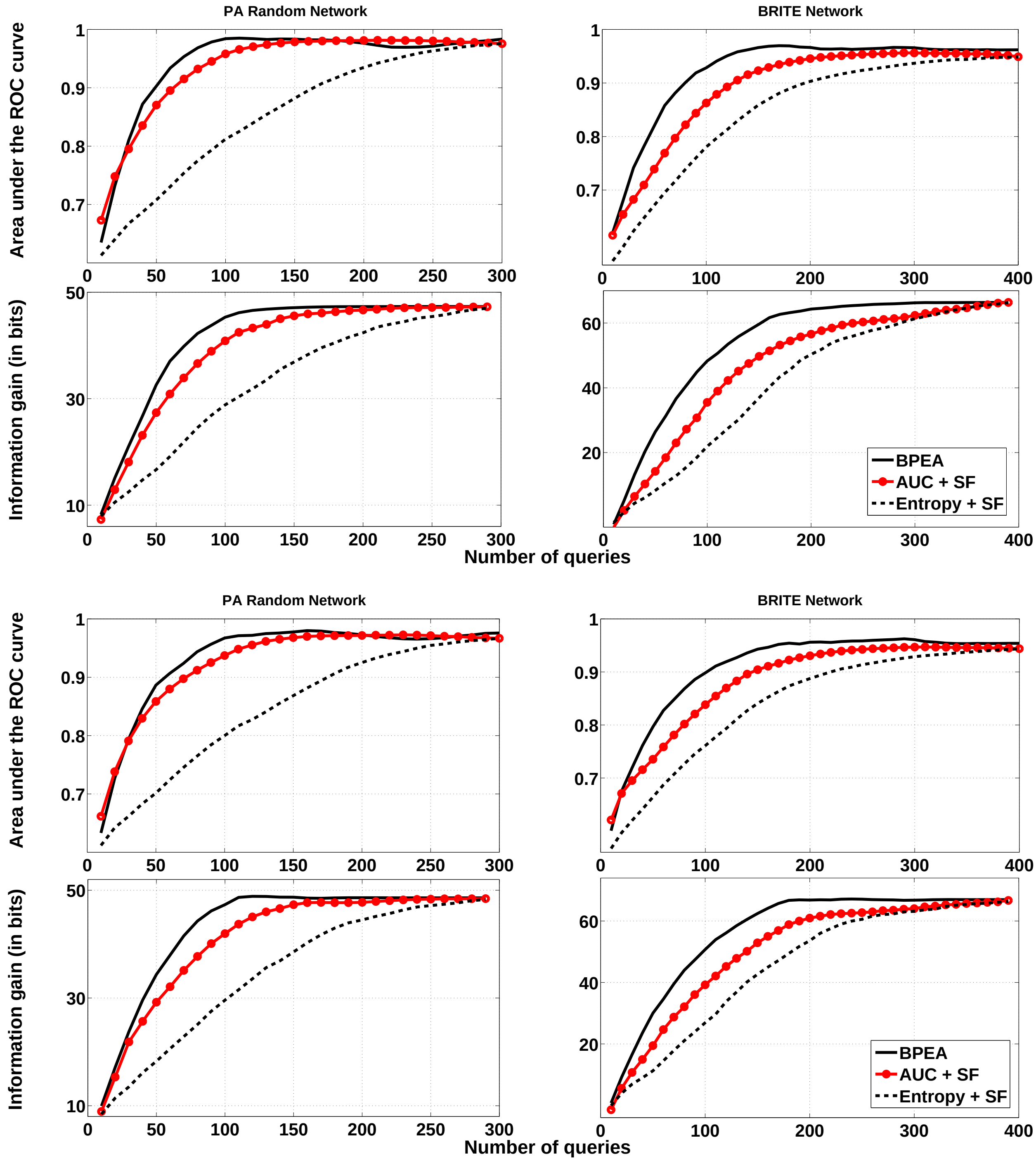

Figure 3: The plots in the first column correspond to a dataset generated using the PA model, and the second column corresponds to a BRITE network. The figure in the top corresponds to $(\alpha, \rho_i, \rho_l) = (0.05, 0.05, 0.05)$, and the figure in the bottom corresponds to $(\alpha, \rho_i, \rho_l) = (0.05, 0.1, 0.1)$.

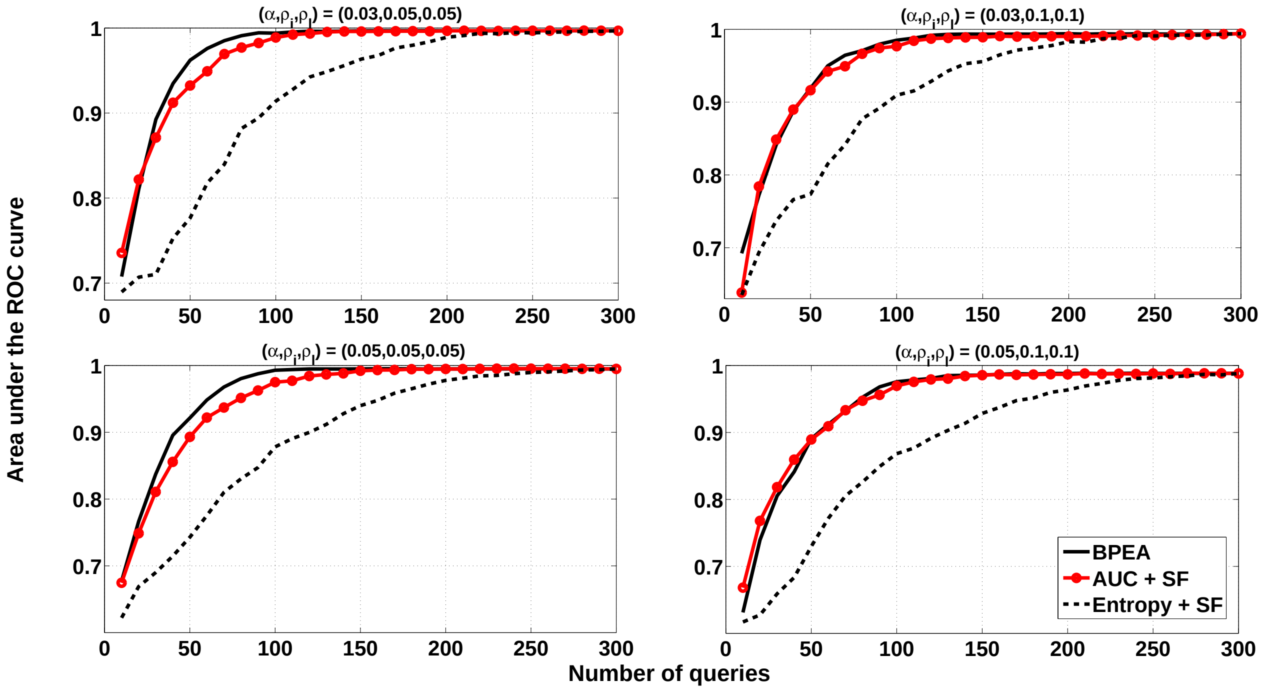

Figure 4: The plots in this figure correspond to a dataset generated using PA model. The AUC is computed by ranking the objects using posterior probabilities obtained from Belief Propagation.

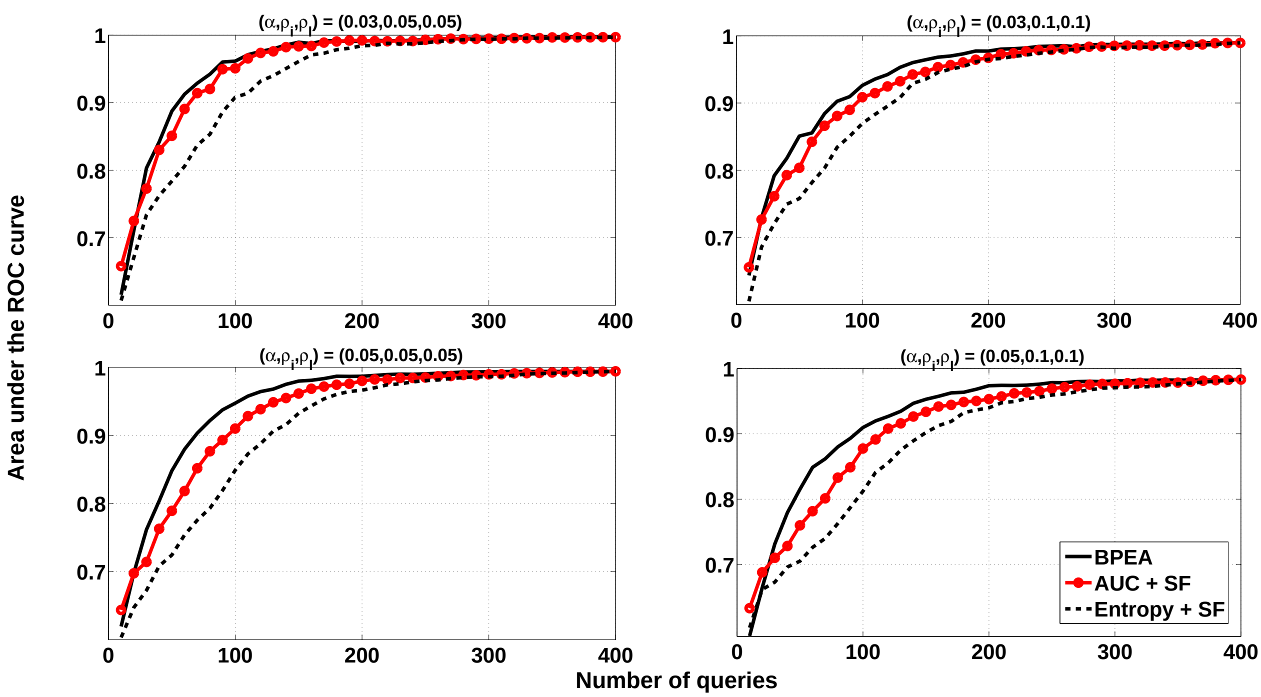

Figure 5: The plots in this figure correspond to a dataset generated using BRITE. The AUC is computed by ranking the objects using posterior probabilities obtained from Belief Propagation.