3 Receiver Operating Characteristic (ROC) Curves and Shortest-Path Examples

The key idea of this paper is that certain path graphs correspond to planar regions with a specific shape (equivalently, the relevant quantities can be computed by counting lattice points). In this section, we illustrate this idea using a few simple examples related to a Cayley graph generated from a binary vector and restricted to the orbit of the multiset $0^{n-k}1^k$ (i.e., the multipermutahedron graph).

In binary classification, True Positives (TP), False Negatives (FN), False Positives (FP), and True Negatives (TN) are computed by comparing model predictions to ground-truth labels. Two standard performance metrics are

$$\begin{aligned}\text{TPR (True Positive Rate)} &= \text{Sensitivity} = \text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}},\\ \text{FPR (False Positive Rate)} &= \frac{\text{FP}}{\text{FP}+\text{TN}}.\end{aligned}$$

Definition 9 (ROC AUC)**.** *Let $P = [p_0, \ldots, p_{n-1}]$ be a vector of distinct real numbers in $[0, 1]$, listed in ascending order. Let v be a binary vector of length n with $(n-k)$ zeros and k ones. Intuitively, P represents n propensity (score) values produced by a binary classifier, and v represents the corresponding ground-truth labels.*

Let $T = P + \epsilon$ be the set of threshold values we consider. Assume ϵ is sufficiently small so that for all j we have $T_j < T_{j+1}$.

For each threshold $T_j \in T$, define a prediction vector PRED_j by assigning 0 to entries with $P < T_j$ and 1 to entries with $P \geq T_j$. From PRED_j and v we compute TP, FP, TN, FN, *and hence* TPR *and* FPR, *obtaining the point $(\text{FPR}_j, \text{TPR}_j)$. Plotting these points yields the ROC curve (TPR vs. FPR). The AUC is the area under this curve.*

Proposition 1 (Path length and area relation)**.** *Let $S(v)$ be the graph generated from a binary vector v with $(n-k)$ zeros and k ones by adjacent transpositions. Let L be the length of the shortest path from v to $\text{sorted}(v) = [0, \ldots, 0, 1, \ldots, 1]$. Then*

$$L = (1 - \text{AUC})(n-k)k.$$

Equivalently, if A is the area above the ROC curve measured in unit squares, then $A = L$.

In the examples below we take $n = 10$ and $k = 5$, and fix

$$\begin{aligned}P &= [0.05, 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, 0.95],\\ T &= P + 0.01,\\ \text{PRED}_j &= \text{int}(P > T_j), \quad \forall j = 0, \ldots, n-1.\end{aligned}$$

3.1 Example 1

$$\begin{aligned} v &= [0,0,0,0,0,1,1,1,1,1], \\ \text{FPR} &= [1.0,0.8,0.6,0.4,0.2,0.0,0.0,0.0,0.0,0.0,0], \\ \text{TPR} &= [1.0,1.0,1.0,1.0,1.0,1.0,0.8,0.6,0.4,0.2,0]. \end{aligned}$$

In this example the area above the ROC curve is $A = 0$. Since $\text{sorted}(v) = v$, we have $L = 0 = A$.

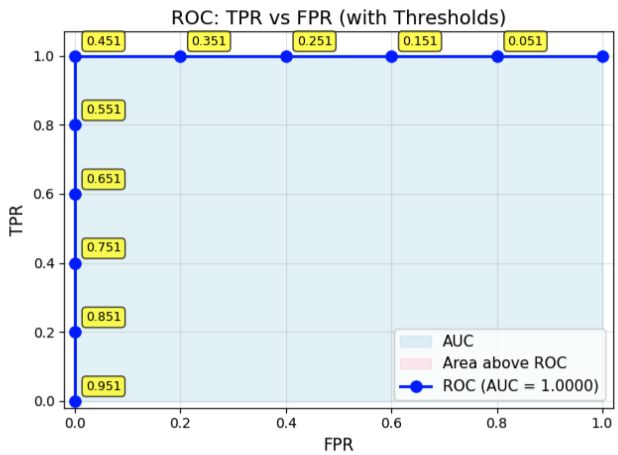

Figure 3: ROC: TPR vs. FPR — perfect predictions.

Now suppose $v = [0,0,0,0,1,0,1,1,1,1]$, i.e., we modify the previous v by swapping one adjacent pair once. Then the shortest path length from this v to $\text{sorted}(v)$ equals 1. In ROC terms, exactly one unit square moves from below the curve to above the curve, so the area above the ROC curve increases by 1 and thus $A = L$.

3.2 Example 2

$$\begin{aligned} v &= [0,0,0,1,0,1,1,0,1,1], \\ \text{FPR} &= [1,0.8,0.6,0.4,0.4,0.2,0.2,0.2,0.0,0.0,0.0], \\ \text{TPR} &= [1,1.0,1.0,1.0,0.8,0.8,0.6,0.4,0.4,0.2,0.0]. \end{aligned}$$

Sorting v using adjacent transpositions requires four swaps. Hence, the shortest path length from $v = [0,0,0,1,0,1,1,0,1,1]$ to $\text{sorted}(v) = [0,0,0,0,0,1,1,1,1,1]$ is $L = 4$. On the other hand, the ROC AUC is 0.84. The maximum possible area is $(n-k)k = 25$ (a 5×5 grid), so the area above the curve is $(1 - 0.84) \cdot 25 = 4$. Thus $A = L$.

One can compute the ROC path (and the shortest path length) directly from the vector v. Start at the lower-left corner of the TPR–FPR plot, i.e. at $(0,0)$. Read v from right to left (starting with the rightmost entry). For each entry: if it is 1, move up by one unit; if it is 0, move right by one unit. The number of

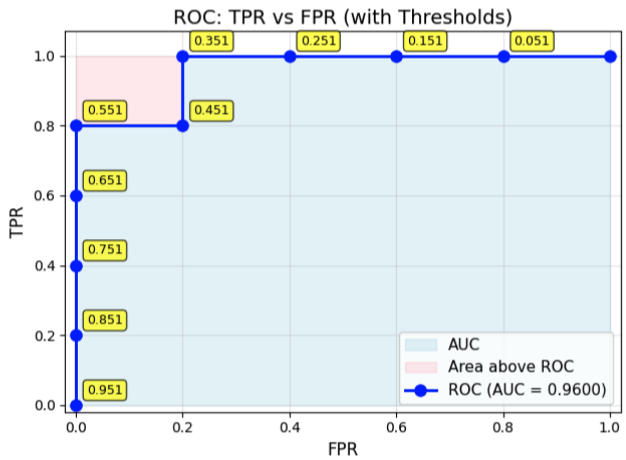

Figure 4: ROC: TPR vs. FPR — one adjacent swap.

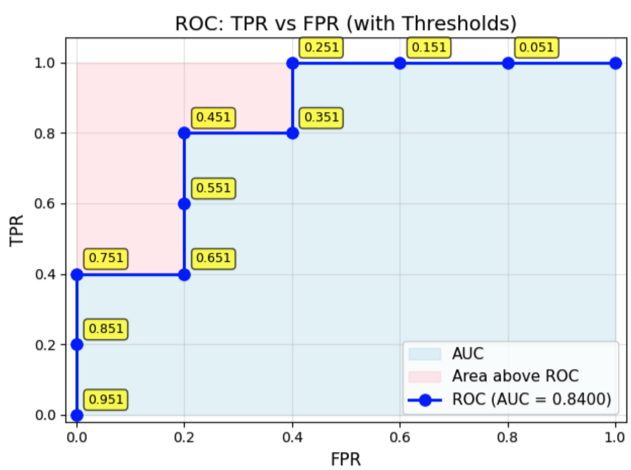

Figure 5: ROC curve showing TPR vs. FPR — four swaps.

unit squares above the resulting path equals the number of inversions in v, i.e. the minimum number of adjacent swaps required to sort it.

Swap	**Transposition**	**Node in** $S(v)$
0	–	$v_0 = v = [0,0,0,1,0,1,1,0,1,1]$
1	$(3,4)$	$v_1 = [0,0,0,0,1,1,1,0,1,1]$
2	$(6,7)$	$v_2 = [0,0,0,0,1,1,0,1,1,1]$
3	$(5,6)$	$v_3 = [0,0,0,0,1,0,1,1,1,1]$
4	$(4,5)$	$v_4 = \text{sorted}(v) = [0,0,0,0,0,1,1,1,1,1]$