1. Proof of Theorem 1

Proof. **Step 1 – the "if" part.** Given a function \( \rho(u) = 1 - \sup\{k \in [0, 1] | \sup_{v \in V_k} (-v^\top u) \leq 0\} \) for some admissible class \( \{V_k\} \), we show that \( \rho(\cdot) \) satisfies all properties required for a CCLF.

Step 1.1 – Complete Classification: If \( u = 0 \), then by \( V_1 = \mathbb{R}_+^m \), we have that \( v^\top u \geq 0 \) for all \( v \in V_1 \), which implies that \( \sup_{v \in V_1} (-v^\top u) \leq 0 \). Hence \( \rho(x) = 0 \). Conversely, if \( u \neq 0 \), without loss of generality we assume \( u_1 < 0 \), then we have \( \sup_{v \in V_1} (-v^\top u) > 0 \).

This, combined with \( V_1 = \text{cl}(\lim_{k \to 1} V_k) \), leads to that \( \exists \delta > 0 \) such that \( \sup_{v \in V_1} (-v^\top u) > 0 \), which implies that \( \rho(u) > 0 \). This shows that \( \rho(\cdot) \) satisfies complete classification.

Step 1.2 – Misclassification avoidance: Fix \( u \) such that \( u < 0 \). We have \( e \in V_0 \) which implies that \( \sup_{v \in V_0} (-v^\top u) \geq (-e^\top u) > 0 \).

Hence \( \rho(u) = 1 \). Thus, \( \rho(\cdot) \) satisfies misclassification avoidance.

Step 1.3 – Monotonicity: If \( u_1 \leq u_2 \), then for any \( k \in [0, 1] \), since \( V_k \subseteq V_1 = \mathbb{R}_+^m \), we have that \( -v^\top u_1 \geq -v^\top u_2 \) for any \( v \in V_k \). Thus,

\[
\{ \sup_{v \in V_k} (-v^\top u_1) \leq 0 \} \implies \{ \sup_{v \in V_k} (-v^\top u_2) \leq 0 \}.
\]

Hence \( \rho(u_1) \geq \rho(u_2) \). Thus, \( \rho(\cdot) \) satisfies monotonicity.

Step 1.4 – Order & scale invariance: Order invariance follows directly from the fact that \( V_k \) is order invariant for all \( k \). Scale invariant holds because for \( \alpha > 0 \) and \( k \in [0, 1] \),

\[
\{ \sup_{v \in V_k} (-v^\top u) \leq 0 \} \iff \{ \sup_{v \in V_k} (-v^\top \alpha u) \leq 0 \}.
\]

Step 1.5 – Quasi-convexity: To show quasi-convexity, let \( c = \max(\rho(u_1), \rho(u_2)) \) and without loss of generality assume \( c < 1 \) since otherwise the claim trivially holds. Thus we have that for any \( \epsilon > 0 \)

\[
\sup_{v \in V_{1-\epsilon}} (-v^\top u_i) \leq 0, \quad i = 1, 2,
\]
which implies that for $\alpha \in [0, 1]$
\[
\sup_{\mathbf{v} \in \mathbf{V}_{1-c, \epsilon}} \{-\mathbf{v}^T [\alpha \mathbf{u}_1 + (1-\alpha)\mathbf{u}_2]\} \leq 0.
\]
Thus $1 - \rho(\alpha \mathbf{u}_1 + (1-\alpha)\mathbf{u}_2) \geq 1 - c$ since $c$ can be arbitrarily close to $0$. The quasi-convexity holds.

Step 2.1 – Lower semi-continuity: We show that $\rho(\mathbf{u}_i) \leq \liminf_i \rho(\mathbf{u}_i)$ for $\mathbf{u}_i \xrightarrow{i} \mathbf{u}_*$. Let $c > \liminf_i \rho(\mathbf{u}_i)$, then there exists an infinite sub-sequence $\{\mathbf{u}_{i_j}\}$ such that $\rho(\mathbf{u}_{i_j}) < c$. That is
\[
-\mathbf{v}^T \mathbf{u}_{i_j} \leq 0; \quad \forall \mathbf{v} \in \mathbf{V}_{1-c}, \forall j.
\]
Note that $\mathbf{u}_{i_j} \rightarrow \mathbf{u}_*$, hence
\[
-\mathbf{v}^T \mathbf{u}_* \leq 0; \quad \forall \mathbf{v} \in \mathbf{V}_{1-c},
\]
i.e., $1 - \rho(\mathbf{u}_*) \geq 1 - c$. Since $c$ can be arbitrarily close to $0$, the semi-continuity follows.

Step 2 – the “only if” part. Given a function $\rho(\cdot)$ which is a CCLF, we show that it can be represented as
\[
\rho(\mathbf{u}) = 1 - \sup \{k \in [0, 1] | \sup_{\mathbf{v} \in \mathbf{V}_k} (-\mathbf{v}^T \mathbf{u}) \leq 0\},
\]
for some admissible class $\{\mathbf{V}_k\}$. This consists of three steps. We first show that $\rho(\cdot)$ can be represented as $\rho(\mathbf{u}) = 1 - \sup \{k \in [0, 1] | \sup_{\mathbf{v} \in \mathbf{V}_k} (-\mathbf{v}^T \mathbf{u}) \leq 0\}$, for some $\{\mathbf{V}_k\}$. Here $\{\mathbf{V}_k\}$ is not necessarily admissible, but satisfies $\mathbf{V}_k \subseteq \mathbf{V}_{k'}$ for all $k \leq k'$. Then we show that we can replace $\mathbf{V}_k$ by a class of closed, convex, order-invariant, cones $\mathbf{V}_k$. Finally we show that $\{\mathbf{V}_k\}$ is admissible to complete the proof.

Step 2.1. The representability of $\rho(\cdot)$ follows from Theorem 2 of (Brown & Sim, 2009). For completeness we re-state the result as a lemma, and provide the proof below.

**Lemma A-1.** Given a CCLF $\rho(\cdot)$, then there exists $\{\mathbf{V}_k\}$ that satisfies $\mathbf{V}_k \subseteq \mathbf{V}_{k'}$ for all $k \leq k'$, such that
\[
\rho(\mathbf{u}) = 1 - \sup \{k \in [0, 1] | \sup_{\mathbf{v} \in \mathbf{V}_k} (-\mathbf{v}^T \mathbf{u}) \leq 0\}.
\]

Step 2.2. We construct $\{\mathbf{V}_k\}$ as follows. Let $\mathbf{\hat{V}}_k \triangleq \text{cl}(\text{or}(\mathbf{V}_k))$. Then we let $\mathbf{V}_k \triangleq \mathbf{\hat{V}}_k$ for $k \in (0, 1)$, and $\mathbf{V}_0 \triangleq \bigcap_{k \in (0, 1)} \mathbf{V}_k$, and $\mathbf{V}_1 \triangleq \text{cl}(\bigcup_{k \in (0, 1)} \mathbf{V}_k)$. Here $\text{or}(\cdot)$ (respectively $\text{cc}(\cdot)$) is the minimal order invariant (respectively, convex cone) superset, defined as
\[
\text{or}(S) = \{P\mathbf{v}|P \in \mathcal{P}_n, \mathbf{v} \in S\}, \quad \text{cc}(S) = \{\sum_{i=1}^k \lambda_i \mathbf{v}_i | k \in \mathbb{N}, \mathbf{v}_i \in S, \lambda_i \geq 0\}.
\]

Let
\[
\rho'(\mathbf{u}) = 1 - \sup \{k \in [0, 1] | \sup_{\mathbf{v} \in \mathbf{V}_k} (-\mathbf{v}^T \mathbf{u}) \leq 0\},
\]
and observe that $\mathbf{V}_k \subseteq \mathbf{\hat{V}}_k$, hence $\rho(\mathbf{u}) \leq \rho'(\mathbf{u})$. To show that $\rho(\mathbf{u}) \geq \rho'(\mathbf{u})$, it suffices to show that for any $k, \epsilon$ and $\mathbf{u}$, the following holds:
\[
\{ \sup_{\mathbf{v} \in \mathbf{V}_k} (-\mathbf{v}^T \mathbf{u}) \leq 0 \} \implies \{ \sup_{\mathbf{v} \in \mathbf{V}_{k-\epsilon}} (-\mathbf{v}^T \mathbf{u}) \leq 0 \}. \tag{A-1}
\]
Note that $\{ \sup_{\mathbf{v} \in \mathbf{V}_k} (-\mathbf{v}^T \mathbf{u}) \leq 0 \}$ implies $k \leq 1 - \rho(\mathbf{u})$, and hence by order invariance of $\rho(\cdot)$, we have $k \leq 1 - \rho(P\mathbf{u})$ for all $P \in \mathcal{P}_n$. This means
\[
\sup_{\mathbf{v} \in \mathbf{V}_{k-\epsilon}} \sup_{P \in \mathcal{P}_n} (-\mathbf{v}^T P\mathbf{u}) \leq 0,
\]
which is equivalent to
\[
\sup_{\mathbf{v} \in \text{or}(\mathbf{V}_{k-\epsilon})} (-\mathbf{v}^T \mathbf{u}) \leq 0.
\]
By definition of $\text{cc}(\cdot)$, this leads to
\[
\sup_{\mathbf{v} \in \text{cc}(\text{or}(\mathbf{V}_{k-\epsilon}))} (-\mathbf{v}^T \mathbf{u}) \leq 0,
\]
which further implies, by continuity of $-v^T u$, that
\[
\sup_{v \in \text{cl}(\text{cc}(\text{or}(\mathbf{V}_{k-1}))))} (-v^T u) \leq 0.
\]
Thus we have $\rho(u) = \rho'(u)$. Finally note that $\hat{V}_k \subseteq \hat{V}_{k'}$ for $k \leq k'$, which leads to the following
\[
\sup_{v \in \mathbf{V}_0} (-v^T u) \leq \sup_{v \in \bigcap_{k \in (0,1)} \hat{V}_k} (-v^T u) \leq \sup_{v \in \mathbf{V}_1} (-v^T u);
\]
\[
\sup_{v \in \mathbf{V}_{1-k}} (-v^T u) \leq \sup_{v \in \bigcup_{k \in (0,1)} \hat{V}_k} (-v^T u) \leq \sup_{v \in \mathbf{V}_1} (-v^T u).
\]
By definitions of $\mathbf{V}_0$ and $\mathbf{V}_1$, together with the fact (due to continuity)
\[
\sup_{v \in \text{cl}(\bigcup_{k \in (0,1)} \hat{V}_k)} (-v^T u) = \sup_{v \in \bigcup_{k \in (0,1)} \hat{V}_k} (-v^T u),
\]
we conclude that
\[
\rho(u) = 1 - \sup_{v \in \mathbf{V}_k} (-v^T u) \leq 0.
\]

Step 2.3. Now we check that $\{\mathbf{V}_k\}$ is indeed admissible. Property 1-3 are straightforward from the definition of $\mathbf{V}_k$. To see that $\mathbf{V}_0$ is closed, recall that the intersection of a class of closes sets is close.

We next show Property 4: $\mathbf{V}_1 = \mathbb{R}^m_+$. By definition of $\mathbf{V}_1$, we have
\[
\lim_{k \to 1} \sup_{v \in \mathbf{V}_k} (-v^T u) = \sup_{v \in \mathbf{V}_1} (-v^T u).
\]
Hence $\rho(u) = 0$ if and only if $\sup_{v \in \mathbf{V}_1} (-v^T u) \leq 0$. Thus by the property of complete classification we have the following
\[
\{ \sup_{v \in \mathbf{V}_1} (-v^T u) \leq 0 \} \iff \{ u \geq 0 \}.
\] (A-2)

Denote the dual cone of a cone $X$ by $X^*$ and recall that for any $k$, $\mathbf{V}_k$ is a closed convex cone, hence we have
\[(\mathbf{V}_1^*)^* = \mathbf{V}_1.
\]
The definition of dual cone states that
\[
\mathbf{V}_1^* = \{ u | u^T v \geq 0; \forall v \in \mathbf{V}_1 \},
\]
which combined with Equation (A-2) implies that
\[
\mathbf{V}_1^* = \mathbb{R}^m_+.
\]
Since $\mathbb{R}^m_+$ is self-dual, we have
\[
\mathbf{V}_1 = \mathbb{R}^m_+.
\]

We now turn to Property 5. Fix $k > 0$. Consider $u = -e$. By misclassification avoidance, $\rho(u^*) = 1$, which means there exists $v^* \in \mathbf{V}_k$ such that $v^T u < 0$, i.e., $\sum_{i=1}^m v_i > 0$. Define a permutation matrix $P \in \mathcal{P}_m$:
\[
P = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1 \\
1 & 0 & 0 & \cdots & 0
\end{bmatrix}.
\]
Thus, by order invariance of $\mathbf{V}_k$, $P^t v^* \in \mathbf{V}_k$ for $t = 0, \cdots, m - 1$. By convexity, this implies $\frac{1}{m} \sum_{t=0}^{m-1} P^t v^* \in \mathbf{V}_k$.
Note that $\frac{1}{m} \sum_{t=0}^{m-1} P^t v^* = [\sum_{i=1}^m v_i^*] e/m$, thus
\[
\frac{1}{m} \sum_{i=1}^m v_i^* e \in \mathbf{V}_k.
\]
Since $\sum_{i=1}^m v_i^* > 0$ and $\mathbf{V}_k$ is a cone, we have $\lambda e \in \mathbf{V}_k$ for all $\lambda \geq 0$ and $k > 0$. By definition of $\mathbf{V}_0$, this implies $\lambda e \in \mathbf{V}_0$. \qed
The rest of this appendix provides a proof to Lemma A-1.

**Proof.** We recall the following results adapted from (Brown & Sim, 2009).

**Definition A-1.** Let $\mathcal{U}$ be the set of random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. A function $\overline{\rho}(\cdot) : \mathcal{U} \to [0,1]$ is a collective satisfying measure if the following holds for all $U, U' \in \mathcal{U}$.

1. If $U \geq 0$, then $\overline{\rho}(U) = 1$;
2. If $U < 0$, then $\overline{\rho}(U) = 0$;
3. If $U \geq U'$ then $\overline{\rho}(U) \geq \overline{\rho}(U')$;
4. $\lim_{\alpha \to 0} \overline{\rho}(U + \alpha) = \overline{\rho}(U)$;
5. If $\lambda \in [0,1]$, then $\overline{\rho}(\lambda U + (1 - \lambda)U') \geq \min(\overline{\rho}(U), \overline{\rho}(U'))$;
6. If $k > 0$, then $\overline{\rho}(kU) = \overline{\rho}(U)$.

**Theorem A-1.** Any collective satisfying measure $\overline{\rho}(\cdot)$ can be represented as

$$\overline{\rho}(U) = \sup\{k \in [0,1] \mid \sup_{Q \in Q_k} \mathbb{E}_Q(-U) \leq 0\},$$

for a class of sets of probability measures $Q_k$ satisfying $Q_k \subseteq Q_{k'}$ for $k \leq k'$.

Given this general result, we focus on a special case where $\Omega = \{1, 2 \cdots, m\}$. Note that in this case each random variable $U : \Omega \to \mathbb{R}$ can be represented as a vector $u \in \mathbb{R}^m$ where $u_i = U(i)$. Given a CCLF $\rho(\cdot) : \mathbb{R}^m \to \mathbb{R}$, we define $\overline{\rho} : \mathcal{U} \to \mathbb{R}$ as following

$$\overline{\rho}(U) = 1 - \rho(u); \quad \text{where } u_i = U(i), \ i = 1, \cdots, m.$$

It is straightforward to check that $\overline{\rho}(\cdot)$ is a collective satisfying measure. Thus, Theorem A-1 states there exists a class of sets of probability measure $Q_k$ such that

$$1 - \rho(u) = \overline{\rho}(U) = \sup\{k \in [0,1] \mid \sup_{Q \in Q_k} \mathbb{E}_Q(-U) \leq 0\}.$$

Note that any probability measure $Q$ on $\Omega = \{1, \cdots, m\}$ can be represented by a vector $v \in \mathbb{R}^m$ such that $v_i = Q(i)$. Thus $\mathbb{E}_Q(-X) = -v^\top x$ where $v$ and $u$ are the vector form for $Q$ and $U$ respectively. Hence we have there exists $V_k$ such that

$$\rho(u) = 1 - \sup\{k \in [0,1] \mid \sup_{v \in V_k} (-v^\top u) \leq 0\}.$$

Note that for $k \leq k'$, $V_k \subseteq V_{k'}$ since $Q_k \subseteq Q_{k'}$. This concludes the proof of Lemma A-1.

2. **Proof of Theorem 2**

**Proof. Claim 1:** We check that all conditions of Definition 1 are satisfied by $\overline{\rho}(\cdot)$. The only condition needs a proof is the semi-continuity. Consider a sequence $u^j \to u^0$, and let $t^0 = \max\{t : \sum_{i=1}^t u^j(i) < 0\}$. Without loss of generality we let $u^0_1 \leq u^0_2 \leq \cdots \leq u^0_m$. Thus we have that $\sum_{i=1}^t u^j(i) < 0$. This implies that $\lim \sup_j \sum_{i=1}^t u^j(i) < 0$, which further leads to $\lim \inf_j \{\max\{t : \sum_{i=1}^t u^j(i) < 0\}\} \geq t^0$. Hence $\lim \inf_j \overline{\rho}(u^j) \geq \overline{\rho}(u^0)$, which established the semi-continuity. Thus, we conclude that $\overline{\rho}(\cdot)$ is a CCLF. Further, observe that $\max\{t : \sum_{i=1}^t u(i) < 0\} \geq \sum_{i=1}^m 1(u_i < 0)$, which established the first claim.

**Claim 2:** It is straightforward to check that $V_k$ satisfies all conditions of Definition 2, and hence is an admissible set. Thus, we proceed to show that $V_k$ is an admissible set corresponding to $\overline{\rho}(\cdot)$, i.e., to show

$$\overline{\rho}(u) = 1 - \sup\{k \in [0,1] \mid \sup_{v \in V_k} (-v^\top u) \leq 0\}.$$
Fix a $\mathbf{u} \in \mathbb{R}^m$. If $\mathbf{u} \geq 0$, then we have $\mathcal{P}(\mathbf{u}) = 0$, as well as $\sup_{\mathbf{v} \in \mathcal{V}_1} (-\mathbf{v}^T \mathbf{u}) \leq 0$, and hence the equivalence holds trivially. Thus we assume $\mathbf{u} \not\geq 0$, and let $t^0 = \max \{ t : \sum_{i=1}^t u(i) < 0 \}$. By definition we have

$$\mathcal{V}_{1-t^0/m} = \text{conv}\{ \lambda \mathbf{e}_N | \lambda > 0, |N| = t^0 + 1 \}.$$ 

Note that by definition of $t^0$

$$\min_{|N| = t^0 + 1} \sum_{i \in N} u_i \geq 0,$$

which implies that

$$\sup_{\mathbf{v} \in \{ \mathbf{e}_N, |N| = t^0 + 1 \}} (-\mathbf{v}^T \mathbf{u}) \leq 0.$$

This leads to

$$\sup_{\mathbf{v} \in \mathcal{V}_{1-t^0/m}} (-\mathbf{v}^T \mathbf{u}) \leq 0. \quad (A-3)$$

On the other hand for arbitrarily small $\epsilon > 0$, by definition

$$\mathcal{V}_{1-t^0/m+\epsilon} = \text{conv}\{ \lambda \mathbf{e}_N | \lambda > 0, |N| = t^0 \}.$$ 

Because $\min_{|N| = t^0} \sum_{i \in N} u_i < 0$, we have

$$\sup_{\mathbf{v} \in \mathcal{V}_{1-t^0/m+\epsilon}} (-\mathbf{v}^T \mathbf{u}) > 0.$$

Combining with Equation $(A-3)$ we established the second claim.

**Claim 3:** Let $\rho'(\cdot)$ be a CCLF satisfying that $\rho'(\mathbf{u}) \geq \rho(\mathbf{u})$ for all $\mathbf{u} \in \mathbb{R}^m$, and let $\{ \mathcal{V}'_k \}$ be its corresponding admissible set. Thus, it suffices to show that $\mathcal{V}_k \subseteq \mathcal{V}'_k$ for all $k$. This holds trivially for $k = 0$, since $\rho'(\mathbf{u}) = 1$ for all $\mathbf{u} < 0$ implies that $\lambda \mathbf{e} \in \mathcal{V}'_0$. When $k > 0$, let $s/m < k \leq (s+1)/m$ for some integer $s$. Then, since $\mathcal{V}'_k$ is an order-invariant convex cone, it suffices to show that $\mathbf{e}_{[1:m-s]} \in \mathcal{V}'_k$ to establish the third claim. Consider $\mathbf{u}^* = -\mathbf{e}_{[1:m-s]}$. Then, by $\rho'(\mathbf{u}^*) \geq \sum_i 1(u_i^* < 0)/m = s/m < k$, we have

$$\sup_{\mathbf{v} \in \mathcal{V}'_k} (-\mathbf{v}^T \mathbf{u}^*) > 0$$

$$\implies \exists \mathbf{v}^* \in \mathcal{V}'_k : \sum_{i=1}^{m-s} v_i^* > 0.$$

Define a permutation matrix $P$:

$$P = \begin{bmatrix} P_1 & 0_{(m-s) \times s} \\ 0_{(m-s) \times s} & 0_{s \times s} \end{bmatrix},$$

where $P_1$ is a $(m-s) \times (m-s)$ matrix:

$$P_1 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Thus, by order invariance of $\mathcal{V}'_k$, $P^t \mathbf{v}^* \in \mathcal{V}'_k$ for $t = 0, \cdots, m-s-1$. By convexity, this implies $\frac{1}{m-s} \sum_{t=0}^{m-s-1} P^t \mathbf{v}^* \in \mathcal{V}'_k$. Note that $\frac{1}{m-s} \sum_{t=0}^{m-s-1} P^t \mathbf{v}^* = [\sum_{i=[1:m-s]} u_i^*] \mathbf{e}_{[1:m-s]} / (m-s)$, thus

$$\frac{\sum_{i=1}^{m-s} v_i^*}{m-s} \mathbf{e}_{[1:m-s]} \in \mathcal{V}'_k.$$

Since $\frac{\sum_{i=1}^{m-s} v_i^*}{m-s}$ is positive, and $\mathcal{V}'_k$ is a cone, we have $\mathbf{e}_{[1:m-s]} \in \mathcal{V}'_k$, which completes the proof. \qed
3. Proof of Theorem 3

Proof. We prove the theorem by constructing such a function \( \rho(\cdot) \). To do this, first consider \( \hat{\rho} : \mathbb{R}^m \mapsto [0, 1] \) defined as

\[
\hat{\rho}(u) = \min_{\gamma \geq 0} \hat{\rho}(u/\gamma).
\]

Then it is easy to check that \( \hat{\rho}(\cdot) \) satisfies complete classification, misclassification avoidance, monotonicity, order invariance, and scale invariance. To see that \( \hat{\rho}(u) \geq \varrho(u) \), note that if \( u \) has \( t \) negative coefficients, than for any \( \gamma > 0 \), \( u/\gamma \) also has \( t \) negative coefficients, which means

\[
\hat{\rho}(u/\gamma) \geq t/m.
\]

Taking minimization over \( \gamma \), we have \( \hat{\rho}(u) \geq \varrho(u) \) holds. Finally, we show quasi-convexity of \( \hat{\rho}(\cdot) \). Fix \( u_1, u_2, \) and \( \alpha \in [0, 1] \), let \( \gamma_1, \gamma_2 \) be \( \epsilon \)-optimal, i.e.,

\[
\hat{\rho}(u_1/\gamma_i) \leq \hat{\rho}(u_i) + \epsilon, \quad i = 1, 2.
\]

Since \( \hat{\rho} \) is quasi-convex, we have

\[
\hat{\rho}(\alpha u_1 + (1 - \alpha)u_2) = \hat{\rho}(\alpha\gamma_1 + (1 - \alpha)\gamma_2) = \hat{\rho}(\alpha\gamma_1 + (1 - \alpha)\gamma_2) \leq \max\{\hat{\rho}(u_1/\gamma_1), \hat{\rho}(u_2/\gamma_2)\}
\]

which implies

\[
\hat{\rho}(\alpha u_1 + (1 - \alpha)u_2) \leq \hat{\rho}(\alpha\gamma_1 + (1 - \alpha)\gamma_2) \leq \max\{\hat{\rho}(u_1/\gamma_1), \hat{\rho}(u_2/\gamma_2)\} \leq \max\{\hat{\rho}(u_1), \hat{\rho}(u_2)\} + \epsilon.
\]

Hence \( \hat{\rho}(\cdot) \) is quasi-convex. Note that the only property that is not satisfied is the semi-continuity. To handle this, define \( \rho : \mathbb{R}^m \mapsto [0, 1] \) as

\[
\rho(u) = \lim_{\epsilon \downarrow 0} \hat{\rho}(u + \epsilon)
\]

Because of monotonicity of \( \hat{\rho}(\cdot) \), \( \rho(\cdot) \) is well-defined. In addition, it can be shown that \( \rho(\cdot) \) is lower-semicontinuous. Complete classification, misclassification avoidance, monotonicity, order invariance, scale invariance, and quasi-convexity all follows easily from the fact that same property holds for \( \hat{\rho}(\cdot) \). Thus, \( \rho(\cdot) \) is a CCLF w.r.t. \( m \). Next, we show that

\[
\rho(u) \geq \varrho(u) \geq \varrho(u + \epsilon).
\]

The first inequality holds due to \( \hat{\rho}(u) \geq \rho(u) \geq \rho(u + \epsilon) \). The second inequality holds because for any \( u \), there exists \( \epsilon > 0 \) small enough such that \( \varrho(u + \epsilon) = \rho(u) \). Thus, taking limit over \( \hat{\rho}(u + \epsilon) \geq \varrho(u + \epsilon) \) establishes the second inequality. Recall that \( \varrho(u) \) is the minimal CCLF, we establish the lemma by

\[
\varrho(u) \leq \rho(u) \leq \rho(u).
\]

4. Proof of Theorem 5

Proof. To prove Theorem 5, we start with establishing the following lemma. Observe that \( \varrho(u) \) only takes value in \( \{0, 1/m, 2/m, \cdots, 1\} \).

Lemma A-2. The level set of Problem (4), i.e., \( U_i \equiv \{(u, w) | \varrho(u) \leq 1 - i/m; \quad f_j(u, w) \leq 0, \forall j \} \) for \( i = 1, \cdots, m \), equals the following

\[
\{(u, w) | 3d - \sum_{i=1}^{m} |d - u_i| \leq (m - i + 1)d; \quad f_j(u, w) \leq 0, \forall j \}.
\]
Proof. From Property 2 of Theorem 2, we have that \( \mathcal{U}_i \) equals to the feasible set of the following program

\[
\sup_{v \in \nabla_{i/m}} (\mathbf{-v}^T \mathbf{u}) \leq 0; \\
f_j(u, w) \leq 0; \quad j = 1, \cdots, n.
\]

Recall that \( \nabla_{i/m} = \text{conv} \{ \lambda e_N | \lambda > 0, |N| = m - i + 1 \} \) we have that \( \sup_{v \in \nabla_{i/m}} (\mathbf{-v}^T \mathbf{u}) \leq 0 \) is equivalent to

\[
\inf_{v: 0 \leq v \leq e, e^T v = m - i + 1} \mathbf{v}^T \mathbf{u} \geq 0,
\]

which left-hand-side by duality theorem is equivalent to the following optimization problem on \((c, d)\)

\[
\text{Maximize: } \sum_{i=1}^{m} c_i + (m - i + 1)d \\
\text{Subject to: } c_i + d \leq u_i \\
\quad c_i \leq 0.
\]

Thus we have \( u \in \mathcal{U}_i \) if and only if there exists \( c, d, \) and \( w \) such that

\[
\mathbf{e}^T \mathbf{c} + (m - i + 1)d \geq 0; \\
\mathbf{c} + de \leq \mathbf{u}; \\
\mathbf{c} \leq \mathbf{0}; \\
f_j(u, w) \leq 0; \quad j = 1, \cdots, n.
\]

Note that this can be further simplified, since optimal \( c_i = -|d - u_i|^+ \), as

\[
\sum_{i=1}^{m} [d - u_i]^+ \leq (m - i + 1)d \\
f_j(u, w) \leq 0; \quad j = 1, \cdots, n.
\]

(A-4)

This establishes the lemma. \( \square \)

Now we turn to prove Theorem 5. Recall the assumption that there are no \( u, w \) such that \( u \geq 0 \), and \( f_j(u, w) \leq 0 \) for all \( j \). Thus any feasible solution to (A-4) must have \( d > 0 \). Hence the feasible set to Problem (A-4) is equivalent to that of

\[
\sum_{i=1}^{m} [1 - u_i/d]^+ \leq (m - i + 1) \\
f_j(u, w) \leq 0; \quad j = 1, \cdots, m.
\]

Thus, finding the optimal solution to Problem (4) is equivalent to solve the following

\[
\text{Minimize: } \sum_{i=1}^{m} [1 - u_i/d]^+ \\
\text{Subject to: } f_j(u, w) \leq 0; \quad j = 1, \cdots, n; \\
d > 0.
\]

(A-5)

By a change of variable where we let \( h = 1/d, s = uh, t = wh \), this is equivalent to

\[
\text{Minimize: } \sum_{i=1}^{m} [1 - s_i]^+ \\
\text{Subject to: } hf_j(s/h, t/h) \leq 0; \quad j = 1, \cdots, n; \\
h > 0.
\]

Hence Theorem 5 is established. \( \square \)
References