APPENDIX A
PROOF OF LEMMA 1

Proof. (Outline) According to the KKT conditions, i.e., Eq.(3-6), we can obtain three cases:

Case 1: $\alpha_i^* = 0, \xi_i^* = 0 \Rightarrow OD(w_i^*, \phi, \boldsymbol{x}_i, \boldsymbol{c}^*)^2 \leq r^{*2}$;

Case 2: $0 < \alpha_i^* < \frac{1}{\nu|X|}, \xi_i^* = 0 \Rightarrow OD(w_i^*, \phi, \boldsymbol{x}_i, \boldsymbol{c}^*)^2 = r^{*2}$;

Case 3: $\alpha_i^* = \frac{1}{\nu|X|}, \xi_i^* \geq 0 \Rightarrow OD(w_i^*, \phi, \boldsymbol{x}_i, \boldsymbol{c}^*)^2 \geq r^{*2}$.

In Case 1, we can observe that $OD(w_i^*, \phi, \boldsymbol{x}_i, \boldsymbol{c}^*) = r^*$ for $\alpha_i^* = 0$. Similarly, in Case 3, it can be obtained that $OD(w_i^*, \phi, \boldsymbol{x}_i, \boldsymbol{c}^*) = r^*$ for $\alpha_i^* = \frac{1}{\nu|X|}$. Hence, it is concluded that α_i^* can take the values 0 and $\frac{1}{\nu|X|}$ when $OD(w_i^*, \phi, \boldsymbol{x}_i, \boldsymbol{c}^*) = r^*$ in Case 2:

$$
\begin{aligned}
U =& \left\{ i : \alpha_i^* = 0,\ OD(w_i^*, \phi, \boldsymbol{x}_i, \boldsymbol{c}^*) < r^* \right\}, \\
R =& \left\{ i : \alpha_i^* \in \left[0, \frac{1}{\nu|X|}\right],\ OD(w_i^*, \phi, \boldsymbol{x}_i, \boldsymbol{c}^*) = r^* \right\}, \\
N =& \left\{ i : \alpha_i^* = \frac{1}{\nu|X|},\ OD(w_i^*, \phi, \boldsymbol{x}_i, \boldsymbol{c}^*) > r^* \right\}.
\end{aligned}
$$

□

APPENDIX B
PROOF OF LEMMA 3

Proof. Let $g(\boldsymbol{\alpha}) = \boldsymbol{\alpha}^\top \boldsymbol{Q} \boldsymbol{\alpha} - \boldsymbol{f}^\top \boldsymbol{\alpha}$ be the matrix form of the objective function in EAMOD($\boldsymbol{\alpha}$)-problem. By Lemma 2, we have

$$
\begin{aligned}
&\langle \nabla g(\boldsymbol{\alpha}_t), \boldsymbol{\alpha}_{t-1} + \boldsymbol{\beta} - \boldsymbol{\alpha}_t \rangle \geq 0, \\
&\langle \nabla g(\boldsymbol{\alpha}_{t-1}), \boldsymbol{\alpha}_t - \boldsymbol{\alpha}_{t-1} \rangle \geq 0,
\end{aligned}
$$

expanding the above two inequalities, then

$$
\begin{aligned}
&2\boldsymbol{\alpha}_t^\top \boldsymbol{Q}\boldsymbol{\alpha}_{t-1} + 2\boldsymbol{\alpha}_{t-1}^\top \boldsymbol{Q}\boldsymbol{\beta} \\
&- 2\boldsymbol{\alpha}_{t-1}^\top \boldsymbol{Q}\boldsymbol{\alpha}_t - \boldsymbol{f}^\top \boldsymbol{\alpha}_{t-1} - \boldsymbol{f}^\top \boldsymbol{\beta} + \boldsymbol{f}^\top \boldsymbol{\alpha}_t \geq 0,
\end{aligned} \tag{11}
$$

$$
2\boldsymbol{\alpha}_{t-1}^\top \boldsymbol{Q}\boldsymbol{\alpha}_t - 2\boldsymbol{\alpha}_{t-1}^\top \boldsymbol{Q}\boldsymbol{\alpha}_{t-1} + \boldsymbol{f}^\top \boldsymbol{\alpha}_{t-1} - \boldsymbol{f}^\top \boldsymbol{\alpha}_t \geq 0. \tag{12}
$$

Adding Eq.(11) and Eq.(12), the inequality below holds

$$
\boldsymbol{\alpha}_t^\top \boldsymbol{Q}\boldsymbol{\alpha}_t - \boldsymbol{\alpha}_t^\top \boldsymbol{Q}(2\boldsymbol{\alpha}_{t-1} + \boldsymbol{\beta}) \leq -\frac{\boldsymbol{f}^\top \boldsymbol{\beta}}{2} - \boldsymbol{\alpha}_{t-1}^\top \boldsymbol{Q}\boldsymbol{\alpha}_{t-1}.
$$

Since $\boldsymbol{c} = \Psi^\top \boldsymbol{\alpha}$, $\boldsymbol{\alpha}_t^\top \boldsymbol{Q}\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_t^\top \Psi\Psi^\top \boldsymbol{\alpha}_t = \boldsymbol{c}_t^\top \boldsymbol{c}_t$ and $\boldsymbol{\alpha}_{t-1}^\top \boldsymbol{Q}\boldsymbol{\alpha}_{t-1} = \boldsymbol{\alpha}_{t-1}^\top \Psi\Psi^\top \boldsymbol{\alpha}_{t-1} = \boldsymbol{c}_{t-1}^\top \boldsymbol{c}_{t-1}$ will be obtained. Let $\boldsymbol{\theta}_{t-1} = \left(2\boldsymbol{c}_{t-1} + \boldsymbol{\beta}^\top \Psi\right)/2$. And then, we have $\boldsymbol{\alpha}_t^\top \boldsymbol{Q}(2\boldsymbol{\alpha}_{t-1} + \boldsymbol{\beta}) = \boldsymbol{c}_t^\top \boldsymbol{\theta}_{t-1}$. Both sides of the inequality are simultaneously added by $\boldsymbol{\theta}_{t-1}^\top \boldsymbol{\theta}_{t-1}$, the inequality is transformed as

$$
\boldsymbol{c}_t^\top \boldsymbol{c}_t - 2\boldsymbol{c}_t^\top \boldsymbol{\theta}_{t-1} + \boldsymbol{\theta}_{t-1}^\top \boldsymbol{\theta}_{t-1} \leq -\frac{\boldsymbol{f}^\top \boldsymbol{\beta}}{2} + \boldsymbol{\theta}_{t-1}^\top \boldsymbol{\theta}_{t-1} - \boldsymbol{c}_{k-1}^\top \boldsymbol{c}_{k-1}.
$$

Finally, we can get $dis(\boldsymbol{c}_t, \boldsymbol{\theta}_{t-1})^2 \leq \zeta^2$. □

APPENDIX C
PROOF OF THEOREM 1

Proof.

$$
\begin{aligned}
&OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{c}_t^*) \\
\leq& OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{\theta}_{t-1}^*) + dis(\boldsymbol{c}_t^*, \boldsymbol{\theta}_{t-1}^*) \\
\leq& OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{\theta}_{t-1}^*) + \zeta,
\end{aligned}
$$

and

$$
\begin{aligned}
&OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{c}_t^*) \\
\geq& OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{\theta}_{t-1}^*) - dis(\boldsymbol{c}_t^*, \boldsymbol{\theta}_{t-1}^*) \\
\geq& OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{\theta}_{t-1}^*) - \zeta.
\end{aligned}
$$

□

APPENDIX D
PROOF OF THEOREM 2

Proof. (Outline) From the property of ν_t, we know that $|R \cup N| \geq \nu_t|X|$, which indicates the bound should separate at least $\lfloor \nu_t|X| \rfloor$ errors. Meanwhile, $|N| \leq \nu_t|X|$, which means the hypersphere should separate at most $\lceil \nu_t|X| \rceil$ errors.

Based on the above analysis, r^* is bounded by the upper and lower boundaries outside the hypersphere. The radius r should fulfill $OD^*_{\lfloor \nu_t|X| \rfloor} \leq r^*$ and $OD^*_{\lceil \nu_t|X| \rceil} \geq r^*$. For each OD_i^* $(i \in [1, |X|])$, OD^{up} and OD^{low} satisfy $OD_i^{low} \leq OD_i^* \leq OD_i^{up}$. Therefore, r satisfies $r^{up} = OD^{up}_{\lceil \nu_t|X| \rceil} \geq r_*$ and $r^{low} = OD^{low}_{\lfloor \nu_t|X| \rfloor} \leq r^*$. □

APPENDIX E
PROOF OF THEOREM 3

Proof. (Outline) According to Lemma 1, for an instance $\boldsymbol{x}_i$, $OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{c}_t^*) < r^* \Rightarrow \alpha_i^{t*} = 0$. Based on Theorems 1 and 2, we have

$$
OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{c}_t^*) \leq OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{\theta}_{t-1}^*) + \zeta.
$$

Therefore, the inequality $OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{c}_t^*) \leq OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{\theta}_{t-1}^*) + \zeta < r^{low} \leq r^*$ can be obtained. If $OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{\theta}_{t-1}^*) + \zeta < r^{low}$, then $OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{c}_t^*) < r^*$. Similarly, we can get $OD(w_i^*\phi, \boldsymbol{x}_i, \boldsymbol{\theta}_{t-1}^*) - \zeta > r^{up}$. □