Appendix for paper "FedCav: Contribution-aware Model Aggregation on Distributed Heterogeneous Data for Federated Learning"

1 PROOF OF THEOREM 2

We assume that for all client j, $f_j(w)$ is convex, and $\forall t \in [0,1]$, we have

$$f_j(tw_1 + (1-t)w_2) \leq tf_j(w_1) + (1-t)f(w_2) \tag{1}$$

Proof. we first note that if we want to prove $F(w)$ is convex, $F(w)$ is required to satisfy all requirements in Theorem 1.

For first requirement of Theorem 1, $dom(F) \subset R^z$ should be a convex set. We have $F(w) = \ln\left(\sum_i^n e^{f_i(w)}\right)$, for $f_j(w) \geq 0$, then we can get $F(w) \geq 0$, the $dom(F) = \{x|x \geq 0\}$, according to the definition of convex set, for any $x_1, x_2 \in dom(F)$ and two real number $\theta_1, \theta_2 \geq 0, \theta_1 + \theta_2 = 1$, we all have $\theta_1 x_1 + \theta_2 x_2 \in dom(F)$, which proves $dom(F)$ is convex set.

For second part, we define $g_1(t) = tF(w_1) + (1-t)F(w_2)$, and $g_2(t) = F(tw_1 + (1-t)w_2)$, then for $g_1(t)$ we get:

$$\begin{aligned} g_1(t) &= t\ln\sum_i^n e^{f_i(w_1)} + (1-t)\ln\sum_i^n e^{f_i(w_2)} \\ &= \ln\left(\sum_i^n e^{f_i(w_1)}\right)^t + \ln\left(\sum_i^n e^{f_i(w_2)}\right)^{(1-t)} \\ &= \ln\left(\left(\sum_i^n e^{f_i(w_1)}\right)^t \left(\sum_i^n e^{f_i(w_2)}\right)^{(1-t)}\right) \end{aligned}$$

For $e^{f_i(w)} \geq 1$, we can get

$$g_1(t) \geq \ln\left(\sum_i^n e^{tf_i(w_1)+(1-t)f_i(w_2)}\right) = g_2(t)$$

according to assumption in Theorem 1, we can easily get

$$f_i(tw_1 + (1-t)w_2) \leq tf_i(w_1) + (1-t)f_i(w_2)$$

so we have

$$tF(w_1) + (1-t)F(w_2) > F(tw_1 + (1-t)w_2) \tag{2}$$

which proves that $F(w)$ satisfies the second requirement. Another simple way is to verify $\nabla_w^2 F(w) \geq 0$, in our mathematical deduction, it still satisfies.

So we conclude that $F(w)$ we defined is convex. □