A PROOFS

Theorem 2. *Uniform federation is modular and therefore satisfies the $2 \cdot c + 1$ egalitarian fairness bound from Theorem 1.*

Proof. In this proof, we will denote the uniform federation error by err^u. Because Definition 3 has five components, this proof will have five sections.

Property 1:
For the first property, we wish to show that (for any federating coalition C), the large player always has strictly lower error than the small player, or:

$$\frac{err_s(C)}{err_l(C)} \leq 1 \quad \Leftrightarrow \quad err_s(C) \geq err_l(C)$$

Using the form of error found in Lemma 1 (Equation 3), we can rewrite this as:

$$\frac{\mu_e}{\sum_{i \in C} n_i} + \sigma^2 \cdot \frac{\sum_{i \in C} n_i^2 - n_s^2 + (\sum_{i \in C} n_i - n_s)^2}{(\sum_{i \in C} n_i)^2}$$

$$\geq \frac{\mu_e}{\sum_{i \in C} n_i} + \sigma^2 \cdot \frac{\sum_{i \in C} n_i^2 - n_\ell^2 + (\sum_{i \in C} n_i - n_\ell)^2}{(\sum_{i \in C} n_i)^2}$$

Cancelling common terms:

$$\sum_{i \in C} n_i^2 - n_s^2 + \left(\sum_{i \in C} n_i - n_s\right)^2 \geq \sum_{i \in C} n_i^2 - n_\ell^2 + \left(\sum_{i \in C} n_i - n_\ell\right)^2$$

$$-n_s^2 + \left(\sum_{i \in C} n_i\right)^2 - 2 \cdot n_s \cdot \left(\sum_{i \in C} n_i\right) + n_s^2 \geq -n_\ell^2 + \left(\sum_{i \in C} n_i\right)^2 - 2 \cdot n_\ell \cdot \left(\sum_{i \in C} n_i\right) + n_\ell^2$$

$$-2 \cdot n_s \cdot \left(\sum_{i \in C} n_i\right) \geq -2 \cdot n_\ell \cdot \left(\sum_{i \in C} n_i\right)$$

$$n_s \leq n_\ell$$

This means that the small player has higher error whenever it has fewer samples than the large player- and strictly higher error whenever it has strictly fewer samples.

Property 2:
For the second property, we wish to show that the worst case ratio of errors occurs in the two-player case. That is,

$$\frac{err_s^u(C)}{err_l^u(C)} \leq \frac{err_s^u(\{n_s, n_\ell\})}{err_l^u(\{n_s, n_\ell\})}$$

In order to prove this, we'll show something stronger. Take any player k (with $k \neq s, l$). Then, we will show that that the derivative of the ratio with respect to the size of player k (n_k) is always negative:

$$\frac{\partial}{\partial n_k} \frac{err_s^u(C)}{err_l^u(C)} < 0$$

For uniform federation, we know that

$$\lim_{n_k \to 0} err_s^u(C) = err_s^u(C \setminus \{n_k\})$$

and similarly for the large player. Then, we can convert C into $\{n_s, n_\ell\}$ by sending the size of every other player to 0: by the result we are trying to prove, this will only ever increase the ratio of their errors.

Next, we will start the proof.

$$\frac{\partial}{\partial n_k} \frac{err_s^u(C)}{err_l^u(C)} = \frac{err_s^u(C)' \cdot err_l^u(C) - err_s^u(C) \cdot err_l^u(C)'}{(err_l^u(C))^2}$$

This is negative whenever:

$$err_s^u(C)' \cdot err_l^u(C) < err_s^u(C) \cdot err_l^u(C)'$$

$$\frac{err_s^u(C)'}{err_s^u(C)} < \frac{err_l^u(C)'}{err_l^u(C)}$$

To show this result, we will calculate the lefthand side (ratio relating to the small player's error) and then show that it is less than the equivalent ratio for the large player. The error of the small player can be written:

$$\frac{\mu_e}{N} + \sigma^2 \frac{\sum_{i \in C} n_i^2 + N^2 - 2N \cdot n_s}{N^2}$$

The derivative with respect to n_k is:

$$-\frac{\mu_e}{N^2} + 2 \cdot \sigma^2 \frac{(n_k + N - n_s) \cdot N - (\sum_{i \in C} n_i^2 + N^2 - 2N \cdot n_s)}{N^3}$$

$$= \frac{-\mu_e \cdot N + 2 \cdot \sigma^2 \cdot \left((n_k + N - n_s) \cdot N - (\sum_{i \in C} n_i^2 + N^2 - 2N \cdot n_s)\right)}{N^3}$$

The term we're interested in is the ratio of the derivative of the small player's error (which we just calculated) to the error of the small player. Note that the error itself can be written as:

$$\frac{\mu_e \cdot N + \sigma^2 \left(\sum_{i \in C} n_i^2 + N^2 - 2N \cdot n_s\right)}{N^2}$$

So the ratio of the derivative to the overall error is:

$$\frac{1}{N} \cdot \frac{-\mu_e \cdot N + 2 \cdot \sigma^2 \cdot \left((n_k + N - n_s) \cdot N - (\sum_{i \in C} n_i^2 + N^2 - 2N \cdot n_s)\right)}{\mu_e \cdot N + \sigma^2 \left(\sum_{i \in C} n_i^2 + N^2 - 2N \cdot n_s\right)}$$

What we would like to show is that the above term is less than the analogous term for the n_ℓ variant, which can be symmetrically written as:

$$\frac{1}{N} \cdot \frac{-\mu_e \cdot N + 2 \cdot \sigma^2 \cdot \left((n_k + N - n_\ell) \cdot N - (\sum_{i \in C} n_i^2 + N^2 - 2N \cdot n_\ell)\right)}{\mu_e \cdot N + \sigma^2 \left(\sum_{i \in C} n_i^2 + N^2 - 2N \cdot n_\ell\right)}$$

This is equivalent to proving:

$$\frac{B + 2\sigma^2 \cdot N \cdot n_s}{A - 2 \cdot \sigma^2 \cdot N \cdot n_s} < \frac{B + 2\sigma^2 \cdot N \cdot n_\ell}{A - 2 \cdot \sigma^2 \cdot N \cdot n_\ell}$$

for

$$A = \mu_e \cdot N + \sigma^2 \cdot (\sum_{i \in C} n_i^2 + N^2) \quad B = -\mu_e \cdot N + 2\sigma^2 \cdot ((n_k + N) \cdot N - (\sum_{i \in C} n_i^2 + N^2))$$

We can cross multiply the inequality to get:

$$(B + 2\sigma^2 \cdot N \cdot n_s) \cdot (A - 2 \cdot \sigma^2 \cdot N \cdot n_\ell) < (B + 2\sigma^2 \cdot N \cdot n_\ell) \cdot (A - 2 \cdot \sigma^2 \cdot N \cdot n_s)$$

$$0 < (B + 2\sigma^2 \cdot N \cdot n_\ell) \cdot (A - 2 \cdot \sigma^2 \cdot N \cdot n_s) - (B + 2\sigma^2 \cdot N \cdot n_s) \cdot (A - 2 \cdot \sigma^2 \cdot N \cdot n_\ell)$$

Which simplifies to:

$$0 < 2 \cdot (A + B) \cdot (n_\ell - n_s) \cdot N \cdot \sigma^2$$

Because we have assumed that $n_\ell > n_s$, this is true if $A+B > 0$. We can evaluate $A+B$ as being:

$$=\mu_e \cdot N + \sigma^2 \cdot (\sum_{i\in C} n_i^2 + N^2) - \mu_e \cdot N + 2\sigma^2 \cdot ((n_k + N) \cdot N - (\sum_{i\in C} n_i^2 + N^2))$$

$$=2\sigma^2 \cdot (n_k + N) \cdot N) - \sigma^2 \cdot \left(\sum_{i\in C} n_i^2 - N^2\right)$$

$$=2\sigma^2 \cdot n_k \cdot N + 2\sigma^2 \cdot N^2 - \sigma^2 \cdot \left(\sum_{i\in C} n_i^2 - N^2\right)$$

$$=2\sigma^2 \cdot n_k \cdot N + \sigma^2 \cdot \left(N^2 - \sum_{i\in C} n_i^2\right)$$

$$>0$$

as desired.

The remaining properties all relate to the ratio of errors for the two-player group $\{n_s, n_\ell\}$. For uniform federation, this ratio can be written as:

$$\frac{err_s^u(\{n_s, n_\ell\})}{err_l^u(\{n_s, n_\ell\})} = \frac{\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_\ell^2}{\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_s^2}$$

Property 3:

This property relates to the derivative of the ratio with respect to the size of the large player (n_ℓ):

$$\frac{\partial}{\partial n_\ell} \frac{\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_\ell^2}{\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_s^2}$$

This derivative is given by:

$$\frac{(\mu_e + 4 \cdot \sigma^2 \cdot n_\ell) \cdot (\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_s^2) - (\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_\ell^2) \cdot \mu_e}{(\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_s^2)^2}$$

The derivative is *positive* whenever:

$$(\mu_e + 4 \cdot \sigma^2 \cdot n_\ell) \cdot (\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_s^2) - (\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_\ell^2) \cdot \mu_e > 0$$

Pulling over terms and expanding gives:

$$\mu_e^2 \cdot (n_s + n_\ell) + 2 \cdot \mu_e \cdot \sigma^2 \cdot n_s^2 + 4 \cdot \sigma^2 \cdot n_\ell \cdot \mu_e \cdot (n_s + n_\ell) + 8 \cdot \sigma^4 \cdot n_\ell \cdot n_s^2 >$$
$$\mu_e^2 \cdot (n_s + n_\ell) + 2\sigma^2 \cdot \mu_e \cdot n_\ell^2$$

Which simplifies to:

$$2 \cdot \mu_e \cdot \sigma^2 \cdot n_s^2 + 2 \cdot \sigma^2 \cdot n_\ell \cdot \mu_e \cdot (2n_s + n_\ell) + 8 \cdot \sigma^4 \cdot n_\ell \cdot n_s^2 > 0$$

Property 4:

This property relates to the derivative of the ratio with respect to the size of the small player (n_s):

$$\frac{\partial}{\partial n_s} \frac{\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_\ell^2}{\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_s^2}$$

This derivative is given by:

$$\frac{\mu_e \cdot (\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_s^2) - (\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_\ell^2) \cdot (\mu_e + 4 \cdot \sigma^2 \cdot n_s)}{(\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_s^2)^2}$$

The derivative is *negative* whenever:

$$\mu_e \cdot (\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_s^2) < (\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_\ell^2) \cdot (\mu_e + 4 \cdot \sigma^2 \cdot n_s)$$

Expanding:

$$\mu_e \cdot (\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_s^2) < \mu_e^2 \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_\ell^2 \cdot \mu_e + 4 \cdot \sigma^2 \cdot n_s \cdot \mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot \mu_e \cdot n_\ell^2$$

Simplifying:

$$0 < 2\sigma^2 \cdot n_\ell^2 \cdot \mu_e + 2 \cdot \sigma^2 \cdot n_s \cdot \mu_e \cdot (n_s + 2n_\ell) + 2\sigma^2 \cdot \mu_e \cdot n_\ell^2$$

which is satisfied.

Property 5:

This can be found by rewriting and applying the limit.

$$\begin{aligned}
\lim_{\frac{n_s}{n_\ell} \to 0} \frac{err_s(\{n_s, n_\ell\})}{err_l(\{n_s, n_\ell\})} &= \lim_{\frac{n_s}{n_\ell} \to 0} \frac{\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_\ell^2}{\mu_e \cdot (n_s + n_\ell) + 2\sigma^2 \cdot n_s^2} \\
&= \lim_{\frac{n_s}{n_\ell} \to 0} \frac{\mu_e \cdot \frac{n_s}{n_\ell} + \mu_e + 2\sigma^2 \cdot n_\ell}{\mu_e \cdot \frac{n_s}{n_\ell} + \mu_e + 2\sigma^2 \cdot n_s \cdot \frac{n_s}{n_\ell}} \\
&= \frac{\mu_e + 2\sigma^2 \cdot n_\ell}{\mu_e} \\
&= \frac{\frac{\mu_e}{n_\ell} + 2\sigma^2}{\frac{\mu_e}{n_\ell}}
\end{aligned}$$

□

Lemma 4. *Consider a set of M federating players, using optimal fine-grained federation. The error player j (with n_j samples) experiences can be given by:*

$$\frac{\frac{\mu_e}{n_j}}{V_j \cdot T} \cdot \left(1 + \sigma^2 \left(T - \frac{1}{V_j}\right)\right)$$

with $V_i = \sigma^2 + \frac{\mu_e}{n_i}$ and $T = \sum_{i=1}^{M} \frac{1}{V_i}$.

Proof. Lemma 4.2 in [5] gives the expected error as:

$$\mu_e \sum_{i=1}^{M} v_{ji}^2 \cdot \frac{1}{n_i} + \left(\sum_{i\neq j} v_{ji}^2 + \left(\sum_{i\neq j} v_{ji}\right)^2\right) \cdot \sigma^2$$

Fine-grained federation requires that $v_{jj} + \sum_{i\neq j} v_{ji} = 1$, so $v_{jj} = 1 - \sum_{i\neq j} v_{ji}$, or $\sum_{i\neq j} v_{ji} = 1 - v_{jj}$. We can then write the error out as:

$$\mu_e \sum_{i=1}^{M} \frac{v_{ji}^2}{n_i} + \left(\sum_{i\neq j} v_{ji}^2 + (1 - v_{jj})^2\right) \cdot \sigma^2$$

Or:

$$\mu_e \cdot \frac{v_{jj}^2}{n_j} + (1 - v_{jj})^2 \cdot \sigma^2 + \sum_{i\neq j} v_{ji}^2 \cdot \left(\frac{\mu_e}{n_i} + \sigma^2\right)$$

We note that this last form has three components (the first involving μ_e and the latter two involving σ^2).

Lemma 7.1 from the same work shows that the optimal v_{ji} weights can be calculated as follows: Define $V_i = \sigma^2 + \frac{\mu_e}{n_i}$. Then, the value of $\{v_{ji}\}$ that minimizes player j's error is:

$$v_{jj} = \frac{1 + \sigma^2 \sum_{i\neq j} \frac{1}{V_i}}{1 + V_j \sum_{i\neq j} \frac{1}{V_i}}$$

$$v_{jk} = \frac{1}{V_k} \cdot \frac{V_j - \sigma^2}{1 + V_j \sum_{i\neq j} \frac{1}{V_i}} \quad k \neq j$$

This analysis follows by plugging in the optimal weights and then simplifying. The first component of the error becomes:

$$\mu_e \cdot \frac{v_{jj}^2}{n_j} = \frac{\mu_e}{n_j} \cdot \frac{(1+\sigma^2 \sum_{i\neq j} \frac{1}{V_i})^2}{(1+V_j \sum_{i\neq j} \frac{1}{V_i})^2}$$

The next component becomes:

$$(1-v_{jj})^2 \cdot \sigma^2 = \sigma^2 \cdot \left(\frac{1+V_j \sum_{i\neq j} \frac{1}{V_i} - 1 - \sigma^2 \sum_{i\neq j} \frac{1}{V_i}}{1+V_j \sum_{i\neq j} \frac{1}{V_i}} \right)^2$$

$$= \sigma^2 \cdot \frac{\frac{\mu_e^2}{n_j^2} \cdot (\sum_{i\neq j} \frac{1}{V_i})^2}{(1+V_j \sum_{i\neq j} \frac{1}{V_i})^2}$$

The last component:

$$\sum_{i\neq j} v_{ji}^2 \cdot \left(\frac{\mu_e}{n_i} + \sigma^2\right) = \sum_{i\neq j} \left(\frac{1}{V_i} \cdot \frac{\frac{\mu_e}{n_j}}{1+V_j \sum_{k\neq j} \frac{1}{V_k}} \right)^2 \cdot V_i$$

$$= \frac{\mu_e^2}{n_j^2} \cdot \sum_{i\neq j} \frac{1}{V_i} \cdot \frac{1}{(1+V_j \sum_{k\neq j} \frac{1}{V_k})^2} = \frac{\mu_e^2}{n_j^2} \cdot \frac{\sum_{i\neq j} \frac{1}{V_i}}{(1+V_j \cdot \sum_{k\neq j} \frac{1}{V_k})^2}$$

When we combine them together, we can pull out a common factor of:

$$\frac{\frac{\mu_e}{n_j}}{(1+V_j \cdot \sum_{i\neq j} \frac{1}{V_i})^2}$$

The other terms become:

$$\left(1+\sigma^2 \sum_{i\neq j} \frac{1}{V_i}\right)^2 + \sigma^2 \cdot \frac{\mu_e}{n_j} \cdot \left(\sum_{i\neq j} \frac{1}{V_i}\right)^2 + \frac{\mu_e}{n_j} \cdot \sum_{i\neq j} \frac{1}{V_i}$$

Combining these together gives:

$$\frac{\frac{\mu_e}{n_j}}{\left(1+V_j \cdot \left(T - \frac{1}{V_j}\right)\right)^2} \cdot \left(\left(1+\sigma^2 \left(T - \frac{1}{V_j}\right)\right)^2 \right.$$

$$\left. + \sigma^2 \cdot \frac{\mu_e}{n_j} \cdot \left(T - \frac{1}{V_j}\right)^2 + \frac{\mu_e}{n_j} \cdot \left(T - \frac{1}{V_j}\right) \right)$$

for $T = \sum_{i=1}^{M} \frac{1}{V_i}$. Next, we begin simplifying, noting that:

$$\frac{\frac{\mu_e}{n_j}}{\left(1+V_j \cdot \left(T - \frac{1}{V_j}\right)\right)^2} = \frac{\frac{\mu_e}{n_j}}{(1+V_j \cdot T - 1)^2} = \frac{\frac{\mu_e}{n_j}}{V_j^2 \cdot T^2}$$

We can also simplify the other coefficient:

$$\left(1+\sigma^2 \left(T - \frac{1}{V_j}\right)\right)^2 + \sigma^2 \cdot \frac{\mu_e}{n_j} \cdot \left(T - \frac{1}{V_j}\right)^2 + \frac{\mu_e}{n_j} \cdot \left(T - \frac{1}{V_j}\right)$$

$$= \left(1+\sigma^2 \left(T - \frac{1}{V_j}\right)\right)^2 + \frac{\mu_e}{n_j} \cdot \left(T - \frac{1}{V_j}\right) \cdot \left(\sigma^2 \cdot \left(T - \frac{1}{V_j}\right) + 1\right)$$

$$= \left(1+\sigma^2 \left(T - \frac{1}{V_j}\right)\right) \cdot \left(1+\sigma^2 \left(T - \frac{1}{V_j}\right) + \frac{\mu_e}{n_j} \cdot \left(T - \frac{1}{V_j}\right)\right)$$

$$= \left(1+\sigma^2 \left(T - \frac{1}{V_j}\right)\right) \cdot \left(1+\left(\sigma^2 + \frac{\mu_e}{n_j}\right) \cdot \left(T - \frac{1}{V_j}\right)\right)$$

$$= \left(1+\sigma^2 \left(T - \frac{1}{V_j}\right)\right) \cdot \left(1+V_j \cdot \left(T - \frac{1}{V_j}\right)\right)$$

$$= \left(1+\sigma^2 \left(T - \frac{1}{V_j}\right)\right) \cdot (1+V_j \cdot T - 1)$$

$$= \left(1+\sigma^2 \left(T - \frac{1}{V_j}\right)\right) \cdot V_j \cdot T$$

Next, we combine the two terms together to get:

$$\frac{\frac{\mu_e}{n_j}}{V_j \cdot T} \cdot \left(1+\sigma^2 \left(T - \frac{1}{V_j}\right)\right)$$

as desired. □

Theorem 3. *Fine-grained federation is modular and therefore satisfies the $2 \cdot c + 1$ egalitarian fairness bound from Theorem 1.*

Proof. In this proof, we will denote fine federation error by err^f. This proof will be similar to that of Lemma ?? and will again have five components corresponding to the five sections of Definition 3.

Property 1:

For the first property, we wish to show that, for any federating coalition, the small player gets higher error than the large player. That is, we wish to show:

$$err_s^f(C) \geq err_l^f(C)$$

From Lemma 4, we know we can write each player's error as:

$$\frac{\mu_e}{n_s \cdot V_s \cdot T} \cdot \left(1+\sigma^2 \left(S + \frac{1}{V_l}\right)\right) \geq \frac{\mu_e}{n_\ell \cdot V_l \cdot T} \cdot \left(1+\sigma^2 \left(S + \frac{1}{V_s}\right)\right)$$

where we have $V_i = \frac{\mu_e}{n_i} + \sigma^2$, $T = \sum_{i \in C} \frac{1}{V_i}$ and $S = T - \frac{1}{V_s} - \frac{1}{V_l}$. We know that $V_i, T > 0$, so we can cancel common terms:

$$\frac{1}{n_s \cdot V_s} \cdot \left(1+\sigma^2 \left(S + \frac{1}{V_l}\right)\right) \geq \frac{1}{n_\ell \cdot V_l} \cdot \left(1+\sigma^2 \left(S + \frac{1}{V_s}\right)\right)$$

$$n_\ell \cdot V_l \left(1+\sigma^2 \left(S + \frac{1}{V_l}\right)\right) \geq n_s \cdot V_s \cdot \left(1+\sigma^2 \left(S + \frac{1}{V_s}\right)\right)$$

$$n_\ell \cdot V_l + \sigma^2 \cdot S \cdot n_\ell \cdot V_l + \sigma^2 \cdot n_\ell \geq n_s \cdot V_s + \sigma^2 \cdot S \cdot n_s \cdot V_s + \sigma^2 \cdot n_s$$

Next, we can plug in for $V_i = \frac{\mu_e}{n_i} + \sigma^2$:

$$\mu_e + n_\ell \cdot \sigma^2 + \sigma^2 \cdot S \cdot \left(\mu_e + n_\ell \cdot \sigma^2\right) + \sigma^2 \cdot n_\ell \geq \mu_e + n_s \cdot \sigma^2 + \sigma^2 \cdot S \cdot \left(\mu_e + n_s \cdot \sigma^2\right)$$

$$2 \cdot n_\ell \cdot \sigma^2 + n_\ell \cdot S \cdot \sigma^4 \geq 2 \cdot n_s \cdot \sigma^2 + S \cdot n_s \cdot \sigma^4$$

$$n_\ell \geq n_s$$

This result tells us that the small player has higher error whenever it has fewer samples than the large player - and strictly higher error whenever it has strictly fewer samples.

Property 2:
For the second property, we again wish to show that the worst case ratio of errors occurs in the two-player case, or

$$\frac{err_s^f(C)}{err_l^f(C)} \le \frac{err_s^f(\{n_s, n_\ell\})}{err_l^f(\{n_s, n_\ell\})}$$

From Lemma 4, we know that we can write the ratio of errors as:

$$\frac{\frac{\mu_e}{n_s \cdot V_s \cdot T} \cdot \left(1 + \sigma^2 \left(S + \frac{1}{V_l}\right)\right)}{\frac{\mu_e}{n_\ell \cdot V_l \cdot T} \cdot \left(1 + \sigma^2 \left(S + \frac{1}{V_s}\right)\right)}$$

where we have $V_i = \frac{\mu_e}{n_i} + \sigma^2$, $T = \sum_{i \in C} \frac{1}{V_i}$ and $S = T - \frac{1}{V_s} - \frac{1}{V_l}$. The ratio of errors can be simplified further:

$$= \frac{n_\ell \cdot V_l \cdot \left(1 + \sigma^2 \left(S + \frac{1}{V_l}\right)\right)}{n_s \cdot V_s \cdot \left(1 + \sigma^2 \left(S + \frac{1}{V_s}\right)\right)} = \frac{n_\ell \cdot V_l + \sigma^2 \cdot S \cdot n_\ell \cdot V_l + n_\ell \cdot \sigma^2}{n_s \cdot V_s + \sigma^2 \cdot S \cdot n_s \cdot V_s + n_s \cdot \sigma^2}$$

Again, we will prove something stronger: that this ratio decreases as S increases. Note that $S = \sum_{i \in C, i \neq s,l} \frac{1}{V_i} = \sum_{i \in C, i \neq s,l} \frac{1}{\frac{\mu_e}{n_i} + 2\sigma^2}$, so as n_i decreases, S decreases as well. We can rewrite the equation as:

$$\frac{a + b \cdot S}{c + d \cdot S}$$

for:

$$a = n_\ell \cdot (V_l + \sigma^2) \quad b = \sigma^2 \cdot n_\ell \cdot V_l$$
$$c = n_s \cdot (V_s + \sigma^2) \quad d = \sigma^2 \cdot n_s \cdot V_s$$

The key question relates to the derivative of the equation with respect to S:

$$\frac{\partial}{\partial S} \frac{a + b \cdot S}{c + d \cdot S} = \frac{b \cdot (c + d \cdot S) - (a + b \cdot S) \cdot d}{(c + dS)^2} = \frac{b \cdot c - a \cdot d}{(c + d \cdot S)^2}$$

This is negative if:

$$b \cdot c < a \cdot d$$

Plugging in for the values gives:

$$\sigma^2 \cdot n_\ell \cdot n_s \cdot V_l \cdot (V_s + \sigma^2) < \sigma^2 \cdot n_\ell \cdot n_s \cdot V_s \cdot (V_l + \sigma^2)$$
$$\sigma^4 \cdot n_\ell \cdot V_l \cdot n_s < \sigma^4 \cdot n_s \cdot n_\ell \cdot V_s$$
$$V_l < V_s$$
$$\frac{\mu_e}{n_\ell} + \sigma^2 < \frac{\mu_e}{n_s} + \sigma^2$$

which is satisfied because $n_\ell > n_s$.

The remaining properties all relate to the ratio of errors for the two-player group $\{n_s, n_\ell\}$. For fine-grained federation, this ratio can be written as:

$$\frac{err_s^f(\{n_s, n_\ell\})}{err_l^f(\{n_s, n_\ell\})} = \frac{n_\ell \cdot V_l \cdot \left(1 + \sigma^2 \cdot \frac{1}{V_l}\right)}{n_s \cdot V_s \cdot \left(1 + \sigma^2 \cdot \frac{1}{V_s}\right)}$$

$$= \frac{n_\ell \cdot V_l + \sigma^2 \cdot n_\ell}{n_s \cdot V_s + \sigma^2 \cdot n_s} = \frac{n_\ell}{n_s} \cdot \frac{2\sigma^2 + \frac{\mu_e}{n_\ell}}{2\sigma^2 + \frac{\mu_e}{n_s}} = \frac{2\sigma^2 \cdot n_\ell + \mu_e}{2\sigma^2 \cdot n_s + \mu_e}$$

The remaining three properties for fine-grained federation are extremely straightforward:

Property 3:
This property relates to the derivative of the ratio of the errors with respect to the size of the large player (n_ℓ):

$$\frac{\partial}{\partial n_\ell} \frac{2\sigma^2 \cdot n_\ell + \mu_e}{2\sigma^2 \cdot n_s + \mu_e} = 2\sigma^2 > 0$$

Property 4:
This property relates to the derivative of the ratio of the errors with respect to the size of the small player (n_s):

$$\frac{\partial}{\partial n_s} \frac{2\sigma^2 \cdot n_\ell + \mu_e}{2\sigma^2 \cdot n_s + \mu_e} = \frac{0 - 2\sigma^2 \cdot (2\sigma^2 \cdot n_\ell + \mu_e)}{(2\sigma^2 \cdot n_s + \mu_e)^2} < 0$$

Property 5:
Finally, this property can be seen by applying the limit:

$$\lim_{\frac{n_s}{n_\ell} \to 0} \frac{err_s(\{n_s, n_\ell\})}{err_l(\{n_s, n_\ell\})} = \lim_{\frac{n_s}{n_\ell} \to 0} \frac{2\sigma^2 \cdot n_\ell + \mu_e}{2\sigma^2 \cdot n_s + \mu_e}$$

$$= \lim_{\frac{n_s}{n_\ell} \to 0} \frac{2\sigma^2 + \frac{\mu_e}{n_\ell}}{2\sigma^2 \cdot \frac{n_s}{n_\ell} + \frac{\mu_e}{n_\ell}} = \frac{\frac{\mu_e}{n_\ell} + 2\sigma^2}{\frac{\mu_e}{n_\ell}}$$

□

THEOREM 4. *Optimal fine-grained federation always has $\alpha \le 1$: the smaller player experiences* lower *error than proportionality would suggest.*

PROOF. As mentioned previously, the expected error for the fine-grained case looks like:

$$\mu_e \sum_{i=1}^{M} \frac{v_{ji}^2}{n_i} + \left(\sum_{i \neq j} v_{ji}^2 + \left(\sum_{i \neq j} v_{ji}\right)^2\right) \cdot \sigma^2$$

We require that $v_{jj} + \sum_{i \neq j} v_{ji} = 1$, so $v_{jj} = 1 - \sum_{i \neq j} v_{ji}$, or $\sum_{i \neq j} v_{ji} = 1 - v_{jj}$. We can then write the error out as:

$$\mu_e \sum_{i=1}^{M} \frac{v_{ji}^2}{n_i} + \left(\sum_{i \neq j} v_{ji}^2 + (1 - v_{jj})^2\right) \cdot \sigma^2$$

Or:

$$\mu_e \cdot \frac{v_{jj}^2}{n_j} + (1 - v_{jj})^2 \cdot \sigma^2 + \sum_{i \neq j} v_{ji}^2 \cdot \left(\frac{\mu_e}{n_i} + \sigma^2\right)$$

For optimal fine-grained federation, we have that the weights are given by:

$$v_{jj} = \frac{1 + \sigma^2 \sum_{i \neq j} \frac{1}{V_i}}{1 + V_j \sum_{i \neq j} \frac{1}{V_i}}$$

$$v_{jk} = \frac{1}{V_k} \cdot \frac{V_j - \sigma^2}{1 + V_j \sum_{i \neq j} \frac{1}{V_i}} \quad k \neq j$$

where we have $V_i = \sigma^2 + \frac{\mu_e}{n_i}$. In order for sub-proportional error ($\alpha \le 1$) to be satisfied, we need:

$$err_s(\Pi) \le err_l(\Pi) \cdot \frac{n_l}{n_s} \quad n_s \le n_\ell$$

$$err_s(\Pi) \cdot n_s \le err_l(\Pi) \cdot n_l \quad n_s \le n_\ell$$

We will start by taking the (weighted) error form for fine-grained federation, plugging in for the v_{jj}, v_{jk} weights, and then simplifying.

The first component of the error (involving μ_e and v_{jj}) becomes:

$$n_j \cdot \mu_e \cdot \frac{v_{jj}^2}{n_j} = \mu_e \cdot \frac{(1 + \sigma^2 \sum_{i \neq j} \frac{1}{V_i})^2}{(1 + V_j \sum_{i \neq j} \frac{1}{V_i})^2}$$

The component involving σ^2 and v_{jj} becomes:

$$\begin{aligned} n_j \cdot (1 - v_{jj})^2 \cdot \sigma^2 &= n_j \cdot \sigma^2 \cdot \left(\frac{1 + V_j \sum_{i \neq j} \frac{1}{V_i} - 1 - \sigma^2 \sum_{i \neq j} \frac{1}{V_i}}{1 + V_j \sum_{i \neq j} \frac{1}{V_i}} \right)^2 \\ &= n_j \cdot \sigma^2 \cdot \frac{\frac{\mu_e^2}{n_j^2} \cdot (\sum_{i \neq j} \frac{1}{V_i})^2}{(1 + V_j \sum_{i \neq j} \frac{1}{V_i})^2} \\ &= \sigma^2 \cdot \frac{\mu_e^2}{n_j} \frac{(\sum_{i \neq j} \frac{1}{V_i})^2}{(1 + V_j \sum_{i \neq j} \frac{1}{V_i})^2} \end{aligned}$$

The last component, involving v_{ji} becomes:

$$\begin{aligned} n_j \cdot \sum_{i \neq j} v_{ji}^2 \cdot \left(\frac{\mu_e}{n_i} + \sigma^2 \right) &= n_j \cdot \sum_{i \neq j} \left(\frac{1}{V_i} \cdot \frac{\frac{\mu_e}{n_j}}{1 + V_j \sum_{k \neq j} \frac{1}{V_k}} \right)^2 \cdot V_i \\ &= \frac{\mu_e^2}{n_j} \cdot \sum_{i \neq j} \frac{1}{V_i} \cdot \frac{1}{(1 + V_j \sum_{k \neq j} \frac{1}{V_k})^2} \\ &= \frac{\mu_e^2}{n_j} \cdot \frac{\sum_{i \neq j} \frac{1}{V_i}}{(1 + V_j \cdot \sum_{k \neq j} \frac{1}{V_k})^2} \end{aligned}$$

Taken together, these components become:

$$\mu_e \cdot \frac{(1 + \sigma^2 \sum_{i \neq j} \frac{1}{V_i})^2}{(1 + V_j \sum_{i \neq j} \frac{1}{V_i})^2} + \sigma^2 \cdot \frac{\mu_e^2}{n_j} \frac{(\sum_{i \neq j} \frac{1}{V_i})^2}{(1 + V_j \sum_{i \neq j} \frac{1}{V_i})^2} + \frac{\mu_e^2}{n_j} \cdot \frac{\sum_{i \neq j} \frac{1}{V_i}}{(1 + V_j \cdot \sum_{k \neq j} \frac{1}{V_k})^2}$$

When we combine them together, we can pull out a common factor to rewrite this as:

$$\frac{\mu_e}{(1 + V_j \cdot \sum_{i \neq j} \frac{1}{V_i})^2} \cdot \left(\left(1 + \sigma^2 \sum_{i \neq j} \frac{1}{V_i} \right)^2 + \sigma^2 \cdot \frac{\mu_e}{n_j} \cdot \left(\sum_{i \neq j} \frac{1}{V_i} \right)^2 + \frac{\mu_e}{n_j} \cdot \sum_{i \neq j} \frac{1}{V_i} \right)$$

We could rewrite this as:

$$\frac{\mu_e}{\left(1 + V_j \cdot \left(T - \frac{1}{V_j}\right)\right)^2} \cdot \left(\left(1 + \sigma^2 \left(T - \frac{1}{V_j} \right) \right)^2 \right.$$

$$\left. + \sigma^2 \cdot \frac{\mu_e}{n_j} \cdot \left(T - \frac{1}{V_j} \right)^2 + \frac{\mu_e}{n_j} \cdot \left(T - \frac{1}{V_j} \right) \right)$$

for $T = \sum_{i=1}^M \frac{1}{V_i}$. Next, we begin simplifying, noting that:

$$\frac{\mu_e}{\left(1 + V_j \cdot \left(T - \frac{1}{V_j}\right)\right)^2} = \frac{\mu_e}{(1 + V_j \cdot T - 1)^2} = \frac{\mu_e}{V_j^2 \cdot T^2}$$

Simplifying the other coefficient gives:

$$\begin{aligned} &\left(1 + \sigma^2 \left(T - \frac{1}{V_j} \right) \right)^2 + \sigma^2 \cdot \frac{\mu_e}{n_j} \cdot \left(T - \frac{1}{V_j} \right)^2 + \frac{\mu_e}{n_j} \cdot \left(T - \frac{1}{V_j} \right) \\ &= \left(1 + \sigma^2 \left(T - \frac{1}{V_j} \right) \right)^2 + \frac{\mu_e}{n_j} \cdot \left(T - \frac{1}{V_j} \right) \cdot \left(\sigma^2 \cdot \left(T - \frac{1}{V_j} \right) + 1 \right) \\ &= \left(1 + \sigma^2 \left(T - \frac{1}{V_j} \right) \right) \cdot \left(1 + \sigma^2 \left(T - \frac{1}{V_j} \right) + \frac{\mu_e}{n_j} \cdot \left(T - \frac{1}{V_j} \right) \right) \\ &= \left(1 + \sigma^2 \left(T - \frac{1}{V_j} \right) \right) \cdot \left(1 + \left(\sigma^2 + \frac{\mu_e}{n_j} \right) \cdot \left(T - \frac{1}{V_j} \right) \right) \\ &= \left(1 + \sigma^2 \left(T - \frac{1}{V_j} \right) \right) \cdot \left(1 + V_j \cdot \left(T - \frac{1}{V_j} \right) \right) \\ &= \left(1 + \sigma^2 \left(T - \frac{1}{V_j} \right) \right) \cdot (1 + V_j \cdot T - 1) \\ &= \left(1 + \sigma^2 \left(T - \frac{1}{V_j} \right) \right) \cdot V_j \cdot T \end{aligned}$$

Thus, both terms together simplify down to:

$$\frac{\mu_e}{V_j \cdot T} \cdot \left(1 + \sigma^2 \left(T - \frac{1}{V_j} \right) \right)$$

Now, that we have simplified the form of the error, we can start looking at sub-proportionality. In this federating group, we are comparing two players $n_s \leq n_\ell$ with

$$V_s = \sigma^2 + \frac{\mu_e}{n_s} \quad V_l = \sigma^2 + \frac{\mu_e}{n_\ell}$$

We will find it useful to rewrite the weighted error of the small player using $T' = T - \frac{1}{V_s} - \frac{1}{V_l}$:

$$\frac{\mu_e}{V_s \cdot T} \cdot \left(1 + \sigma^2 \left(T' + \frac{1}{V_l} \right) \right)$$

And for the large player:

$$\frac{\mu_e}{V_l \cdot T} \cdot \left(1 + \sigma^2 \left(T' + \frac{1}{V_s} \right) \right)$$

What we want to be true is that the small player has a lower weighted error, or:

$$\begin{aligned} \frac{\mu_e}{V_s \cdot T} \cdot \left(1 + \sigma^2 \left(T' + \frac{1}{V_l} \right) \right) &\leq \frac{\mu_e}{V_l \cdot T} \cdot \left(1 + \sigma^2 \left(T' + \frac{1}{V_s} \right) \right) \\ \frac{1}{V_s} \cdot \left(1 + \sigma^2 \left(T' + \frac{1}{V_l} \right) \right) &\leq \frac{1}{V_l} \cdot \left(1 + \sigma^2 \left(T' + \frac{1}{V_s} \right) \right) \\ V_l \cdot \left(1 + \sigma^2 \left(T' + \frac{1}{V_l} \right) \right) &\leq V_s \cdot \left(1 + \sigma^2 \left(T' + \frac{1}{V_s} \right) \right) \\ V_l + \sigma^2 \cdot T' \cdot V_l + \sigma^2 &\leq V_s + \sigma^2 \cdot T' \cdot V_s + \sigma^2 \\ V_l + \sigma^2 \cdot T' \cdot V_l &\leq V_s + \sigma^2 \cdot T' \cdot V_s \\ V_l &\leq V_s \end{aligned}$$

Note that because $V_i = \sigma^2 + \frac{\mu_e}{n_i}$:

$$n_s \leq n_\ell \quad \Rightarrow \quad V_s \geq V_l$$

This means that the overall proof is concluded: using optimal fine-grained federation, sub-proportional error is always satisfied. □

Theorem 5. *Any individually rational coalition using vanilla federation always has $\alpha \leq 1$: the smaller player experiences lower error than proportionality would suggest.*

Proof. We will show this result by proving the contrapositive: if $\alpha > 1$, as defined in Definition 2, then at least one player wishes to leave that coalition for local learning (so stability is violated).

In our analysis, we will consider two players $n_s < n_\ell$, though the players could be arbitrarily situated with respect to $\frac{\mu_e}{\sigma^2}$. We will assume that all of the players are federating together in coalition A. For notational convenience, we will write the federating coalition A of interest as $A = C \cup \{n_\ell\}$, so C refers to all players except n_ℓ.

For later portions of this proof, we will find it useful to rewrite the difference in error between the small and large player:

$$err_S(C \cup \{n_\ell\}) - err_L(C \cup \{n_\ell\}) = 2\sigma^2 \frac{n_\ell - n_s}{N_C + n_\ell}$$

We note that both terms have the same μ_e coefficient, so they differ only in the σ^2 component:

$$\begin{aligned}
&\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2 - n_s^2 + n_\ell^2 + (N_C - n_s + n_\ell)^2}{(N_C + n_\ell)^2} - \sigma^2 \frac{\sum_{i\in C} n_i^2 + N_C^2}{(N_C + n_\ell)^2} \\
=&\sigma^2 \cdot \frac{-n_s^2 + n_\ell^2 + N_C^2 + n_s^2 + n_\ell^2 - 2n_s \cdot N_C + 2n_\ell \cdot N_C - 2n_\ell \cdot n_s - N_C^2}{(N_C + n_\ell)^2} \\
=&\sigma^2 \cdot \frac{2n_\ell^2 - 2n_s \cdot N_C + 2 \cdot n_\ell \cdot N_C - 2n_\ell \cdot n_s}{(N_C + n_\ell)^2}
\end{aligned}$$

Note that we can rewrite the numerator:

$$2(N_C + n_\ell) \cdot (n_\ell - n_s) = 2 \cdot (N_C \cdot n_\ell - n_s \cdot N_C + n_\ell^2 - n_s \cdot n_\ell)$$

So the overall difference reduces to:

$$= \sigma^2 \cdot \frac{2 \cdot (n_\ell - n_s)}{N_C + n_\ell}$$

as desired.

Next, we will begin the formal proof. At a high level, what the proof does is first derive conditions where $\alpha > 1$, and then show that, when these conditions are satisfied, at least one player wishes to leave C for local learning.

$\alpha > 1$ is implied whenever:

$$\frac{err_S(C \cup \{n_\ell\})}{err_L(C \cup \{n_\ell\})} \geq \frac{n_\ell}{n_s}$$

$$err_S(C \cup \{n_\ell\}) \geq \frac{n_\ell}{n_s} \cdot err_L(C \cup \{n_\ell\})$$

(Note that our definition of α is defined with respect to the smaller player.) Our goal will be to get a lower bound on n_ℓ, given that this occurs. Rewriting the subproportionality violation:

$$2\sigma^2 \frac{n_\ell - n_s}{N_C + n_\ell} \geq \left(\frac{n_\ell}{n_s} - 1\right) \cdot err_L(C)$$

Plugging in for the value of the large player's error:

$$2\sigma^2 \frac{n_\ell - n_s}{N_C + n_\ell} \geq \left(\frac{n_\ell}{n_s} - 1\right) \cdot \left(\frac{\mu_e}{N_C + n_\ell} + \sigma^2 \frac{\sum_{i\in C} n_i^2 + N_C^2}{(N_C + n_\ell)^2}\right)$$

Cancelling a $N_C + n_\ell$ in the denominator:

$$2\sigma^2(n_\ell - n_s) \geq \left(\frac{n_\ell}{n_s} - 1\right) \cdot \left(\mu_e + \sigma^2 \frac{\sum_{i\in C} n_i^2 + N_C^2}{N_C + n_\ell}\right)$$

Rearranging a $\frac{1}{n_s}$ in the right hand side:

$$2\sigma^2 \cdot (n_\ell - n_s) \geq (n_\ell - n_s) \cdot \left(\frac{\mu_e}{n_s} + \frac{\sigma^2}{n_s} \frac{\sum_{i\in C} n_i^2 + N_C^2}{N_C + n_\ell}\right)$$

Cancelling $n_\ell - n_s$ from both sides:

$$2\sigma^2 \geq \frac{\mu_e}{n_s} + \frac{\sigma^2}{n_s} \frac{\sum_{i\in C} n_i^2 + N_C^2}{N_C + n_\ell}$$

Bringing over common terms:

$$0 \geq \frac{\mu_e}{n_s} - 2\sigma^2 + \frac{\sigma^2}{n_s} \frac{\sum_{i\in C} n_i^2 + N_C^2}{N_C + n_\ell}$$

Multiplying by $N_C + n_\ell$:

$$0 \geq \frac{\mu_e}{n_s} \cdot (N_C + n_\ell) - 2\sigma^2 \cdot (N_C + n_\ell) + \frac{\sigma^2}{n_s} \left(\sum_{i\in C} n_i^2 + N_C^2\right)$$

Collecting terms around n_ℓ:

$$0 \geq N_C \cdot \left(\frac{\mu_e}{n_s} - 2\sigma^2\right) + \frac{\sigma^2}{n_s} \left(\sum_{i\in C} n_i^2 + N_C^2\right) + \left(\frac{\mu_e}{n_s} - 2\sigma^2\right) \cdot n_\ell$$

Note that if $\frac{\mu_e}{n_s} \geq 2\sigma^2$, then this inequality can never hold (all of the terms on the righthand side are positive, where we need at least one to be negative). So, in order for it to even be possible to have $\alpha \geq 1$, we need $\frac{\mu_e}{n_s} < 2\sigma^2$, or $\frac{\mu_e}{2\sigma^2} < n_s$. Knowing this, we can pull over terms and flip the sign of the inequality:

$$\begin{aligned}
\left(2\sigma^2 \cdot N_C - \frac{\mu_e}{n_s} \cdot N_C - \frac{\sigma^2}{n_s} \cdot \left(\sum_{i\in C} n_i^2 + N_C^2\right)\right) &\geq \left(\frac{\mu_e}{n_s} - 2\sigma^2\right) \cdot n_\ell \\
\frac{2\sigma^2 \cdot N_C - \frac{\mu_e}{n_s} \cdot N_C - \frac{\sigma^2}{n_s} \cdot \left(\sum_{i\in C} n_i^2 + N_C^2\right)}{\frac{\mu_e}{n_s} - 2\sigma^2} &\leq n_\ell \\
\frac{-2\sigma^2 \cdot N_C + \frac{\mu_e}{n_s} \cdot N_C + \frac{\sigma^2}{n_s} \cdot \left(\sum_{i\in C} n_i^2 + N_C^2\right)}{2\sigma^2 - \frac{\mu_e}{n_s}} &\leq n_\ell
\end{aligned}$$

This gives us a lower bound on n_ℓ for when sub-proportionality is violated.

Next, let's look at the condition for when the larger player would choose to leave.

$$err_L(C \cup \{n_\ell\}) \geq err_L(\{n_\ell\})$$

The form of the error in this case is given by:

$$\frac{\mu_e}{N_C + n_\ell} + \sigma^2 \cdot \frac{\sum_{i\in C} n_i^2 + N_C^2}{(N_C + n_\ell)^2} \geq \frac{\mu_e}{n_\ell}$$

Simplifying:

$$\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2 + N_C^2}{(N_C+n_\ell)^2} \geq \frac{\mu_e}{n_\ell} - \frac{\mu_e}{N_C+n_\ell}$$
$$\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2 + N_C^2}{(N_C+n_\ell)^2} \geq \frac{\mu_e}{n_\ell \cdot (N_C+n_\ell)} \cdot (N_C+n_\ell-n_\ell)$$
$$\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2 + N_C^2}{N_C+n_\ell} \geq \frac{\mu_e}{n_\ell} \cdot N_C$$
$$\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2 + N_C^2}{N_C} \geq \frac{\mu_e}{n_\ell} \cdot (N_C+n_\ell)$$
$$\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2 + N_C^2}{N_C^2} \geq \frac{\mu_e}{n_\ell \cdot N_C} \cdot (N_C+n_\ell)$$
$$\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2}{N_C^2} + \sigma^2 \geq \frac{\mu_e \cdot N_S}{n_\ell \cdot N_C} + \frac{\mu_e \cdot n_\ell}{n_\ell \cdot N_C}$$
$$\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2}{N_C^2} + \sigma^2 \geq \frac{\mu_e}{n_\ell} + \frac{\mu_e}{N_C}$$
$$\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2}{N_C^2} + \sigma^2 - \frac{\mu_e}{N_C} \geq \frac{\mu_e}{n_\ell}$$
$$n_\ell \cdot \left(\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2}{N_C^2} + \sigma^2 - \frac{\mu_e}{N_C}\right) \geq \mu_e$$

We have a similar situation to before: if this coefficent is negative, then the inequality can never hold (no large player will ever wish to leave). The coefficient is negative when:

$$\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2}{N_C^2} + \sigma^2 - \frac{\mu_e}{N_C} \leq 0$$

We're going to show that this can never happen so long as $n_s > \frac{\mu_e}{2\sigma^2}$ (necessarily for sub-proportionality to be violated). Rewriting:

$$\sigma^2 \cdot \sum_{i\in C} n_i^2 + \sigma^2 \cdot N_C^2 < \mu_e \cdot N_C$$

For simplicity, we will rewrite with $C' = C \setminus \{n_s\}$. Our term becomes:

$$\sigma^2 \cdot \sum_{i\in C'} n_i^2 + \sigma^2 \cdot n_s^2 + \sigma^2 \cdot (N_{C'} + n_s)^2 < \mu_e \cdot N_{C'} + \mu_e \cdot n_s$$

Expanding:

$$\sigma^2 \cdot \sum_{i\in C'} n_i^2 + \sigma^2 \cdot n_s^2 + \sigma^2 \cdot N_{C'}^2 + \sigma^2 \cdot n_s^2 + 2\sigma^2 \cdot n_s \cdot N_{C'} < \mu_e \cdot N_{C'} + \mu_e \cdot n_s$$
$$\sigma^2 \cdot \sum_{i\in C'} n_i^2 + \sigma^2 \cdot N_{C'}^2 + 2\sigma^2 \cdot n_s \cdot N_{C'} + 2\sigma^2 \cdot n_s^2 < \mu_e \cdot N_{C'} + \mu_e \cdot n_s$$

We can show that this never holds by breaking it into two parts. The first one is to show that:

$$2\sigma^2 \cdot n_s^2 > \mu_e \cdot n_s$$

which holds by the assumption that $n_s > \frac{\mu_e}{2\sigma^2}$. Next, we can show that

$$\sigma^2 \cdot \sum_{i\in C'} n_i^2 + \sigma^2 \cdot N_{C'}^2 + 2\sigma^2 \cdot n_s \cdot N_{C'} > \mu_e \cdot N_{C'}$$

This holds again by the assumption that $n_s > \frac{\mu_e}{2\sigma^2}$. Taken together, this shows that the coefficients on the original equation can never be negative.

Returning to our original equation:

$$n_\ell \cdot \left(\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2}{N_C^2} + \sigma^2 - \frac{\mu_e}{N_C}\right) \geq \mu_e$$

We know that the coefficient can never be negative, so we can rewrite it as:

$$\frac{\mu_e}{\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2}{N_C^2} + \sigma^2 - \frac{\mu_e}{N_C}} \leq n_\ell \tag{7}$$

As a reminder, our condition for $\alpha \geq 1$ is given by:

$$\frac{-2\sigma^2 \cdot N_C + \frac{\mu_e}{n_s} \cdot N_C + \frac{\sigma^2}{n_s} \cdot \left(\sum_{i\in C} n_i^2 + N_C^2\right)}{2\sigma^2 - \frac{\mu_e}{n_s}} \leq n_\ell \tag{8}$$

What we would like to show is that the bound in Equation 7 is lower than the bound in Equation 8: any situation where the n_ℓ term is large enough so that $\alpha > 1$, we know that the large player already wishes to leave.

So, what we want to show is:

$$\frac{\mu_e}{\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2}{N_C^2} + \sigma^2 - \frac{\mu_e}{N_C}} \leq \frac{-2\sigma^2 \cdot N_C + \frac{\mu_e}{n_s} \cdot N_C + \frac{\sigma^2}{n_s} \cdot \left(\sum_{i\in C} n_i^2 + N_C^2\right)}{2\sigma^2 - \frac{\mu_e}{n_s}}$$

Multiply the top and bottom of the righthand side by $\frac{n_s}{N_C}$:

$$\frac{\mu_e}{\sigma^2 \cdot \frac{\sum_{i\in C} n_i^2 + N_C^2}{N_C^2} - \frac{\mu_e}{N_C}} \leq \frac{\mu_e - 2\sigma^2 \cdot n_s + \frac{\sigma^2}{N_C} \cdot \left(\sum_{i\in C} n_i^2 + N_C^2\right)}{2\sigma^2 \cdot \frac{n_s}{N_C} - \frac{\mu_e}{N_C}}$$

Next, we multiply the LHS by N_C^2 on both top and bottom and the RHS by N_C on both top and bottom.

$$\frac{\mu_e \cdot N_C^2}{\sigma^2 \cdot B - \mu_e \cdot N_C} \leq \frac{\mu_e \cdot N_C - 2\sigma^2 \cdot n_s \cdot N_C + \sigma^2 \cdot B}{2\sigma^2 \cdot n_s - \mu_e}$$

Next, we cross multiply. On the lefthand side, we obtain:

$$2\mu_e \cdot N_C^2 \cdot \sigma^2 \cdot n_s - \mu_e^2 \cdot N_C^2$$

On the righthand side:

$$\mu_e \cdot N_C \cdot \sigma^2 \cdot B - \mu_e^2 \cdot N_C^2 - 2(\sigma^2)^2 \cdot n_s \cdot N_C \cdot B + 2\mu_e \cdot \sigma^2 \cdot n_s \cdot N_C^2 + (\sigma^2)^2 \cdot B^2 - \mu_e \cdot \sigma^2 \cdot B \cdot N_C$$
$$= -\mu_e^2 \cdot N_C^2 - 2(\sigma^2)^2 \cdot n_s \cdot N_C \cdot B + 2\mu_e \cdot \sigma^2 \cdot n_s \cdot N_C^2 + (\sigma^2)^2 \cdot B^2$$

Bringing the two sides together, what we wish to show is:

$$2\mu_e \cdot N_C^2 \cdot \sigma^2 \cdot n_s - \mu_e^2 \cdot N_C^2$$
$$\leq -\mu_e^2 \cdot N_C^2 - 2(\sigma^2)^2 \cdot n_s \cdot N_C \cdot B + 2\mu_e \cdot \sigma^2 \cdot n_s \cdot N_C^2 + (\sigma^2)^2 \cdot B^2$$

or:

$$0 \leq -2(\sigma^2)^2 \cdot n_s \cdot N_C \cdot B + (\sigma^2)^2 \cdot B^2$$

Dropping the common (positive) terms of B, $(\sigma^2)^2$, gives:

$$0 \leq -2 \cdot n_s \cdot N_C + B$$
$$2 \cdot n_s \cdot N_C \leq \sum_{i\in C} n_i^2 + N_C^2$$

We'll use the same $C' = C \setminus \{n_s\}$ rewriting as before. This gives us:

$$2 \cdot n_s \cdot N_{C'} + 2 \cdot n_s^2 \leq \sum_{i \in C'} n_i^2 + n_s^2 + (N_{C'} + n_s)^2$$

$$2 \cdot n_s \cdot N_{C'} + 2 \cdot n_s^2 \leq \sum_{i \in C'} n_i^2 + n_s^2 + N_{C'}^2 + n_s^2 + 2n_s \cdot N_C'$$

Cancelling:

$$0 \leq \sum_{i \in C'} n_i^2 + N_{C'}^2$$

This is always satisfied. This tells us that the inequality is satisfied and is strict whenever $C' \neq \emptyset$. If $C' = \emptyset$, then the bound for each becomes:

$$\frac{\mu_e}{2\sigma^2 - \frac{\mu_e}{n_s}} = n_\ell$$

If this equation holds exactly, the exact proportionality holds and the large player gets equal error by federating with n_s or doing local learning. □

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009