

# Technical Report

No Author Given

No Institute Given

## 1 Proof

### 1.1 Proof of Theorem 1

*Proof.* Let  $\Delta \mathbf{x} = \tilde{\mathbf{x}}_j^i - \mathbf{x}_j^i = \text{clamp}(\tilde{\mathbf{y}}_j^i, 0, t) - \text{ReLU}(\mathbf{y}_j^i)$  be the difference of two neurons in QNN and DNN after applying the activation function. Note that, here we use  $\tilde{\mathbf{y}}_j^i$  (resp.  $\mathbf{y}_j^i$ ) to denote the value of neuron before an activation function in QNN (resp. DNN). Let  $\Delta = \tilde{\mathbf{y}}_j^i - \mathbf{y}_j^i$ .

First, considering **Case 1** ( $\text{UB}(S^{in}(\mathbf{x}_j^i)) \leq 0$ ), the neuron of DNN is always deactivated as 0. Hence, the output difference  $\delta_{i,j} = S(\tilde{\mathbf{x}}_j^i)$ .

Next, we consider following cases when the neuron in DNN is always activated, i.e.,  $\text{LB}(S^{in}(\mathbf{x}_j^i)) > 0$ :

- **Case 2-1&2-2** ( $\text{UB}(S^{in}(\tilde{\mathbf{x}}_j^i)) \leq 0$  or  $\text{LB}(S^{in}(\tilde{\mathbf{x}}_j^i)) \geq t$ ): Since the neuron in QNN is always deactivated or clamped to  $t$ , we can get  $\delta_{i,j} = -S^{in}(\mathbf{x}_j^i)$  or  $\delta_{i,j} = t - S^{in}(\mathbf{x}_j^i)$ .
- **Case 2-3** ( $\text{LB}(S^{in}(\tilde{\mathbf{x}}_j^i)) \geq 0$  and  $\text{UB}(S^{in}(\tilde{\mathbf{x}}_j^i)) \leq t$ ): Both of two neurons in QNN and DNN are activated and without clamped. Therefore, we have  $\delta_{i,j} = \delta_{i,j}^{in}$ .
- **Case 2-4** ( $\text{LB}(S^{in}(\tilde{\mathbf{x}}_j^i)) < 0$  and  $0 \leq \text{UB}(S^{in}(\tilde{\mathbf{x}}_j^i)) \leq t$ ): Since  $\tilde{\mathbf{y}}_j^i$  is always smaller than  $t$ , we have  $\Delta \mathbf{x} = \max(\tilde{\mathbf{y}}_j^i, 0) - \mathbf{y}_j^i = \max(\tilde{\mathbf{y}}_j^i - \mathbf{y}_j^i, -\mathbf{y}_j^i)$ . Then,  $\delta = \max(\delta_{i,j}^{in}, -S^{in}(\mathbf{x}_j^i))$ .
- **Case 2-5** ( $0 \leq \text{LB}(S^{in}(\tilde{\mathbf{x}}_j^i)) < t$  and  $\text{UB}(S^{in}(\tilde{\mathbf{x}}_j^i)) > t$ ): Since  $\tilde{\mathbf{y}}_j^i$  is always larger than 0, we have  $\Delta \mathbf{x} = \min(\tilde{\mathbf{y}}_j^i, t) - \mathbf{y}_j^i = \min(\tilde{\mathbf{y}}_j^i - \mathbf{y}_j^i, t - \mathbf{y}_j^i)$ . Hence, we have  $\delta_{i,j} = \min(\delta_{i,j}^{in}, t - S^{in}(\mathbf{x}_j^i))$ .
- **Case 2-6** (Otherwise): By case 2-4 & 2-5, we have  $\Delta \mathbf{x} = \max(\min(\tilde{\mathbf{y}}_j^i, t), 0) - \mathbf{y}_j^i = \max(\min(\Delta, t - \mathbf{y}_j^i), -\mathbf{y}_j^i)$ . Hence, we have  $\delta_{i,j} = \max(\min(\delta_{i,j}^{in}, t - S^{in}(\mathbf{x}_j^i)), -S^{in}(\mathbf{x}_j^i))$ .

Finally, we consider following cases when the neuron in DNN can be either activated or deactivated:

- **Case 3-1&3-2** ( $\text{UB}(S^{in}(\tilde{\mathbf{x}}_j^i)) \leq 0$  or  $\text{LB}(S^{in}(\tilde{\mathbf{x}}_j^i)) \geq t$ ): Similar to above, we have  $\delta_{i,j} = -S(\mathbf{x}_j^i)$  or  $t - S(\mathbf{x}_j^i)$  directly.
- **Case 3-3** ( $\text{LB}(S^{in}(\tilde{\mathbf{x}}_j^i)) \geq 0$  and  $\text{UB}(S^{in}(\tilde{\mathbf{x}}_j^i)) \leq t$ ):  $\Delta \mathbf{x}_j^i = \tilde{\mathbf{y}}_j^i - \max(\mathbf{y}_j^i, 0) = \tilde{\mathbf{y}}_j^i + \min(-\mathbf{y}_j^i, 0) = \min(\tilde{\mathbf{y}}_j^i, \Delta)$ . Then, we have  $\delta_{i,j} = \min(S^{in}(\tilde{\mathbf{x}}_j^i), \delta_{i,j}^{in})$ .
- **Case 3-4** ( $\text{LB}(S^{in}(\tilde{\mathbf{x}}_j^i)) < 0$  and  $0 \leq \text{UB}(S^{in}(\tilde{\mathbf{x}}_j^i)) \leq t$ ): We rewrite  $\Delta \mathbf{x}$  as  $\max(\tilde{\mathbf{y}}_j^i) - \max(\mathbf{y}_j^i)$ :

- If  $\Delta \leq 0$ ,  $\Delta \mathbf{x} = \max(\mathbf{y}_j^i + \Delta) - \max(\mathbf{y}_j^i) \leq 0$ . Then, we have  $\Delta \mathbf{x} = 0$  when  $\mathbf{y}_j^i \leq 0$ , and  $\Delta \mathbf{x} = \max(\mathbf{y}_j^i + \Delta, 0) - \mathbf{y}_j^i = \max(\Delta, -\mathbf{y}_j^i) \leq 0$  when  $\mathbf{y}_j^i \geq 0$ . Therefore, we have  $\max(\text{LB}(\delta_{i,j}^{in}), -\text{UB}(S^{in}(\mathbf{x}_j^i))) \leq \Delta \mathbf{x} \leq 0$
- If  $\Delta \geq 0$ ,  $\Delta \mathbf{x} = \max(\tilde{\mathbf{y}}_j^i) - \max(\tilde{\mathbf{y}}_j^i - \Delta) \geq 0$ . Then, we have  $\Delta \mathbf{x} = 0$  when  $\tilde{\mathbf{y}}_j^i \leq 0$ , and  $\Delta \mathbf{x} = \tilde{\mathbf{y}}_j^i - \max(\tilde{\mathbf{y}}_j^i - \Delta, 0) = \tilde{\mathbf{y}}_j^i + \min(\Delta - \tilde{\mathbf{y}}_j^i, 0) = \min(\Delta, \tilde{\mathbf{y}}_j^i) \geq 0$  when  $\mathbf{y}_j^i \geq 0$ . Therefore, we have  $0 \leq \Delta \mathbf{x} \leq \min(\text{UB}(\delta_{i,j}^{in}), \text{UB}(S^{in}(\tilde{\mathbf{x}}_j^i)))$

Therefore, we have  $\text{LB}(\delta_{i,j}) = 0$  if  $\text{LB}(\delta_{i,j}^{in}) \geq 0$ , and  $\max(\text{LB}(\delta_{i,j}^{in}), -\text{UB}(S^{in}(\mathbf{x}_j^i)))$  otherwise.  $\text{UB}(\delta_{i,j}) = 0$  if  $\text{UB}(\delta_{i,j}^{in}) \leq 0$ , and  $\min(\text{UB}(\delta_{i,j}^{in}), \text{UB}(S^{in}(\tilde{\mathbf{x}}_j^i)))$  otherwise.

- **Case 3-5** ( $0 \leq \text{LB}(S^{in}(\tilde{\mathbf{x}}_j^i)) \leq t$  and  $\text{UB}(S^{in}(\tilde{\mathbf{x}}_j^i)) > t$ ): We rewrite  $\Delta \mathbf{x}$  as  $\min(\tilde{\mathbf{y}}_j^i, t) - \max(\mathbf{y}_j^i, 0)$ . Then,  $\Delta \mathbf{x} = t - \max(\mathbf{y}_j^i, 0) = \min(t - \mathbf{y}_j^i, t)$  when  $\tilde{\mathbf{y}}_j^i \geq t$ , and  $\Delta \mathbf{x} = \tilde{\mathbf{y}}_j^i - \max(\tilde{\mathbf{y}}_j^i - \Delta, 0) = \tilde{\mathbf{y}}_j^i + \min(\Delta - \tilde{\mathbf{y}}_j^i, 0) = \min(\Delta, \tilde{\mathbf{y}}_j^i)$  when  $\tilde{\mathbf{y}}_j^i \leq t$ . Specifically, when  $\tilde{\mathbf{y}}_j^i \geq t$ :
  - If  $\Delta \leq t$ , then  $\mathbf{y}_j^i = \tilde{\mathbf{y}}_j^i - \Delta \geq 0$ , and we will have  $\Delta \mathbf{x} = t - \mathbf{y}_j^i = t - \tilde{\mathbf{y}}_j^i + \Delta \leq \Delta \leq t$ ;
  - If  $\Delta \geq t$ , then  $\Delta \mathbf{x} = t - \max(\mathbf{y}_j^i, 0) \leq t \leq \Delta$ .
 Then,  $t - \text{UB}(S^{in}(\mathbf{x}_j^i)) \leq \Delta \mathbf{x} \leq \{\Delta, t\}$  for  $\tilde{\mathbf{y}}_j^i \leq t$ , and  $\min(\text{LB}(\delta_{i,j}^{in}), \text{LB}(S^{in}(\tilde{\mathbf{x}}_j^i))) \leq \Delta \mathbf{x} \leq \{\Delta, t\}$  for  $\tilde{\mathbf{y}}_j^i \geq t$ . Finally,  $\text{LB}(\delta_{i,j}) = \min(\text{LB}(\delta_{i,j}^{in}), \text{LB}(S^{in}(\tilde{\mathbf{x}}_j^i)), t - \text{UB}(S^{in}(\mathbf{x}_j^i)))$  and  $\text{UB}(\delta_{i,j}) = \min(\text{UB}(\delta_{i,j}^{in}), t)$ .
- **Case 3-6** (Otherwise):
  - If  $\tilde{\mathbf{y}}_j^i < 0$ ,  $\text{UB}(\delta_{i,j}) = 0$  by case 3-1.
  - If  $\tilde{\mathbf{y}}_j^i \geq 0$ ,  $\text{UB}(\delta_{i,j}) = \min(\text{UB}(\delta_{i,j}^{in}), t)$  by case 3-5.
  - If  $\tilde{\mathbf{y}}_j^i \geq t$ ,  $\text{LB}(\delta_{i,j}) = t - \text{UB}(S^{in}(\mathbf{x}_j^i))$  by case 3-2;
  - If  $\tilde{\mathbf{y}}_j^i \leq t$ ,  $\text{LB}(\delta_{i,j}) = 0$  if  $\text{LB}(\delta_{i,j}^{in}) \geq 0$ , and  $\max(\text{LB}(\delta_{i,j}^{in}), -\text{UB}(S^{in}(\mathbf{x}_j^i)))$  otherwise by case 3-4.

Then, we get the lower bound as

$$\text{LB}(\delta_{i,j}) = \min(t - \text{UB}(S^{in}(\mathbf{x}_j^i)), 0, \max(\text{LB}(\delta_{i,j}^{in}), -\text{UB}(S^{in}(\mathbf{x}_j^i)))),$$

and upper bound as

$$\text{UB}(\delta_{i,j}) = \max(\min(\text{UB}(\delta_{i,j}^{in}), t), 0) = \text{clamp}(\text{UB}(\delta_{i,j}^{in}), 0, t).$$

## 2 Examples

Now we give the illustration examples for Section 3 and Section 4, both of which are based on DNN  $\mathcal{N}_e$  and QNN  $\hat{\mathcal{N}}_e$  given in Fig. 1. The quantization configurations for the weights, output of the input layer and hidden layer are  $\mathcal{C}_w = \langle \pm, 4, 2 \rangle$ ,  $\mathcal{C}_{in} = \langle +, 4, 4 \rangle$  and  $\mathcal{C}_h = \langle +, 4, 2 \rangle$ . Note that, the weights of DNN  $\mathcal{N}_e$  and QNN  $\hat{\mathcal{N}}_e$  used here is little different from that in the paper (i.e., examples in Section 2) for simplicity consideration. The input region for all the examples in this section for QNN  $\hat{\mathcal{N}}_e$  is  $R((9, 6), 3) = \{(x, y) \in \mathbb{Z}^2 \mid 6 \leq x \leq 12, 3 \leq y \leq 9\}$ .

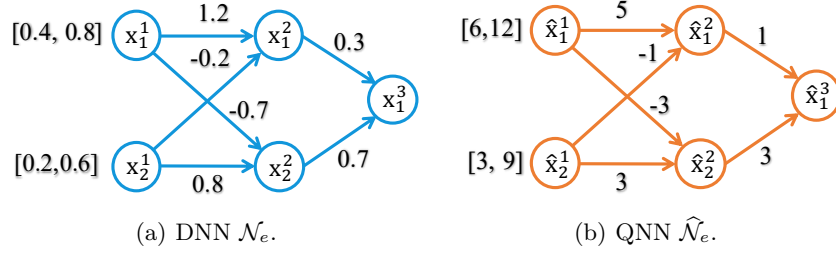


Fig. 1. A 3-layer DNN  $\mathcal{N}_e$  and its quantized version  $\hat{\mathcal{N}}_e$ .

## 2.1 Example for Section 3

*Example 1.* Consider the DNN  $\mathcal{N}_e$  and QNN  $\hat{\mathcal{N}}_e$  given in Fig. 1(a) and Fig. 1(b), where  $2^{-F_{in}} = 1/16$ ,  $\mathcal{C}_h^{ub} = 1/15$ . We set the input region for QNN  $\hat{\mathcal{N}}_e$  as  $R((9, 6), 3) = \{(x, y) \in \mathbb{Z}^2 \mid 6 \leq x \leq 12, 3 \leq y \leq 9\}$ .

By symbolic interval analysis, we have  $S(\mathbf{x}_1^1) = S^{in}(\mathbf{x}_1^1) = [0.4, 0.8]$ ,  $S(\mathbf{x}_2^1) = S^{in}(\mathbf{x}_2^1) = [0.2, 0.6]$ ,  $S(\mathbf{x}_1^2) = S^{in}(\mathbf{x}_1^2) = [0.36, 0.92]$ ,  $S(\mathbf{x}_2^2) = [0, 0.2]$ ,  $S^{in}(\mathbf{x}_2^2) = [-0.4, 0.2]$  for DNN  $\mathcal{N}_e$ , and by interval analysis, we have  $S(\hat{\mathbf{x}}_1^1) = S^{in}(\hat{\mathbf{x}}_1^1) = [6, 12]$ ,  $S(\hat{\mathbf{x}}_2^1) = S^{in}(\hat{\mathbf{x}}_2^1) = [3, 9]$ ,  $S(\hat{\mathbf{x}}_1^2) = S^{in}(\hat{\mathbf{x}}_1^2) = [1, 4]$ ,  $S(\hat{\mathbf{x}}_2^2) = [0, 1]$ ,  $S^{in}(\hat{\mathbf{x}}_2^2) = [-2, 1]$  for QNN  $\hat{\mathcal{N}}_e$ . According to Definition 3 and Algorithm 1, we then have  $\delta_{1,1} = [-0.05, -0.025]$ ,  $\delta_{1,2} = [-0.0375, -0.0125]$ ,  $\Delta \mathbf{b}_1^2 = \Delta \mathbf{b}_2^2 = 0$ , and rounding error  $\xi = 0.125$ .

Then, we get the input difference intervals  $\delta_{2,1}^{in}$  and  $\delta_{2,2}^{in}$  as follows based on Algorithm 2:

- $\text{LB}(\delta_{2,1}^{in}) = 1.25 \times \text{LB}(\delta_{1,1}) + (1.25 - 1.2) \times \text{LB}(S(\mathbf{x}_1^1)) + (-0.25) \times \text{UB}(\delta_{1,2}) + (-0.25 + 0.2) \times \text{UB}(S(\mathbf{x}_2^1)) - \xi = -0.194375$ ;
- $\text{UB}(\delta_{2,1}^{in}) = 1.25 \times \text{UB}(\delta_{1,1}) + (1.25 - 1.2) \times \text{UB}(S(\mathbf{x}_1^1)) + (-0.25) \times \text{LB}(\delta_{1,2}) + (-0.25 + 0.2) \times \text{LB}(S(\mathbf{x}_2^1)) + \xi = 0.133125$ ;
- $\text{LB}(\delta_{2,2}^{in}) = -0.75 \times \text{UB}(\delta_{1,1}) + (-0.75 + 0.7) \times \text{UB}(S(\mathbf{x}_1^1)) + 0.75 \times \text{LB}(\delta_{1,2}) + (0.75 - 0.8) \times \text{UB}(S(\mathbf{x}_2^1)) - \xi = -0.204375$ ;
- $\text{UB}(\delta_{2,2}^{in}) = -0.75 \times \text{LB}(\delta_{1,1}) + (-0.75 + 0.7) \times \text{LB}(S(\mathbf{x}_1^1)) + 0.75 \times \text{UB}(\delta_{1,2}) + (0.75 - 0.8) \times \text{LB}(S(\mathbf{x}_2^1)) + \xi = 0.123125$ ;

According to Algorithm 3, we have

- $\delta_{2,1}^{in} = \delta_{2,1} = [-0.194375, 0.133125]$ ;
- $\delta_{2,2}^{in} = [-0.2, 0.123125]$ :
  - $\text{LB}(\delta_{2,2}) = \max(\text{LB}(\delta_{2,2}^{in}), -\text{UB}(S^{in}(\mathbf{x}_2^2))) = \max(-0.204375, -0.2) = -0.2$ ;
  - $\text{UB}(\delta_{2,2}) = \min(\text{UB}(\delta_{2,2}^{in}), \text{UB}(S^{in}(\hat{\mathbf{x}}_2^2))) = \min(0.123125, 0.25) = 0.123125$ .

Therefore, we get the input/output difference interval  $\delta_{3,1} = \delta_{3,1}^{in}$  as follows:

- $\text{LB}(\delta_{3,1}^{in}) = 0.25 \times \text{LB}(\delta_{2,1}) + (0.25 - 0.3) \times \text{UB}(S(\mathbf{x}_1^2)) + 0.75 \times \text{LB}(\delta_{2,2}) + (0.75 - 0.7) \times \text{LB}(S(\mathbf{x}_2^2)) = -0.24459375$ ;
- $\text{UB}(\delta_{3,1}^{in}) = 0.25 \times \text{UB}(\delta_{2,1}) + (0.25 - 0.3) \times \text{LB}(S(\mathbf{x}_1^2)) + 0.75 \times \text{UB}(\delta_{2,2}) + (0.75 - 0.7) \times \text{UB}(S(\mathbf{x}_2^2)) = 0.117625$ ;

Finally, the quantization error interval is  $[-0.24459375, 0.117625]$ .

## 2.2 Example for Section 4

*Example 2.* Consider the DNN  $\mathcal{N}_e$  and QNN  $\widehat{\mathcal{N}}_e$  given in Fig. 1(a) and Fig. 1(b) again, where  $2^{-F_{in}} = 1/16$ ,  $C_h^{ub} = 1/15$ . We set the input region for QNN  $\widehat{\mathcal{N}}_e$  as  $R((9, 6), 3) = \{(x, y) \in \mathbb{Z}^2 \mid 6 \leq x \leq 12, 3 \leq y \leq 9\}$ .

Different from example 1, we use symbolic-based method given in Section 4 to compute the intervals  $\delta_{2,1}^{in}$  and  $\delta_{2,2}^{in}$ .

Firstly, we get the abstract element  $\mathcal{A}_{j,s}^2 = \langle \mathbf{a}_{j,s}^{2,\leq}, \mathbf{a}_{j,s}^{2,\geq}, l_{j,s}^2, u_{j,s}^2 \rangle$  for  $j \in \{1, 2\}$  and  $s \in \{0, 1\}$  for DNN  $\mathcal{N}_e$  as follows:

- $\mathbf{x}_1^2 \rightarrow \mathbf{x}_{1,0}^2, \mathbf{x}_{1,1}^2$ :
  - $\mathcal{A}_{1,0}^2 = \langle 1.2\mathbf{x}_1^1 - 0.2\mathbf{x}_2^1, 1.2\mathbf{x}_1^1 - 0.2\mathbf{x}_2^1, 0.36, 0.92 \rangle$ ;
  - $\mathcal{A}_{1,1}^2 = \langle \mathbf{x}_{1,0}^2, \mathbf{x}_{1,0}^2, 0.36, 0.92 \rangle$ .
- $\mathbf{x}_2^2 \rightarrow \mathbf{x}_{2,0}^2, \mathbf{x}_{2,1}^2$ :
  - $\mathcal{A}_{2,0}^2 = \langle -0.7\mathbf{x}_1^1 + 0.8\mathbf{x}_2^1, -0.7\mathbf{x}_1^1 + 0.8\mathbf{x}_2^1, -0.4, 0.2 \rangle$ ;
  - $\mathcal{A}_{2,1}^2 = \langle 0, \frac{1}{3}\mathbf{x}_{2,0}^2 + \frac{0.4}{3}, 0, 0.2 \rangle$ .

After substituting every variable in  $\mathbf{a}_{1,1}^{2,\leq}$ ,  $\mathbf{a}_{1,1}^{2,\geq}$ ,  $\mathbf{a}_{2,1}^{2,\leq}$ , and  $\mathbf{a}_{2,1}^{2,\geq}$  until no further substitution is possible, we have the following forms of linear combination of the input variables:

- $\mathcal{A}_{1,0}^{2,*} = \langle 1.2\mathbf{x}_1^1 - 0.2\mathbf{x}_2^1, 1.2\mathbf{x}_1^1 - 0.2\mathbf{x}_2^1, 0.36, 0.92 \rangle$ ;
- $\mathcal{A}_{1,1}^{2,*} = \langle 1.2\mathbf{x}_1^1 - 0.2\mathbf{x}_2^1, 1.2\mathbf{x}_1^1 - 0.2\mathbf{x}_2^1, 0.36, 0.92 \rangle$ ;
- $\mathcal{A}_{2,0}^{2,*} = \langle -0.7\mathbf{x}_1^1 + 0.8\mathbf{x}_2^1, -0.7\mathbf{x}_1^1 + 0.8\mathbf{x}_2^1, -0.4, 0.2 \rangle$ ;
- $\mathcal{A}_{2,1}^{2,*} = \langle 0, -\frac{0.7}{3}\mathbf{x}_1^1 + \frac{0.8}{3}\mathbf{x}_2^1 + \frac{0.4}{3}, 0, 0.2 \rangle$ .

Then, we get the abstract element  $\widehat{\mathcal{A}}_{j,p}^2 = \langle \hat{\mathbf{a}}_{j,p}^{2,\leq}, \hat{\mathbf{a}}_{j,p}^{2,\geq}, \hat{l}_{j,p}^2, \hat{u}_{j,p}^2 \rangle$  for  $j \in \{1, 2\}$  and  $p \in \{0, 1, 2\}$  for QNN  $\widehat{\mathcal{N}}_e$  as follows:

- $\hat{\mathbf{x}}_1^2 \rightarrow \hat{\mathbf{x}}_{1,0}^2, \hat{\mathbf{x}}_{1,1}^2, \hat{\mathbf{x}}_{1,2}^2$ :
  - $\widehat{\mathcal{A}}_{1,0}^2 = \langle \frac{1}{16}(5\hat{\mathbf{x}}_1^1 - \hat{\mathbf{x}}_2^1) - 0.5, \frac{1}{16}(5\hat{\mathbf{x}}_1^1 - \hat{\mathbf{x}}_2^1) + 0.5, 0.8125, 4.0625 \rangle$ ;
  - $\widehat{\mathcal{A}}_{1,1}^2 = \langle \hat{\mathbf{x}}_{1,0}^2, \hat{\mathbf{x}}_{1,0}^2, 0.8125, 4.0625 \rangle$ ;
  - $\widehat{\mathcal{A}}_{1,2}^2 = \langle \hat{\mathbf{x}}_{1,1}^2, \hat{\mathbf{x}}_{1,1}^2, 0.8125, 4.0625 \rangle$ .
- $\hat{\mathbf{x}}_2^2 \rightarrow \hat{\mathbf{x}}_{2,0}^2, \hat{\mathbf{x}}_{2,1}^2, \hat{\mathbf{x}}_{2,2}^2$ :
  - $\widehat{\mathcal{A}}_{2,0}^2 = \langle \frac{1}{16}(-3\hat{\mathbf{x}}_1^1 + 3\hat{\mathbf{x}}_2^1) - 0.5, \frac{1}{16}(-3\hat{\mathbf{x}}_1^1 + 3\hat{\mathbf{x}}_2^1) + 0.5, -2.1875, 1.0625 \rangle$ ;
  - $\widehat{\mathcal{A}}_{2,1}^2 = \langle 0, \frac{17(\hat{\mathbf{x}}_{2,0}^2 + 2.1875)}{52}, 0, 1.0625 \rangle$ ;
  - $\widehat{\mathcal{A}}_{2,2}^2 = \langle \hat{\mathbf{x}}_{2,1}^2, \hat{\mathbf{x}}_{2,1}^2, 0, 1.0625 \rangle$ .

After substituting every variable in  $\hat{\mathbf{a}}_{1,1}^{2,\leq}$ ,  $\hat{\mathbf{a}}_{1,1}^{2,\geq}$ ,  $\hat{\mathbf{a}}_{1,2}^{2,\leq}$ ,  $\hat{\mathbf{a}}_{1,2}^{2,\geq}$ ,  $\hat{\mathbf{a}}_{2,1}^{2,\leq}$ ,  $\hat{\mathbf{a}}_{2,1}^{2,\geq}$ ,  $\hat{\mathbf{a}}_{2,2}^{2,\leq}$ , and  $\hat{\mathbf{a}}_{2,2}^{2,\geq}$  until no further substitution is possible, we have the following forms of linear combination of the input variables:

- $\widehat{\mathcal{A}}_{1,0}^{2,*} = \langle \frac{1}{16}(5\hat{\mathbf{x}}_1^1 - \hat{\mathbf{x}}_2^1) - 0.5, \frac{1}{16}(5\hat{\mathbf{x}}_1^1 - \hat{\mathbf{x}}_2^1) + 0.5, 0.8125, 4.0625 \rangle$ ;
- $\widehat{\mathcal{A}}_{1,1}^{2,*} = \langle \frac{1}{16}(5\hat{\mathbf{x}}_1^1 - \hat{\mathbf{x}}_2^1) - 0.5, \frac{1}{16}(5\hat{\mathbf{x}}_1^1 - \hat{\mathbf{x}}_2^1) + 0.5, 0.8125, 4.0625 \rangle$ ;

$$\begin{aligned}
- \hat{\mathcal{A}}_{1,2}^{2,*} &= \langle \frac{1}{16}(5\hat{\mathbf{x}}_1^1 - \hat{\mathbf{x}}_2^1) - 0.5, \frac{1}{16}(5\hat{\mathbf{x}}_1^1 - \hat{\mathbf{x}}_2^1) + 0.5, 0.8125, 4.0625 \rangle; \\
- \hat{\mathcal{A}}_{2,0}^{2,*} &= \langle \frac{1}{16}(-3\hat{\mathbf{x}}_1^1 + 3\hat{\mathbf{x}}_2^1) - 0.5, \frac{1}{16}(-3\hat{\mathbf{x}}_1^1 + 3\hat{\mathbf{x}}_2^1) + 0.5, -2.1875, 1.0625 \rangle; \\
- \hat{\mathcal{A}}_{2,1}^{2,*} &= \langle 0, \frac{\frac{17}{16}(-3\hat{\mathbf{x}}_1^1 + 3\hat{\mathbf{x}}_2^1) + 45.6875}{52}, 0, 1.0625 \rangle; \\
- \hat{\mathcal{A}}_{2,2}^{2,*} &= \langle 0, \frac{\frac{17}{16}(-3\hat{\mathbf{x}}_1^1 + 3\hat{\mathbf{x}}_2^1) + 45.6875}{52}, 0, 1.0625 \rangle;
\end{aligned}$$

Therefore, we have the lower bounds  $\Delta l_{1,0}^{2,*}$ ,  $\Delta l_{2,0}^{2,*}$  and upper bounds  $\Delta u_{1,0}^{2,*}$ ,  $\Delta u_{2,0}^{2,*}$  of the difference interval  $\delta_{2,1}^{in}$ ,  $\delta_{2,2}^{in}$  for the hidden neurons based on the input region as well as  $\hat{\mathbf{x}}_j^1 = 15\mathbf{x}_j^1$  for  $j \in \{1, 2\}$  as follows:

$$\begin{aligned}
- \Delta l_{1,0}^{2,*} &= 2^{-2}(\hat{\mathbf{a}}_{1,0}^{2,\leq,*}) - \mathbf{a}_{1,0}^{2,\geq,*} = 2^{-2}(\frac{1}{16}(5\hat{\mathbf{x}}_1^1 - \hat{\mathbf{x}}_2^1) - 0.5) - (1.2\mathbf{x}_1^1 - 0.2\mathbf{x}_2^1), \\
&\text{and } \text{LB}(\Delta l_{1,0}^{2,*}) = -0.168125; \\
- \Delta u_{1,0}^{2,*} &= 2^{-2}(\hat{\mathbf{a}}_{1,0}^{2,\geq,*}) - \mathbf{a}_{1,0}^{2,\leq,*} = 2^{-2}(\frac{1}{16}(5\hat{\mathbf{x}}_1^1 - \hat{\mathbf{x}}_2^1) + 0.5) - (1.2\mathbf{x}_1^1 - 0.2\mathbf{x}_2^1), \\
&\text{and } \text{UB}(\Delta u_{1,0}^{2,*}) = 0.106875. \\
- \Delta l_{2,0}^{2,*} &= 2^{-2}(\hat{\mathbf{a}}_{2,0}^{2,\leq,*}) - \mathbf{a}_{2,0}^{2,\geq,*} = 2^{-2}(\frac{1}{16}(-3\hat{\mathbf{x}}_1^1 + 3\hat{\mathbf{x}}_2^1) - 0.5) - (-0.7\mathbf{x}_1^1 + 0.8\mathbf{x}_2^1), \\
&\text{and } \text{LB}(\Delta l_{2,0}^{2,*}) = -0.185625; \\
- \Delta u_{2,0}^{2,*} &= 2^{-2}(\hat{\mathbf{a}}_{2,0}^{2,\geq,*}) - \mathbf{a}_{2,0}^{2,\leq,*} = 2^{-2}(\frac{1}{16}(-3\hat{\mathbf{x}}_1^1 + 3\hat{\mathbf{x}}_2^1) + 0.5) - (-0.7\mathbf{x}_1^1 + 0.8\mathbf{x}_2^1), \\
&\text{and } \text{UB}(\Delta u_{2,0}^{2,*}) = 0.104375.
\end{aligned}$$

Note that, based on above, we can compute  $S^{in}(\hat{\mathbf{x}}_1^2) = [0.8125, 3.0625]$  and  $S^{in}(\hat{\mathbf{x}}_2^2) = [-2.1875, 1.0625]$  via symbolic interval analysis on QNN  $\hat{\mathcal{N}}_e$ . Then, according to Algorithm 3, we have:

$$\begin{aligned}
- \delta_{2,1}^{in} &= \delta_{2,1} = [-0.168125, 0.106875]; \\
- \delta_{2,2}^{in} &= [-0.185625, 0.104375]; \\
&\bullet \text{LB}(\delta_{2,2}) = \max(\text{LB}(\delta_{2,2}^{in}), -\text{UB}(S^{in}(\mathbf{x}_2^2))) = \max(-0.185625, -0.2) = -0.185625; \\
&\bullet \text{UB}(\delta_{2,2}) = \min(\text{UB}(\delta_{2,2}^{in}), \text{UB}(S^{in}(\tilde{\mathbf{x}}_2^2))) = \min(0.104375, 1.0625/4) = 0.104375.
\end{aligned}$$

We remark that for the output layers in DNN  $\mathcal{N}_e$  and QNN  $\hat{\mathcal{N}}_e$ , we also have  $\mathbf{x}_1^3 = \mathbf{x}_{1,0}^3 = \mathbf{x}_{1,1}^2 + 3\mathbf{x}_{2,1}^2$  and  $\hat{\mathbf{x}}_1^3 = \hat{\mathbf{x}}_{1,0}^3 = \hat{\mathbf{x}}_{1,2}^2 + 3\hat{\mathbf{x}}_{2,2}^2$ . Hence, for the output layer, we have:

$$\begin{aligned}
- \mathbf{a}_{1,0}^{3,\leq,*} &= 0.3 \times \mathbf{a}_{1,1}^{2,\leq,*} + 0.7 \times \mathbf{a}_{2,1}^{2,\leq,*}, \mathbf{a}_{1,0}^{3,\geq,*} = 0.3 \times \mathbf{a}_{1,1}^{2,\geq,*} + 0.7 \times \mathbf{a}_{2,1}^{2,\geq,*}; \\
- \hat{\mathbf{a}}_{1,0}^{3,\leq,*} &= 2^{-2}(\hat{\mathbf{a}}_{1,2}^{2,\leq,*} + 3 \times \hat{\mathbf{a}}_{2,2}^{2,\leq,*}), \hat{\mathbf{a}}_{1,0}^{3,\geq,*} = 2^{-2}(\hat{\mathbf{a}}_{1,2}^{2,\geq,*} + 3 \times \hat{\mathbf{a}}_{2,2}^{2,\geq,*}).
\end{aligned}$$

Finally, we get the lower bound  $\Delta l_{1,0}^{3,*}$  and upper bound  $\Delta u_{1,0}^{3,*}$  for the output neurons as follows:

$$\begin{aligned}
- \Delta l_{1,0}^{3,*} &= 2^{-2}(\hat{\mathbf{a}}_{1,0}^{3,\leq,*}) - \mathbf{a}_{1,0}^{3,\geq,*} = 2^{-4}(\hat{\mathbf{a}}_{1,2}^{2,\leq,*} + 3 \times \hat{\mathbf{a}}_{2,2}^{2,\leq,*}) - (0.3 \times \mathbf{a}_{1,1}^{2,\geq,*} + 0.7 \times \mathbf{a}_{2,1}^{2,\geq,*}), \text{ and } \text{LB}(\Delta l_{1,0}^{3,*}) = -0.19721875; \\
- \Delta u_{1,0}^{3,*} &= 2^{-2}(\hat{\mathbf{a}}_{1,0}^{3,\geq,*}) - \mathbf{a}_{1,0}^{3,\leq,*} = 2^{-4}(\hat{\mathbf{a}}_{1,2}^{2,\geq,*} + 3 \times \hat{\mathbf{a}}_{2,2}^{2,\geq,*}) - (0.3 \times \mathbf{a}_{1,1}^{2,\leq,*} + 0.7 \times \mathbf{a}_{2,1}^{2,\leq,*}), \text{ and } \text{UB}(\Delta u_{1,0}^{3,*}) = 0.2045.
\end{aligned}$$

### 2.3 Analysis of Results

Note that, although in the Examples 1 and 2, the quantization error interval computed by `QEBVerif (sym)` is looser than that by `QEBVerif (Con)`, `QEBVerif` performs better on computing the difference intervals for the hidden neurons, and such a comparison result is quite similar to the cases of P1-8 and P1-10 in Table 2. However, we remark that `QEBVerif (sym)` works better in most cases, especially when there are more than 1 hidden layer, as shown in Table 2.