Theorem .1. *Let H' denote the transformed hidden layer matrix of an RdNN, computed as $H' = \phi(XW' + B)$, where $X \in \mathbb{R}^{m\times d}$ is the input matrix, $W' \in \mathbb{R}^{d\times h}$ is the regulated random weight matrix with spectral norm $\|W'\|_2 = \rho_{desired}$, and $\phi(\cdot)$ is a Lipschitz continuous activation function with Lipschitz constant L_ϕ. The bias matrix $B \in \mathbb{R}^{m\times h}$ consists of replicated rows of a random row vector $b \in \mathbb{R}^{1\times h}$, i.e., $B = \mathbf{1}_m b$. Then, the spectral norm of H' is bounded as:*

$$\|H'\|_2 \leq L_\phi \cdot (\|X\|_2 \cdot \rho_{desired} + \sqrt{m} \cdot \|b\|_2).$$

Proof. The hidden layer matrix H' is given by $H' = \phi(XW' + B)$. To analyze the spectral norm of H', we begin with its definition as:

$$\|H'\|_2 = \|\phi(XW' + B)\|_2. \tag{1}$$

Since, $\phi(\cdot)$ is Lipschitz continuous with constant L_ϕ. Therefore:

$$\|\phi(A)\|_2 \leq L_\phi \cdot \|A\|_2, \tag{2}$$

for any matrix A. Applying this property to $\phi(XW' + B)$, we obtain:

$$\|\phi(XW' + B)\|_2 \leq L_\phi \cdot \|XW' + B\|_2. \tag{3}$$

Using the triangle inequality:

$$\|XW' + B\|_2 \leq \|XW'\|_2 + \|B\|_2. \tag{4}$$

Using sub-multiplicativity of spectral norm:

$$\|XW'\|_2 \leq \|X\|_2 \cdot \|W'\|_2 = \|X\|_2 \cdot \rho_{\text{desired}}. \tag{5}$$

To bound $\|B\|_2$, recall that $B = \mathbf{1}_m b$. Since B is the product of a column and a row vector, the result is a rank-1 matrix — all rows are scalar multiples of the same vector (i.e., they are linearly dependent). So:

$$\|B\|_2 = \|\mathbf{1}_m b\|_2 = \|\mathbf{1}_m\|_2 \cdot \|b\|_2 = \sqrt{m} \cdot \|b\|_2. \tag{6}$$

Finally, combining all the results from equations (8), (10), (11), (12), and (13) , we obtain:

$$\|H'\|_2 \leq L_\phi \cdot (\|X\|_2 \cdot \rho_{\text{desired}} + \sqrt{m} \cdot \|b\|_2).$$

This completes the proof. □

Corollary .2. *In practical settings where the bias vector b is initialized using uniform random values in $[0, 1]$ or $[-1, 1]$ scaled by a small constant s (e.g., $s = 1$), the norm $\|b\|_2$ remains bounded. For instance, $\|b\|_2 \leq \sqrt{h}$. Therefore, the term $\sqrt{m} \cdot \|b\|_2 \leq \sqrt{mh}$, and the upper bound becomes:*

$$\|H'\|_2 \leq L_\phi(\|X\|_2 \cdot \rho_{desired} + \sqrt{mh}).$$

This justifies that, under standard initialization practices, the impact of B on the spectral norm remains controlled and the SNoRe framework retains its stabilizing properties.

Corollary .3. *For commonly used activation functions such as ReLU and sigmoid, the Lipschitz constant $L_\phi \leq 1$ (Goodfellow, 2016), which simplifies the bound to:*

$$\|H'\|_2 \leq \|X\|_2 \cdot \rho_{desired} + \sqrt{mh}. \tag{7}$$