

Gradient Norm

2.0

1.5

1.0

0

5000

10000

Steps

Base  
VCD  
VORD

The graph displays the Gradient Norm on the y-axis (ranging from 1.0 to 2.0+) against the number of Steps on the x-axis (ranging from 0 to 10,000). Three models are compared: Base (blue line), VCD (orange line), and VORD (green line). The VCD model starts with a high gradient norm of approximately 1.8, peaks at 2.4 around step 1,000, and then fluctuates between 1.9 and 2.2. The Base model starts at 0.9, peaks at 1.2 around step 1,000, and then fluctuates between 1.0 and 1.2. The VORD model starts at 1.1, peaks at 1.4 around step 1,000, and then fluctuates between 1.1 and 1.4. All three models show an upward trend in gradient norm after 5,000 steps, with VCD reaching the highest value of approximately 2.4 and VORD reaching approximately 1.4 at 10,000 steps.

Steps	Base	VCD	VORD
0	0.9	1.8	1.1
1000	1.2	2.4	1.4
2000	1.1	2.1	1.2
3000	1.0	2.0	1.1
4000	1.1	2.0	1.2
5000	1.1	2.1	1.2
6000	1.2	2.2	1.3
7000	1.2	2.2	1.3
8000	1.2	2.3	1.3
9000	1.2	2.3	1.3
10000	1.3	2.4	1.4