

KL Divergence

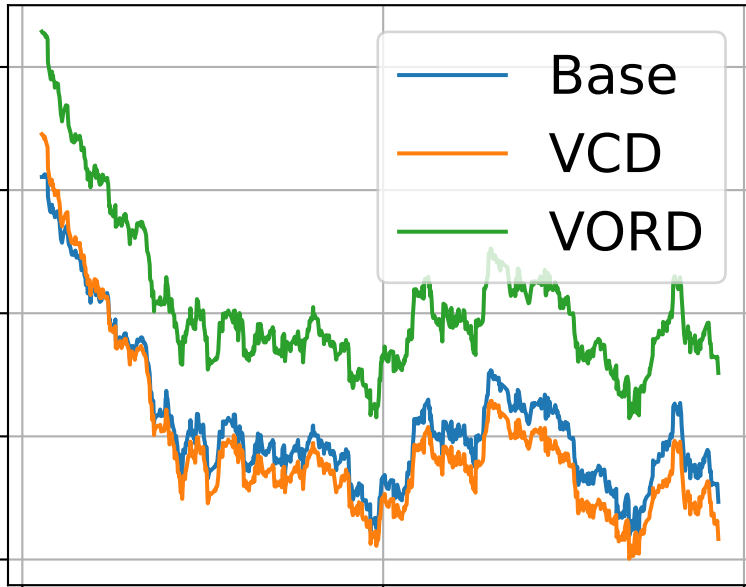
1.8
1.7
1.6
1.5
1.4

0

5000
Steps

10000

Base
VCD
VORD



The graph displays the KL Divergence on the y-axis (ranging from 1.4 to 1.8) 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). All models start with high KL Divergence (around 1.7-1.8) and decrease over time. The VORD model (green) remains the highest throughout, fluctuating between 1.5 and 1.65. The Base (blue) and VCD (orange) models track closely, with VCD generally being slightly lower than Base after the initial steps. Both Base and VCD show significant fluctuations, with VCD reaching a minimum near 1.4 around 8,500 steps.

Steps	Base	VCD	VORD
0	1.71	1.75	1.83
1000	1.58	1.57	1.68
2000	1.50	1.48	1.58
3000	1.48	1.47	1.57
4000	1.49	1.48	1.58
5000	1.46	1.44	1.56
6000	1.51	1.50	1.60
7000	1.53	1.51	1.62
8000	1.47	1.45	1.57
9000	1.53	1.50	1.62
10000	1.46	1.43	1.56