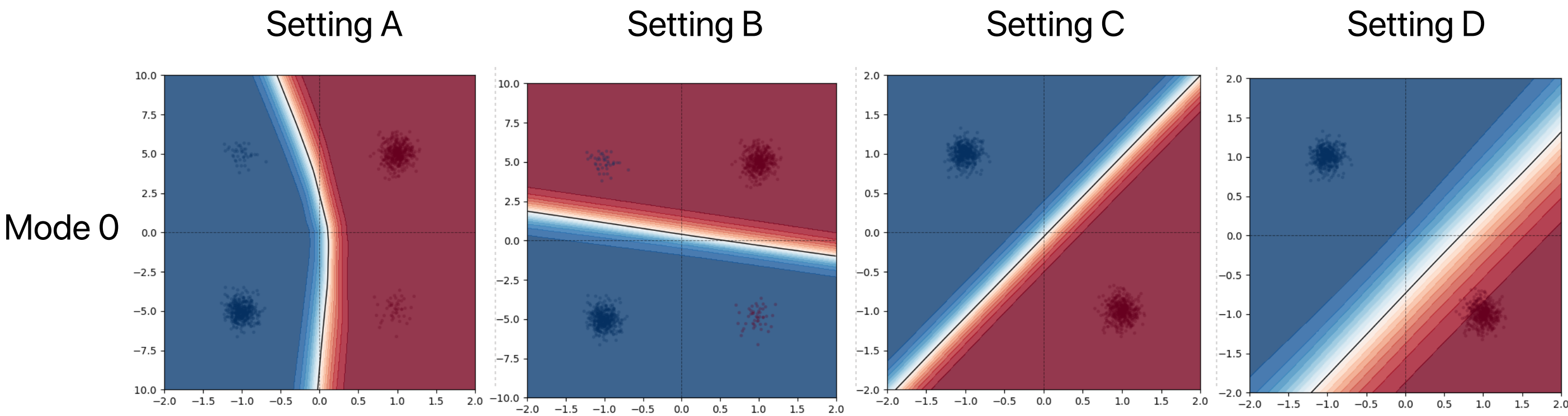


Setups:

	core corr	core gamma	spu corr	spu gamma	noise dim	noise gamma
Setting A	1	1	0.9	5	0	n/a
Setting B	1	1	0.9	5	40	1
Setting C	1	1	0	1	0	n/a
Setting D	1	1	0	1	40	1

Mode 0	Standard rich training; layer-wise representation rank computed
Mode 1	Lazy training; layer-wise representation rank computed
Mode 2	No training; Shred the NTRF matrix into chunks and computing the rank
Mode 3	No training; NTRF rank of separate l -layer networks are computed for $0 \leq l < 6$

Data and the decision boundary at Mode 0:



Rank curves:

