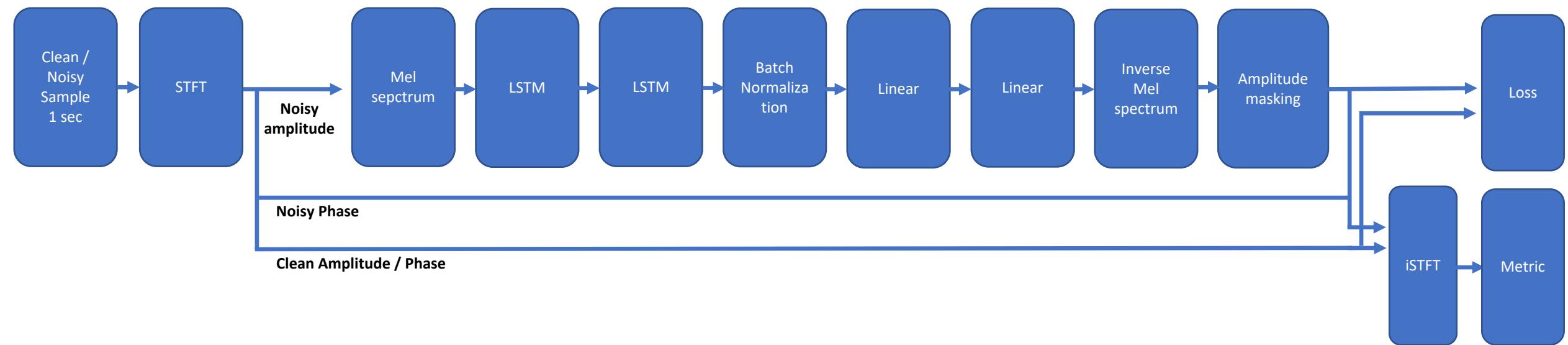


Block Diagram



Default Setting

Total number of dataset	84249 , 1 seg segmentation
Normalization	z-score
Segmentation Normalization	Entire wav file
FFT normalized by fft size	True
Loss function	Phase sensitive approximation
Dense normalized method	glorot_uniform
Batch size	512
LSTM layer	256
Batch Normalization after LSTM layer	True

metric: loss function, SISDR

every epoch consists of 1000 step filled with dataset

- If 2048 samples is divided by 64/256 batch, dataset can be divided as 8, 4 case.
- Then it will filled until 1000 steps
- After "preprocess include stft" and "process stft during training", it keeps to use time sample data, not stft

Result

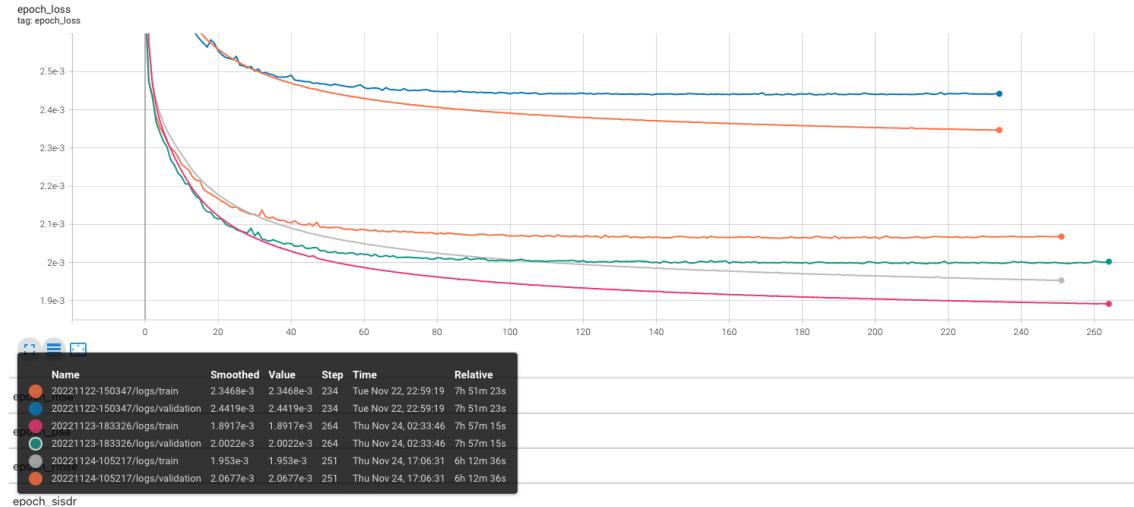
		Loss Function				Dataset Split ratio(Train : Valid)				Dense		Time period of Normalization		Batch size		Batch Norm	
		rmse (stft)	mse (stft)	psa	rmse (mag)	90:10	80:20	70:30	50:50	uniform	normal	entire	1 sec	512	64	O	X
Train	Loss at Convex (1e-3)	2.3468	0.1329	1.8917	1.953	1.8917	1.9003	1.8947	1.9023	1.8917	1.8963	1.8917	1.85544	1.8917	1.904	1.8917	1.9194
	SISDR at Convex	14.96	14.51	14.98	14.87	14.98	14.95	14.97	14.9	14.98	14.96	14.98	14.93	14.98	14.73	14.98	14.89
Valid	Loss at Convex (1e-3)	2.4419	0.1390	2.0022	2.0677	2.0022	2.0222	2.0537	2.0824	2.0022	2.0023	2.0022	1.971	2.0022	1.9721	2.0022	2.0355
	SISDR at Convex	14.71	14.42	14.73	14.68	14.73	14.71	14.65	14.63	14.73	14.74	14.73	14.68	14.73	14.64	14.73	14.67
	Epoch at Convex 1 Epoch = 1000 iter	234	210	264	251	264	190	194	134	264	244	264	252	264	385	264	223
	Time at Convex	7h 51m 23s	7h 32m 26s	7h 57m 15s	6h 12m 36s	7h 57m 15s	5h 5m 16s	5h 8m 30s	3h 30m 37s	7h 57m 15s	6h 11m 56s	7h 57m 15s	6h 19m 26s	7h 57m 15s	2h 14m 24s	7h 57m 15s	7h 8m 54s
Diff bew/ train and valid	Loss at Convex (1e-3)	0.0951	0.0061	0.1105	0.1147	0.1105	0.1219	0.159	0.1801	0.1105	0.106	0.1105	0.11556	0.1105	0.0681	0.1105	0.1161
	SISDR at Convex	0.25	0.09	0.25	0.19	0.25	0.24	0.32	0.27	0.25	0.22	0.25	0.25	0.25	0.09	0.25	0.22

Result

		FFT size		LSTM Layer		Hop size		Hop size / LSTM Layer	
		512	256	256	128	256	128	256/256	128/128
Train	Loss at Convex (1e-3)	1.8917	3.0974	1.8917	1.928	1.8917	1.8738	1.8917	1.9064
	SISDR at Convex	14.98	13.48	14.98	14.91	14.98	15.28	14.98	15.22
Valid	Loss at Convex (1e-3)	2.0022	3.2427	2.0022	1.9768	2.0022	2.0039	2.0022	1.9451
	SISDR at Convex	14.73	13.31	14.73	14.76	14.73	14.92	14.73	15.1
	Epoch at Convex 1 Epoch = 1000 iter	264	203	264	399	264	279	264	399
	Time at Convex	7h 57m 15s	6h 36m 46s	7h 57m 15s	12h 14m 24s	7h 57m 15s	16h 12m 23s	7h 57m 15s	22h 26m 24s
Diff bew/ train and valid	Loss at Convex (1e-3)	0.1105	0.1453	0.1105	0.0487	0.1105	0.1300	0.1105	0.0387
	SISDR at Convex	0.25	0.17	0.25	0.15	0.25	0.36	0.25	0.12

- rmse of stft : 20221122-150347
- mse of stft : 20221123-093354
- psa : 20221123-183326
- rmsa of amplitude : 20221124-105217

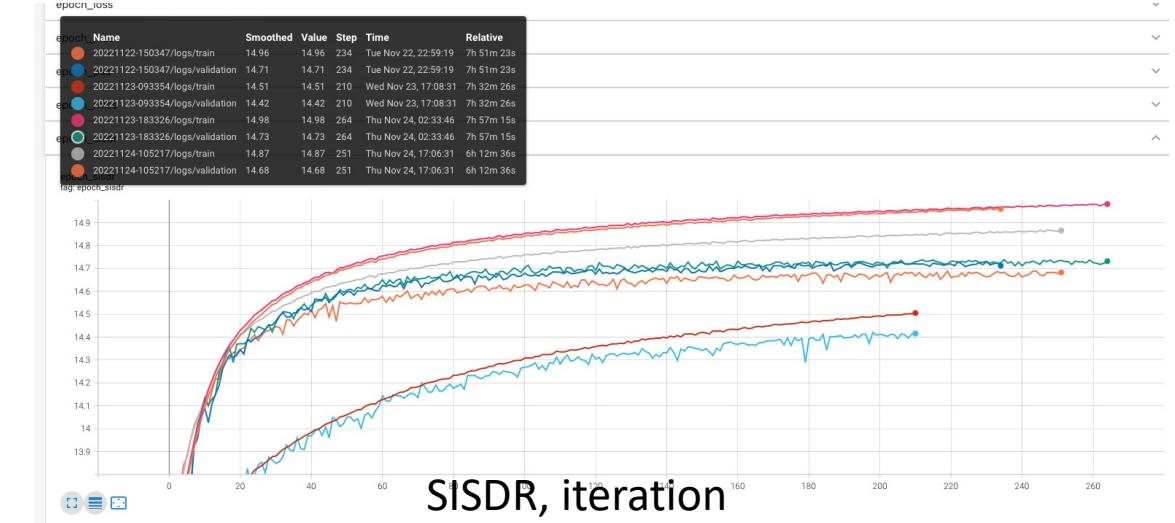
Loss, epoch



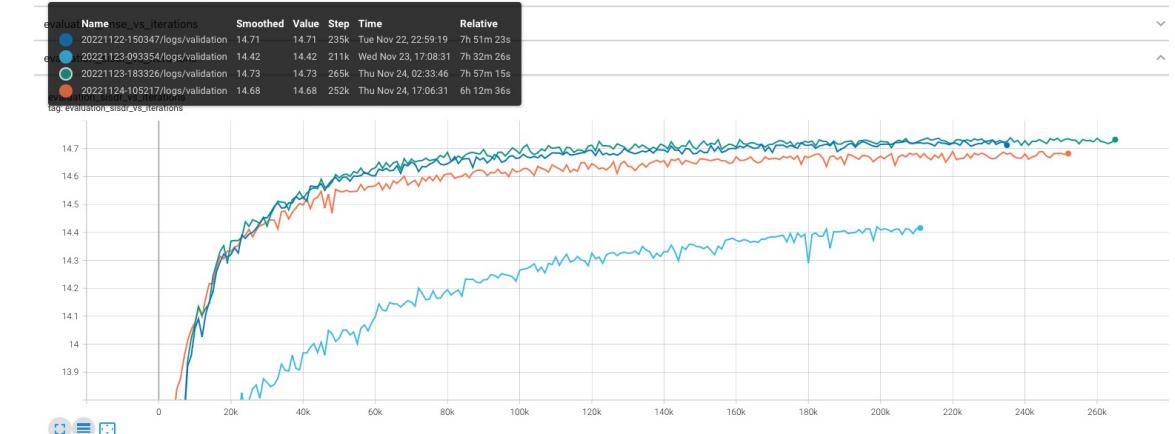
RMSE(stft) vs MSE(stft) vs PSA vs RMSE(amplitude)

Loss function

SISDR, epoch



evaluation_psa_vs_iterations



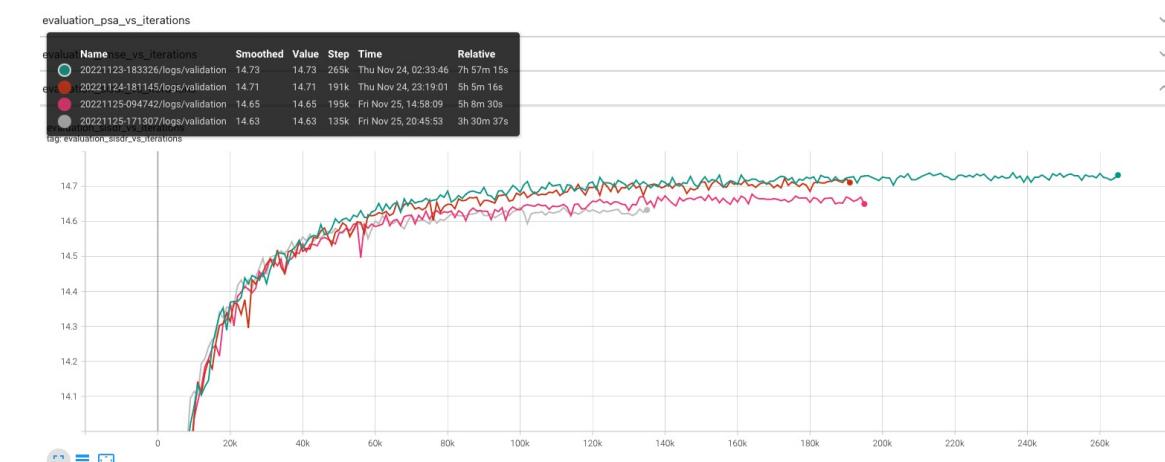
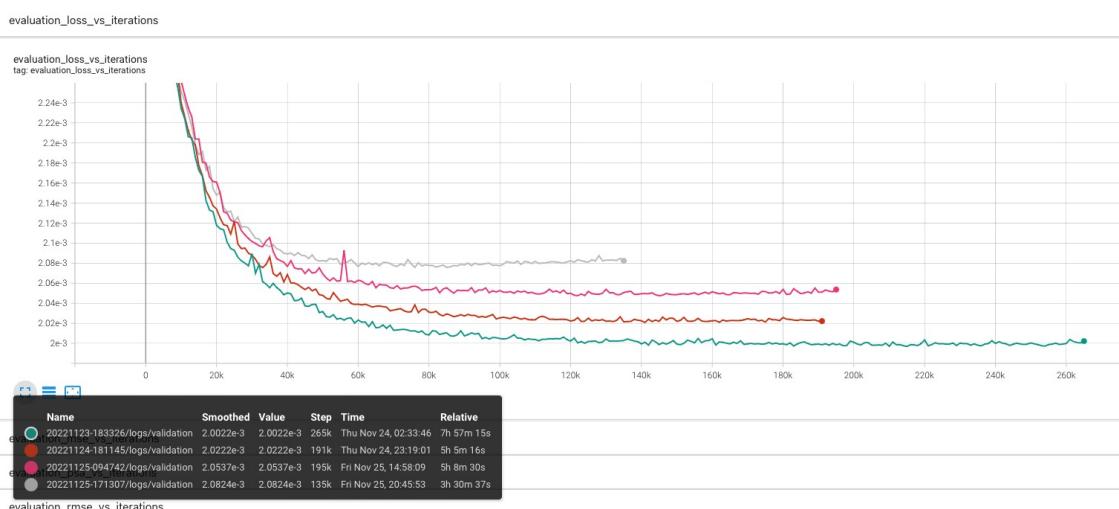
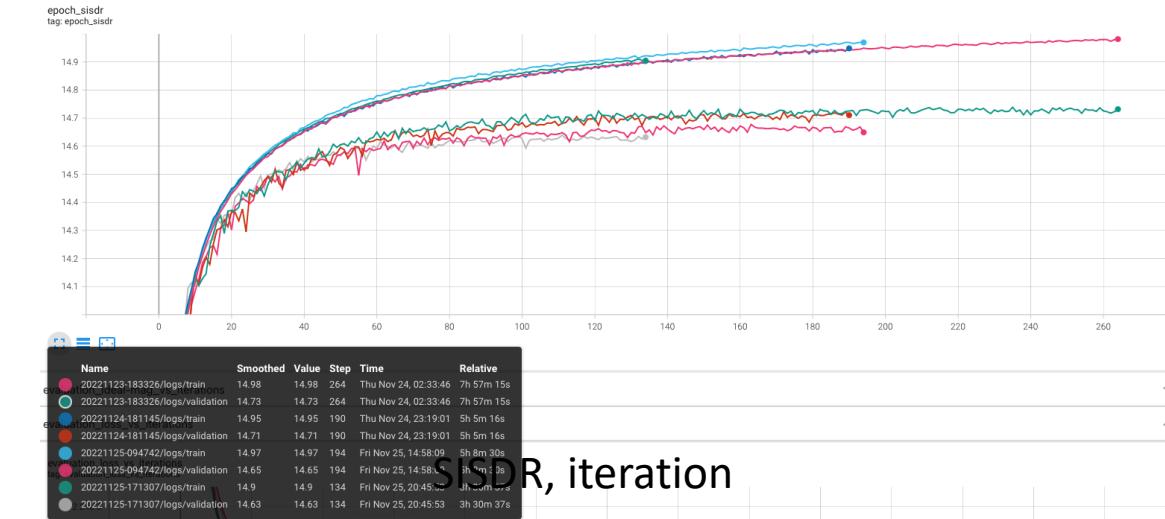
- 90:10: 20221123-183326
- 80:20: 20221124-181145
- 70:30: 20221125-094742
- 50:50: 20221125-171307

Loss, epoch

Train vs Validation Ratio

90:10, 80:20, 70:30, 50:50

SISDR, epoch



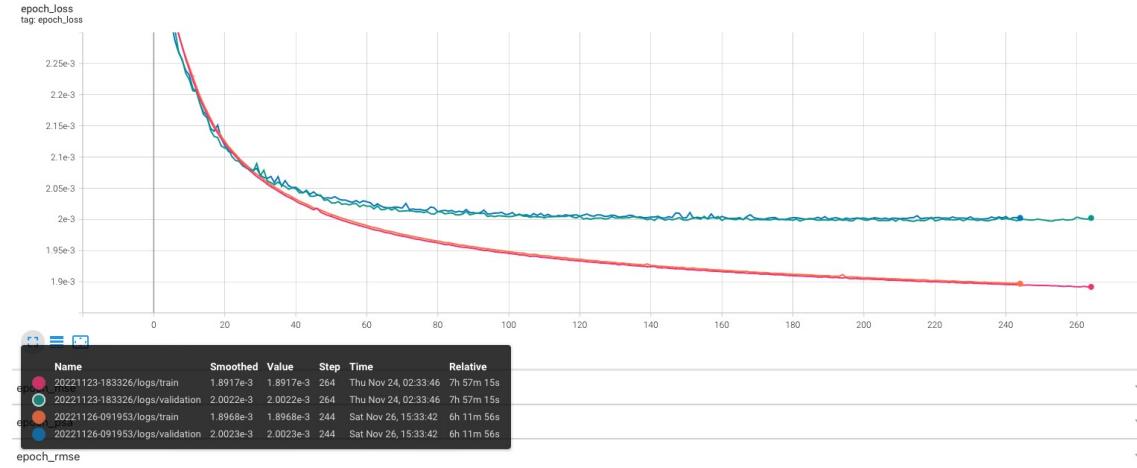
Dense layer initialization

O / X

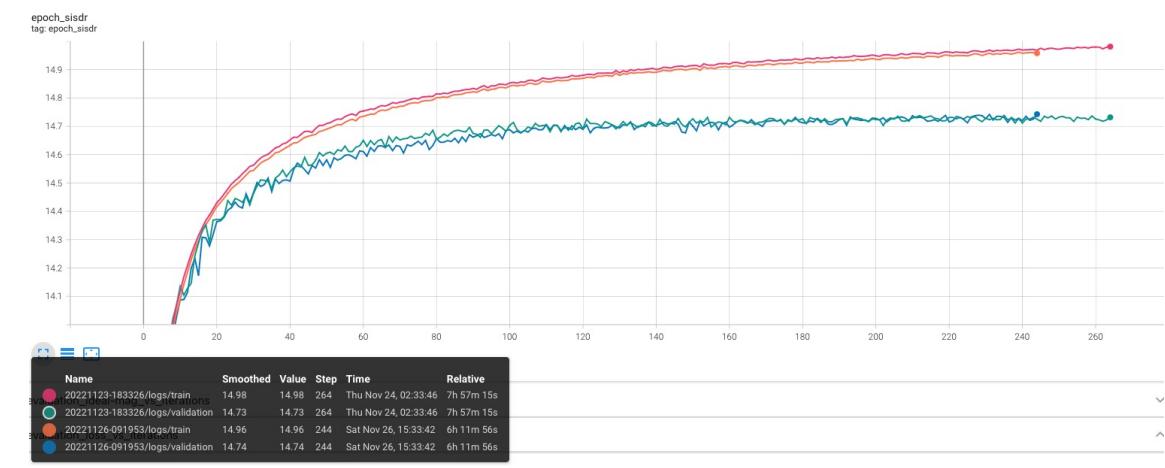
20221123-183326 / 20221126-091953

Loss, epoch

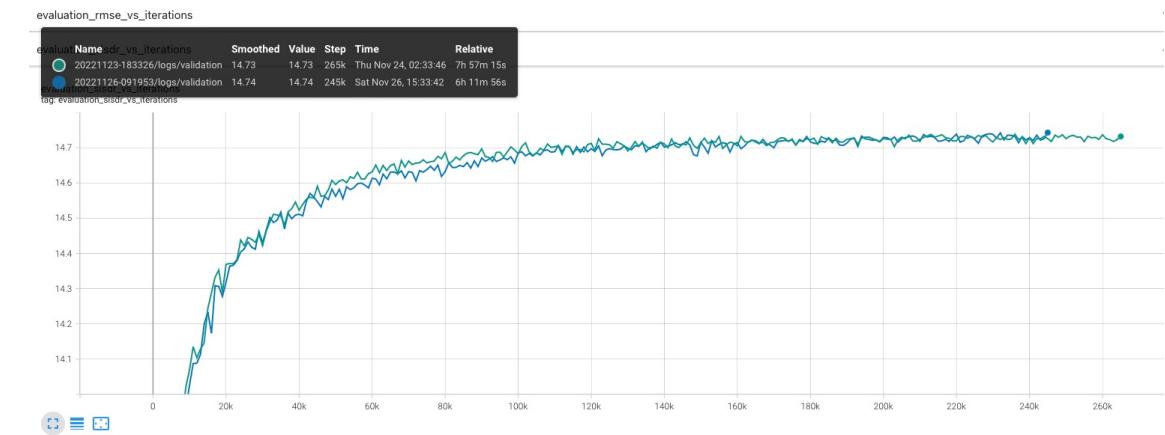
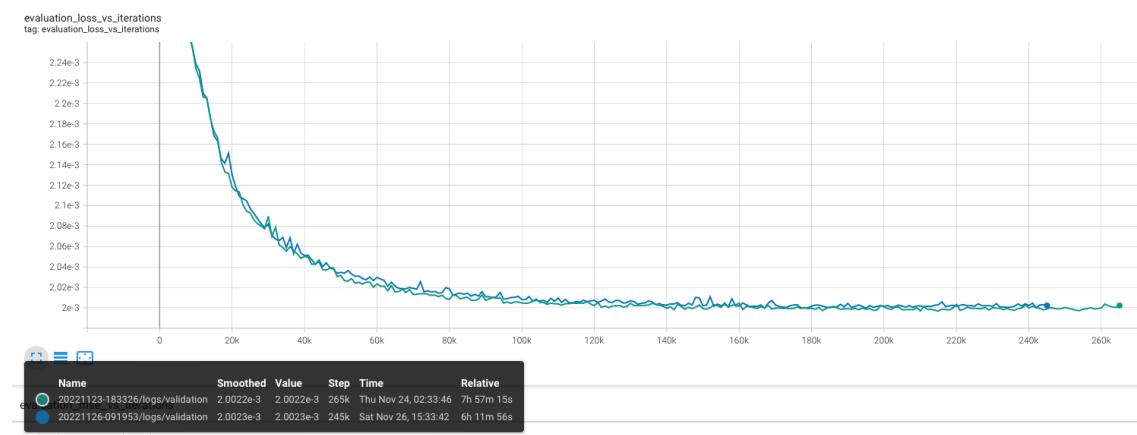
SISDR, epoch



Loss, iteration



SISDR, iteration



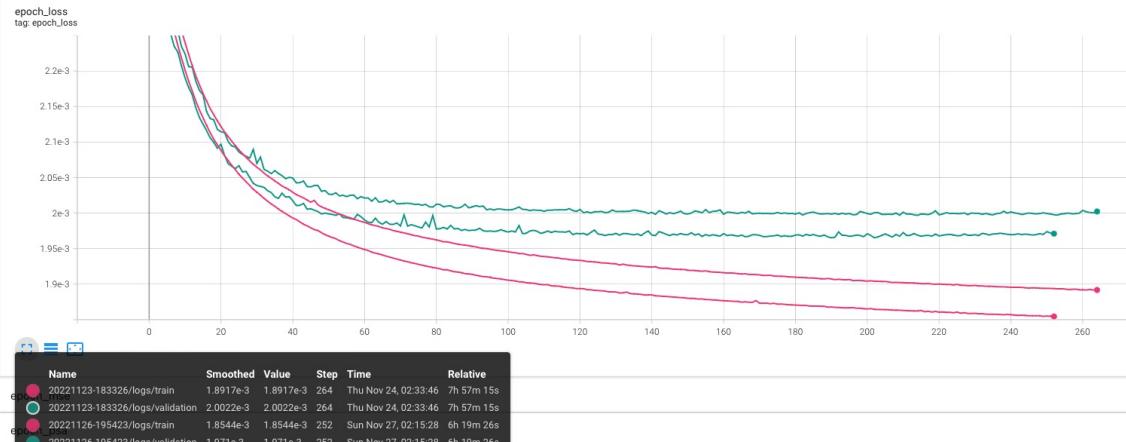
Time Period of Normalization

Entire / 1 sec

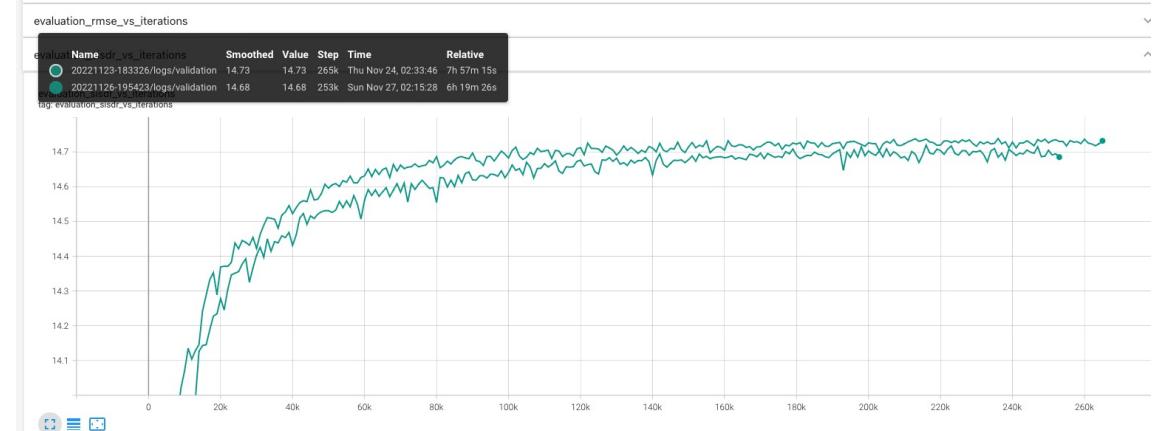
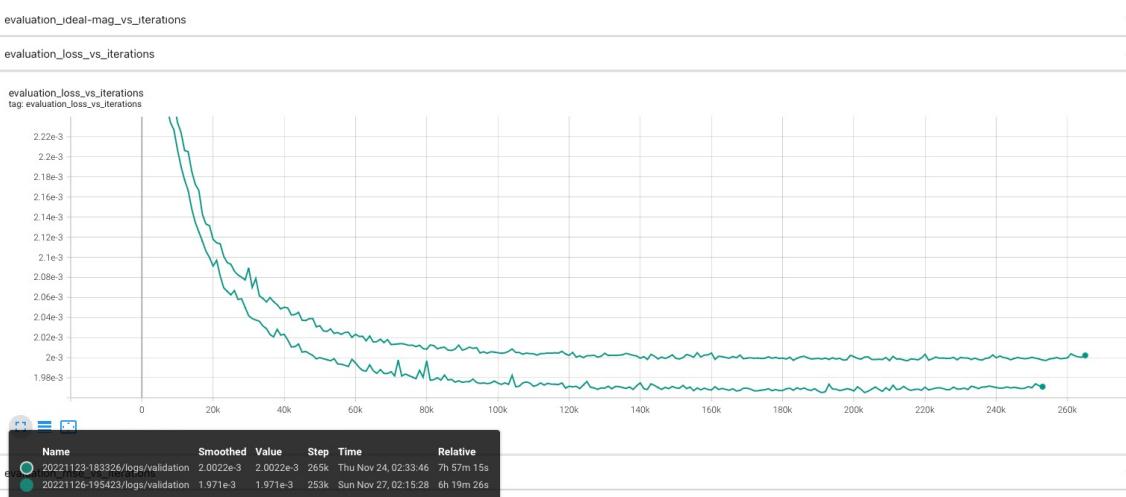
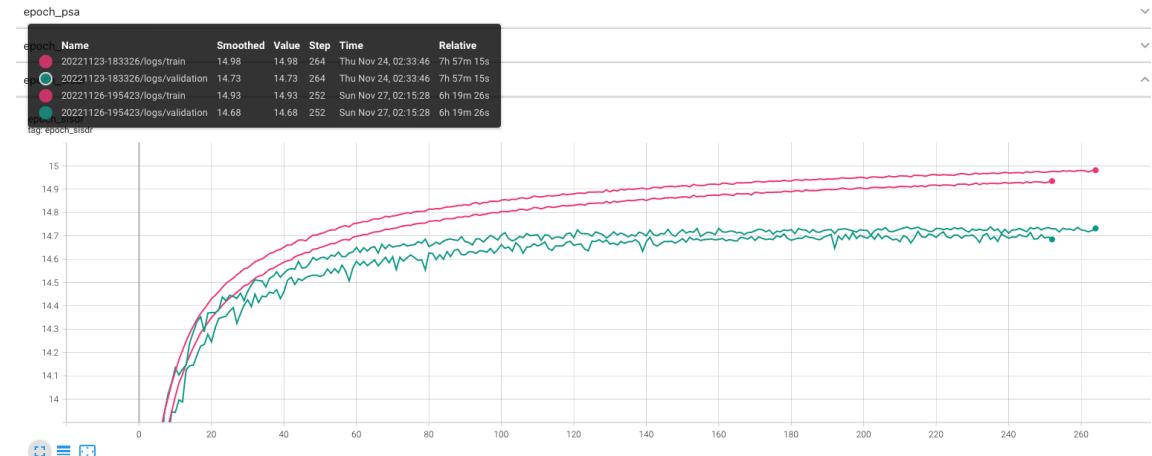
20221123-183326 / 20221126-195423

Loss, epoch

SISDR, epoch



Loss, iteration



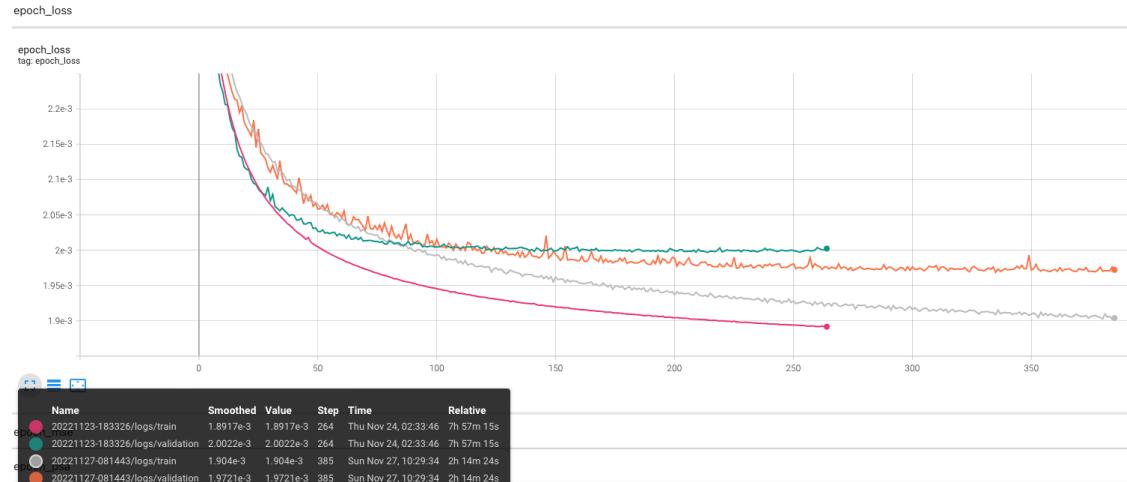
Loss, epoch

Batch size

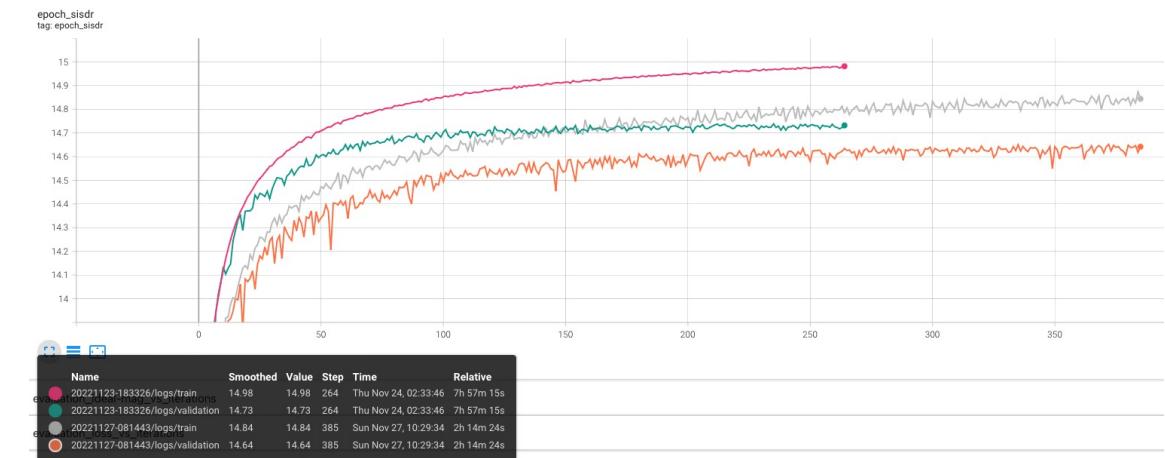
512 / 64

20221123-183326 / 20221127-081443

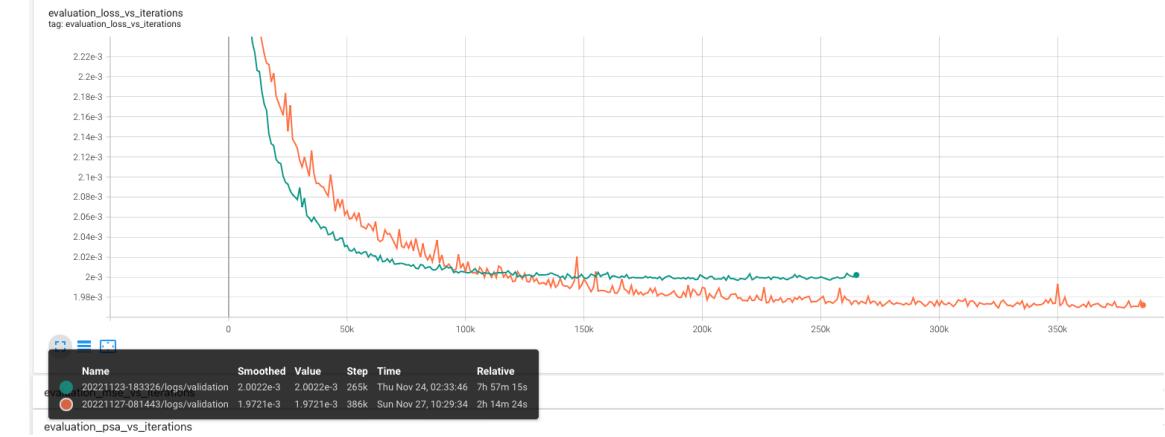
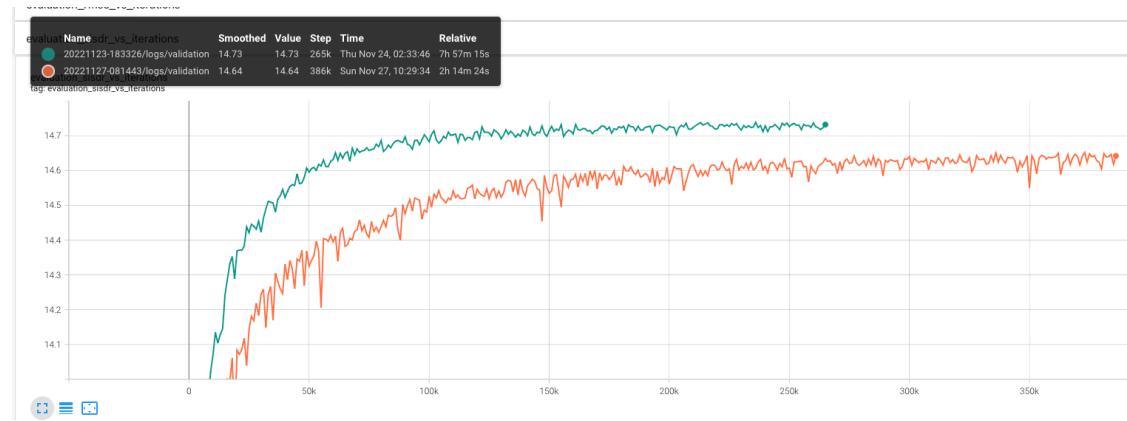
SISDR, epoch



Loss, iteration



SISDR, iteration



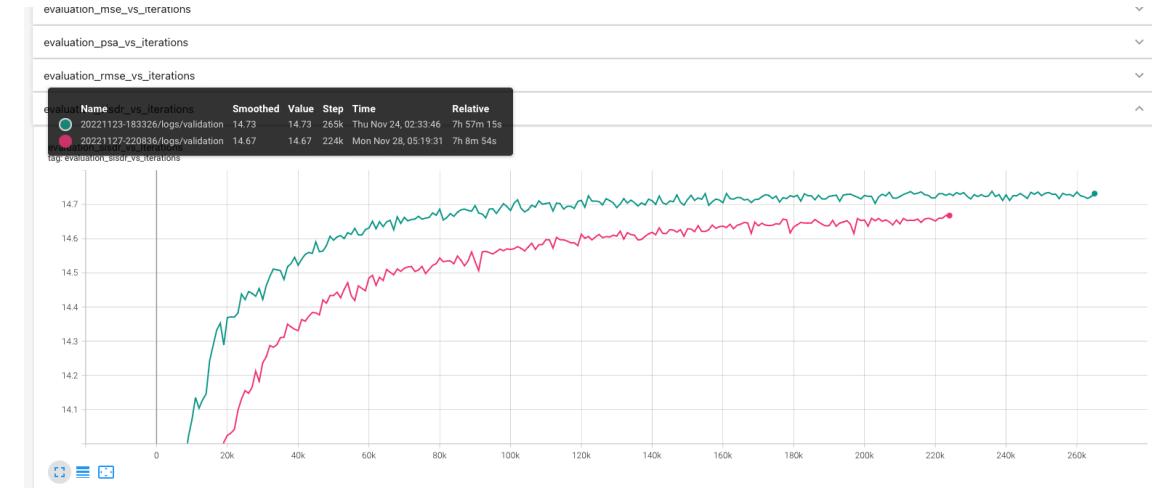
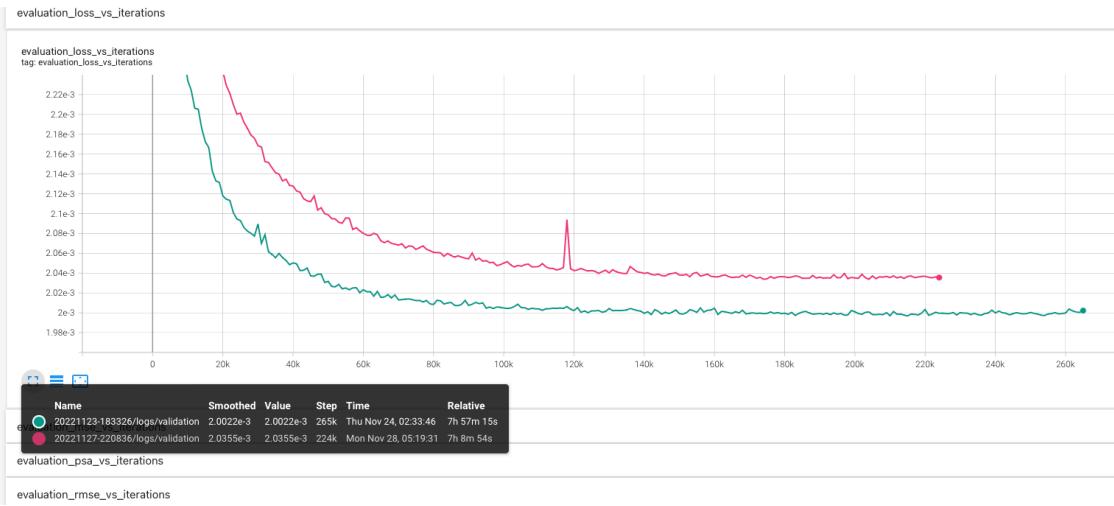
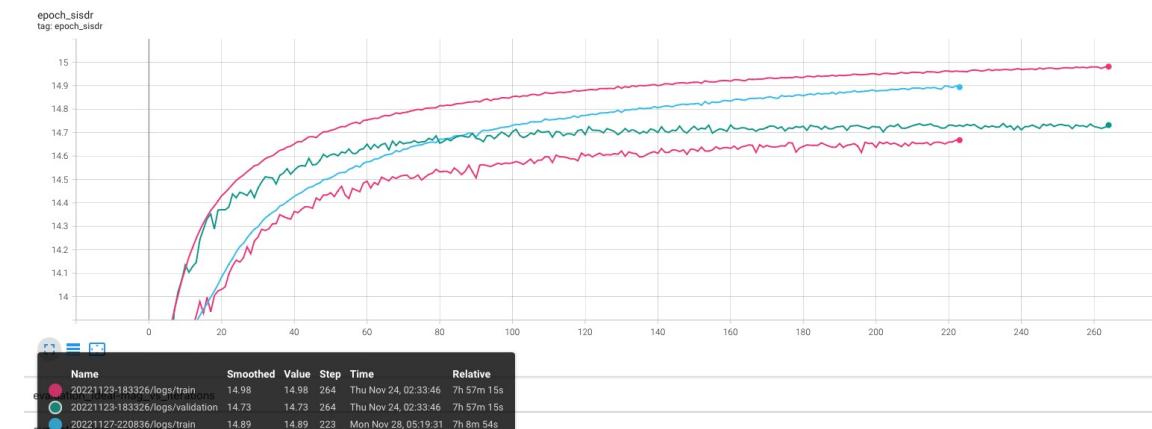
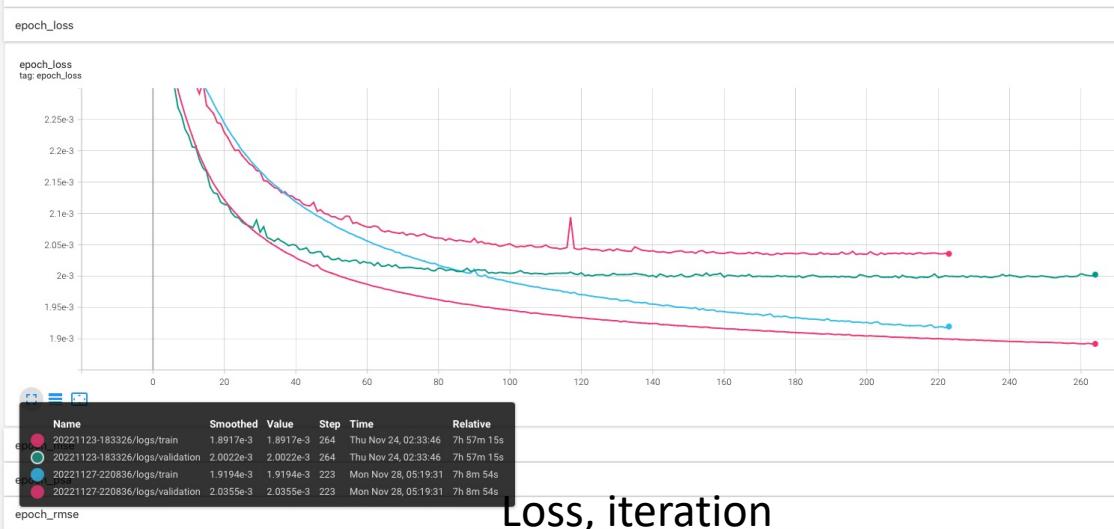
Batch Normalization

O / X

20221123-183326, 20221127-220836

Loss, epoch

SISDR, epoch



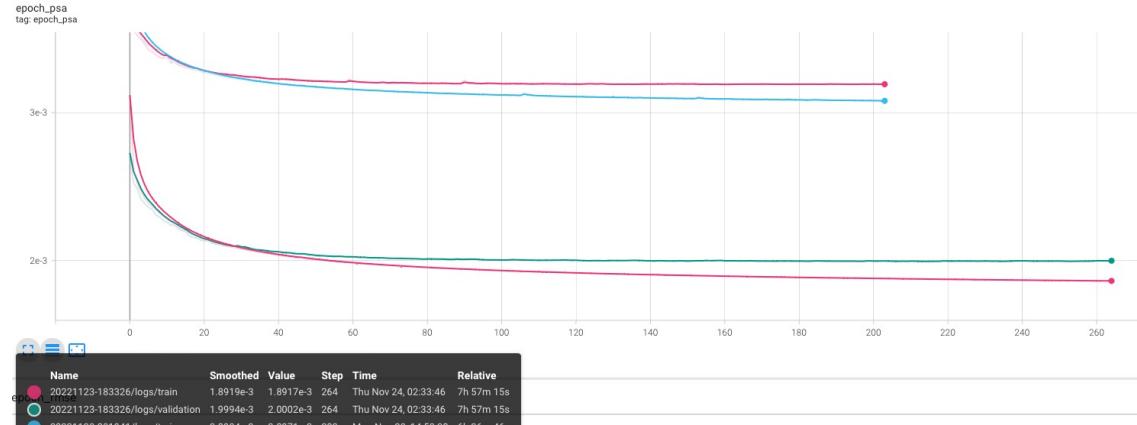
Loss, epoch

FFT Size

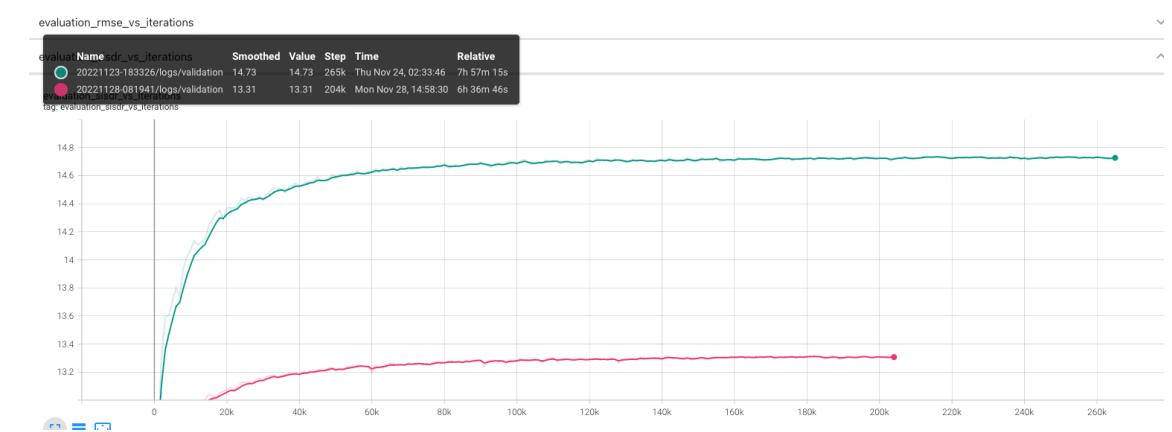
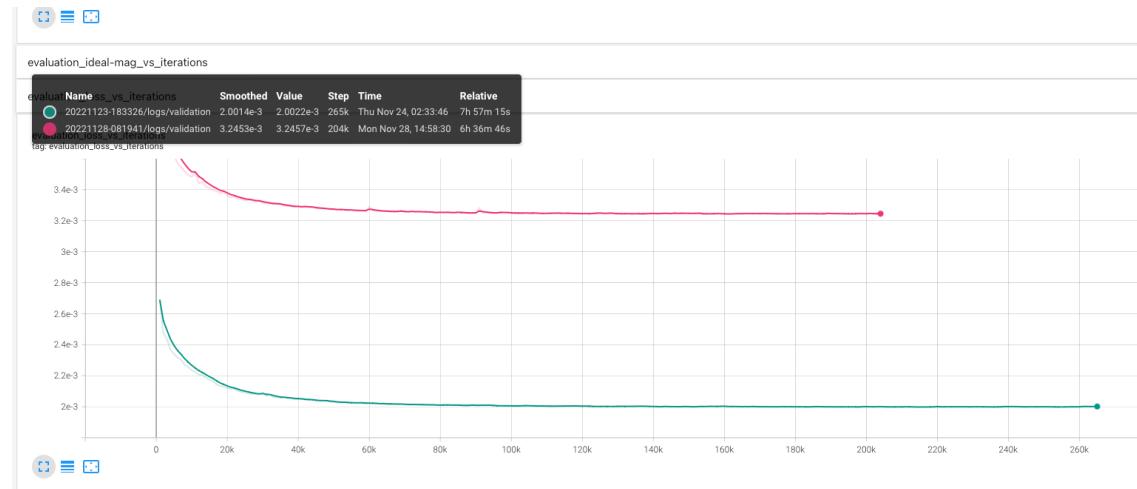
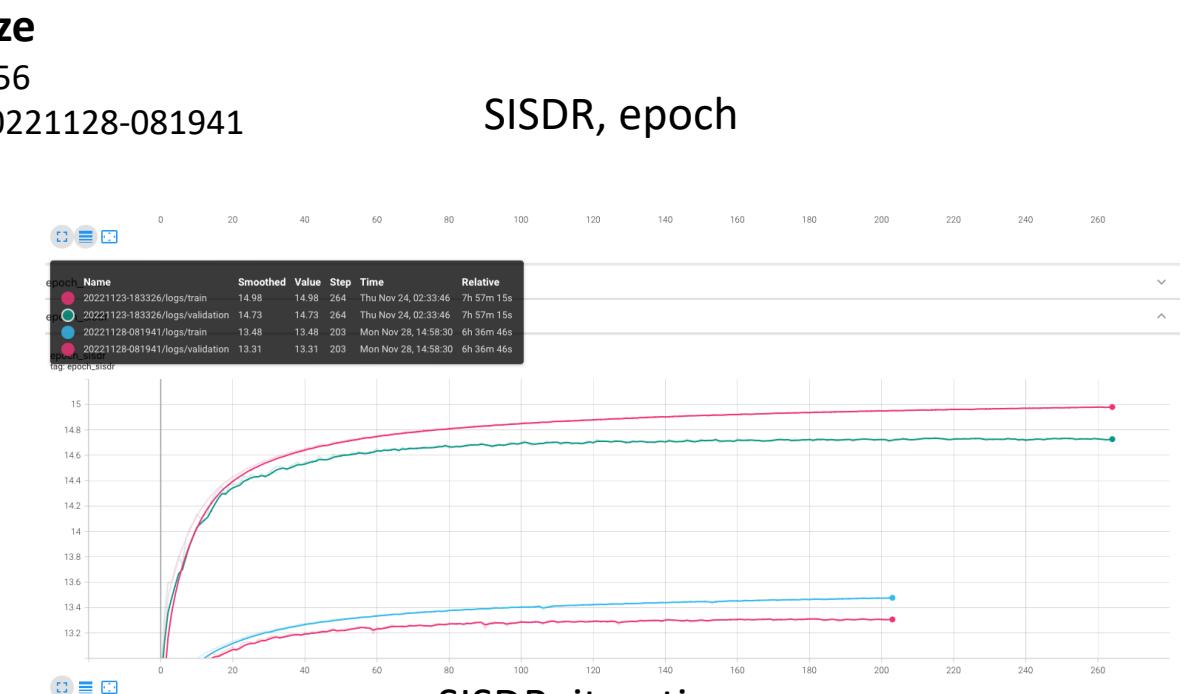
512 / 256

20221123-183326, 20221128-081941

SISDR, epoch



Loss, iteration



Loss, epoch

LSTM Layer

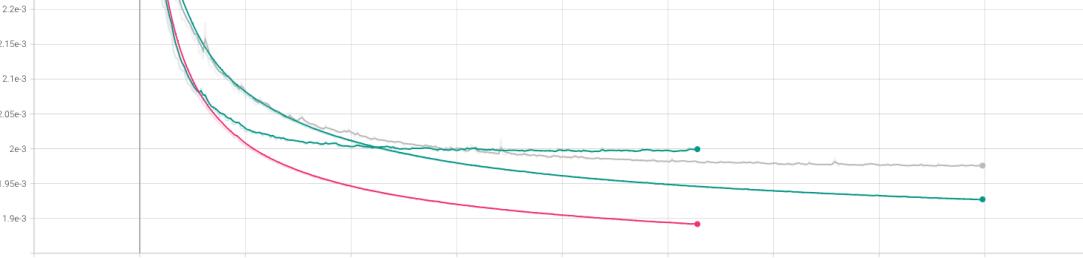
O / X

20221123-183326, 20221128-151953

SISDR, epoch

epoch_psa

epoch_psa
tag: epoch_psa

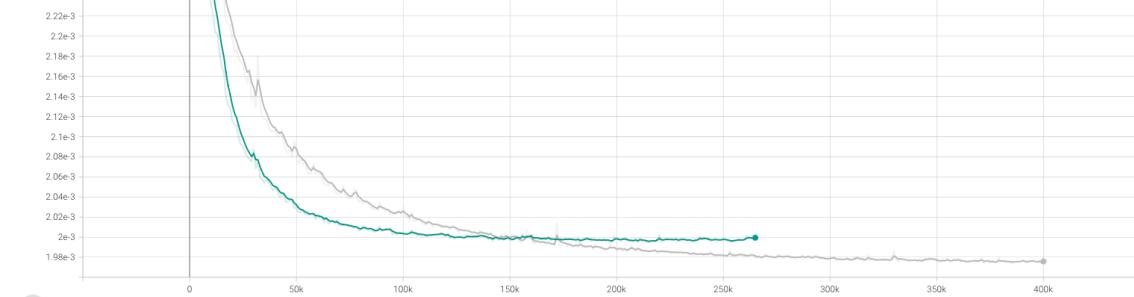


Name	Smoothed	Value	Step	Time	Relative
20221123-183326/logs/train	1.8919e-3	1.8917e-3	264	Thu Nov 24, 02:33:46	7h 57m 15s
20221123-183326/logs/validation	1.9994e-3	2.0002e-3	264	Thu Nov 24, 02:33:46	7h 57m 15s
20221128-151953/logs/train	1.9274e-3	1.928e-3	399	Tue Nov 29, 04:03:16	12h 41m 24s
20221128-151953/logs/validation	1.9759e-3	1.9768e-3	399	Tue Nov 29, 04:03:16	12h 41m 24s

Loss, iteration

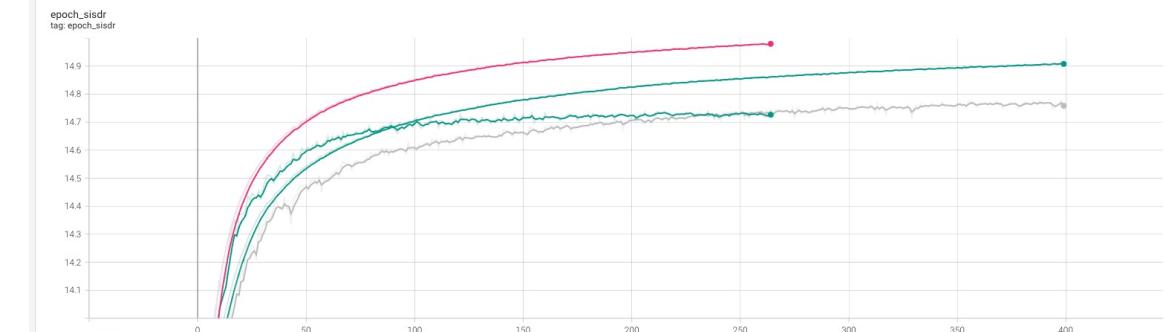
evaluation_psa_vs_iterations

evaluation_psa_vs_iterations
tag: evaluation_psa_vs_iterations



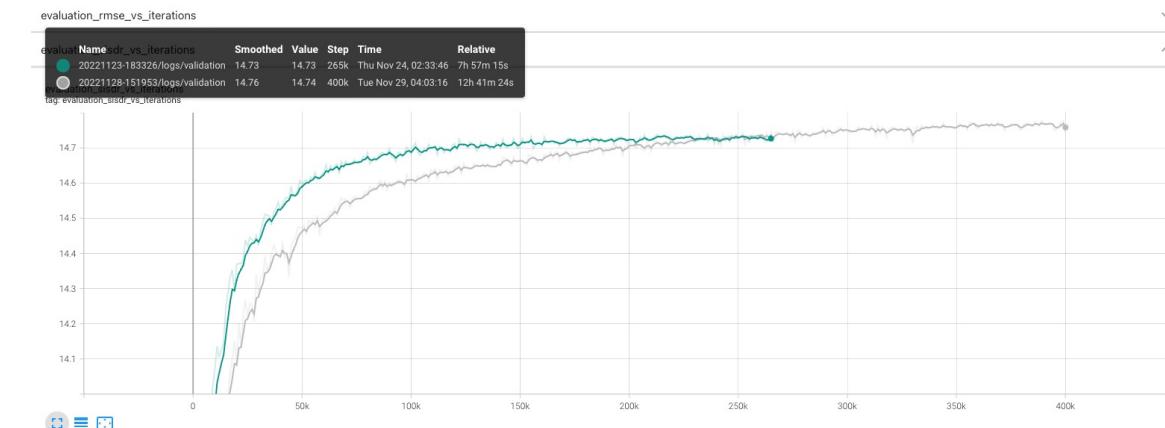
Name	Smoothed	Value	Step	Time	Relative
20221123-183326/logs/validation	1.9994e-3	2.0002e-3	265k	Thu Nov 24, 02:33:46	7h 57m 15s
20221128-151953/logs/validation	1.9759e-3	1.9768e-3	400k	Tue Nov 29, 04:03:16	12h 41m 24s

evaluation_sisdr_vs_iterations

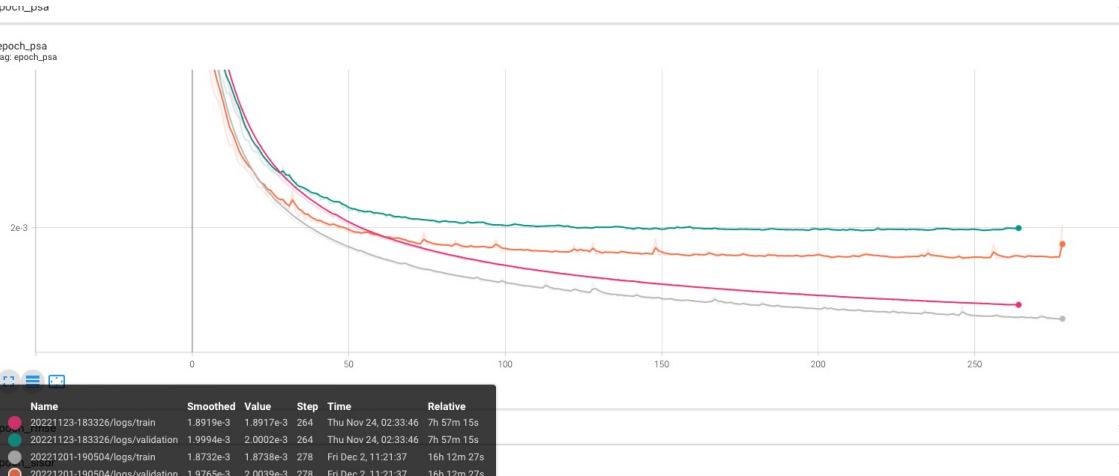


Name	Smoothed	Value	Step	Time	Relative
20221123-183326/logs/train	14.98	14.98	264	Thu Nov 24, 02:33:46	7h 57m 15s
20221123-183326/logs/validation	14.73	14.73	264	Thu Nov 24, 02:33:46	7h 57m 15s
20221128-151953/logs/train	14.91	14.91	399	Tue Nov 29, 04:03:16	12h 41m 24s
20221128-151953/logs/validation	14.76	14.74	399	Tue Nov 29, 04:03:16	12h 41m 24s

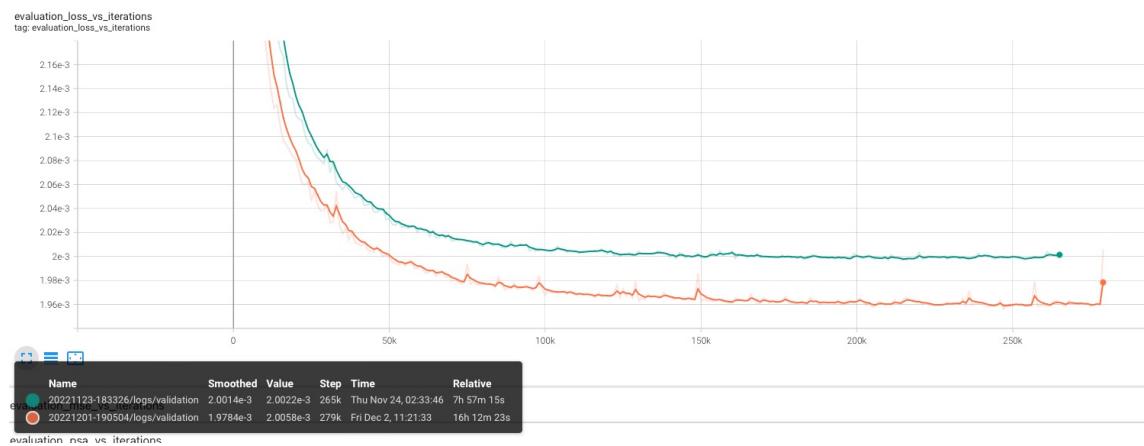
SISDR, iteration



Loss, epoch



Loss, iteration

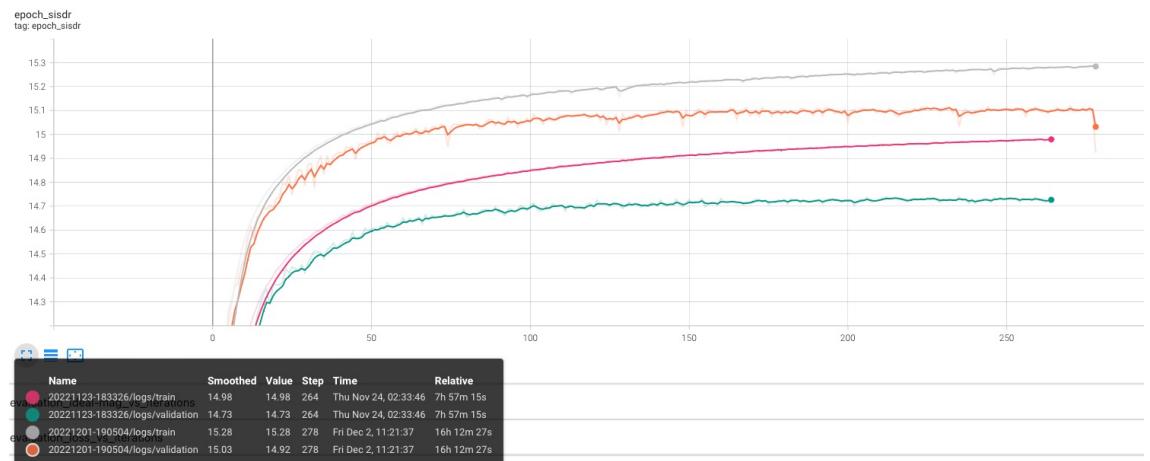


Hop size

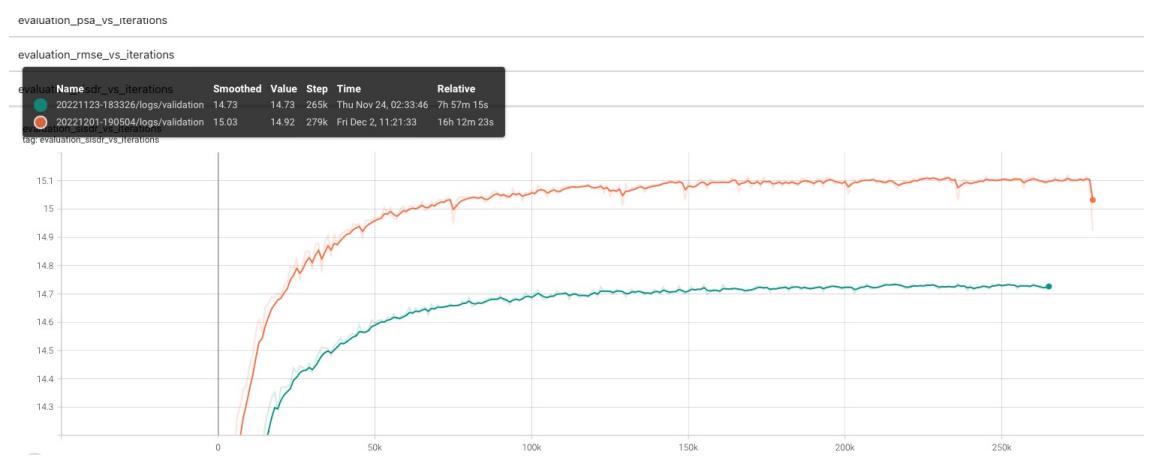
256 / 128

20221123-183326, 20221201-190504

SISDR, epoch



SISDR, iteration



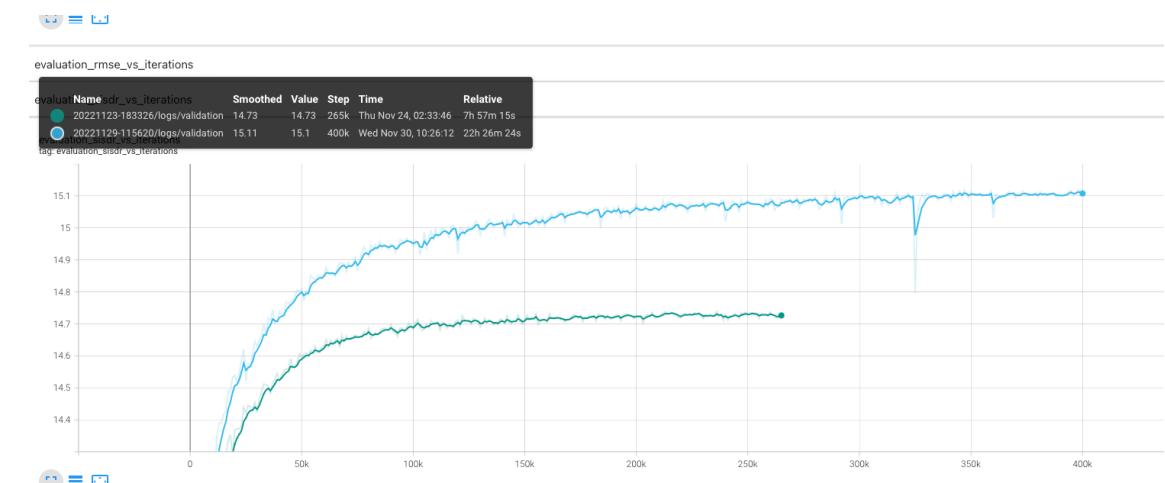
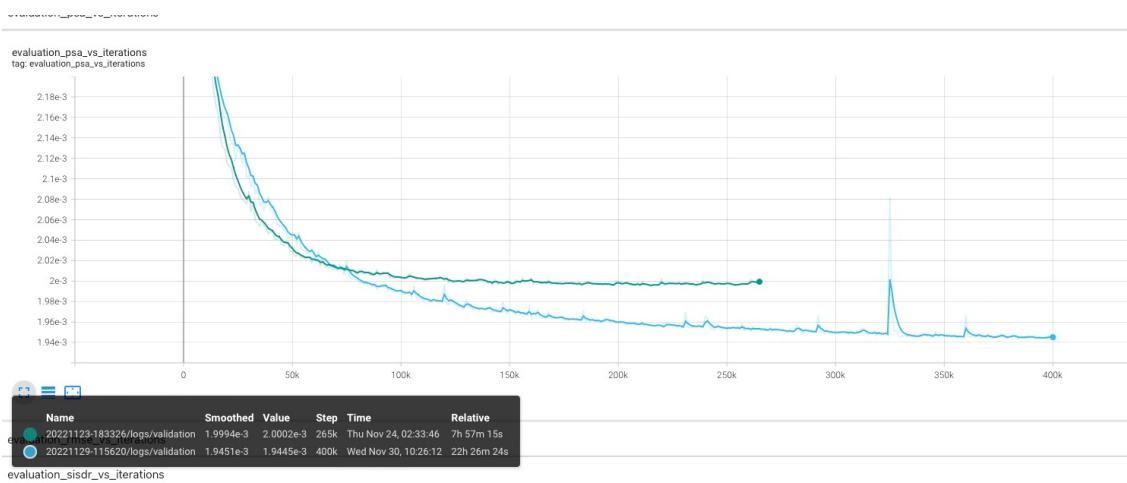
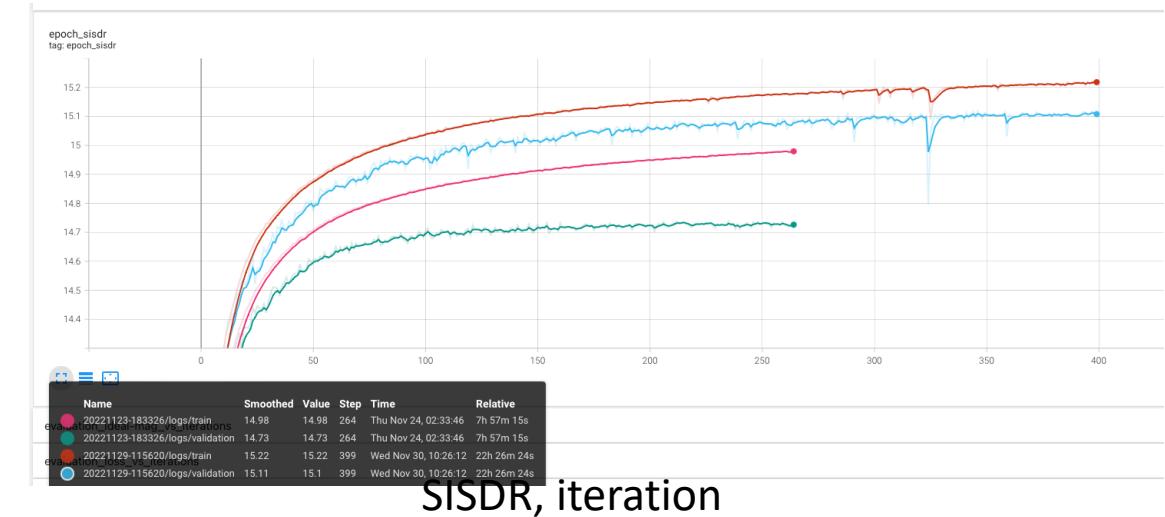
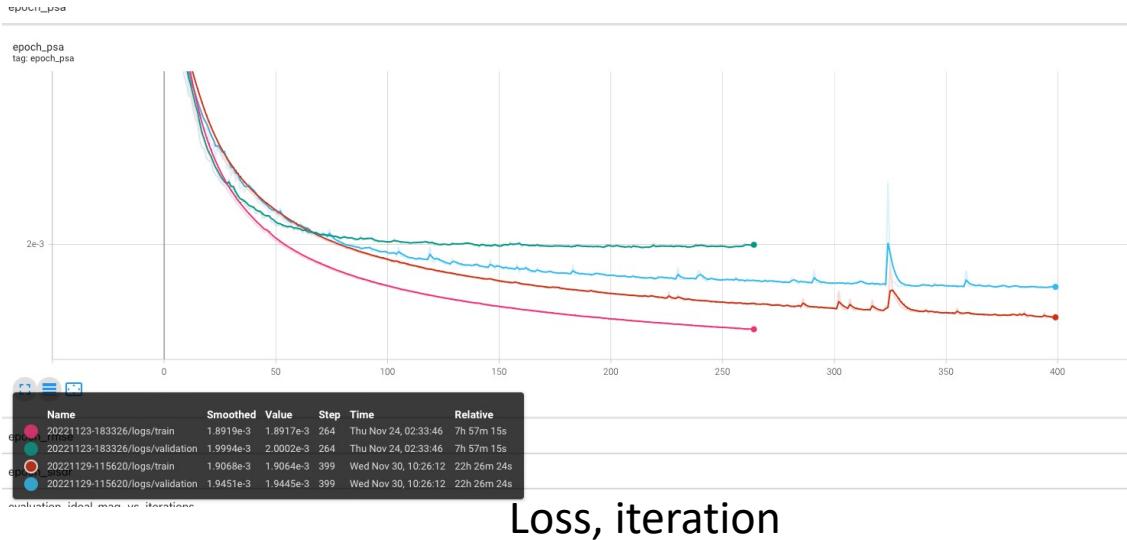
LSTM layer + Hop size

256,256 / 128, 128

20221123-183326, 20221129-115620

Loss, epoch

SISDR, epoch



Detail

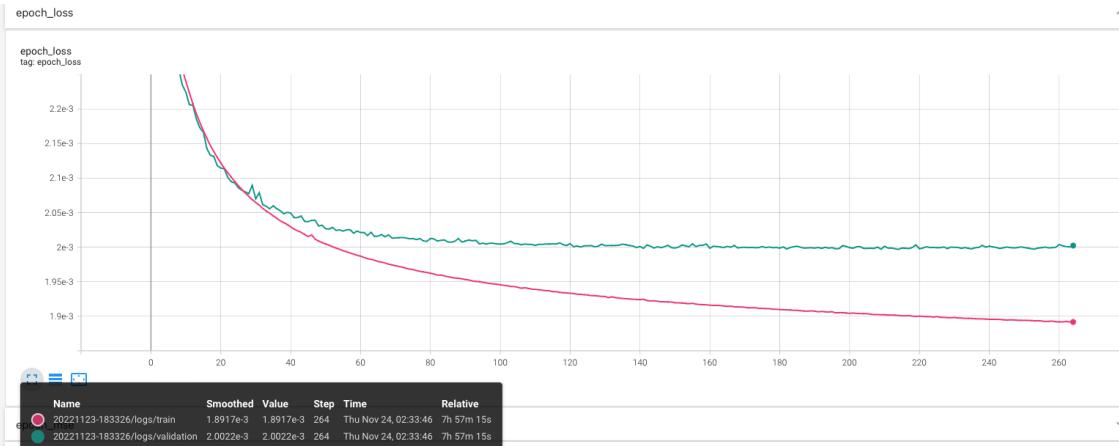
Train vs Validation Ratio

90:10, 80:20, 70:30, 50:50

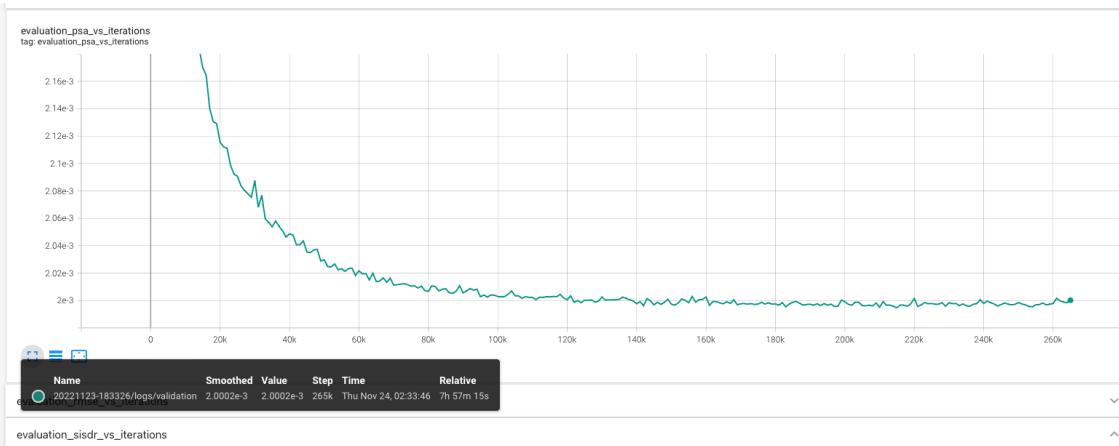
90:10

20221123-183326

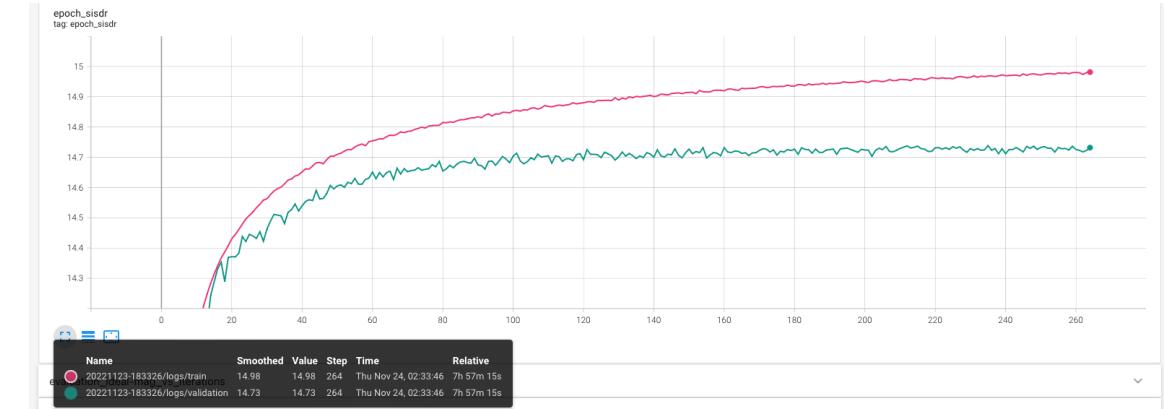
Loss, epoch



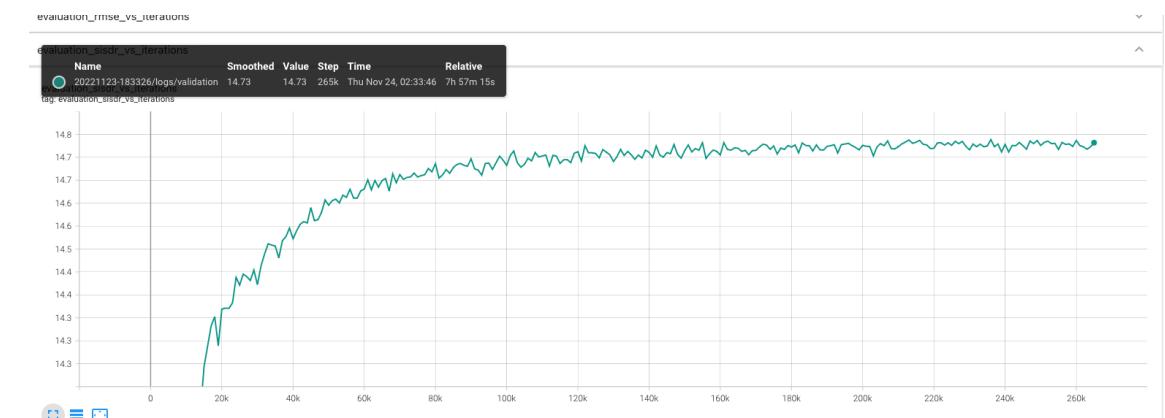
Loss, iteration



SISDR, epoch



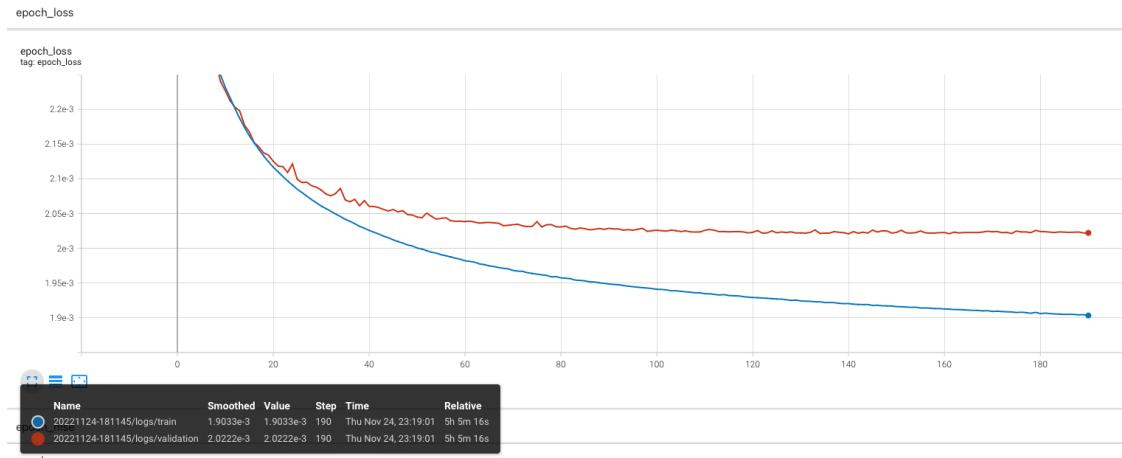
SISDR, iteration



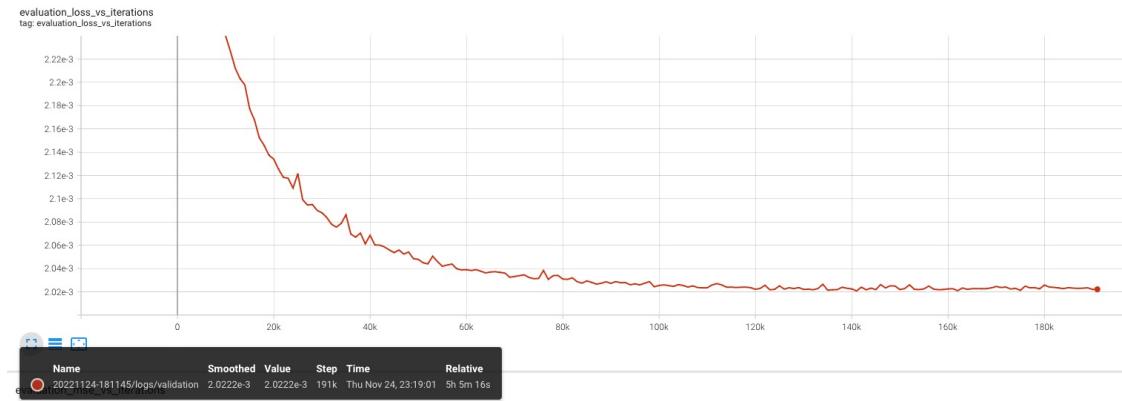
80:20

20221124-181145

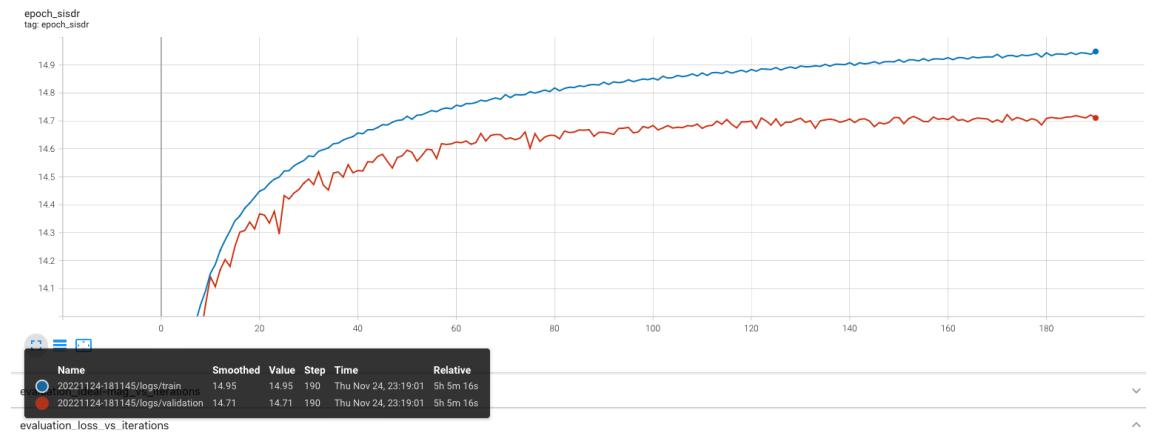
Loss, epoch



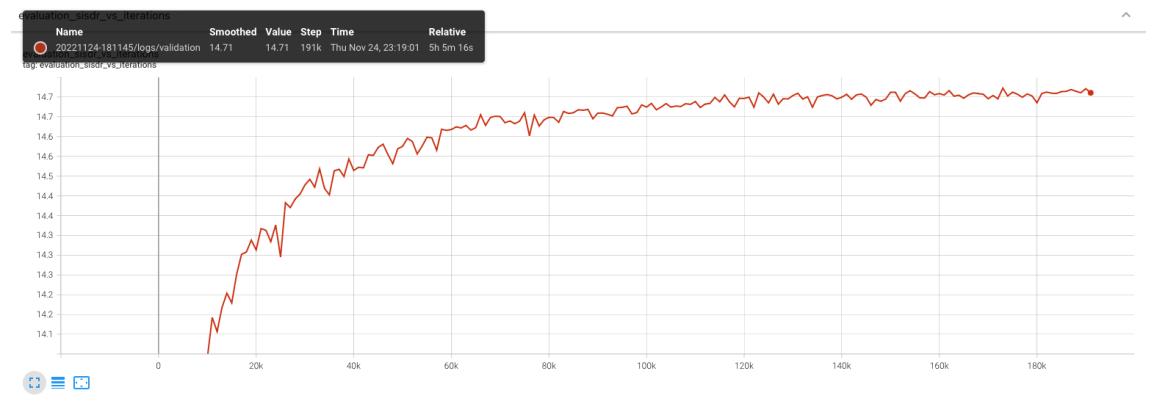
Loss, iteration



SISDR, epoch



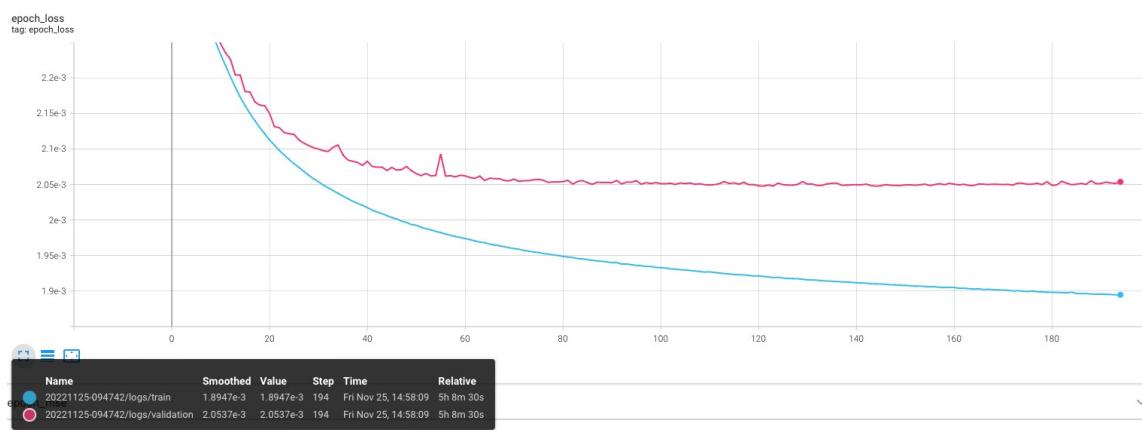
SISDR, iteration



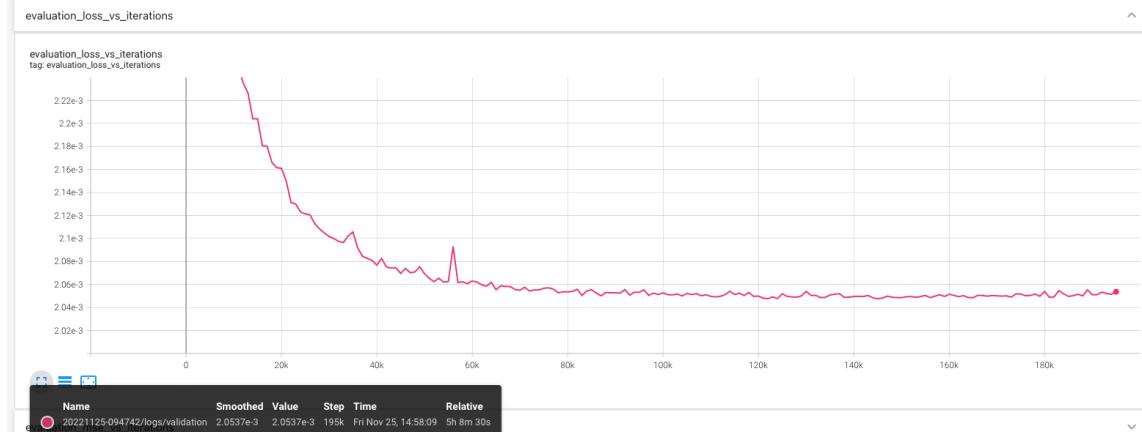
70:30

20221125-094742

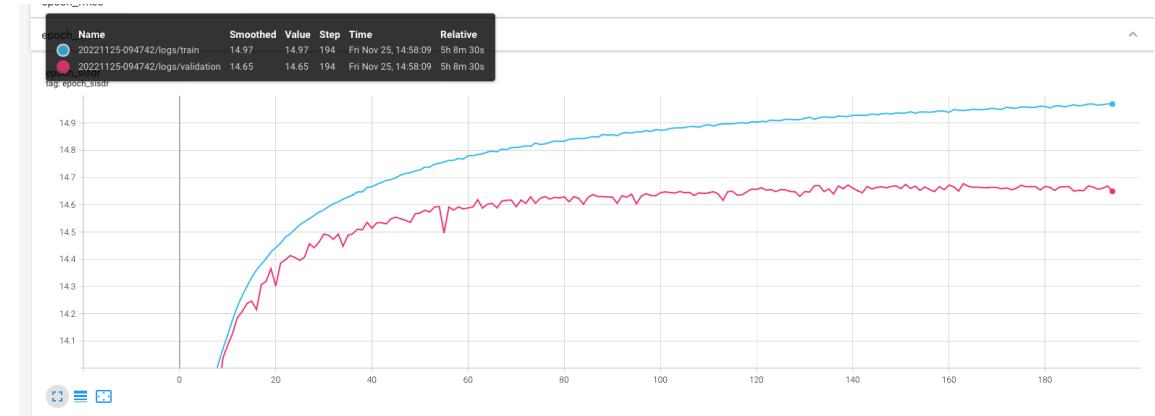
Loss, epoch



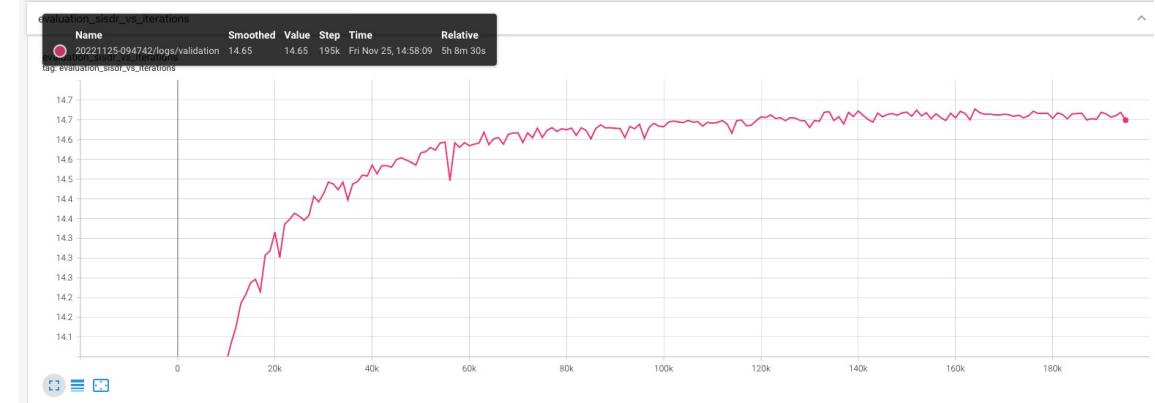
Loss, iteration



SISDR, epoch



SISDR, iteration

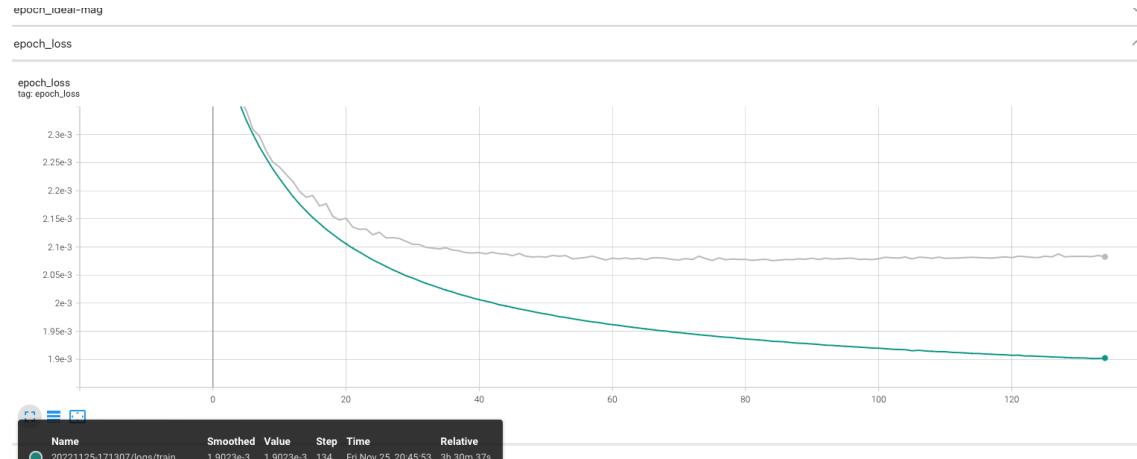


50:50

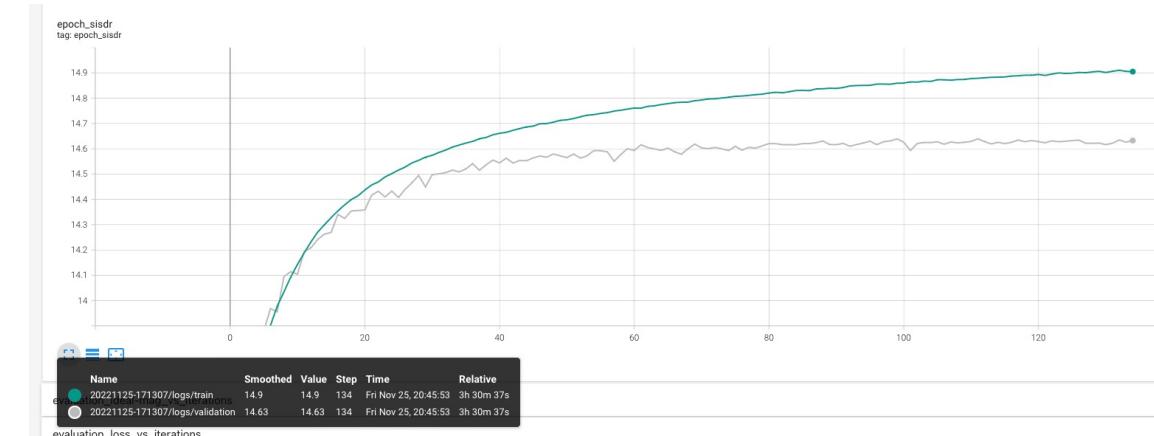
20221125-171307

Loss, epoch

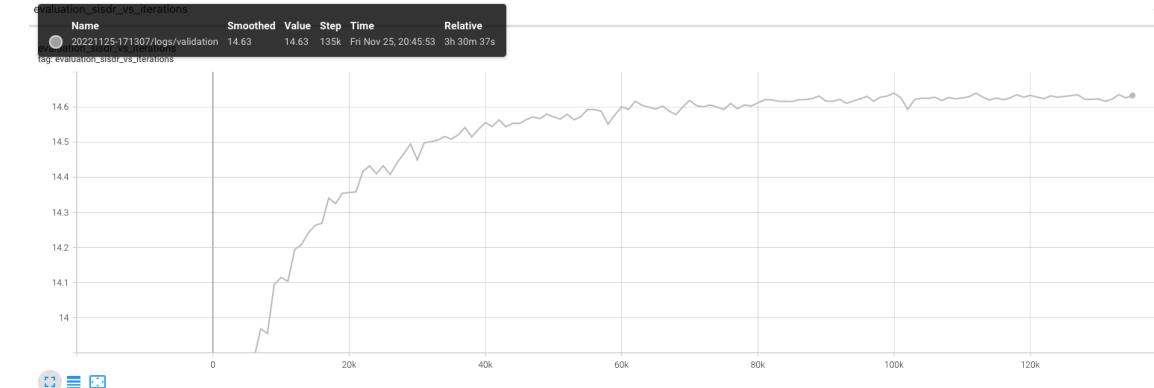
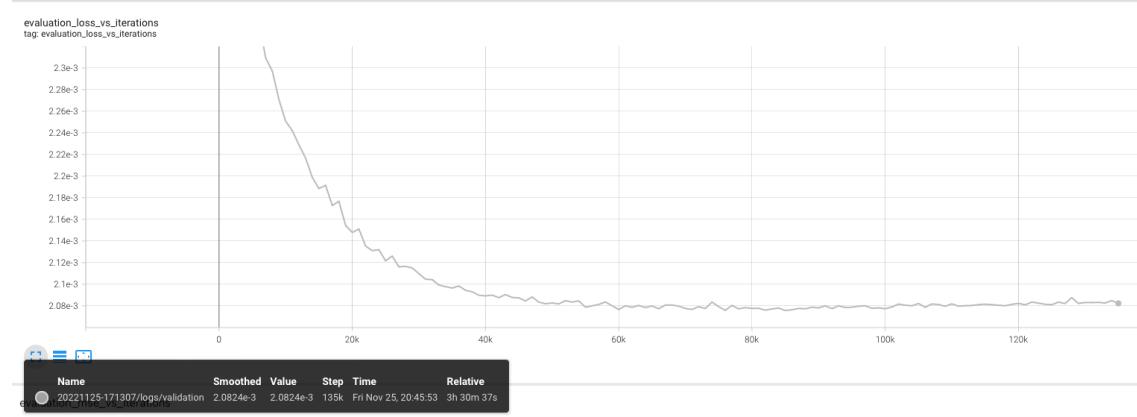
SISDR, epoch



Loss, iteration



SISDR, iteration



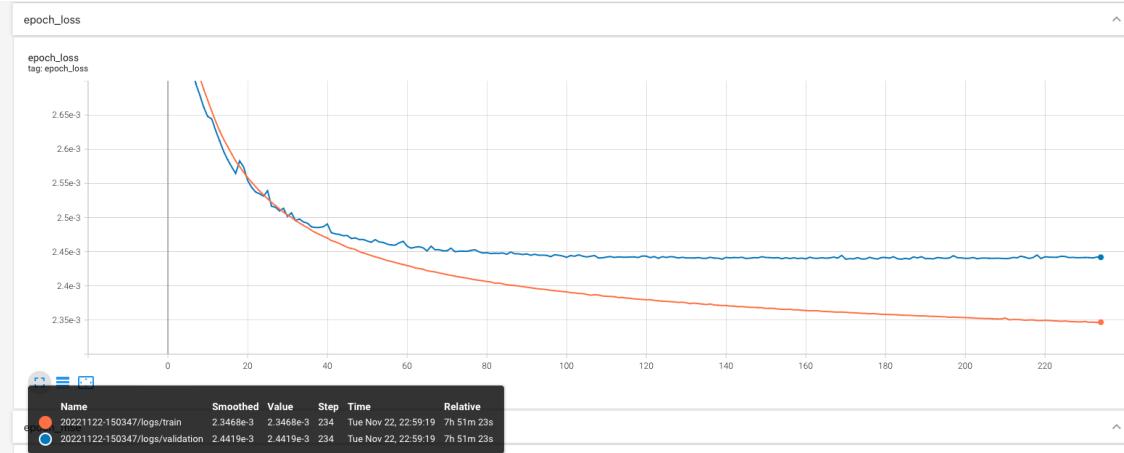
Detail
Loss function

RMSE(stft) vs MSE(stft) vs PSA vs RMSE(amplitude)

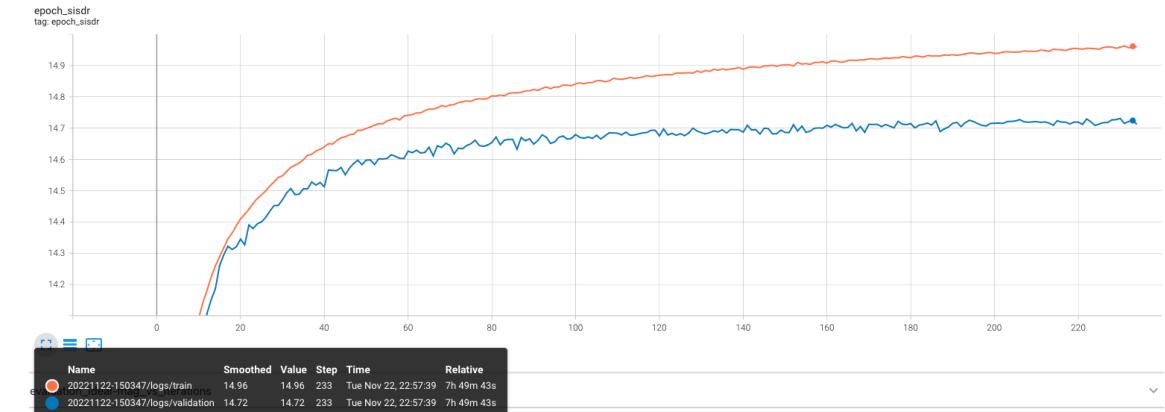
RMSE(stft)

20221122-150347

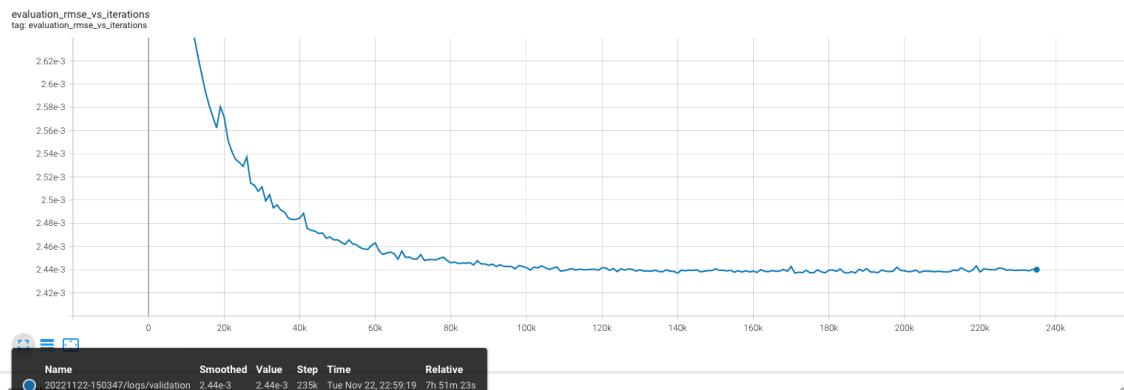
Loss, epoch



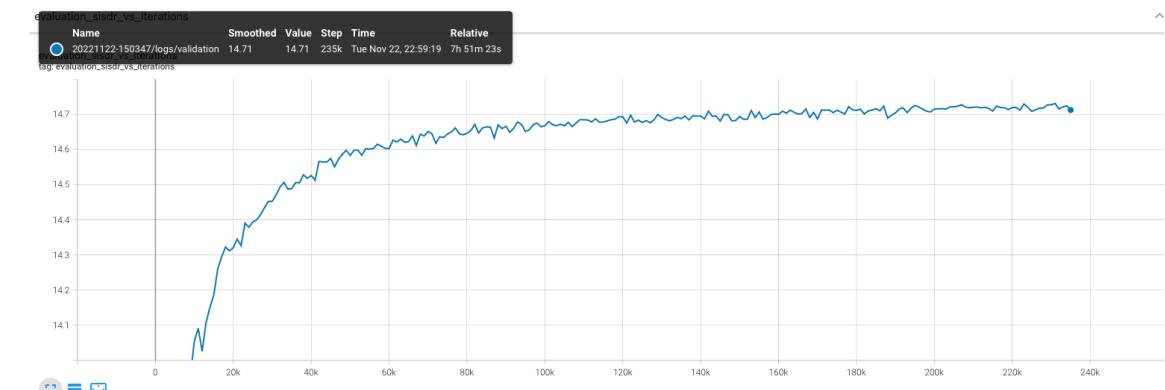
SISDR, epoch



Loss, iteration



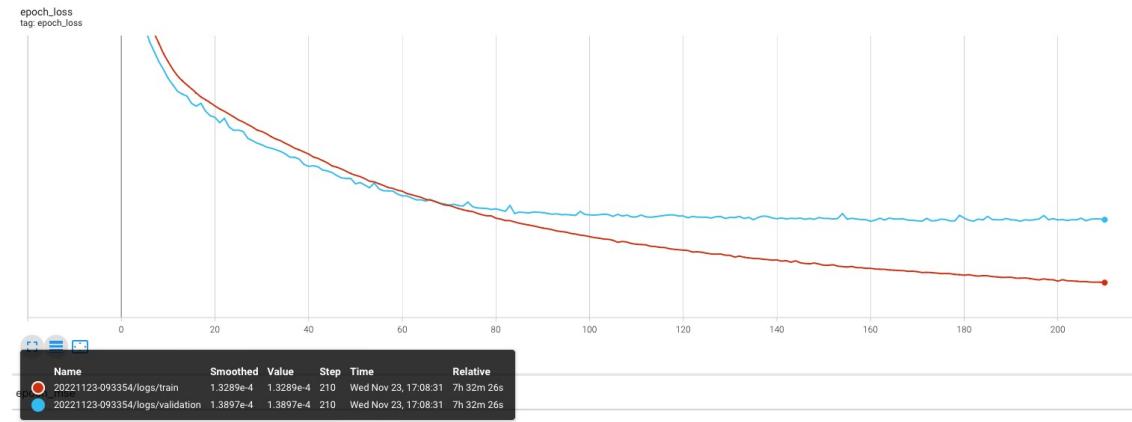
SISDR, iteration



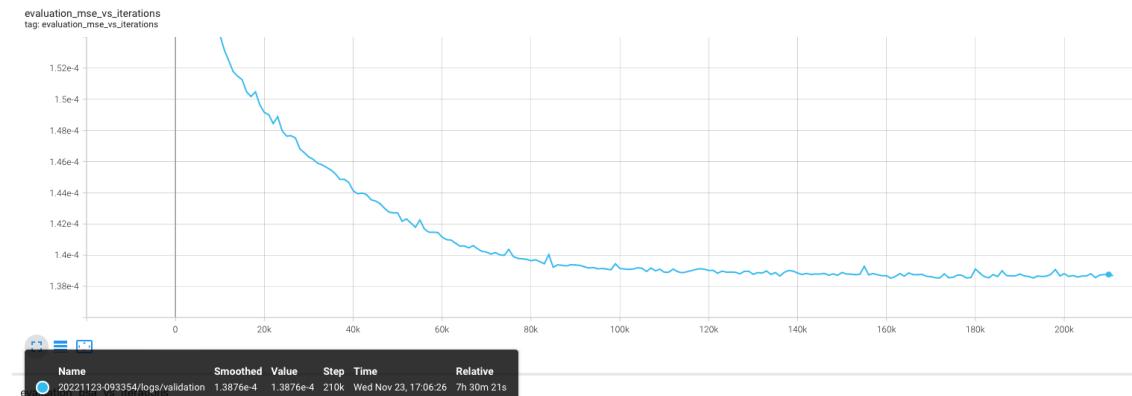
MSE(stft)

221123-093354

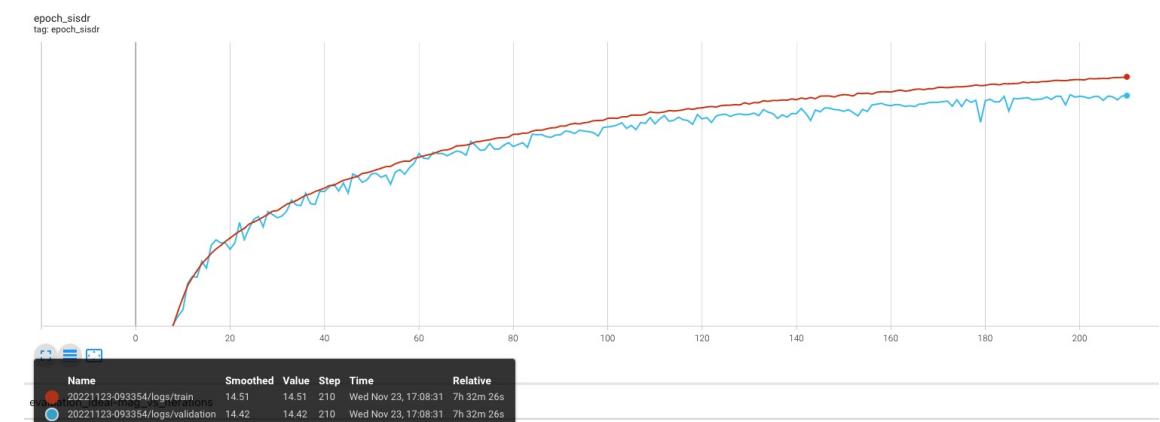
Loss, epoch



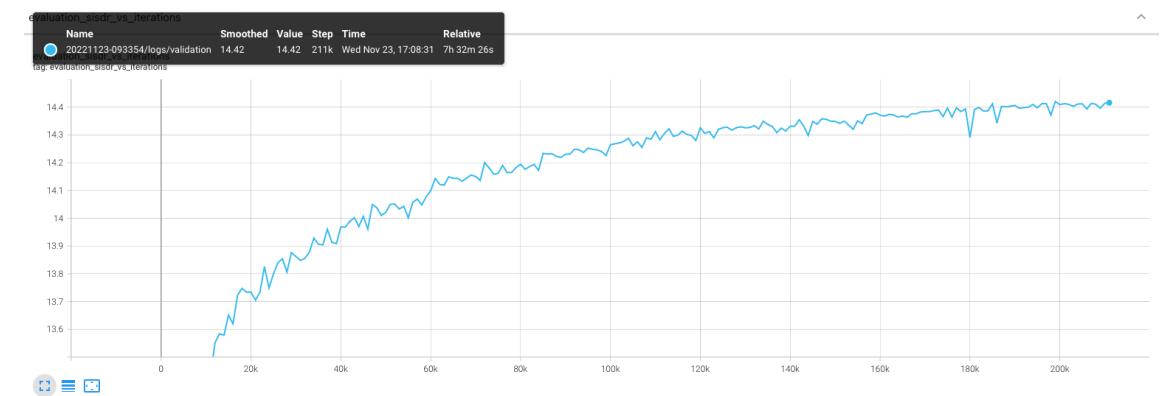
Loss, iteration



SISDR, epoch



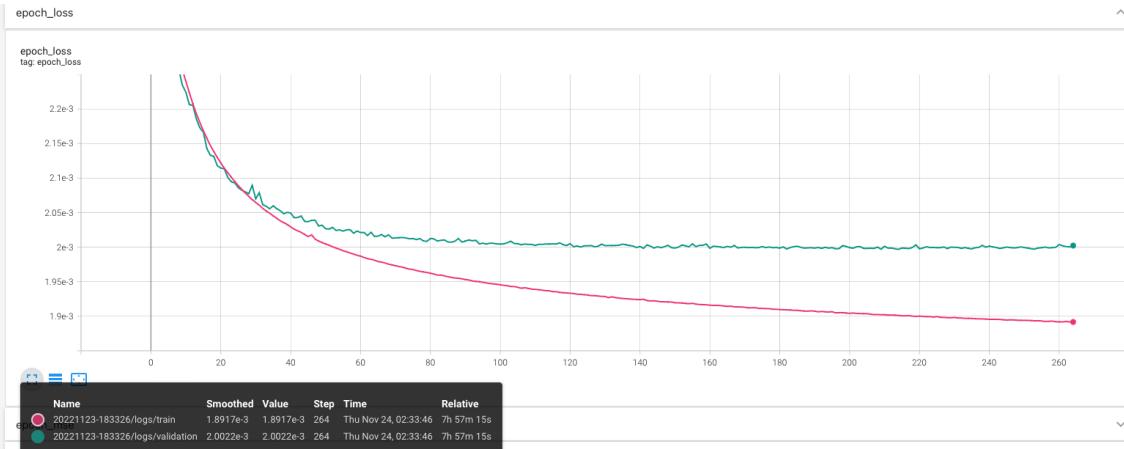
SISDR, iteration



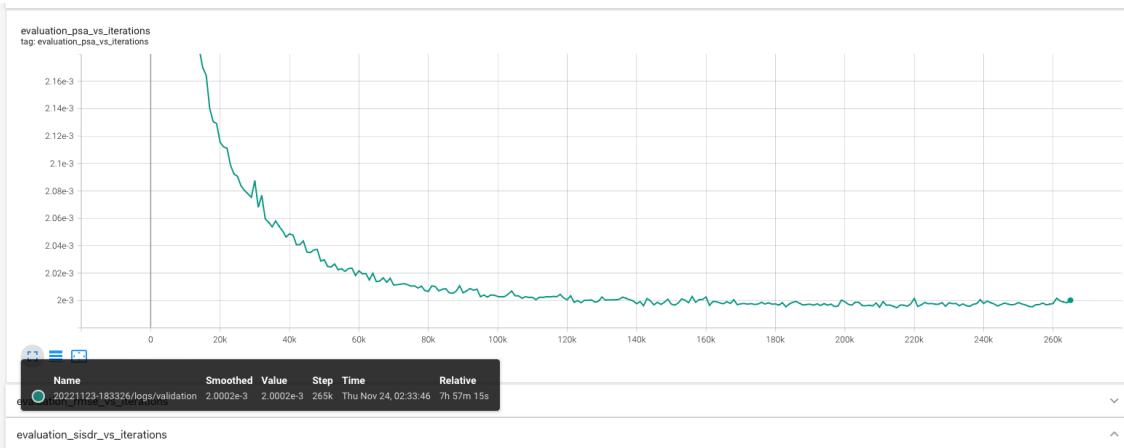
PSA

20221123-183326

Loss, epoch

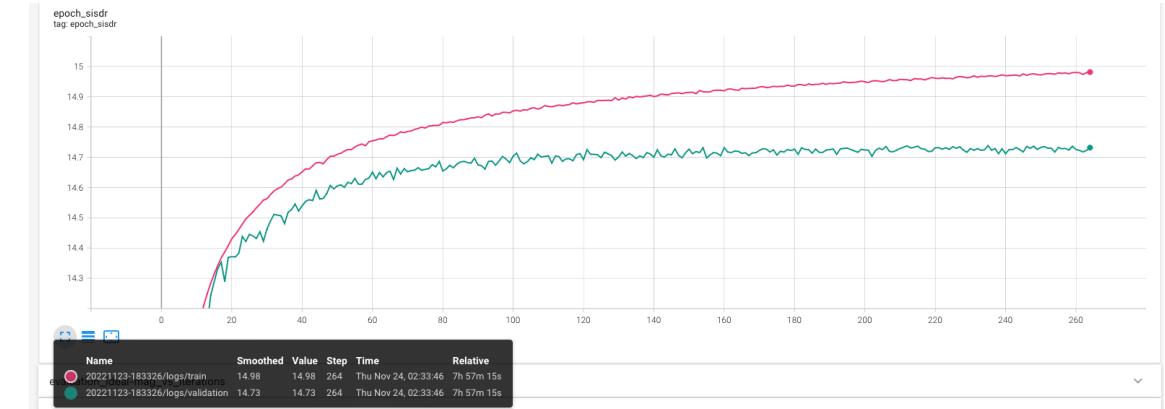


Loss, iteration

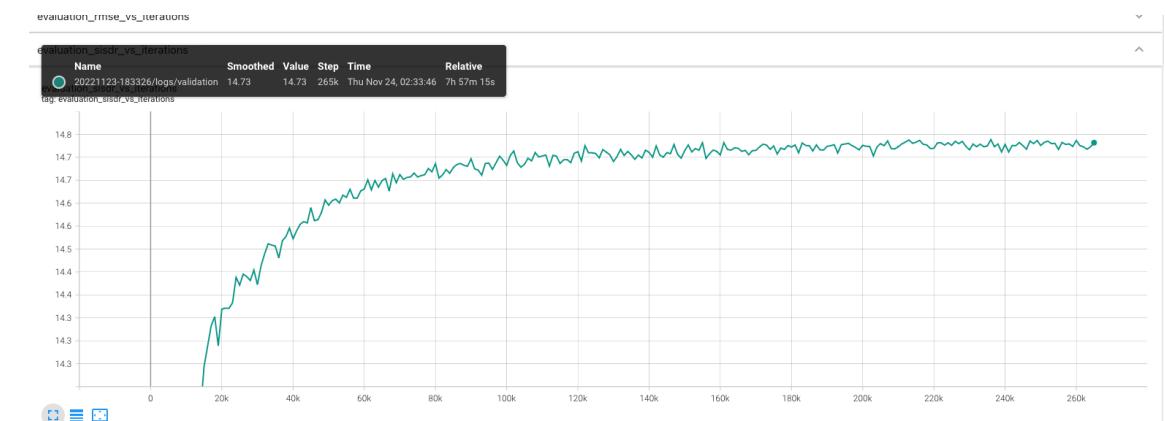


evaluation_sisdr_vs_iterations

SISDR, epoch



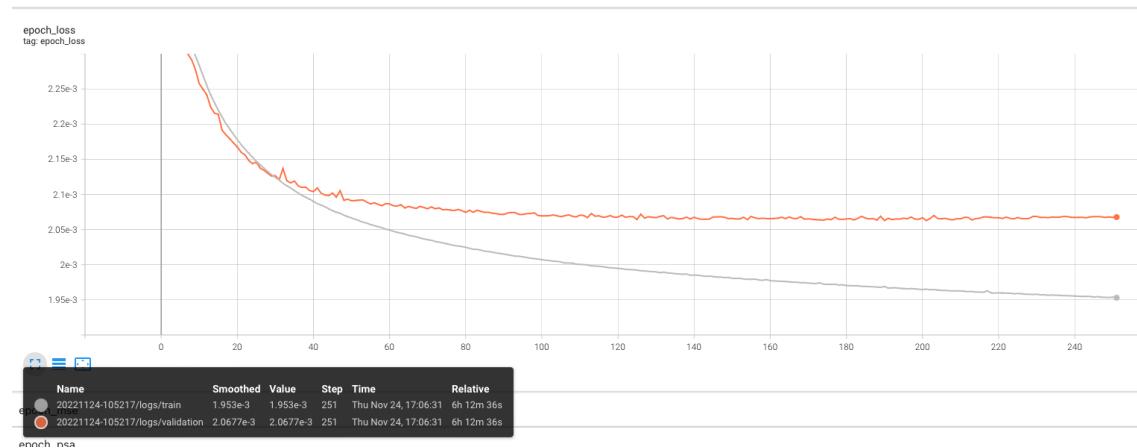
SISDR, iteration



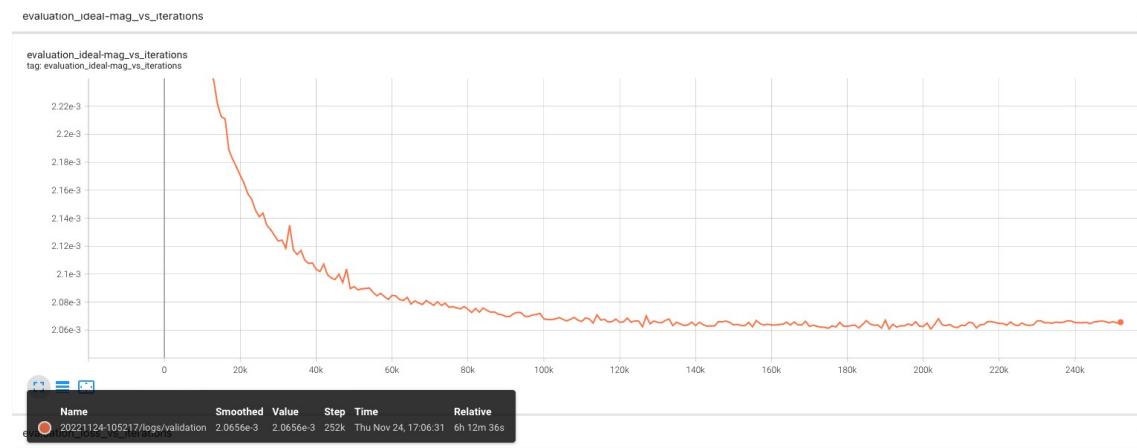
RMSE(amplitude)

20221124-105217

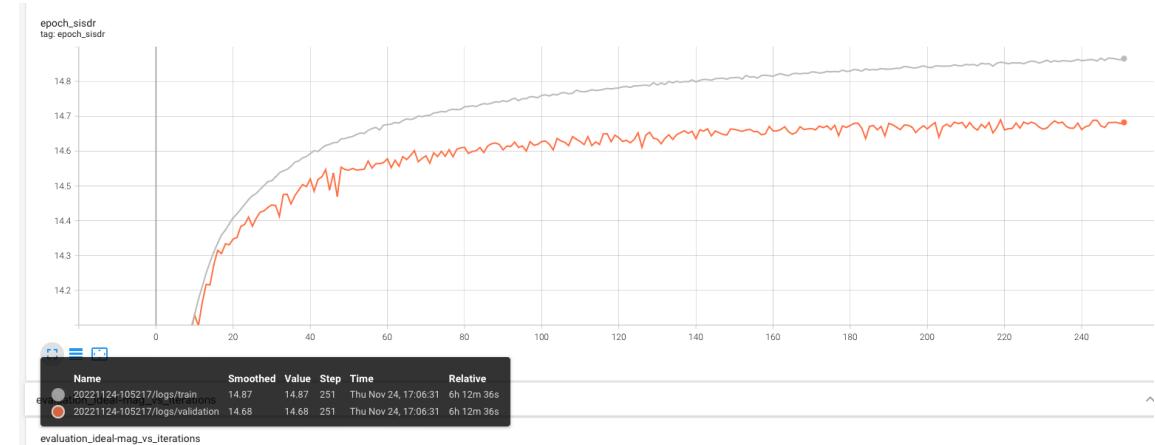
Loss, epoch



Loss, iteration



SISDR, epoch



SISDR, iteration

