Motivations

MSS empirical upper bound

- Common assumption:
  - The magnitude of source spectrogram < mixture spectrogram
  - |cIRM| < 1 or |IRM| < 1

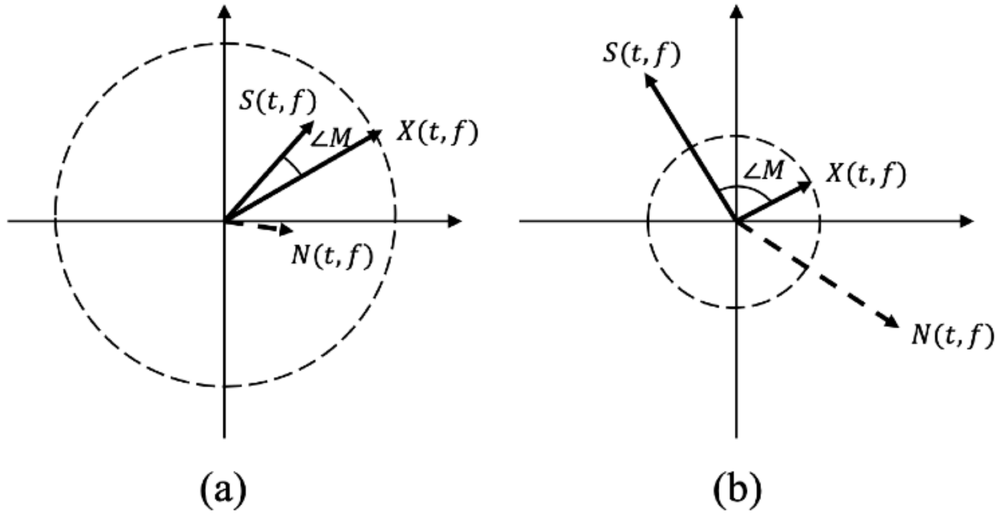

Figure 1. Illustrations of a source signal s, a noise n, and mixture x on a complex plain. (a) is an example when $|M(t,f)|$ smaller than 1 and (b) is an example when $|M(t,f)|$ larger than 1.

Motivations

MSS empirical upper bound

Table 1. The empirical upper bounds of MSS systems on MUSDB18. 'acc.' indicates accompaniment. On the top row, numbers indicate the limit of the magnitude masks.

	Mixture	IBM	IRM (1)	IRM (inf)	cIRM (1)	cIRM (2)	cIRM (5)	cIRM (10)	cIRM (inf)
vocals	-5.69	10.59	10.04	10.42	19.84	31.02	41.04	47.62	54.50
acc.	-5.68	16.10	15.31	15.97	26.54	37.62	47.33	53.51	60.63
bass	-6.36	7.17	6.05	6.07	17.99	27.88	37.86	44.30	54.12
drums	-4.30	8.75	8.03	8.61	19.10	30.38	39.91	46.45	56.08
other	-4.92	8.20	7.28	7.37	18.97	28.91	39.08	45.64	56.00

The magnitude of cIRMs are not always smaller than one.

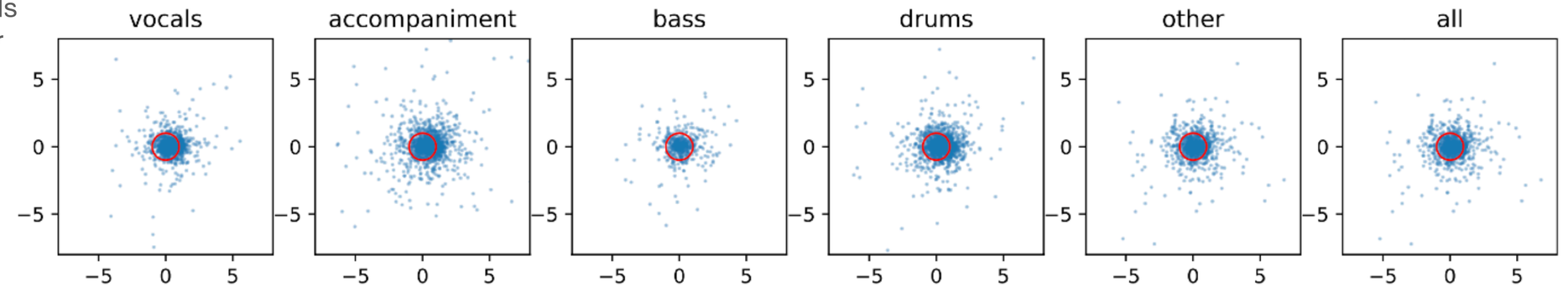

Figure 2. cIRMs of vocals, accompaniment, bass, drums, other, and all sources, on the complex 2D plain. Unit circles are drawn in red.