LIST OF ABBREVIATIONS

AM	Acoustic Model
ASR	Automatic Speech Recognition
ATF	Acoustic Transfer Function
BAN	Blind Analytic Normalization
BLSTM	Bi-directional LSTM
BSS	Blind Source Separation
CACGMM	Complex Angular Central GMM
CD	Cepstral Distortion
CE	Cross Entropy
CNN	Convolutional Neural Network
DAN	Deep Atractor Network
DC	Deep Clustering
DER	Diarization Error Rate
DL	Deep Learning
DNN	Deep Neural Network
DOA	Direction-Of-Arrival
DSP	Digital Signal Processing
EM	Expectation-Maximization
FF	Feed Forward
FWSSNR	Frequency-Weighted Segmental SNR
GEV	Generalized Eigenvalue Decomposition
GMM	Gaussian Mixture Model
ICA	Independent Component Analysis
IVA	Independent Vector Analysis
ILRMA	Independent Low-Rank Matrix Analysis
LP	Linear Prediction
LSTM	Long-Short Term Memory
ML	Maximum Likelihood
MMSE	Minimum Mean Squared Error
MPDR	Minimum Power Distortionless Response

MSE	Mean Squared Error
MVDR	Minimum Variance Distortionless Response
MWF	Multichannel Wiener Filter
NMF	Nonnegative Matrix Factorization
NN	Neural Network
PESQ	Perceptual Evaluation of Speech Quality
PIT	Permutation Invariant Training
PLDA	Probabilistic Linear Discriminant Analysis
PSD	Power Spectral Density
RIR	Room Impulse Response
RNN	Recurrent Neural Network
RSAN	Recursive Selective Attention Network
RTF	Relative Transfer Function
SCER	Speaker Confusion Error Rate
SDW	Speech Distortion Weighted
SDR	Signal to Distortion Ratio
SDW-MWF	Speech Distortion Weighted MWF
SNR	Signal to Noise Ratio
SPP	Speech Presense Probability
STFT	Short-Time Fourier Transformation
STOI	Short-Time Objective Intelligibility
TasNet	Time Domain Audio Separation Network
TF	Time-Frequency
TDOA	Time Difference Of Arrival
TDNN	Time-Delay Neural Network
VAD	Voice Activity Detection
WER	Word Error Rate
WPE	Weighted Prediction Error
WSJ	Wall Street Journal

LIST OF NOTATIONS

Mathematical expressions and operations	
$\top$ and H	Non-conjugate and conjugate transpose.
a	A scalar variable.
$\mathbf{a}$	A column vactor.
$\mathbf{A}$	A matrix.
D	A constant.
σ	A scalar parameter, such as a power spectral density (PSD) of a source.
Ψ	A matrix parameter, such as a spatial covariance matrix.
$\mathbb{E}[X]$	Expectation operator.
$\Pr(A = a)$	Probability
$p(x)$	Probability density function
$\mathcal{N}(\mathbf{x}; \mathbf{m}, \mathbf{R})$	Probability distribution of (multi-dimensional) (complex) normal distribution
$\mathrm{tr}\{\mathbf{\Phi}\}$	Trace of a matrix
$\|\cdot\|_2$	Eucredean norm of a vector
$\mathbb{R}$ and $\mathbb{C}$	A set of real scalars, and a set of complex scalars.
$\mathbb{R}^M$ and $\mathbb{R}^{M\times M}$	A set of M dimentional real vectors, and a set of $M \times M$ dimentional real matrices. $\mathbb{C}^M$ and $\mathbb{C}^{M\times M}$ are defined similarly.
$\nabla_{\mathbf{w}} J(\mathbf{w}) \in \mathbb{R}^{N\times 1}$	Gradient in denominator layout: Gradient is a column vector; Note: $\nabla_{\mathbf{w}} J(\mathbf{w}) = \dfrac{\partial}{\partial \mathbf{w}} J(\mathbf{w})$

Symbols for Short Time Fourier Transformation (STFT) domain signals	
t, f, m, and i	Indicies of time frames, frequency bins, microphones, and sources.
T, F, M, and I	The numbers of time frames, frequency bins, microphones, and sources.
$s^{(i)}_{t,f} \in \mathbb{C}$	A clean signal for the i-th source.
$x^{(i)}_{m,t,f} \in \mathbb{C}$	A microphone image of the i-th source at the m-th microphone, i.e, noiseless reverberant signal for the source captured at the microphone.
$n_{m,t,f} \in \mathbb{C}$	Diffuse noise.
$y_{m,t,f} \in \mathbb{C}$	A signal captured at the m-th microphone. When I sources and diffuse noise are included, it is typically modeled by $y_{m,t,f} = \sum_{i=1}^{I} x^{(i)}_{m,t,f} + n_{m,t,f}$.
$d^{(i)}_{m,t,f} \in \mathbb{C}$	A part of $x^{(i)}_{m,t,f}$ composed of its direct signal and early reflections.
$r^{(i)}_{m,t,f} \in \mathbb{C}$	A part of $x^{(i)}_{m,t,f}$ composed of its late reverberation.
$\mathbf{y}_{t,f} \in \mathbb{C}^M$	A vector composed of $y_{m,t,f}$ for all m, i.e., $\mathbf{y}_{t,f} = (y_{1,t,f}, \ldots, y_{M,t,f})^\top$. $\mathbf{n}_{t,f}$, $\mathbf{x}^{(i)}_{t,f}$, $\mathbf{d}^{(i)}_{n,f}$, and $\mathbf{r}^{(i)}_{n,f}$ are defined similarly.
$\mathbf{x}_{t,f} \in \mathbb{C}^M$	Sum of $\mathbf{x}^{(i)}_{t,f}$ for all i, namely $\mathbf{x}_{t,f} = \sum_{i=1}^{I} \mathbf{x}^{(i)}_{t,f}$.
Symbols for time domain signals	
$\tilde{t}$ and $\tilde{T}$	A time sample index and the number of time samples in time domain. The same symbols as those for STFT domain signals are used for m, i, M, and I.
$y_m[\tilde{t}]$	A signal captured at the m-th microphone. $x^{(i)}_m[\tilde{t}]$ and $n_m[\tilde{t}]$ are defined similarly.