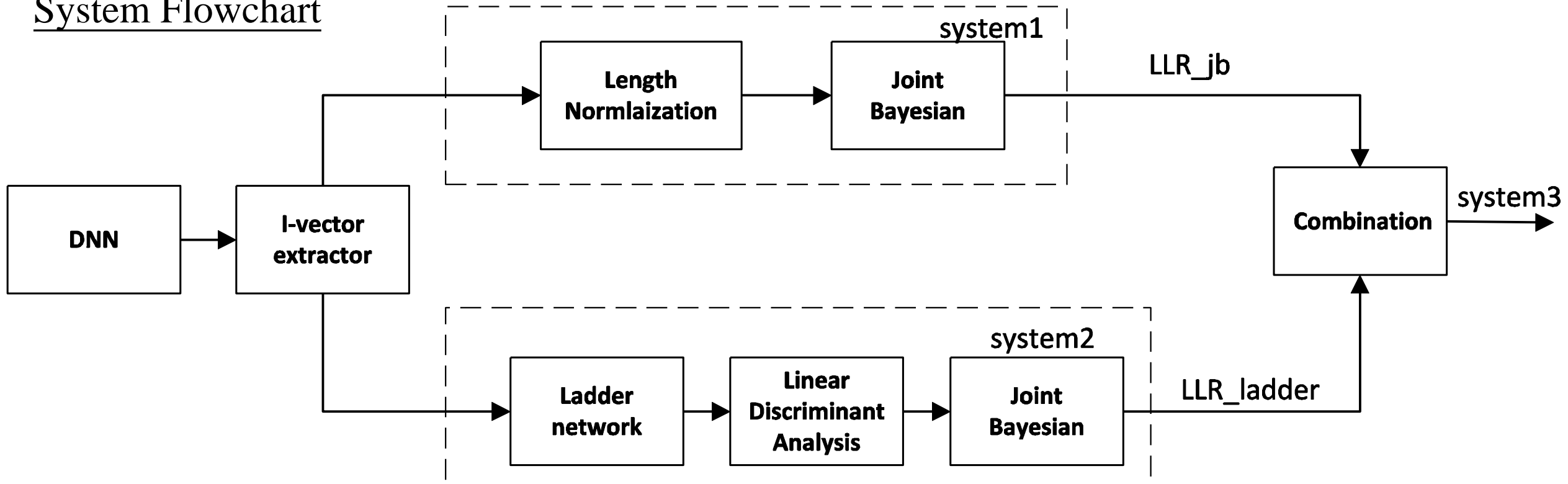# The THU-SPMI SRE-16 System with Joint Bayesian Scoring and Ladder Network based Feature Learning

Yiyan Wang, Haotian Xu, **Zhijian Ou**

Speech Processing and Machine Intelligence (SPMI) Lab, Tsinghua University, Beijing, China
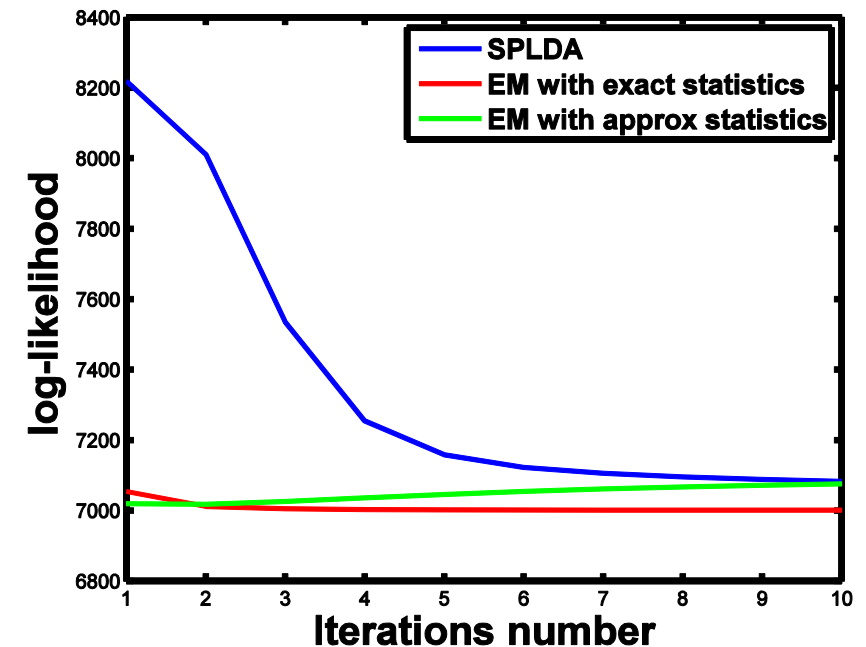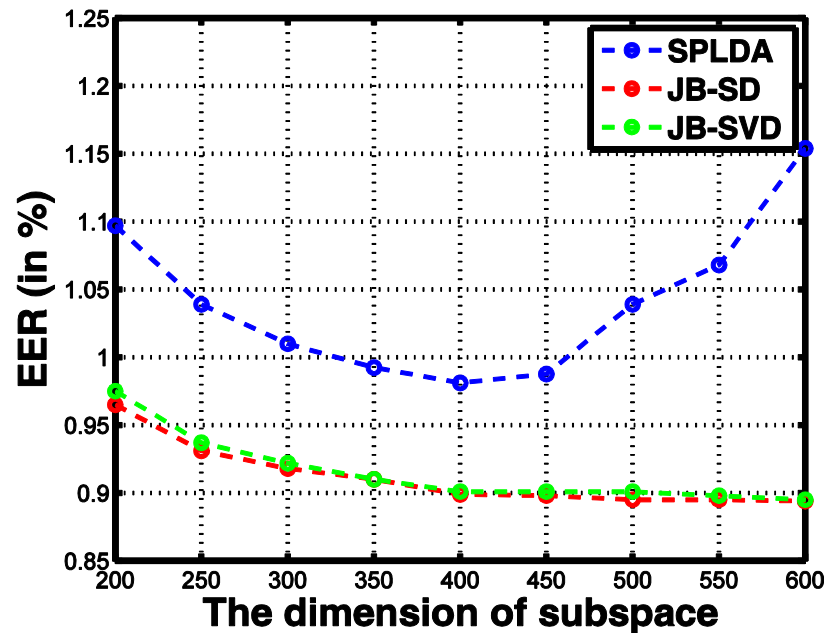
System Flowchart

$$x_{ij} = \mu_i + \varepsilon_{ij}$$

Drawbacks of PLDAs

- It is difficult to determine the subspace dimension which is crucial for performance.
- Suffer from slow convergence of their implemented EM iterations.
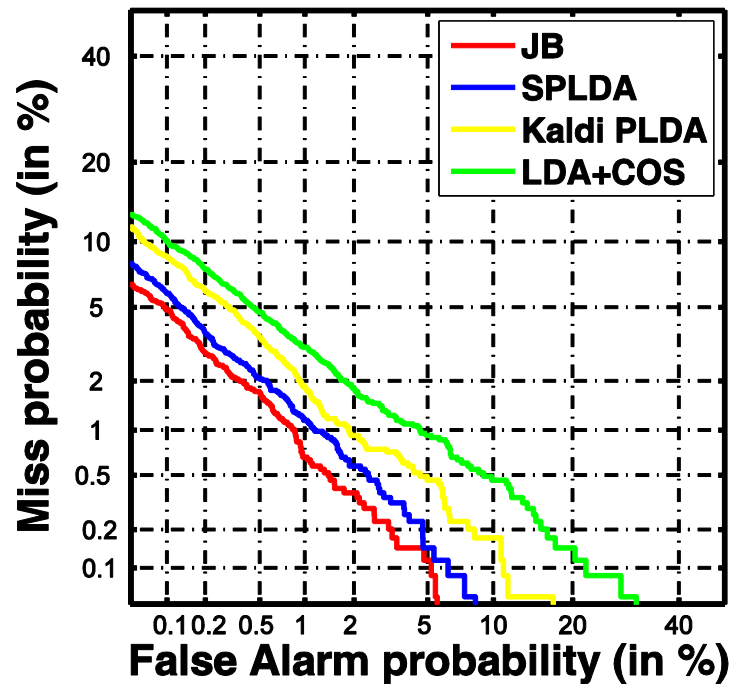
Experiment results on SRE10

D. Chen, X. Cao, D. Wipf, F. Wen, and J. Sun,   "An efficient joint formulation for bayesian face verification,"   IEEE Trans. PAMI, 2016.

Our Novelties   2. Ladder Network based Feature Learning

leveraging both labeled (SWD-2 and previous SREs) and un-labeled i-vectors (CallMyNet un-labeled)

$$L(\boldsymbol{x}, \boldsymbol{y}) = -logp(\tilde{\boldsymbol{y}}|\boldsymbol{x}) + \sum_{i=1}^{N}||\boldsymbol{z}^{(i)} - \hat{\boldsymbol{z}}_{BN}^{(i)}||^2 + \alpha||\boldsymbol{h}^{N-1} - \boldsymbol{c}_y||^2$$

$$\boldsymbol{h}^{N-1} = f(\gamma^{N-1}(\tilde{\boldsymbol{z}}^{N-1} + \beta^{N-1}))$$
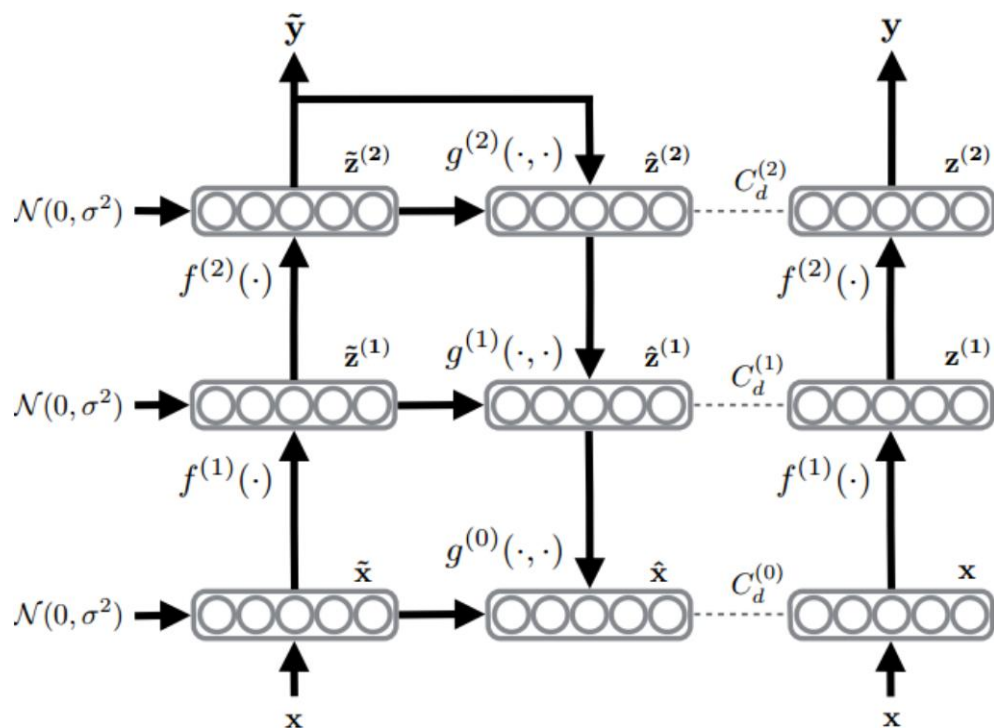
SRE16 results

	System 1(secondary)	System 2	System 3(primary)
dev			
Equalized			
eer	20.29	19.13	19.18
min_Cprimary	0.8243	0.8225	0.8080
act_Cprimary	0.9095	0.9992	0.9986
Un-equalized			
eer	19.41	19.78	19.55
min_Cprimary	0.8170	0.8213	0.8045
act_Cprimary	0.9209	0.9994	0.9989
eval			
Equalized			
eer	15.39	17.24	14.71
min_Cprimary	0.7826	0.8570	0.7747
act_Cprimary	0.8993	0.9154	0.8860
Un-equalized			
eer	15.19	17.90	14.58
min_Cprimary	0.8025	0.8795	0.7949
act_Cprimary	0.9205	0.9474	0.9089

A. Rasmus, M. Berglund, M. Honkala, H. Valpola, and T. Raiko, "Semi-supervised learning with ladder networks," NIPS 2015.