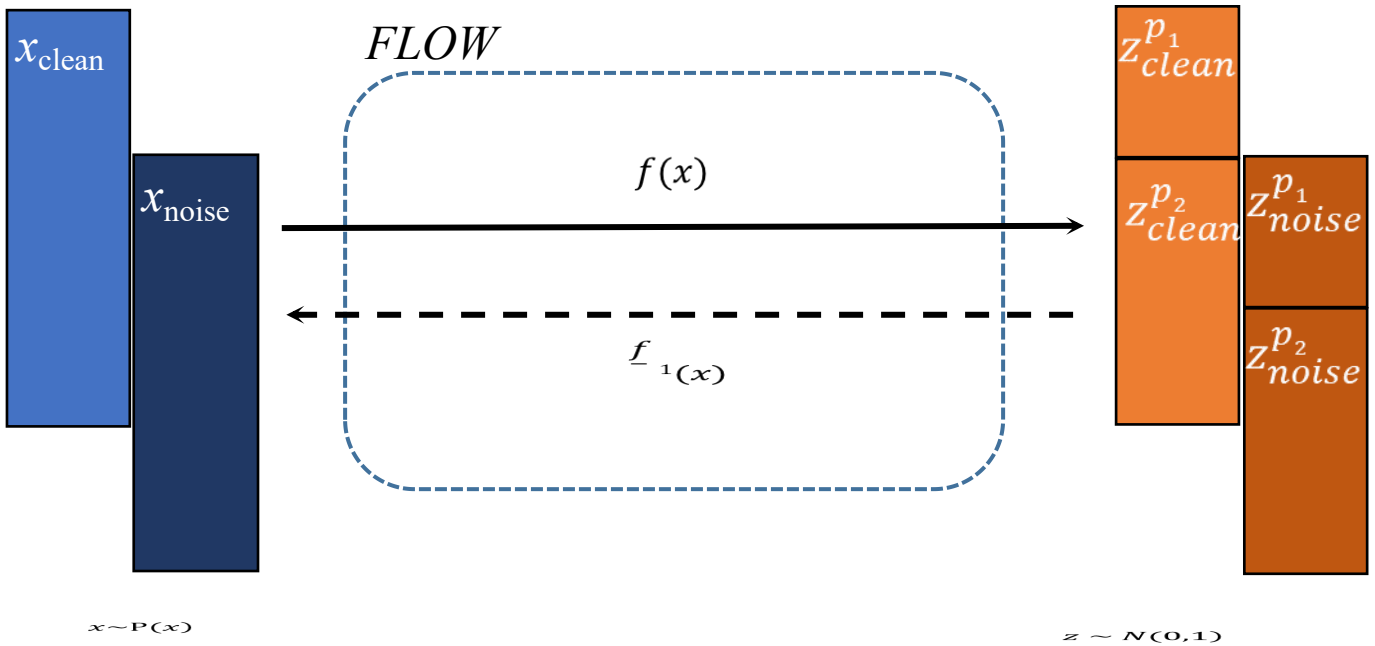
Data: /work4/sunhaoran/spec-flow/data/timit/phone_train/
Model: MAF (code from Yunqi Cai)
Test data: /work4/sunhaoran/spec-flow/data/timit/phone_test/ (sample 19 utterance correspond to 19 vowels)



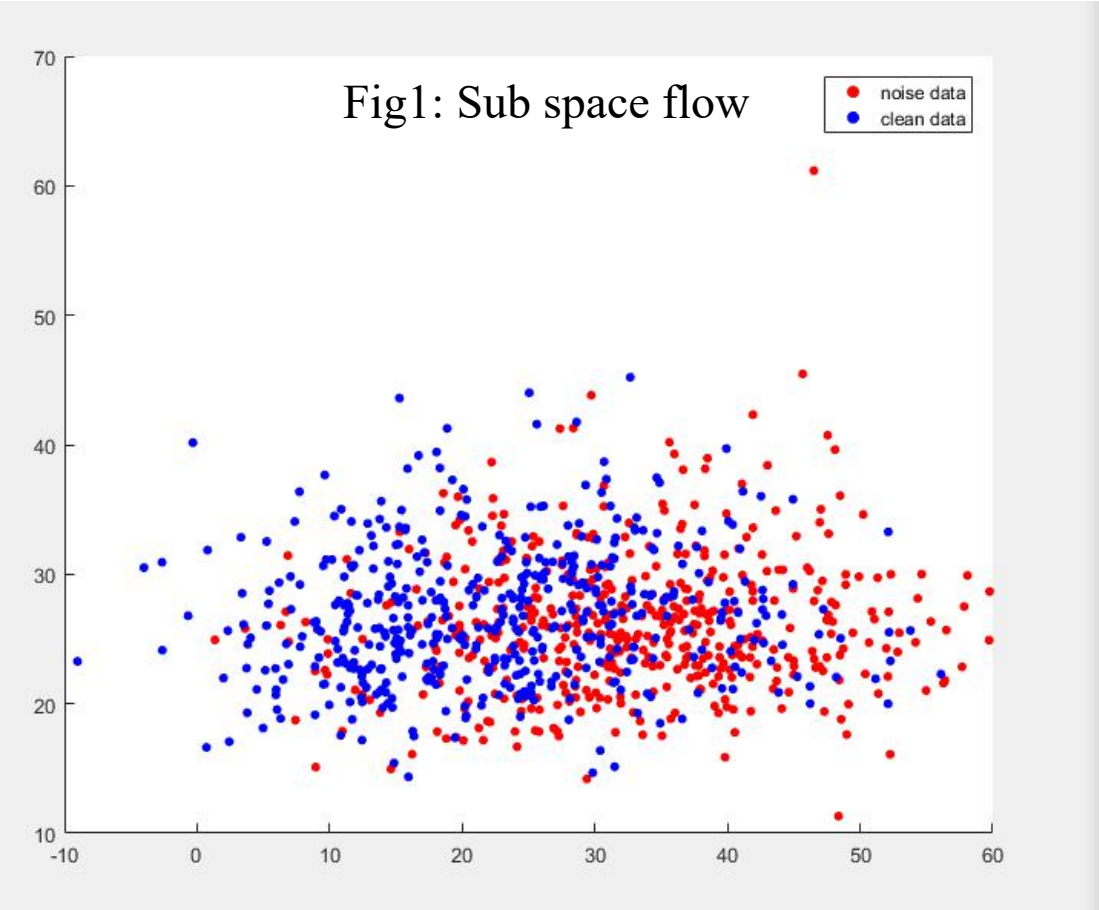
Maximize: $|\ (z^{p_1}_{clean})^2 - (z^{p_1}_{noise})^2\ |$

Minimize: $|\ (z^{p_2}_{clean})^2 - (z^{p_2}_{noise})^2\ |$

Loss function

$$log\mathrm{P}(x) = \Sigma_x\, log\left(\mathrm{P}z(f(x))\right) + \Sigma_x\, log\left|det\,\frac{\partial f}{\partial x^{\mathrm{T}}}\right| + \lambda_1|\ (z^{p_1}_{clean})^2 - (z^{p_1}_{noise})^2\ | - \lambda_2|\ (z^{p_2}_{clean})^2 - (z^{p_2}_{noise})^2\ | - \lambda_3|\ (z^{p_1}_{clean})^2$$

$$x \in \{clean,\ noise\}$$



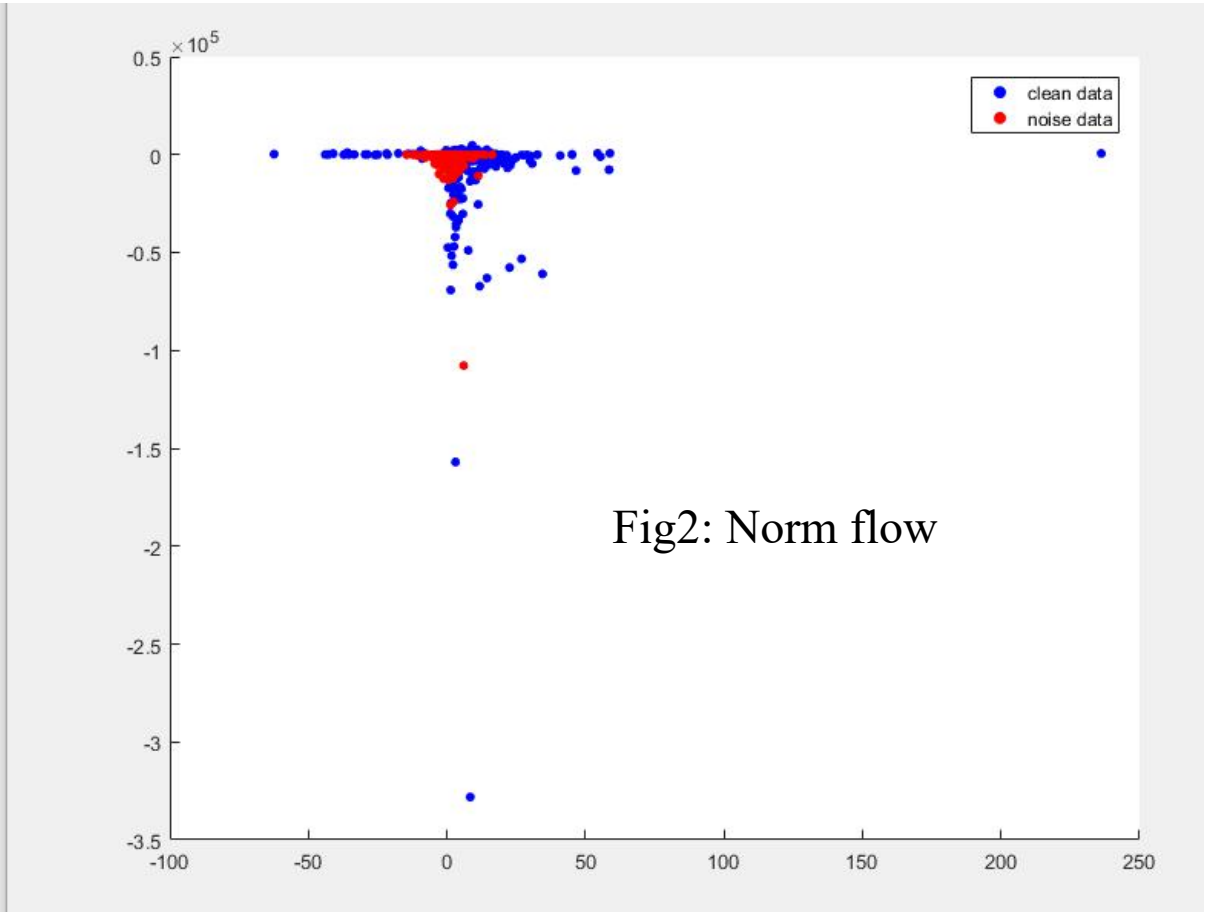


Fig1: Dimension of Z space is equal to X which is 771 (-1, 0, 1) In this fig. The axis x is z[50] (the first 100th. **P1** Clean & noise space) And the axis y is z[250] (dimensions after 100th **P2**. Which is Semantic space) According to our assumption the clean & noise space should be wide at the same time the clean data should be more close to origin(0, 0). While the semantic space should be narrow.

In Fig1. the distribution of axis x is wide the range is from -10 to 60 but the clean data is not close enough to origin(0, 0). However the distribution of axis y is narrow the x range is from 10 to 50.

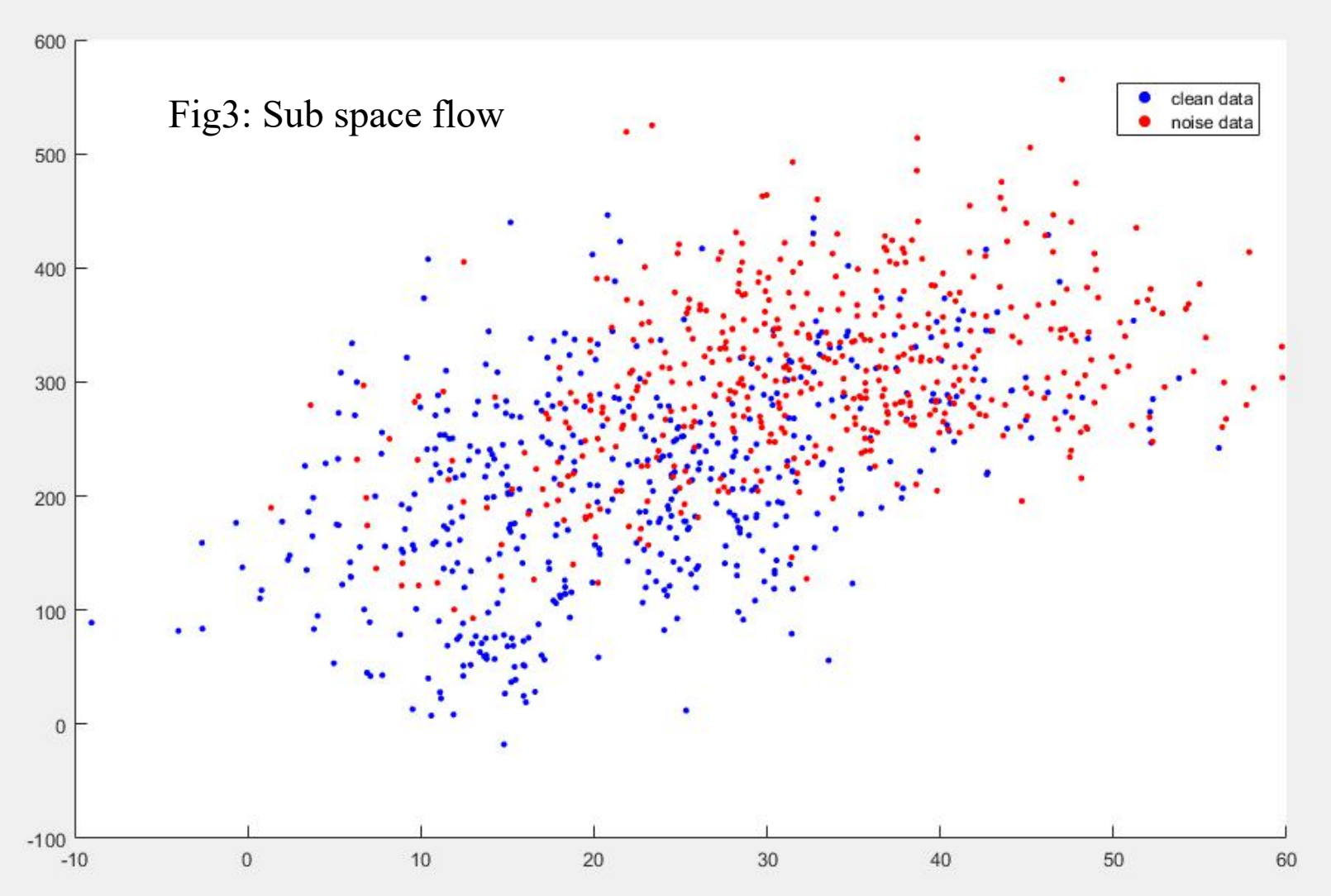
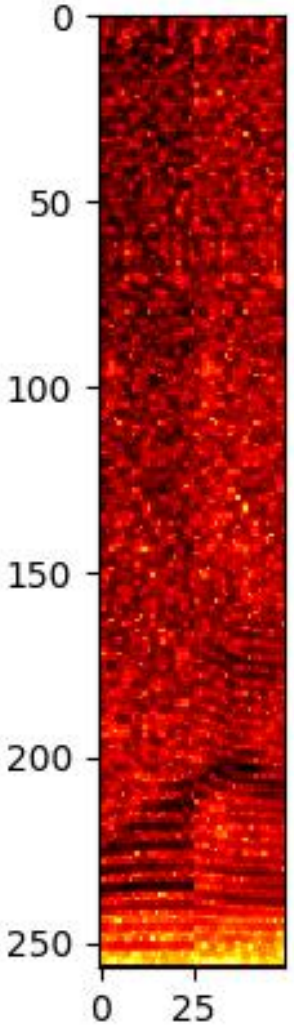
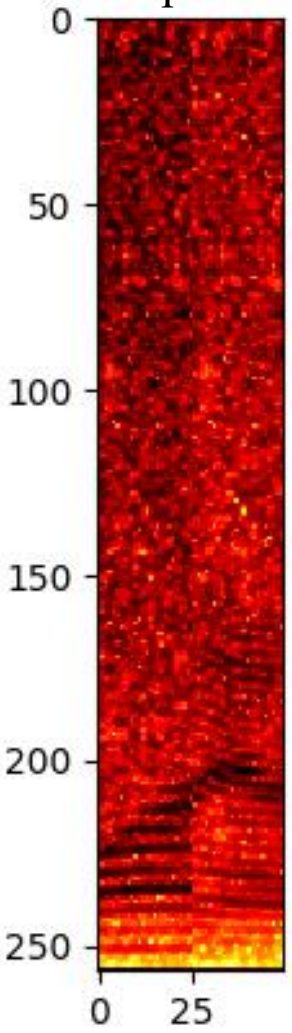


Fig3: Both axis x and axis y is select from P1 x = z[50] y=z[80]
Clean data is more close to (0, 0) so the loss function is useful.

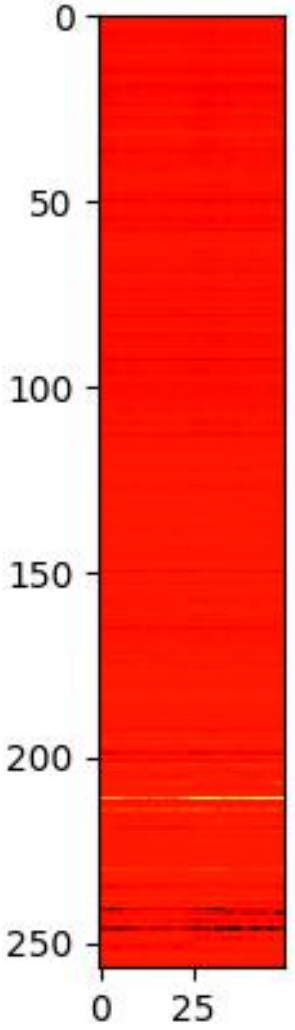
Org noise spectrum



Recover from z noise spectrum



Recover from z but set p1 to **0** noise spectrum



Summary: Analysis from the figs. The loss function is useful.
In some respects it's not powerful enough because in P1 space the clean data is not very close to (0,0)
In other respects it's to powerful because the loss function destroy the org. likelihood structure so that when we modify the P1 space the z can't be transform back to x space.
In my mind the under current structure the P1 is not the sub space of flow. Because the distribution in P1 is strange and weak. Maybe we should control the covariance matrix of P1.