Confusion Matrix - noisy labels

January 18, 2016

1 White noise

Labels distribution

	Test	$Train^{(s)}$	N1	N2	N3	N4	N5	N6
size	4661	4152	..	..	..	..	..	..
1	246	214	220	195	207	614	757	1015
0	4415	3938	3932	3957	3945	3538	3395	3137

True confusion matrices

$$Q_1 = \begin{pmatrix} .99 & .01 \\ .19 & .81 \end{pmatrix} Q_2 = \begin{pmatrix} .98 & .02 \\ .38 & .62 \end{pmatrix} Q_3 = \begin{pmatrix} .97 & .03 \\ .57 & .43 \end{pmatrix}$$
$$Q_4 = \begin{pmatrix} .9 & .1 \\ .01 & .99 \end{pmatrix} Q_5 = \begin{pmatrix} .85 & .015 \\ .15 & .985 \end{pmatrix} Q_6 = \begin{pmatrix} .8 & .02 \\ .2 & .98 \end{pmatrix}$$

Learning

Training for 10000 iters ($\approx$ 150 epochs)- Architecture: $3 \times \{Conv, MaxPool, ReLU\}$ (the baseline) $\rightarrow$ InfoGain loss ($H := Q$).

$$\mathcal{L} = -\frac{1}{N}\sum_i \left(H_{l^{(n)},0}.\log \hat{p}_0^{(n)} + H_{l^{(n)},1}.\log \hat{p}_1^{(n)} \right)$$

Where $l^{(n)}$ and $\hat{p}^{(n)}$ are the label and class probabilities of sample n

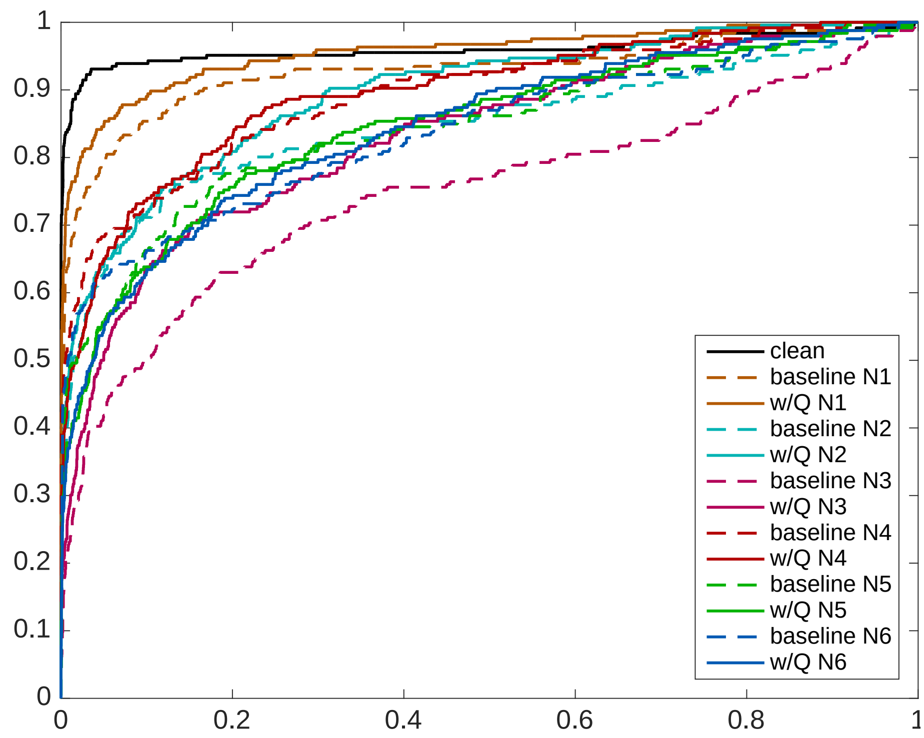

Figure 1: Roc curves: variant noise levels - baseline vs. Q

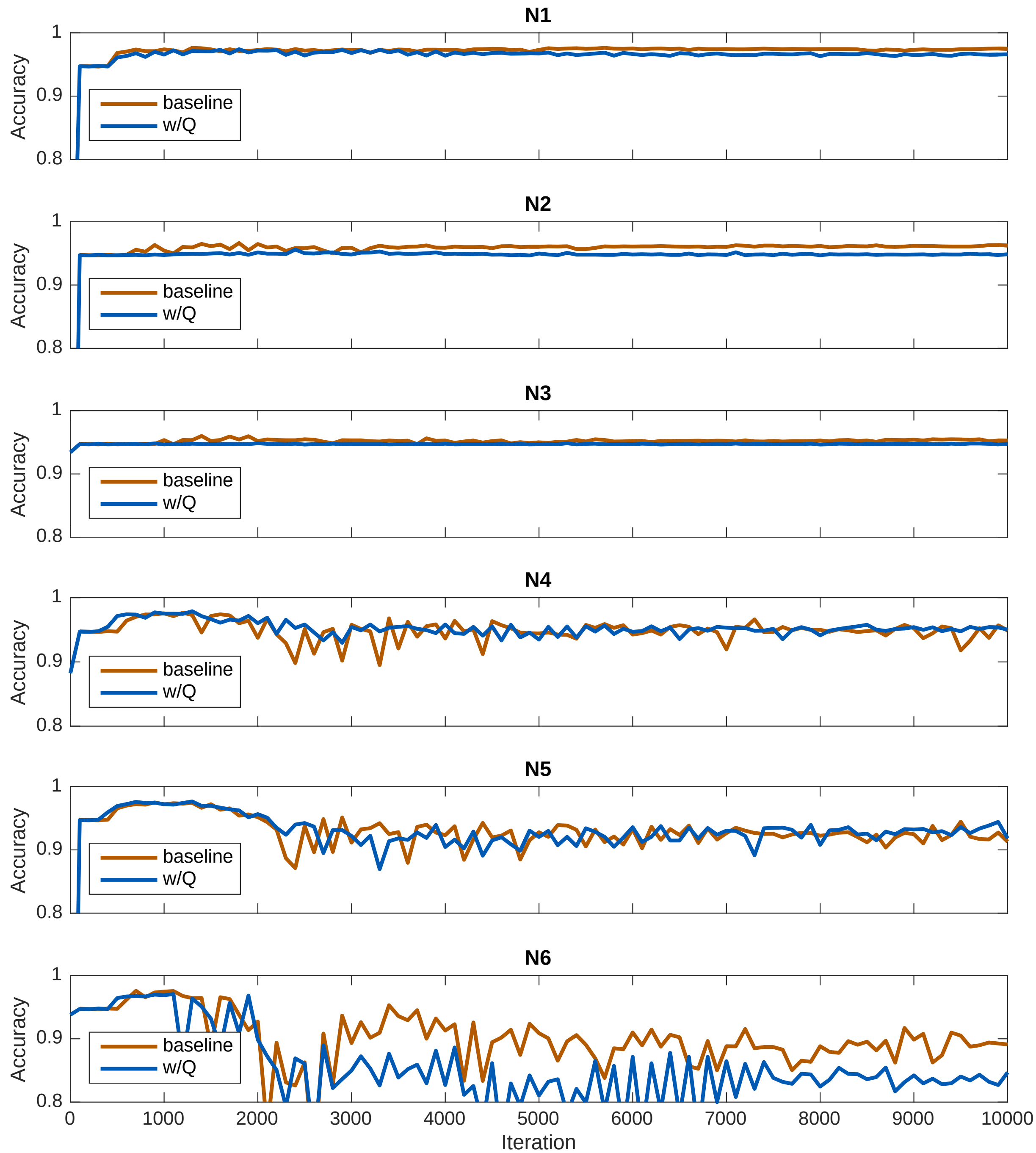

Figure 2: Accuracies: variant noise levels - baseline vs. Q

2 Image dependent noise

Thresholding the old classifier scores : 0 if score < th1 , 1 if score > th2

N1: th1 = 10, th2 = 30

N2: th1 = th2 = 20

N3: th1 = th2 = 30

	$Test$	$Train^{(s)}$	$Train^{(f)}$	N1	N2	N3
size	4661	4152	381949	249683	381916	381942
1	246	214	-	15862	32132	15827
0	4415	3938	-	233821	349784	366115

We estimate Q on a hand-labeled subset $Train^{(s)} \subset Train^{(f)}$

$$Q_1 = \begin{pmatrix} .979 & .021 \\ .08 & .92 \end{pmatrix} \; Q_2 = \begin{pmatrix} .9502 & .0498 \\ .2103 & .7897 \end{pmatrix} \; Q_3 = \begin{pmatrix} .9963 & .0037 \\ .3925 & .6075 \end{pmatrix}$$

Training for 10000 iters ($\approx$ 2 epochs):

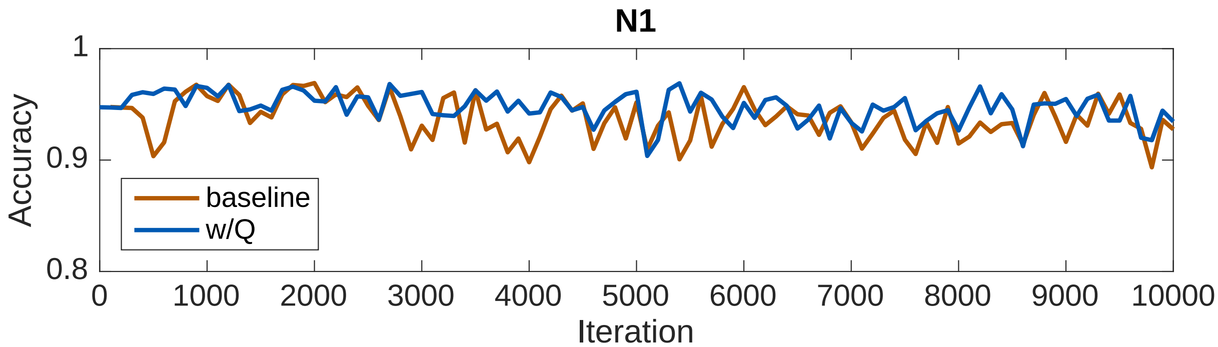

Figure 3: Accuracies: baseline vs. Q

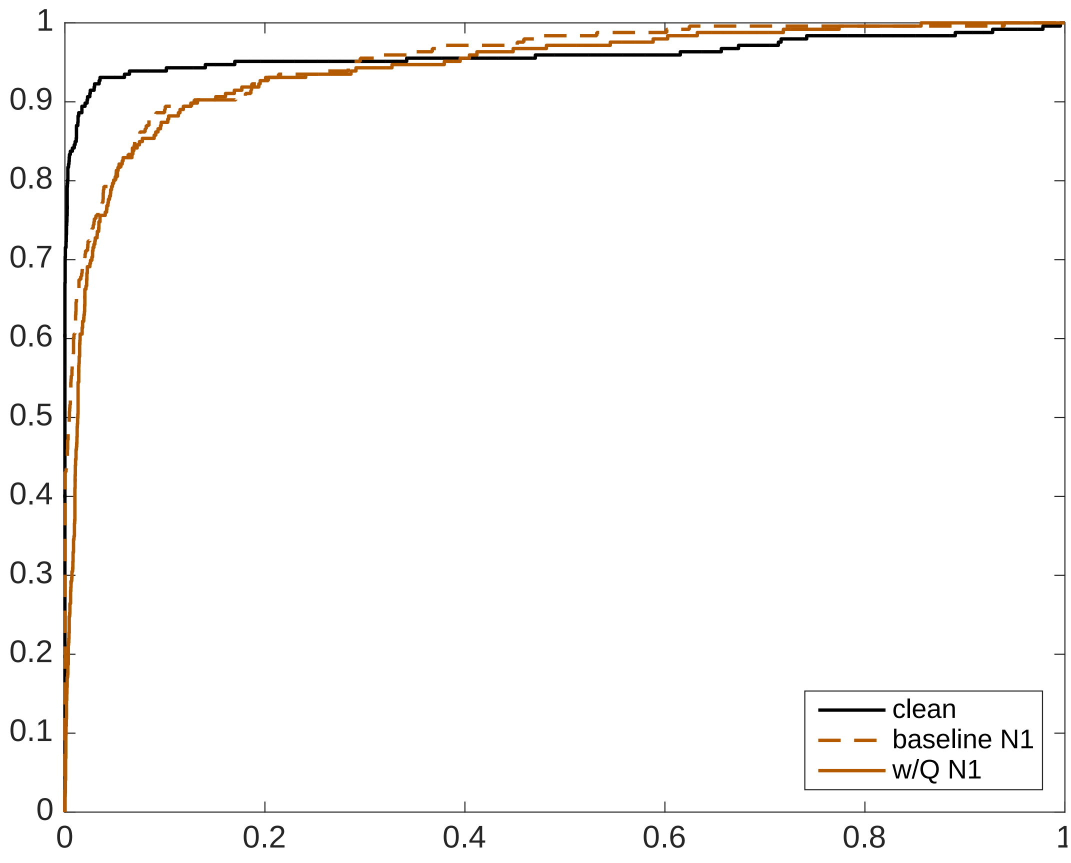

Figure 4: Roc curves: baseline vs. Q