

```

[1] "# of discriminatory genes = 50"

[1] "fold change for discriminatory genes: 0.3"

[1] "fold change for nondiscriminatory genes: 0"

choose_k
  2
100

```

The optimal number of clusters K is set as the most frequently found K from the 100 simulations tabulated above. It is found by using the BIC criterion after running the unpenalized EM algorithm on K spanning from 2 to 8.

Using the last set of simulated counts, I ran a grid search across varying tuning parameters. As done in Pan et al, I fixed $\lambda_1 = 1$, and searched over $\lambda_2 = (0.1, 0.2, \dots, 2)$ and $\tau = (0.1, 0.2, \dots, 2)$:

```

[1] "tau= 0.1"

[1] "lambda1= 1"

[1] "lambda2= 0.4"

```

The results of the final run based on optimal tuning parameters are below:
Below are the summary of results:

```

[1] "Mean pi: 0.603913043478261" "Mean pi: 0.396086956521739"

[1] "First 3 genes:"

      [,1]      [,2]
[1,] 3.577234 4.149592
[2,] 4.794702 5.390808
[3,] 4.304627 4.898223

[1] "Last 3 genes:"

      [,1]      [,2]
[98,] 6.399984 6.399536
[99,] 5.951904 5.948828
[100,] 6.378082 6.377756

[1] "Mean % of nondiscriminatory genes: 0.548"

[1] "Mean ARI: 1"

[1] "Mean sensitivity: 0.904"

[1] "Mean false positive rate: 0"

```