

```

[1] "# of discriminatory genes = 50"

[1] "fold change for discriminatory genes: 0.5"

[1] "fold change for nondiscriminatory genes: 0.1"

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] "lambda1, lambda2, tau, BIC:"

[1]      1.00      0.40      0.10 19579.62

```

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

```

[1] "Mean pi: 0.608695652173913" "Mean pi: 0.391304347826087"

[1] "First 3 genes:"

      [,1]      [,2]
[1,] 3.371845 4.350029
[2,] 4.599084 5.595473
[3,] 4.108733 5.098699

[1] "Last 3 genes:"

      [,1]      [,2]
[98,] 6.308045 6.491506
[99,] 5.862732 6.040337
[100,] 6.287035 6.473635

[1] "Mean % of nondiscriminatory genes: 0.3647"

[1] "Mean ARI: 1"

[1] "Mean sensitivity: 0.9884"

[1] "Mean false positive rate: 0.2822"

```