

```

[1] "# of discriminatory genes = 50"
[1] "fold change for discriminatory genes: 1"
[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] [,2] [,3] [,4]
[1,]    1  0.3  0.3 19355.82
[2,]    1  0.3  0.4 19355.82
[3,]    1  0.3  0.5 19355.82
[4,]    1  0.3  0.8 19355.82
[5,]    1  0.3  0.9 19355.82
[6,]    1  0.4  0.7 19355.82

```

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

```

[1] "Mean pi: 0.621739130434783" "Mean pi: 0.378260869565217"
[1] "First 3 genes:"
      [,1] [,2]
[1,] 2.890858 4.852961
[2,] 4.103816 6.089300
[3,] 3.622098 5.597359
[1] "Last 3 genes:"
      [,1] [,2]
[98,] 6.314458 6.487724
[99,] 5.871518 6.036315
[100,] 6.293817 6.468361
[1] "Mean % of nondiscriminatory genes: 0.4978"
[1] "Mean ARI: 1"
[1] "Mean sensitivity: 1"
[1] "Mean false positive rate: 0.0044"

```