

```

[1] "# of discriminatory genes = 25"

[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]      1.00      0.30      0.10 19336.78

```

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

```

[1] "Mean pi: 0.599130434782609" "Mean pi: 0.400869565217391"

[1] "First 3 genes:"

      [,1]      [,2]
[1,] 2.903470 4.848901
[2,] 4.106535 6.089263
[3,] 3.625700 5.597039

[1] "Last 3 genes:"

      [,1]      [,2]
[98,] 6.312364 6.482265
[99,] 5.867003 6.033000
[100,] 6.292706 6.463111

[1] "Mean % of nondiscriminatory genes: 0.7486"

[1] "Mean ARI: 1"

[1] "Mean sensitivity: 1"

[1] "Mean false positive rate: 0.001866666666666667"

```