

```

[1] "# of discriminatory genes = 10"

[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] [,2] [,3]      [,4]
[1,]      1 0.3 0.1 19202.9
[2,]      1 0.3 0.2 19202.9
[3,]      1 0.3 0.3 19202.9
[4,]      1 0.3 0.4 19202.9
[5,]      1 0.3 0.5 19202.9
[6,]      1 0.4 0.1 19202.9
[7,]      1 0.4 0.2 19202.9
[8,]      1 0.4 0.3 19202.9
[9,]      1 0.4 0.4 19202.9
[10,]     1 0.4 0.5 19202.9
[11,]     1 0.5 0.1 19202.9
[12,]     1 0.5 0.2 19202.9
[13,]     1 0.5 0.3 19202.9
[14,]     1 0.5 0.4 19202.9
[15,]     1 0.6 0.1 19202.9
[16,]     1 0.6 0.2 19202.9
[17,]     1 0.6 0.3 19202.9
[18,]     1 0.7 0.1 19202.9
[19,]     1 0.7 0.2 19202.9
[20,]     1 0.8 0.1 19202.9

```

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

```

[1] "Mean pi: 0.609565217391304" "Mean pi: 0.390434782608696"

[1] "First 3 genes:"

```

```

      [,1]      [,2]
[1,] 3.371681 4.350396
[2,] 4.598244 5.593376
[3,] 4.107202 5.098627

[1] "Last 3 genes:"

      [,1]      [,2]
[98,] 6.311017 6.484711
[99,] 5.869878 6.030200
[100,] 6.291170 6.461884

[1] "Mean % of nondiscriminatory genes: 0.8981"

[1] "Mean ARI: 1"

[1] "Mean sensitivity: 1"

[1] "Mean false positive rate: 0.002111111111111111"

```