

```

[1] "# of discriminatory genes = 50"

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

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

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= 0.6"

```

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

```

[1] "Mean pi: 0.611304347826087" "Mean pi: 0.388695652173913"

[1] "First 3 genes:"

      [,1]      [,2]
[1,] 2.892689 4.844944
[2,] 4.107289 6.091442
[3,] 3.623112 5.600358

[1] "Last 3 genes:"

      [,1]      [,2]
[98,] 5.901085 6.897152
[99,] 5.453914 6.448912
[100,] 5.879825 6.878043

[1] "Mean % of nondiscriminatory genes: 6e-04"

[1] "Mean ARI: 1"

[1] "Mean sensitivity: 1"

[1] "Mean false positive rate: 0.9988"

```