

```

[1] "# of discriminatory genes = 10"
[1] "fold change for discriminatory genes: 0.5"
[1] "fold change for nondiscriminatory genes: 0"

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= 0.2" "tau= 0.3" "tau= 0.4" "tau= 0.5" "tau= 0.6" "tau= 0.7" "tau= 0.1"
[1] "lambda1= 1"
[1] "lambda2= 0.1" "lambda2= 0.2" "lambda2= 0.3" "lambda2= 0.4" "lambda2= 0.5"
[6] "lambda2= 0.6" "lambda2= 0.7"

```

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

```

[1] "Mean pi: 0.609130434782609" "Mean pi: 0.390869565217391"
[1] "First 3 genes:"

      [,1]      [,2]
[1,] 3.372743 4.349844
[2,] 4.595638 5.592601
[3,] 4.112293 5.103663

[1] "Last 3 genes:"

      [,1]      [,2]
[98,] 6.399703 6.399942
[99,] 5.951753 5.949702
[100,] 6.377799 6.377593

[1] "Mean % of nondiscriminatory genes: 0.8921"
[1] "Mean ARI: 1"
[1] "Mean sensitivity: 1"
[1] "Mean false positive rate: 0.00877777777777778"

```