

```

[1] "# of discriminatory genes = 1"
[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= 0.1" "tau= 0.2" "tau= 0.3" "tau= 0.4" "tau= 0.5" "tau= 0.6" "tau= 0.7"
[8] "tau= 0.8"
[1] "lambda1= 1"
[1] "lambda2= 0.2" "lambda2= 0.3" "lambda2= 0.4" "lambda2= 0.5" "lambda2= 0.6"
[6] "lambda2= 0.7" "lambda2= 0.8" "lambda2= 0.9"

```

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

```

[1] "Mean pi: 0.596084289419111" "Mean pi: 0.40391571058089"
[1] "First 3 genes:"
      [,1]      [,2]
[1,] 3.368560 4.349557
[2,] 5.091587 5.093447
[3,] 4.608483 4.606823
[1] "Last 3 genes:"
      [,1]      [,2]
[98,] 6.401470 6.400205
[99,] 5.951471 5.952577
[100,] 6.378398 6.380115
[1] "Mean % of nondiscriminatory genes: 0.9884"
[1] "Mean ARI: 0.991289148214381"
[1] "Mean sensitivity: 1"
[1] "Mean false positive rate: 0.00161616161616162"

```