

```

[1] "# of discriminatory genes = 1"
[1] "fold change for discriminatory genes: 0.3"
[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"
[7] "tau= 0.8" "tau= 0.9" "tau= 1" "tau= 1.1" "tau= 1.2" "tau= 1.3"
[13] "tau= 1.4" "tau= 1.5" "tau= 1.6" "tau= 1.7" "tau= 1.8" "tau= 1.9"
[19] "tau= 2" "tau= 0.1"
[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" "lambda2= 1" "lambda2= 1.1"
[11] "lambda2= 1.2" "lambda2= 1.3" "lambda2= 1.4" "lambda2= 1.5" "lambda2= 1.6"
[16] "lambda2= 1.7" "lambda2= 1.8" "lambda2= 1.9" "lambda2= 2"

```

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

Below are the summary of results:

```

[1] "Mean pi: 0.560008053926193" "Mean pi: 0.439991946073806"
[1] "First 3 genes:"
      [,1]      [,2]
[1,] 3.634130 4.029085
[2,] 5.095036 5.099472
[3,] 4.606607 4.602592
[1] "Last 3 genes:"
      [,1]      [,2]
[98,] 6.402102 6.398765
[99,] 5.951465 5.951638
[100,] 6.380186 6.377789
[1] "Mean % of nondiscriminatory genes: 0.9914"
[1] "Mean ARI: 0.531398847059508"
[1] "Mean sensitivity: 0.74"
[1] "Mean false positive rate: 0.00121212121212121"

```