Small RNA dataset

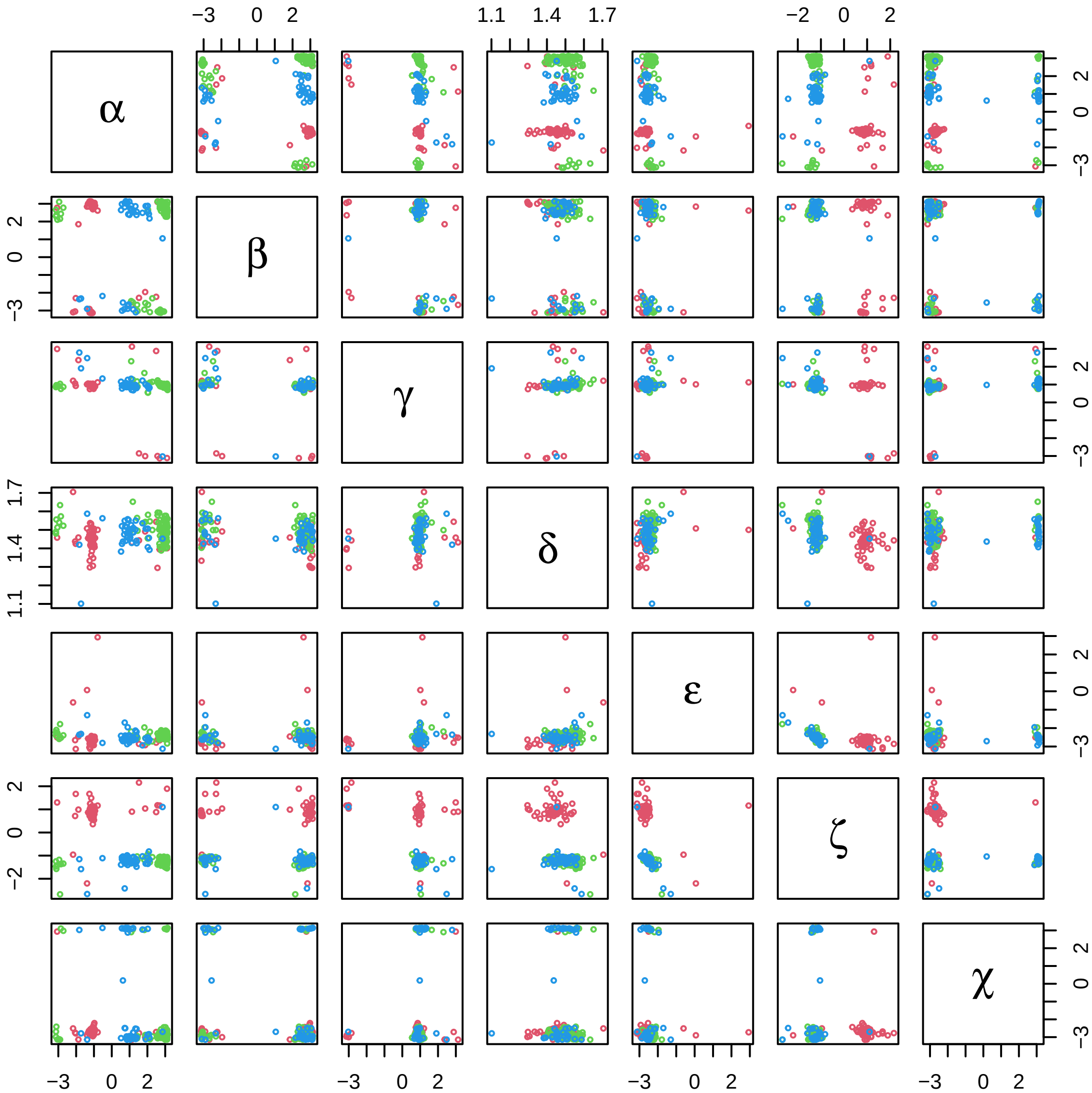
Luis Ángel Rodríguez García

25-05-2022

The objective in this case study is to recover clusters identified in `smallrna$clusters` using only the information on `smallrna$angles`, a 7-dimensional matrix of angles (i.e., data on $(\mathbb{S}^1)^7$)). The clusters have been constructed using the information in `smallrna$torsion`. If a dimension-reduction technique is able to successfully identify clusters, then it will be doing a good job in terms of identifying the underlying structure of the data. Section 5.2 in Zoubouloglou et al. (2021) describes the history of the "Small RNA" dataset and its construction.

Let's begin by importing the data.

```r
# Load package
library(pscsne)
stopifnot(packageVersion("pscsne") >= "0.0.1.900005")

# Load dataset
data("smallrna")

# Visualize original data
pairs(smallrna$angles, col = smallrna$clusters + 1, cex = 0.5,
      labels = c(expression(alpha), expression(beta), expression(gamma),
                 expression(delta), expression(epsilon), expression(zeta),
                 expression(chi)))
```



We can now run psc-SNE. First, we transform the data and obtain the ρ's giving the prescribed perplexity.

```
# Data to Cartesian coordinates
smallrna_X <- sphunif::Theta_to_X(Theta = smallrna$angles)

# Obtain rhos for given perplexity
rho_psc_list <- rho_optim_bst(x = smallrna_X, perp_fixed = 30)
```

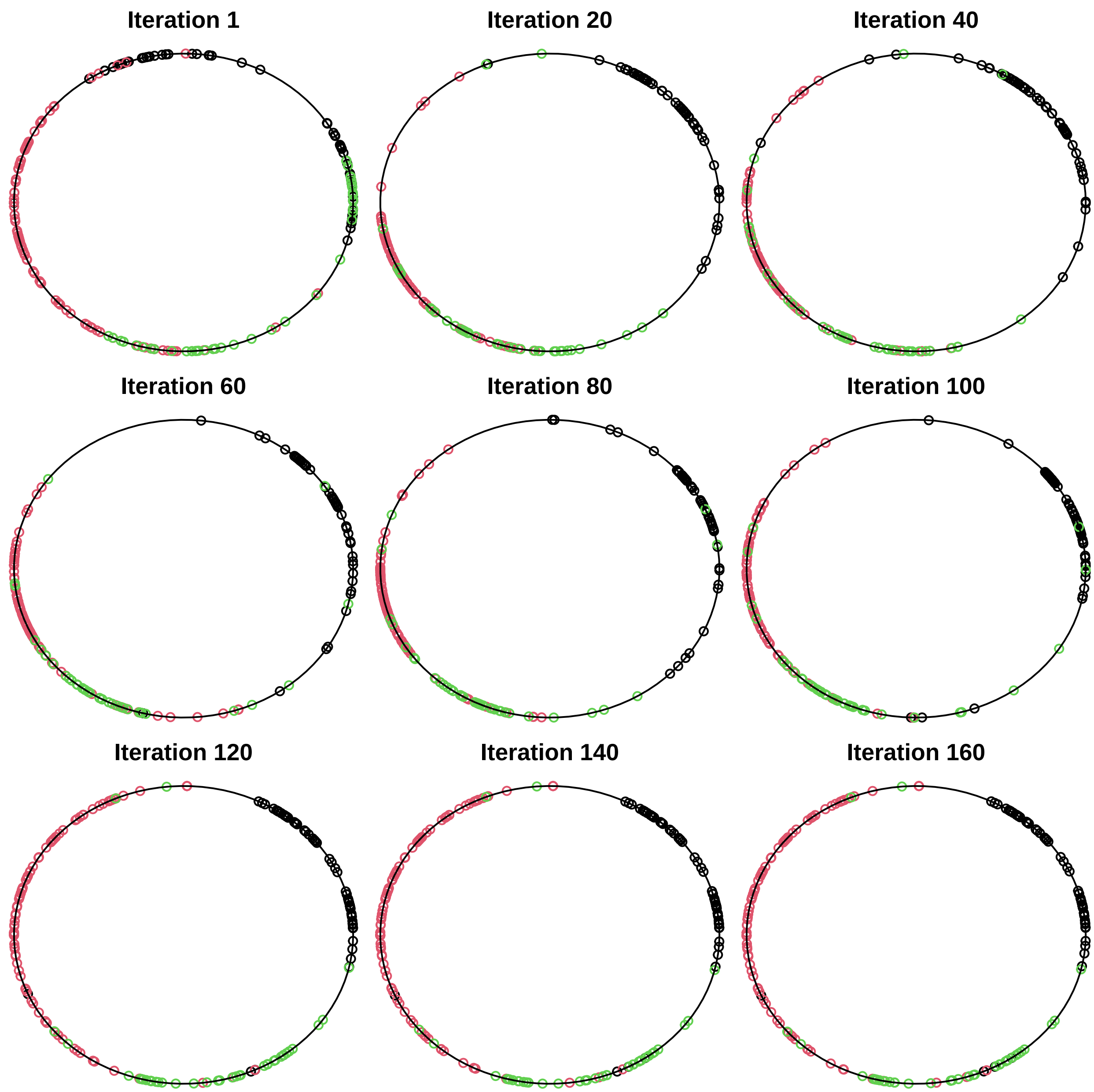
```
## Time difference of 1.127477 secs
```

We run psc-SNE for $d = 1$ with its default η.

```
# Default
fit_1 <- psc_sne(X = smallrna_X, d = 1, rho_psc_list = rho_psc_list,
                 eta = 200, maxit = 1e3, tol = 1e-6, show_prog = 10,
                 colors = smallrna$clusters)
```

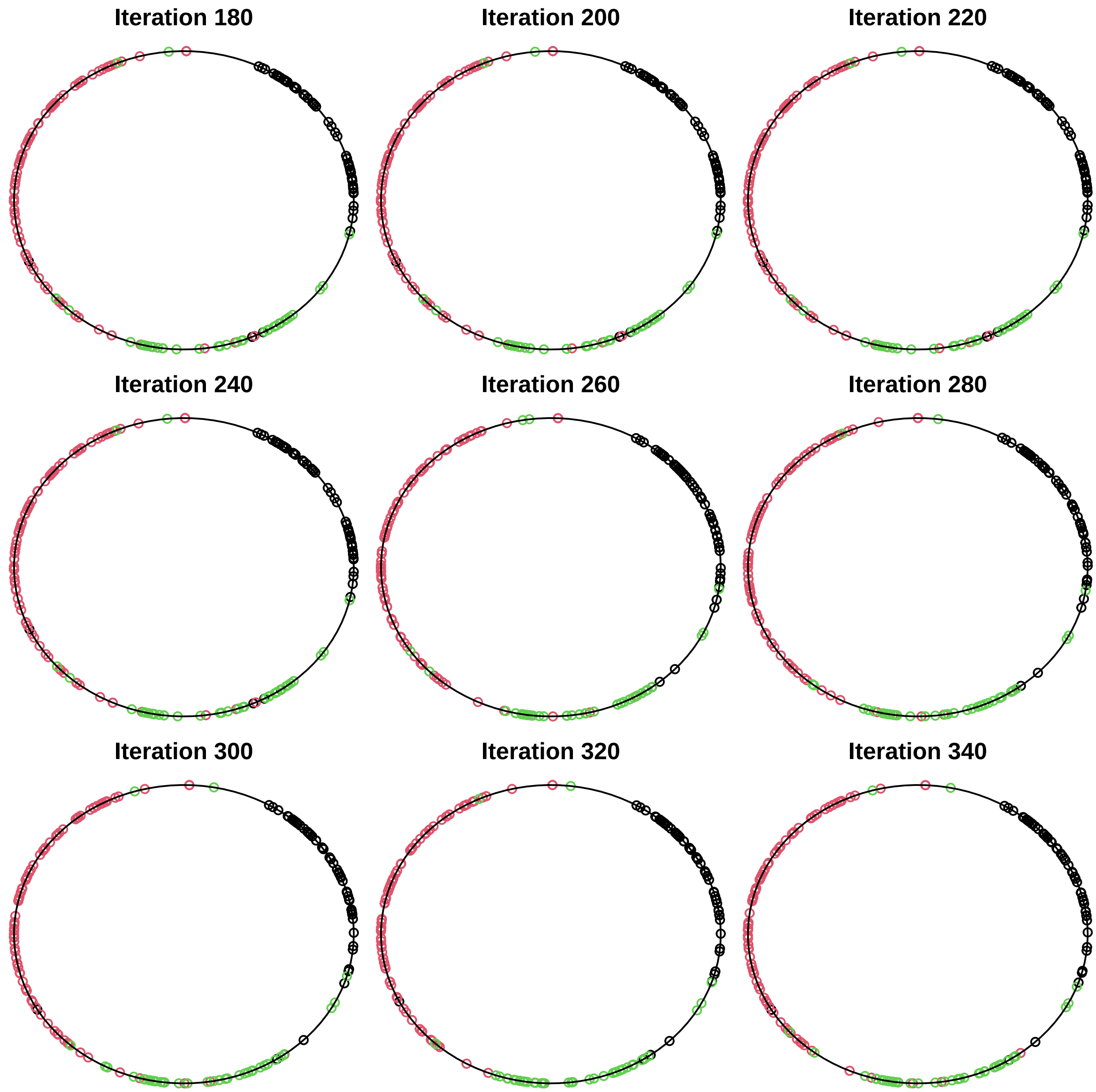
```
## It: 1 (best: 1); obj: 1.20e+01 (best: 1.20e+01); abs: 0.0e+00; rel: 0.0e+00; norm: 3.2e-01; mom: 0.0e+00

## It: 10 (best: 6); obj: 1.05e+01 (best: 1.05e+01); abs: 1.7e+00; rel: 1.4e-01; norm: 2.6e-01; mom: 7.7e+00

## It: 20 (best: 14); obj: 1.05e+01 (best: 1.03e+01); abs: 1.8e+00; rel: 1.4e-01; norm: 2.5e-01; mom: 7.8e+00

## It: 30 (best: 14); obj: 1.05e+01 (best: 1.03e+01); abs: 1.8e+00; rel: 1.5e-01; norm: 2.6e-01; mom: 7.5e+00

## It: 40 (best: 14); obj: 1.07e+01 (best: 1.03e+01); abs: 1.5e+00; rel: 1.2e-01; norm: 2.5e-01; mom: 7.4e+00
```

```
## It: 50 (best: 14); obj: 1.06e+01 (best: 1.03e+01); abs: 1.6e+00; rel: 1.3e-01; norm: 2.5e-01; mom: 7.6e+00
## It: 60 (best: 14); obj: 1.05e+01 (best: 1.03e+01); abs: 1.9e+00; rel: 1.6e-01; norm: 2.5e-01; mom: 7.7e+00
## It: 70 (best: 14); obj: 1.04e+01 (best: 1.03e+01); abs: 2.0e+00; rel: 1.6e-01; norm: 2.4e-01; mom: 7.6e+00
## It: 80 (best: 14); obj: 1.04e+01 (best: 1.03e+01); abs: 2.0e+00; rel: 1.6e-01; norm: 2.5e-01; mom: 7.6e+00
## It: 90 (best: 14); obj: 1.04e+01 (best: 1.03e+01); abs: 2.1e+00; rel: 1.7e-01; norm: 2.4e-01; mom: 7.7e+00
## It: 100 (best: 14); obj: 1.05e+01 (best: 1.03e+01); abs: 1.9e+00; rel: 1.5e-01; norm: 2.6e-01; mom: 7.5e+00
## It: 110 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 2.3e-02; rel: 2.1e-02; norm: 8.8e-02; mom: 4.4e+00
## It: 120 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 2.0e-02; rel: 1.7e-02; norm: 8.9e-02; mom: 4.4e+00
## It: 130 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 2.2e-02; rel: 1.9e-02; norm: 8.8e-02; mom: 4.4e+00
## It: 140 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 1.5e-02; rel: 1.4e-02; norm: 8.9e-02; mom: 4.5e+00
## It: 150 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 2.2e-02; rel: 1.9e-02; norm: 8.8e-02; mom: 4.4e+00
## It: 160 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 1.5e-02; rel: 1.3e-02; norm: 8.9e-02; mom: 4.5e+00
## It: 170 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 2.2e-02; rel: 1.9e-02; norm: 8.8e-02; mom: 4.4e+00
## It: 180 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 1.5e-02; rel: 1.3e-02; norm: 8.9e-02; mom: 4.5e+00
```



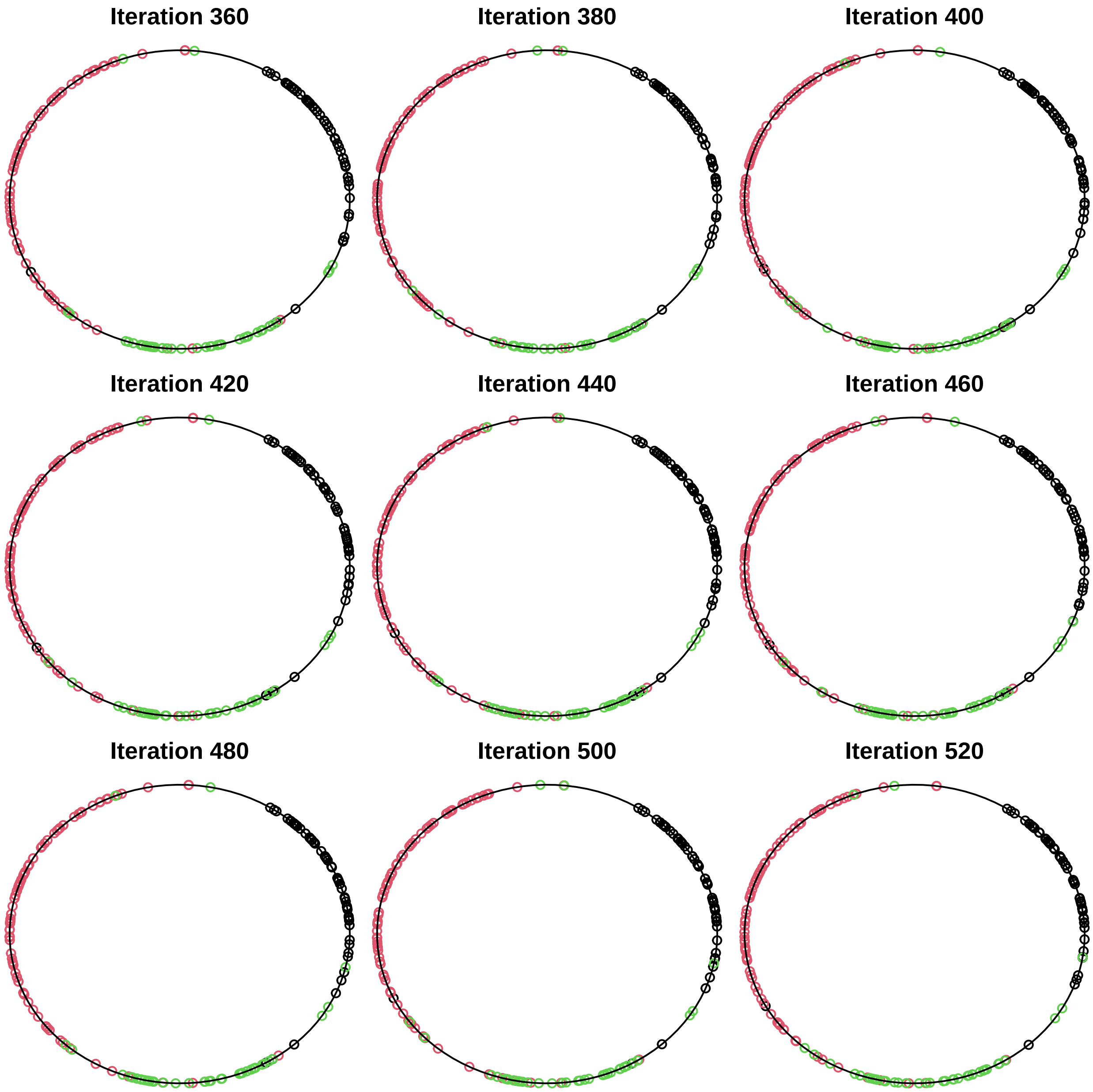

```
## It: 190 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 2.2e-02; rel: 1.9e-02; norm: 8.8e-02; mom: 4.4e+00
## It: 200 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 1.5e-02; rel: 1.3e-02; norm: 8.9e-02; mom: 4.5e+00
## It: 210 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 2.2e-02; rel: 1.9e-02; norm: 8.8e-02; mom: 4.4e+00
## It: 220 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 1.5e-02; rel: 1.3e-02; norm: 8.9e-02; mom: 4.5e+00
## It: 230 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 2.2e-02; rel: 1.9e-02; norm: 8.8e-02; mom: 4.4e+00
## It: 240 (best: 101); obj: 1.15e+00 (best: 1.08e+00); abs: 1.5e-02; rel: 1.3e-02; norm: 8.9e-02; mom: 4.5e+00
## It: 250 (best: 250); obj: 1.02e+00 (best: 1.02e+00); abs: 1.1e-01; rel: 9.4e-02; norm: 8.8e-02; mom: 7.1e+00
## It: 260 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 6.7e-03; rel: 6.1e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 270 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 4.3e-03; rel: 3.9e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 280 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 9.3e-03; rel: 8.5e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 290 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 7.3e-03; rel: 6.7e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 300 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 3.3e-03; rel: 3.0e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 310 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 4.2e-03; rel: 3.8e-03; norm: 8.6e-02; mom: 6.2e+00
```

```
## It: 320 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 1.1e-02; rel: 9.8e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 330 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 5.3e-03; rel: 4.8e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 340 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 2.0e-03; rel: 1.8e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 350 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 2.8e-03; rel: 2.6e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 360 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 2.8e-03; rel: 2.6e-03; norm: 8.6e-02; mom: 6.1e+00
```



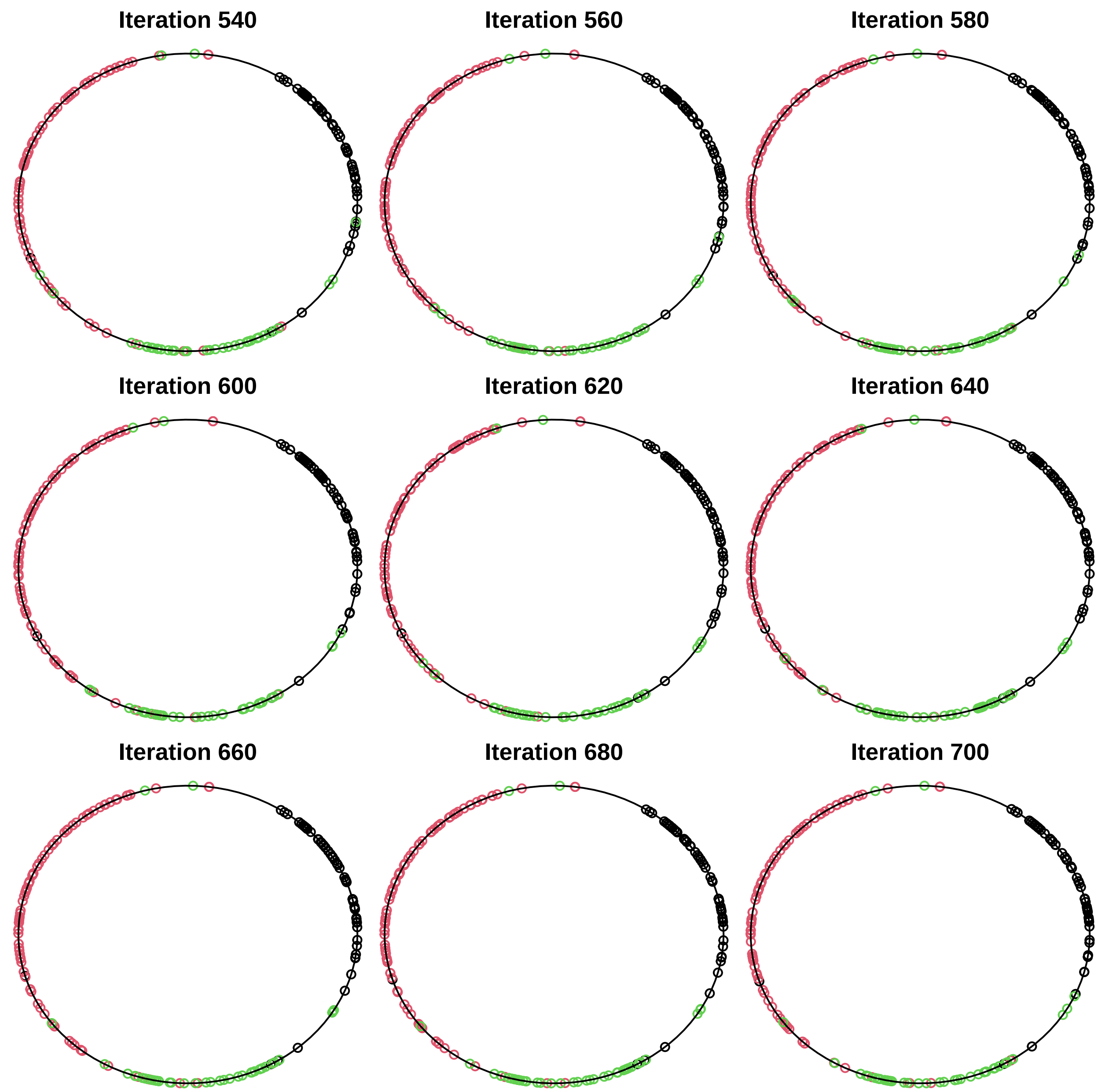

```
## It: 370 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 3.5e-03; rel: 3.2e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 380 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 7.9e-03; rel: 7.3e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 390 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 4.6e-04; rel: 4.3e-04; norm: 8.6e-02; mom: 6.1e+00
## It: 400 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 1.2e-02; rel: 1.1e-02; norm: 8.6e-02; mom: 6.1e+00
## It: 410 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 1.0e-02; rel: 9.2e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 420 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 2.3e-03; rel: 2.1e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 430 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 4.2e-04; rel: 3.8e-04; norm: 8.6e-02; mom: 6.1e+00
## It: 440 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 5.3e-03; rel: 4.9e-03; norm: 8.6e-02; mom: 6.0e+00
```

```
## It: 450 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 2.6e-04; rel: 2.4e-04; norm: 8.6e-02; mom: 6.1e+00
## It: 460 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 1.3e-04; rel: 1.2e-04; norm: 8.6e-02; mom: 6.1e+00
## It: 470 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 2.6e-03; rel: 2.4e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 480 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 1.9e-03; rel: 1.7e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 490 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 2.3e-03; rel: 2.1e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 500 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 6.4e-03; rel: 5.8e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 510 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 9.8e-03; rel: 9.0e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 520 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 9.6e-03; rel: 8.8e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 530 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 2.9e-03; rel: 2.6e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 540 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 7.3e-03; rel: 6.7e-03; norm: 8.7e-02; mom: 6.1e+00
```



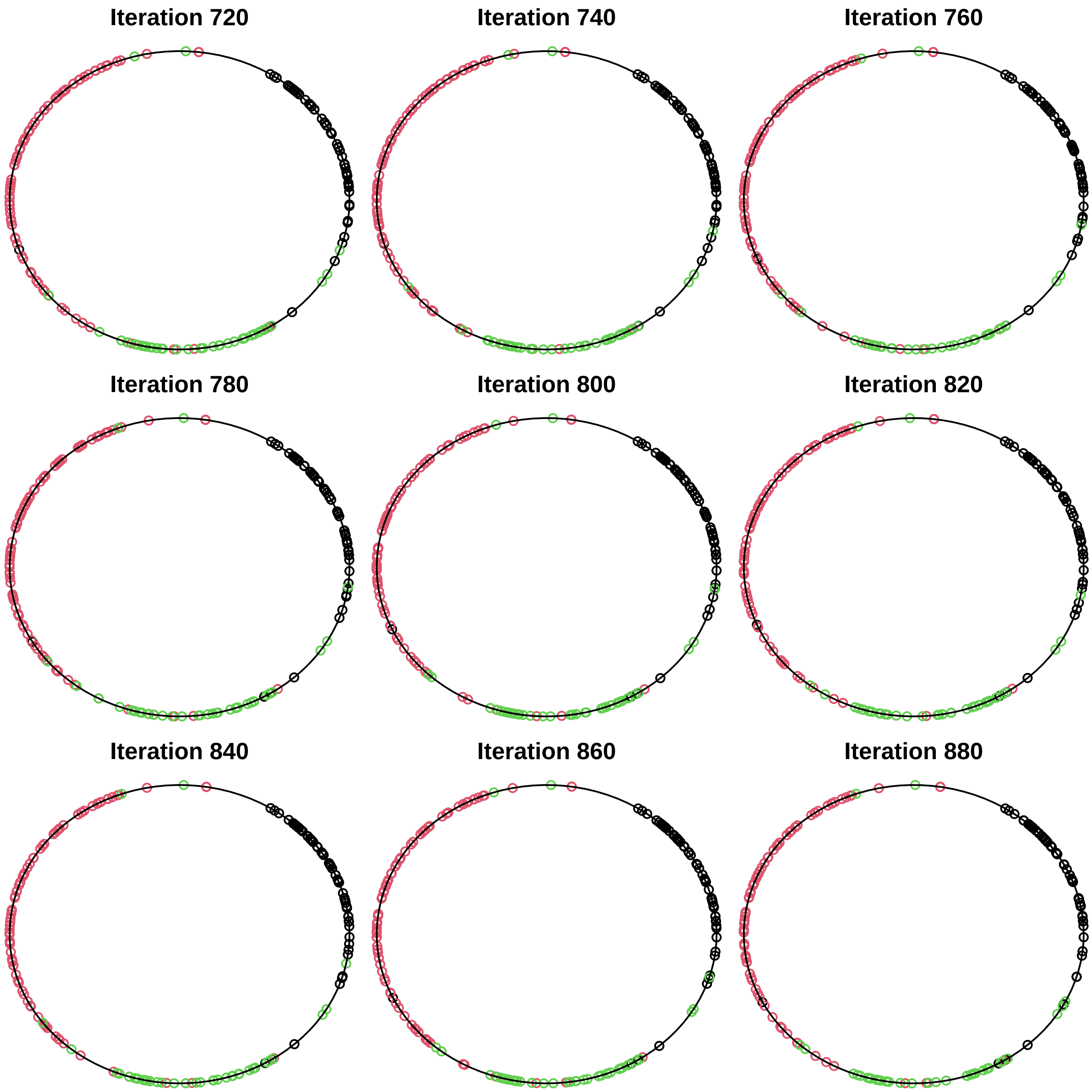

```
## It: 550 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 6.9e-03; rel: 6.3e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 560 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 8.0e-03; rel: 7.3e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 570 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 3.2e-03; rel: 2.9e-03; norm: 8.7e-02; mom: 6.1e+00
```

```
## It: 580 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 5.2e-04; rel: 4.8e-04; norm: 8.6e-02; mom: 6.1e+00
## It: 590 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 1.1e-02; rel: 1.0e-02; norm: 8.6e-02; mom: 6.2e+00
## It: 600 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 3.1e-03; rel: 2.8e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 610 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 4.1e-03; rel: 3.8e-03; norm: 8.6e-02; mom: 6.2e+00
## It: 620 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 5.5e-03; rel: 5.1e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 630 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 9.0e-04; rel: 8.2e-04; norm: 8.6e-02; mom: 6.1e+00
## It: 640 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 4.5e-03; rel: 4.1e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 650 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 8.2e-03; rel: 7.5e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 660 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 4.9e-04; rel: 4.5e-04; norm: 8.6e-02; mom: 6.1e+00
## It: 670 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 7.6e-03; rel: 7.0e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 680 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 1.4e-04; rel: 1.3e-04; norm: 8.6e-02; mom: 6.1e+00
## It: 690 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 6.8e-03; rel: 6.3e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 700 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 4.7e-04; rel: 4.3e-04; norm: 8.6e-02; mom: 6.1e+00
## It: 710 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 3.0e-03; rel: 2.8e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 720 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 2.7e-03; rel: 2.5e-03; norm: 8.6e-02; mom: 6.1e+00
```



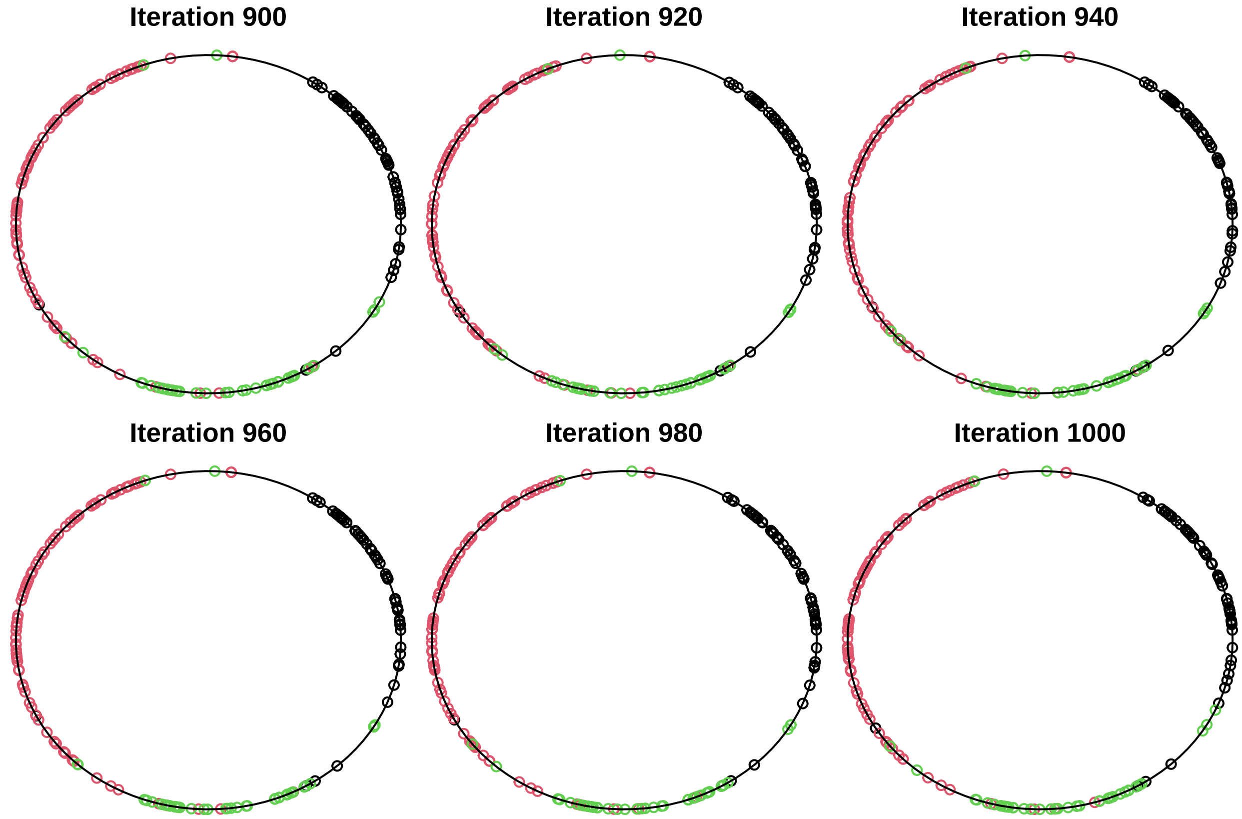

```
## It: 730 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 6.3e-03; rel: 5.8e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 740 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 1.1e-03; rel: 9.9e-04; norm: 8.6e-02; mom: 6.1e+00
## It: 750 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 7.0e-03; rel: 6.5e-03; norm: 8.6e-02; mom: 6.0e+00
## It: 760 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 6.6e-03; rel: 6.0e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 770 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 4.7e-03; rel: 4.3e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 780 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 3.9e-03; rel: 3.6e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 790 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 5.1e-04; rel: 4.7e-04; norm: 8.6e-02; mom: 6.1e+00
## It: 800 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 8.9e-04; rel: 8.2e-04; norm: 8.7e-02; mom: 6.1e+00
## It: 810 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 6.9e-03; rel: 6.3e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 820 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 4.3e-03; rel: 4.0e-03; norm: 8.6e-02; mom: 6.0e+00
## It: 830 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 1.1e-02; rel: 9.7e-03; norm: 8.7e-02; mom: 6.2e+00
## It: 840 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 1.3e-03; rel: 1.2e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 850 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 2.2e-03; rel: 2.1e-03; norm: 8.6e-02; mom: 6.1e+00
```

```
## It: 860 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 2.3e-03; rel: 2.1e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 870 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 7.6e-03; rel: 7.0e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 880 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 2.6e-03; rel: 2.4e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 890 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 7.7e-03; rel: 7.1e-03; norm: 8.7e-02; mom: 6.1e+00
## It: 900 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 6.9e-03; rel: 6.4e-03; norm: 8.6e-02; mom: 6.1e+00
```




```
## It: 910 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 5.8e-03; rel: 5.3e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 920 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 6.1e-04; rel: 5.6e-04; norm: 8.6e-02; mom: 6.1e+00
## It: 930 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 4.5e-03; rel: 4.2e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 940 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 2.1e-03; rel: 1.9e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 950 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 5.4e-03; rel: 4.9e-03; norm: 8.6e-02; mom: 6.1e+00
## It: 960 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 1.3e-02; rel: 1.2e-02; norm: 8.6e-02; mom: 6.1e+00
## It: 970 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 3.5e-03; rel: 3.2e-03; norm: 8.6e-02; mom: 6.0e+00
## It: 980 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 9.3e-03; rel: 8.5e-03; norm: 8.6e-02; mom: 6.1e+00
```

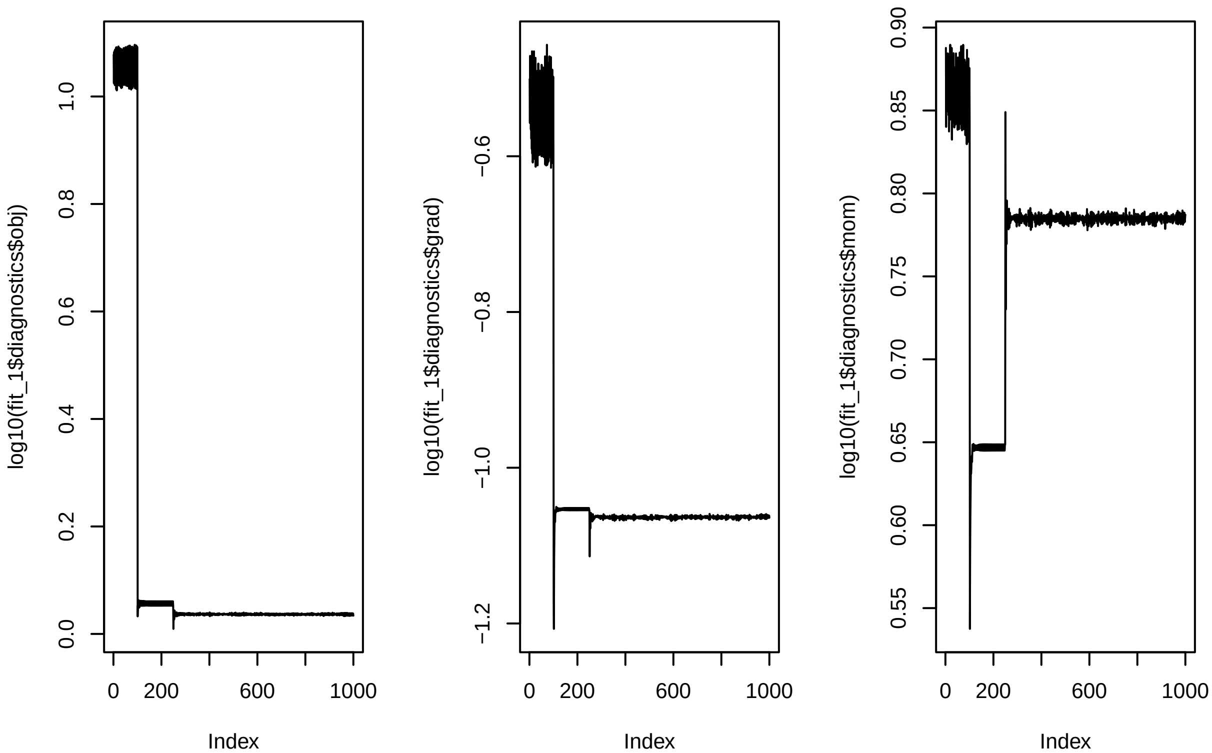
```
## It: 990 (best: 250); obj: 1.09e+00 (best: 1.02e+00); abs: 7.6e-04; rel: 7.0e-04; norm: 8.6e-02; mom: 6.0e+00
## It: 1000 (best: 250); obj: 1.08e+00 (best: 1.02e+00); abs: 9.7e-03; rel: 8.9e-03; norm: 8.6e-02; mom: 6.1e+00
## **NO** CONVERGENCE. Decrease eta? Change init? Increase maxit?
```




```
# Does not converge
fit_1$convergence
```

```
## [1] FALSE
```

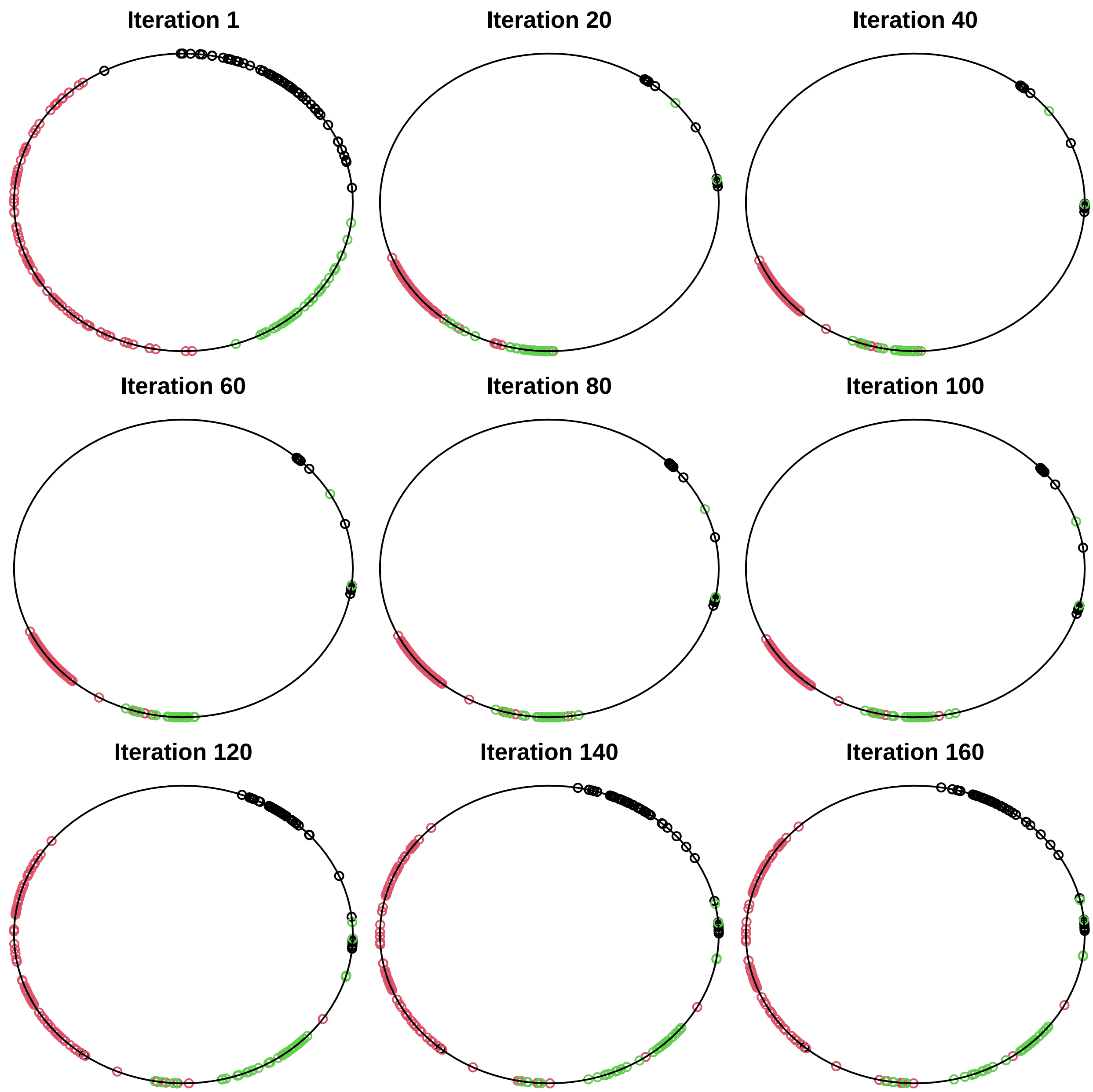
```
par(mfrow = c(1, 3))
plot(log10(fit_1$diagnostics$obj), type = "l")
plot(log10(fit_1$diagnostics$grad), type = "l")
plot(log10(fit_1$diagnostics$mom), type = "l")
```



The employed η seems to be too large. Let's reduce it.

```
# Lower eta
fit_2 <- psc_sne(X = smallrna_X, d = 1, rho_psc_list = rho_psc_list,
                 eta = 10, maxit = 1e3, tol = 1e-6, show_prog = 10,
                 colors = smallrna$clusters)
```

```
## It: 1 (best: 1); obj: 1.03e+01 (best: 1.03e+01); abs: 0.0e+00; rel: 0.0e+00; norm: 3.2e-01; mom: 0.0e+00

## It: 10 (best: 10); obj: 9.62e+00 (best: 9.62e+00); abs: 1.1e-02; rel: 1.2e-03; norm: 6.0e-02; mom: 4.8e-01

## It: 20 (best: 13); obj: 9.68e+00 (best: 9.57e+00); abs: 1.1e-03; rel: 1.1e-04; norm: 2.0e-02; mom: 2.0e-01

## It: 30 (best: 13); obj: 9.64e+00 (best: 9.57e+00); abs: 1.9e-03; rel: 1.9e-04; norm: 2.3e-03; mom: 2.5e-02

## It: 40 (best: 13); obj: 9.63e+00 (best: 9.57e+00); abs: 2.2e-05; rel: 2.3e-06; norm: 1.9e-03; mom: 1.9e-02

## It: 50 (best: 13); obj: 9.63e+00 (best: 9.57e+00); abs: 1.2e-04; rel: 1.2e-05; norm: 1.8e-03; mom: 1.8e-02

## It: 60 (best: 13); obj: 9.64e+00 (best: 9.57e+00); abs: 1.5e-04; rel: 1.5e-05; norm: 1.6e-03; mom: 1.6e-02

## It: 70 (best: 13); obj: 9.64e+00 (best: 9.57e+00); abs: 1.8e-04; rel: 1.9e-05; norm: 1.5e-03; mom: 1.5e-02

## It: 80 (best: 13); obj: 9.64e+00 (best: 9.57e+00); abs: 2.4e-04; rel: 2.5e-05; norm: 1.3e-03; mom: 1.3e-02

## It: 90 (best: 13); obj: 9.64e+00 (best: 9.57e+00); abs: 3.1e-04; rel: 3.2e-05; norm: 1.1e-03; mom: 1.2e-02

## It: 100 (best: 13); obj: 9.65e+00 (best: 9.57e+00); abs: 3.5e-04; rel: 3.6e-05; norm: 9.7e-04; mom: 1.0e-02

## It: 110 (best: 110); obj: 8.89e-01 (best: 8.89e-01); abs: 5.4e-03; rel: 6.1e-03; norm: 1.6e-02; mom: 2.0e-01

## It: 120 (best: 120); obj: 8.66e-01 (best: 8.66e-01); abs: 1.1e-03; rel: 1.3e-03; norm: 7.3e-03; mom: 8.4e-02

## It: 130 (best: 130); obj: 8.61e-01 (best: 8.61e-01); abs: 2.0e-04; rel: 2.4e-04; norm: 3.1e-03; mom: 3.8e-02

## It: 140 (best: 140); obj: 8.60e-01 (best: 8.60e-01); abs: 2.9e-05; rel: 3.4e-05; norm: 1.2e-03; mom: 1.4e-02

## It: 150 (best: 150); obj: 8.60e-01 (best: 8.60e-01); abs: 4.3e-06; rel: 5.0e-06; norm: 4.5e-04; mom: 5.5e-03

## It: 160 (best: 160); obj: 8.60e-01 (best: 8.60e-01); abs: 6.5e-07; rel: 7.6e-07; norm: 1.8e-04; mom: 2.1e-03

## It: 170 (best: 170); obj: 8.60e-01 (best: 8.60e-01); abs: 1.0e-07; rel: 1.2e-07; norm: 6.9e-05; mom: 8.4e-04

## It: 180 (best: 180); obj: 8.60e-01 (best: 8.60e-01); abs: 1.6e-08; rel: 1.8e-08; norm: 2.7e-05; mom: 3.3e-04
```

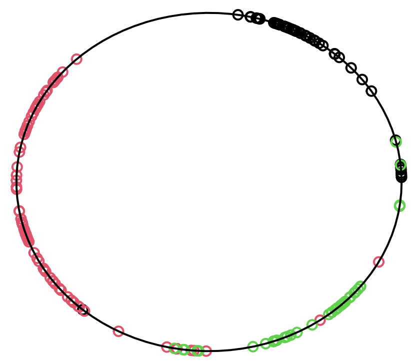



```
## It: 190 (best: 190); obj: 8.60e-01 (best: 8.60e-01); abs: 2.6e-09; rel: 3.0e-09; norm: 1.1e-05; mom: 1.3e-04
## It: 200 (best: 200); obj: 8.60e-01 (best: 8.60e-01); abs: 4.9e-10; rel: 5.7e-10; norm: 4.8e-06; mom: 5.7e-05
## It: 210 (best: 210); obj: 8.60e-01 (best: 8.60e-01); abs: 1.2e-10; rel: 1.3e-10; norm: 2.4e-06; mom: 2.7e-05
## It: 220 (best: 220); obj: 8.60e-01 (best: 8.60e-01); abs: 3.7e-11; rel: 4.3e-11; norm: 1.3e-06; mom: 1.5e-05
## It: 227 (best: 227); obj: 8.60e-01 (best: 8.60e-01); abs: 2.0e-11; rel: 2.3e-11; norm: 9.8e-07; mom: 1.1e-05
## CONVERGENCE!
```

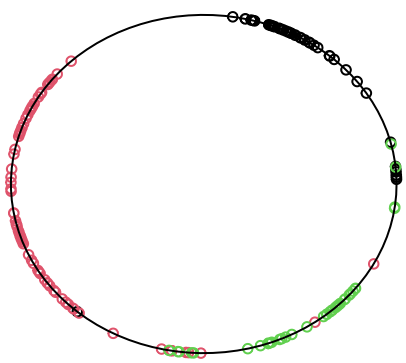
Iteration 180

Iteration 200

Iteration 220



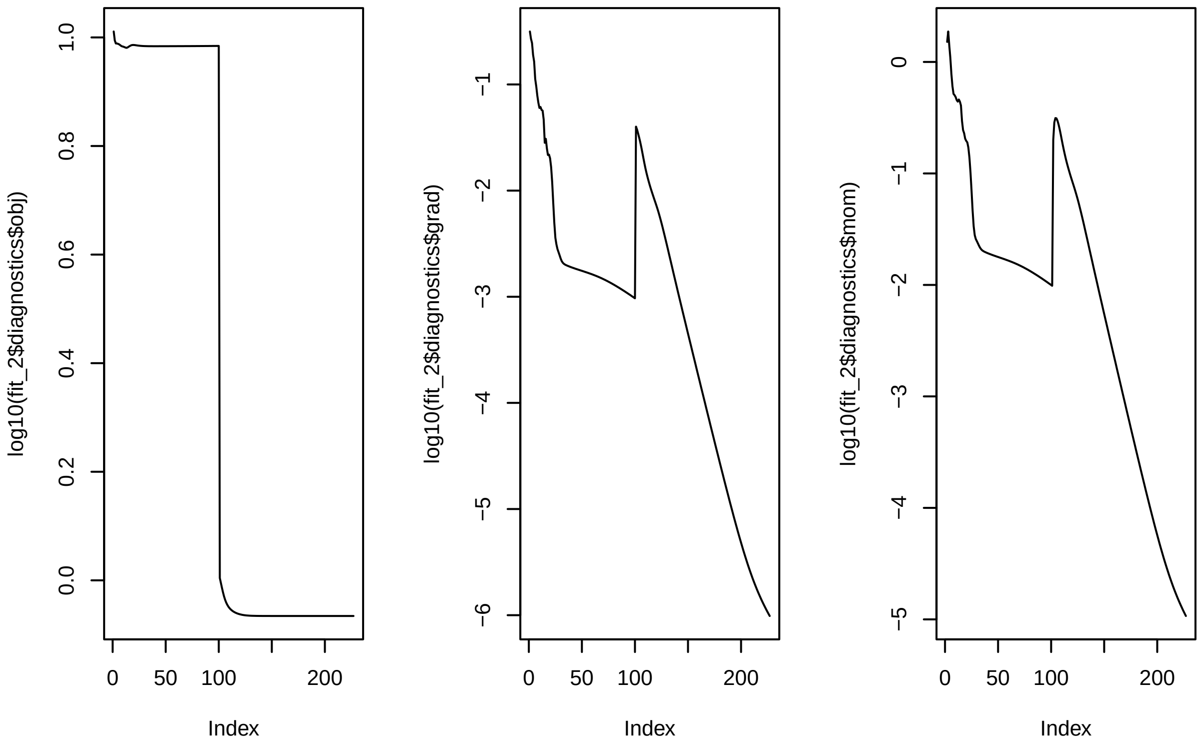
Iteration 227



```
# Converges
fit_2$convergence
```

```
## [1] TRUE
```

```
par(mfrow = c(1, 3))
plot(log10(fit_2$diagnostics$obj), type = "l")
plot(log10(fit_2$diagnostics$grad), type = "l")
plot(log10(fit_2$diagnostics$mom), type = "l")
```



Convergence is attained in the second run, yet it is weird that the objective function takes exactly the zero value.

Let's see the recovery of the clusters.

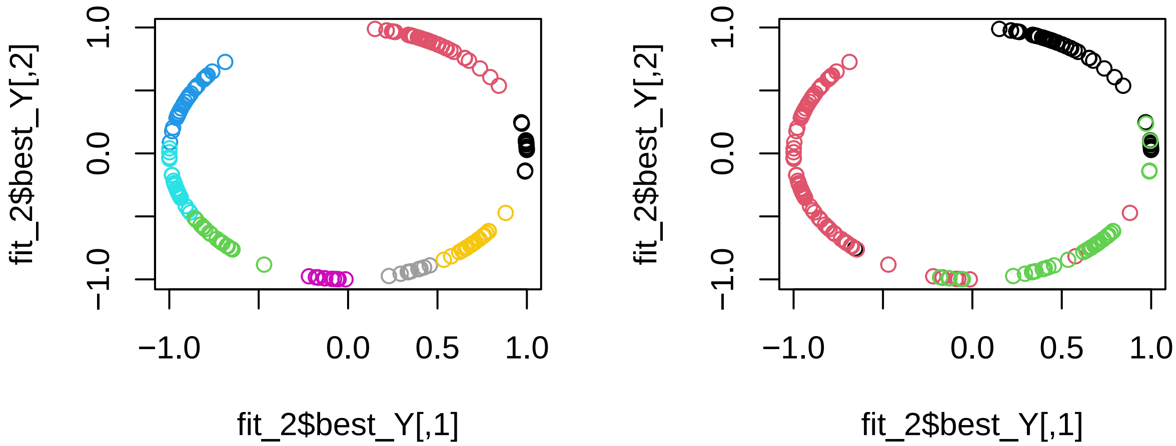
```
# Kernel mean shift clustering
n <- nrow(fit_2$best_Y)
d <- ncol(fit_2$best_Y) - 1
fit_mix <- DirStats::bic_vmf_mix(data = fit_2$best_Y, kappa_max = 1e3)
h <- DirStats::bw_dir_emi(data = fit_2$best_Y, fit_mix = fit_mix)$h_opt *
      n^(1 / (d + 4)) * n^(-1 / (d + 6))
kms <- kms_dir(data = fit_2$best_Y, h = h)
```

```
##   |                                                                                  |
```

```
# Detects 8 clusters by splitting the 3 original clusters
length(unique(kms$cluster))
```

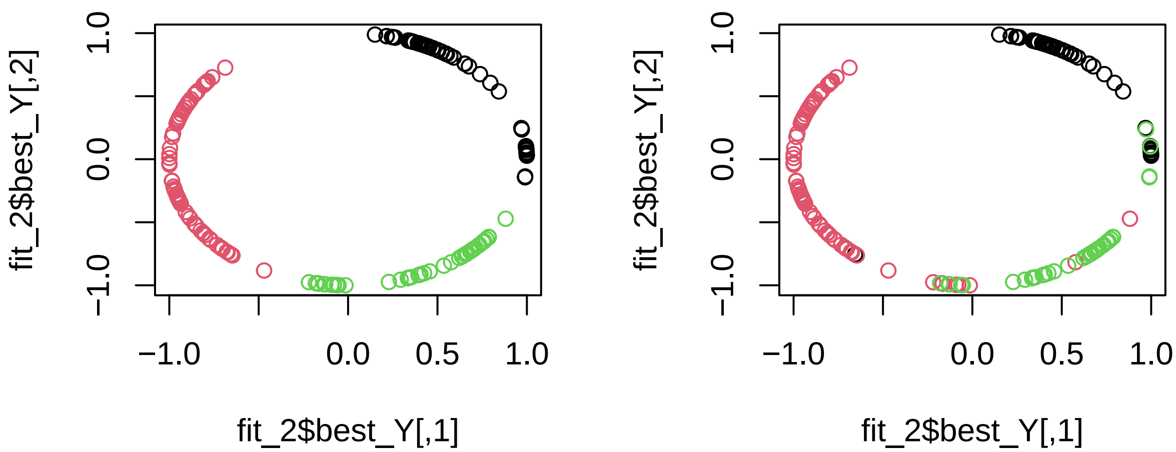
```
## [1] 8
```

```
# Fully-automatically recovered clusters vs. real clusters
par(mfrow = c(1, 2))
plot(fit_2$best_Y, col = kms$cluster)
plot(fit_2$best_Y, col = smallrna$clusters)
```



The original clusters are not fully recovered, in the sense that more clusters are obtained. However, the three-cluster structure is present, as the new clusters appear dividing the three main ones. This can be checked by cutting the hierarchical clustering tree behind kernel mean shift clustering exactly at three groups. Or, in other words, by merging the 8 groups into 3.

```
# Recovered clusters with three clusters vs. real clusters
par(mfrow = c(1, 2))
labels <- cutree(kms$tree, k = 3)
plot(fit_2$best_Y, col = labels)
plot(fit_2$best_Y, col = smallrna$clusters)
```




```
# Correct classification rate: 90%
mean(labels == smallrna$clusters)
```

```
## [1] 0.9
```

```
# 19 incorrectly classified observations
sum(labels != smallrna$clusters)
```

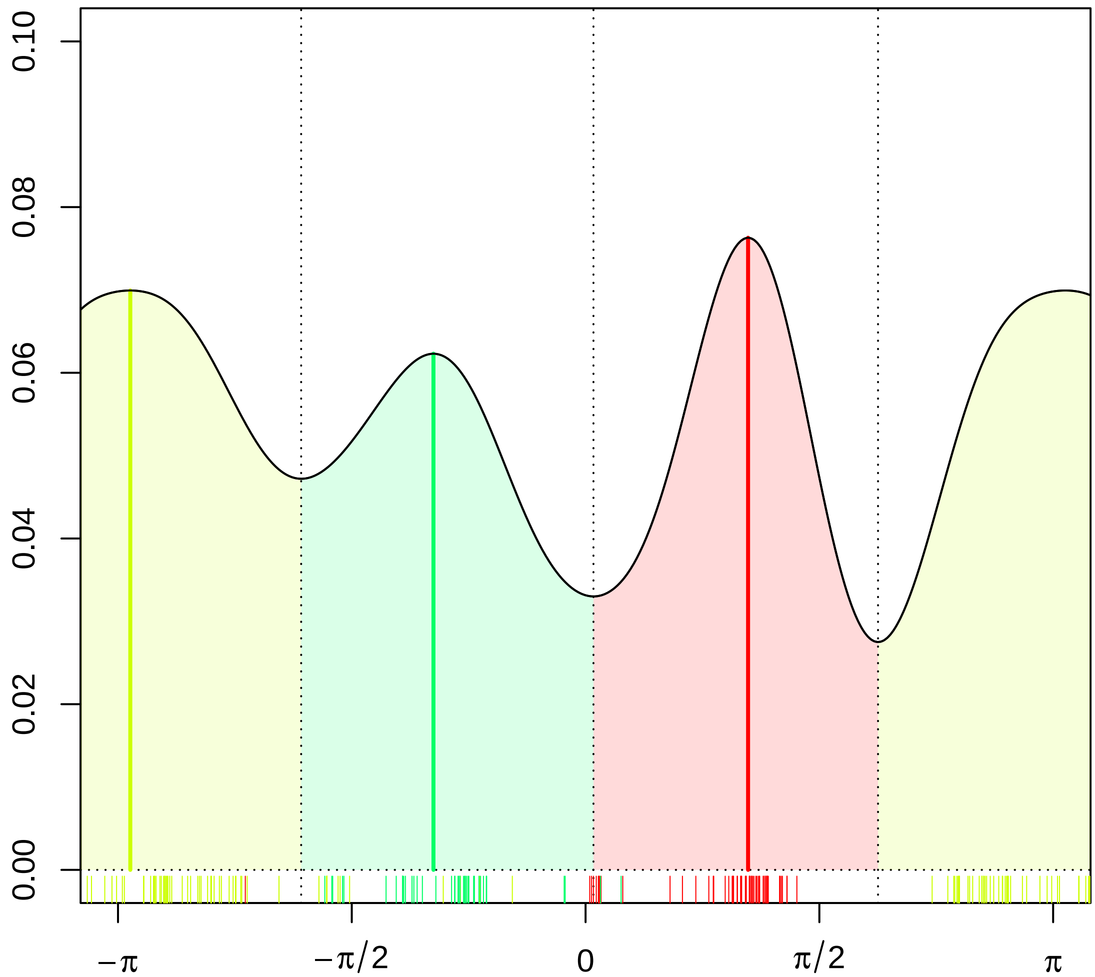
```
## [1] 19
```

The classification accuracy is on-par with Zoubouloglou et al. (2021), which misclassifies 16 points and has a classification rate of 0.916.

There is another way to cluster the scores provided by the psc-SNE. These values, x and y, can be mapped to a parameter θ encoded in radians, and later they can be used with the kernel mean shift clustering for linear data to obtain these groups and plot them in a density graph.

```r
# Convert to radians the scores
x <- DirStats::to_rad(fit_2$best_Y)

# Obtain optimal plug-in bandwidth for density derivative estimation,
# disregarding periodicity
h <- ks::hpi(x, deriv.order = 1)

pscsne::plot_kde(x, smallrna$clusters, h)
```

```
##   |                                                                              |
```



The results are similar to those one obtained above: classification rate of 89.47% and 20 observations wrongly classified.