realnetdata

Etienne Côme

10 septembre 2019

```
library(greed)
library(future)
plan(multisession)
library(ggplot2)
```

```
## Registered S3 methods overwritten by 'ggplot2':
##   method         from
##   [.quosures     rlang
##   c.quosures     rlang
##   print.quosures rlang
```

```
library(ggpubr)
```

```
## Loading required package: magrittr
```

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(Matrix)

set.seed(1234)
```

```
data("Blogs")
fit_blogs = greed(Blogs$X,model=new("dcsbm"),alg=new("hybrid",pop_size=40))
data("Books")
fit_books = greed(Books$X,model=new("dcsbm"),alg=new("hybrid",pop_size=40))
data("Football")
fit_foot = greed(Football$X,model=new("dcsbm"),alg=new("hybrid",pop_size=40))
data("Jazz")
fit_jazz = greed(Jazz,model=new("dcsbm"),alg=new("hybrid",pop_size=40))
```

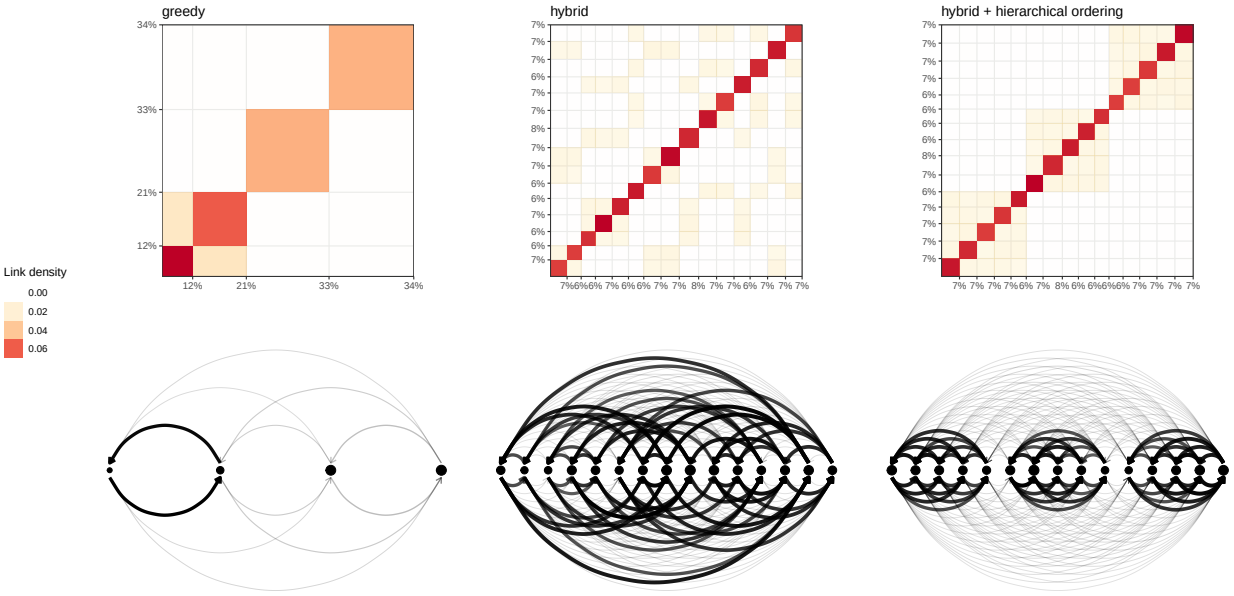
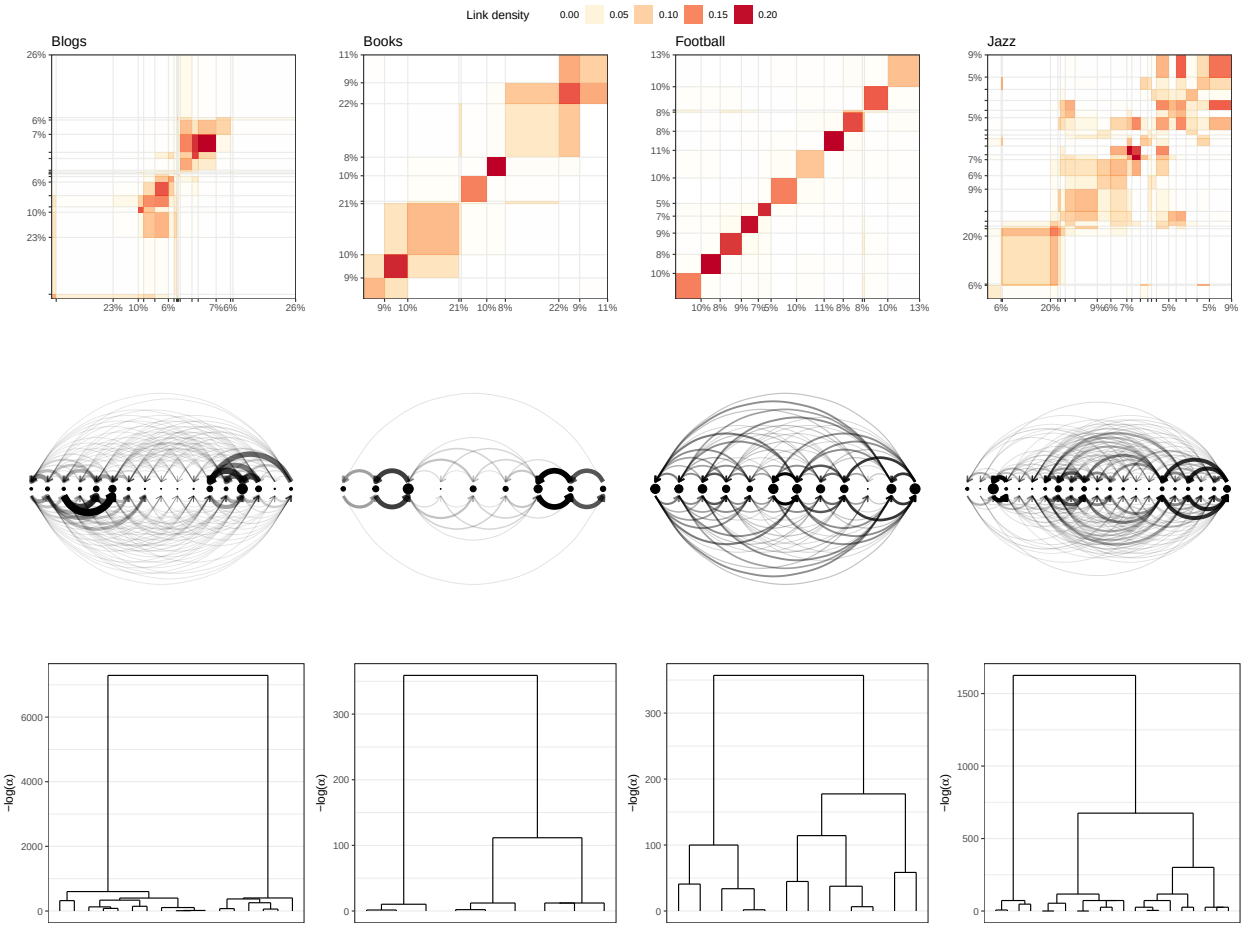
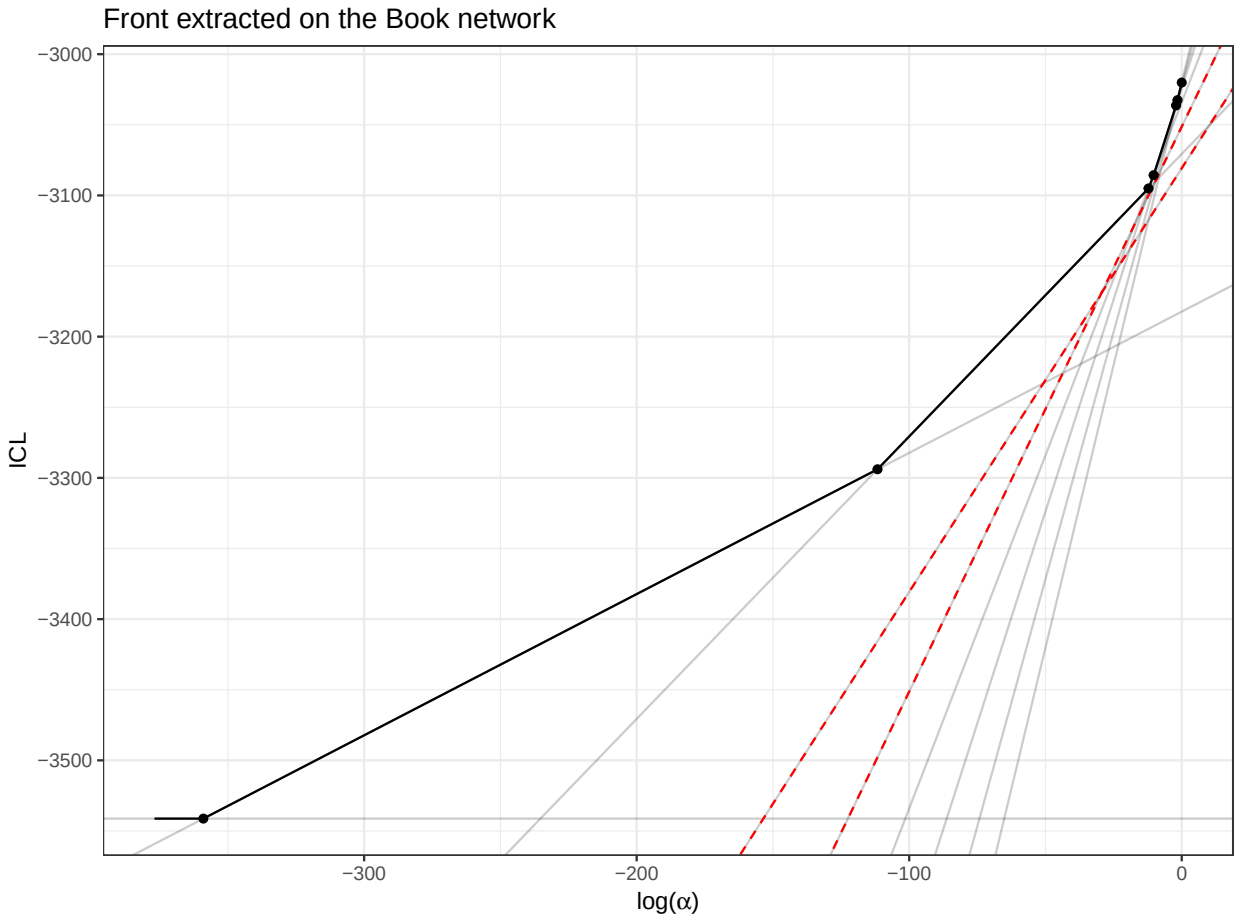
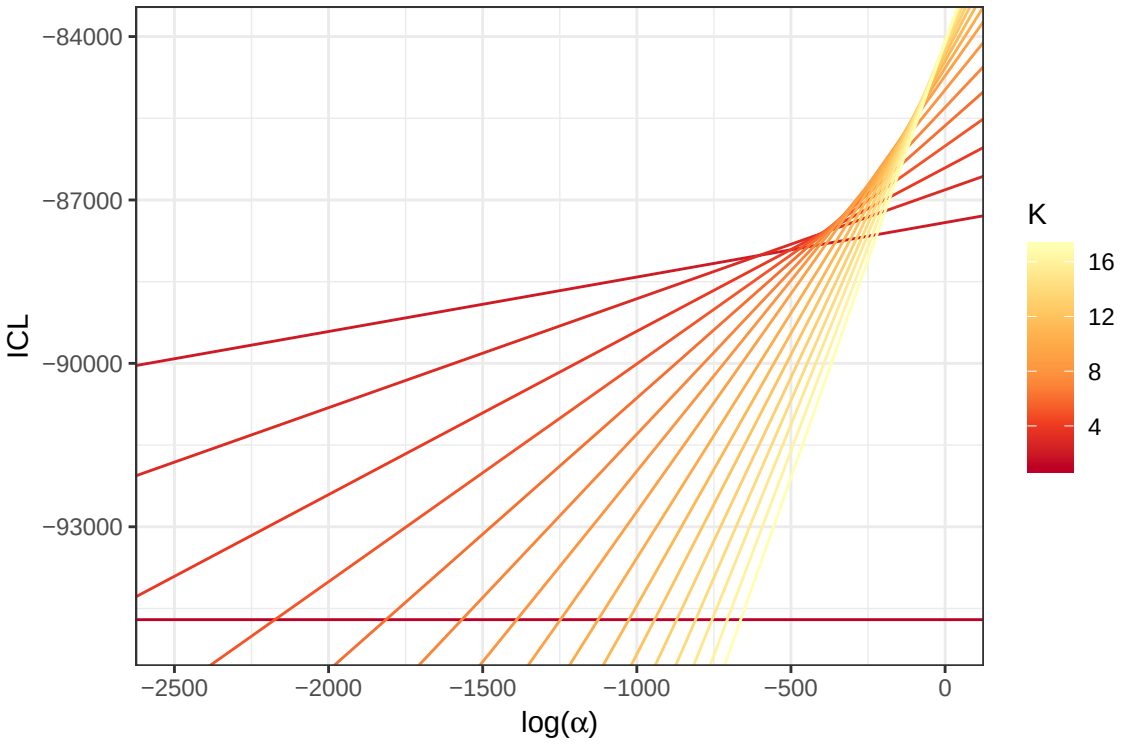


Figure 1: Motivating example of the proposed algorithm. Block matrix representation of the solutions (upper row) and cluster node link diagram (bottom row).

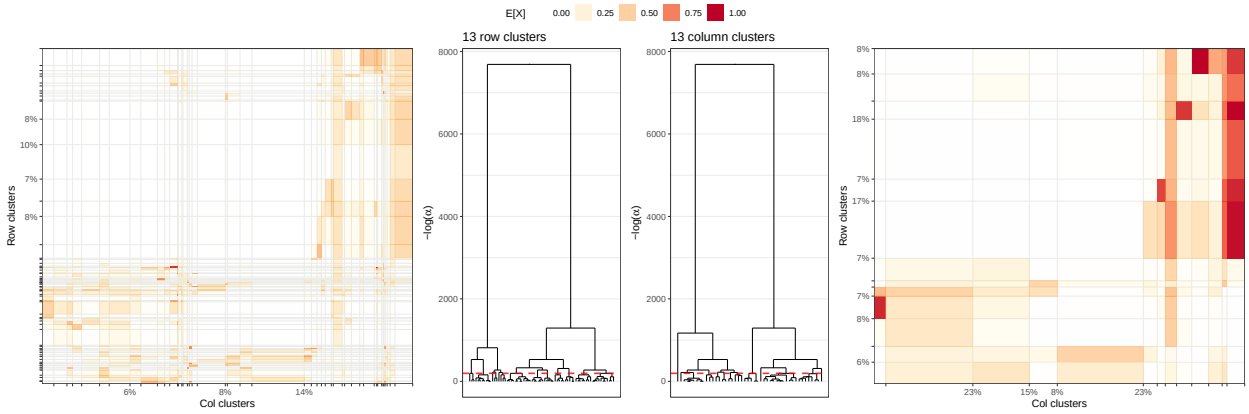


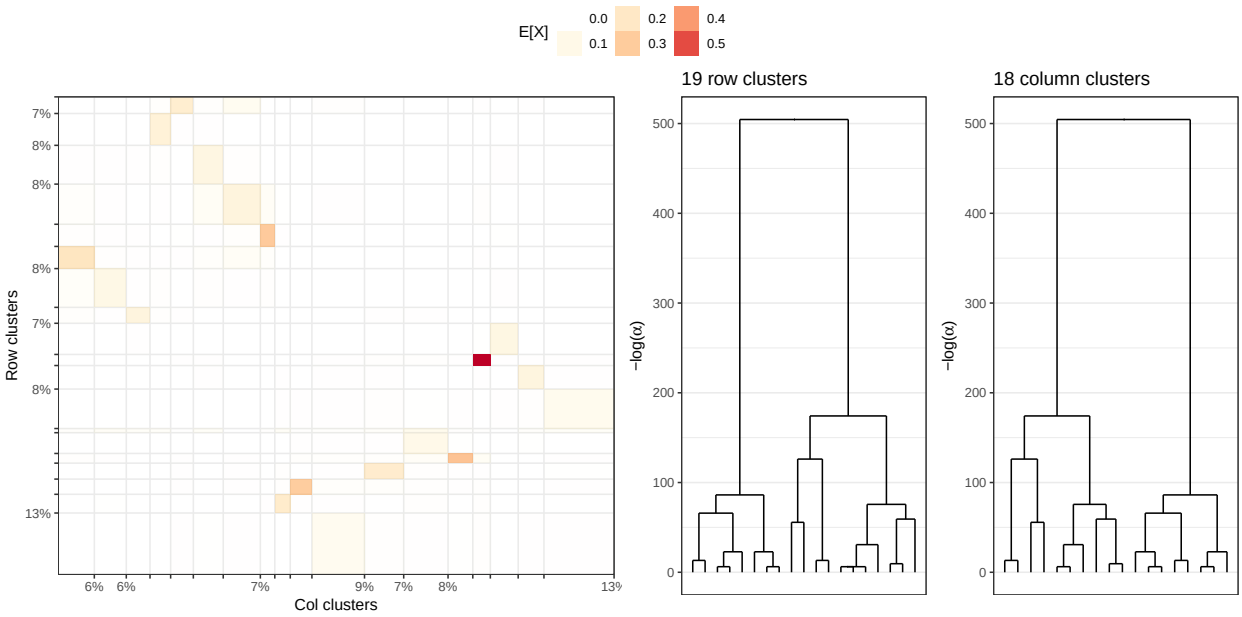
```
## Joining, by = c("icl", "K")
```






```
data("Xvlegislature")
fit_pol = greed(Xvlegislature$X,alg=new("hybrid",pop_size=40),K=100)
data("Jazz_full")
Xc=Jazz_full$X[rowSums(Jazz_full$X)>2,]
Xc=Xc[,colSums(Xc)>2]
fit_cojazz = greed(Xc,alg=new("hybrid",pop_size=40),K=60)
```




E[X]
0.0
0.1
0.2
0.3
0.4
0.5
19 row clusters
18 column clusters
Row clusters
Col clusters
−log(α)
7%
8%
8%
8%
7%
8%
13%
6%
6%
7%
9%
7%
8%
13%
500
400
300
200
100
0