

AnchorFCI: Enhancing Causal Discovery with Reliable Anchors

```
library(anchorFCI)
library(pcalg)
library(FCI.Utils)

#####
# Generating a random MAG with n_vars + n_prec_vars nodes, #
# with n_prec_vars \prec n_vars #
#####

n_vars = 5
n_prec_vars = 3
dir_edges_prob = 0.2
bidir_edges_prob = 0.3

#cur_seed <- sample(1:Machine$integer.max, 1)
cur_seed <- 421953804
set.seed(cur_seed)

true.amat.mag <- getRandomMAG(n_nodes = n_vars, dir_edges_prob = dir_edges_prob,
                             bidir_edges_prob = bidir_edges_prob)

for (prec_v in 1:n_prec_vars) {
  true.amat.mag <- cbind(true.amat.mag, rep(0, dim(true.amat.mag)[1]))
  true.amat.mag <- rbind(true.amat.mag, rep(0, dim(true.amat.mag)[2]))
  n_anchored_vars <- min(sample(1:(ncol(true.amat.mag)-1), 1), 2)
  anchored_vars <- sample(1:(ncol(true.amat.mag)-1), n_anchored_vars)
  true.amat.mag[anchored_vars, n_vars + prec_v] <- 3
  true.amat.mag[n_vars + prec_v, anchored_vars] <- 2
}
colnames(true.amat.mag) <- row.names(true.amat.mag) <-
  c(colnames(true.amat.mag)[1:n_vars], paste0("G", 1:n_prec_vars))

renderAG(true.amat.mag, add_index = FALSE, fileid="trueMAG")

true.amat.pag <- getTruePAG(pcalg::pcalg2dagitty(
  true.amat.mag, colnames(true.amat.mag), type="mag"))@amat

renderAG(true.amat.pag, add_index = FALSE, fileid="truePAG")

#####
# Generating a random dataset following the true PAG #
#####

f.args <- list()
var_names <- c("A", "B", "C", "D", "E")
prec_var_names <- c("G1", "G2", "G3")
```

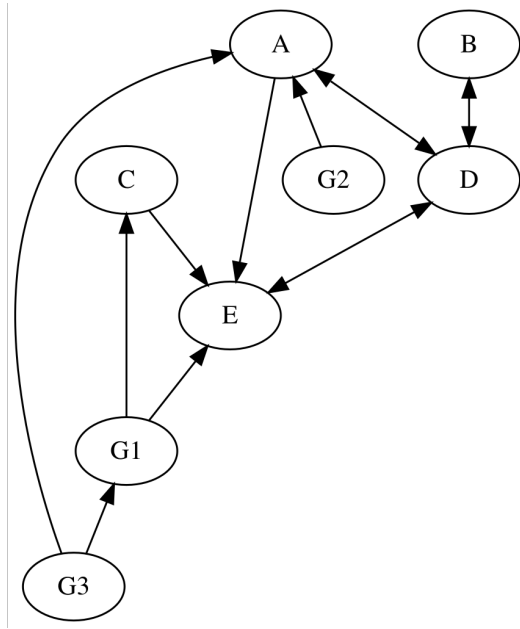


Figure 1: True MAG.

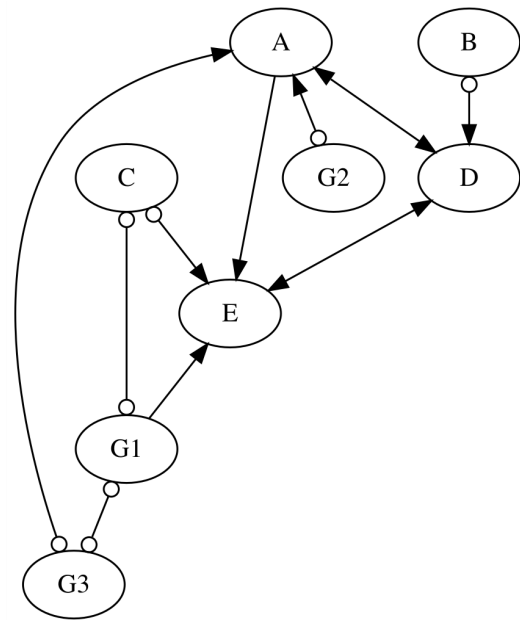


Figure 2: True PAG.

```

for (lab in var_names) {
  f.args[[lab]] <- list(levels = 1) # simulating continuous variables (mimics phenotypes)
}
for (lab in prec_var_names) {
  f.args[[lab]] <- list(levels = 3) # simulating anchors with three levels (mimics SNPs)
}

#cur_seed <- sample(1:.Machine$integer.max, 1)
cur_seed <- 1070843257
set.seed(cur_seed)

N = 1000
dat_out <- generateDatasetFromPAG(apag = true.amat.pag, N=N, type="mixed", f.args=f.args)
#> Warning: replacing previous import 'gam::gam' by 'mgcv::gam' when loading
#> 'simMixedDAG'
#> Warning: replacing previous import 'gam::gam.control' by 'mgcv::gam.control'
#> when loading 'simMixedDAG'
#> Warning: replacing previous import 'gam::gam.fit' by 'mgcv::gam.fit' when
#> loading 'simMixedDAG'
#> Warning: replacing previous import 'gam::s' by 'mgcv::s' when loading
#> 'simMixedDAG'
dataset <- dat_out$dat

#####
# Causal Discovery using AnchorFCI #
#####

output_folder <- "./tmp/"
if (!file.exists(output_folder)) {
  dir.create(output_folder, recursive = TRUE)
}

all_vars <- c(var_names, prec_var_names)
sel_dat <- dataset[,all_vars]
indepTest <- mixedCITest
suffStat <- getMixedCISuffStat(sel_dat, all_vars)

suffStat$citestResults <- getAllCITestResults( sel_dat, indepTest, suffStat,
                                              m.max=2,
                                              citestResults_folder=output_folder)

#> Loading required package: foreach
#> Loading required package: future

out_anchorFCI <- anchorFCI(suffStat, indepTest,
                           var_names=var_names, var_labels=var_names,
                           prec_var_names = prec_var_names,
                           prec_var_labels = prec_var_names,
                           select_anchors = TRUE,
                           m.max=2, alpha = 0.05, fileid="fit",
                           conservative = FALSE, maj.rule = TRUE,
                           renderAll = FALSE, width=800, height = 450,
                           output_folder=output_folder)

```

```
#out_anchorFCI$amb_triplets_df
```

```
renderAG(out_anchorFCI$fit_anchor_fci@amat, add_index = FALSE, fileid="anchorFCI_outputPAG")
```

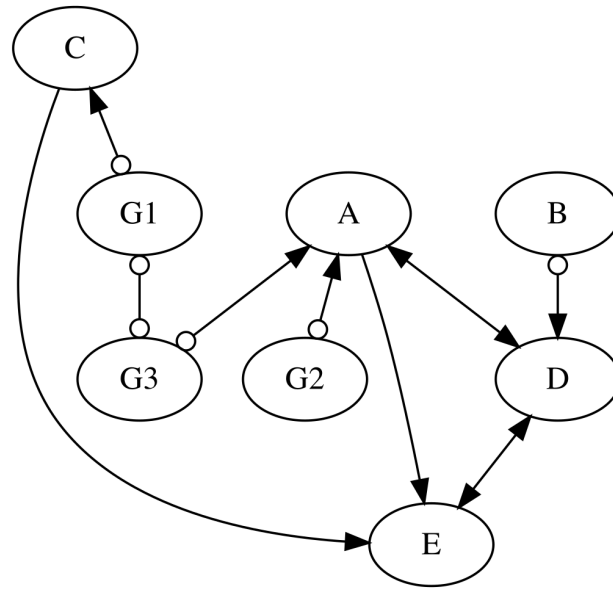


Figure 3: AnchorFCI's PAG.

```
#####
# For the sake of comparison with RFCI: #
#####
```

```
fit_rfci <- pcalg::rfci(suffStat, indepTest = indepTest,
  skel.method = "stable",
  labels = c(var_names, prec_var_names), m.max=2,
  Ndelete = FALSE, #type = "normal",
  alpha = 0.05,
  verbose = TRUE, conservative = FALSE,
  maj.rule = TRUE)
```

```
renderAG(fit_rfci@amat, add_index = FALSE, fileid="RFCI_outputPAG") # RFCI's PAG
```

```
# Uncomment the following lines to render all PAGs
# renderAG(true.amat.mag, add_index = FALSE) # True PAG
# renderAG(true.amat.pag, add_index = FALSE) # True PAG
# renderAG(fit_rfci@amat, add_index = FALSE) # RFCI's PAG
# renderAG(out_anchorFCI$fit_anchor_fci@amat, add_index = FALSE) # anchorFCI's PAG
```

```
#####
# Computing SHD #
#####
```

```
# comparing true pag with true mag considering only var_names
sel_vars <- var_names
```

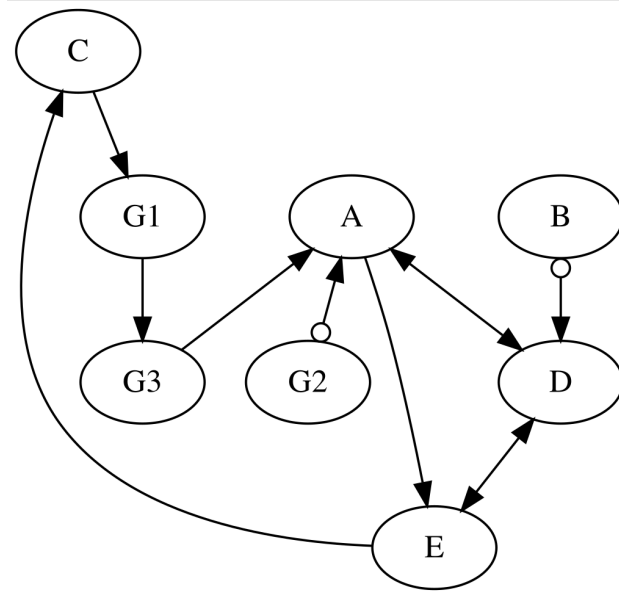


Figure 4: RFCT's PAG.

```

mec_shd <- shd_PAG(true.amat.mag[sel_vars, sel_vars], true.amat.pag[sel_vars, sel_vars])
anchor_fci_shd <- shd_PAG(true.amat.mag[sel_vars, sel_vars],
                          out_anchorFCI$fit_anchor_fci@amat[sel_vars, sel_vars]) - mec_shd
rfci_shd <- shd_PAG(true.amat.mag[sel_vars, sel_vars],
                    fit_rfci@amat[sel_vars, sel_vars]) - mec_shd

c(anchor_fci_shd=anchor_fci_shd, rfci_shd=rfci_shd)
#> anchor_fci_shd      rfci_shd
#>          -1          1

```