Simulation Results: Power

Contents

We review results from our simulations for power.

0.1 Prepare Working Environment

We begin by loading the relevant R packages:

```
pkgs_to_load <- c('dplyr', 'ggplot2', 'tidyr', 'viridis', 'cowplot',
                  'kableExtra')
lapply(X = pkgs_to_load, FUN = library, character.only = TRUE)
options(knitr.table.format = 'html')
```

We then define directories for relevant scripts and data:

```
dir_main <- dirname(dirname(rstudioapi::getActiveDocumentContext()$path))
dir_src <- file.path(dir_main, 'source_scripts')
source(file.path(dir_src, 'define_directories.R'))
```

Next, we define custom functions and other objects used for plotting:

```
source(file.path(dir_src, "define_plot_functions.R"))
source(file.path(dir_src, "define_plot_settings.R"))
```

All scenarios in this simulation setting share the following parameter values:

```
size_or_power <- "power"
alpha <- 0.05
num_permutations <- 1000
num_replicates <- 1000
error_distribution <- "normal"
signal_strength <- 1
```

0.2 Continuous Features

This section includes scenarios where $\boldsymbol{X}$ was simulated as a continuous random vector.

```
x_type <- 'cts'
```

0.2.1 Feature Dimension $p = 100$

We consider the case where the dimension of $\boldsymbol{X}$ is $p = 100$. We include simulated results for the following set of scenarios:

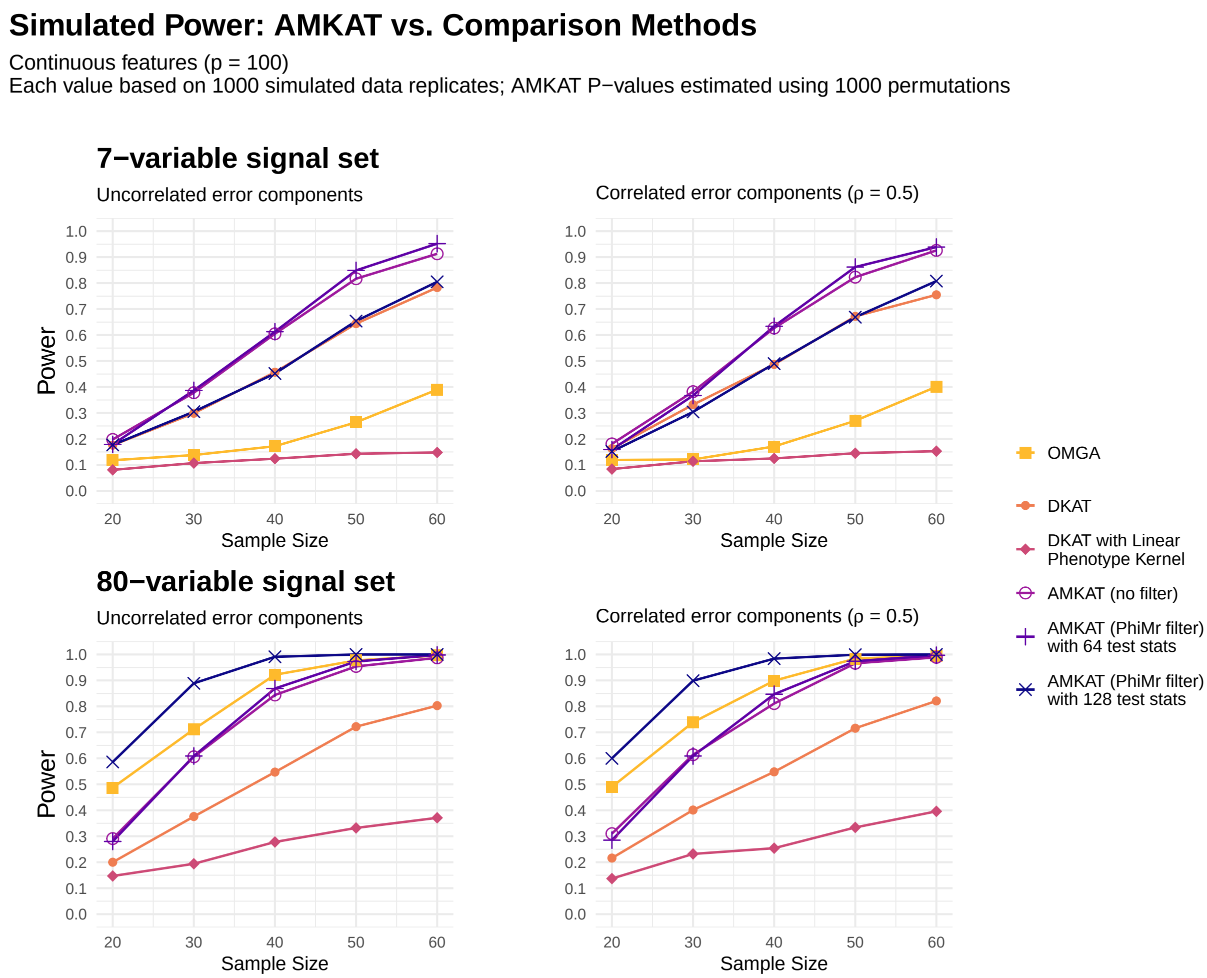
```
p <- 100
values_for_signal_density <- c('sparse', 'dense')
values_for_error_corr_strength <- c(0, 0.5)
values_for_sample_size <- c(20, 30, 40, 50, 60)
```

Each combination of values from the above parameter vectors corresponds to a distinct scenario. We load the combined and formatted results for this set of scenarios:

```
plot_files <- powerPlotFiles()
load(plot_files$plotdata)
```

0.2.1.1 AMKAT vs. Comparison Methods We plot the empirical power of AMKAT and of the comparison methods included in our simulations. We include three variations of AMKAT in this comparison: one without the PhiMr filter and two with the PhiMr filter (one using 64 test statistics and the other using 128).

```
makeCompoundPowerPlot(plot_data_main)
```



Across scenarios involving the 7-variable signal set (top row of plot), the base AMKAT method (without the PhiMr filter) had substantially higher power than each of the three non-AMKAT comparison methods.

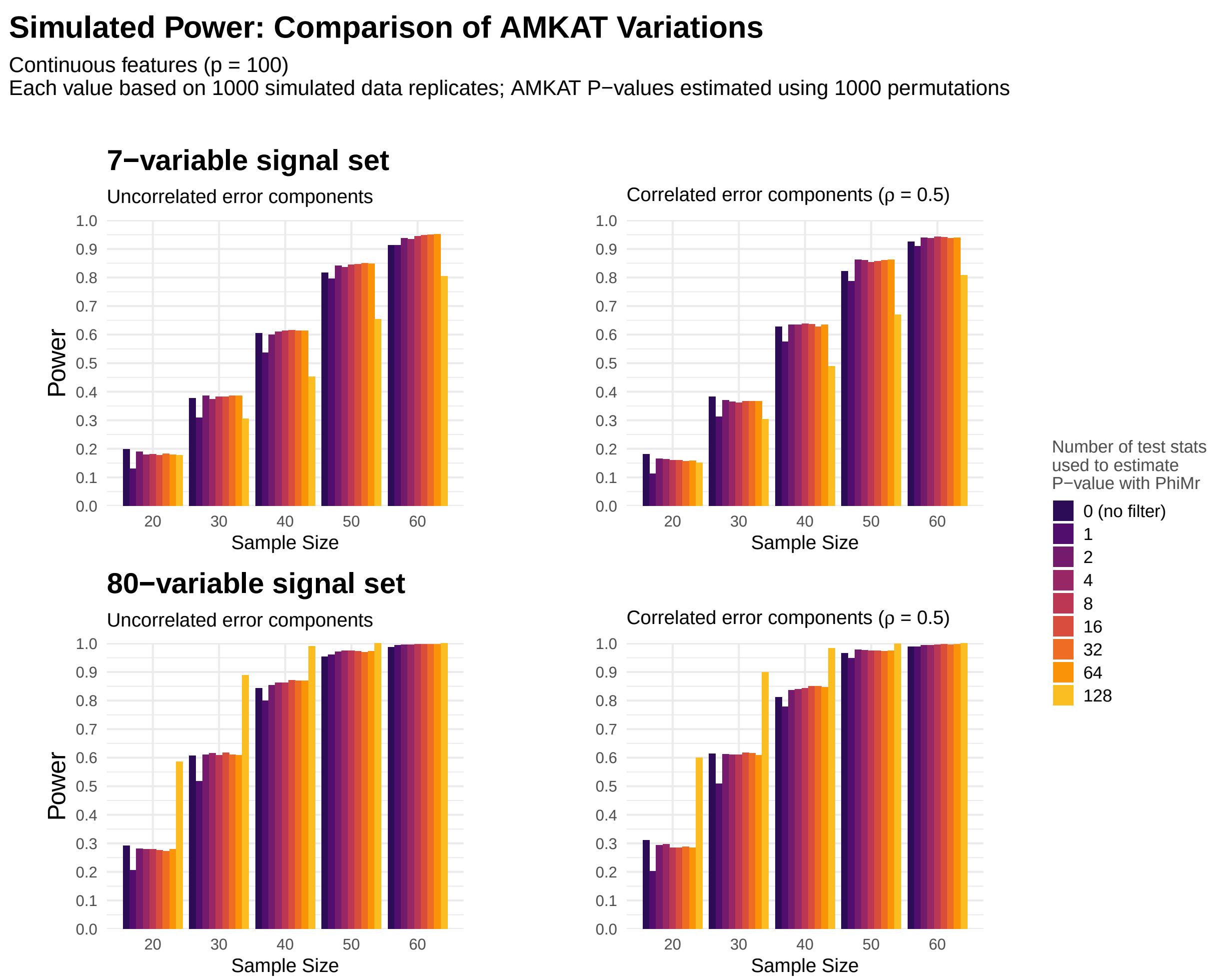
For the 80-variable signal (bottom row of plot), OMGA's power was greater than that of AMKAT; however, subsequent plots reveal this was not the case at greater values of p. The 80-variable signal is fixed across scenarios, such that increasing p serves only to increase the number of noise variables. Thus, OMGA's advantage for the 80-variable signal when $p = 100$ appears to be a consequence of $\boldsymbol{X}$ containing a dramatically higher ratio of signal to noise variables (4 to 1) compared to other cases (all of which had ratios well below 1 to 1).

With 64 test statistics used for P-value estimation, using the PhiMr filter with AMKAT yielded only a small improvement in power compared to AMKAT without PhiMr. When using 128 test statistics for P-value estimation, the story was more intriguing: for the 7-variable signal, power for AMKAT with PhiMr was considerably lower than without PhiMr; for the 80-variable signal, however, AMKAT's power was much higher with PhiMr than without (in fact, the $Q = 128$ variation of AMKAT with PhiMr had the highest power among all methods in the 80-variable signal setting by a large margin). The following plot explores this dynamic in greater detail.

While the method with the most power in each of the two signal settings was an AMKAT method, no single AMKAT method was more powerful than all three non-AMKAT comparison methods under both settings: Under the 7-variable signal setting, AMKAT without PhiMr (as well as with PhiMr at $Q = 64$) resulted in higher power than non-AMKAT methods, but AMKAT with PhiMr at $Q = 128$ performed no better than DKAT; under the 80-variable signal setting, AMKAT with PhiMr at $Q = 128$ was more powerful than non-AMKAT methods, but other AMKAT methods were less powerful than OMGA.

0.2.1.2 Comparison of AMKAT Methods We compare the empirical power of all AMKAT variations included in this simulation setting; these include AMKAT without the PhiMr filter and AMKAT with PhiMr using different variations for the number Q of test statistics used for P-value estimation.

```
makeCompoundPowerBarPlot(
  plot_data_amkat,
  theme_settings = theme_power(show_legend_title = TRUE,
                               legend_title_size = rel(0.8)))
```



Compared to AMKAT without testing subset selection, incorporating PhiMr with a single test statistic ($Q = 1$) resulted in a modest loss of power in essentially every scenario.

Increasing to $Q = 2$ was sufficient to achieve power at or slightly above that achieved without PhiMr. Each successive variation of AMKAT with PhiMr doubled the value of Q; up through $Q = 64$, this did not consistently yield a noticeable improvement in power across scenarios when compared to $Q = 2$.

Doubling Q once more, from 64 to 128, resulted in a stark difference in power that showed a consistent pattern across scenarios for the same signal set but a highly divergent pattern between the two signal sets: for the 7-variable signal set, using $Q = 128$ resulted in substantially lower power compared to $Q = 64$; meanwhile, for the 80-variable signal set, $Q = 128$ yielded much higher power under the same comparison.

Another interesting pattern is that for the 7-variable signal, the power deficit of $Q = 128$ compared to $Q = 64$ was trivial at $n = 20$ and grew as n increased, while for the 80-variable signal, the power surplus of $Q = 128$ over $Q = 64$ was at its greatest when $n = 20$ and diminished as n grew. Thus, despite the difference in the relative power of $Q = 128$ to $Q = 64$ when comparing the two signal sets, we do see a consistent pattern for both settings in that this relative power is greater for smaller sample sizes.

The reason for the vast difference in power of the $Q = 128$ variation between the 7-variable and 80-variable signal settings is unclear; so too is the reason for its sharp deviation in power compared to the near-constant power levels observed across variations from $Q = 2$ to $Q = 64$.

If we refer to results from the simulations for setting $\mathbb{S}_2$ (exploring the distribution of AMKAT statistics and P-values on fixed data sets), which considered values of Q up to 512, we observe for each data set that the empirical distribution of the P-value estimator exhibited a smooth progression of change both in its variance and in its mean as Q increased from 2 to 512, without any kinks or anomalies that would help to explain why our power results were constant while successively doubling Q from 2 up to 64 yet changed so dramatically on the final doubling to $Q = 128$.

This presents motivation for further study by conducting power simulations that consider a greater range of values for Q and provide more granularity in the variations of its value, in order to simultaneously attain a broader and more textured view of the relationship between Q and empirical power across different settings.

0.2.1.3 Full Table of Values

We display the complete table of values for the results in both of the preceding plots:

```
load(file.path(
  dir_tables,
  paste0(size_or_power, '_a', alpha * 1000, '_b', num_permutations,
         '_m', num_replicates, '_', x_type, '_p', p, '.Rdata')))
kbl(compiled_results, escape = F, booktabs = T,
    caption = caption) %>%
  # kable_styling(full_width = T) %>%
  # kable_styling(latex_options = 'scale_down') %>%
  kable_styling() %>%
  add_header_above(header_row) %>%
  pack_rows(index = row_group_index)
```

0.2.2 Feature Dimension $p = 500$

We consider the case where the dimension of $\boldsymbol{X}$ is $p = 500$. We include simulated results for the following set of scenarios:

```
p <- 500
values_for_signal_density <- c('sparse', 'dense')
values_for_error_corr_strength <- c(0, 0.5)
values_for_sample_size <- c(50, 75, 100, 125, 150)
```

Each combination of values from the above parameter vectors corresponds to a distinct scenario. We load the combined and formatted results for this set of scenarios:

```
plot_files <- powerPlotFiles()
load(plot_files$plotdata)
```

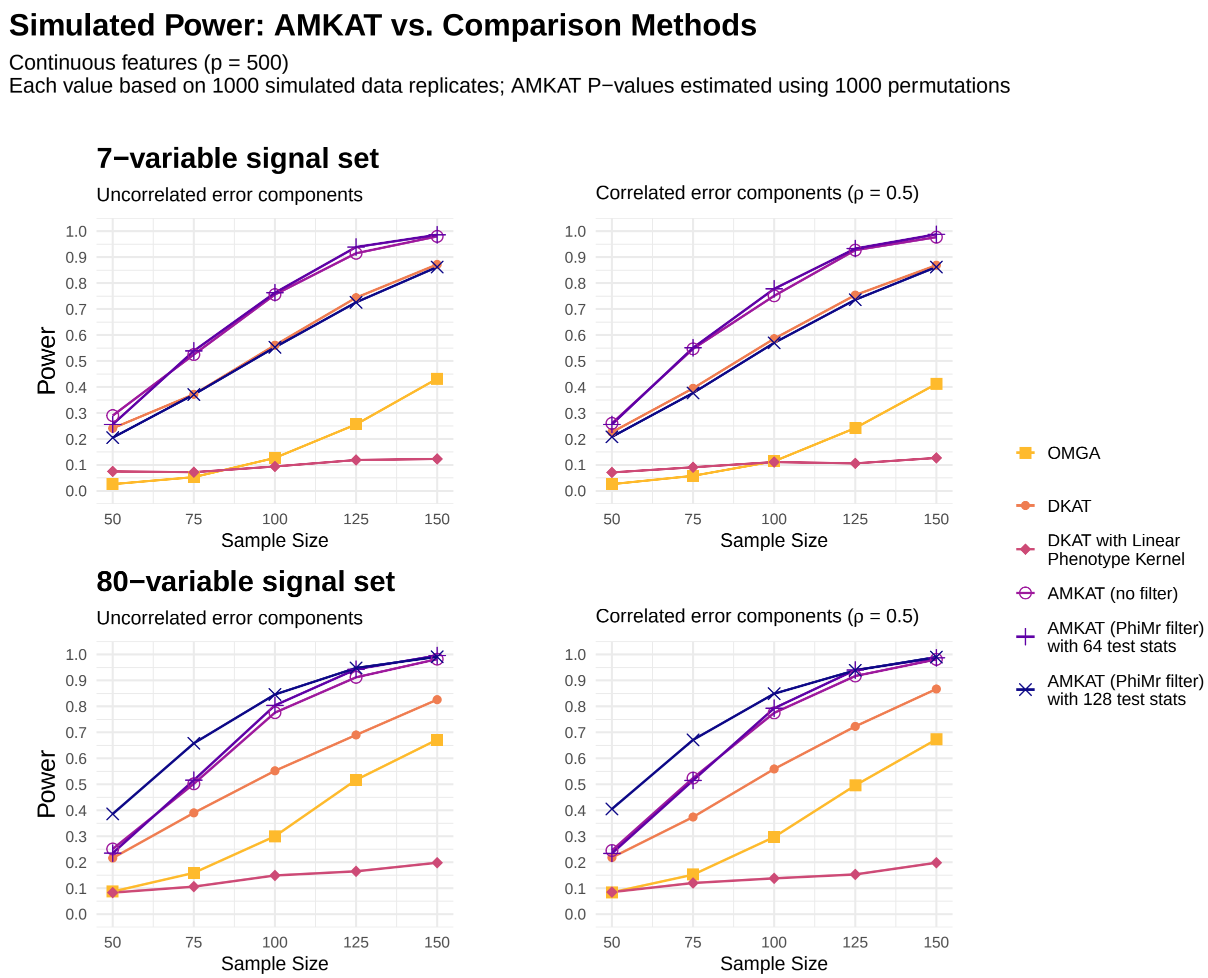
0.2.2.1 AMKAT vs. Comparison Methods

We plot the empirical power of AMKAT and of the comparison methods included in our simulations. We include three variations of AMKAT in this comparison: one without the PhiMr filter and two with the PhiMr filter (one using 64 test statistics and the other using 128).

```
makeCompoundPowerPlot(plot_data_main)
```

Table 1: Empirical power (percentage) across 1000 simulated data replicates using continuous features ($p = 100$). All tests performed at significance level $\alpha = 0.05$. AMKAT P-values estimated using 1000 permutations. AMKAT-Phimr-Qm denotes AMKAT with PhiMr filter using $Q = m$ test statistics for P-value estimation; DKAT-LPK denotes DKAT with linear phenotype kernel.

	Uncorrelated error components					Correlated error components				
	$n=20$	$n=30$	$n=40$	$n=50$	$n=60$	$n=20$	$n=30$	$n=40$	$n=50$	$n=60$
7-variable signal set										
AMKAT (no filter)	19.8	37.8	60.5	81.7	91.3	18.1	38.2	62.7	82.3	92.6
AMKAT-PhiMr-Q1	13.0	30.9	53.6	79.6	91.3	11.3	31.2	57.6	78.8	91.0
AMKAT-PhiMr-Q2	19.0	38.7	59.9	84.1	93.7	16.6	37.0	63.5	86.3	94.0
AMKAT-PhiMr-Q4	18.0	37.4	61.0	83.6	93.4	16.3	36.5	63.5	86.1	93.7
AMKAT-PhiMr-Q8	18.1	38.3	61.3	84.5	94.5	16.0	36.1	63.9	85.4	94.3
AMKAT-PhiMr-Q16	17.8	38.3	61.6	84.6	94.8	16.0	36.6	63.7	85.8	94.1
AMKAT-PhiMr-Q32	18.3	38.6	61.3	85.0	95.0	15.6	36.6	62.8	86.0	93.8
AMKAT-PhiMr-Q64	17.9	38.7	61.3	84.9	95.2	15.9	36.7	63.4	86.2	93.9
AMKAT-PhiMr-Q128	17.7	30.5	45.2	65.4	80.5	15.2	30.4	49.0	66.9	80.8
OMGA	11.8	13.8	17.2	26.4	39.0	11.9	12.1	17.1	27.0	40.1
DKAT	17.6	29.9	45.7	64.4	78.3	16.3	33.2	48.7	67.2	75.5
DKAT-LPK	8.1	10.7	12.4	14.3	14.8	8.4	11.4	12.5	14.5	15.3
80-variable signal set										
AMKAT (no filter)	29.1	60.6	84.4	95.4	98.7	31.0	61.4	81.1	96.6	98.9
AMKAT-PhiMr-Q1	20.6	51.8	80.0	96.1	99.3	20.3	50.8	77.9	94.8	98.9
AMKAT-PhiMr-Q2	28.1	61.1	85.4	97.1	99.5	29.3	61.2	83.7	97.8	99.4
AMKAT-PhiMr-Q4	28.0	61.6	86.3	97.5	99.6	29.6	61.1	83.9	97.7	99.4
AMKAT-PhiMr-Q8	27.9	60.9	86.2	97.5	99.7	28.5	61.0	84.4	97.4	99.6
AMKAT-PhiMr-Q16	27.6	61.7	87.2	97.3	99.7	28.4	61.8	85.0	97.4	99.7
AMKAT-PhiMr-Q32	27.3	61.1	86.9	97.0	99.8	28.8	61.5	85.0	97.3	99.6
AMKAT-PhiMr-Q64	28.0	60.9	86.9	97.3	99.8	28.5	60.9	84.7	97.4	99.7
AMKAT-PhiMr-Q128	58.6	88.9	99.1	100.0	100.0	60.0	89.9	98.4	99.9	100.0
OMGA	48.7	71.2	92.2	97.7	99.6	49.0	73.9	89.9	98.4	99.5
DKAT	20.0	37.6	54.7	72.2	80.3	21.6	40.1	54.8	71.6	82.1
DKAT-LPK	14.7	19.4	27.8	33.2	37.1	13.7	23.2	25.4	33.4	39.6



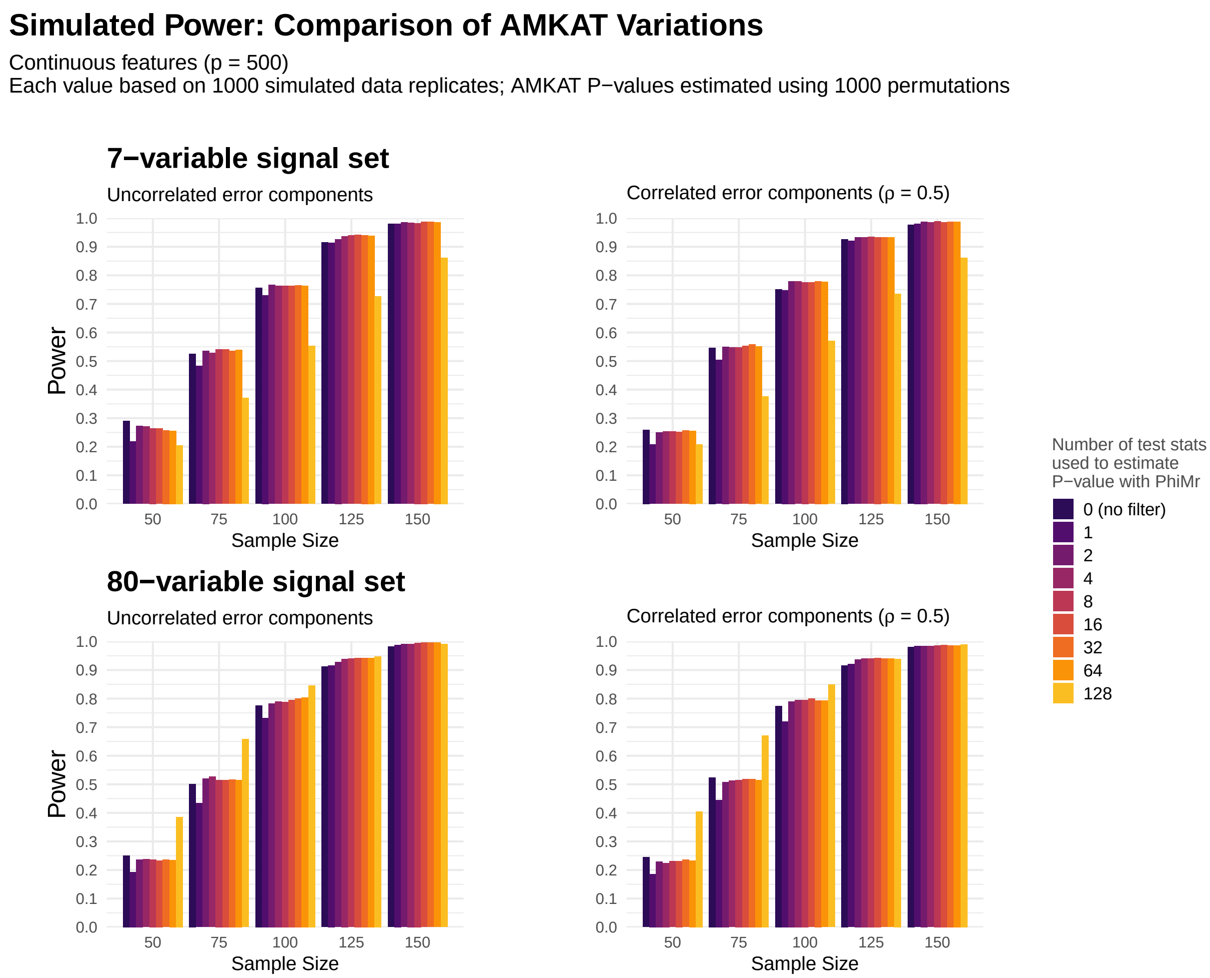
For these scenarios, AMKAT without PhiMr (as well as with PhiMr using $2 \leq Q \leq 64$ test statistics) had substantially higher power than all non-AMKAT methods under both the 7-variable signal and 80-variable signal settings.

Following the pattern observed when $p = 100$, AMKAT with PhiMr and $Q = 128$ test statistics had lower power than the other two AMKAT methods for the 7-variable signal, while for the 80-variable signal it had higher power.

Our results were consistent with the trend of OMGA's power (relative to other methods) decreasing with the ratio of signal to noise variables: its relative power for the 80-variable signal, with a signal-to-noise-variable ratio of 0.16, was dramatically lower than when $p = 100$ (for which the ratio was 4).

0.2.2.2 Comparison of AMKAT Methods We compare the empirical power of all AMKAT variations included in this simulation setting; these include AMKAT without the PhiMr filter and AMKAT with PhiMr using various numbers of test statistics.

```
makeCompoundPowerBarPlot(
  plot_data_amkat,
  theme_settings = theme_power(show_legend_title = TRUE,
                               legend_title_size = rel(0.8)))
```



We observe similar patterns as those seen for the case where $p = 100$: AMKAT with PhiMr and $2 \leq Q \leq 64$ test statistics had comparable power to AMKAT without PhiMr, AMKAT with PhiMr and $Q = 1$ test statistic had modestly reduced power; also, as observed in the previous plot, AMKAT with PhiMr and $Q = 128$ test statistics had lower power for the 7-variable signal and higher power for the 80-variable signal. However, its advantage for the 80-variable signal was more modest than the same advantage we observed when $p = 100$.

0.2.2.3 Full Table of Values We display the complete table of values for the results in both of the preceding plots:

```
load(file.path(
  dir_tables,
  paste0(size_or_power, '_a', alpha * 1000, '_b', num_permutations,
         '_m', num_replicates, '_', x_type, '_p', p, '.Rdata')))
kbl(compiled_results, escape = F, booktabs = T,
    caption = caption) %>%
  # kable_styling(full_width = T) %>%
  # kable_styling(latex_options = 'scale_down') %>%
  kable_styling() %>%
  add_header_above(header_row) %>%
  pack_rows(index = row_group_index)
```

Table 2: Empirical power (percentage) across 1000 simulated data replicates using continuous features ($p = 500$). All tests performed at significance level $\alpha = 0.05$. AMKAT P-values estimated using 1000 permutations. AMKAT-Phimr-Qm denotes AMKAT with PhiMr filter using $Q = m$ test statistics for P-value estimation; DKAT-LPK denotes DKAT with linear phenotype kernel.

	Uncorrelated error components					Correlated error components				
	$n = 50$	$n = 75$	$n = 100$	$n = 125$	$n = 150$	$n = 50$	$n = 75$	$n = 100$	$n = 125$	$n =$ 1
7-variable signal set										
AMKAT (no filter)	29.0	52.5	75.6	91.5	98.0	26.0	54.7	75.1	92.7	9
AMKAT-PhiMr-Q1	21.8	48.3	73.1	91.4	98.0	20.8	50.4	74.8	92.1	9
AMKAT-PhiMr-Q2	27.3	53.6	76.7	92.7	98.5	25.0	55.0	78.0	93.3	9
AMKAT-PhiMr-Q4	27.1	52.8	76.3	93.6	98.4	25.3	54.8	77.9	93.3	9
AMKAT-PhiMr-Q8	26.5	54.1	76.3	94.0	98.3	25.3	54.8	77.6	93.5	9
AMKAT-PhiMr-Q16	26.4	54.1	76.3	94.2	98.8	25.2	55.3	77.5	93.4	9
AMKAT-PhiMr-Q32	25.8	53.5	76.5	94.1	98.7	25.7	55.9	78.0	93.3	9
AMKAT-PhiMr-Q64	25.6	53.9	76.3	93.9	98.6	25.6	55.1	77.8	93.3	9
AMKAT-PhiMr-Q128	20.5	37.1	55.3	72.6	86.2	20.8	37.7	57.0	73.6	8
OMGA	2.6	5.3	12.7	25.6	43.2	2.6	5.8	11.4	24.2	4
DKAT	24.1	37.2	56.1	74.4	87.2	22.6	39.5	58.6	75.4	8
DKAT-LPK	7.5	7.2	9.4	11.9	12.3	7.1	9.1	11.1	10.6	1
80-variable signal set										
AMKAT (no filter)	25.1	50.2	77.6	91.2	98.2	24.5	52.4	77.5	91.7	9
AMKAT-PhiMr-Q1	19.3	43.4	73.3	91.6	98.8	18.6	44.5	72.1	92.1	9
AMKAT-PhiMr-Q2	23.7	52.0	78.3	92.8	99.1	23.0	50.9	79.1	93.7	9
AMKAT-PhiMr-Q4	23.8	52.7	79.1	93.9	99.2	22.5	51.3	79.6	94.1	9
AMKAT-PhiMr-Q8	23.7	51.6	78.9	94.0	99.5	23.1	51.6	79.6	94.1	9
AMKAT-PhiMr-Q16	23.4	51.6	79.6	94.3	99.6	23.2	51.9	80.0	94.3	9
AMKAT-PhiMr-Q32	23.6	51.7	80.1	94.3	99.6	23.6	51.9	79.4	94.0	9
AMKAT-PhiMr-Q64	23.5	51.6	80.4	94.3	99.6	23.4	51.5	79.3	94.0	9
AMKAT-PhiMr-Q128	38.6	65.8	84.6	94.8	99.1	40.5	67.1	84.9	93.9	9
OMGA	8.7	15.9	30.0	51.7	67.2	8.4	15.2	29.7	49.6	6
DKAT	21.6	39.0	55.2	69.0	82.6	21.8	37.4	55.9	72.3	8
DKAT-LPK	8.3	10.6	14.9	16.5	19.8	8.5	12.0	13.8	15.3	1

0.2.3 Feature Dimension $p = 1000$

We consider the case where the dimension of $\boldsymbol{X}$ is $p = 1000$. We include simulated results for the following set of scenarios:

```
p <- 1000
values_for_signal_density <- c('sparse', 'dense')
values_for_error_corr_strength <- c(0, 0.5)
values_for_sample_size <- c(50, 100, 150, 200)
```

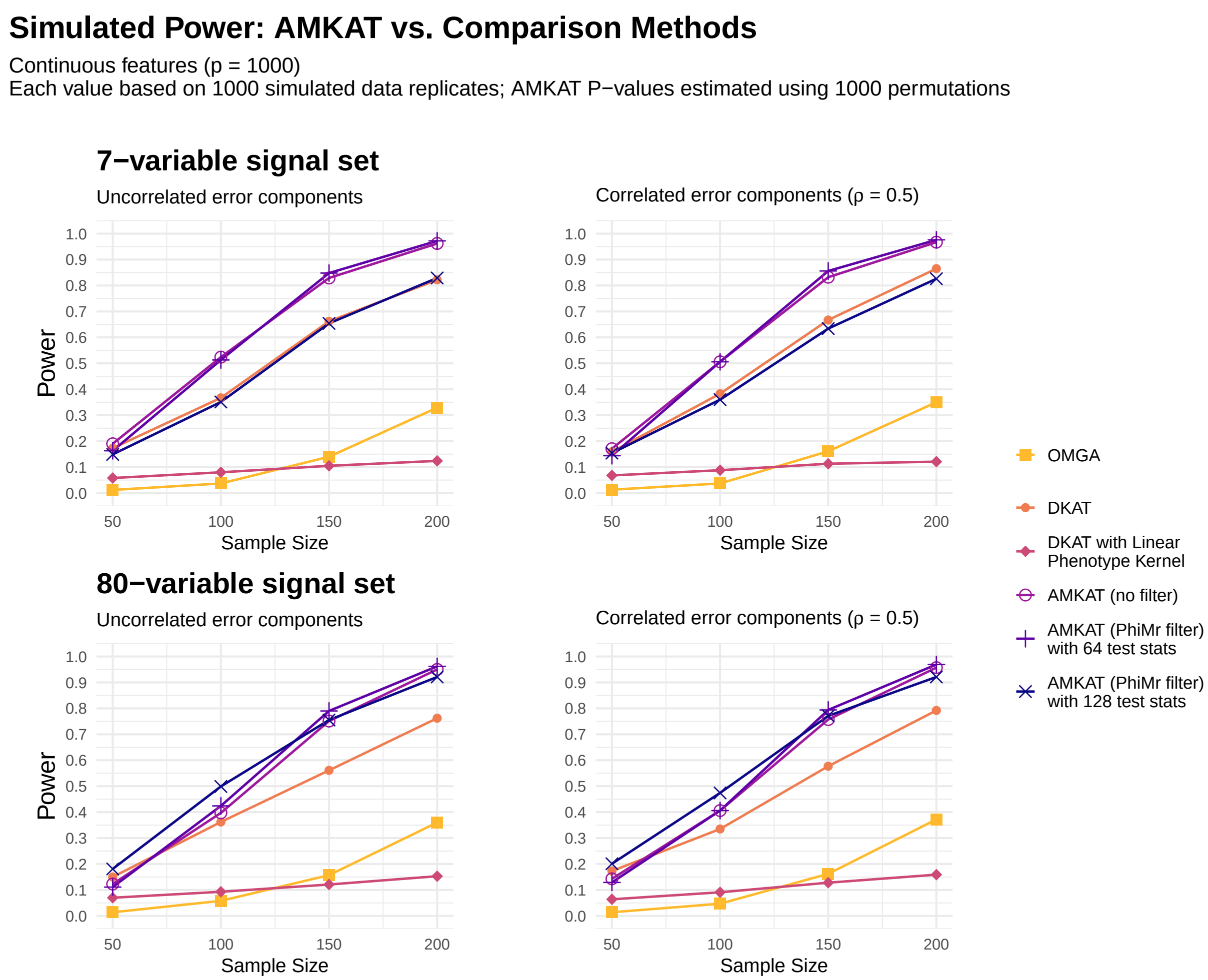
Each combination of values from the above parameter vectors corresponds to a distinct scenario. We load the combined and formatted results for this set of scenarios:

```
plot_files <- powerPlotFiles()
load(plot_files$plotdata)
```

0.2.3.1 AMKAT vs. Comparison Methods

We plot the empirical power of AMKAT and of the comparison methods included in our simulations. We include three variations of AMKAT in this comparison: one without the PhiMr filter and two with the PhiMr filter (one using 64 test statistics and the other using 128).

```
makeCompoundPowerPlot(plot_data_main)
```



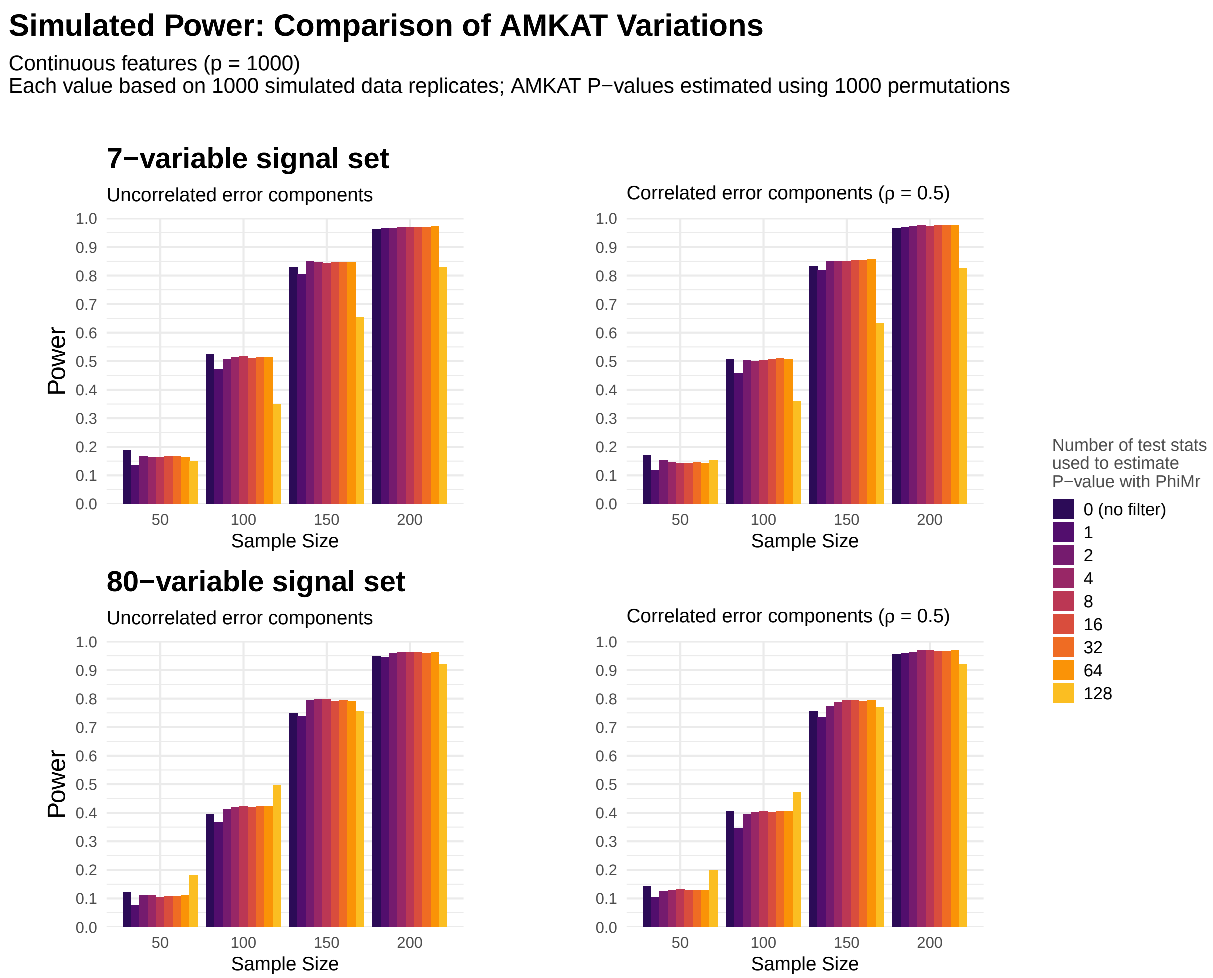
We see results quite similar to those for $p = 500$, with only relatively minor differences.

For the 80-variable signal, the relative power advantage of AMKAT with PhiMr and $Q = 128$ test statistics over other variations of AMKAT was smaller compared to the case where $p = 500$, continuing the trend seen there (where the same advantage was smaller compared to the case where $p = 100$). Here, this advantage actually disappears for sufficiently large sample size, with $Q = 128$ yielding slightly lower power than other AMKAT variations when $n \geq 150$.

For the 7-variable signal, however, the deficit in power with $Q = 128$ (relative to other AMKAT methods) observed at $p = 100$ was similar to that observed at $p = 500$ and $p = 1000$.

0.2.3.2 Comparison of AMKAT Methods We compare the empirical power of all AMKAT variations included in this simulation setting; these include AMKAT without the PhiMr filter and AMKAT with PhiMr using various numbers of test statistics.

```
makeCompoundPowerBarPlot(
  plot_data_amkat,
  theme_settings = theme_power(show_legend_title = TRUE,
                               legend_title_size = rel(0.8)))
```



0.2.3.3 Full Table of Values We display the complete table of values for the results in both of the preceding plots:

```
load(file.path(
  dir_tables,
```

Table 3: Empirical power (percentage) across 1000 simulated data replicates using continuous features ($p = 1000$). All tests performed at significance level $\alpha = 0.05$. AMKAT P-values estimated using 1000 permutations. AMKAT-Phimr-Qm denotes AMKAT with PhiMr filter using $Q = m$ test statistics for P-value estimation; DKAT-LPK denotes DKAT with linear phenotype kernel.

	Uncorrelated error components				Correlated error components			
	$n = 50$	$n = 100$	$n = 150$	$n = 200$	$n = 50$	$n = 100$	$n = 150$	$n = 200$
7-variable signal set								
AMKAT (no filter)	19.0	52.4	82.9	96.2	17.1	50.6	83.2	96.7
AMKAT-PhiMr-Q1	13.6	47.4	80.5	96.6	11.8	46.0	82.0	97.0
AMKAT-PhiMr-Q2	16.7	50.6	85.1	96.7	15.4	50.4	84.9	97.5
AMKAT-PhiMr-Q4	16.4	51.6	84.7	97.0	14.6	49.9	85.1	97.6
AMKAT-PhiMr-Q8	16.3	51.8	84.4	97.1	14.4	50.4	85.2	97.5
AMKAT-PhiMr-Q16	16.6	51.2	84.8	97.1	14.3	50.8	85.4	97.6
AMKAT-PhiMr-Q32	16.7	51.5	84.7	97.1	14.5	51.1	85.5	97.6
AMKAT-PhiMr-Q64	16.3	51.3	84.8	97.2	14.4	50.6	85.6	97.6
AMKAT-PhiMr-Q128	14.9	35.1	65.4	82.9	15.4	36.0	63.4	82.6
OMGA	1.2	3.7	14.0	32.9	1.3	3.7	16.1	35.0
DKAT	17.3	36.7	66.2	82.2	15.9	38.3	66.7	86.5
DKAT-LPK	5.8	8.0	10.5	12.4	6.8	8.8	11.3	12.1
80-variable signal set								
AMKAT (no filter)	12.3	39.7	75.1	95.0	14.3	40.6	75.7	95.7
AMKAT-PhiMr-Q1	7.6	36.9	73.8	94.5	10.4	34.6	73.7	95.8
AMKAT-PhiMr-Q2	11.2	41.3	79.4	95.9	12.6	39.7	77.5	96.3
AMKAT-PhiMr-Q4	11.2	42.2	79.8	96.2	12.8	40.3	78.8	96.9
AMKAT-PhiMr-Q8	10.6	42.4	79.7	96.2	13.2	40.7	79.6	97.1
AMKAT-PhiMr-Q16	11.0	42.2	79.3	96.3	13.1	40.2	79.6	96.7
AMKAT-PhiMr-Q32	10.9	42.5	79.4	96.1	12.8	40.8	79.1	96.8
AMKAT-PhiMr-Q64	11.1	42.4	79.0	96.2	12.9	40.6	79.4	96.9
AMKAT-PhiMr-Q128	18.1	49.9	75.5	92.1	20.1	47.4	77.1	92.1
OMGA	1.4	5.8	15.7	36.0	1.4	4.7	16.2	37.2
DKAT	15.0	36.2	56.1	76.2	17.3	33.5	57.7	79.2
DKAT-LPK	7.0	9.3	12.1	15.3	6.4	9.1	12.8	15.9

```
  paste0(size_or_power, '_a', alpha * 1000, '_b', num_permutations,
         '_m', num_replicates, '_', x_type, '_p', p, '.Rdata')))
kbl(compiled_results, escape = F, booktabs = T,
    caption = caption) %>%
  # kable_styling(full_width = T) %>%
  # kable_styling(latex_options = 'scale_down') %>%
  kable_styling() %>%
  add_header_above(header_row) %>%
  pack_rows(index = row_group_index)
```

0.2.4 Feature Dimension $p = 2000$

We consider the case where the dimension of $\boldsymbol{X}$ is $p = 2000$. We include simulated results for the following set of scenarios:

```
p <- 2000
values_for_signal_density <- c('sparse', 'dense')
```

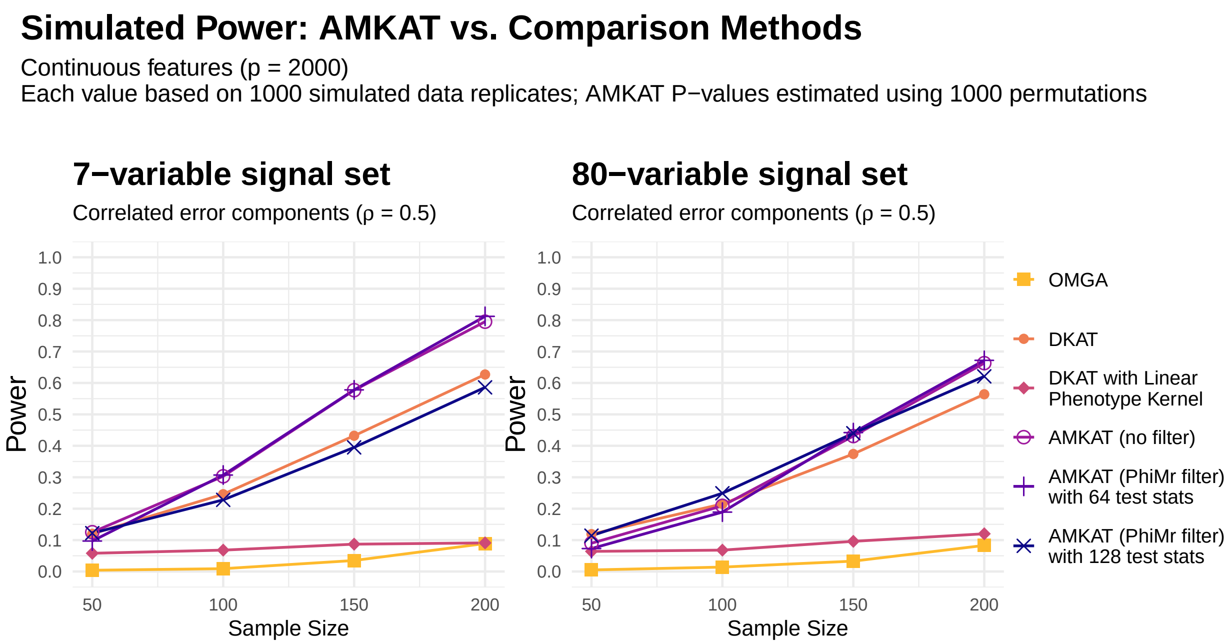
```r
values_for_error_corr_strength <- c(0.5)
values_for_sample_size <- c(50, 100, 150, 200)
```

Each combination of values from the above parameter vectors corresponds to a distinct scenario. We load the combined and formatted results for this set of scenarios:

```r
plot_files <- powerPlotFiles()
load(plot_files$plotdata)
```

0.2.4.1 AMKAT vs. Comparison Methods We plot the empirical power of AMKAT and of the comparison methods included in our simulations. We include three variations of AMKAT in this comparison: one without the PhiMr filter and two with the PhiMr filter (one using 64 test statistics and the other using 128).

```r
makeCompoundPowerPlot(plot_data_main)
```

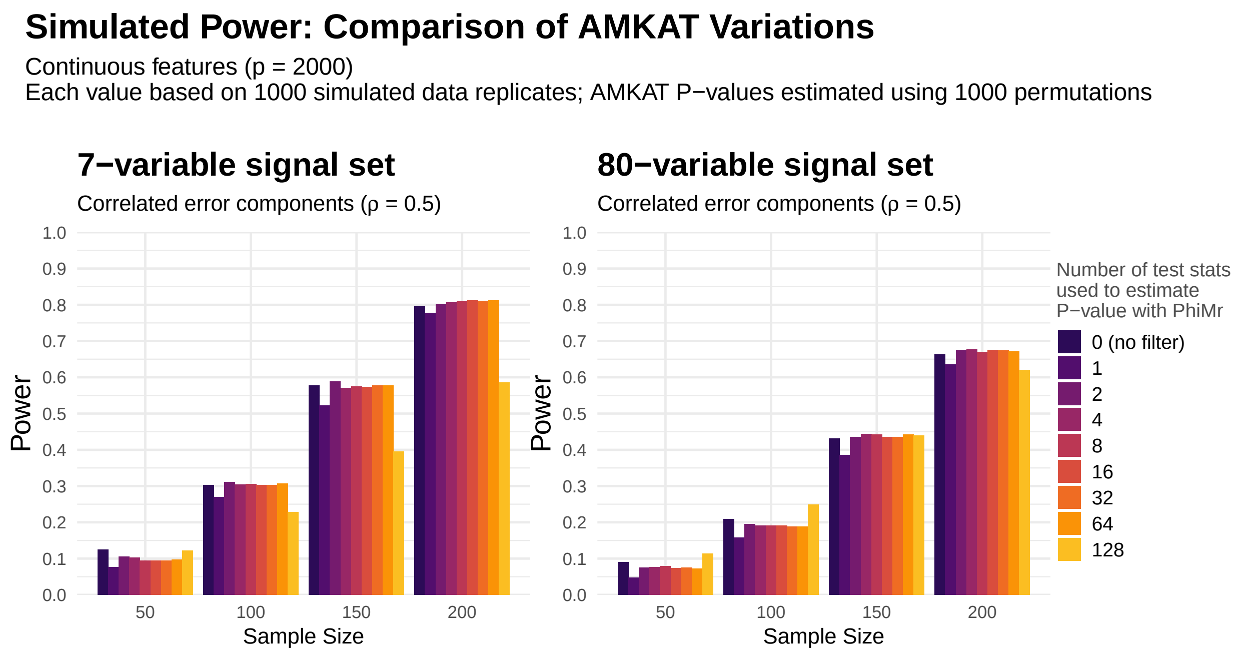


Compared to the case with $p = 1000$, AMKAT's advantage over DKAT (the non-AMKAT method with the greatest power) was similar for the 7-variable signal, but more modest for the 80-variable signal.

Comparing the results in this plot to those at other values of p, we continue to observe the trend of OMGA's power relative to other methods decreasing with the ratio of signal to noise variables.

0.2.4.2 Comparison of AMKAT Methods We compare the empirical power of all AMKAT variations included in this simulation setting; these include AMKAT without the PhiMr filter and AMKAT with PhiMr using various numbers of test statistics.

```r
makeCompoundPowerBarPlot(
  plot_data_amkat,
  theme_settings = theme_power(show_legend_title = TRUE,
                               legend_title_size = rel(0.8)))
```



0.2.4.3 Full Table of Values We display the complete table of values for the results in both of the preceding plots:

```
load(file.path(
  dir_tables,
  paste0(size_or_power, '_a', alpha * 1000, '_b', num_permutations,
         '_m', num_replicates, '_', x_type, '_p', p, '.Rdata')))
kbl(compiled_results, escape = F, booktabs = T,
    caption = caption) %>%
  # kable_styling(full_width = T) %>%
  # kable_styling(latex_options = 'scale_down') %>%
  kable_styling() %>%
  add_header_above(header_row) %>%
  pack_rows(index = row_group_index)
```

0.3 Discrete Features

This section includes scenarios where $\boldsymbol{X}$ was simulated as a discrete random vector representing additive-encoded SNP-set data.

0.3.1 Weakly Correlated Signal Variables

For both of the the signal sets used in this setting, correlation between signal variables tended to be low. Details on the signal sets, including correlation heatmaps, can be found here.

```
signal_correlation <- "low"
```

We include simulated results for the following set of scenarios:

```
x_type <- 'snp'
p <- 567
values_for_signal_density <- c('sparse', 'dense')
```

Table 4: Empirical power (percentage) across 1000 simulated data replicates using continuous features ($p = 2000$). All tests performed at significance level $\alpha = 0.05$. AMKAT P-values estimated using 1000 permutations. AMKAT-Phimr-Qm denotes AMKAT with PhiMr filter using $Q = m$ test statistics for P-value estimation; DKAT-LPK denotes DKAT with linear phenotype kernel.

	Correlated error components			
	$n = 50$	$n = 100$	$n = 150$	$n = 200$
7-variable signal set				
AMKAT (no filter)	12.5	30.3	57.7	79.5
AMKAT-PhiMr-Q1	7.6	27.0	52.3	77.8
AMKAT-PhiMr-Q2	10.6	31.1	58.8	80.1
AMKAT-PhiMr-Q4	10.3	30.5	57.0	80.6
AMKAT-PhiMr-Q8	9.5	30.6	57.5	81.0
AMKAT-PhiMr-Q16	9.5	30.3	57.4	81.2
AMKAT-PhiMr-Q32	9.5	30.3	57.8	81.1
AMKAT-PhiMr-Q64	9.7	30.7	57.8	81.2
AMKAT-PhiMr-Q128	12.2	22.8	39.5	58.6
OMGA	0.4	0.9	3.5	8.9
DKAT	12.0	24.6	43.2	62.7
DKAT-LPK	5.8	6.8	8.7	9.1
80-variable signal set				
AMKAT (no filter)	9.0	20.9	43.1	66.3
AMKAT-PhiMr-Q1	4.7	15.8	38.6	63.6
AMKAT-PhiMr-Q2	7.5	19.5	43.5	67.5
AMKAT-PhiMr-Q4	7.6	19.1	44.4	67.7
AMKAT-PhiMr-Q8	7.9	19.1	44.2	67.0
AMKAT-PhiMr-Q16	7.4	19.1	43.6	67.6
AMKAT-PhiMr-Q32	7.5	18.9	43.6	67.4
AMKAT-PhiMr-Q64	7.3	18.9	44.2	67.2
AMKAT-PhiMr-Q128	11.4	24.9	44.0	62.1
OMGA	0.5	1.4	3.3	8.3
DKAT	11.9	21.5	37.4	56.4
DKAT-LPK	6.4	6.8	9.6	12.0

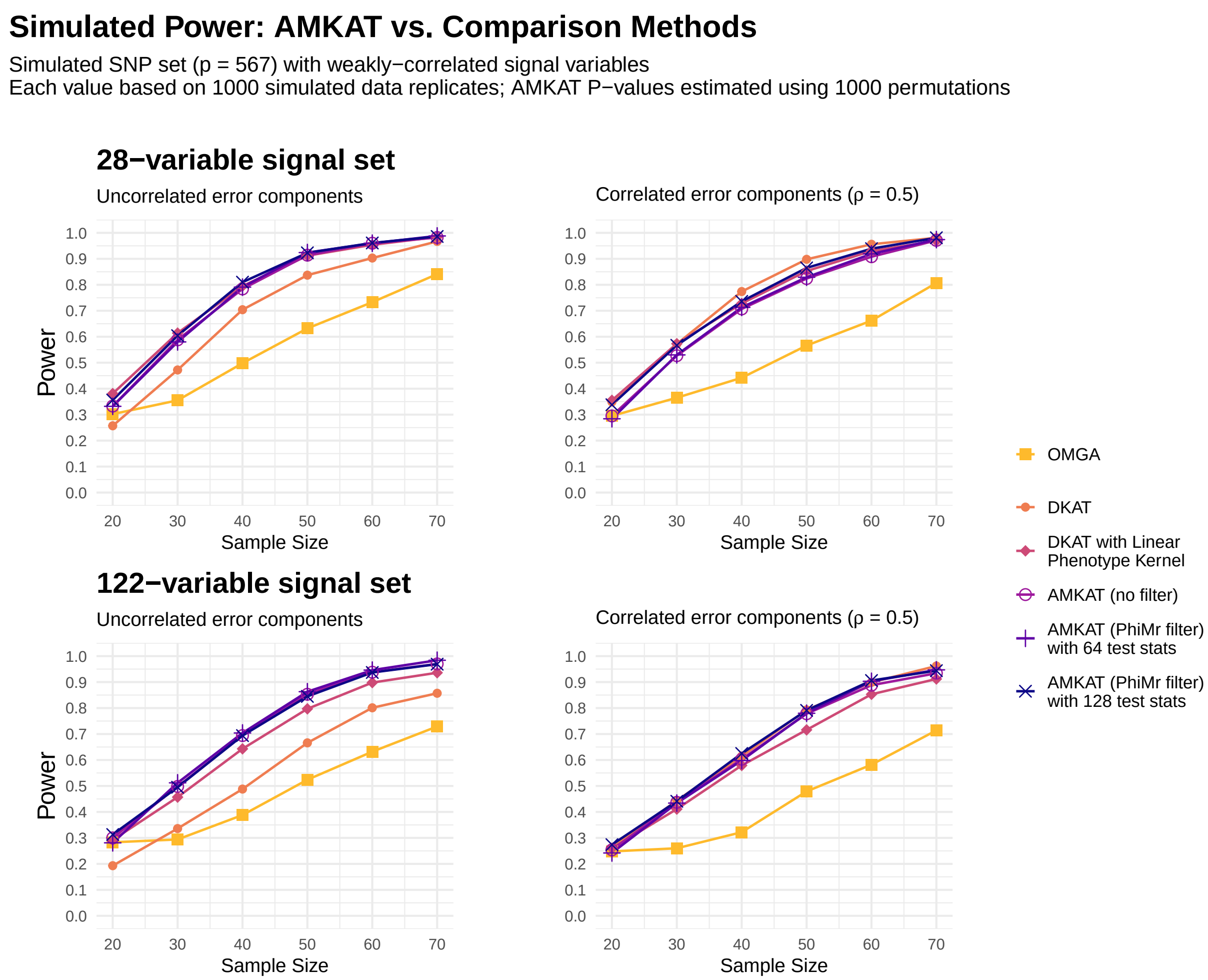
```
values_for_error_corr_strength <- c(0, 0.5)
values_for_sample_size <- c(20, 30, 40, 50, 60, 70)
```

Each combination of values from the above parameter vectors corresponds to a distinct scenario. We load the combined and formatted results for this set of scenarios:

```
plot_files <- powerPlotFiles()
load(plot_files$plotdata)
```

0.3.1.1 AMKAT vs. Comparison Methods We plot the empirical power of AMKAT and of the comparison methods included in our simulations. We include three variations of AMKAT in this comparison: one without the PhiMr filter and two with the PhiMr filter (one using 64 test statistics and the other using 128).

```
makeCompoundPowerPlot(plot_data_main)
```



Among the non-AMKAT methods, across both signal sets, DKAT with a linear phenotype kernel had the highest power when the error components were uncorrelated, while DKAT using the DKAT phenotype kernel had slightly higher power when the scenario parameter $\rho = 0.5$.

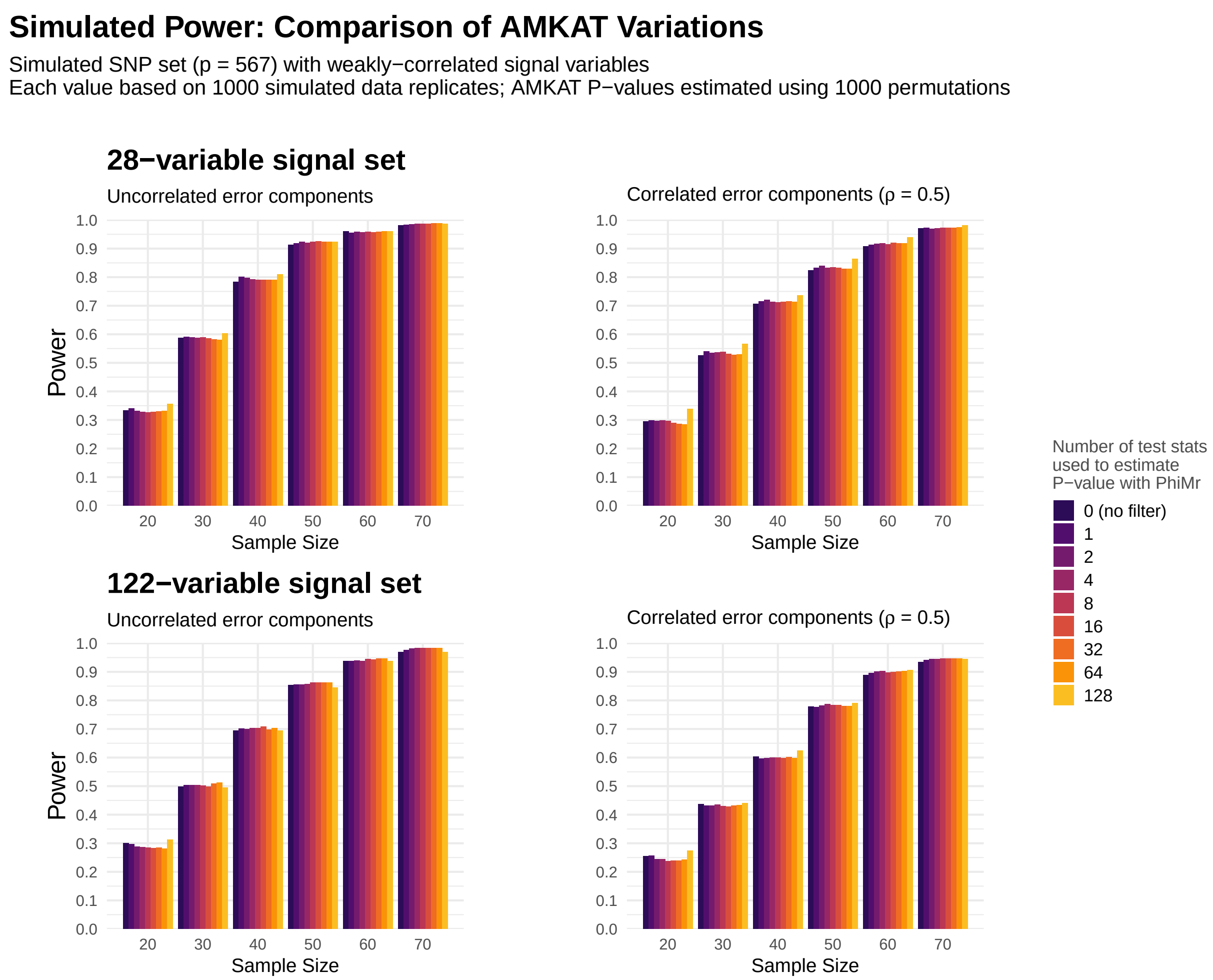
For the 28-variable signal, the power of each AMKAT method was competitive with that of DKAT using the linear phenotype kernel, having power superior to that of the remaining non-AMKAT methods when error components were uncorrelated and slightly below DKAT with the DKAT phenotype kernel when $\rho = 0.5$.

For the 122-variable signal, when the error components were uncorrelated, each AMKAT method had superior power to all non-AMKAT methods, with power modestly above that of DKAT with the linear phenotype kernel; when $\rho = 0.5$, each AMKAT method had power competitive with DKAT using the DKAT phenotype kernel, which was superior to that of the remaining non-AMKAT methods.

In summary, AMKAT was competitive with the most powerful non-AMKAT method in each setting seen here, with modestly higher power for the 122-variable signal when $\rho = 0$, and with marginally lower power for the 28-variable signal when $\rho = 0.5$. The most powerful non-AMKAT method differed across scenarios, such that no non-AMKAT method had power equal to AMKAT's across all four of the settings depicted in the plot.

0.3.1.2 Comparison of AMKAT Methods We compare the empirical power of all AMKAT variations included in this simulation setting; these include AMKAT without the PhiMr filter and AMKAT with PhiMr using various numbers of test statistics.

```
makeCompoundPowerBarPlot(
  plot_data_amkat,
  theme_settings = theme_power(show_legend_title = TRUE,
                               legend_title_size = rel(0.8)))
```



All variations of AMKAT had similar power under this setting, with the variation of AMKAT using PhiMr and $Q = 128$ test statistics having slightly higher power in some scenarios.

0.3.1.3 Full Table of Values We display the complete table of values for the results in both of the preceding plots:

Table 5: Empirical power (percentage) across 1000 simulated data replicates using a simulated SNP-set ($p = 567$)with mildly-correlated signal variables. All tests performed at significance level $\alpha = 0.05$. AMKAT P-values estimated using 1000 permutations. AMKAT-Phimr-Qm denotes AMKAT with PhiMr filter using $Q = m$ test statistics for P-value estimation; DKAT-LPK denotes DKAT with linear phenotype kernel.

	Uncorrelated error components						Correlated error componen			
	$n = 20$	$n = 30$	$n = 40$	$n = 50$	$n = 60$	$n = 70$	$n = 20$	$n = 30$	$n = 40$	$n = 50$
28-variable signal set										
AMKAT (no filter)	33.3	58.8	78.4	91.4	96.0	98.2	29.5	52.7	70.7	82.4
AMKAT-PhiMr-Q1	34.0	59.1	80.1	91.9	95.5	98.4	29.9	54.0	71.6	83.2
AMKAT-PhiMr-Q2	33.2	59.0	79.7	92.3	95.9	98.5	29.6	53.5	72.1	83.9
AMKAT-PhiMr-Q4	32.9	58.7	79.3	92.0	95.7	98.6	29.8	53.7	71.4	83.3
AMKAT-PhiMr-Q8	32.7	58.9	79.0	92.3	95.8	98.7	29.6	53.8	71.2	83.5
AMKAT-PhiMr-Q16	32.9	58.6	79.1	92.6	95.7	98.7	28.9	53.1	71.3	83.2
AMKAT-PhiMr-Q32	33.0	58.2	79.1	92.4	95.9	98.8	28.6	52.8	71.5	83.0
AMKAT-PhiMr-Q64	33.2	58.0	79.1	92.4	96.0	98.8	28.4	53.0	71.3	82.9
AMKAT-PhiMr-Q128	35.6	60.4	81.0	92.3	96.1	98.6	33.8	56.7	73.6	86.5
OMGA	30.2	35.5	49.8	63.2	73.3	84.1	29.5	36.5	44.2	56.6
DKAT	25.7	47.2	70.4	83.7	90.3	96.7	34.9	57.2	77.4	89.8
DKAT-LPK	38.1	61.4	79.5	91.2	95.4	98.4	35.5	57.3	72.8	85.3
122-variable signal set										
AMKAT (no filter)	30.0	49.8	69.4	85.3	93.8	97.0	25.4	43.7	60.3	77.8
AMKAT-PhiMr-Q1	29.6	50.3	70.1	85.5	93.8	97.6	25.7	43.1	59.6	77.7
AMKAT-PhiMr-Q2	28.8	50.4	69.9	85.6	94.0	98.2	24.4	43.1	59.8	78.2
AMKAT-PhiMr-Q4	28.6	50.4	70.4	85.7	93.8	98.3	24.5	43.5	60.0	78.8
AMKAT-PhiMr-Q8	28.4	50.1	70.3	86.2	94.5	98.3	23.7	43.0	59.9	78.4
AMKAT-PhiMr-Q16	28.3	49.9	70.9	86.2	94.3	98.3	23.9	42.9	59.8	78.4
AMKAT-PhiMr-Q32	28.5	50.9	69.8	86.2	94.6	98.4	23.9	43.2	60.1	78.0
AMKAT-PhiMr-Q64	28.1	51.2	70.4	86.3	94.6	98.4	24.2	43.4	59.8	78.0
AMKAT-PhiMr-Q128	31.3	49.5	69.4	84.5	93.7	96.9	27.4	44.1	62.5	79.1
OMGA	28.3	29.4	38.8	52.4	63.2	72.9	24.8	26.0	32.2	47.9
DKAT	19.3	33.6	48.8	66.6	80.1	85.7	26.2	44.4	61.5	79.3
DKAT-LPK	29.2	45.7	64.3	79.7	89.8	93.6	26.0	41.1	57.9	71.6

```r
load(file.path(
  dir_tables,
  paste0(size_or_power, '_a', alpha * 1000, '_b', num_permutations,
         '_m', num_replicates, '_', x_type, '_sc-',
         signal_correlation, '.Rdata')))
kbl(compiled_results, escape = F, booktabs = T,
    caption = caption) %>%
  # kable_styling(full_width = T) %>%
  # kable_styling(latex_options = 'scale_down') %>%
  kable_styling() %>%
  add_header_above(header_row) %>%
  pack_rows(index = row_group_index)
```

0.3.2 Strongly Correlated Signal Variables

For both of the the signal sets used in this setting, correlation between signal variables tended to be very strong, with perfect colinearity between some signals. Details on the signal sets, including correlation heatmaps, can be found here.

```
signal_correlation <- "high"
```

We include simulated results for the following set of scenarios:

```
x_type <- 'snp'
p <- 567
values_for_signal_density <- c('sparse', 'dense')
values_for_error_corr_strength <- c(0, 0.5)
values_for_sample_size <- c(20, 30, 40, 50, 60, 70)
```

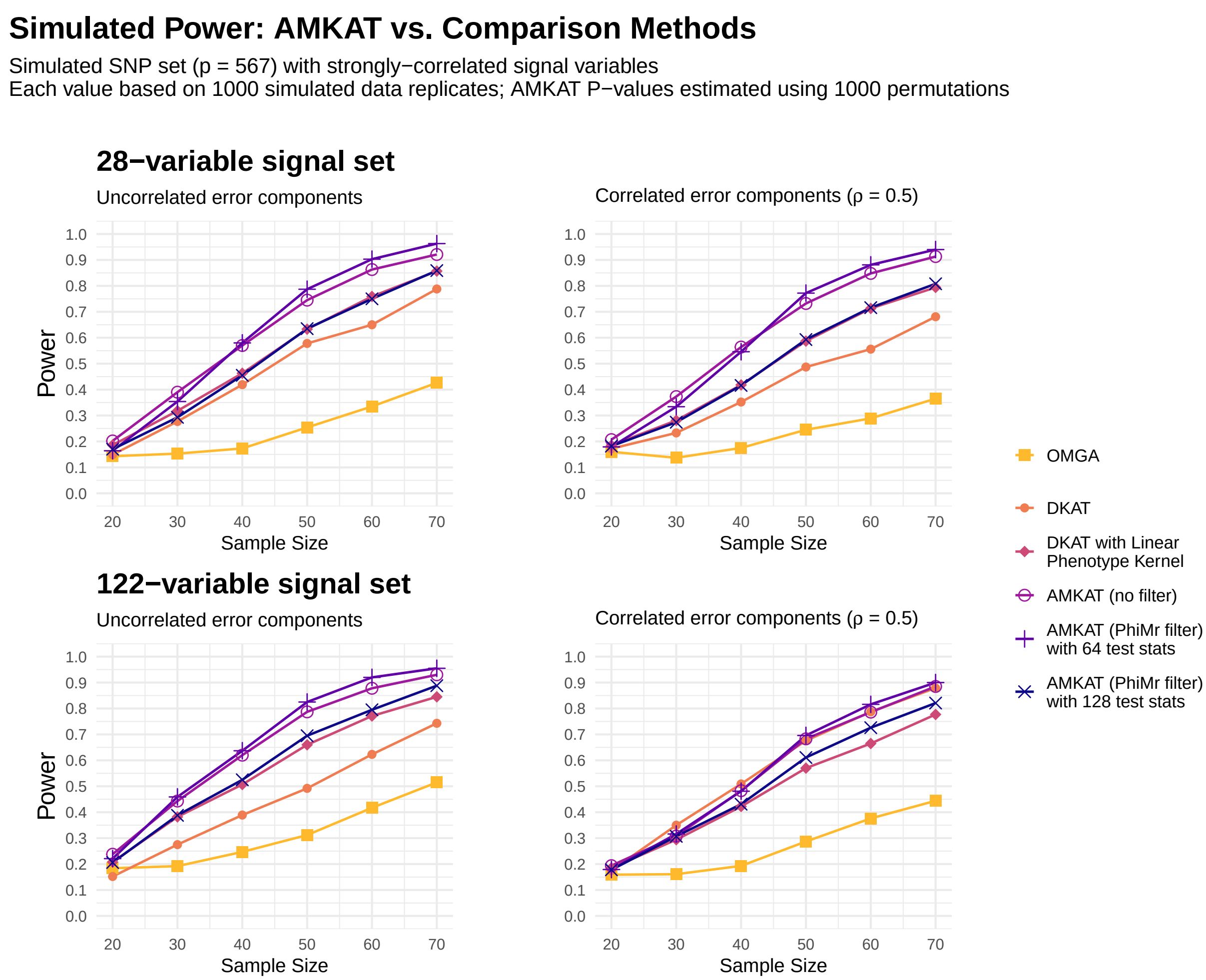
Each combination of values from the above parameter vectors corresponds to a distinct scenario. We load the combined and formatted results for this set of scenarios:

```
plot_files <- powerPlotFiles()
load(plot_files$plotdata)
```

0.3.2.1 AMKAT vs. Comparison Methods

We plot the empirical power of AMKAT and of the comparison methods included in our simulations. We include three variations of AMKAT in this comparison: one without the PhiMr filter and two with the PhiMr filter (one using 64 test statistics and the other using 128).

```
makeCompoundPowerPlot(plot_data_main)
```



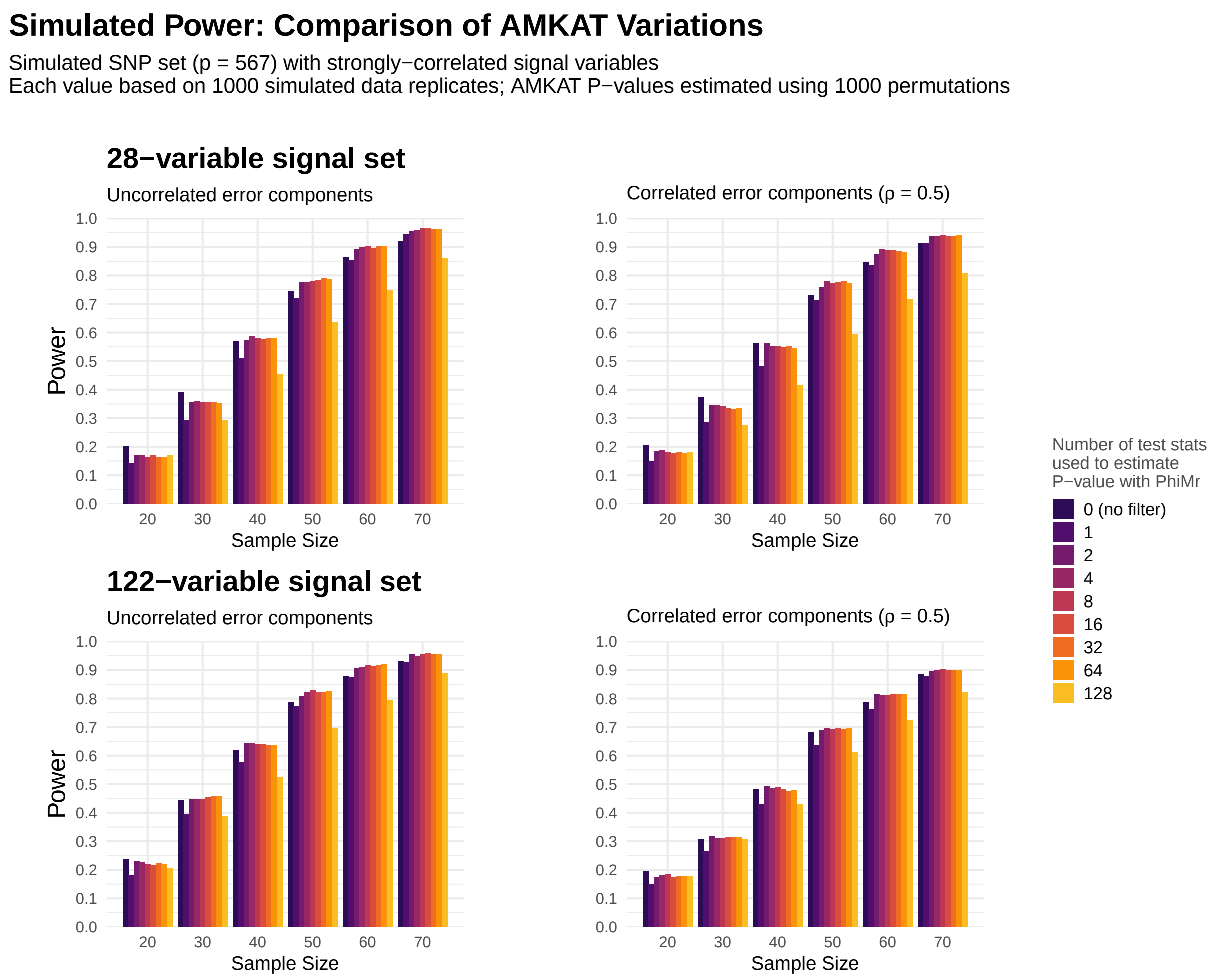
For the 28-variable signal, AMKAT without PhiMr was considerably more powerful than non-AMKAT methods, as was AMKAT using PhiMr with $Q = 64$ test statistics; the former was slightly more powerful at very small sample sizes, while the latter was modestly more powerful for $n \geq 50$. AMKAT using PhiMr with $Q = 128$ test statistics was comparable to DKAT with the linear phenotype kernel, which had the highest power among the non-AMKAT methods. Interestingly, DKAT with the linear phenotype kernel had higher power than DKAT with the DKAT phenotype kernel even when $\rho = 0.5$, despite the fact that the DKAT phenotype kernel accounts for correlations among error components while the linear phenotype kernel ignores them (in other settings with $\rho = 0.5$, the DKAT phenotype kernel yielded higher power than the linear phenotype kernel).

For the 122-variable signal, when error components were uncorrelated, relative power among methods was similar to that for the 28-variable signal, though here the power of DKAT with the DKAT phenotype kernel was farther below that of DKAT with the linear phenotype kernel. When $\rho = 0.5$, DKAT had greater power with the DKAT phenotype kernel than the linear phenotype kernel, with power comparable to that of AMKAT.

In summary, AMKAT was generally more powerful than non-AMKAT methods across these settings, with the PhiMr filter yielding a small increase in power when $Q = 64$ and $n \geq 50$, but yielding considerably lower power when $Q = 128$. DKAT with the DKAT phenotype kernel was competitive with AMKAT for the 122-variable signal and correlated error components, but had much lower power in the other settings considered here.

0.3.2.2 Comparison of AMKAT Methods We compare the empirical power of all AMKAT variations included in this simulation setting; these include AMKAT without the PhiMr filter and AMKAT with PhiMr using various numbers of test statistics.

```
makeCompoundPowerBarPlot(
  plot_data_amkat,
  theme_settings = theme_power(show_legend_title = TRUE,
                               legend_title_size = rel(0.8)))
```



Here we notice that PhiMr yields a small drop in power compared to AMKAT without PhiMr when using only $Q = 1$ test statistic for P-value estimation, similar to the typical case we observed when $\boldsymbol{X}$ was continuous. Power was similar for PhiMr with $2 \leq Q \leq 64$ test statistics, while increasing to $Q = 128$ yielded a noticeable drop in power for both the 28-variable and 122-variable signal sets.

0.3.2.3 Full Table of Values We display the complete table of values for the results in both of the preceding plots:

```
load(file.path(
  dir_tables,
  paste0(size_or_power, '_a', alpha * 1000, '_b', num_permutations,
         '_m', num_replicates, '_', x_type, '_sc-',
         signal_correlation, '.Rdata')))
kbl(compiled_results, escape = F, booktabs = T,
    caption = caption) %>%
  # kable_styling(full_width = T) %>%
  # kable_styling(latex_options = 'scale_down') %>%
  kable_styling() %>%
  add_header_above(header_row) %>%
  pack_rows(index = row_group_index)
```

Table 6: Empirical power (percentage) across 1000 simulated data replicates using a simulated SNP-set ($p = 567$)with highly-correlated signal variables. All tests performed at significance level $\alpha = 0.05$. AMKAT P-values estimated using 1000 permutations. AMKAT-Phimr-Qm denotes AMKAT with PhiMr filter using $Q = m$ test statistics for P-value estimation; DKAT-LPK denotes DKAT with linear phenotype kernel.

	Uncorrelated error components						Correlated error componen			
	$n = 20$	$n = 30$	$n = 40$	$n = 50$	$n = 60$	$n = 70$	$n = 20$	$n = 30$	$n = 40$	$n = 50$
28-variable signal set										
AMKAT (no filter)	20.2	39.0	57.0	74.5	86.3	92.1	20.7	37.3	56.4	73.2
AMKAT-PhiMr-Q1	14.2	29.4	51.0	71.9	85.4	94.6	15.0	28.6	48.3	71.4
AMKAT-PhiMr-Q2	16.9	35.8	57.5	77.8	89.3	95.4	18.4	34.7	56.2	76.0
AMKAT-PhiMr-Q4	17.1	36.0	58.9	77.7	90.0	96.0	18.7	34.6	55.2	78.0
AMKAT-PhiMr-Q8	16.3	35.8	58.0	78.1	90.1	96.4	18.0	34.3	55.3	77.4
AMKAT-PhiMr-Q16	16.9	35.7	57.6	78.5	89.7	96.5	17.9	33.4	54.9	77.5
AMKAT-PhiMr-Q32	16.3	35.8	58.0	79.1	90.3	96.3	18.0	33.2	55.3	77.9
AMKAT-PhiMr-Q64	16.4	35.4	58.0	78.7	90.3	96.3	17.9	33.4	54.6	77.2
AMKAT-PhiMr-Q128	17.0	29.3	45.5	63.5	75.0	85.9	18.2	27.4	41.6	59.3
OMGA	14.3	15.3	17.3	25.4	33.5	42.6	16.0	13.7	17.5	24.5
DKAT	14.9	27.7	41.9	57.8	65.0	78.8	17.2	23.3	35.2	48.7
DKAT-LPK	18.6	31.8	46.3	63.3	75.9	85.7	18.5	28.1	41.8	58.7
122-variable signal set										
AMKAT (no filter)	23.8	44.3	62.0	78.7	87.8	93.0	19.4	30.8	48.3	68.3
AMKAT-PhiMr-Q1	18.2	39.6	57.7	77.4	87.5	92.9	14.9	26.7	43.1	63.6
AMKAT-PhiMr-Q2	22.9	44.7	64.4	81.0	90.7	95.5	17.6	31.9	49.3	69.1
AMKAT-PhiMr-Q4	22.6	44.9	64.3	82.1	91.1	94.8	18.1	31.1	48.6	69.8
AMKAT-PhiMr-Q8	21.9	44.8	64.2	82.8	91.7	95.5	18.4	31.0	49.1	69.2
AMKAT-PhiMr-Q16	21.5	45.5	64.0	82.4	91.5	95.8	17.4	31.4	48.3	69.7
AMKAT-PhiMr-Q32	22.2	45.7	63.8	82.1	91.7	95.7	17.7	31.3	47.7	69.4
AMKAT-PhiMr-Q64	22.1	45.9	63.7	82.5	92.0	95.5	17.9	31.6	48.1	69.6
AMKAT-PhiMr-Q128	20.6	38.8	52.5	69.5	79.5	88.8	17.8	30.7	43.1	61.1
OMGA	18.4	19.2	24.7	31.2	41.7	51.5	15.9	16.1	19.3	28.7
DKAT	15.2	27.5	38.9	49.2	62.3	74.3	17.5	35.0	50.9	67.6
DKAT-LPK	20.8	38.2	50.7	66.0	77.1	84.5	18.5	29.4	42.2	57.0