# A Complete Generalized Adjustment Criterion

Emilija Perković[1], Johannes Textor[2], Markus Kalisch[1], and Marloes H. Maathuis[1]

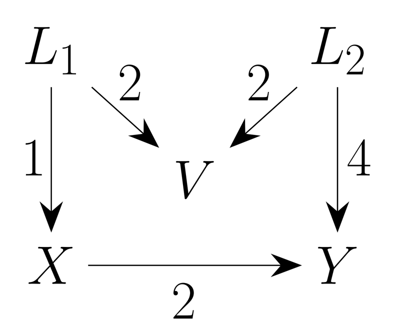
[1]Seminar for Statistics, ETH Zurich, Switzerland
[2]Theoretical Biology & Bioinformatics, Utrecht University, The Netherlands

**Problem** Determining valid adjustment sets for estimating total causal effects from observational data.

**Background** In practice, causal effects are often estimated by adjusted regression. It is non-trivial to determine the variables one should adjust for. This depends on the causal structure, which can be represented by a DAG, MAG, CPDAG or PAG. Several graphical criteria for adjustment exist, but none are complete for all graph classes.

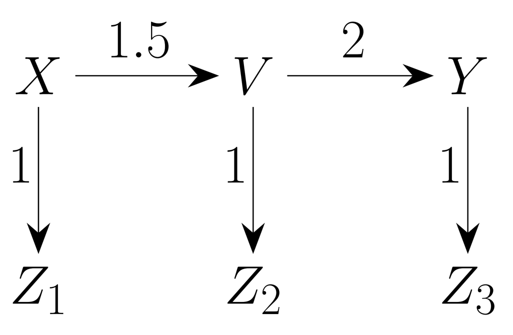
**Contribution** A complete graphical criterion for adjustment in DAGs, MAGs, CPDAGs and PAGs. Our criterion subsumes the existing ones and unifies adjustment set construction for a large set of graph classes.

## Covariate adjustment in DAGs




Total effect of $X$ on $Y$ is $2$.
$\{V\}$ is not a valid adjustment set.

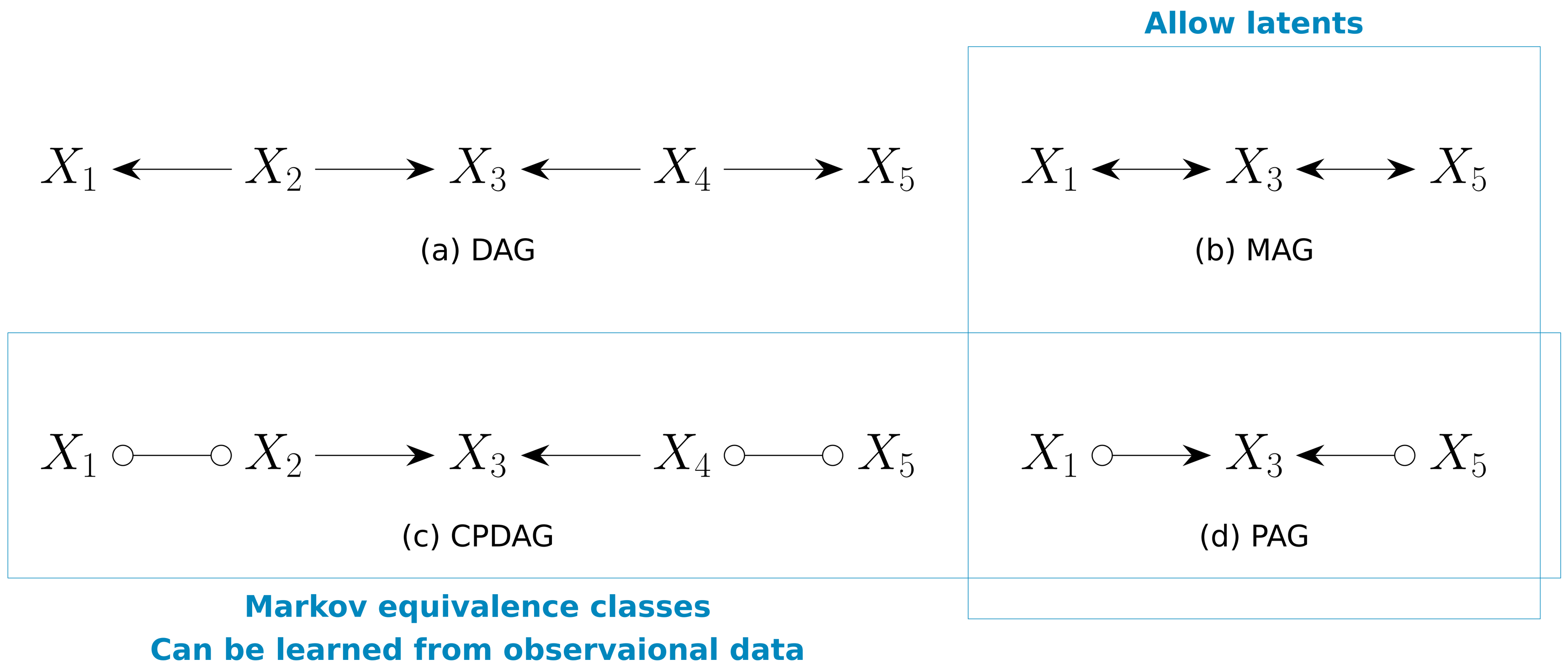
```
lm(Y~X)$coef[2] #1.987355
lm(Y~X+V)$coef[2] #0.8683616
```




Total effect of $X$ on $Y$ is $3$.
$\emptyset, \{Z_1\}$ are the only valid adjustment sets.

```
lm(Y~X+Z1)$coef[2] #3.006502
lm(Y~X+Z2)$coef[2] #1.471494
lm(Y~X+Z3)$coef[2] #0.4919003
```

## Graph types for causal structure




## Graphical criteria for covariate adjustment

| | DAG | MAG | CPDAG | PAG |
|---|---|---|---|---|
| Back-door [?] | ✓ | | | |
| Adjustment [?] | **✓** | | | |
| Adjustment [?] | **✓** | **✓** | | |
| Generalized back-door [?] | ✓ | ✓ | ✓ | ✓ |
| **Generalized adjustment** | **✓** | **✓** | **✓** | **✓** |

✓ - Sound
**✓** - Sound and Complete

## Main result

**Definition** (**Adjustment set**; [**?**]) Let $\mathcal{G}$ represent a DAG, MAG, CPDAG or PAG. Then $\mathbf{Z}$ is an adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ if for any density $f$ consistent with $\mathcal{G}$

$$f(\mathbf{y}|do(\mathbf{x})) = \begin{cases} f(\mathbf{y}|\mathbf{x}) & \text{if } \mathbf{Z} = \emptyset, \\ \int_{\mathbf{Z}} f(\mathbf{y}|\mathbf{x},\mathbf{z}) f(\mathbf{z}) d\mathbf{z} = E_{\mathbf{Z}}\{f(\mathbf{y}|\mathbf{z},\mathbf{x})\} & \text{otherwise.} \end{cases}$$

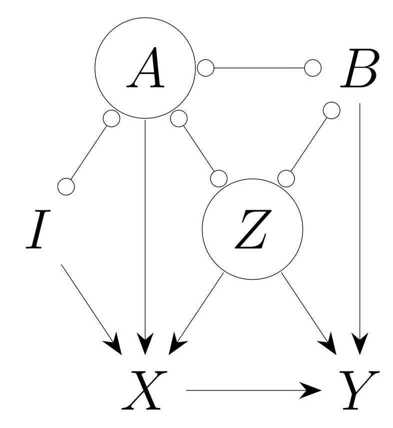
**Definition** (**Generalized Adjustment Criterion (GAC)**) Let $\mathcal{G}$ represent a DAG, MAG, CPDAG or PAG. Then $\mathbf{Z}$ satisfies the GAC relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ if

(0) $\mathcal{G}$ is adjustment amenable relative to $(\mathbf{X}, \mathbf{Y})$, and
(1) $\mathbf{Z} \cap \mathbf{F}_{\mathcal{G}}(\mathbf{X}, \mathbf{Y}) = \emptyset$, where $\mathbf{F}_{\mathcal{G}}(\mathbf{X}, \mathbf{Y})$ is the set of possible descendants of all $W \in \mathbf{V} \setminus \mathbf{X}$ that lie on a proper possibly causal path from $\mathbf{X}$ to $\mathbf{Y}$ in $\mathcal{G}$, and
(2) all proper definite status non-causal paths in $\mathcal{G}$ from $\mathbf{X}$ to $\mathbf{Y}$ are blocked by $\mathbf{Z}$.

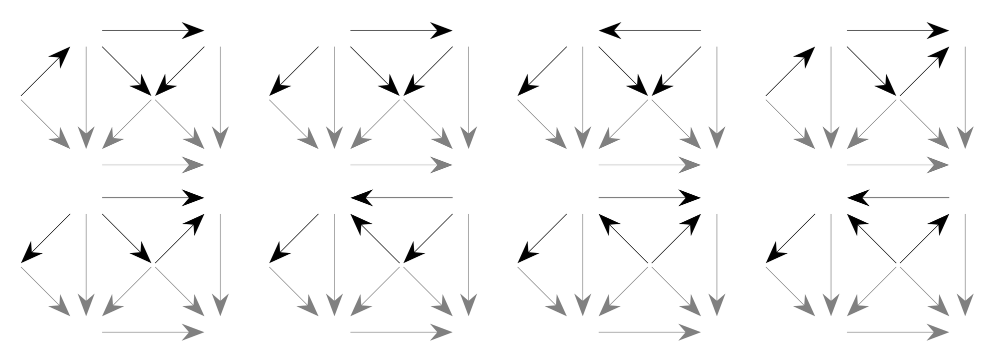
**Theorem** Let $\mathcal{G}$ represent a DAG, MAG, CPDAG or PAG. Then $\mathbf{Z}$ is an adjustment set relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$ if and only if $\mathbf{Z}$ satisfies the generalized adjustment criterion relative to $(\mathbf{X}, \mathbf{Y})$ in $\mathcal{G}$.

**Software** Function `gac` in R package `pcalg` [**?**].
Input: Graph $\mathcal{G}$, intervention variables $\mathbf{X}$, response variables $\mathbf{Y}$, covariate set $\mathbf{Z}$.
Output: Three booleans indicating whether $\mathbf{Z}$ satisfies the GAC conditions, and $\mathbf{F}_{\mathcal{G}}(\mathbf{X}, \mathbf{Y})$.

## Examples

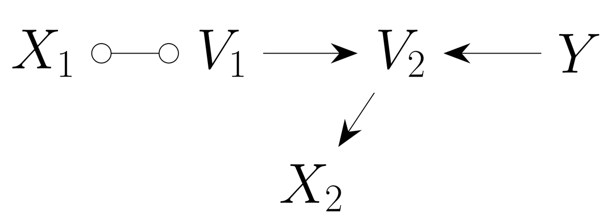



(a) A CPDAG.

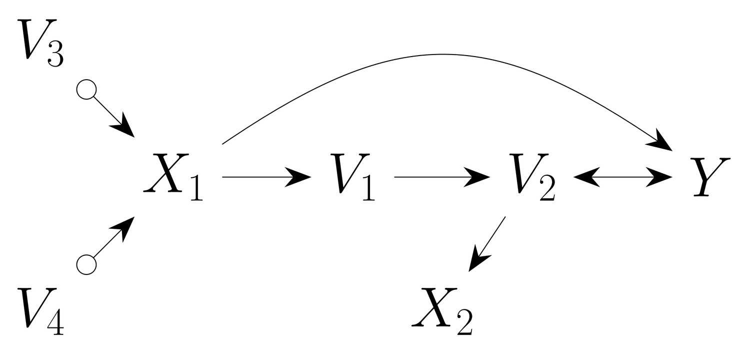


(b) DAGs in the Markov equivalence class of (a).

$\{A, Z\}$ satisfies the GAC relative to $(X, Y)$ in the CPDAG (a) and in every DAG in (b).



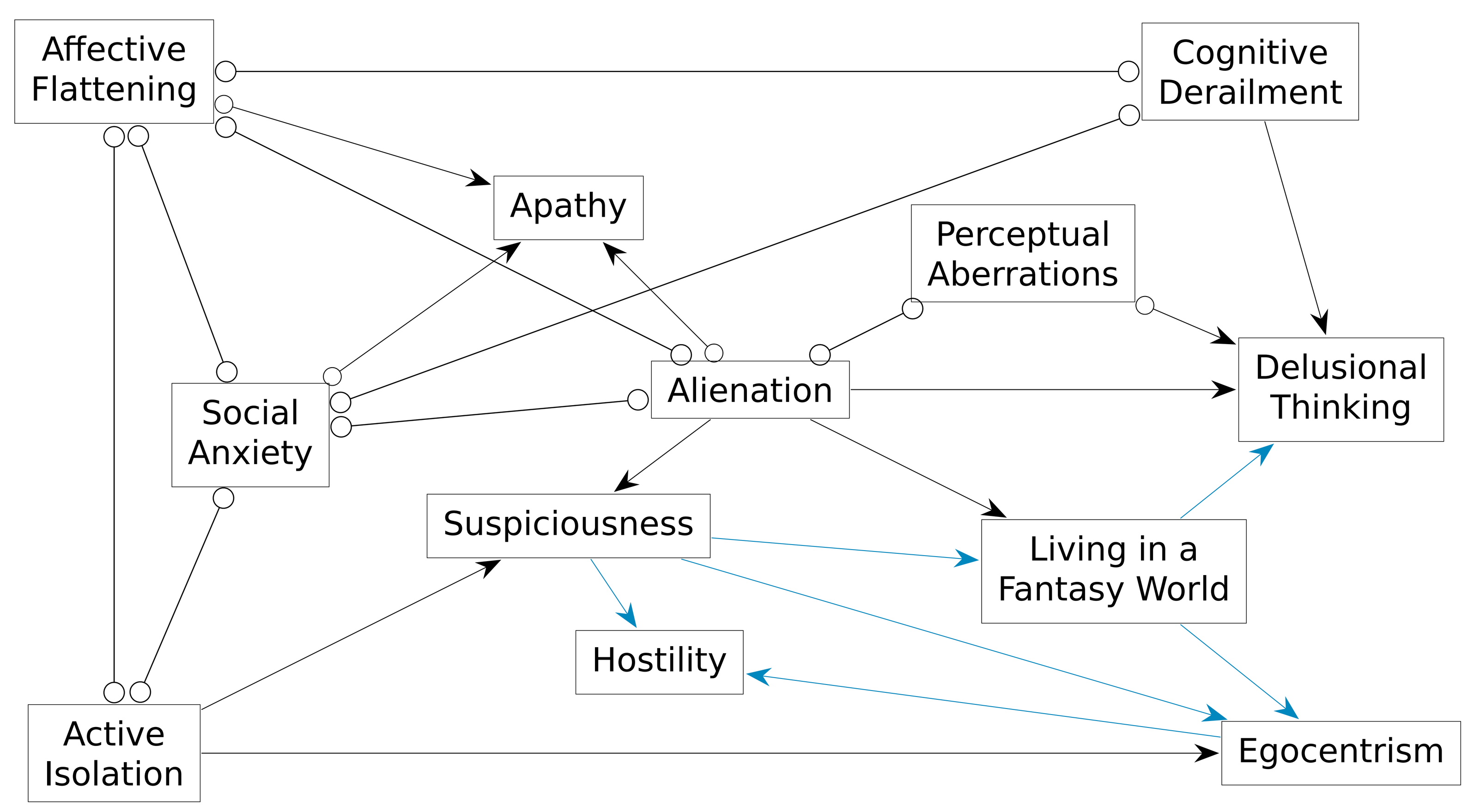

(a) A CPDAG.




(b) A PAG.

In both (a) and (b), $\{V_1, V_2\}$ satisfies the GAC relative to $((X_1, X_2), Y)$.
No set satisfies the generalized back-door criterion relative to $((X_1, X_2), Y)$ in (a) or (b).

## Application




A PAG of the truncated SSQ model of schizophrenic unfolding [**?**]. Edges in blue are visible.

- $X$ = Suspiciousness, $Y$ = Delusional Thinking.
  $\mathbf{F}_{\mathcal{G}}(\mathbf{X}, \mathbf{Y})$ = {Living in a Fantasy World, Egocentricism, Hostility, Delusional Thinking}.
  Some adjustment sets: {Alienation, Act. Isolation}, {Alienation, Cogn. Derailment}.
- $\mathbf{X}$ = {Suspiciousness, Living in a Fantasy World}, $Y$ = Hostility.
  $\mathbf{F}_{\mathcal{G}}(\mathbf{X}, \mathbf{Y})$ = {Egocentricism, Hostility}.
  Some adjustment sets: {Active Isolation}, {Active Isolation, Delusional Thinking}.
- $X$ = Alienation, $Y$ = Hostility.
  $\mathbf{F}_{\mathcal{G}}(\mathbf{X}, \mathbf{Y})$ includes all nodes except {Alienation, Perceptual Aberrations}.
  There is no set that fulfills the GAC and hence no adjustment set.

## Limitations

- We only consider causal effects that are identifiable through covariate adjustment.
- We do not allow for cycles nor for selection variables.

## Future work

- Studying the relation between the GAC and the generalized back-door criterion.
- Developing algorithms to quickly determine if there is an adjustment set.
- Developing algorithms to find all minimal adjustment sets.

## References

[1] M. Kalisch, M. Mächler, D. Colombo, M. Maathuis, and P. Bühlmann. Causal inference using graphical models with the R package `pcalg`. *J. Stat. Softw.*, 47:1–26, 2012.
[2] M. H. Maathuis and D. Colombo. A generalized back-door criterion. *Ann. Stat.*, 43:1060–1088, 2015.
[3] J. Pearl. Comment: Graphical models, causality and intervention. *Stat. Sci.*, 8:266–269, 1993.
[4] I. Shpitser, T. VanderWeele, and J. M. Robins. On the validity of covariate adjustment for estimating causal effects. In *Proceedings of UAI 2010*, 2010.
[5] B. van der Zander, M. Liśkiewicz, and J. Textor. Constructing separators and adjustment sets in ancestral graphs. In *Proceedings of UAI 2014*, 2014.
[6] D. van Kampen. The SSQ model of schizophrenic prodromal unfolding revised: An analysis of its causal chains based on the language of directed graphs. *Eur. Psychiat.*, 2013.