Algebraic geometry with a view towards applications

Dr. Eliana Duarte

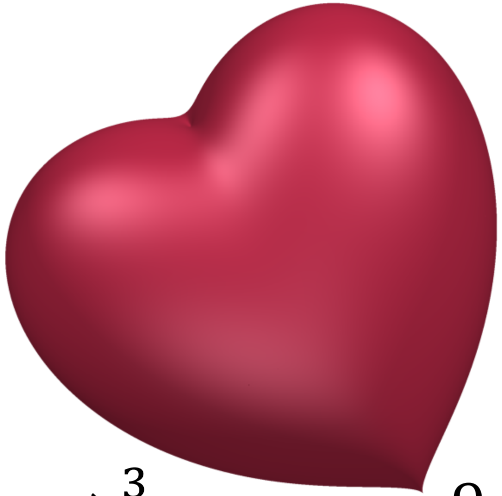
https://emduart2.github.io/

Max-Planck-Institut für Mathematik in den Naturwissenschaften

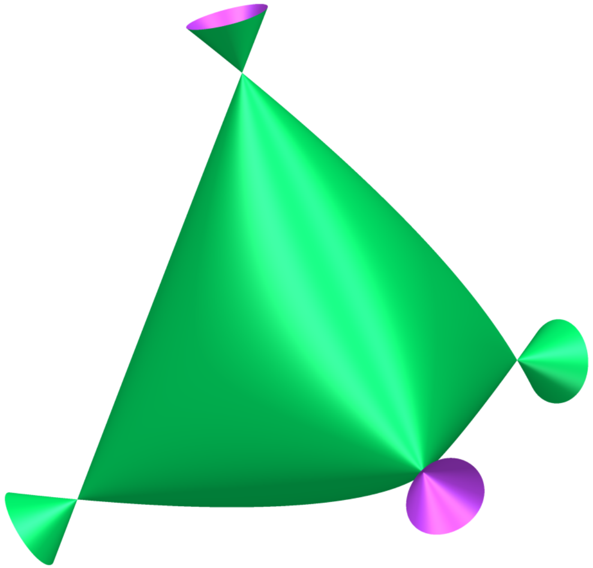
OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG

DFG-Graduiertenkolleg MATHEMATISCHE KOMPLEXITÄTSREDUKTION

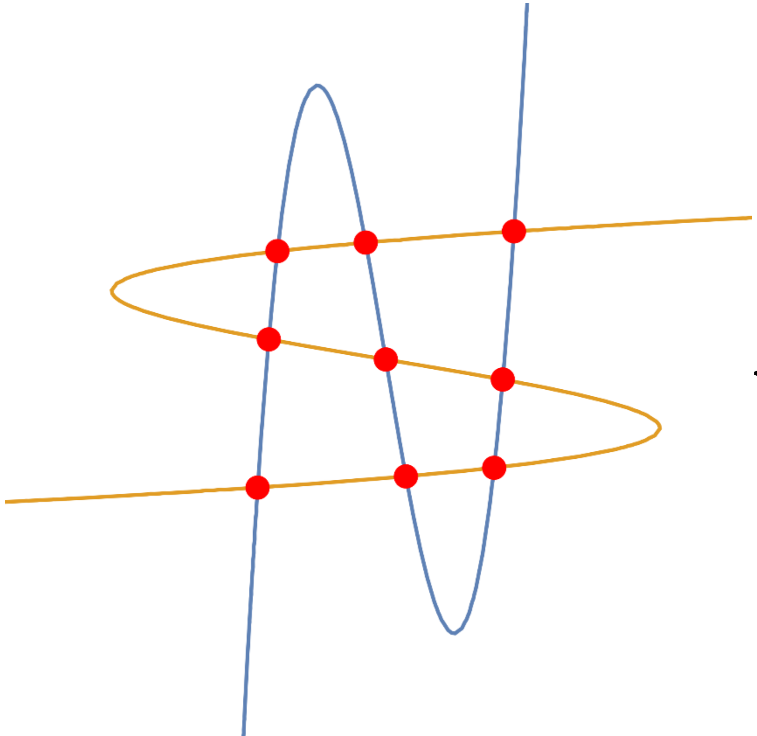
Algebraic varieties



$$\left(x^2 + \frac{9}{4y^2} + z^2 - 1\right)^3 - x^2 z^3 - \frac{9}{80} y^2 z^3 = 0$$



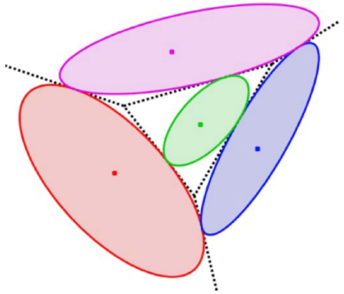
$$1 - x^2 - y^2 + 2\,x\,y\,z - z^2 = 0$$



$$\begin{cases} y - x^3 + 6\,x &= 0 \\ x - y^3 + 6y &= 0 \end{cases}$$

Algebraic varieties in applications

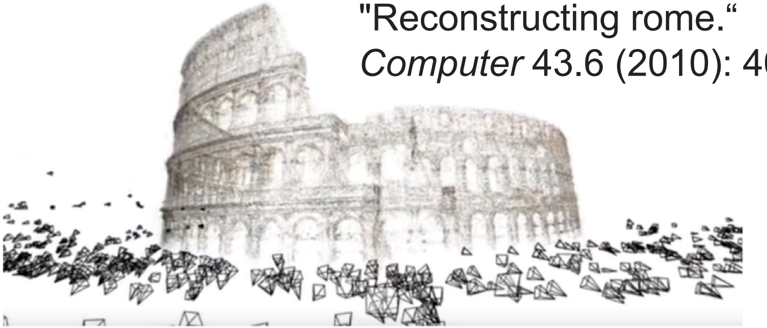
Polynomial optimization



Cifuentes, Harris, and Sturmfels. "The geometry of SDP-exactness in quadratic optimization.“ *Mathematical Programming* (2018): 1-30.

Computer vision

Agarwal, Sameer, et al. "Reconstructing rome.“ *Computer* 43.6 (2010): 40-47.



Does the system have solutions?

What is the dimension of the solution set?

Can we write them explicitly?

Algebraic varieties in applications

Maple™ Macaulay2

Homotopy Continuation.jl

Computational Algebraic Geometry

CoCoA SINGULAR

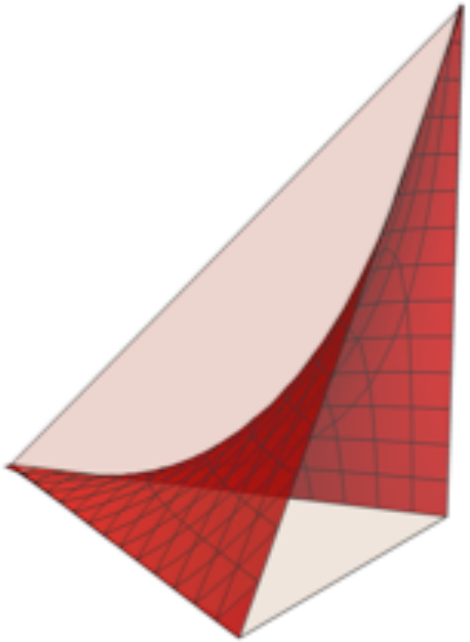
WOLFRAM MATHEMATICA 11

Does the system have solutions?

What is the dimension of the solution set?

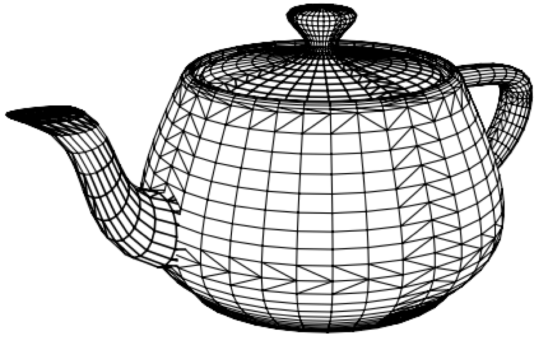
Can we write them explicitly?

My Research Profile



Algebraic Statistics

Ananiadi and **Duarte** (J. Algebr. Stat.: *in Rev.*) 2020
Duarte, Marigliano and Sturmfels (Bernoulli: *in Rev.*) 2019
Duarte and Görgen (J. Symb. Comput.) 2019
Guerra, Delgado-Baquerizo, **Duarte**, et al. (*in Rev.*)



Geometric Modeling

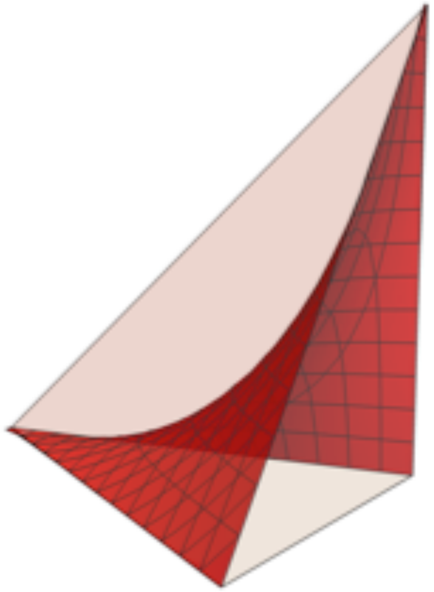
Duarte and Seceleanu (Math. of Comp.: *accepted*) 2020
Duarte (J. Algebra its Appl.) 2016
Duarte and Schenck (Proc. Am. Math. Soc.) 2014

Duarte and Francis (Conf. Proc. Transformables) 2013

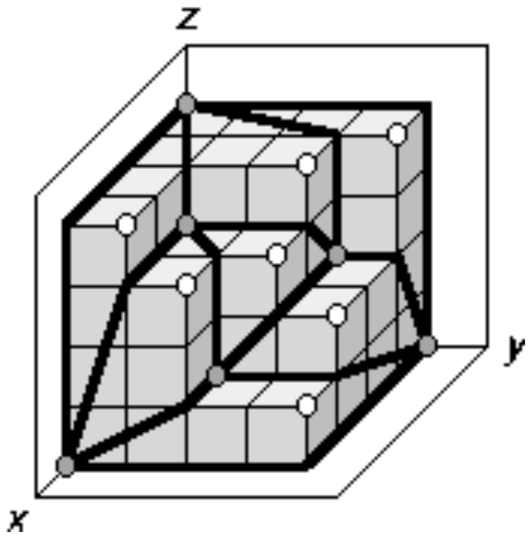
Rigidity theory

Algebraic Statistics

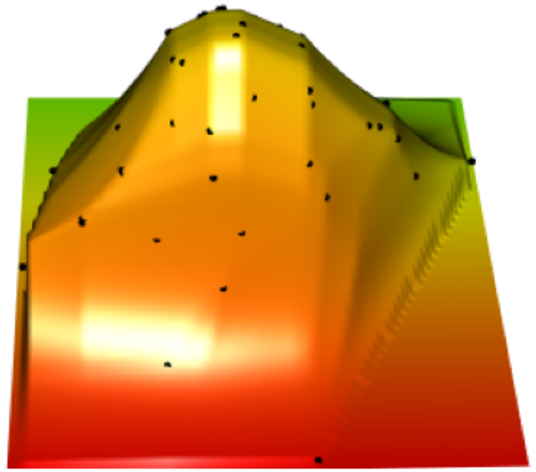
Branch of mathematical statistics employing and developing tools from…



Algebraic Geometry



Commutative Algebra



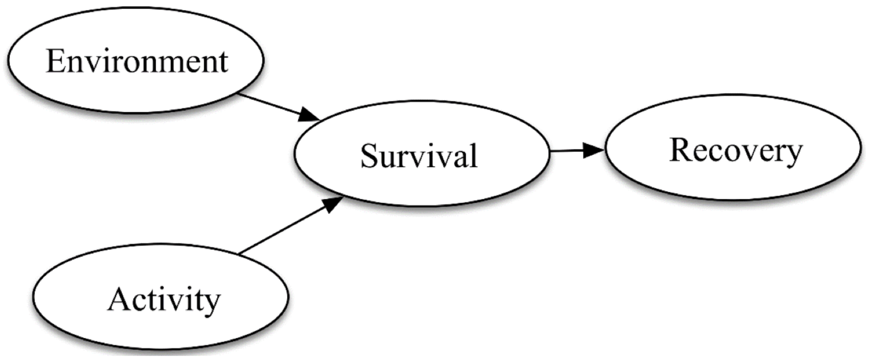
Discrete Geometry

…to tackle problems in statistics

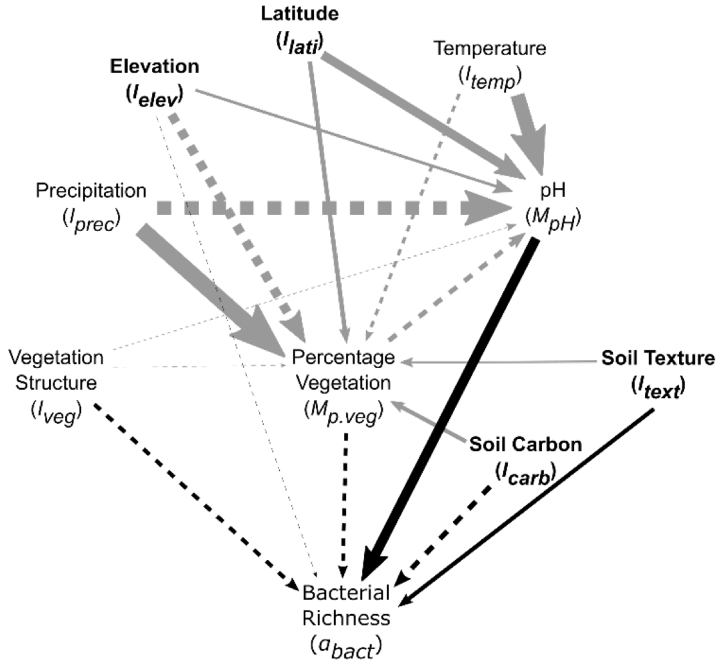
Graphical Models

family of statistical models that describes dependency relations among variables using a graph

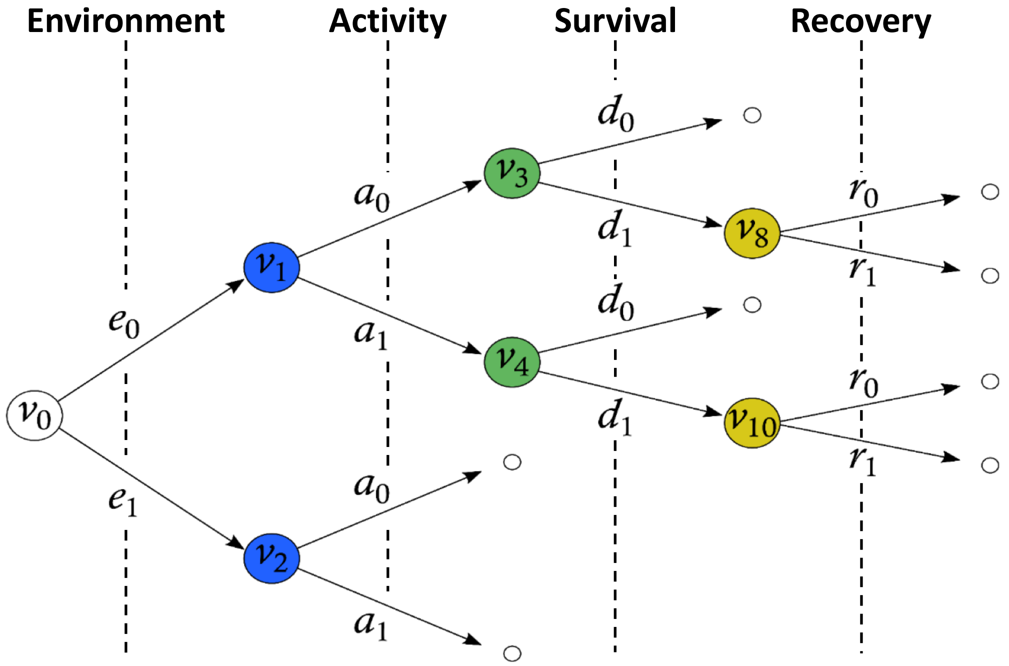
Bayesian networks



Structural equation models

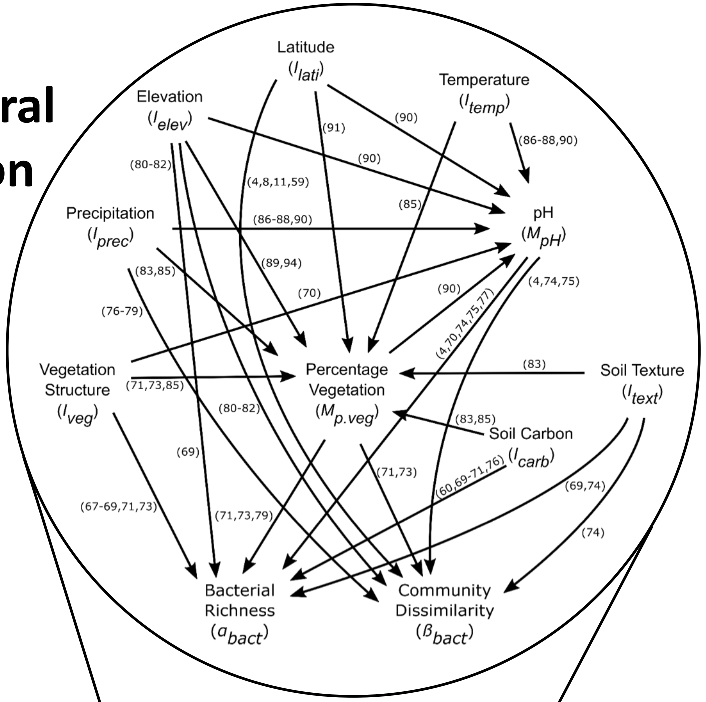


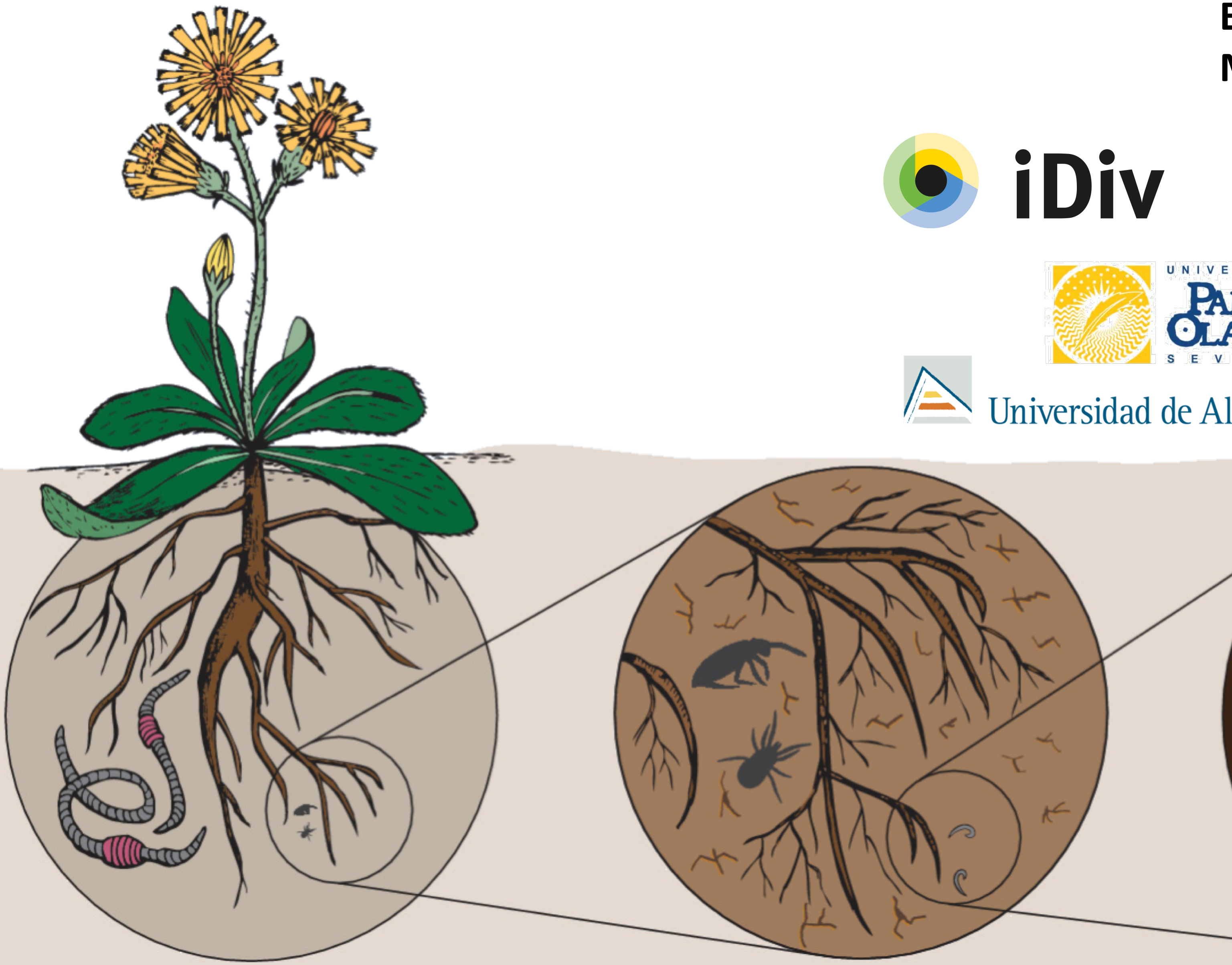
Staged tree models



The soil microbiome

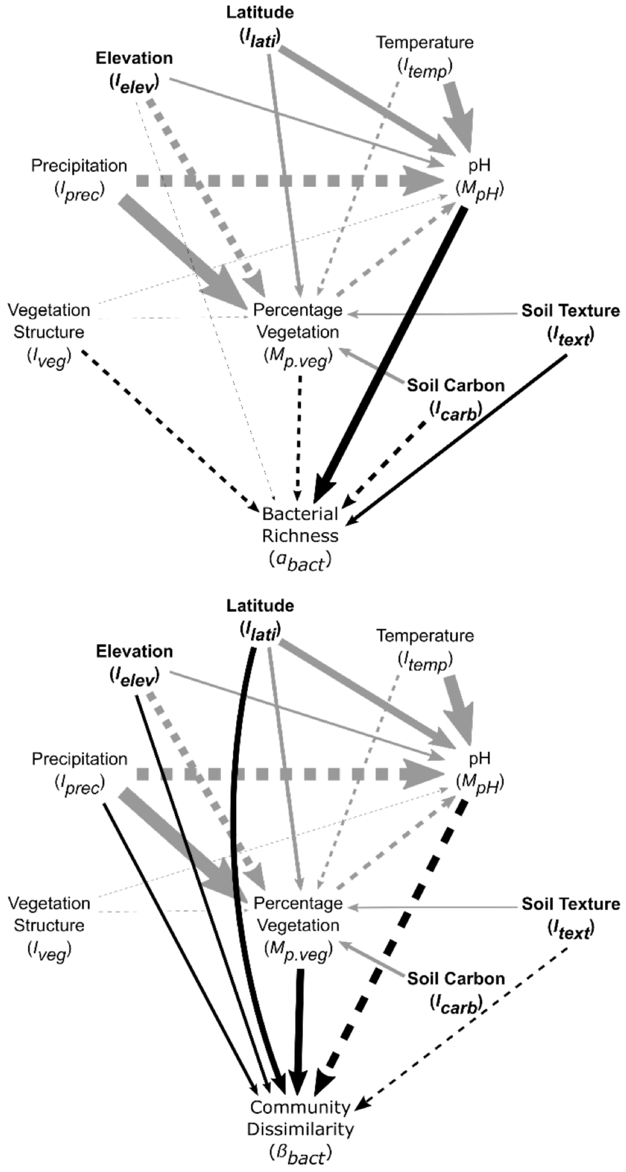
Structural Equation Model

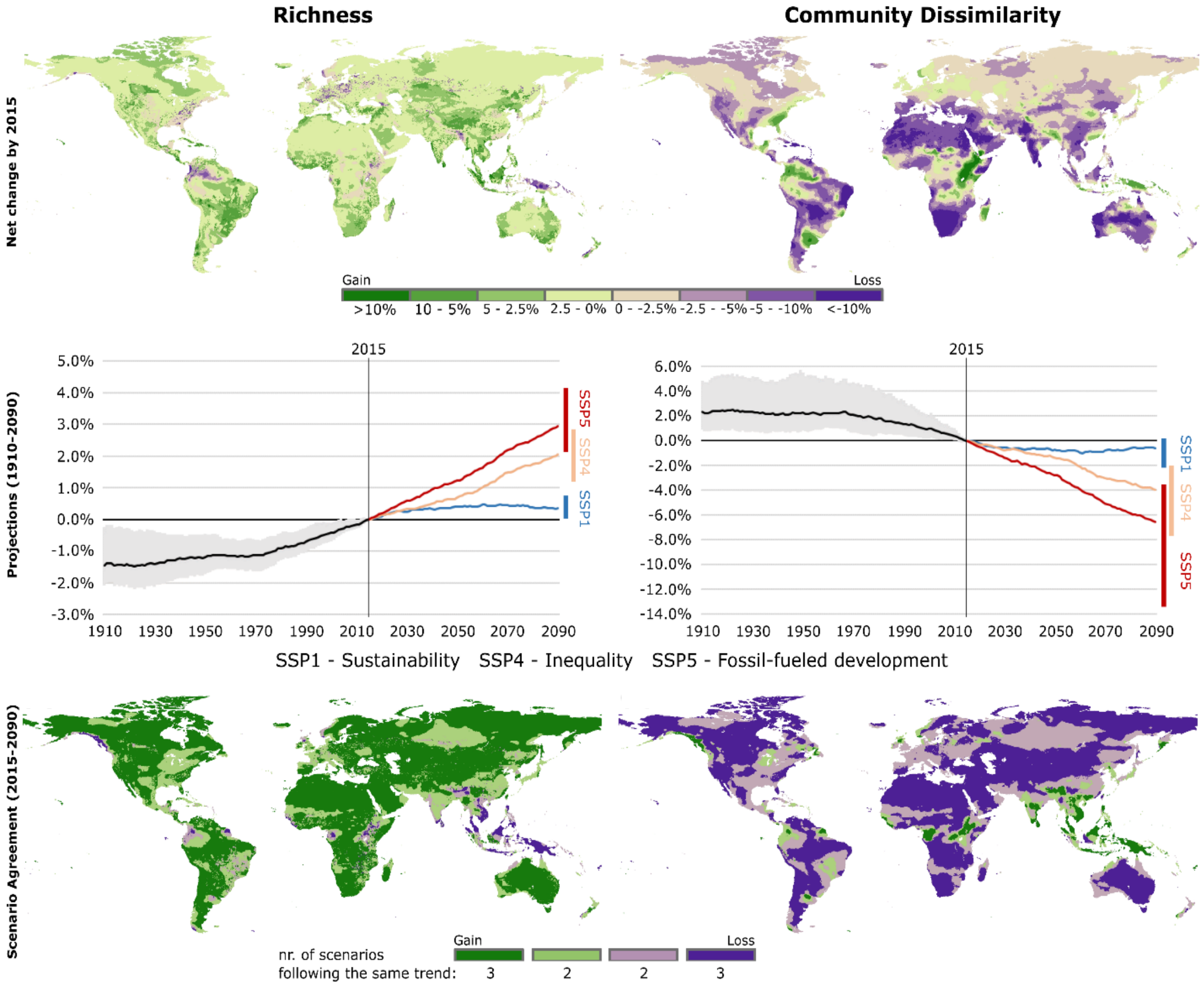




Thakur, et al., (Biol. Cons.) 2019

The soil microbiome



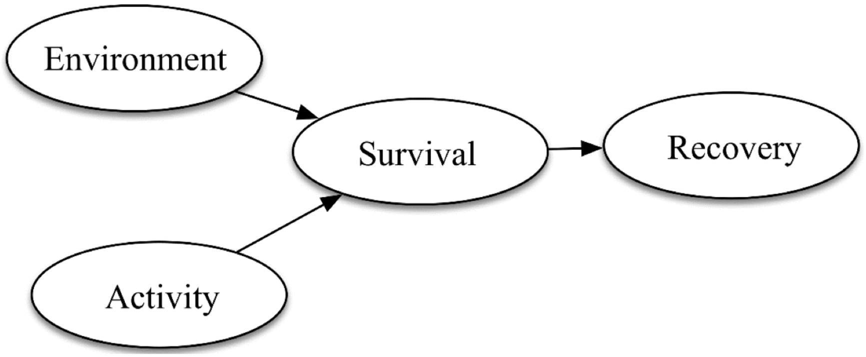


Guerra, Delgado-Baquerizo, **Duarte**, et al. (*in Rev.*)

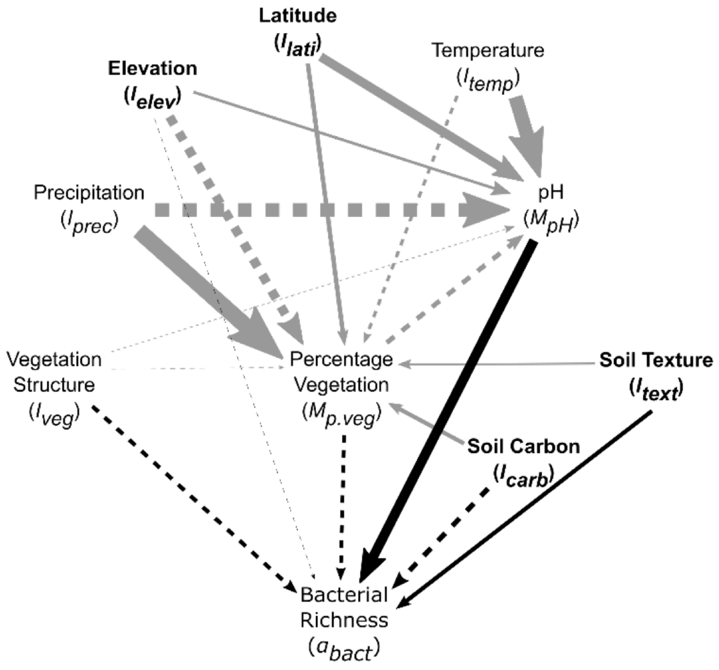
Graphical Models

family of statistical models that describes dependency relations among variables using a graph

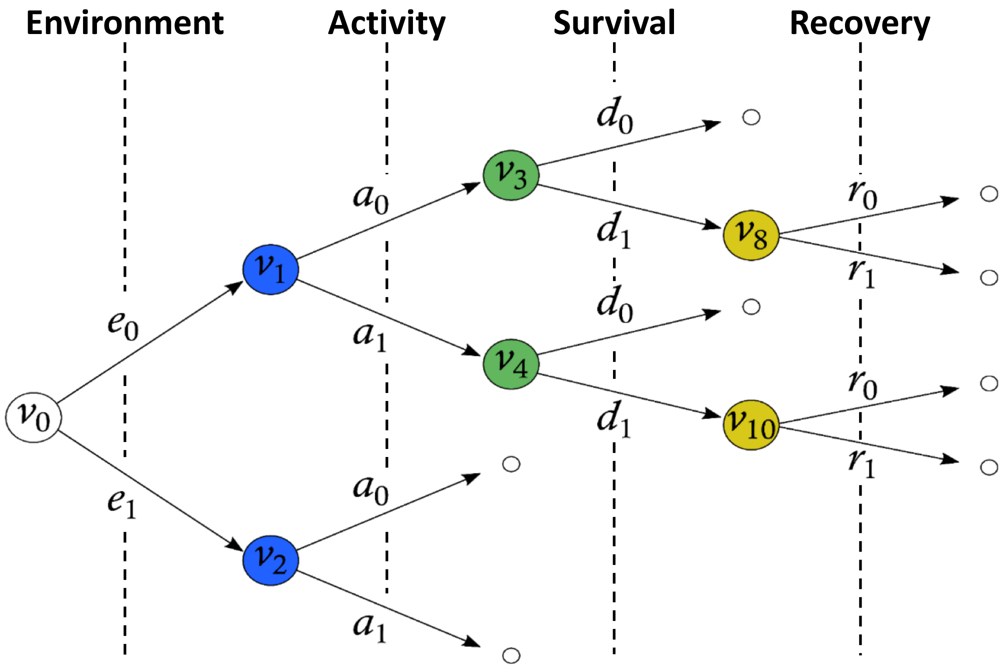
Bayesian networks



Structural equation models

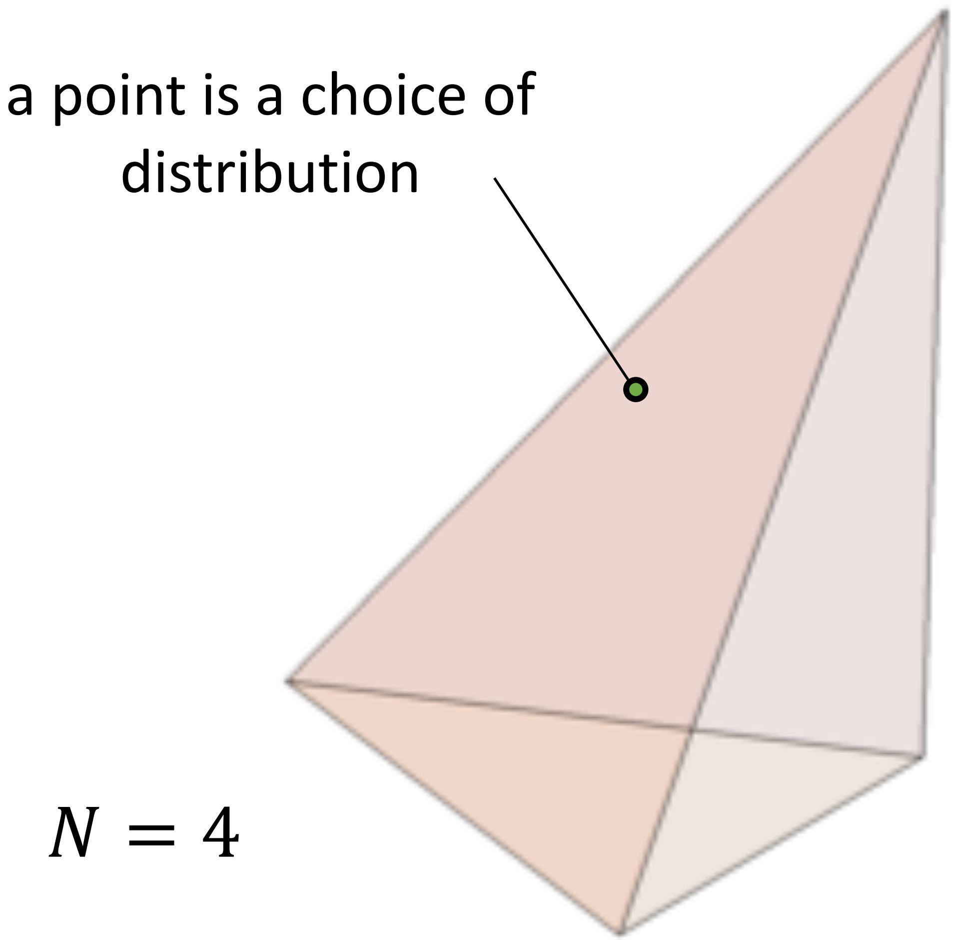


Staged tree models



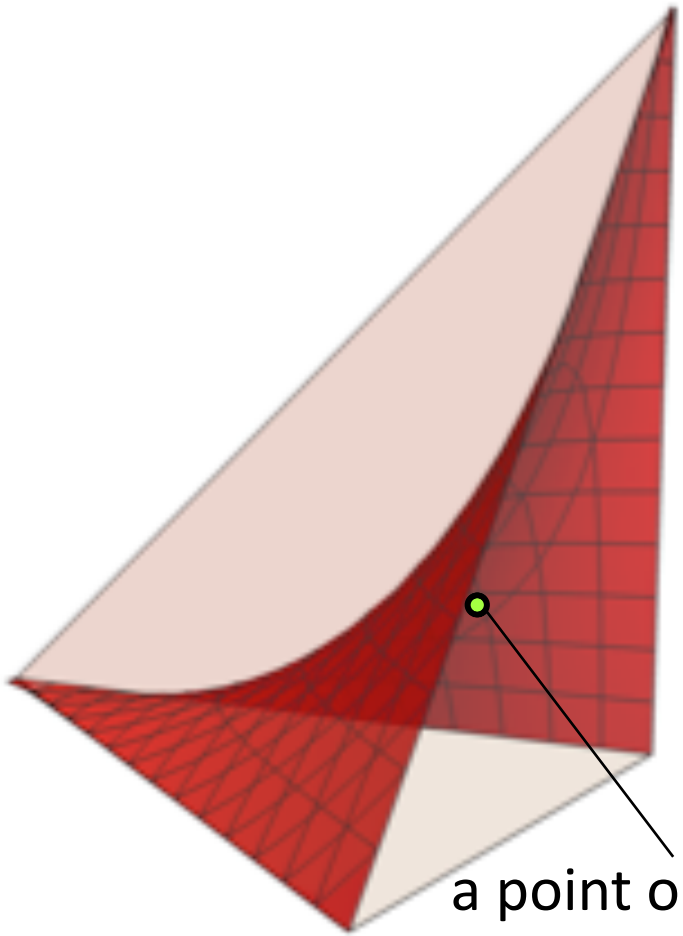
Discrete statistical models

$$\Delta_{N-1} = \{(p_1, \ldots, p_N) \colon p_i \geq 0, p_1 + \cdots + p_N = 1\}$$



A discrete statistical model with N outcomes is a family of distributions inside the N-1 dimensional probability simplex

The independence model



a point on the surface is a distribution in the model

X and Y are independent binary random variables

$$P_{00} + P_{01} + P_{10} + P_{11} = 1$$

$$P_{ij} \geq 0$$

$$P_{00}P_{11} - P_{01}P_{10} = 0$$

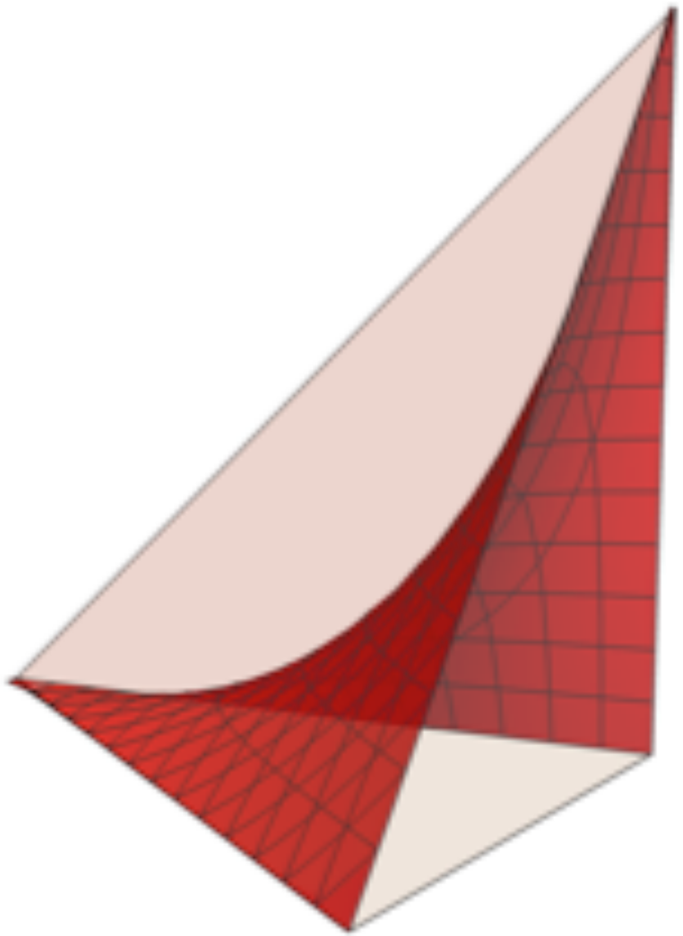
$$\Delta_1 \times \Delta_1 \longrightarrow \Delta_3$$

$$(s_0, s_1) \times (t_0, t_1) \longmapsto (s_0 t_0, s_0 t_1, s_1 t_0, s_1 t_1)$$

$$s_0 + s_1 = 1$$

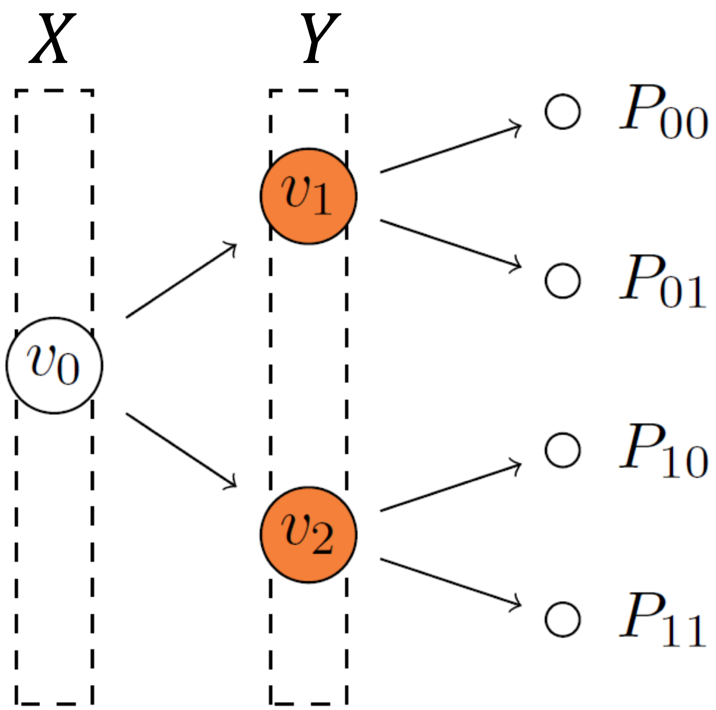
$$t_0 + t_1 = 1$$

The independence model is a graphical model



Bayesian network

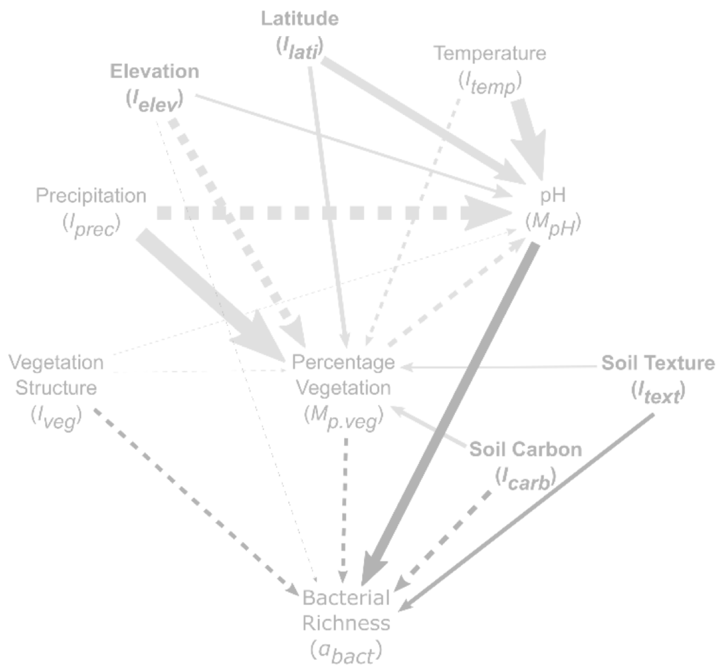
Staged tree Model



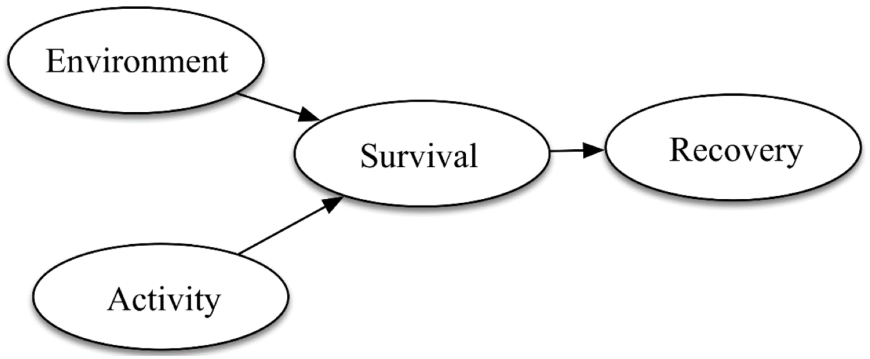
Graphical Models

family of statistical models that describes dependency relations among variables using a graph

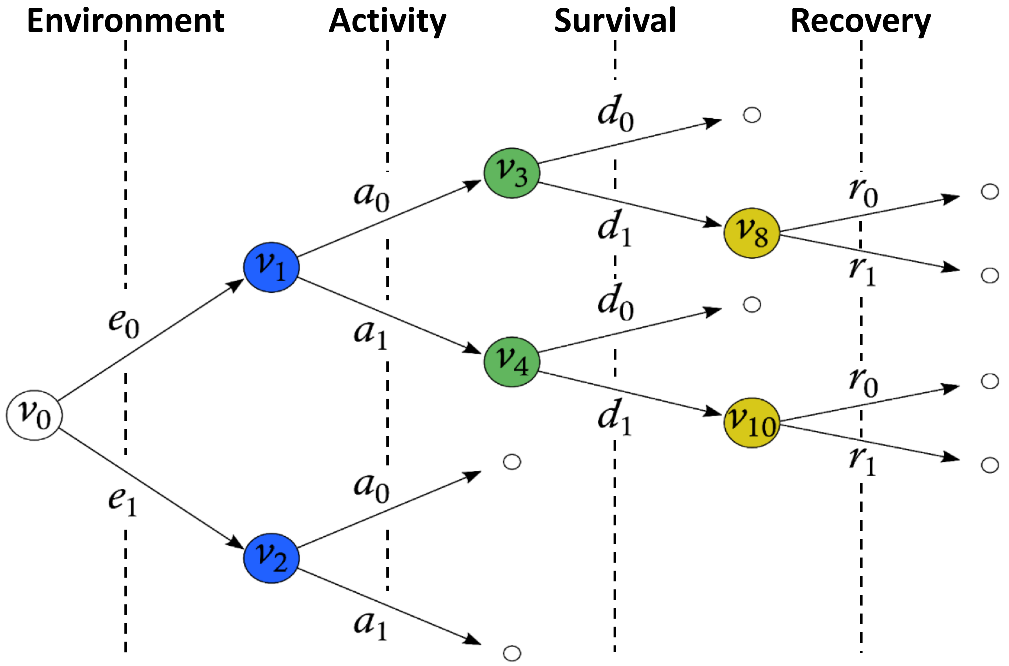
Structural equation models



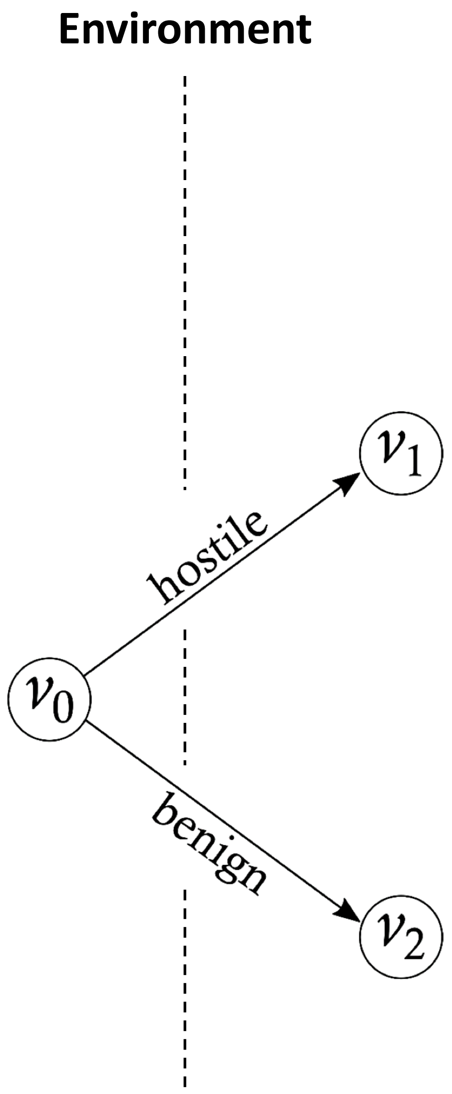
Bayesian networks



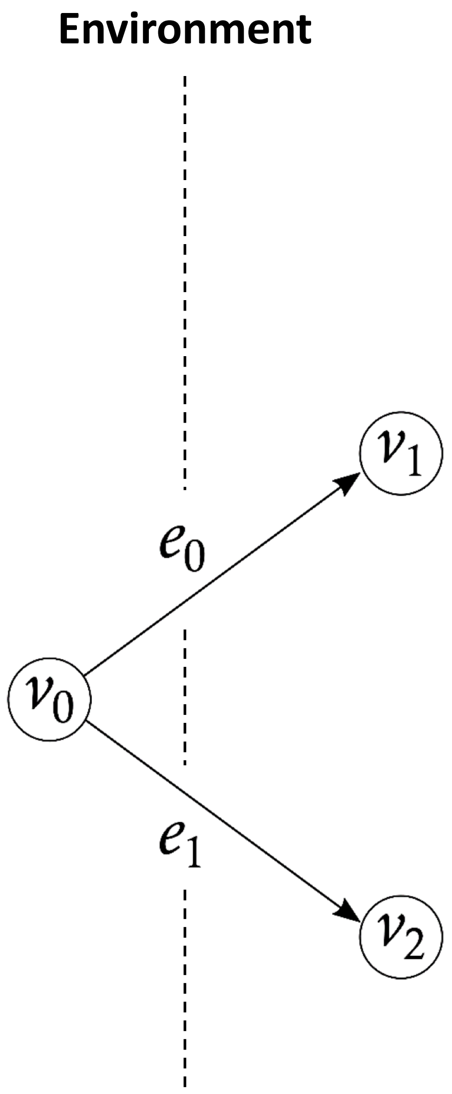
Staged tree models



Unfolding of events in a cell culture

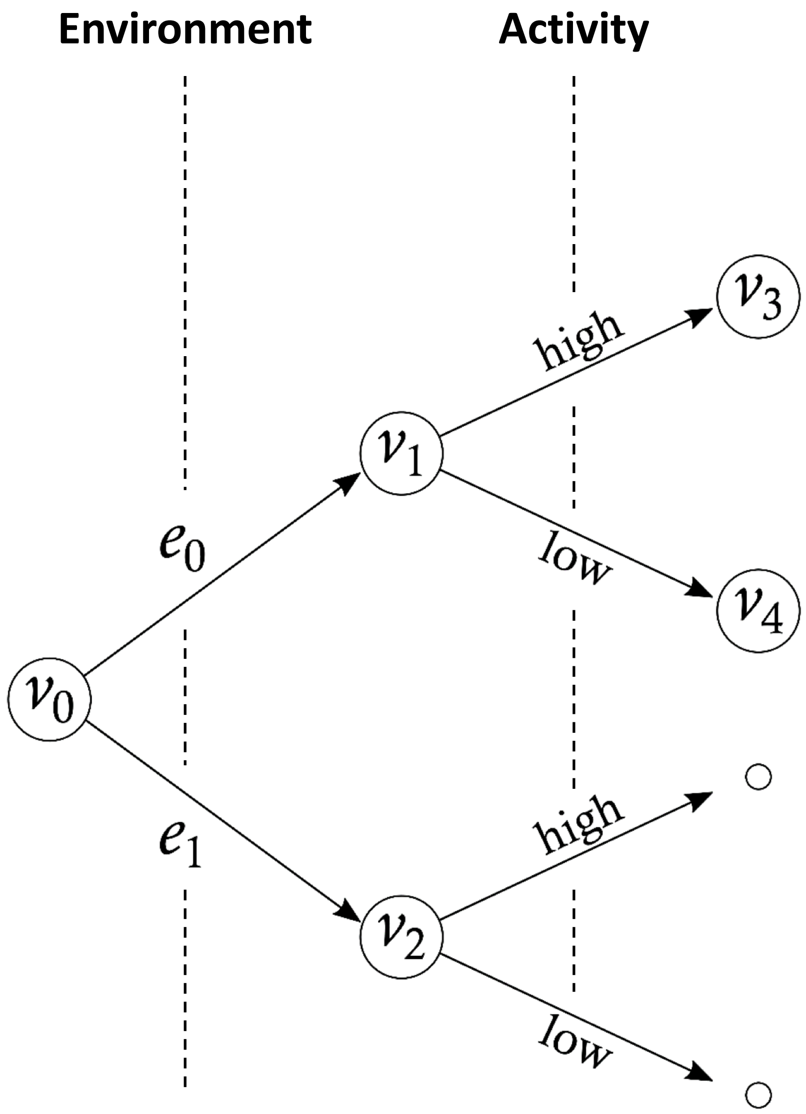


Unfolding of events in a cell culture

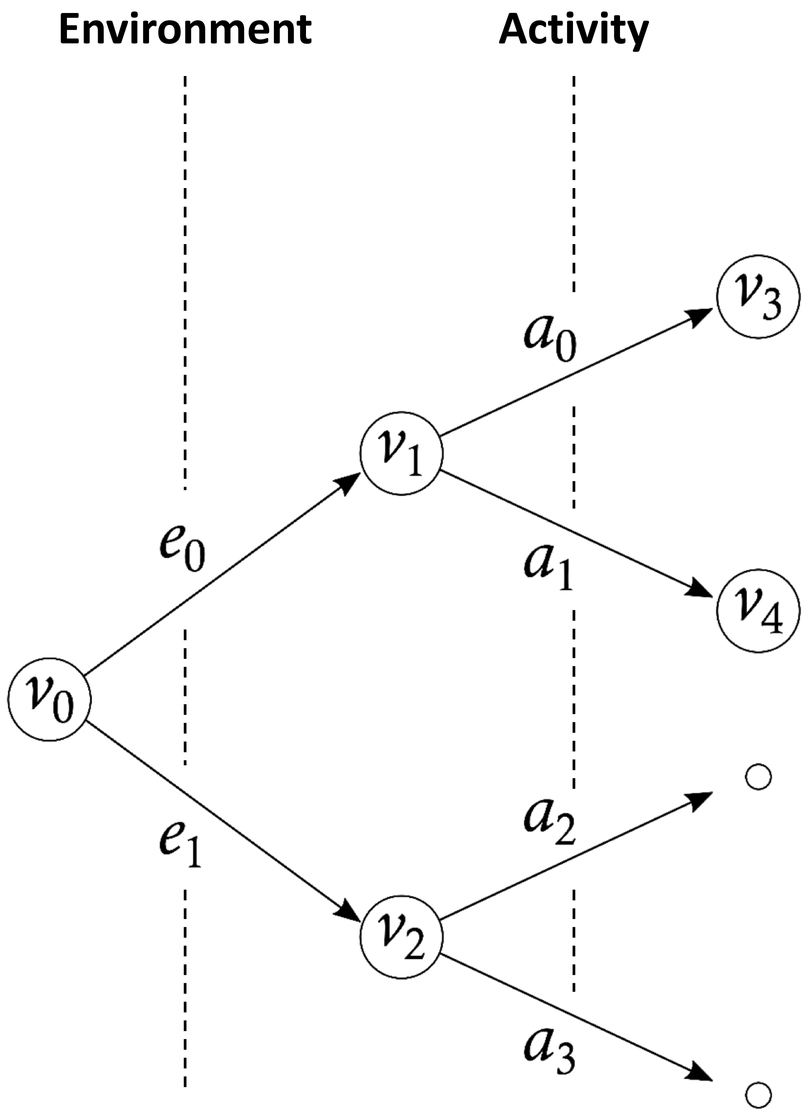


$$e_0 + e_1 = 1$$

Unfolding of events in a cell culture

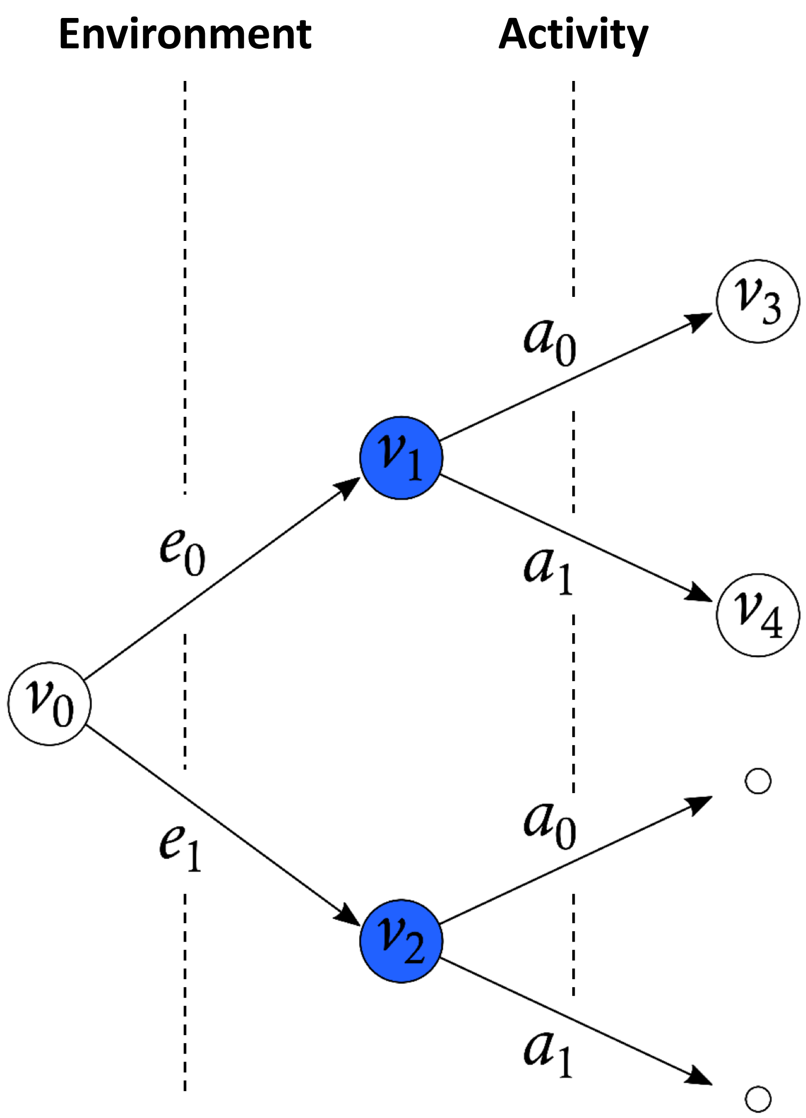


Unfolding of events in a cell culture



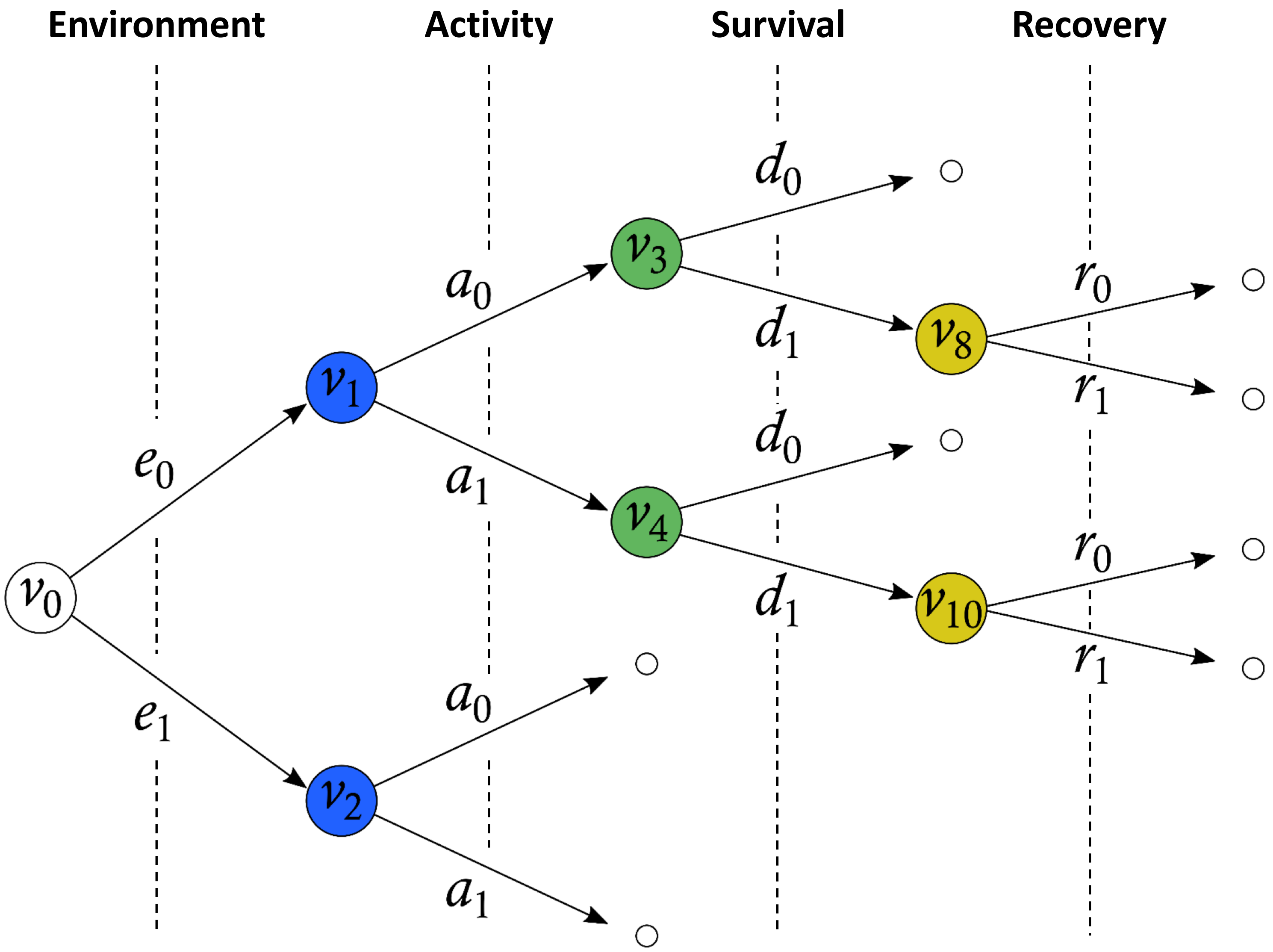
$$a_0 + a_1 = a_2 + a_3 = 1$$

Unfolding of events in a cell culture



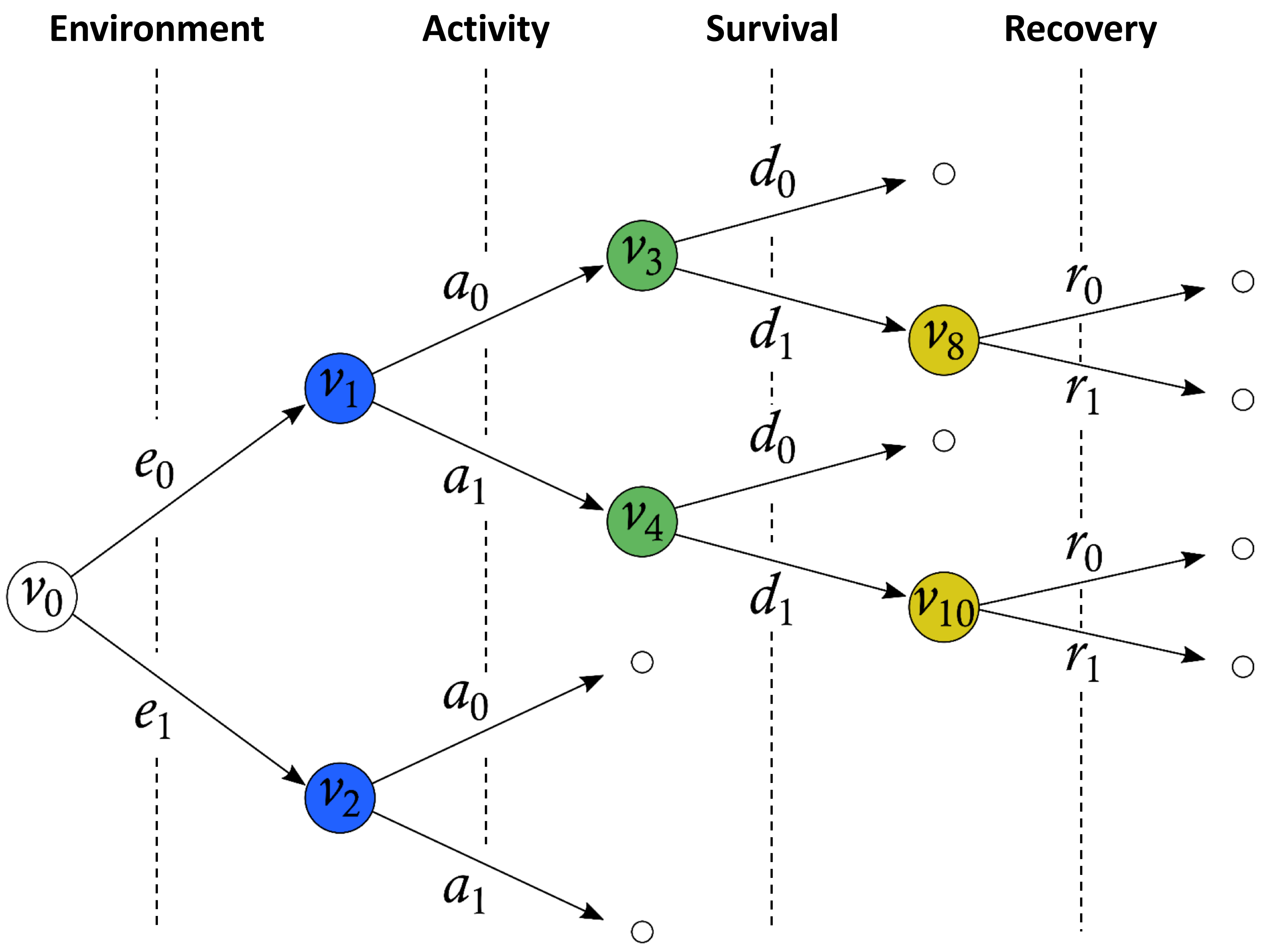
Activity is independent of the environment

Unfolding of events in a cell culture



- Activity is independent of the environment
- In a hostile environment a cell gets damaged and might die or survive. Survival does not depend on activity
- After survival, recovery is not affected by history

Unfolding of events in a cell culture



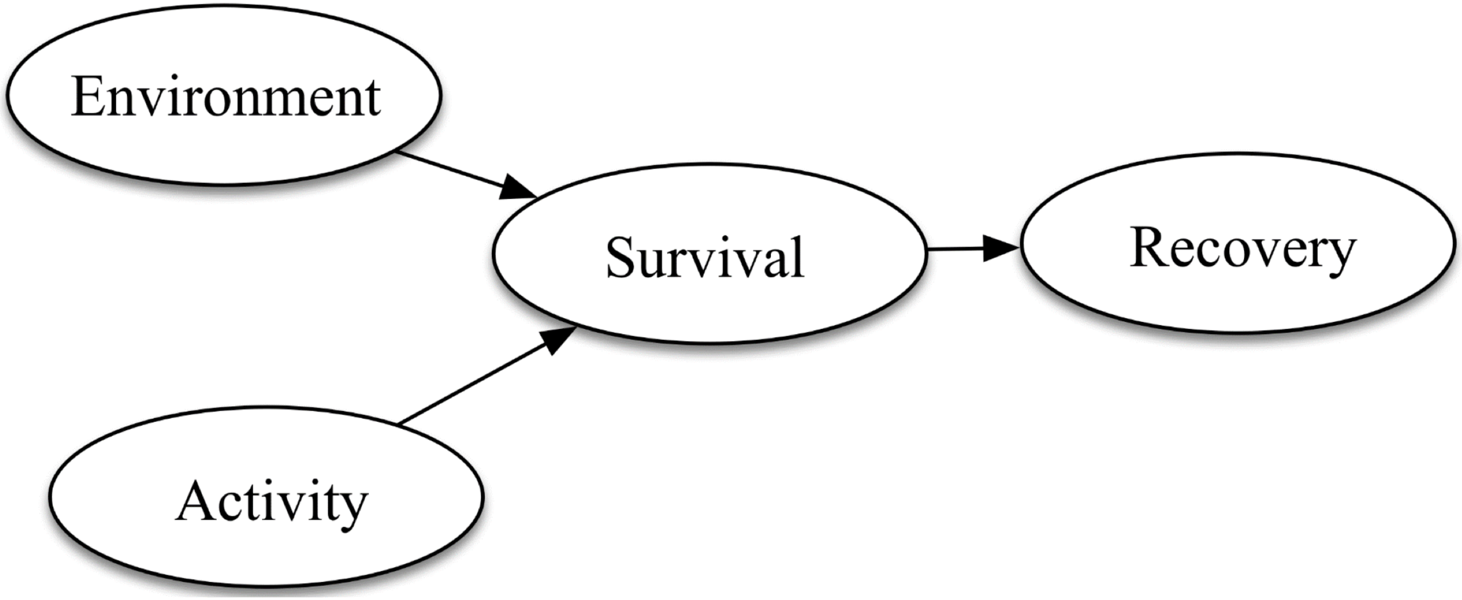
The staged tree model $\mathcal{M}_{\mathcal{T}}$ is the image of

$$\psi \ : [0,1]^4 \longrightarrow \Delta_7$$

$$(e_0, a_0, d_0, r_0) \ \longmapsto (e_0 a_0 d_0, \ldots)$$

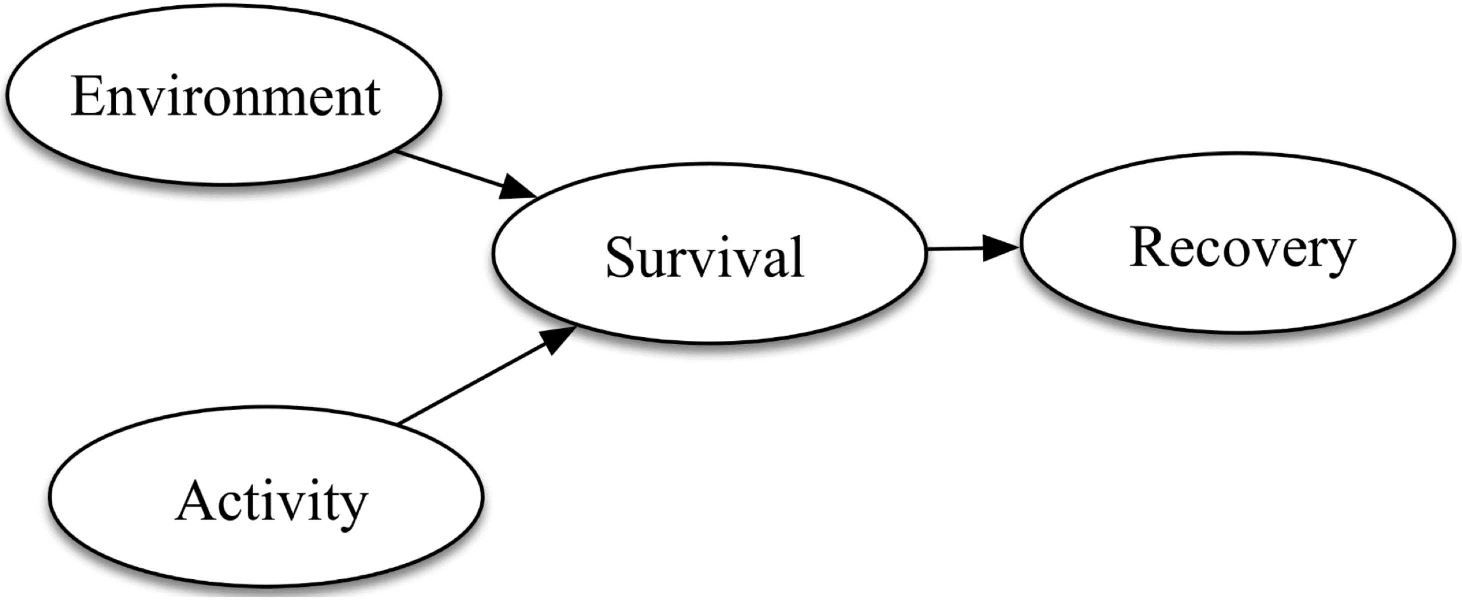
$$\begin{cases} P_0 + \cdots + P_7 - 1, \\ P_5 P_6 - P_2 P_7, \\ P_4 P_6 - P_1 P_7, \\ P_3 P_6 - P_0 P_7, \\ P_2 P_4 - P_1 P_5, \\ P_2 P_3 - P_0 P_5, \\ P_1 P_3 - P_0 P_4 \end{cases}$$

Bayesian network approach



- Activity is independent of the environment
- In a hostile environment a cell gets damaged and might die or survive. Survival does not depend on activity
- After survival, recovery is not affected by history

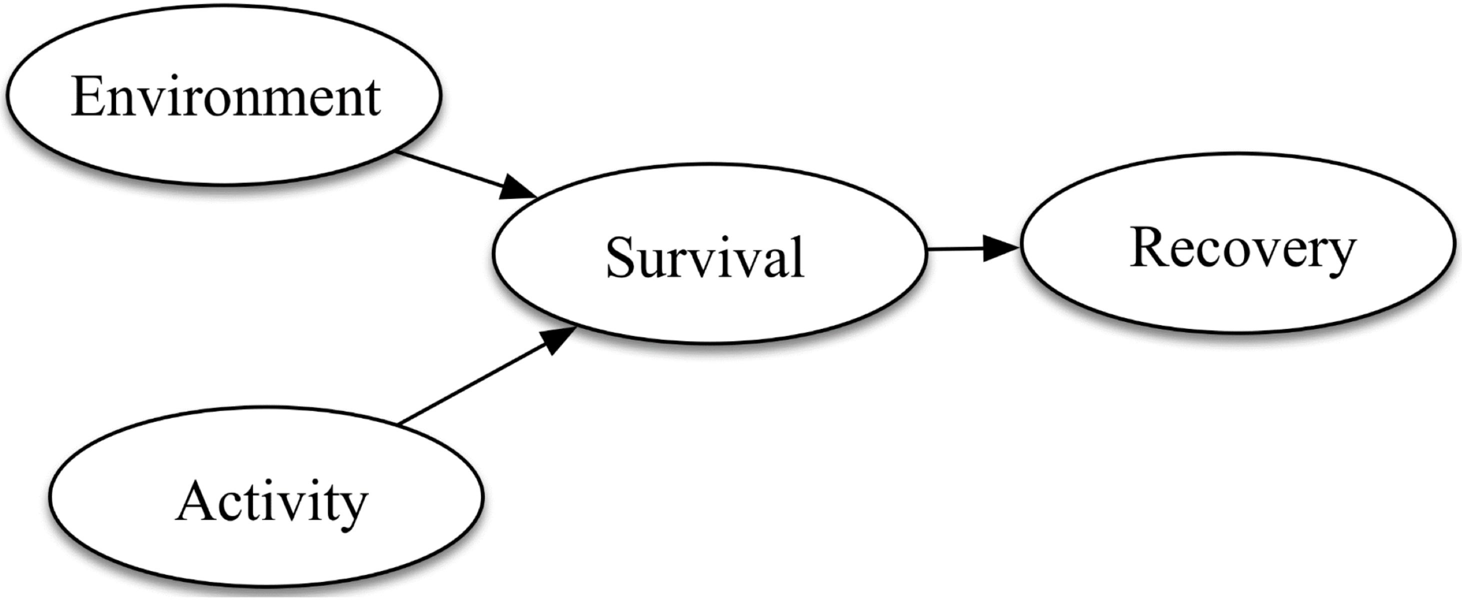
Bayesian network approach



- Activity is independent of the environment
- In a hostile environment a cell gets damaged and might die or survive. Survival does not depend on activity
- After survival, recovery is not affected by history

- Environment $\perp\!\!\!\perp$ Activity
  - Recovery $\perp\!\!\!\perp$ (Activity , Environment) | Survival

Bayesian network approach

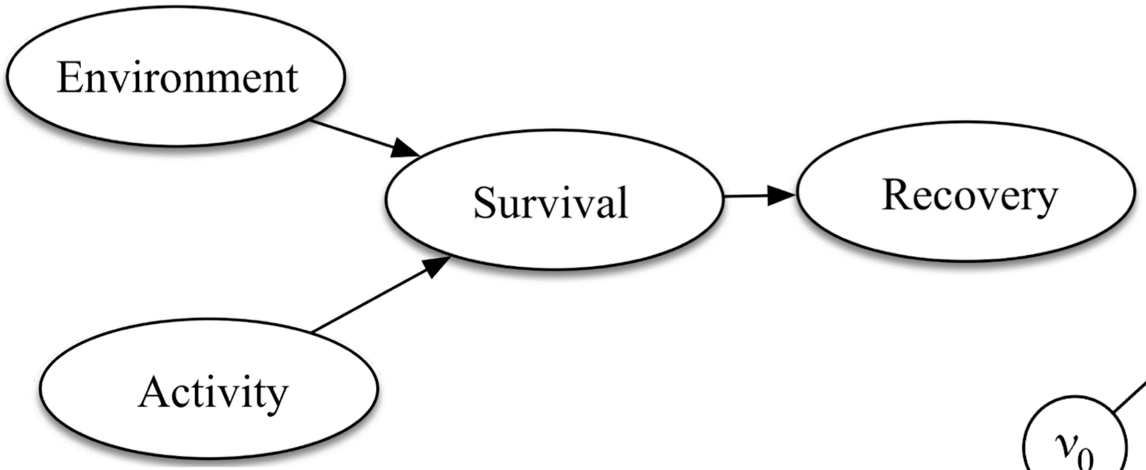


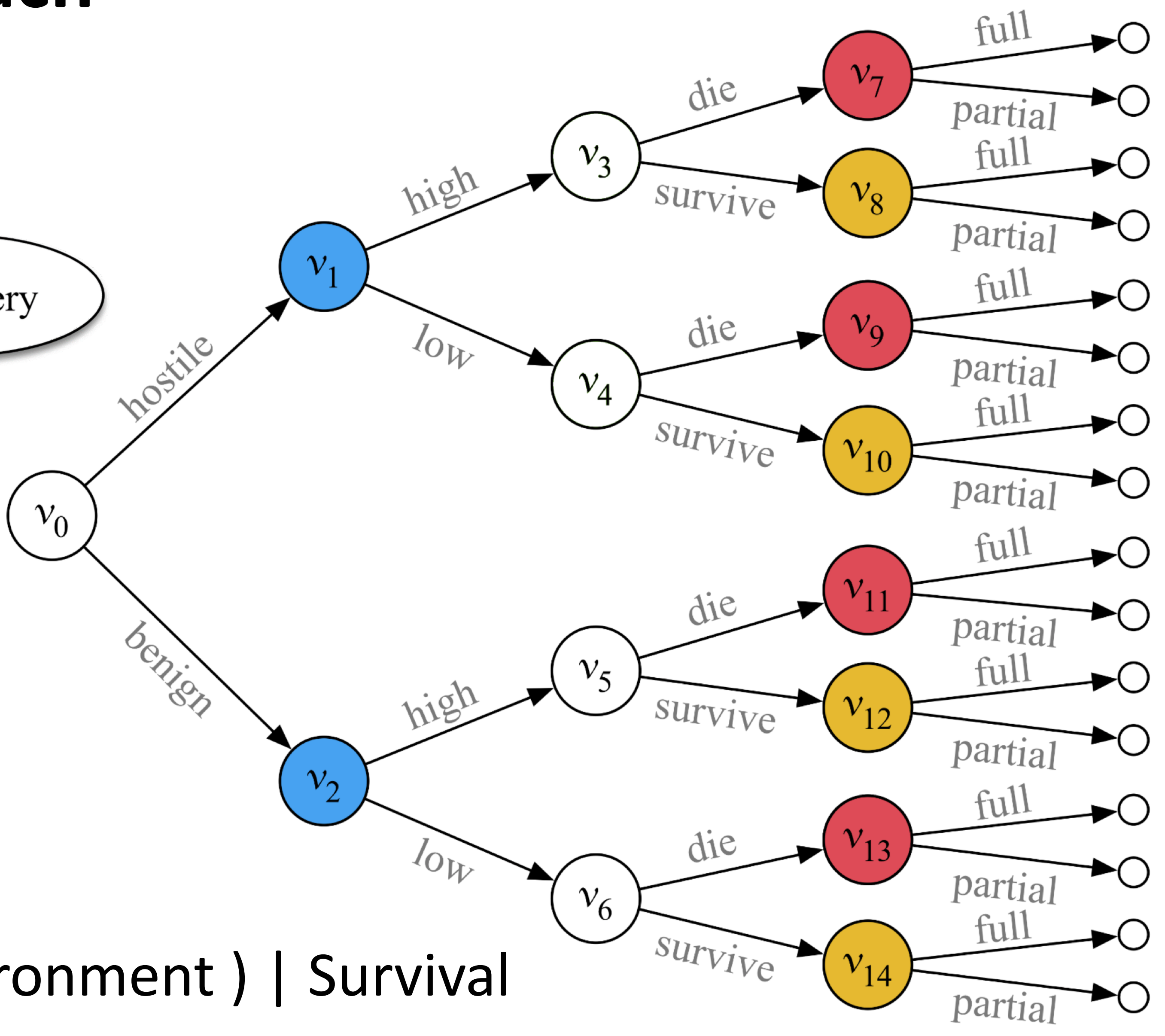
- Activity is independent of the environment
- ~~In a hostile environment a cell gets damaged and might die or survive. Survival does not depend on activity~~
- After survival, recovery is not affected by history

Environment $\perp\!\!\!\perp$ Activity

Recovery $\perp\!\!\!\perp$ (Activity , Environment) | Survival

Bayesian network approach

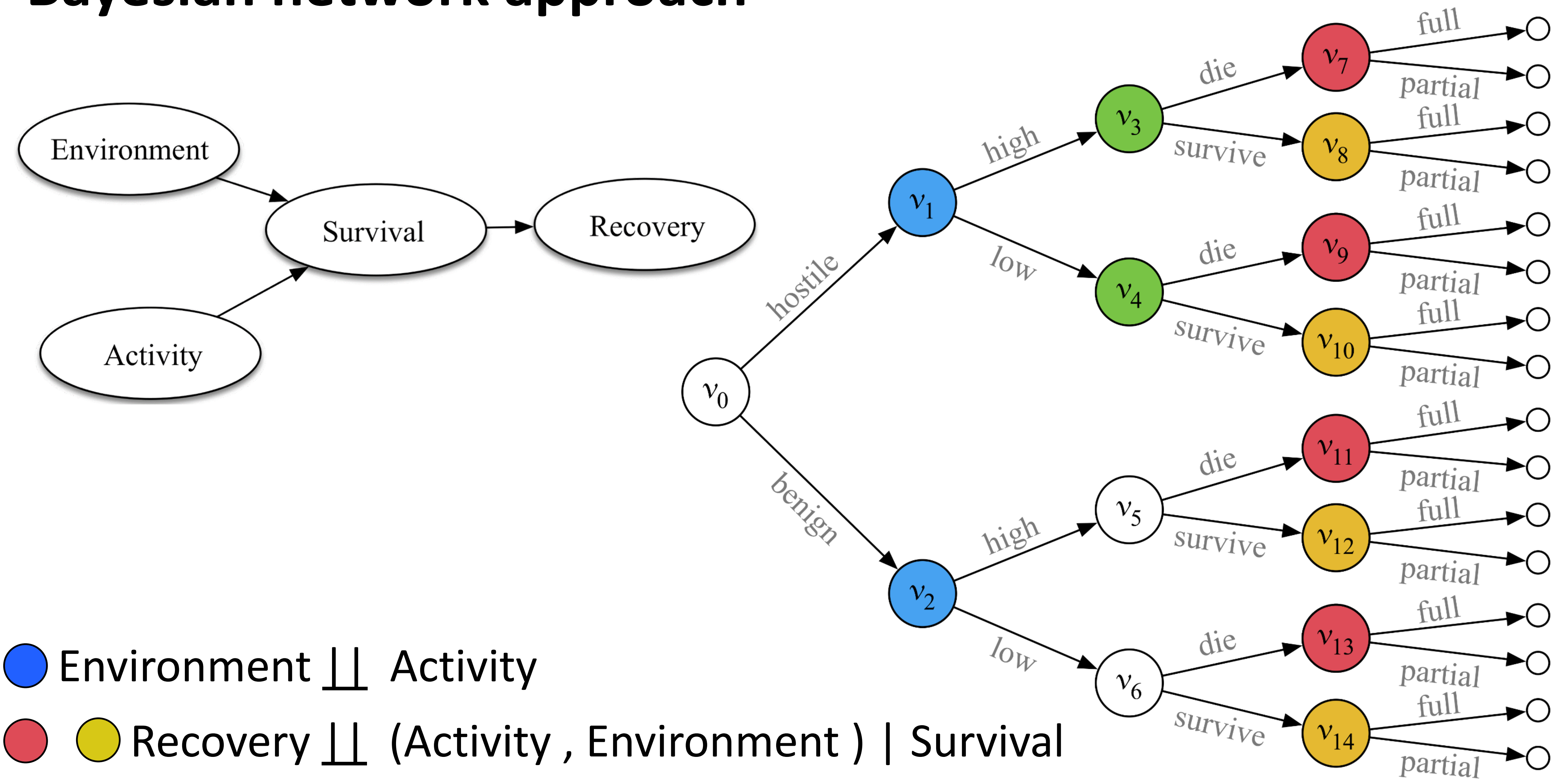




Environment $\perp\!\!\!\perp$ Activity

Recovery $\perp\!\!\!\perp$ (Activity , Environment) | Survival

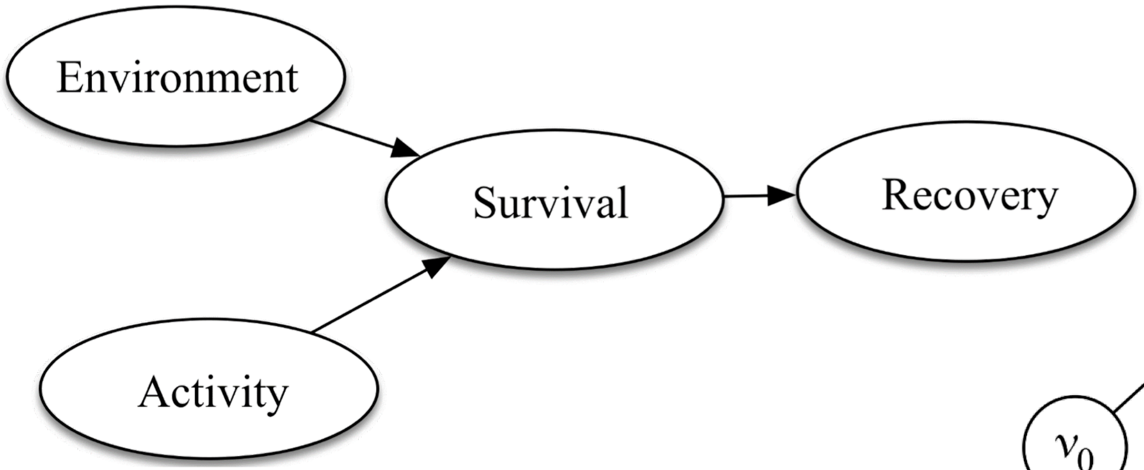
Bayesian network approach

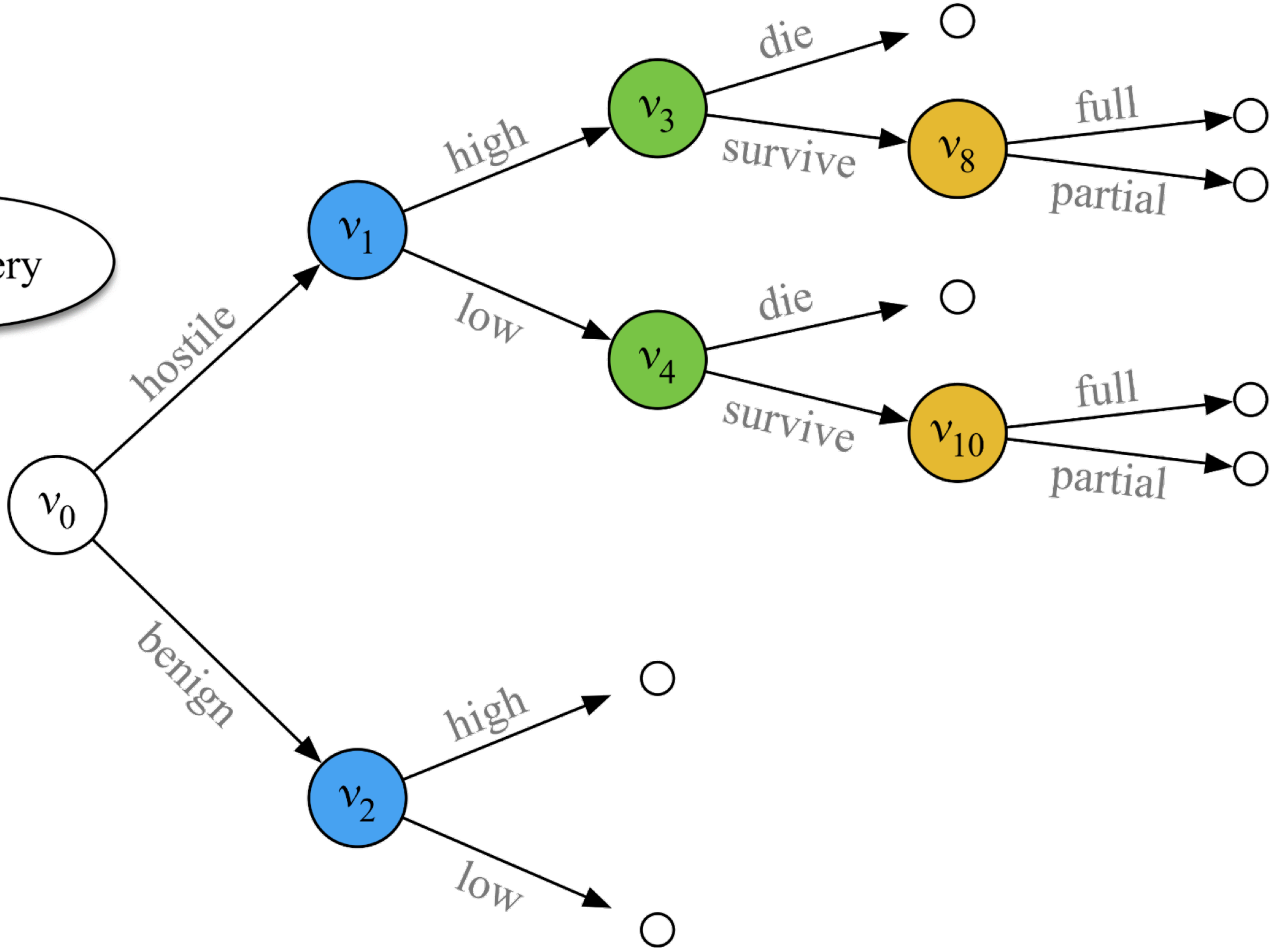


Environment $\perp\!\!\!\perp$ Activity

Recovery $\perp\!\!\!\perp$ (Activity , Environment) | Survival

Bayesian network approach



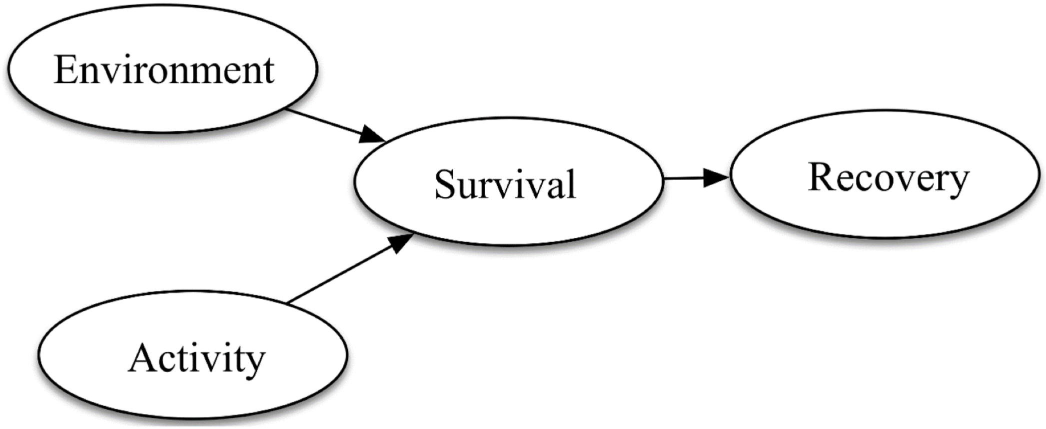


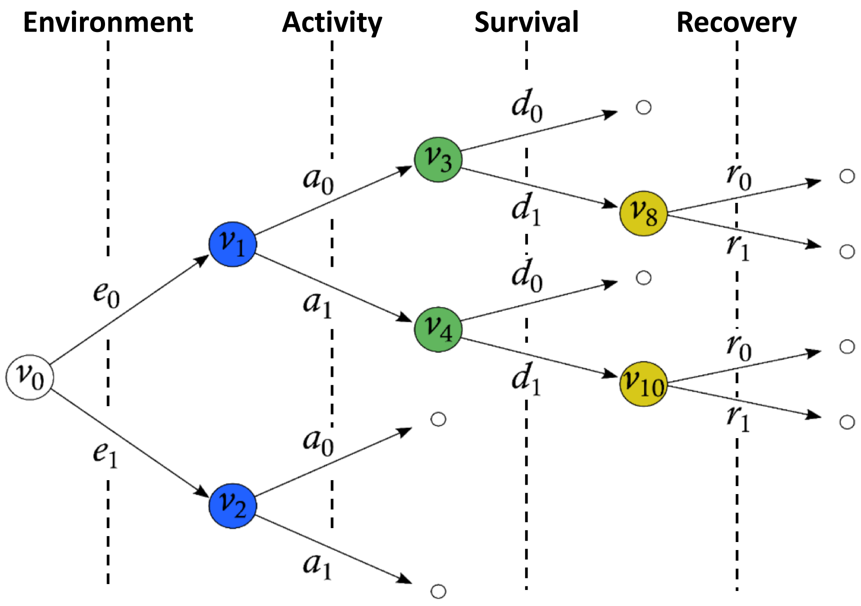
Environment $\perp\!\!\!\perp$ Activity

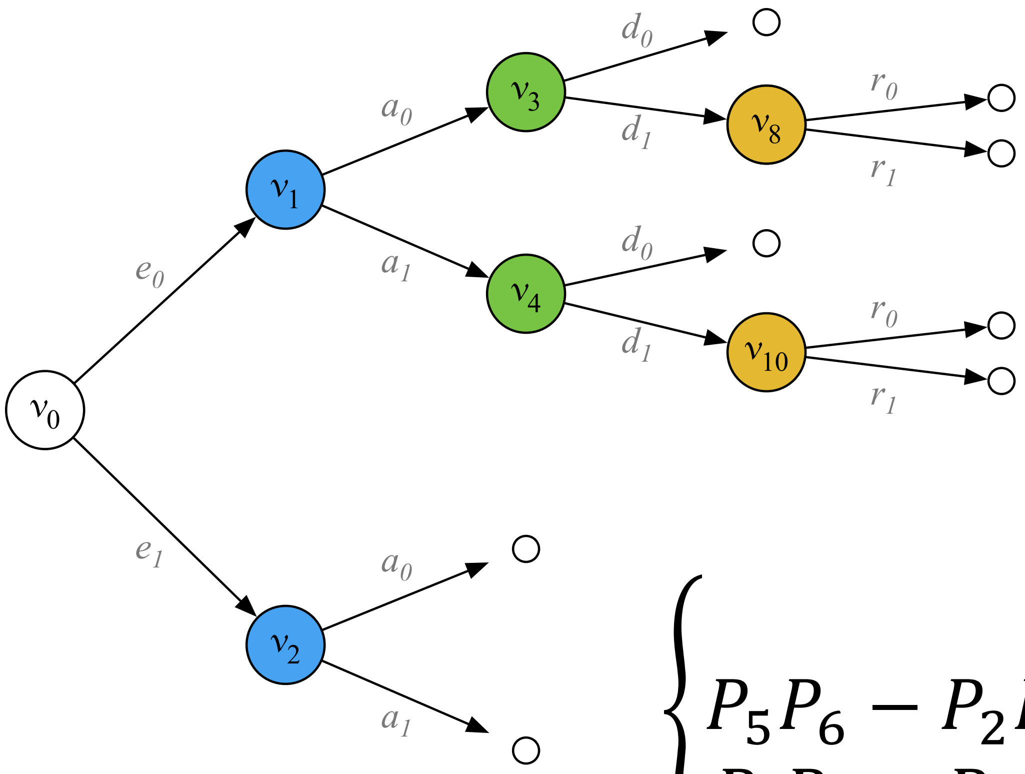
Recovery $\perp\!\!\!\perp$ (Activity , Environment) | Survival

Properties of staged tree models

- Encode conditional independence statements
- Generalize discrete Bayesian networks
- They capture context specific information
- The outcome space is not necessarily a cartesian product
- 3-valent phylogenetic trees are staged tree models (Améndola et al. 2019)







Toric varieties are a class of algebraic varieties whose defining equations only have two terms

$$\begin{cases} & P_0 + \cdots + P_7 - 1, & \\ P_5P_6 - P_2P_7, & P_3P_6 - P_0P_7, & P_2P_3 - P_0P_5, \\ P_4P_6 - P_1P_7, & P_2P_4 - P_1P_5, & P_1P_3 - P_0P_4 \end{cases}$$

- Discrete exponential families are toric varieties
- Discrete Bayesian networks with chordal graphs are toric varieties

What are the algebraic and geometric properties of staged tree models?

When is a staged tree model a toric variety/discrete exponential family?

Duarte and Görgen (J. Symb. Comput.) 2019

What are the algebraic and geometric properties of staged tree models?

What are the implicit equations that define the model?

Ananiadi and **Duarte** (J. Algebr. Stat.: *in Rev.*)

What is the geometry of interventional distributions?

Duarte and Solus (*in Prep.*)

When is a staged tree model a toric variety/discrete exponential family?

Duarte and Görgen (J. Symb. Comput.) 2019

Unifying principle to study equations of discrete graphical models

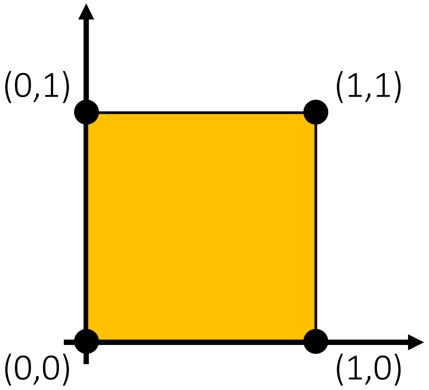
What are the implicit equations that define the model?

Ananiadi and **Duarte** (J. Algebr. Stat.: *in Rev.*)

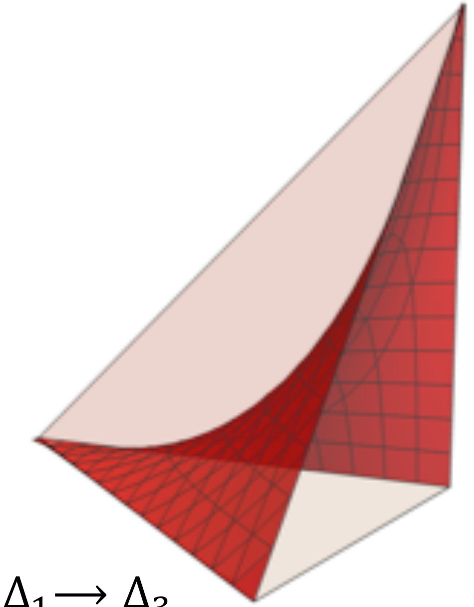
What is the geometry of interventional distributions?

Duarte and Solus (*in Prep.*)

Geometric Modeling and Algebraic Statistics



Toric Geometry



Polytope (2D)

Polytopes with rational linear precision

$$\Delta_1 \times \Delta_1 \longrightarrow \Delta_3$$

$$(s_0, s_1) \times (t_0, t_1) \longmapsto (s_0 t_0, s_0 t_1, s_1 t_0, s_1 t_1)$$

Rational Maximum Likelihood Estimator (MLE)

$$(u_{00}, u_{01}, u_{10}, u_{11}) \longmapsto \left(\frac{u_{0+}u_{+0}}{u_{++}^2}, \frac{u_{0+}u_{+1}}{u_{++}^2}, \frac{u_{1+}u_{+0}}{u_{++}^2}, \frac{u_{1+}u_{+1}}{u_{++}^2} \right)$$

?

Polytopes in higher dimensions

Duarte, Marigliano and Sturmfels (Bernoulli: *in Rev.*)

Staged tree models have Rational MLE

Are there any questions?

https://emduart2.github.io/

Max-Planck-Institut für Mathematik in den Naturwissenschaften

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG

DFG-Graduiertenkolleg MATHEMATISCHE KOMPLEXITÄTSREDUKTION

Dankeschön

Thank you

Gracias