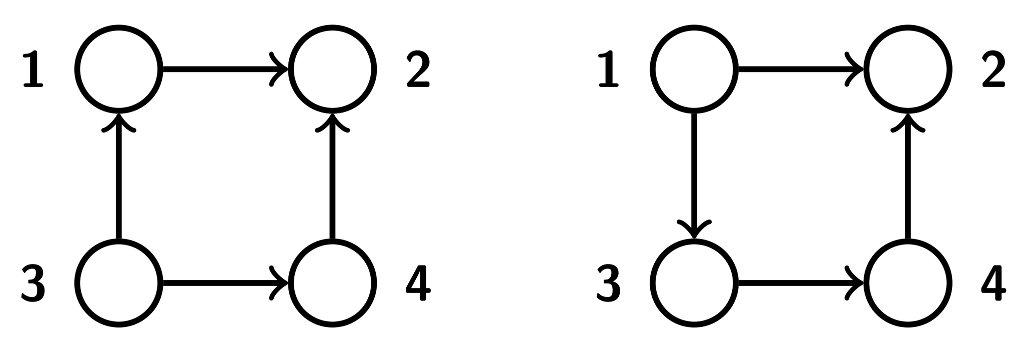
Uniform sampling of decomposable Gaussian graphical models

Irene Córdoba, Gherardo Varando, Concha Bielza, Pedro Larrañaga
Universidad Politécnica de Madrid, University of Copenhagen

Gaussian graphical models (GGMs)

Representation of **conditional independence** relationships via separation in an **acyclic digraph** (DAG)



- **Multivariate Gaussian** distributions

$$\boldsymbol{X} = (X_1, \ldots, X_p) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

- **Markov property** (ancestral order)

$$X_i \perp\!\!\!\perp \boldsymbol{X}_{\{1,\ldots,i-1\}\setminus \mathrm{pa}(i)} | \boldsymbol{X}_{\mathrm{pa}(i)}$$
$$X_i = \sum_{j<i} \beta^{ji|\,\mathrm{pa}(i)} X_j + \epsilon_i$$

Usual simulation set-up: given a DAG G:

- Sample $\beta^{ji|\,\mathrm{pa}(i)} \sim U(0.1, 1)$
- Recursively generate data via Markov property

$\rightarrow$ **problem**: bias when exploring the space of GGMs

$\rightarrow$ can mislead synthetic validation of GGM algorithms

$\rightarrow$ *our proposal*: sample GGMs uniformly

Our proposed GGMs parametrization: theoretical results

We propose to use the **upper Cholesky** factorization of $\boldsymbol{\Sigma}^{-1}$ as a parametrization Φ of the model:

$$\boldsymbol{\Sigma}^{-1} = \boldsymbol{U}\boldsymbol{U}^t = (\boldsymbol{I} - \boldsymbol{B})^t \boldsymbol{V}^{-1} (\boldsymbol{I} - \boldsymbol{B})$$
$$\Phi(\boldsymbol{U}) \rightarrow \boldsymbol{\Sigma}^{-1} = \boldsymbol{\Omega}$$

- $\boldsymbol{B}$ is lower triangular and contains the regression coefficients $\beta^{ji|\,\mathrm{pa}(i)}$
- $\boldsymbol{V}$ contains the variances of $\epsilon_i \rightarrow$ diagonal elements in $\boldsymbol{U}$ are strictly positive
- need unit diagonal on $\boldsymbol{\Omega}$ in order to bound the sample space $\rightarrow$ correlation matrix

$\rightarrow$ *idea*:

- consider the case of **decomposable** DAGs (no v-structures)
- zeroes are preserved between $\boldsymbol{U}$ and $\boldsymbol{\Omega} \rightarrow$ the Jacobian $J\Phi(\boldsymbol{U})$ is a square matrix
- sampling from $\propto \det J\Phi(\boldsymbol{U})$ yields the uniform distribution over decomposable GGMs

Lemma. *The elements of the Jacobian matrix $J\Phi(\boldsymbol{U})$ with respect to an acyclic digraph G are*

$$\frac{\partial \omega_{ji}}{\partial u_{st}} = \begin{cases} u_{jt} & s = i \wedge t \in \mathrm{ch}(j), \\ u_{it} & s = j \wedge t \in \mathrm{ch}(i), \\ u_{ii} & s = j \wedge t = i, \\ 0 & \textit{otherwise}, \end{cases} \quad \frac{\partial \omega_{ii}}{\partial u_{st}} = \begin{cases} 2u_{it} & s = i \wedge t \in \mathrm{ch}(i) \\ 2u_{ii} & s = t = i, \\ 0 & \textit{otherwise}. \end{cases}$$

Theorem. *If $J\Phi$ is the Jacobian matrix of Φ with u_{st} and ω_{ji} ordered with the row order we have that $J\Phi$ is an upper triangular matrix whose determinant is*

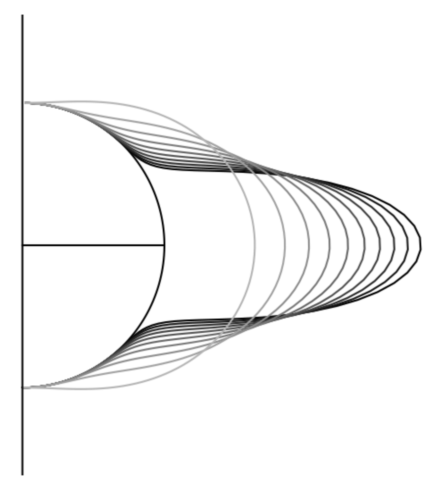
$$\det(J\Phi(\boldsymbol{U})) = 2^p \prod_i u_{ii}^{|\,\mathrm{pa}(i)|+1}.$$

Metropolis-Hastings (MH) sampling algorithm

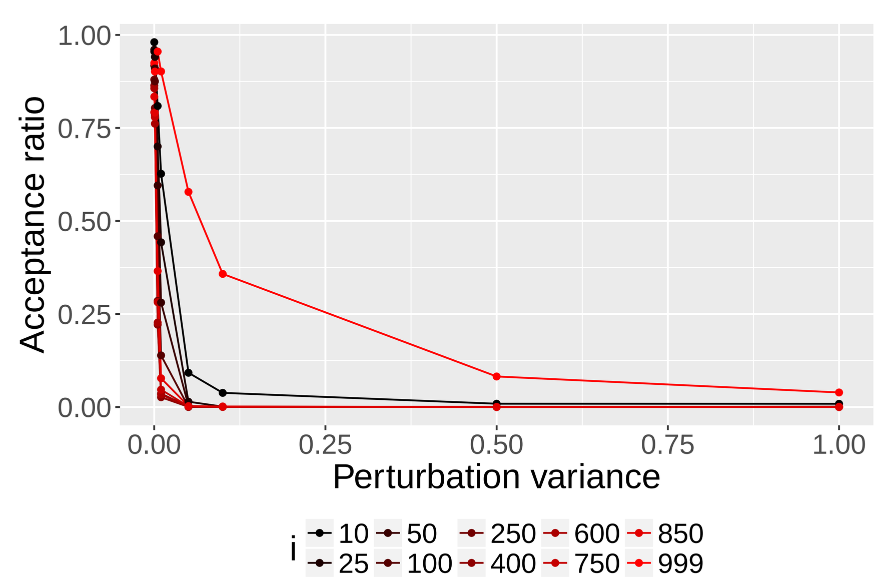
Input: Sample size N
Output: N random GGMs corresponding to a given graph

1: **for** $n = 1, \ldots, N$ **do**
2: $\quad \boldsymbol{U}^n \leftarrow \boldsymbol{0}_p$
3: $\quad$ **for** $i = 1, \ldots, p$ **do**
4: $\quad\quad \boldsymbol{u}_i^n \leftarrow$ sample from $f(\boldsymbol{u}_i) \propto u_{ii}^{|\,\mathrm{pa}(i)|+1}$ on $\mathcal{S}_+^{|\,\mathrm{ch}(i)|}$ using MH
5: $\quad$ **end for**
6: **end for**
7: **return** $\{\Phi(\boldsymbol{U}^1), \ldots, \Phi(\boldsymbol{U}^N)\}$

- target density for each run of MH: $\propto u_{ii}^{|\,\mathrm{pa}(i)|+1}$
- $2D$ example $\rightarrow$ **circular density**



- MH new proposed value is a normalized perturbation of the previous state
- Results in the **projected normal** proposal distribution $\rightarrow$ similar to the target one $\rightarrow$ good theoretical convergence properties

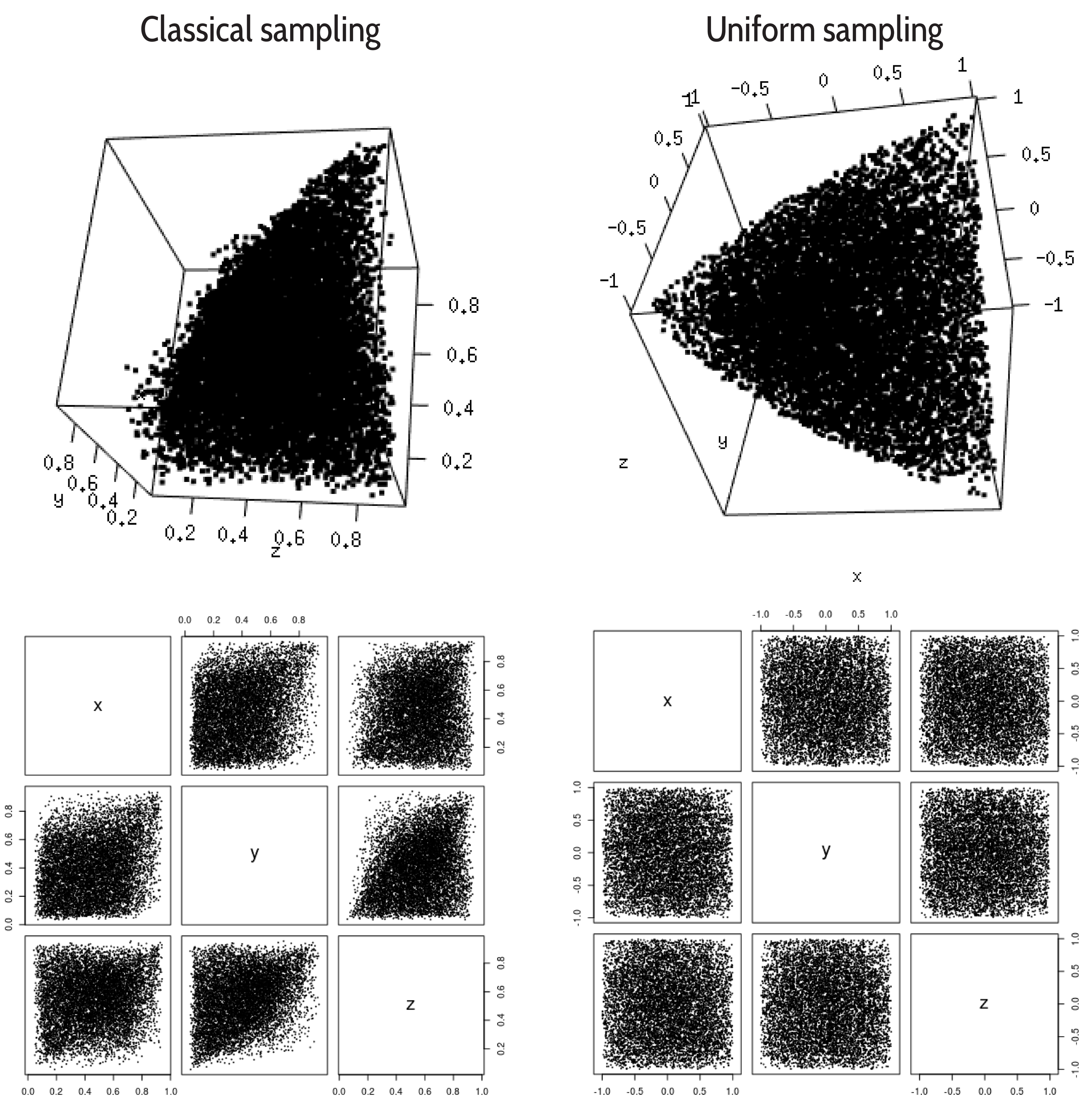


$\rightarrow$ moderately small perturbations for fast convergence

An illustrative example

$\rightarrow$ $\boldsymbol{\Omega}$ of dimension 3×3 vectorized

$\rightarrow$ can be represented as a convex $3D$ body called **elliptope**



A **whole region** of the elliptope is **missed** by the classical sampling scheme.

Conclusions

- The proposed method allows for uniform sampling over **decomposable GGMs**.
- In future work, we want to extend the results to **non-decomposable** graphs. We will also further analyze the relationship between the **topology and sparsity** of the graph and the behaviour of the MH algorithm, both theoretically and empirically.