

---

# Mixed high-dimensional network inference via the Gaussian copula

Ekaterina Tomilina\*<sup>1,2</sup>, Gildas Mazo<sup>1</sup>, and Florence Jaffrézic<sup>3</sup>

<sup>1</sup>Mathématiques et Informatique Appliquées du Génome à l'Environnement [Jouy-En-Josas] – Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement – France

<sup>2</sup>Génétique Animale et Biologie Intégrative – Université Paris-Saclay, Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement – France

<sup>3</sup>Génétique Animale et Biologie Intégrative – AgroParisTech, Université Paris-Saclay, Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement, Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement : UMR1313 – France

## Résumé

L'intégration de données hétérogènes à grande échelle pour l'inférence de réseaux est un défi majeur, en particulier dans le contexte de l'analyse de données multi-omiques. Nous proposons ici une nouvelle méthode basée sur les copules afin d'analyser conjointement des données de différents types (continues, discrètes, etc.). La procédure d'estimation proposée est semi-paramétrique et ne nécessite donc aucune hypothèse explicite concernant les distributions marginales des données, ce qui offre une grande flexibilité pour l'analyse des données biologiques. Nous présentons également un résultat théorique montrant l'équivalence entre l'indépendance par blocs dans la matrice de corrélation de la copule et dans la structure de corrélation des données réelles. Nous illustrons via une étude de simulation les capacités de l'estimateur du maximum de vraisemblance par paires à estimer la matrice de corrélation de la copule, même pour un nombre assez important de variables (plusieurs centaines) et un nombre assez faible de réplicats (plusieurs dizaines). Enfin, nous appliquons la méthode à un jeu de données réel ICGC concernant le cancer du sein.

**Mots-Clés:** copule, inférence de réseaux, données mixtes, grande dimension, semi paramétrique

---

\*Intervenant