

# Analyse de réseaux trophiques : Comparaison d'algorithmes pour l'échantillonage uniforme de polytope

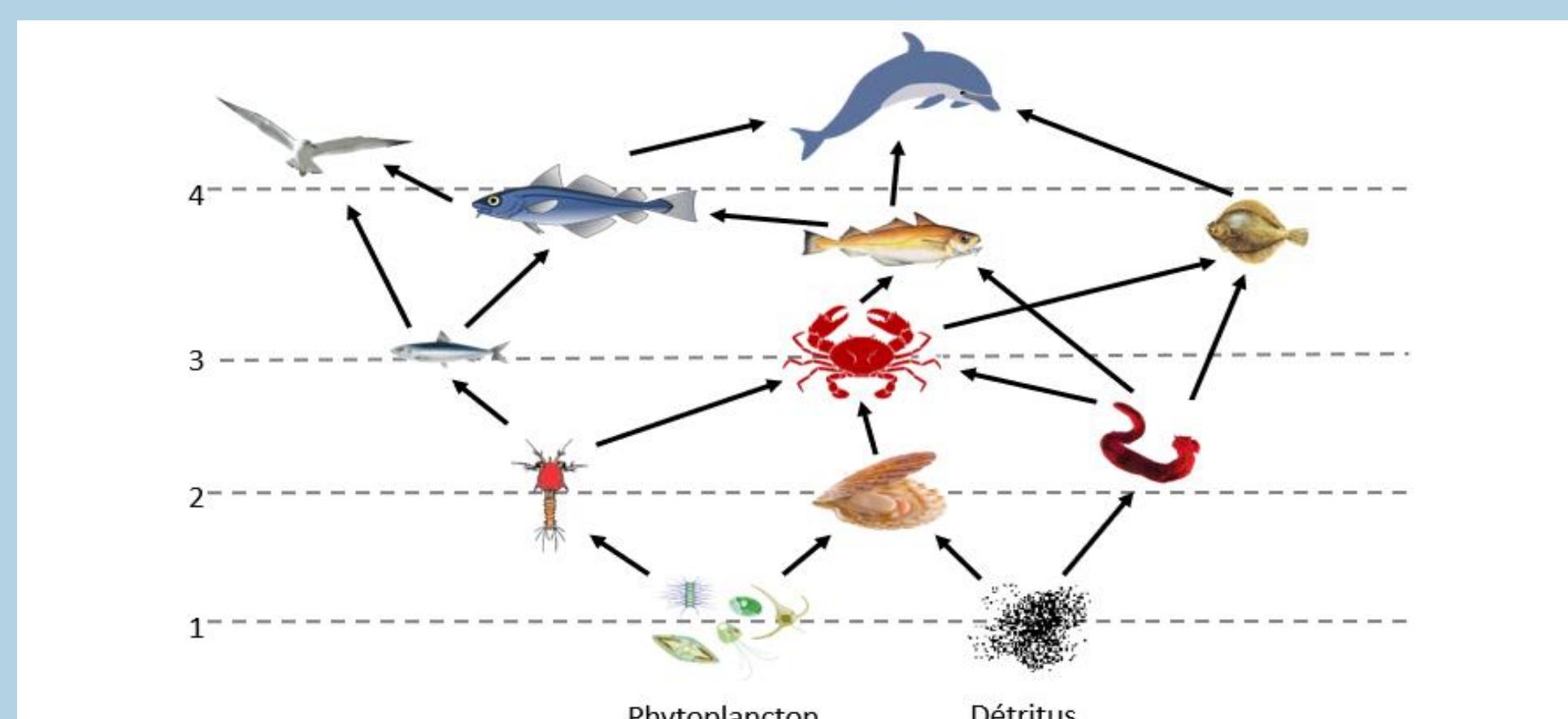
T. Grente<sup>1,2</sup> M.Dien<sup>3</sup> V. Girardin<sup>2</sup> N. Niquil<sup>4</sup> P. Regnault<sup>5</sup>

1. France Energies Marines 2. Normandie Université, Laboratoire de Mathématiques Nicolas Oresme (UMR 6139 CNRS) 3. Normandie Université, Groupe de Recherche en Informatique, Image et Instrumentation de Caen (UMR 6072 CNRS)  
4. Normandie Université, Laboratoire de Biologie des organismes et des écosystèmes aquatiques (UMR 8067 CNRS) 5. Université de Reims Champagne-Ardenne, Laboratoire de Mathématiques de Reims (UMR 9008 CNRS)

## Les réseaux trophiques

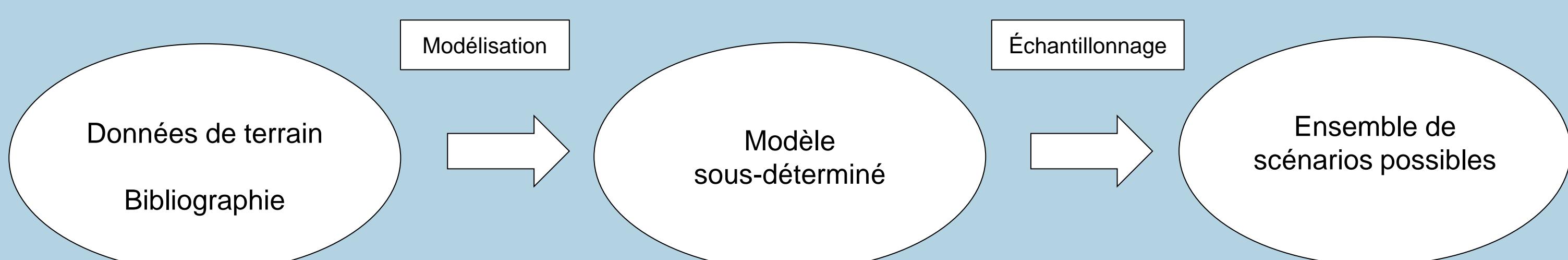
Utilisés notamment en écologie marine, les réseaux trophiques sont une représentation sous forme de graphes des interactions proies/prédateurs d'un écosystème.

Les nœuds du graphe représentent les espèces et les arêtes leurs interactions sous forme d'échanges de matière organique appelés flux.



## Le problème inverse linéaire (LIM)

Dans le but d'estimer la valeur de ces flux, une classe de méthodes appelée Modélisation Linéaire Inverse (LIM) a été développée. Ces méthodes consistent à établir un système de contraintes sur ces flux à partir de mesures réalisées sur le terrain, d'expériences en laboratoire et de connaissances issues de la bibliographie.



Les équations et les inéquations de ce système de contraintes définissent un polytope (généralisation à toutes dimensions de la notion de polygone en deux dimensions) à l'intérieur duquel se trouve l'ensemble des solutions du système de contraintes. Pour obtenir un ensemble de scénarios possibles représentatif, les écologues utilisent actuellement le package R `limsolve`. La fonction d'échantillonage de ce package est basée sur une exploration du polytope par une méthode de Monte-Carlo par chaînes de Markov où les bordures du polytope agissent comme des miroirs sur lesquels se réfléchit une trajectoire.

## Exploration et échantillonage du polytope par la méthode du miroir

Le système de contraintes définissant le polytope  $\mathbf{S}$  des solutions peut se résumer sous la forme du système :

$$Ax = b$$

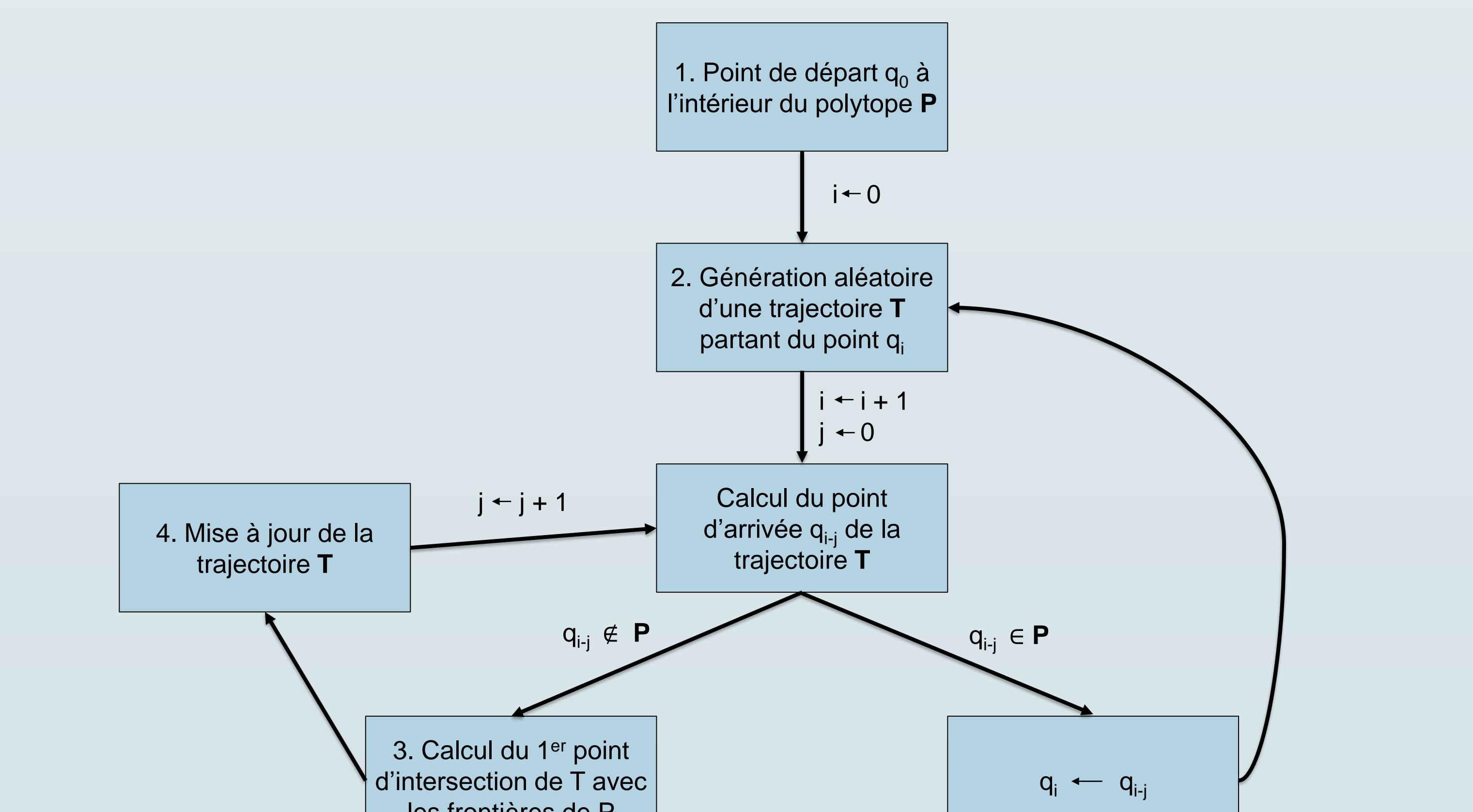
$$Gx \leq h$$

L'échantillonage est alors réalisé dans le polytope  $\mathbf{P}$  à pleine dimension défini par le système de contraintes :

$$G'x \leq h'$$

avec  $G' = GZ$  et  $h' = Gx_0 - h$  où  $Z$  est la base de l'espace nul à droite de  $A$  et  $x_0$  est une solution particulière de  $\mathbf{S}$ .

Le diagramme ci-dessous décrit la méthode de Monte-Carlo par chaînes de Markov utilisée pour explorer le polytope.



1. Le point de départ  $q_0$  de l'exploration du polytope par la méthode du miroir peut être obtenu par l'utilisation d'un solveur de contraintes linéaires.

2. La trajectoire  $T$  est générée suivant une loi normale d'espérance 0 et un vecteur d'écart type donné en entrée. Elle peut se décomposer sous la forme  $T = L \times \eta$  où  $L$  est sa longueur et  $\eta$  son vecteur direction.

3. Pour chaque frontière (hyperplan) du polytope, un scalaire  $\alpha_p$  est calculé tel que  $G'_{(p)}(q_{i-j} + \alpha_p \eta) + h'_p = 0$ . L'hyperplan avec le plus petit  $\alpha_p$  positif, noté  $\alpha_s$ , est le premier hyperplan traversé par la trajectoire  $T$ .

4. La trajectoire est ensuite mise à jour pour démarer au point d'intersection entre l'hyperplan  $s$  et le polytope  $P$ .

La longueur devient  $L - \alpha_s$  et sa direction est réfléchie selon l'hyperplan comme un rayon de lumière sur un miroir (l'angle d'incidence est égal à l'angle de réflexion).

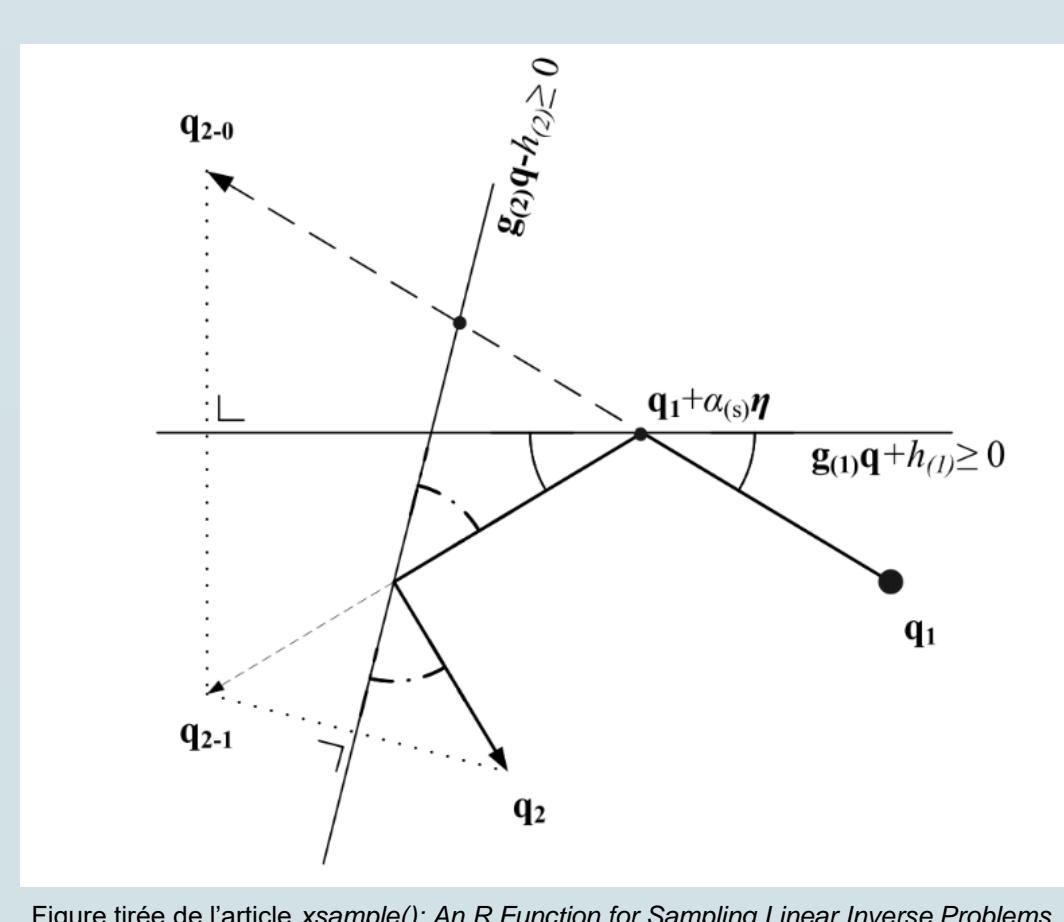


Figure tirée de l'article `xsample`: An R Function for Sampling Linear Inverse Problems  
K. Van den Meersche, K. Soetaert, D. Van Oevelen, <https://doi.org/10.18637/jss.v030.c01>

## limsolve : un package R pour l'échantillonage de polytope

### Avantages :

- Bonne qualité des échantillons obtenus : très faible corrélation entre deux points consécutifs
- Permet une exploration adaptative du polytope : écart-type de la loi normale proportionnel à l'étendue de la variable

### Inconvénients :

- Dans un polytope non isotrope de dimension importante, la fonction d'échantillonage (`xsample`) est très lente
- Le point de départ de l'exploration est généralement obtenu par la méthode des moindres carrés et se trouve donc dans un coin du polytope
- Présence de plusieurs bugs, liés notamment à sa dépendance avec de vieux packages

### Conclusion :

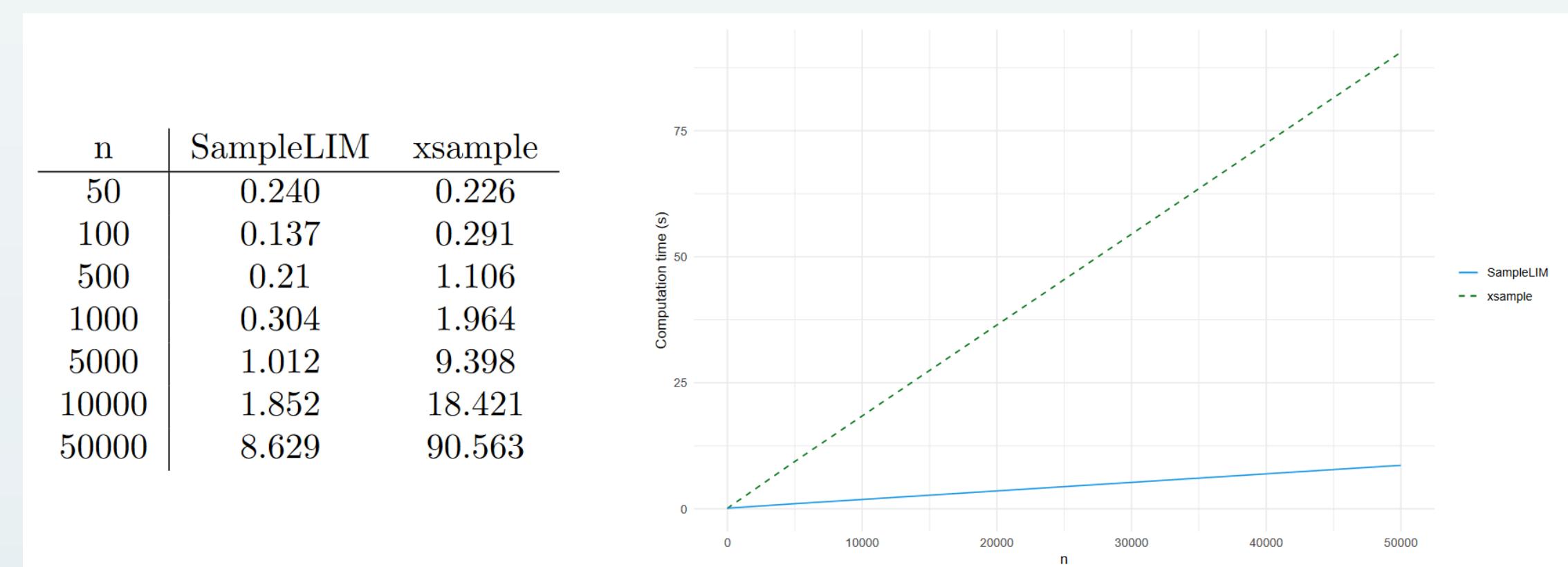
- `limsolve` a été développé il y a plus de 10 ans → Besoin d'une mise à jour

## SampleLIM : une mise à jour du package limsolve

- Basé sur le package R `volesti` utilisé pour calculer le volume de polytopes.
- Utilise la même méthode d'exploration (miroir) que `xsample`.
- Le centre du polytope est choisi comme point de départ de l'exploration.
- Entièrement programmé en C++ avec une interface R : transparent pour l'utilisateur mais temps de calcul considérablement réduit.
- Correction de fonctions annexes présentes dans `limsolve` (`xranges`).

## Comparaison de performances entre limsolve et SampleLIM

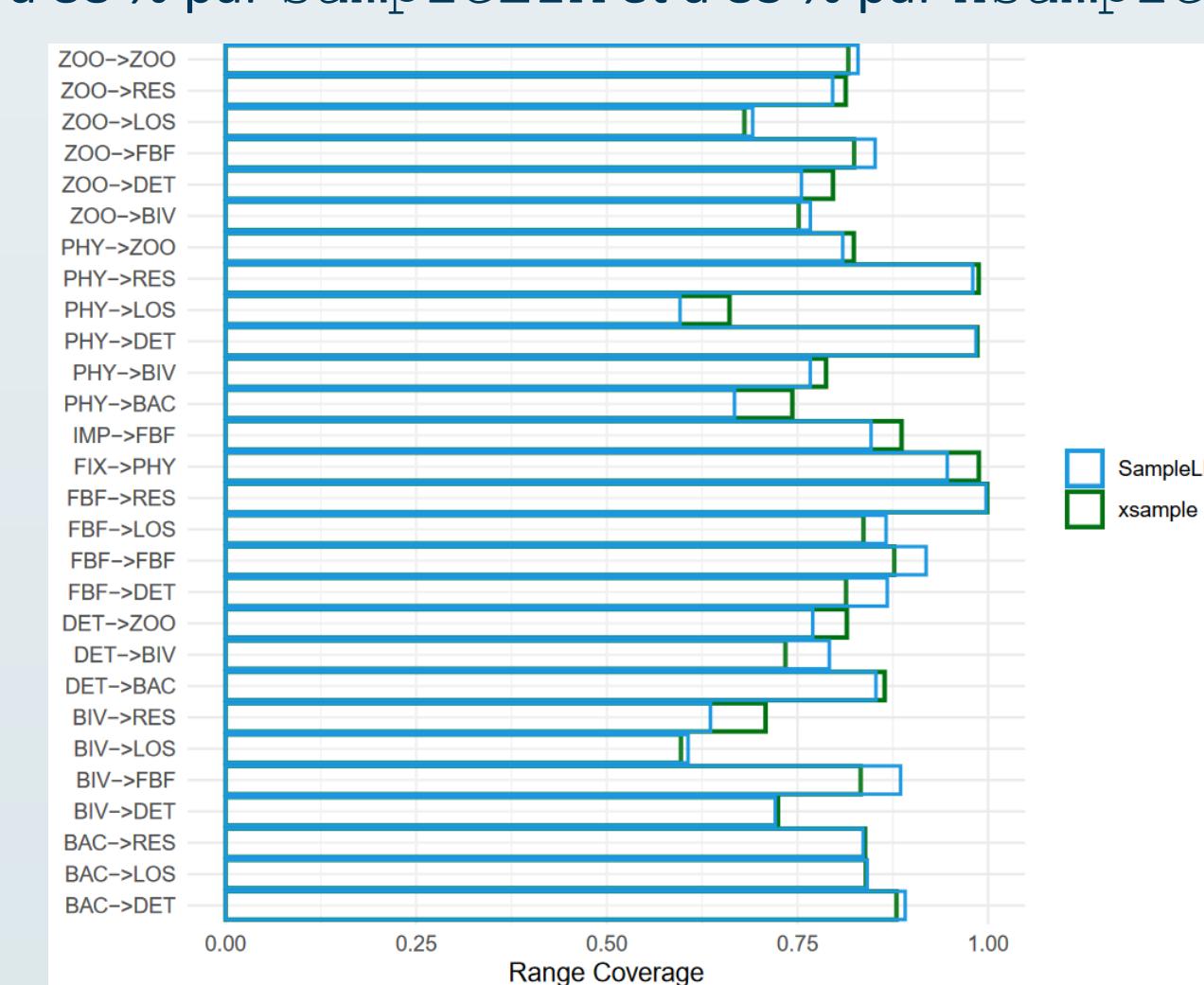
Les performances ci-dessous ont été obtenues en faisant tourner `limsolve` et `SampleLIM` sur un modèle trophique simplifié afin d'éviter des temps de calcul trop importants de la part de `limsolve`.



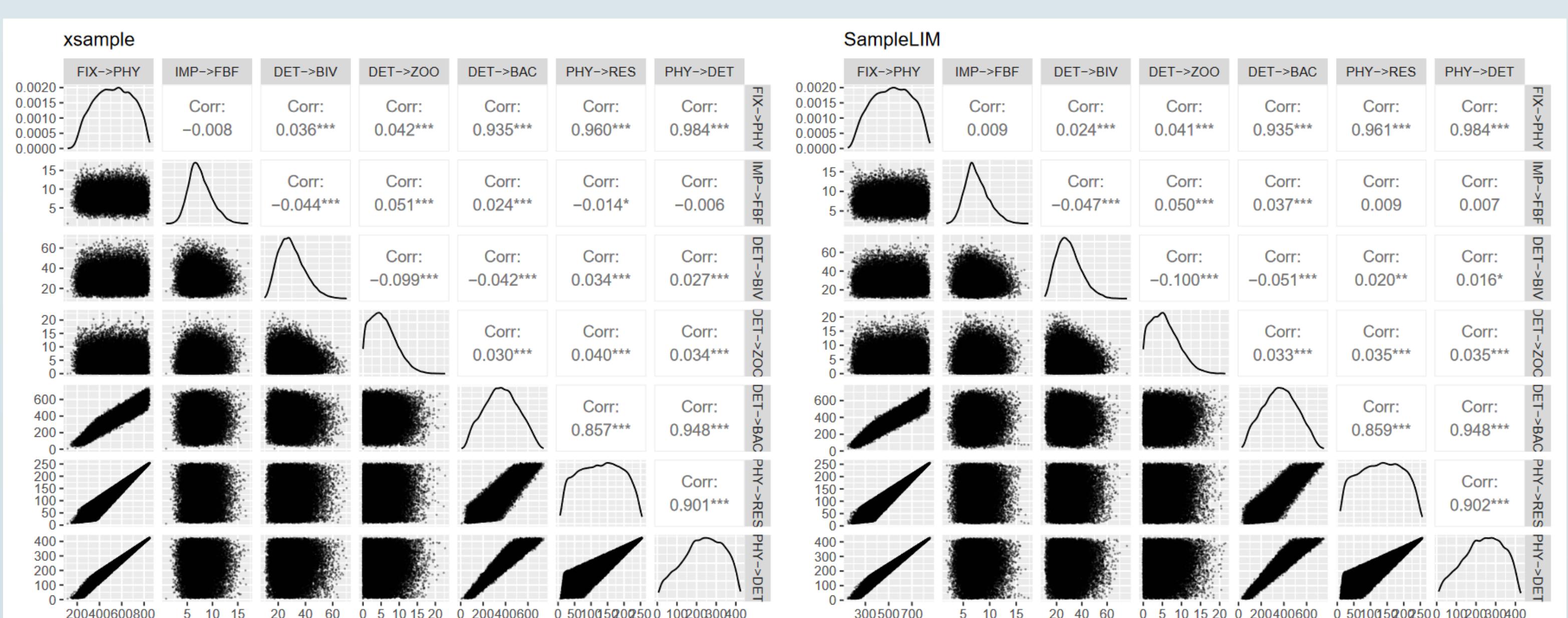
La *range coverage* est un indice que nous avons développé pour mesurer la qualité de l'échantillon obtenu.

Il s'agit du rapport entre l'étendue des flux observé dans l'échantillon et l'étendue théorique de ces flux.

Par exemple, l'étendue du flux `PHY -> DET` est couverte par les deux algorithmes à plus de 98% tandis que l'étendue du flux `BIV -> FBF` est couverte à 88 % par `SampleLIM` et à 83 % par `xsample`.



Pour s'assurer que `SampleLIM` génère des échantillons de même qualité que `limsolve`, on peut projeter les échantillons obtenus sur 2 dimensions (2 flux) avec une estimation de la corrélation entre ces deux flux.



## Perspectives

La réduction du temps de calcul pour obtenir un échantillon de scénarios ouvre de nouvelles perspectives :

- Gain de temps important pour les écologues ;
- Étude plus efficace de l'impact de la dégradation des données de terrain sur le modèle ;
- Possibilité de lancer plusieurs instances de LIM sur un maillage spatial pour étudier des modèles avec un couplage de paramètres biologiques/physiques.