

# OPTIMISATION SOUS CONTRAINTES EN CHANCES RARES :

## Un Cadre Robuste pour la Gestion du Risque de Queue des Actions Piliers de l'Intelligence Artificielle

Georges Marc-Arthur Sébastien

December 27, 2025

### Contents

<b>1 INTRODUCTION</b>	<b>4</b>
1.1 Motivation et Contexte Général . . . . .	4
1.1.1 Contexte Sectoriel : Le Risque de Queue dans l'IA . . . . .	4
1.2 Contribution et Objectifs de l'Étude . . . . .	4
1.2.1 Problématique . . . . .	4
1.2.2 Objectifs du Projet . . . . .	4
1.3 Fondements Mathématiques Clés . . . . .	5
1.3.1 Théories de l'Extrême et de l'Approximation . . . . .	5
<b>2 REVUE DE LITTÉRATURE : DE L'INTRACTABILITÉ À LA THÉORIE DES GRANDES DÉVIATIONS</b>	<b>5</b>
2.1 L'Intractabilité des Contraintes en Chance Rares : Généralités . . . . .	5
2.1.1 Difficultés Théoriques des Contraintes en Chance . . . . .	5
Première Famille : Exploitations Structurelles . . . . .	5
Deuxième famille : méthodes basées sur les échantillons . . . . .	6
Méthodes basées sur la réduction de variance . . . . .	6
Importance Sampling . . . . .	6
Approche par scénarios . . . . .	6
Méthodes elliptiques radiales . . . . .	6
Troisième famille : méthodes de fiabilité (FORM/SORM) . . . . .	6
2.1.2 Synthèse des limitations de la littérature . . . . .	6
2.2 L'Intractabilité des Contraintes en Chance Rares : Généralités . . . . .	7
2.2.1 Décroissance Exponentielle et Échec des Méthodes Classiques . . . . .	7
Justification de l'Échec de Monte Carlo . . . . .	7
2.2.2 La Fonction de Taux ( $I(\mathcal{A})$ ) et l'Entropie Relative . . . . .	7
Définition Formelle . . . . .	7
Le Parallèle avec l'Entropie Relative . . . . .	8
Interprétation physique et mathématique . . . . .	8
Avantages Majeurs . . . . .	8
2.2.3 Le Point Dominant ( $\xi^*$ ) . . . . .	8
Interprétation . . . . .	8
Lien avec la Méthode de Fiabilité . . . . .	9
2.2.4 Amélioration de l'Estimation par Linéarisation Locale (FORM/SORM) . . . . .	9
Rôle des Développements de Taylor . . . . .	9
Exploitation des Propriétés Gaussiennes . . . . .	9
2.2.5 Pseudo-Code de Résolution OCCR par l'Approche LDT . . . . .	9
2.3 Fondements Théoriques et Conditions de Validité . . . . .	10
2.3.1 Théorème 3.2 : Approximation LDT de la probabilité rare . . . . .	10
Conclusion du théorème . . . . .	11

2.3.2	Théorème 3.3 : Conditions suffisantes pour le Théorème 3.2 . . . . .	11
	Condition principale. . . . .	11
	Conditions additionnelles. . . . .	11
	Conclusion. . . . .	11
2.3.3	Théorème 4.1 : Validité pour les mélanges gaussiens . . . . .	11
	Conclusion du théorème. . . . .	11
	Interprétation. . . . .	11
2.4	Littérature Empirique : Optimisation et Risque du Secteur de l'Intelligence Artificielle . . . . .	12
2.4.1	Optimisation et Marché de l'Intelligence Artificielle (Général) . . . . .	12
2.4.2	Optimisation en Chance Rare et AI (Gestion du Risque Extrême) . . . . .	12
2.4.3	Caractère Unique de la Contribution et Gestion du Risque d'Effondrement . . . . .	12
	Fusion Méthodologique LDT-Fiabilité . . . . .	12
	Gestion Directe vs. Indirecte du Risque d'Effondrement Sectoriel . . . . .	13
<b>3</b>	<b>DÉMARCHE MÉTHODOLOGIQUE</b>	<b>13</b>
3.1	Choix des actifs et justification : les six piliers de l'écosystème IA . . . . .	13
3.1.1	Contexte général . . . . .	13
3.1.2	Définition des piliers IA . . . . .	13
3.1.3	Critères de sélection . . . . .	13
3.1.4	Présentation des six titres retenus . . . . .	14
3.1.5	Positionnement des six piliers IA parmi les grands acteurs technologiques . . . . .	14
3.2	Modèle mathématique, hypothèses et formulations LDT . . . . .	14
3.2.1	Distribution du vecteur aléatoire . . . . .	14
3.2.2	Fonction de perte et contrainte en chance . . . . .	15
3.2.3	Point dominant LDT . . . . .	15
3.2.4	Approximation de premier ordre $P_1$ . . . . .	15
3.2.5	Approximation de second ordre $P_2$ . . . . .	15
3.2.6	Formulation LDT du problème d'optimisation . . . . .	15
3.3	Résultats attendus . . . . .	16
3.3.1	Estimation efficace de la probabilité rare . . . . .	16
3.3.2	Problème d'optimisation tractable . . . . .	16
3.3.3	Caractérisation du scénario extrême . . . . .	16
3.3.4	Structure du portefeuille optimal . . . . .	16
3.3.5	Comparaison qualitative avec les méthodes classiques . . . . .	17
<b>4</b>	<b>RÉSULTATS EMPIRIQUES</b>	<b>17</b>
4.1	Modélisation et description des données . . . . .	17
4.2	Portefeuille et pondérations . . . . .	17
4.2.1	Modélisation des rendements . . . . .	18
4.2.2	Valeur du portefeuille et variable de perte . . . . .	18
4.2.3	Méthodologie de shringage Ledoit–Wolf . . . . .	18
4.2.4	Estimation de la moyenne $\mu$ et de la matrice de covariance $\Sigma$ . . . . .	19
4.3	Structure de corrélation des actifs IA . . . . .	20
4.4	Évaluation du risque : estimation d'une probabilité rare . . . . .	20
4.4.1	Estimateur Monte Carlo standard (SAA) . . . . .	21
4.4.2	Approximation de grandes déviations de premier ordre (P1) . . . . .	21
4.5	Allocation optimale sous contraintes de chance rares . . . . .	22
4.5.1	Portefeuilles de références (baseline) . . . . .	22
4.5.2	Value at Risk (VaR), Conditional Value at Risk (CVaR) et contrainte en chance . . . . .	22
4.5.3	Problème d'optimisation sous contrainte en chance . . . . .	23
4.5.4	Résultats numériques . . . . .	24
4.6	Analyse des points dominants . . . . .	24
4.6.1	Définition et interprétation économique . . . . .	24
4.6.2	Analyse empirique des points dominants par portefeuille . . . . .	24
	Portefeuille équipondéré. . . . .	25
	Portefeuille max-rendement. . . . .	25
	Portefeuille minimum variance. . . . .	25
	Portefeuille optimisé sous contrainte de chance rare. . . . .	26

<b>5 DISCUSSION</b>	<b>26</b>
5.1 Synthèse des résultats empiriques . . . . .	26
5.2 Implications conceptuelles pour la gestion du risque . . . . .	26
5.3 Limites de l'étude . . . . .	26
<b>6 CONCLUSION</b>	<b>27</b>

# 1 INTRODUCTION

## 1.1 Motivation et Contexte Général

L'optimisation sous contraintes en chances rares (OCCR) gagne aujourd'hui en importance et devient un outil essentiel dans les domaines où les décisions doivent satisfaire des exigences de fiabilité extrême. Historiquement, la modélisation du risque financier a longtemps reposé sur des distributions simples (telles que la loi Normale, utilisée notamment dans le modèle Black-Scholes de 1973), qui sous-estimaient la probabilité et l'impact des mouvements extrêmes.

Le Krach boursier de 1987, et surtout la crise financière mondiale de 2008 (marquée par le concept de Cygne Noir), ont révélé l'insuffisance des outils conventionnels, comme la Value at Risk (VaR), à capter le risque systémique et les défaillances de faible probabilité mais à fort impact. Ces événements ont souligné l'importance d'outils qui assurent la survie et la fiabilité extrême.

C'est dans ce contexte que l'OCCR s'est imposée. Elle exige un contrôle strict de la probabilité de défaillance, souvent exprimée par :

$$\mathbb{P}(g(X, \xi) \leq 0) \leq \delta \quad (1)$$

où  $\delta$  est un seuil de probabilité extrêmement petit (typiquement de  $10^{-4}$  à  $10^{-9}$ ). Ce régime nécessite un cadre mathématique robuste, car les événements de défaillance ne sont pratiquement jamais observés sur l'horizon de données classique, rendant les simulations de type Monte Carlo classiques inopérantes .

### 1.1.1 Contexte Sectoriel : Le Risque de Queue dans l'IA

Ces méthodes, bien que nées dans l'ingénierie et l'énergie, trouvent aujourd'hui une application cruciale dans les marchés financiers technologiques. Le secteur de l'Intelligence Artificielle connaît une expansion spectaculaire, mais il est caractérisé par :

1. Des valorisations extrêmes et une forte concentration du marché.
2. Une volatilité et une interdépendance prononcées entre les titres leaders.

Ces caractéristiques créent un environnement où le risque de krach est faible en probabilité mais colossal en impact, faisant du secteur IA un terrain d'étude idéal pour l'OCCR.

## 1.2 Contribution et Objectifs de l'Étude

Le présent travail n'a pas pour objectif de reproduire l'intégralité des résultats de Tong et al. (2022), mais plutôt d'adapter leur cadre au contexte du marché IA, en développant un prototype fonctionnel basé uniquement sur l'approximation FORM ( $P_1$ ), cohérente avec les contraintes temporelles du projet.

### 1.2.1 Problématique

Comment utiliser l'optimisation sous Contraintes en Chances Rares pour quantifier, anticiper et contrôler le risque de krach au sein des principales actions considérées comme les piliers de l'écosystème IA moderne ?

### 1.2.2 Objectifs du Projet

Ce projet vise à :

- Modéliser un portefeuille constitué des principales actions leaders de l'IA (NVIDIA, Microsoft, AMD, Meta, Broadcom et Oracle).
- Formuler une contrainte de chance rare représentant un crash sévère du portefeuille.
- Utiliser la théorie des Grandes Déviations (LDT) pour approximer la probabilité de ce crash extrême.
- Identifier le *dominating point* du portefeuille, c'est-à-dire le scénario extrême le plus probable déclenchant la crise.
- Comparer les résultats de la LDT à une analyse Monte Carlo classique, afin d'illustrer l'efficacité de la méthodologie dans le régime rare.

### 1.3 Fondements Mathématiques Clés

La mise en œuvre de l'OCCR dans la gestion de portefeuille se heurte au défi majeur de la **non-convexité** de la contrainte de chance, surtout lorsque les distributions financières (à queues lourdes) ne respectent pas l'hypothèse de log-concavité. Pour contourner cette intractabilité et adresser le régime rare, nous nous appuyons sur des avancées théoriques majeures :

#### 1.3.1 Théories de l'Extrême et de l'Approximation

- **Théorie des Grandes Déviations (LDT)** : Cruciale pour l'estimation efficace de la probabilité d'événements très rares. Les travaux fondateurs de Varadhan (1966, 1984) et son application aux problèmes de risque par Glasserman and Kim (2001) sont essentiels pour déterminer le scénario de défaillance le plus probable.
- **Théorie des Valeurs Extrêmes (EVT)** : Indispensable pour la modélisation statistique des queues de distribution financière. Les contributions clés de Pickands (1975) et l'application à la finance par Embrechts et al. (1997) permettent de caractériser rigoureusement le comportement des rendements extrêmes.
- **Approximation par le CVaR** : Pour garantir la traitabilité numérique, l'optimisation du Conditional Value at Risk (CVaR), démontrée comme étant convexe par Rockafellar and Uryasev (2000), sert de référence pour le contrôle de la queue de distribution, offrant ainsi un cadre robuste pour notre optimisation.

## 2 REVUE DE LITTÉRATURE : DE L'INTRACTABILITÉ À LA THÉORIE DES GRANDES DÉVIATIONS

### 2.1 L'Intractabilité des Contraintes en Chance Rares : Généralités

Les contraintes en chances sont utilisées pour modéliser des problèmes d'optimisation dans lesquels la solution doit rester sûre ou fiable avec une forte probabilité. Elles apparaissent notamment dans les systèmes où des événements rares ayant un impact significatif peuvent se produire (ingénierie, aéronautique, réseaux électriques, finance, etc.). Une contrainte en chance typique consiste à imposer :

$$\mathbb{P}(F(u, \xi) \geq z) \leq \alpha \quad (2)$$

où  $u$  désigne la décision,  $F(u, \xi)$  un indicateur de performance soumis à l'incertitude,  $z$  un seuil critique, et  $\alpha$  un niveau de risque acceptable, souvent très faible. Ces formulations permettent de limiter la probabilité d'événements extrêmes, mais elles posent des défis computationnels majeurs.

#### 2.1.1 Difficultés Théoriques des Contraintes en Chance

La raison principale de la difficulté est que l'ensemble des solutions satisfaisant une telle contrainte est, en général, **non convexe** Tong et al. (2022). Cette non-convexité conduit à des problèmes dits *NP-hard* (non tractables). Pour cette raison, la littérature a historiquement cherché des reformulations ou des approximations permettant de rendre ces problèmes tractables. Deux grands axes d'approches ont émergé:

**Première Famille : Exploitations Structurelles** Cette première famille repose sur des hypothèses fortes visant à restaurer la convexité du problème. L'idée est d'exploiter des propriétés de la fonction  $F(u, \xi)$  ou de la distribution de  $\xi$ . Parmi les hypothèses couramment rencontrées, on trouve :

- L'hypothèse que  $F(u, \xi)$  est convexe ou linéaire en  $u$ . Nemirovski and Shapiro (2007).
- L'hypothèse selon laquelle la probabilité  $\mathbb{P}$  est **quasi-concave** Prékopa (2014), ce qui se produit notamment lorsque  $\xi$  suit une distribution **log-concave** (ex. : la distribution Gaussienne). Sous cette condition, l'ensemble faisable peut devenir convexe.

Malheureusement, ces hypothèses structurelles sont rarement satisfaites dans les applications financières réelles.

- Les rendements d'actifs financiers sont caractérisés par des **queues lourdes** (*heavy-tailed*), des propriétés qui **violent l'hypothèse de log-concavité**.
- L'analyse analytique de la convexité est rendue encore plus complexe en **hautes dimensions**, limitant la portée de ces méthodes pour des portefeuilles complexes.

**Deuxième famille : méthodes basées sur les échantillons** Ces méthodes remplacent la contrainte probabiliste par une approximation empirique basée sur des scénarios simulés :

$$\mathbb{P}(F(u, \xi) \geq z) \approx \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{F(u, \xi_i) \geq z}. \quad (3)$$

Elles constituent la catégorie la plus utilisée en pratique, mais souffrent de plusieurs limitations. On distingue plusieurs approches :

**Méthodes basées sur la réduction de variance** Les approches Monte Carlo ou de réduction de variance peuvent estimer des probabilités rares, mais elles sont en général conçues pour un vecteur de décision  $u$  fixe. Elles deviennent inefficaces lorsque  $u$  varie à chaque itération d'un algorithme d'optimisation.

**Importance Sampling** L'*importance sampling* permet de diriger l'échantillonnage vers les événements rares. Cependant, cette technique suppose souvent que les composantes de  $\xi$  sont indépendantes, une hypothèse souvent violée. De plus, sa performance se dégrade fortement lorsque la dimension augmente ou lorsque  $\alpha$  devient extrêmement petit.

**Approche par scénarios** L'approche par scénarios consiste à imposer que  $F(u, \xi_i) \leq z$  pour un ensemble de scénarios représentatifs. Cette méthode produit un problème déterministe, conservation de la convexité comprise. Cependant, le nombre de scénarios nécessaires explose à mesure que  $\alpha$  diminue. La croissance est de l'ordre de  $N = \mathcal{O}(\alpha^{-1})$ , ce qui rend cette approche impraticable pour des niveaux de risque très petits Shapiro et al. (2014).

**Méthodes elliptiques radiales** Pour certaines distributions elliptiques (comme la loi normale multivariée), des transformations radiales peuvent simplifier la contrainte. Toutefois, ces méthodes deviennent extrêmement coûteuses en calcul et exigent des hypothèses structurelles fortes sur la distribution Lagoa et al. (2005); Nemirovski and Shapiro (2007).

**Troisième famille : méthodes de fiabilité (FORM/SORM)** Ces méthodes, couramment utilisées en ingénierie, reposent sur l'identification d'un « point de défaillance le plus probable » (*design point*). Elles utilisent une approximation locale de  $F(u, \xi)$  autour de ce point et une linéarisation ou quadratrisation du domaine de défaillance.

Bien qu'efficaces pour des distributions gaussiennes ou presque gaussiennes, ces méthodes :

- manquent de précision pour des événements très rares ( $\alpha < 10^{-4}$ ),
- supposent souvent une structure trop rigide de la distribution,
- s'adaptent mal à des distributions multimodales ou lourdes,
- peuvent échouer dans des dimensions élevées.

### 2.1.2 Synthèse des limitations de la littérature

Face à ces limites, les travaux de Tong et al. (2022) dans leur article *Optimization under Rare Chance Constraints* proposent une approche dite « *sampling free* » (sans échantillonnage) pour la résolution des OCCR. L'idée principale est d'éviter complètement l'échantillonnage massif en s'appuyant sur des développements récents de la **Théorie des Grandes Déviations (LDT)**. La probabilité rare est alors remplacée par une approximation asymptotique explicite, dépendant uniquement :

- du minimiseur LDT (*point dominant*),
- des dérivées locales de la fonction  $F(u, \xi)$ ,

- et d'une distribution flexible de  $\xi$  (modélisée comme mélange gaussien).

Cette approche contourne à la fois la malédiction de la dimension et l'explosion en  $\alpha^{-1}$ , ce qui la distingue des méthodes précédentes et en fait une contribution originale pour les problèmes d'optimisation sous contraintes en chance extrêmes.

## 2.2 L'Intractabilité des Contraintes en Chance Rares : Généralités

La méthodologie « *sampling-free* » est fondée sur la Théorie des Grandes Déviations (LDT), qui transforme le problème difficile de l'estimation des probabilités très rares en un problème d'optimisation convexe.

### 2.2.1 Décroissance Exponentielle et Échec des Méthodes Classiques

Le LDT part de l'observation que la probabilité d'un événement extrême  $\mathcal{A}$  ne s'annule pas de manière arbitraire, mais à un taux exponentiellement rapide en fonction d'un paramètre d'échelle. Ce principe, pour une séquence de variables aléatoires  $X_n$  et un événement rare  $\mathcal{A}$ , stipule que la probabilité de  $\mathcal{A}$  peut être estimée par :

$$\mathbb{P}(X_n \in \mathcal{A}) \approx \exp(-n \cdot I(\mathcal{A}))$$

ou, plus formellement, par une limite logarithmique :

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \mathbb{P}(X_n \in \mathcal{A}) = -I(\mathcal{A})$$

où :

- $n$  est un facteur d'échelle (souvent lié à l'inverse de la probabilité  $\alpha$  ou à la dimension). Plus  $n$  est grand, plus la décroissance est rapide,
- $I(\mathcal{A})$  est la fonction de taux (ou rate function). Elle est le taux auquel la probabilité décroît.

**Justification de l'Échec de Monte Carlo** Cette nature de décroissance exponentielle est la raison fondamentale pour laquelle les méthodes d'échantillonnage classiques, comme Monte Carlo (MC) ou SAA, échouent dans le régime rare :

- **Exigence Numérique** : Pour garantir qu'une estimation de probabilité  $\hat{\alpha}$  soit proche de la vraie probabilité  $\alpha$  avec une certaine confiance, le nombre d'échantillons  $N$  requis est de l'ordre de  $\mathcal{O}(\alpha^{-1})$  Shapiro et al. (2014).
- **Intractabilité** : Si le niveau de risque ciblé est très faible, par exemple  $\alpha = 10^{-6}$  (un risque critique en finance ou ingénierie), l'optimisation nécessiterait des millions d'échantillons. Cette explosion dimensionnelle rend le problème d'optimisation computationnellement impraticable à chaque itération.

La LDT contourne cette difficulté en transformant le problème de l'estimation de la probabilité minuscule  $\mathbb{P}(\mathcal{A})$  en un problème d'optimisation de son exposant  $I(\mathcal{A})$ , qui est beaucoup plus stable et facile à calculer.

### 2.2.2 La Fonction de Taux ( $I(\mathcal{A})$ ) et l'Entropie Relative

La fonction de taux  $I(\mathcal{A})$  est la pierre angulaire de la LDT. Elle est bien plus qu'un simple exposant : elle est directement liée à un concept d'information et d'entropie.

**Définition Formelle** La fonction de taux  $I(\xi)$  est mathématiquement définie comme la transformée de Legendre-Fenchel de la fonction génératrice des cumulants (Théorème 3.2 de l'article de Tong et al. (2022) qui énonce les conditions d'application du LDP).

**Le Parallèle avec l'Entropie Relative** La fonction de taux  $I(\xi)$  est égale à la divergence de Kullback-Leibler (DKL) minimale entre la distribution d'origine  $P$  de l'incertitude  $\xi$  et toute distribution alternative  $\tilde{P}$  dont l'espérance est  $\xi$ .

$$I(\xi) = \min_{\tilde{P} \text{ t.q. } \mathbb{E}_{\tilde{P}}[\xi] = \xi} D(\tilde{P} \| P)$$

où  $D(\tilde{P} \| P)$  est l'entropie relative.

### Interprétation physique et mathématique

- **L'Entropie Relative (DKL) :** La DKL mesure la "distance" ou la quantité d'information perdue lorsque l'on utilise une distribution  $\tilde{P}$  pour modéliser une distribution réelle  $P$ .
- **Signification:** Le taux de décroissance  $I(\xi)$  représente le coût minimal d'information (ou coût entropique) pour que la variable aléatoire  $\xi$  s'écarte de son comportement normal (régi par  $P$ ) pour atteindre la valeur extrême  $\xi^*$ .

### Avantages Majeurs

- **Convexité :** La divergence de Kullback-Leibler est une fonction convexe. En transformant le problème de probabilité non convexe en un problème de minimisation d'entropie, la LDT permet de travailler avec une fonction objectif (le taux  $I(\xi)$ ) qui est convexe.
- **Tractabilité:** Les Théorèmes 3.3 et 4.1 garantissent que cette fonction de taux peut être exprimée sous une forme analytique, même pour des distributions complexes comme les Mélanges Gaussiens, ce qui permet la résolution du problème d'optimisation sans échantillonnage.

#### 2.2.3 Le Point Dominant ( $\xi^*$ )

L'utilisation de la LDT dans l'optimisation s'articule autour de la recherche du Point Dominant ( $\xi^*$ ).

Le point dominant  $\xi^*$  est le scénario d'incertitude qui, parmi tous les scénarios causant la défaillance ( $F(u, \xi) \geq z$ ), minimise la fonction de taux  $I(\xi)$  :

$$\xi^*(u) = \arg \min_{\xi: F(u, \xi) \geq z} I(\xi)$$

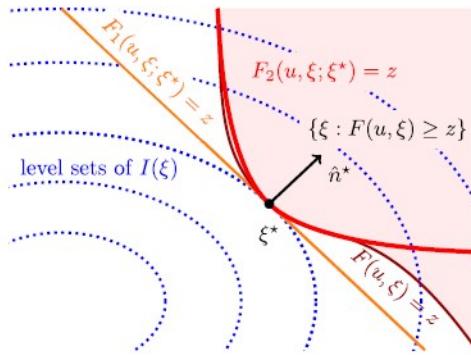


Figure 1: Visualisation du Point Dominant ( $\xi^*$ ) dans l'espace des incertitudes. Les lignes de niveau représentent la fonction de taux  $I(\xi)$  (distance entropique).  $\xi^*$  est le point qui minimise  $I(\xi)$  tout en satisfaisant la contrainte de défaillance ( $F(u, \xi) \geq z$ ) **source :** Tong et al. (2022)

**Interprétation** Puisque  $I(\xi)$  est le coût entropique,  $\xi^*$  est le scénario de crise qui se trouve à la distance d'entropie relative la plus courte de la distribution normale du marché. Par conséquent c'est le scénario de défaillance le moins improbable : celui qui nécessite la plus petite "pression" ou la plus petite déviation de la distribution normale pour que la contrainte  $F(u, \xi) \geq z$  soit violée. Il domine tous les autres événements de défaillance en termes de probabilité. En trouvant  $\xi^*$ , la méthode LDT identifie et caractérise le pire scénario crédible, permettant ainsi de reformuler la contrainte de chance rare.

**Lien avec la Méthode de Fiabilité** L’approche de Tong et al. (2022) utilise ce point  $\xi^*$  comme le point de défaillance le plus probable (design point), établissant ainsi que leur méthode est une méthode de fiabilité dont le point de départ est trouvé de manière optimale et analytique par la LDT.

#### 2.2.4 Amélioration de l’Estimation par Linéarisation Locale (FORM/SORM)

Une fois  $\xi^*$  trouvé, la méthode utilise la linéarisation locale de type FORM pour calculer l’estimation finale de la probabilité en exploitant les propriétés Gaussiennes. Cela corrige l’imprécision inhérente aux approximations purement asymptotiques de la LDT pour des valeurs  $\alpha$  qui ne sont pas infiniment petites.

**Rôle des Développements de Taylor** Dans l’approche de Tong et al., l’objectif est de transformer la contrainte de chance rare :

$$\mathbb{P}(F(u, \xi) \geq z) \leq \alpha$$

en une contrainte tractable. Pour cela, on utilise le point dominant  $\xi^*$  fourni par la LDT comme le point idéal pour l’approximation.

- **Développement au Premier Ordre (FORM) :** Le développement de Taylor au premier ordre (FORM) linéarise la fonction de contrainte  $F(u, \xi)$  autour du point dominant  $\xi^*$ .

$$F(u, \xi) \approx F(u, \xi^*) + \nabla_{\xi} F(u, \xi^*) \cdot (\xi - \xi^*)$$

Ce choix est fait pour la tractabilité. Le problème d’optimisation global utilise cette approximation linéaire, car elle préserve la structure de quasi-convexité nécessaire à la résolution par des optimiseurs standards. Le FORM permet de calculer l’estimation de la probabilité de défaillance via l’indice de fiabilité ( $\beta$ ) qui sert de mesure de la distance (en écarts-types) entre le point dominant et le seuil critique  $z$  dans l’espace des variables aléatoires.

- **Développement du Second Ordre (SORM) : Amélioration de la Précision** Le développement de Taylor du second ordre (SORM) inclut la courbure de la fonction de contrainte  $F(u, \xi)$  grâce à la matrice Hesseenne ( $\nabla^2 F$ ).

$$F(u, \xi) \approx F(u, \xi^*) + \nabla_{\xi} F(u, \xi^*) \cdot (\xi - \xi^*) + \frac{1}{2}(\xi - \xi^*)^T \nabla_{\xi}^2 F(u, \xi^*)(\xi - \xi^*)$$

Son rôle est d’améliorer la précision de l’estimation de la probabilité  $\mathbb{P}(\mathcal{A})$  en capturant la courbure de la surface de défaillance, surtout lorsque  $F$  est fortement non-linéaire. Le SORM est généralement utilisé a posteriori pour affiner la solution, et non pour définir la contrainte dans le cœur du problème d’optimisation, afin de ne pas compromettre la convexité.

**Exploitation des Propriétés Gaussiennes** En modélisant  $\xi$  comme un mélange gaussien, la linéarisation locale de  $F(u, \xi)$  autour de  $\xi^*$  ramène le calcul de la probabilité à la détermination d’une probabilité d’un événement sous une distribution Gaussienne (ou quasi-Gaussienne). Le calcul de l’estimation de la probabilité se fait alors analytiquement, sans simulation, en utilisant la fonction de répartition de la loi normale centrée réduite ( $\Phi$ ) :

$$\mathbb{P}(F(u, \xi) \geq z) \approx 1 - \Phi(\beta)$$

#### 2.2.5 Pseudo-Code de Résolution OCCR par l’Approche LDT

L’approche « sampling free » de l’Optimisation Sous Contrainte en Chance Rare (OCCR) de Tong et al. (2022) repose sur une séquence d’étapes qui transforment le problème probabiliste intractables en un problème d’optimisation bi-niveau et analytique. Le processus se déploie comme suit :

1. **Problème Initial (OCCR) :** Définir le problème d’optimisation avec la contrainte en chance rare :

$$\min_{u \in \mathcal{U}} C(u) \quad \text{s.c.} \quad \mathbb{P}(F(u, \xi) \geq z) \leq \alpha$$

où  $u$  est le vecteur de décision,  $C(u)$  la fonction objectif,  $F(u, \xi)$  la fonction de contrainte, et  $\alpha$  le niveau de risque cible. Ce problème est, en l’état, intractable.

- 2. Substitution par la Théorie des Grandes Déviations (LDT) :** Utiliser l'approximation asymptotique de la LDT pour remplacer la probabilité rare.

$$\mathbb{P}(F(u, \xi) \geq z) \approx \exp(-I(\xi^*(u)))$$

où  $I(\xi)$  est la fonction de taux (l'entropie relative, ou Divergence de Kullback-Leibler) et  $\xi^*$  est le Point Dominant.

- 3. Résolution du Problème Interne (Point Dominant) :** Le cœur de la méthode consiste à résoudre le problème d'optimisation interne (ou problème bi-niveau) pour trouver le scénario de défaillance le moins improbable  $\xi^*$ . Ce problème est indépendant de l'échantillonnage :

$$\xi^*(u) = \arg \min_{\xi} \{I(\xi) \text{ s.c. } F(u, \xi) \geq z\}$$

- 4. Linéarisation Locale (FORM) :** Utiliser le  $\xi^*$  trouvé comme point de conception (\*design point\*) pour appliquer l'approximation d'ordre un (FORM). Cette étape rend la contrainte analytique pour les  $\alpha$  non asymptotiquement petits.

$$F(u, \xi) \approx F(u, \xi^*) + \nabla_{\xi} F(u, \xi^*)^T (\xi - \xi^*)$$

- 5. Problème Final Tractable :** Remplacer la contrainte OCCR par une contrainte analytique basée sur l'indice de fiabilité  $\beta(\xi^*)$ , dérivé de la LDT et de l'approximation FORM.

$$\min_{u \in \mathcal{U}} C(u) \text{ s.c. } \beta(\xi^*(u)) \geq \beta_{\text{cible}}$$

Ce problème final est soluble par des algorithmes d'optimisation non linéaire standards.

## 2.3 Fondements Théoriques et Conditions de Validité

La validité de la reformulation asymptotique des contraintes en chance rares repose sur des résultats issus de la théorie des grandes déviations (Large Deviations Theory, LDT). Nous présentons ici trois théorèmes fondamentaux qui garantissent l'existence du minimiseur LDT, la précision de l'approximation asymptotique de la probabilité rare, ainsi que les conditions structurelles assurant la validité de cette reformulation, en particulier pour les distributions de type mélange gaussien.

### 2.3.1 Théorème 3.2 : Approximation LDT de la probabilité rare

Soit  $u \in U_0$ . Supposons que les hypothèses suivantes soient satisfaites :

(A1) (**Existence et unicité**) L'optimisation

$$\xi^*(u, z) = \arg \min_{\xi \in \Xi} \{I(\xi) : F(u, \xi) \geq z\}$$

possède un unique minimiseur pour tout  $z \geq z_0$ .

(A2) (**Comportement asymptotique**) La fonction  $\xi^*(u, z)$  est continûment différentiable et l'on a, pour  $z \rightarrow \infty$ ,

$$\|\xi^*(u, z)\| \rightarrow \infty, \quad \|\nabla_{\xi} F(u, \xi^*(u, z))\| \geq K_0 > 0.$$

(A3) (**Condition de domination**) Pour tout  $\varepsilon > 0$ ,

$$\{\xi : F(u, \xi) \geq z\} \subseteq \{\xi : \langle \nabla_{\xi} F(u, \xi^*(u, z)), \xi - \xi^*(u, z) \rangle \geq 0\}^{\varepsilon}.$$

(A4) (**Condition intégrale**) Il existe  $K_1 > 0$  tel que

$$\lim_{z \rightarrow \infty} \log \int_{A(u, z, K_1)} \exp(I(\xi^*(u, z)) + \langle \nabla_{\xi} F(u, \xi^*(u, z)), \xi - \xi^*(u, z) \rangle) p(\xi) d\xi = 0.$$

**Conclusion du théorème.** Sous les hypothèses (A1)–(A4), on obtient :

$$\lim_{z \rightarrow \infty} \frac{\log \mathbb{P}(F(u, \xi) \geq z)}{I(\xi^*(u, z))} = -1.$$

Autrement dit,

$$\mathbb{P}(F(u, \xi) \geq z) \approx \exp(-I(\xi^*(u, z))), \quad z \rightarrow \infty.$$

Ce résultat justifie le rôle central de  $\xi^*(u, z)$  comme point dominant (minimiseur LDT).

### 2.3.2 Théorème 3.3 : Conditions suffisantes pour le Théorème 3.2

Soit  $\Xi = \mathbb{R}^n$  et supposons que la fonction  $F(u, \xi)$  soit concave en  $\xi$ .

**Condition principale.**

$$\lim_{z \rightarrow \infty} \frac{\log \mathbb{P}(F(u, \xi) \geq z)}{I(\xi^*(u, z))} \leq -1.$$

**Conditions additionnelles.** Le Théorème 3.2 est vérifié si, de plus :

- (C1)  $F(u, \xi)$  est concave en  $\xi$ .
- (C2)  $\nabla_\xi F(u, \xi)$  est Lipschitz continu sur  $\Xi$ .
- (C3) Pour tout  $r > 0$ ,

$$\log \int_{B(\xi^*, r)} \exp(I(\xi^*) + \langle \nabla I(\xi^*), \xi - \xi^* \rangle) p(\xi) d\xi = o(I(\xi^*)) \quad \text{lorsque } \|\xi^*\| \rightarrow \infty,$$

où  $B(\xi^*, r)$  désigne une boule fermée centrée en  $\xi^*$ .

**Conclusion.** Sous (C1)–(C3), l'approximation asymptotique du Théorème 3.2 est rigoureusement validée.

### 2.3.3 Théorème 4.1 : Validité pour les mélanges gaussiens

Supposons que la variable aléatoire  $\xi$  suive une distribution *mélange gaussien*

$$\xi \sim \sum_{k=1}^M w_k \mathcal{N}(\mu_k, \Sigma_k), \quad w_k > 0, \quad \sum_{k=1}^M w_k = 1.$$

**Conclusion du théorème.** Dans ce cas, la condition (C3) du Théorème 3.3 est automatiquement satisfaite.

**Interprétation.** Les mélanges gaussiens constituent donc une classe de distributions pour laquelle les conditions de validité de l'approximation LDT sont garanties, sans restrictions supplémentaires.

## 2.4 Littérature Empirique : Optimisation et Risque du Secteur de l'Intelligence Artificielle

L'application des méthodes d'optimisation aux marchés boursiers est un domaine classique, mais son croisement avec la gestion du risque de queue des entreprises technologiques à forte croissance (l'Intelligence Artificielle) est une aire de recherche plus récente et hautement pertinente.

### 2.4.1 Optimisation et Marché de l'Intelligence Artificielle (Général)

Les travaux dans cette catégorie utilisent l'Intelligence Artificielle (IA) principalement comme un outil pour améliorer les paramètres d'optimisation de portefeuille ou pour définir la stratégie d'allocation elle-même.

- **Allocation par Apprentissage Profond (Deep Learning) :** Une partie significative de la littérature utilise des réseaux neuronaux récurrents (RNN) ou des mémoires à long terme (LSTM) pour prédire les rendements attendus ( $\mu$ ) ou la matrice de covariance ( $\Sigma$ ). Ces prédictions IA sont ensuite insérées dans des modèles d'optimisation classiques. Par exemple, Chen et al. (2020) ont montré l'efficacité des réseaux LSTM pour capturer la dynamique non linéaire des marchés financiers, offrant des paramètres d'entrée plus précis que les modèles historiques simples.
- **Optimisation Directe par Apprentissage par Renforcement (RL) :** Les modèles RL représentent une approche plus radicale. L'agent RL apprend directement la stratégie d'allocation optimale en fonction des états du marché. Des auteurs comme Jiang et al. (2017) ont exploré l'utilisation d'algorithmes RL pour gérer des portefeuilles avec une efficacité remarquable, contournant la nécessité de modèles analytiques intermédiaires.

### 2.4.2 Optimisation en Chance Rare et AI (Gestion du Risque Extrême)

Les méthodes d'optimisation en chance rare (OCCR) ou les techniques connexes de gestion du risque de queue sont appliquées à l'écosystème IA pour mitiger les pertes extrêmes, un risque accru par la concentration et la volatilité du secteur.

- **Approches Basées sur la CVaR :** La gestion du risque de queue en finance est historiquement dominée par la Valeur à Risque Conditionnelle (CVaR), rendue populaire par Rockafellar and Uryasev (2000). Cette mesure est couramment utilisée pour construire des portefeuilles robustes contre les chocs.
- **Apprentissage Automatique Robuste (Robust ML) et OCCR :** Le lien théorique entre OCCR et la LDT est également utilisé pour l'apprentissage automatique robuste. Des travaux comme ceux de Xu et al. (2009) ont utilisé les principes de la programmation stochastique robuste pour rendre les modèles d'apprentissage automatique plus résilients face aux données extrêmes, partageant la même fondation théorique que la méthode utilisée ici.

### 2.4.3 Caractère Unique de la Contribution et Gestion du Risque d'Effondrement

Notre travail se distingue des approches citées par la fusion méthodologique qu'il opère et par la nature précise du risque géré.

**Fusion Méthodologique LDT-Fiabilité** Contrairement aux approches basées sur l'échantillonnage ou aux approximations CVaR, notre méthode se fonde sur le travail de Tong et al. (2022), qui combine :

1. La modélisation de l'incertitude  $\xi$  avec une distribution flexible (Mélange Gaussien).
2. L'utilisation de la Théorie des Grandes Déviations (LDT) pour identifier le scénario de crise le moins improbable ( $\xi^*$ ) par minimisation d'entropie.
3. La reformulation analytique grâce à l'approximation locale (FORM), évitant l'explosion dimensionnelle.

**Gestion Directe vs. Indirecte du Risque d’Effondrement Sectoriel** L’approche de gestion du risque est subtilement double :

- **Approche Indirecte (Optimisation de la Conséquence)** : Mathématiquement, la contrainte en chance ( $\mathbb{P}(F(u, \xi) \geq z) \leq \alpha$ ) vise à optimiser la conséquence de la défaillance (limiter la probabilité de la perte  $\alpha$ ), plutôt que de modéliser directement la cause (les fondamentaux des entreprises d’IA).
- **Approche Directe (Identification et Défense du Scénario)** : L’analyse LDT confère une précision chirurgicale. Elle identifie le Point Dominant  $\xi^*$ , le scénario de rendements extrêmes le moins improbable qui cause la perte critique. L’optimisation est alors forcée de se défendre directement contre ce  $\xi^*$ . En se concentrant sur les actions pilier de l’IA, le  $\xi^*$  identifié est le scénario qui ressemble le plus à un effondrement sectoriel crédible.

Bien que la littérature ait exploré l’optimisation de portefeuille avec des outils d’IA Chen et al. (2020); Jiang et al. (2017) et la gestion du risque extrême avec des techniques robustes Rockafellar and Uryasev (2000); Xu et al. (2009), aucun travail n’a combiné une méthodologie aussi rigoureuse que l’OCCR basée sur la LDT de Tong et al. (2022) à l’étude spécifique et urgente de la gestion du risque de queue inhérent au marché de l’Intelligence Artificielle lui-même. Notre travail comble cette lacune en appliquant une solution sampling-free et théoriquement robuste à un problème de marché contemporain et critique

### 3 DÉMARCHE MÉTHODOLOGIQUE

#### 3.1 Choix des actifs et justification : les six piliers de l’écosystème IA

##### 3.1.1 Contexte général

L’écosystème de l’intelligence artificielle est structuré autour d’un ensemble d’entreprises dont le rôle est central dans la chaîne de valeur : semi-conducteurs, centres de données, cloud hyperscale, modèles fondamentaux, infrastructures d’interconnexion, et systèmes de gestion de données. Ces entreprises forment un ensemble interdépendant dont la croissance est rapide, mais dont la fragilité est accrue par la concentration technologique et financière du secteur.

##### 3.1.2 Définition des piliers IA

Nous définissons comme *piliers IA* les entreprises satisfaisant les trois critères suivants :

- un rôle indispensable dans le fonctionnement de l’écosystème IA (calcul, cloud, interconnexion, modèles, données) ;
- une sensibilité directe au cycle d’investissement IA, reflétée dans leur dynamique boursière ;
- une importance systémique telle qu’une défaillance pourrait générer un effet de contagion au reste du secteur.

Ces entreprises constituent les points de dépendance critiques de la chaîne de valeur de l’intelligence artificielle moderne.

##### 3.1.3 Critères de sélection

La sélection des actifs repose sur trois critères principaux :

- **Critère technologique** : contribution essentielle à au moins un segment structurants de l’IA (GPU, cloud, modèles, infrastructures).
- **Critère de sensibilité au cycle IA** : capacité de refléter les expansions et contractions du cycle IA.
- **Critère systémique** : impact significatif sur les autres acteurs en cas de choc.

### 3.1.4 Présentation des six titres retenus

Le portefeuille étudié se compose des six entreprises suivantes :

- **Nvidia (NVDA)** : leader technologique absolu des GPU IA, le “sauveur psychologique” du marché (sa performance soutient les indices.)
- **Microsoft (MSFT)** : pilier systémique, partenaire stratégique d’OpenAI et acteur majeur du cloud IA.
- **AMD (AMD)** : challenger technologique de Nvidia, actif à forte volatilité et amplificateur des cycles (monte plus vite en phase bull, chute plus vite en phase bear).
- **Meta (META)** : leader de l’IA open source via Llama, investisseur massif en infrastructures GPU.
- **Broadcom (AVGO)** : acteur clé de l’infrastructure profonde (interconnexions, ASIC hyperscale).
- **Oracle (ORCL)** : pilier cloud d’entreprise, exposé à la demande réelle IA et baromètre du cycle dû au fait qu’il finance son expansion entièrement par la dette.

### 3.1.5 Positionnement des six piliers IA parmi les grands acteurs technologiques

Les six actifs retenus se distinguent d’autres entreprises technologiques également impliquées dans l’IA (Alphabet, Amazon, TSMC, Cisco, Marvell, etc.) par une combinaison unique de centralité technologique, sensibilité cyclique et importance systémique. Les entreprises retenues :

- jouent un rôle direct dans l’ entraînement, l’inférence ou l’hébergement de modèles IA ;
- réagissent fortement aux fluctuations du cycle d’investissement IA ;
- transmettent rapidement les chocs au reste de l’écosystème ;
- représentent des points de fragilité potentielle dans un scénario de crash.

Ainsi, notre portefeuille n’est pas un panier générique de valeurs technologiques, mais un ensemble ciblé d’acteurs structurels dont l’analyse est pertinente pour l’étude des risques de queue.

## 3.2 Modèle mathématique, hypothèses et formulations LDT

Dans cette section, nous présentons le modèle mathématique utilisé pour l’optimisation du portefeuille sous contrainte en chance rare, en suivant fidèlement les formulations explicites développées dans les Sections 4 et 5 de Tong et al. (2022)).

### 3.2.1 Distribution du vecteur aléatoire

L’incertitude du marché est modélisée par un vecteur aléatoire

$$\xi \in \mathbb{R}^m,$$

suivant un *mélange gaussien*

$$\xi \sim \sum_{j=1}^M w_j \mathcal{N}(\mu_j, \Sigma_j), \quad w_j > 0, \quad \sum_{j=1}^M w_j = 1.$$

Cette classe de distributions est essentielle, car par le Théorème 4.1 de Tong et al. (2022) :

*La condition (C3) nécessaire au Théorème 3.3 (validité LDT) est automatiquement satisfaite par tout mélange gaussien.*

Le modèle reste donc mathématiquement cohérent pour l’utilisation d’approximation LDT de la probabilité rare.

### 3.2.2 Fonction de perte et contrainte en chance

Nous considérons une perte linéaire :

$$F(u, \xi) = -u^\top \xi,$$

où  $u \in \mathbb{R}^m$  est le vecteur des pondérations du portefeuille, soumis aux contraintes usuelles :

$$u \geq 0, \quad \mathbf{1}^\top u = 1.$$

La contrainte en chance prend la forme :

$$\mathbb{P}(F(u, \xi) \geq z) \leq \alpha, \quad \alpha \ll 1.$$

Suivant la section 5.1, cette contrainte est directement équivalente à

$$\mathbb{P}\left(\sum_{i=1}^m u_i \xi_i \leq -z\right) \leq \alpha.$$

### 3.2.3 Point dominant LDT

Pour chaque valeur de  $u$  et  $z$ , le point dominant est défini par

$$\xi^*(u, z) = \arg \min_{\xi \in \mathbb{R}^m} \left\{ I(\xi) : F(u, \xi) \geq z \right\},$$

où  $I(\xi)$  est la fonction de taux du mélange gaussien (voir (4.1)–(4.3) de Tong et al. (2022)) :

$$I(\xi) = \min_{j=1, \dots, M} \frac{1}{2} \|\xi - \mu_j\|_{\Sigma_j^{-1}}^2.$$

Ces points minimiseurs existent et sont uniques pour  $z$  grand, conformément au Théorème 3.2.

### 3.2.4 Approximation de premier ordre $P_1$

Le premier estimateur (Section 4.2) de la probabilité rare est :

$$P_1(u, z, \xi^*) \approx \sum_{j=1}^M w_j \Phi\left(-\frac{(\nabla_\xi F(u, \xi^*))^\top (\xi^* - \mu_j)}{\|\Sigma_j^{1/2} \nabla_\xi F(u, \xi^*)\|}\right), \quad (\text{P1})$$

où  $\Phi$  désigne la fonction de répartition de la loi normale standard.

### 3.2.5 Approximation de second ordre $P_2$

L'estimateur de second ordre (Section 4.3–4.4) est :

$$P_2(u, z, \xi^*) \approx \sum_{j=1}^M w_j \Phi\left(-\|\xi^* - \mu_j\|_{\Sigma_j^{-1}} \det(H_j)^{1/2}\right), \quad (\text{P2})$$

où

$$H_j = I_n - \frac{(\xi^* - \mu_j)(\xi^* - \mu_j)^\top \Sigma_j^{-1}}{(\xi^* - \mu_j)^\top \Sigma_j^{-1} \nabla_\xi F(u, \xi^*)}$$

est la matrice de Hessien préconditionnée caractéristique dans la formule (4.8).

### 3.2.6 Formulation LDT du problème d'optimisation

Suivant l'équation (5.1), l'optimisation du portefeuille sous contrainte en chance rare s'écrit :

$$\begin{aligned} & \max_{u \in \mathbb{R}^m, z \in \mathbb{R}} && z \\ & \text{s.c.} && \mathbf{1}^\top u = 1, \\ & && u \geq 0, \\ & && P_k(u, z, \xi^*(u, z)) \leq \alpha, \quad k \in \{1, 2\}. \end{aligned} \quad (\text{LDT-CC})$$

Le choix entre  $k = 1$  (premier ordre) ou  $k = 2$  (second ordre) correspond à l'arbitrage entre rapidité et précision :

- $P_1$  est rapide, mais moins précis pour des événements extrêmement rares ;
- $P_2$  est plus précis et reste numériquement stable pour  $\alpha \leq 10^{-4}$ .

Pour des raisons computationnelles et de reproductibilité, nous retenons exclusivement l'approximation  $P_1$  (FORM).

### 3.3 Résultats attendus

L'application de l'approximation de premier ordre issue de la théorie des grandes déviations (LDT) au problème d'optimisation sous contrainte en chance doit permettre d'obtenir des résultats quantitatifs et qualitatifs cohérents avec les propriétés des événements rares, tout en restant computationnellement légers. Les résultats attendus se déclinent comme suit.

#### 3.3.1 Estimation efficace de la probabilité rare

En utilisant uniquement l'approximation de premier ordre

$$P_1(u, z, \xi^*) = \Phi\left(-\frac{\|\xi^* - \mu\|_{\Sigma^{-1}}}{\|\Sigma^{1/2} \nabla_\xi F(u, \xi^*)\|}\right),$$

nous nous attendons à obtenir une estimation raisonnablement précise de  $\mathbb{P}(F(u, \xi) \geq z)$  pour des niveaux rares tels que  $\alpha = 10^{-3}$  ou  $10^{-4}$ , tout en évitant l'utilisation de méthodes coûteuses comme le Monte Carlo. L'objectif n'est pas d'obtenir une précision asymptotique parfaite, mais une approximation suffisante pour permettre une optimisation fiable. Pour  $\alpha = 10^{-3}$ , on s'attend à un  $\xi^*$  éloigné de 3 à 4 écarts-types du centre.

#### 3.3.2 Problème d'optimisation tractable

La contrainte en chance

$$\mathbb{P}(F(u, \xi) \geq z) \leq \alpha$$

peut être remplacée par la contrainte déterministe

$$P_1(u, z, \xi^*) \leq \alpha.$$

Cette reformulation génère un programme d'optimisation lisse et de très faible dimension, qui sera résolu avec Ipopt ou Gurobi selon la nature de la formulation.

Le temps de calcul attendu est de l'ordre de la seconde, ce qui rend la méthode adaptée à une implémentation sur une machine standard.

#### 3.3.3 Caractérisation du scénario extrême

Le point dominant  $\xi^*(u, z)$ , obtenu comme solution du problème de taux de Cramér, fournit une interprétation structurelle du risque extrême. Il permet d'identifier :

- les actifs qui contribuent le plus au risque de queue ;
- la direction principale du stress dans l'espace des rendements ;
- le type de choc extrême qui viole la contrainte.

Ce point dominant joue ainsi le rôle d'un *stress test endogène* directement issu du modèle.

#### 3.3.4 Structure du portefeuille optimal

Nous nous attendons à ce que la solution optimale ( $u^*$ ) :

- respecte la contrainte de risque extrême fixée par  $\alpha$  ;
- présente une diversification accrue par rapport au portefeuille naïf ;
- surpondère les actifs aux queues légères et aux variances modérées ;
- sous-pondère les actifs présentant les comportements extrêmes les plus prononcés.

Le portefeuille obtenu devrait ainsi être plus robuste aux chocs rares tout en conservant une performance raisonnable.

### 3.3.5 Comparaison qualitative avec les méthodes classiques

Sans reproduire les expériences numériques complètes de Tong et al., nous pouvons anticiper une comparaison qualitative simple :

- les méthodes Monte Carlo deviennent impraticables pour  $\alpha = 10^{-4}$  en raison du nombre d'échantillons nécessaire ( $N \approx 10^7$ ) ;
- l'approximation LDT de premier ordre fournit une estimation exploitable en temps réel ;
- l'approche LDT permet de stabiliser l'optimisation sous contrainte en chance sans explosion de dimension.

En résumé, les résultats attendus devraient suggérer que la méthode LDT de premier ordre constitue une alternative efficace, stable et computationnellement légère pour traiter les contraintes en chance dans le régime des événements rares.

## 4 RÉSULTATS EMPIRIQUES

Afin de réaliser les objectifs que nous nous sommes fixés dans ce projet, nous avons décidé de diviser cette section en deux parties.

- **Évaluation du risque:** Nous évaluons le risque de pertes extrêmes d'un portefeuille donné afin d'illustrer l'efficacité de la théorie des grandes déviations dans l'estimation des probabilités rare. Plus précisément, nous comparons l'approximation des grandes déviations de premier ordre ( $P_1$ ) à la méthode de Monte Carlo classique basée sur l'approximation par échantillonage (SAA). Cette comparaison a pour but de répondre à la question suivante : À quel point les estimateurs de l'approximation LDT de premier ordre  $P_1$ , du SAA de Monte Carlo  $P_N^{MC}$  approchent-ils la vraie probabilité de queue pour un portefeuille donné?
- **Optimisation sous contrainte en chances :** Afin de mettre en évidence l'apport spécifique d'une optimisation fondée sur la géométrie des événements rares, nous allons exploiter le cadre des grandes déviations pour résoudre un problème d'allocation de portefeuille sous contrainte en chance rare. Nous comparons plusieurs stratégies d'allocation (portefeuille équivalant, portefeuille minimum variance et portefeuille maximum rendement) à une optimisation explicite orientée vers le risque de pertes extrêmes.

Nos formulations, ainsi que les contraintes non linéaires issues de l'approximation des grandes déviations, traitées via la différentiation automatique, sont implémentées en **Julia** à l'aide du package JuMP. Les problèmes d'optimisation sous contrainte en chances rares sont résolus numériquement à l'aide du solveur Ipopt, avec les options par défaut.

### 4.1 Modélisation et description des données

Soit un univers composé de  $n$  actifs représentatifs de l'écosystème de l'intelligence artificielle (IA) : *AMD*, *NVDA*, *META*, *AMZN*, *GOOGL*, *MSFT*, *ORCL*. Les rendements sont observés à une fréquence hebdomadaire sur la période 2018-2024. L'horizon de détention est fixé à  $T = 1$  (une période hebdomadaire).

Nous considérons les rendements sous la forme logarithmique. Ainsi, pour un actif  $i$  donné, le rendement à la date  $t$  est défini par :

$$r_t = \log(P_t) - \log(P_{t-1}),$$

où  $P_t$  désigne le prix ajusté observé à la fin de la semaine  $t$ .

### 4.2 Portefeuille et pondérations

Soit un portefeuille caractérisé par un vecteur de pondération  $u = (u_1, \dots, u_n) \in \mathbb{R}^n$  où  $u_i$  représente la fraction de richesse investie dans l'actif  $i$ . Ces pondérations satisfont les contraintes suivantes :

- $u_i \geq 0$
- $\sum_{i=1}^n u_i = 1$ .

#### 4.2.1 Modélisation des rendements

Soit  $\xi \in \mathbb{R}^n$  le vecteur aléatoire des rendements logarithmiques hebdomadaires des actifs. Nous faisons l'hypothèse suivante :

$$\xi \sim \mathcal{N}(\mu, \Sigma),$$

où :

- $\mu \in \mathbb{R}^n$  est le vecteur des rendements moyens,
- $\Sigma \in \mathbb{R}^{n \times n}$  est la matrice de covariance

Dans ce présent travail, nous retenons la gaussienne multivariée ( sous famille d'un mélange gaussien) comme modèle de référence pour les raisons suivantes :

- La théorie des grandes déviations admet des expressions fermées et interprétables dans le cas gaussien
- L'estimation du point dominant et la résolution du problème d'optimisation sous contrainte en chance sont nettement plus robustes sous une loi gaussienne simple
- Nous avons volontairement choisi d'alléger notre cadre de modélisation.

#### 4.2.2 Valeur du portefeuille et variable de perte

La valeur du portefeuille en rendement est donnée par :

$$R(u) = u^\top \xi$$

La perte du portefeuille est donnée par :

$$L(u) = -u^\top \xi$$

On définit un événement de perte extrême  $A_c$  par :

$$A_n := \{L(u) \geq c\}$$

pour un seuil  $c > 0$ .

#### 4.2.3 Méthodologie de shrinkage Ledoit–Wolf

Soit l'estimation de la matrice empirique :

$$\hat{\Sigma}_{\text{emp}} = \frac{1}{T-1} \sum_{t=1}^T (r_t - \bar{r})(r_t - \bar{r})^T$$

Cette estimation a plusieurs problèmes connus :

- lorsque  $T$  n'est pas grand, le bruit est élevé
- Les valeurs propres de la matrice sont instables,
- La matrice peut être parfois mal conditionnée,
- L'inversion numérique devient fragile,

Dans notre univers d'IA très corrélé, ces problèmes se trouvent amplifiés ce qui justifie l'idée du shrinkage.

On construit donc une combinaison convexe :

$$\Sigma_{\text{LW}} = (1 - \lambda)\hat{\Sigma}_{\text{emp}} + \lambda\Sigma_{\text{cible}}$$

où :

- $0 \leq \lambda \leq 1$  est le paramètre de shrinkage,

- Les valeurs propres de la matrice sont instables,
- $\Sigma_{\text{cible}}$  est la matrice simple, bien conditionnée

Dans le cas standard Ledoit and Wolf (2004), la cible est :

$$\Sigma_{\text{cible}} = \bar{\sigma}^2 I_n$$

avec :

$$\bar{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \hat{\Sigma}_{ii}$$

Ledoit and Wolf (2004) fournit une formule fermé pour estimer  $\lambda$  de manière optimale en minimisant l'erreur quadratique :

$$\mathbb{E} \left[ \|\Sigma_{\text{LW}} - \Sigma\|_F^2 \right]$$

nous fournissant un  $0 < \lambda < 1$ .

Avec la méthodologie de Ledoit-Wolf, on est assuré que :

- les valeurs propres extrêmes sont régularisées,
- la matrice devient inversible,
- le risque de sur-apprentissage des corrélations bruitées est réduit
- l'optimisation sous contrainte en chance devient numériquement stable

Cependant :

- la structure de dépendance se trouve atténué
- on risque de sous estimé le risque de que si utilisé sans prudence.

Ainsi, nous choisissons d'améliorer la stabilité numérique au prix d'une atténuation partielle de la dépendance extrême.

#### 4.2.4 Estimation de la moyenne $\mu$ et de la matrice de covariance $\Sigma$

Le vecteur des rendements moyens  $\mu$  et la matrice de covariance  $\Sigma$  sont estimés empiriquement à partir des rendements hebdomadaires observés au cours de la période considérée.

Dans le cadre de l'analyse des événements à pertes extrêmes, le rôle de la moyenne est secondaire par rapport à celui de la matrice de covariance étant donné que ces événements sont principalement gouvernés par la structure de dépendance entre actifs. Par conséquent, les résultats sont peu sensibles à de légères variations de  $\mu$ .

Table 1: Statistiques des rendements logarithmiques des actifs IA (2018-2024)

Actif	$\mu$ (hebdo)	$\sigma$ (hebdo)	$\mu$ (annualisé)	$\sigma$ (annualisé)
NVDA	0.00590	0.06897	0.4941	0.4974
AMD	-0.00103	0.06988	-0.0534	0.5039
GOOGL	0.00191	0.04366	0.0992	0.3148
AMZN	0.00169	0.04785	0.0878	0.3451
MSFT	0.00190	0.03515	0.0986	0.2534
META	0.00418	0.06312	0.2171	0.4552
ORCL	0.00432	0.04357	0.2244	0.3142

### 4.3 Structure de corrélation des actifs IA

La Figure 2 présente la matrice de corrélation des rendements logarithmiques hebdomadaires des sept actifs représentant l'écosystème de l'intelligence artificielle sur la période 2018–2024. L'analyse de la matrice de corrélation met en évidence trois structures distinctes:

Premièrement, les actifs NVDA, AMD, GOOGL, AMZN et MSFT forment un noyau fortement corrélé (zone plus sombre du heatmap), avec des coefficients de corrélation généralement compris entre 0,60 et 0,70 traduisant une exposition commune à des facteurs de risque similaires. Ce sous-ensemble d'actifs constitue le cœur du portefeuille IA, dans lequel se concentre une part importante de l'exposition aux innovations, aux investissements et aux anticipations de croissance liées à l'intelligence artificielle. Dans ce contexte de forte corrélation, les événements défavorables affectant ces facteurs communs sont susceptibles d'engendrer des pertes simultanées sur plusieurs actifs du noyau. Il est donc raisonnable de considérer que les contributions dominantes au risque de queue du portefeuille proviennent principalement de ce sous-ensemble fortement corrélé.

Deuxièmement, les corrélations impliquant Oracle et Meta avec ce noyau central sont plus modérées, variant approximativement entre 0,35 et 0,55. Bien que ces entreprises appartiennent également à l'écosystème de l'intelligence artificielle, leurs modèles d'affaires et leurs sources de revenus présentent des spécificités qui atténuent leur mouvement commun avec les acteurs dominants du groupe.

Enfin, la corrélation entre Oracle et Meta est relativement faible (0,27), indiquant une dynamique conjointe limitée entre ces deux actifs. Cette faible dépendance suggère qu'Oracle et Meta, bien qu'appartenant au même secteur, évoluent dans des segments distincts de la chaîne de valeur.

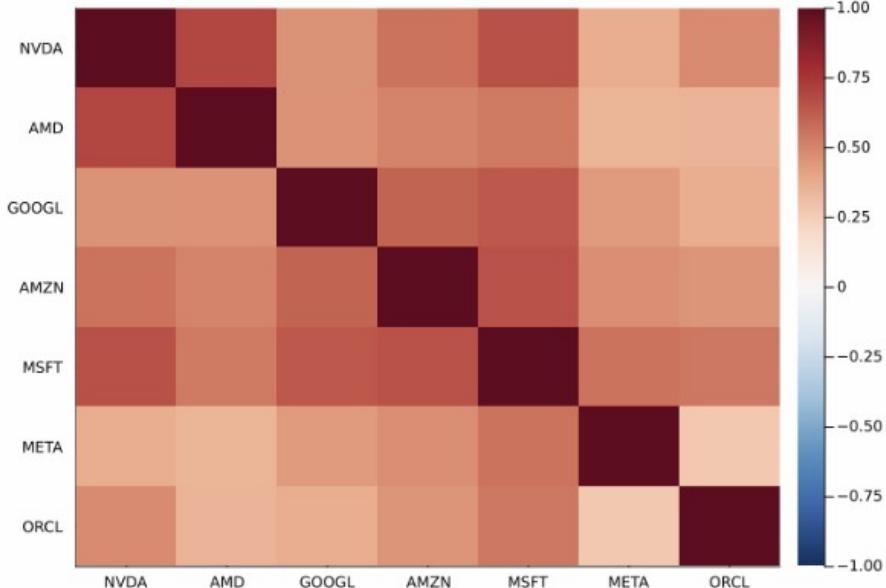


Figure 2: Matrice de corrélation des rendements logarithmiques hebdomadaires des actifs IA (2018–2024).

### 4.4 Évaluation du risque : estimation d'une probabilité rare

Soit un portefeuille de poids  $u = (u_1, \dots, u_d) \in \mathbb{R}_+^d$  tel que  $\sum_{i=1}^d u_i = 1$ . Soit un vecteur de rendements hebdomadaires  $\xi = (\xi_1, \dots, \xi_d)$  tel que  $\xi \sim \mathcal{N}(\mu, \Sigma)$  où  $\mu \in \mathbb{R}^d$  et  $\Sigma \in \mathbb{R}^{d \times d}$  sont estimés empiriquement (avec shrinkage de covariance).

La perte hebdomadaire du portefeuille est définie par

$$L(u, \xi) = -u^\top \xi$$

Pour un seuil  $z \in \mathbb{R}$ , l'événement de perte  $A_z$  est défini par :

$$A_z = \{L(u, \xi) \geq z\}.$$

On cherche à estimer la probabilité rare associée

$$P(u, z) := \mathbb{P}(L(u, \xi) \geq z). \quad (4)$$

Le seuil  $z$  est paramétrisé par un multiple de l'écart-type de  $L(u, \xi)$  comme suit :

$$z(k) = \mu_L + k \sigma_L, \quad k \geq 0, \quad (5)$$

sous l'hypothèse gaussienne,

$$L(u, \xi) \sim \mathcal{N}(\mu_L, \sigma_L^2), \quad \mu_L = -u^\top \mu, \quad \sigma_L^2 = u^\top \Sigma u.$$

Cette paramétrisation permet de faire une comparaison sur une grille de niveau de rareté. Ainsi, la probabilité cible  $P(u, z(k))$  décroît lorsque  $k$  augmente.

#### 4.4.1 Estimateur Monte Carlo standard (SAA)

À titre comparatif, nous allons considérer la méthode du *sample-average approximation (SAA)* pour la probabilité (4).

On génère  $N$  échantillons i.i.d

$$\xi^{(1)}, \dots, \xi^{(N)} \stackrel{i.i.d.}{\sim} \mathcal{N}(\mu, \Sigma),$$

et on définit l'estimateur

$$\hat{P}_N^{MC}(u, z) = \frac{1}{N} \sum_{j=1}^N \mathbf{1}\{L(u, \xi^{(j)}) \geq z\}. \quad (6)$$

On calcule les perte  $L(u, \xi^{(j)})$  et on compte les dépassement ou le nombre d'occurrences de l'événement  $A_z$  (ou *hits*) par :

$$M_N^{MC}(u, z) = \sum_{j=1}^N \mathbf{1}\{L(u, \xi^{(j)}) \geq z\}, \quad \hat{P}_N^{MC}(u, z) = \frac{M_N^{MC}(u, z)}{N}.$$

Un intervalle de confiance binomial (Wilson ou Clopper–Pearson) fournit

$$IC_{lo} \leq P(u, z) \leq IC_{hi}.$$

#### 4.4.2 Approximation de grandes déviations de premier ordre (P1)

Nous comparons donc notre Monte Carlo (6) à une approximation de premier ordre, notée  $P_1(u, z)$ , dérivé de la théorie des grandes déviations (LDT). Pour un seuil  $z(k)$ , on souhaite évaluer la proximité de  $P_1(u, z(k))$  avec une probabilité de référence.

Étant donné qu'on n'a pas accès à la vraie probabilité de nos événements, nous définissons une probabilité de référence  $P_{\text{ref}}(u, z)$  par une Monte Carlo massive (taille  $N_{\text{ref}}$  très grande), faisant office de proxy. Nous rapportons :

- la valeur  $P_1(u, z(k))$  (approximation LDT premier ordre),
- l'estimateur SAA  $\hat{P}_N^{MC}(u, z(k))$  et le nombre de hits  $M_N^{MC}(u, z(k))$ ,
- des erreurs en échelle logarithmique (quand elles sont définies) :

$$\varepsilon_1^{LDT}(u, z) = \log_{10} P_1(u, z) - \log_{10} P_{\text{ref}}(u, z), \quad \varepsilon_N^{MC}(u, z) = \log_{10} \hat{P}_N^{MC}(u, z) - \log_{10} P_{\text{ref}}(u, z).$$

Le tableau 2 synthétise ces comparaisons sur une grille de niveau  $k$ .

Table 2: Comparaison entre la probabilité de référence Monte Carlo  $P(u, z) \approx P_{N_{\text{ref}}}^{MC}(u, z)$ , l'approximation sampling-free de premier ordre ( $P_1$ ) et Monte Carlo standard ( $P_N^{MC}$ )

$k$	$P_{N_{\text{ref}}}^{MC}(u, z)$	$P_1(u, z)$	$P_N^{MC}(u, z)$	$M_N^{MC}(u, z)$	$\epsilon_1^{\text{LDT}}$	$\epsilon_N^{\text{MC}}$
3.0	$1.3501 \times 10^{-3}$	$1.3499 \times 10^{-3}$	$1.32 \times 10^{-3}$	132	$-6.85 \times 10^{-5}$	$-6.10 \times 10^{-3}$
3.5	$2.0923 \times 10^{-4}$	$2.3263 \times 10^{-4}$	$2.10 \times 10^{-4}$	21	$1.55 \times 10^{-2}$	$-4.34 \times 10^{-2}$
4.0	$3.1257 \times 10^{-5}$	$3.1971 \times 10^{-5}$	$2.00 \times 10^{-5}$	2	$9.05 \times 10^{-3}$	$-2.12 \times 10^{-1}$
4.5	$3.4250 \times 10^{-6}$	$3.3977 \times 10^{-6}$	0	0	$-3.95 \times 10^{-3}$	$-\infty$
5.0	$2.9467 \times 10^{-7}$	$2.8665 \times 10^{-7}$	0	0	$-1.01 \times 10^{-2}$	$-\infty$

Le Tableau 2 compare l'approximation de premier ordre ( $P_1$ ) au Monte Carlo standard ( $P_N^{MC}$ ) en utilisant un Monte Carlo massif ( $N_{\text{ref}} = 10^9$ ) comme probabilité de référence. Les résultats montrent que  $P_1$  reste très proche du benchmark numérique sur l'ensemble des seuils considérés, y compris pour des niveaux de rareté élevés ( $k = 5$ ), avec des erreurs logarithmiques faibles. La colonne  $M_N^{MC}(x, c)$  représente le nombre de scénarios (sur  $N$ ) où la perte dépasse  $c$  ( $L(x) \geq c$ ) mettant en évidence la rareté des dépassements observés par Monte Carlo standard, passant de quelques dizaines d'occurrences à zéro lorsque le seuil augmente. Par conséquent, le Monte Carlo standard avec  $N = 10^5$  fournit dès  $k \geq 4.5$  une estimation dégénérée nulle, associée à une erreur logarithmique infinie.

Lorsque  $P(u, z)$  est très petit,  $M_N^{MC}(u, z)$  devient nulle même pour des tailles d'échantillons modérées, ce qui rend  $\hat{P}_N^{MC}(u, z)$  très bruité, voire dégénéré (estimation nulle). De ce fait, nous observons l'apparition d'erreurs logarithmiques importantes dès qu'aucun dépassement  $\log_{10} \hat{P}_N^{MC}(u, z)$  devient numériquement indéfini. Cette limite justifie l'usage d'approximation fondées sur la théorie des grandes déviations.

## 4.5 Allocation optimale sous contraintes de chance rares

### 4.5.1 Portefeuilles de références (baseline)

Avant d'introduire l'optimisation sous contrainte en chance, on définit quatre portefeuilles qui serviront de points de comparaison.

- **Portefeuille équipondéré** : Pour un actif  $i$  le poids associé est :  $u_i = 1/n$ . C'est une diversification "naive" qui ne s'appuie sur aucune information sur les rendements, la covariance ou les événements rares. Ce portefeuille permettra d'illustrer ce qui se produit lors d'une allocation à poids égaux.
- **Portefeuille minimum-variance (modèle de Markowitz)** : On minimise la variance  $u^\top \Sigma u$  sous contrainte :  $u \geq 0, \sum_i u_i = 1$ . On optimise le risque local autour de la moyenne.
- Portefeuille maximum rendement (MaxRet) : On maximise  $u^\top \mu$  sous contrainte:  $u \geq 0, \sum_i u_i = 1$ . Cette allocation agressive se concentre sur l'actif dont le rendement espéré est le plus élevé.
- Portefeuille optimisé sous contrainte en chance rare (OptVaR) : Obtenu par la résolution du problème (7).

### 4.5.2 Value at Risk (VaR), Conditional Value at Risk (CVaR) et contrainte en chance

Pour une perte de portefeuille définie par  $L(u, \xi) = -\sum_{i=1}^n u_i \xi_i$ , la Value-at-Risk (VaR) au niveau  $\alpha \in (0, 1)$  du portefeuille est le seuil  $z$  tel que :

$$\text{VaR}_{1-\alpha}(u) := \inf \{z \in \mathbb{R} : \mathbb{P}(L(u, \xi) \geq z) \leq \alpha\}.$$

La Value-at-Risk (VaR) nous dit avec une probabilité au plus  $\alpha$  que perte dépasse  $z$ . La Conditional Value-at-Risk (CVaR) au même niveau est définie comme l'espérance conditionnelle de la partie au-delà de la VaR:

$$\text{CVaR}_{1-\alpha}(u) = \mathbb{E}[L(u, \xi) | L(u, \xi) \geq \text{VaR}_{1-\alpha}(u)].$$

Lorsque la loi de  $L(u, \xi)$  est continue, la CVaR admet l'expression équivalente suivante :

$$\text{CVaR}_{1-\alpha}(u) = \frac{1}{\alpha} \int_{1-\alpha}^1 \text{VaR}_p(u) dp.$$

La CVaR est implicitement gouvernée par le point dominant  $\xi^*(u, z^*)$ , qui décrit la trajectoire typique menant aux pertes conditionnelles extrêmes.

Le point dominant  $\xi^*(u, z^*)$  est le scénario de marché le plus probable conduisant à une perte  $z$ . Il résout un problème de minimisation quadratique sous contrainte. Il révèle la direction systémique des pertes, les actifs qui dominent le risque extrême et permet de comprendre pourquoi un portefeuille est fragile ou robuste.

#### 4.5.3 Problème d'optimisation sous contrainte en chance

Nous considérons un portefeuille *long-only* décrit par un vecteur de pondérations

$$u = (u_1, \dots, u_n) \in \mathbb{R}_+^n, \quad \sum_{i=1}^n u_i = 1,$$

et un vecteur aléatoire de rendements hebdomadaires

$$\xi = (\xi_1, \dots, \xi_n),$$

modélisé par une distribution gaussienne de référence

$$\xi \sim \mathcal{N}(\mu, \Sigma),$$

où le vecteur de moyenne  $\mu$  et la matrice de covariance  $\Sigma$  sont estimés empiriquement à partir des données historiques.

La perte du portefeuille est définie par

$$L(u, \xi) = - \sum_{i=1}^n u_i \xi_i.$$

Dans ce cadre, on ne cherche pas à minimiser une espérance ou une variance, mais plutôt on cherche le plus grand seuil de perte  $z$  tel que la probabilité de dépasser ce seuil reste inférieure à  $\alpha$ . Autrement dit, on se fixe un niveau  $\alpha \ll 1$ , et on pousse le seuil  $z$  le plus loin possible tout en respectant la contrainte en chance. Ainsi on augmente la tolérance du portefeuille aux pertes extrêmes tout en maintenant sous contrôle la probabilité de dépassement.

Ainsi la Value-at-Risk devient :

$$\text{VaR}_{1-\alpha}(u) = \sup\{z \in \mathbb{R} : P(L(u, \xi) \geq z) \leq \alpha\}.$$

et la CVaR(Expected Shortfall) empirique au même niveau est :

$$\text{CVaR}_{1-\alpha}(u) = \max_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{\alpha} \mathbb{E} [(L(u, \xi) - \eta)_+] \right\}.$$

Le problème d'optimisation considéré devient

$$\begin{aligned} & \max_{u \in \mathbb{R}_+^n, z \in \mathbb{R}} && z \\ & \text{s.c.} && \sum_{i=1}^n u_i = 1, \\ & && \mathbb{P}(L(u, \xi) \geq z) \leq \alpha. \end{aligned} \tag{7}$$

#### 4.5.4 Résultats numériques

Le tableau 3 présente les résultats obtenus pour un niveau de rareté  $\alpha = 0.01\%$ . Pour chaque méthode, le seuil optimal  $z^*$  est calculé à partir de l'approximation P1, puis validé par simulation Monte Carlo. Les probabilités empiriques sont accompagnées d'intervalles de confiance binomiaux. Le seuil  $\alpha$  place l'analyse dans un régime beaucoup plus rare (soit environ 20 "hits" sur 200,000 tirages).

Table 3: Allocation optimale sous contrainte de chance rare ( $\alpha = 10^{-4}$ ).

Méthode	$z^*$	$P_1(u, z^*)$	MC	Hits	$IC_{lo}$	$IC_{hi}$	$VaR^{MC}$	$CVaR^{MC}$	$\Delta$
Uniforme	0.146533	0.0001	9.0e-5	18	5.69325e-5	1.42271e-4	0.144451	0.154326	-0.002083
Max-Ret	0.246898	0.0001	9.5e-5	19	6.08218e-5	1.48381e-4	0.244684	0.258258	-0.002214
Min-Var	0.120573	0.0001	1.05e-4	21	6.86808e-5	1.60522e-4	0.120578	0.129360	+0.000005
OptVaR	0.261993	0.0001	5.5e-5	11	3.07125e-5	9.84924e-5	0.250243	0.267384	-0.011750

Le portefeuille (**OptVaR**) atteint le seuil optimal le plus élevé, avec  $z^* = 0.261993$ , contre 0.246898, 0.146533 et 0.120573 respectivement pour les portefeuilles maximum-rendement, équipondéré et minimum-variance. Cela montre que l'optimisation sous contrainte en chance permet d'augmenter significativement le niveau de perte extrême tolérable. **Min-Var** est le plus fragile en queue.

À  $\alpha = 0.01\%$ , **OptVaR** repousse le plus le seuil extrême admissible  $z^*$ , mais l'évaluation Monte Carlo devient intrinsèquement plus instables (hits très faibles), ce qui reflète un  $\Delta$  plus négatif pour **OptVaR**. L'Expected Shortfall **CVaR** montre que les portefeuilles dominant sur  $z^*$  (**OptVaR** et **Max-Ret**) restent associés à des pertes conditionnelles plus sévères lorsque l'événement rare se produit. En effet, la perte moyenne subie par **OptVaR** et **Max-Ret** sous l'événement rare de probabilité  $\alpha = 0.01\%$  est respectivement d'environ 26.7% et 24.7% contre 14.7% pour **Uniforme** et 12.06% pour **Min-Var**.

## 4.6 Analyse des points dominants

### 4.6.1 Définition et interprétation économique

Dans le cadre de l'approximation de grandes déviations de premier ordre, la probabilité d'un événement de perte extrême est dominée par une trajectoire unique du vecteur de rendements. Cette trajectoire, appelée *point dominant*, correspond au scénario de marché le plus probable conduisant à une perte de niveau donné. Formellement, pour un portefeuille  $u$  et un seuil de perte  $z$ , le point dominant  $\xi^*(u, z)$  est défini comme la solution du problème quadratique contraint

$$\xi^*(u, z) = \arg \min_{\xi: L(u, \xi)=z} \frac{1}{2} (\xi - \mu)^\top \Sigma^{-1} (\xi - \mu).$$

D'un point de vue économique,  $\xi^*$  ne représente le choc systémique *le moins improbable* capable de générer une perte extrême. Il capture ainsi la direction privilégiée des stress de marché, compte tenu de la structure de dépendance entre actifs.

### 4.6.2 Analyse empirique des points dominants par portefeuille

La Table 4 présente la décomposition du point dominant  $\xi^*$  par actif et par portefeuille. Les différentes colonnes s'interprètent comme suit:

- $u_i$  représente le poids de l'actif  $i$  dans le portefeuille considéré. Dans le cas du portefeuille équipondéré, tous les poids sont identiques, tandis que pour les portefeuilles optimisés, ils reflètent la stratégie d'allocation retenue.
- $\xi_i^*$  désigne la composante du point dominant associée à l'actif  $i$ . Elle correspond au rendement extrême que doit subir cet actif dans le scénario rare *le moins improbable* conduisant à une perte totale égale à  $z^*$ .
- $\xi_i^* - \mu_i$  mesure l'ampleur du choc subi par l'actif  $i$  relativement à son rendement moyen. Cette quantité permet d'identifier quels actifs doivent s'écartez le plus fortement de leur comportement normal pour engendrer l'événement extrême.

- $-u_i \xi_i^*$  représente la contribution individuelle de l'actif  $i$  à la perte totale du portefeuille au point dominant, la perte étant définie par  $L(u, \xi) = -u^\top \xi$ . Cette colonne permet de relier directement la géométrie du point dominant à l'impact économique sur le portefeuille.
- **Part abs. (%)** correspond à la part absolue de la contribution de l'actif  $i$  dans la perte extrême totale,  $\frac{|u_i \xi_i^*|}{\sum_j |u_j \xi_j^*|} \times 100$

Table 4: Décomposition du point dominant  $\xi^*$  par actif et par portefeuille ( $\alpha = 10^{-4}$ ).

Actif	$u_i$	$\xi_i^*$	$\text{choc}_i$	$-u_i \xi_i^*$	Part abs. (%)
<b>Portefeuille équipondéré (Equal), <math>z^* = 0.1465</math></b>					
AMD	0.1429	-0.2050	-0.2040	0.0293	20.0
NVDA	0.1429	-0.2045	-0.2140	0.0292	19.9
META	0.1429	-0.1528	-0.1569	0.0218	14.9
AMZN	0.1429	-0.1395	-0.1411	0.0199	13.6
GOOGL	0.1429	-0.1168	-0.1187	0.0167	11.4
MSFT	0.1429	-0.1092	-0.1111	0.0156	10.6
ORCL	0.1429	-0.0980	-0.1024	0.0140	9.6
<b>Portefeuille max-rendement (MaxRet), <math>z^* = 0.2469</math></b>					
NVDA	0.9920	-0.2481	-0.2576	0.2461	99.6
AMD	0.0013	-0.1842	-0.1831	0.0002	0.1
AMZN	0.0013	-0.0991	-0.1008	0.0001	0.1
MSFT	0.0013	-0.0855	-0.0874	0.0001	0.0
GOOGL	0.0013	-0.0742	-0.0761	0.0001	0.0
<b>Portefeuille min-variance (MinVar), <math>z^* = 0.1206</math></b>					
META	0.0597	-0.1220	-0.1262	0.0000	0.0
AMD	0.0000	-0.1460	-0.1450	0.0000	0.0
AMZN	0.0000	-0.1227	-0.1244	0.0000	0.0
<b>Portefeuille optimisé VaR (OptVaR), <math>z^* = 0.2620</math></b>					
AMD	1.0000	-0.2620	-0.2610	0.2620	100.0
NVDA	0.0000	-0.1710	-0.1805	0.0000	0.0
GOOGL	0.0000	-0.0742	-0.0761	0.0000	0.0
AMZN	0.0000	-0.0891	-0.0908	0.0000	0.0
MSFT	0.0000	-0.0689	-0.0707	0.0000	0.0
META	0.0000	-0.0783	-0.0824	0.0000	0.0
ORCL	0.0000	-0.0534	-0.0577	0.0000	0.0

**Portefeuille équipondéré.** Le point dominant n'est pas réparti de manière uniforme. *AMD* et *NVDA* dominent la queue avec un  $\xi_i^*$  plus négatif. *ORCL* apparaît le moins contributeur, malgré les craintes du marché. On peut affirmer que la diversification naïve n'élimine pas le risque de queue mais le redistribue vers les actifs les plus sensibles au chocs systémiques (les actifs créateurs de valeurs).

**Portefeuille max-rendement.** Le portefeuille max-rendement est quasi entièrement concentré sur *NVDA* (99.2%). Les autres titres, bien que soumis à des chocs extrêmes non nuls, ont une contribution négligeable du fait de leurs poids très faibles voir nuls.

**Portefeuille minimum variance.** Le portefeuille minimum variance est extrêmement compressé avec des actifs à poids quasi nuls. Le point dominant existe mathématiquement mais ne porte aucun poids. Ainsi ce portefeuille minimise la variance locale, mais ne structure pas le risque de queue (risque diffus et incontrôlé). C'est un portefeuille rendant aveugle à la direction des chocs rares.

**Portefeuille optimisé sous contrainte de chance rare.** Le portefeuille optimisé sous contrainte de chance rare (OptVaR) présente un comportement qualitativement différent. L’allocation est fortement concentrée sur AMD, mais cette concentration est ici structurellement défensive. En effet **OptVaR** cherche à absorber le choc dans la direction la moins pénalisante au sens probabiliste.

## 5 DISCUSSION

### 5.1 Synthèse des résultats empiriques

Ce travail avait pour objectif d’analyser le risque de queue associé à un portefeuille d’actions piliers de l’écosystème de l’intelligence artificielle, à l’aide d’un cadre d’optimisation sous contraintes en chance rares inspiré de Tong et al. (2022).

Les résultats empiriques montrent que pour un niveau de rareté  $\alpha = 10^{-4}$ , on obtient un seuil admissible de perte extrême  $z^* = 0.262$  pour **OptVaR**, contre 0.247 pour **MaxRet**, 0.147 pour le portefeuille **Uniforme** et seulement 0.121 pour **MinVar**. Cette héarchie est confirmé par l’approximation des grandes déviations de premier ordre  $P_1$  et par les estimations de Monte Carlo.

Ensuite, l’écart

$$\Delta = \text{VaR}_{1-\alpha}^{MC} - z^*$$

reste faible et légèrement négatif pour l’ensemble des portefeuilles. Ce comportement indique que l’approximation  $P_1$  est conservatrice, conformément aux résultats théoriques de la littérature, et qu’elle fournit une borne prudente du risque extrême.

Enfin, l’analyse des points dominants montre que les pertes extrêmes sont générées par des chocs conjoints et cohérents affectant simultanément l’ensemble des actifs IA. On voit clairement comment le portefeuille **MinVar** s’effondre lorsqu’on considère les événements rares. Ainsi on conclu que le risque de queue n’est pas gouverné par la variance ni par la diversification naïve, mais par la direction dominante des chocs extrêmes.

### 5.2 Implications conceptuelles pour la gestion du risque

Sans chercher à formuler des recommandations opérationnelles, ce travail apporte plusieurs enseignements conceptuels importants pour la gestion du risque extrême.

Premièrement, il montre que la diversification intra-sectorielle est insuffisante pour maîtriser le risque de pertes extrêmes lorsque les actifs partagent des facteurs de risque dominants communs. Le portefeuille equipondéré, bien qu’il répartisse uniformément les poids, reste exposé à des pertes extrêmes importants car la direction dominante des chocs reste inchangée.

Deuxièmement, le portefeuille **MinVar** bien que conçu pour résurer la dispersion autour de la moyenne, apparaît comme particulièrement fragile, avec le plus faible  $z^*$  montrant ainsi la limite de critère d’optimisation classique.

Enfin, les résultats mettent en évidence le rôle central de la structure de dépendance entre actifs dans la formation des queues de distribution. En régime extrême, les pondérations individuelles deviennent secondaires face à la direction dominante des chocs, capturée par les points dominants. le risque de queue est donc gouverné par des directions de stress communs, et non par de considérations locales de volatilité ou de rendement moyen.

### 5.3 Limites de l’étude

Plusieurs limites doivent être soulignées afin de bien circonscrire la portée des résultats. Tout d’abord, la distribution des rendements est modélisée par une loi gaussienne, ce qui constitue une approximation. Bien que l’objectif principal soit l’analyse directionnelle du risque de queue plutôt que l’estimation exacte des probabilités, des extensions à des modèles non gaussiens pourraient enrichir l’analyse.

Ensuite, l’étude est volontairement restreint à un seul secteur fortement corrélé. Ce choix permet d’isoler le risque systémique propre à l’écosystème IA mais limite la généralisation à des portefeuilles plus diversifiés.

Enfin, l'étude repose sur l'approximation de grandes déviations de premier ordre. Bien que validée empiriquement par simulation, des corrections de second ordre pourraient améliorer la précision quantitative dans des contextes encore plus extrêmes.

## 6 CONCLUSION

Ce projet montre que l'optimisation sous contrainte en chance rares constitue un cadre particulièrement pertinent pour analyser le risque de perte extrême dans des portefeuilles fortement corrélés. En combinant des approximations de grandes déviations et une analyse géométrique des points dominant, il met en exergue les mécanismes de risques invisibles aux approches classiques.

Les résultats confirment que le risque de queue est avant tout gouverné par des facteurs sustémiques et par la direction dominante des chocs extrêmes, plutôt que par des considérations locales de volatilité et de rendement moyens. L'analyse des points dominant apporte une information complémentaire et plus parlante que les seules mesures scalaires comme la VaR ou la CVaR.

## References

- Chen, W., Zhou, W., Zhang, Y., Chen, S., and Li, X. (2020). Deep learning for financial signal representation and trading: a systematic review. *Expert Systems with Applications*, 155:113380.
- Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events: For Insurance and Finance*. Springer-Verlag.
- Glasserman, P. and Kim, K. (2001). Importance sampling for portfolio value-at-risk. *Journal of Banking & Finance*, 25(10):1783–1804.
- Jiang, F., Xu, J., Liang, F., Wang, C., Li, Y., and Yang, Z. (2017). A deep reinforcement learning framework for the financial market. In *IJCAI*, pages 4776–4782.
- Lagoa, C. M., Li, X., and Sznaier, M. (2005). Probabilistically constrained linear programs and risk-adjusted controller design. *SIAM Journal on Optimization*, 15(4):938–951.
- Ledoit, O. and Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2):365–411.
- Nemirovski, A. and Shapiro, A. (2007). Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996.
- Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1):119–131.
- Prékopa, A. (2014). *Stochastic Programming*. Springer, 3 edition.
- Rockafellar, R. T. and Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, 2(3):21–41.
- Shapiro, A., Dentcheva, D., and Ruszczyński, A. (2014). *Lectures on Stochastic Programming: Modeling and Theory*. SIAM, Philadelphia.
- Tong, S., Subramanyam, A., and Rao, V. (2022). Optimization under rare chance constraints. *SIAM Journal on Optimization*, 32(2):930–958.
- Varadhan, S. R. S. (1966). Asymptotic probabilities and differential equations. *Communications on Pure and Applied Mathematics*, 19(3):261–286.
- Varadhan, S. R. S. (1984). *Large deviations and applications*, volume 46 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM.
- Xu, H., Caramanis, C., and Mannor, S. (2009). Robust support vector machine training via convex risk measures. *IEEE Transactions on Neural Networks*, 20(11):1715–1728.