



HAL
open science

Phénoménologie du LHC, prédictions théoriques et leurs incertitudes dans un contexte Bayésien

Nicolas Houdeau

► **To cite this version:**

Nicolas Houdeau. Phénoménologie du LHC, prédictions théoriques et leurs incertitudes dans un contexte Bayésien. Physique des Hautes Energies - Phénoménologie [hep-ph]. Université Pierre et Marie Curie - Paris VI, 2011. Français. NNT: . tel-00640266

HAL Id: tel-00640266

<https://theses.hal.science/tel-00640266>

Submitted on 11 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE PIERRE ET MARIE CURIE
UFR 925 de Physique fondamentale et appliquée

THÈSE

pour l'obtention du diplôme de

DOCTEUR DE L'UNIVERSITÉ PIERRE ET MARIE CURIE
Constituants élémentaires - Systèmes complexes

présentée par

Nicolas Houdeau

**Phénoménologie du LHC, prédictions théoriques et leurs incertitudes
dans un contexte Bayésien.**

soutenance le 30 septembre 2012 devant la commission d'examen composée de

M. M. Cacciari (directeur de thèse)
M. G. D'Agostini (rapporteur)
M. A. Deandrea (rapporteur)
M. A. Djouadi
M. J-P. Guillet
M. J-B. Zuber

Phénoménologie du LHC, prédictions théoriques et leurs incertitudes dans un contexte Bayésien

Le récent démarrage du LHC appelle à la mise à jour et à l'amélioration des prédictions théoriques obtenues à partir du Modèle Standard. Les éventuels signes d'une nouvelle physique devraient apparaître dans un premier temps sous la forme de légères différences entre ces résultats et les observations. Les sections efficaces recherchées doivent donc être ré-estimées à l'énergie de collision du nouvel accélérateur, puis calculées à des ordres supérieurs en perturbation. La complexité des manipulations mises en jeu impose le développement de nouvelles techniques, notamment numériques. Une fois les prédictions théoriques obtenues, leur précision doit également être évaluée de la manière la plus juste possible. Une différence notable avec les résultats expérimentaux ne pourra être mise en évidence qu'à cette seule condition. Cette thèse présente des travaux effectués sur ces trois aspects. D'une part, l'outil numérique FONLL déjà existant a été utilisé pour actualiser les prédictions de production de quarks lourds au premier ordre en perturbation à l'énergie de collision de 7 TeV du LHC. L'étude d'une approche alternative de la gestion d'un type de divergences apparaissant dans la procédure de calcul des sections efficaces, les divergences infrarouges, est également présentée. Enfin, un modèle de confiance (ou probabilités bayésiennes) permettant de décrire rigoureusement la contribution de l'effet de troncature des séries perturbatives à l'incertitude théorique est détaillé. Une discussion sur les notions de mesure de probabilité et de mesure de confiance introduit cette étude.

Incertitude, Bayes, LHC, Probabilité, Confiance, Statistique

LHC phenomenology, theoretical predictions and their uncertainties in a Bayesian framework

The recent starting up of the LHC calls for an update and an improvement of the theoretical predictions computed from the Standard Model. The possible signs of a new physic should at first appear as slight discrepancies between those results and the observations. As a consequence, the cross sections must be estimated anew at the LHC energy of collision then calculated at higher perturbative orders. The complexity of the manipulations involved requires the development of new, possibly numerical, techniques. Once the theoretical predictions are obtained, their precision has also to be evaluated as accurately as possible. This thesis presents developments on these three avenues. First, the already existing FONLL tool has been used to update the predictions of the heavy quark production rate at first order in perturbation theory at the energy of collision of 7 TeV. An alternative approach to the handling of a type of divergences arising in the computation of the cross section, namely the infrared divergences, is also presented. Finally, a confidence model (or bayesian model) allowing a rigorous description of the contribution of the truncation of the perturbative series to the theoretical uncertainty is detailed. A discussion on the notions of probability measure and confidence measure provides an introduction to this study.

Uncertainty, Bayes, LHC, Probability, Confidence, Statistics

Je tiens à remercier le Laboratoire de Physique Théorique et Hautes Énergies (LPTHE - UMR 7589) pour m'avoir accueilli ces trois dernières années. Je remercie également les membres du jury et rapporteurs pour leur participation. Il me faut également remercier mon directeur de thèse, Matteo Cacciari, pour m'avoir supporté tout au long de cette période.

Cette thèse a indéniablement été une expérience inoubliable. Si celle-ci a débuté comme un projet personnel, c'est à toutes les personnes qui m'ont soutenu que je dois de l'avoir terminée. Mes pensées se tournent donc aujourd'hui vers tous ceux sans qui ce manuscrit n'aurait pas vu le jour. Merci en premier lieu à tous les doctorants et post-doctorants pour avoir réchauffé l'atmosphère du laboratoire. Merci en particulier à Nicolas Méric, Damien Levis, Julius Bonart, Gauthier Solard et Anna Carolina Ribeiro Teixeira pour les pauses cafés, les diners, sorties et discussions parfois endiablées que nous avons pu partager. Je n'aurais jamais pû survivre à cette dernière année sans vous. Merci ensuite à toute ma famille pour m'avoir soutenu de loin et permis de m'évader lorsque j'en avais le plus besoin. Rassurez-vous, c'est terminé ! Merci également à Xavier Epiard, Tomas Simovic, Ludovic Geber, Sébastien Guido, Clémentine Camosseto, Bertrand Moreau, François-Marie Julio, Jennifer Ségard, Mathieu Houdeau, ainsi que tous ceux que je n'ai pas nommé mais qui se reconnaîtront, pour m'avoir aidé à me recentrer et à me retrouver alors que je m'étais perdu en route.

Enfin, je tiens à remercier tout particulièrement Frédéric Hublou, Caroline Gaché, Julien Gélébart, Antoine Devillers, Nicolas Guezennec, Solène Théry et Laure Feuillebois pour m'avoir prêté une oreille attentive tout au long de ces années, dans tous ces moments où la potence ne me paraissait plus être une option si inenvisageable. Je ne sais pas où j'en serais aujourd'hui sans vous. Merci pour votre amitié et votre soutien.

Table des matières

I	Manipulation et estimation d’incertitudes théoriques	13
1	Probabilité et confiance	15
1.1	Notions de probabilité et de confiance	18
1.1.1	Exemple du jet de dé	18
1.1.1.1	Probabilité d’une expérience de jet de dé	18
1.1.1.2	Confiance sur le résultat d’un jet de dé	19
1.1.2	Probabilité et reproductibilité	19
1.1.2.1	Reproductibilité	20
1.1.2.2	Espace probabilisé et variable aléatoire	20
1.1.3	Confiance et incertitude	21
1.1.3.1	Incertitude	21
1.1.3.2	Modèle de confiance et variable inconnue	21
1.1.3.3	Utilisation pratique d’un modèle de confiance	23
1.2	Confiance sur le résultat d’une expérience	26
1.2.1	Mesure de confiance et information	26
1.2.2	Mesures de confiance et de probabilité	26
1.2.3	Définition de l’expérience considérée	29
1.3	Confiance sur la valeur d’un paramètre inconnu	30
1.3.1	Estimation statistique : approche “fréquentiste”	31
1.3.1.1	Expérience statistique	31
1.3.1.2	Estimation de paramètre	31
1.3.1.3	Intervalle de confiance statistique	32
1.3.1.4	Composition d’intervalles de confiance statistiques	34
1.3.1.5	Intervalle de confiance et information supplémentaire	35
1.3.2	Mesure continue de confiance : approche “bayésienne”	36
1.3.2.1	Modèle de confiance	36
1.3.2.2	Normalisation du prior. Information insuffisante	37
1.3.2.3	Variable privilégiée	38
1.3.3	Choix de méthode	39
1.3.3.1	Test répété	39
1.3.3.2	Incertitude statistique et biais	39
1.4	Applications : manipulations des incertitudes théoriques	41
1.4.1	Incertitudes expérimentales et théoriques	41
1.4.2	Extraction de paramètre	42
1.4.2.1	Construction du modèle de confiance	42
1.4.2.2	Expressions des densités	43

1.4.2.3	Discussion du modèle	44
1.4.2.4	Extraction de paramètre	45
1.5	Conclusion	46
2	Incertitudes théoriques	51
2.1	Incertitude théorique sur une série perturbative tronquée	55
2.1.1	Notations	55
2.1.2	Estimation conventionnelle de l'incertitude	55
2.1.3	Variation de l'échelle de renormalisation	56
2.1.4	Exemple du processus $e^+e^- \rightarrow$ hadrons	57
2.1.5	Ordre de grandeur des coefficients de la section	59
2.2	Modèle de confiance	60
2.2.1	Définition du modèle	60
2.2.2	Comportement du modèle	62
2.2.2.1	Densité $f_\epsilon(\bar{c}, c_0, c_1, \dots)$ de la mesure de confiance	62
2.2.2.2	Densité conditionnelle $f(\bar{c} c_0, \dots, c_k)$	62
2.2.2.3	Densité conditionnelle $f(c_n c_0, \dots, c_k)$	63
2.2.2.4	Comportement du modèle en pratique	64
2.3	Discussion du modèle	65
2.3.1	Paramètres inconnus	65
2.3.2	Coefficients de même ordre de grandeur	66
2.3.3	Information partagée	67
2.3.3.1	Informations incompatibles	67
2.3.3.2	Premier modèle de prise en charge du signe de la section efficace	68
2.3.3.3	Deuxième modèle de prise en charge du signe de la section efficace	68
2.3.3.4	Liens directs et indirects	69
2.3.4	Incertitude résiduelle	70
2.3.4.1	Confiance et certitude	70
2.3.4.2	Poids des coefficients calculés	71
2.3.4.3	Choix de variable privilégiée	72
2.3.5	Nature du paramètre caché	73
2.4	Estimation de l'incertitude de troncature	74
2.4.1	Densité conditionnelle $f(\Delta_k c_0, \dots, c_k)$ approchée	74
2.4.2	Validité de l'approximation	78
2.4.3	Évaluation numérique de $f(\Delta_k c_0, \dots, c_k)$	80
2.4.4	Choix de densité $f(c_n \bar{c})$ non uniforme	81
2.5	Comparaison du modèle et de la méthode conventionnelle	81
2.5.1	Niveaux de confiance des intervalles conventionnels	81
2.5.2	Évolution des valeurs de confiance	84
2.5.3	Exemple du processus $e^+e^- \rightarrow$ hadrons	88
2.6	Approximation des ordres supérieurs	89
2.6.1	Ajustement du modèle	90
2.6.1.1	Informations supplémentaires	90
2.6.1.2	Mesure de confiance	90
2.6.1.3	Dérivation des densités (2.114)	91

TABLE DES MATIÈRES

2.6.2	Comportement du modèle	92
II	Phénoménologie au LHC	99
3	Production de quarks lourds au LHC	101
3.1	Production et évolution de quarks lourds	104
3.1.1	Production d'un quark lourd. Factorisation	104
3.1.1.1	Quarks légers et quarks lourds	104
3.1.1.2	Théorème de factorisation	105
3.1.2	Production de quarks massifs	107
3.1.2.1	Prédiction à ordre fixe en théorie des perturbations . . .	107
3.1.2.2	Processus dur de production de quark non massif et frag- mentation perturbative	109
3.1.2.3	Comparaison des deux approches	111
3.1.3	Fragmentation non perturbative	112
3.1.3.1	Fonction de fragmentation non perturbative	112
3.1.3.2	Extraction dans l'espace de Mellin	112
3.1.4	Désintégration	112
3.2	Prédictions théoriques FONLL pour le LHC	113
3.2.1	Paramètres utilisés	113
3.2.1.1	Mésons B	113
3.2.1.2	Mésons D	113
3.2.2	Incertitudes théoriques	114
3.2.3	Production de quarks lourds au LHC, 7 TeV	115
3.2.3.1	Production de mésons D	115
3.2.3.2	Production de mésons B	115
3.2.3.3	Production $B \rightarrow J/\Psi$	117
4	Annulation numérique des divergences infrarouges et collinéaires	121
4.1	Annulation numérique des divergences infrarouges et collinéaires	124
4.1.1	Principe de l'annulation	124
4.1.2	Cas du processus $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$	124
4.1.3	Manipulation des distributions	128
4.2	Annulation numérique des divergences dans l'espace de Mellin	129
4.2.1	Transformée de Mellin	129
4.2.1.1	Définition	129
4.2.1.2	Convolution de Mellin	130
4.2.2	Cas du processus $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$	130
4.2.2.1	Transformée de Mellin	130
4.2.2.2	Annulation des divergences	131
4.2.2.3	Transformée de Mellin inverse	135
4.3	Limitations et utilité de l'approche	136
4.3.1	Dépendance explicite en ϵ	136
4.3.2	Prolongement analytique dans l'espace conjugué	136
4.3.3	Transformée de Mellin en k dimensions et position des pôles	137
4.3.4	Domaine d'application et utilité de la méthode	137

Introduction

Le LHC (*Large Hadron Collider*) a finalement été lancé et, aucun trou noir n'ayant englouti la Terre, la course à la collecte de données et à leur analyse a maintenant commencé pour de bon. Tandis que l'objectif à court terme est de redécouvrir le Modèle Standard, celui à long terme est bien entendu de trouver des signaux de "nouvelle physique", qu'il s'agisse du boson de Higgs, de la supersymétrie ou de quelque chose d'autre, plus exotique, et potentiellement inattendu. Bien que chacun espère que des découvertes s'annonceront sous la forme de signaux clairs, il est concevable, et probablement également inévitable dans un premier temps, qu'elles se présentent plutôt cachées sous la forme d'un décalage subtile entre théorie et expérience. Si c'est le cas, la qualité des prédictions et le contrôle de leur prédictions devient d'une importance cruciale.

Trois pistes de développements phénoménologiques sont par conséquent d'actualité. Dans un premier temps, les prédictions du Modèle Standard doivent être actualisées de manière à correspondre à l'énergie de collision du LHC. Un travail important doit être ensuite accompli dans l'amélioration de ces calculs. Face à la quantité de données recueillies au nouvel accélérateur, la précision des études théoriques risque en effet d'être le point limitant de la recherche de nouvelle physique. Enfin, tout effort dans cette voie ne pourra porter ses fruits que lorsque cette incertitude sur les prédictions sera pleinement décrite et l'amélioration de leur précision quantifiée. Comparant une mesure expérimentale à un calcul théorique, il doit être possible de dire s'ils sont en accord ou non, et avec quel degré de confiance une telle déclaration peut être effectuée. Ceci est impossible à réaliser sans que les mesures expérimentales et les prédictions théoriques ne soient fournies avec un degré, sensé et communément accepté, d'incertitude.

Cette thèse présente des travaux effectués sur ces trois aspects. En deuxième partie, les prédictions de production de quarks lourds en collision hadronique au premier ordre en perturbation sont actualisées pour une énergie de 7TeV (chapitre 3). Le chapitre 4 présente une tentative d'amélioration de la gestion des divergences infrarouges. Ces singularités limitent l'automatisation des calculs de section efficace

Première partie

Manipulation et estimation
d'incertitudes théoriques

Chapitre 1

Probabilité et confiance

Sommaire

1.1	Notions de probabilité et de confiance	18
1.1.1	Exemple du jet de dé	18
1.1.2	Probabilité et reproductibilité	19
1.1.3	Confiance et incertitude	21
1.2	Confiance sur le résultat d'une expérience	26
1.2.1	Mesure de confiance et information	26
1.2.2	Mesures de confiance et de probabilité	26
1.2.3	Définition de l'expérience considérée	29
1.3	Confiance sur la valeur d'un paramètre inconnu	30
1.3.1	Estimation statistique : approche "fréquentiste"	31
1.3.2	Mesure continue de confiance : approche "bayésienne"	36
1.3.3	Choix de méthode	39
1.4	Applications : manipulations des incertitudes théoriques . . .	41
1.4.1	Incertainces expérimentales et théoriques	41
1.4.2	Extraction de paramètre	42
1.5	Conclusion	46

Introduction

Les domaines des probabilités et des statistiques sont aujourd'hui des outils indispensables dans de nombreuses parties de la société. De l'utilisation de la *règle de trois* en gestion de stock au calcul stochastique en analyse financière, leur utilisation à tous les niveaux de complexité a depuis longtemps dépassé les frontières de la recherche fondamentale. Pour autant, ces outils font toujours l'objet de nouveaux développements en science en générale et en physique en particulier. De nouveaux modèles adaptés aux sujets d'étude contemporains doivent être développés et des points de détail conceptuels peuvent par moment nécessiter une certaine clarification.

Ce chapitre introduit ainsi succinctement les outils que je choisis de désigner dans la suite sous les noms de mesures de probabilité (en lieu et place de probabilité *fréquentiste*) et de confiance (correspondant essentiellement à la notion de probabilité *bayésienne*, encore appelée probabilité *subjective*, ou encore de degré de croyance). Son objectif est double. Il permet tout d'abord d'introduire le vocabulaire, non conventionnel, utilisé dans le chapitre 2 dédié à l'étude de la contribution de l'effet de troncature à l'incertitude sur la prédiction théorique d'une section efficace. Ce chapitre est également l'occasion de présenter l'état d'esprit avec lequel la notion de confiance a été utilisée dans cette étude. Celui-ci diffère légèrement de l'interprétation bayésienne généralement trouvée dans la littérature (*e.g.* [1]) sans toutefois proposer un point de vue réellement différent tel que peuvent le faire d'autres approches (*e.g.* [2]).

Les deux mesures sont introduites dans la section 1.1. Elles s'écrivent toutes les deux sous la forme de mesures finies dans la théorie de la mesure, ce qui peut prêter à confusion. Un exemple concret, le jet d'un dé à six faces, permet de mettre plus facilement en évidence leurs différences. La mesure de confiance peut être utilisée en particulier pour estimer les issues possibles d'une expérience dont la loi de probabilité est connue. La section 1.2 présente le lien particulier qu'il existe entre les deux notions dans ce cas. La section 1.3 présente le rapport entre les méthodes statistiques et la mesure de confiance. Elle étudie plus particulièrement le problème de l'estimation de la valeur inconnue d'un paramètre. Elle traite également à titre d'exemple de la composition de l'incertitude statistique et d'un biais inconnu. Enfin, la dernière section s'intéresse brièvement à deux sujets particuliers : celui de la composition des incertitudes théoriques et expérimentales ainsi que celui de l'extraction de paramètre.

1.1 Notions de probabilité et de confiance

La notion de probabilité est aujourd'hui couramment associée à la notion de fréquence, ce que certains appellent l'approche fréquentiste. Une autre approche de la notion coexiste avec celle-ci : celle dite bayésienne. À la lecture d'articles ou de livres sur cette dernière, on peut constater une certaine difficulté sur ce que l'on doit mesurer lorsque l'on souhaite dire qu'une chose, ou un événement est "probable" ou non. Le problème de la mesure de la "probabilité" est en fait bien plus délicat qu'il semble de prime abord.

Les "approches" bayésiennes et fréquentistes sont souvent présentées en opposition. L'approche bayésienne est même parfois présentée comme une généralisation de la seconde. Ces points de vue me semblent trop extrémistes de premier abord. Les deux méthodes m'ont toujours semblé à la fois valides et complémentaires. La suite de ce chapitre ne cherche donc pas à valider une approche au dépend de l'autre. Il s'agit plutôt de chercher à comprendre plus précisément chacune des deux approches, leurs sens, leurs différences et leurs domaines d'application. C'est en approfondissant dans ce sens que je me suis rendu compte que deux notions, souvent confondues, doivent être soigneusement distinguées : celles que j'appelle dans la suite probabilité et confiance. Les points de vue bayésien et fréquentistes n'apparaissent plus alors que comme deux manières différentes d'imbriquer ces deux concepts en vue d'estimer des incertitudes dans le cas particulier où l'on étudie les résultats de la répétition d'une expérience.

1.1.1 Exemple du jet de dé

1.1.1.1 Probabilité d'une expérience de jet de dé

La loi de probabilité du jet d'un dé à six faces est obtenue par la répétition, en théorie infinie, de cette expérience. Un dé pipé est un dé qui fournit une distribution de probabilité non uniforme. Un tel objet est créé en brisant la symétrie du cube, que ce soit par sa densité de masse ou par sa forme. Un coin du dé peut par exemple être limé. Réciproquement, si la loi de probabilité de l'expérience du jet du dé était précisément connue, et si cette loi était non uniforme, il serait possible de déduire que le dé est non symétrique.

La loi de probabilité contient une information sur la forme du dé. Elle traite de la réalité des choses. Plus précisément, elle caractérise l'expérience complète du jet du dé. Elle est fonction du dé lui-même, mais aussi de possibles restrictions sur la manière dont il est lancé, du nombre de rebonds réalisés, etc. De la même manière qu'une particule est caractérisée par sa masse ou qu'un solide peut l'être par ses dimensions, une expérience est caractérisée par sa loi de probabilité.

Le terme de "propension" est également utilisé pour désigner cette mesure. Celle-ci décrit en effet la tendance d'une expérience à donner un résultat plutôt qu'un autre. Le terme de "probabilité" est pourtant le plus communément utilisé aujourd'hui, que ce soit dans la littérature spécialisée, en enseignement ou dans la vie de tous les jours. Il porte pourtant à confusion. En effet, utilisé dans une expression telle que : "il est probable qu'il pleuve aujourd'hui", ce mot se rapporte plutôt à la notion de confiance décrite plus bas qu'à celle de propension.

1.1 Notions de probabilité et de confiance

1.1.1.2 Confiance sur le résultat d'un jet de dé

Un joueur pourrait être, quant à lui, intéressé par le résultat d'un *unique* jet. S'il ne connaît ni le dé ni la manière dont il va être lancé, il pourra supposer que tous les résultats ont une chance égale d'aboutir. Il peut également connaître précisément le dé, l'ayant utilisé pendant de nombreuses années, et prendre en compte une légère asymétrie. Il peut même connaître exactement la loi de probabilité de l'expérience mais décider de miser toutes ses économies sur la sortie du six. Cette valeur n'étant pas apparue une seule fois au cours des dix derniers lancers, ce joueur superstitieux pourrait considérer que ce résultat *doit* arriver.

Ces différents choix ne modifient pas l'expérience. Ils peuvent d'ailleurs être effectués simultanément par différentes personnes. L'issue du jet du dé n'est pas affectée par la confiance d'un observateur sur le résultat. Cette dernière n'a pas de rapport avec la réalité mais uniquement avec les informations considérées. En d'autres termes, ce n'est pas parce qu'une personne considère que le numéro six doit sortir que cela va être le cas.

À ce stade, rien n'empêche de chercher à *décrire* la confiance de n'importe quel observateur sur la valeur que va fournir le jet de dé parmi toutes les possibilités envisageables. Par exemple, la confiance de l'observateur superstitieux précédent sera *décrite* en associant une confiance de 0% aux valeurs un à cinq et 100% à la valeur six. Il s'agit d'une description de la confiance de cet individu, une mesure de sa confiance. Elle est finie : l'observateur est certain que le jet de dé va fournir un résultat.

Plusieurs mesures de confiance, décrivant les confiances de différents individus¹, peuvent être définies simultanément sur le même problème du jet de dé. Une mesure de confiance ne peut donc pas caractériser une expérience. Elle traite en fait d'*une* question particulière (ici, l'issue du jet de dé) et dérive les conséquences des informations considérées (la loi de probabilité de l'expérience, l'absence du numéro six dans les derniers résultats, ...) sur les réponses envisageables. Cette question peut être ou ne pas être liée au résultat d'une expérience. Les informations considérées ne sont quant à elles pas nécessairement véridiques, objectives ou rigoureuses. Seule la dérivation de leurs conséquences, si dérivation il y a, l'est.

Le terme de "croyance" ("*belief*") est souvent utilisé pour désigner cette mesure. J'ai décidé de ne pas l'utiliser ici pour ses connotations bien trop encrées dans le domaine spirituel.

1.1.2 Probabilité et reproductibilité

Puisque les mesures de probabilités et de confiance ne sont pas construites dans le même but, les interprétations de leurs valeurs doivent être également différentes. Les manières dont elles sont construites ou modélisées n'ont aucune raison d'être identiques.

1. La mesure de confiance n'est pas restreinte à la description de la confiance des personnes superstitieuses ou non rigoureuses. Toute personne, même éduquée ou scientifique, basera ses décisions sur sa propre vision du monde, sur sa propre confiance et non sur ce que l'on souhaiterait appeler la "réalité". La difficulté qui se présente à l'homme de science est bien entendu de minimiser l'impact de ses préconceptions sur ses choix.

1.1.2.1 Reproductibilité

Une loi de probabilité donne une information sur l'expérience dont elle est issue. Sa distribution est obtenue par des réalisations répétées. En particulier, la loi obtenue se rapproche d'une distribution de Dirac lorsque les résultats de l'expérience sont quasiment tous identiques. Elle est *exactement* égale à une loi de Dirac si tous les résultats sont *exactement* identiques. L'expérience est alors reproductible. Une loi de probabilité s'écartant de la loi de Dirac décrit une expérience donnant des résultats différents d'une réalisation à l'autre. L'étalement d'une loi de probabilité, donnée en partie par sa variance, est une mesure de la reproductibilité de l'expérience associée².

La fonction entropie peut s'appliquer à toute forme de densité. Elle mesure, dans le cas d'une loi de probabilité, le *manque de contraintes* définissant l'expérience considérée. Lorsque l'entropie est nulle, l'expérience est totalement contrainte : elle est reproductible.

1.1.2.2 Espace probabilisé et variable aléatoire

Dans de nombreux cas, la physique mise en jeu dans une expérience est connue pour être déterministe. Des états initiaux différents fournissent des résultats différents. Mais lorsque ces états sont trop proches pour être distingués, ils sont désignés de la même manière. Il arrive pourtant que des états initiaux voisins donnent des résultats visiblement différents. L'expérience apparaît alors comme étant aléatoire.

La loi de probabilité est un outil qui permet de modéliser cet effet. L'état initial fait partie de la définition de l'expérience et il est, à la précision utilisée, toujours le même. D'une réalisation à l'autre, il correspond pourtant à des états physiques exacts ω (inconnus) différents. La physique mise en jeu étant déterministe, les valeurs des quantités observables sont des fonctions de l'état physique ω . On appelle ces fonctions des variables aléatoires. Elles peuvent prendre des valeurs différentes d'une réalisation à une autre de l'expérience. On note Ω l'ensemble des états ω initiaux différents mais indistingués.

Dans le cas du jet de dé, l'état initial considéré peut être :

le dé est dans la main

Cet état est très peu précis. En fait, la définition du reste de l'expérience l'est généralement aussi. Elle peut être par exemple :

le dé est jeté sur la table

Le déroulement exact du jet lui-même change d'un cas sur l'autre. Un élément ω de Ω ne désigne donc pas simplement un état initial mais plus généralement toute une réalisation de l'expérience. L'ensemble Ω est celui des déroulements possibles qui vérifient une définition générale choisie. Plus la définition générale est contrainte, plus l'expérience devrait être reproductible.

L'espace Ω n'est pas nécessairement bien défini. C'est un outil qui permet de modéliser le lien entre les variables aléatoires. Il permet de rappeler qu'une réalisation précise ω de l'expérience leur donne à chacune une valeur simultanément. En pratique, on construit empiriquement la mesure de probabilité sur les résultats des variables aléatoires.

2. Et en aucun cas une mesure d'incertitude!

1.1.3 Confiance et incertitude

1.1.3.1 Incertitude

La mesure de confiance a pour objectif premier de décrire les caractéristiques de la confiance d'un observateur sur la réponse à une question. Par extension, une telle mesure peut être créée en encodant un jeu d'informations. Si ces données sont suffisantes, la mesure doit pouvoir encore une fois décrire une confiance sur le résultat d'une question, étant donné cette fois le jeu d'informations qui lui a été fourni à sa création. Cette mesure de confiance est en effet maintenant une conséquence directe des informations employées pour la définir.

Une mesure de confiance décrit donc puis dérive les conséquences d'informations qui sont codées en elle lorsqu'elle est définie. Ces dernières correspondent typiquement à l'ensemble des connaissances qu'une personne possède en relation à une question donnée. La distribution de confiance prend la forme d'une distribution de Dirac si et seulement si cette personne possède suffisamment d'informations pour être certaine de la réponse à la question étudiée. L'étalement de la distribution traduit l'étendue des réponses que cette personne peut envisager étant données ses connaissances. Sa valeur, partiellement donnée par sa variance, estime l'écart à la certitude. L'étalement d'une distribution de confiance est donc une mesure d'incertitude.

La mesure de confiance ne permet pas de valider ou d'infirmer les informations utilisées. Qu'une personne soit certaine de posséder la réponse à une question ne signifie pas qu'elle ait raison. Elle permet par contre de dériver rigoureusement les conséquences des informations considérées et d'éviter des incohérences de raisonnement. Elle fournit un cadre clair qui permet de contrôler les hypothèses utilisées.

La fonction entropie s'applique à toute densité de confiance. Elle mesure le *manque d'information* de la mesure de confiance. Lorsque l'entropie est nulle, l'opérateur est certain de l'issue de la question étudiée (qu'il ait raison ou tort). L'information considérée est complète.

1.1.3.2 Modèle de confiance et variable inconnue

La mesure de confiance est, comme son nom l'indique, une mesure. Elle est également, comme précisé plus tôt, finie. Elle partage le formalisme des mesures finies de la théorie de la mesure avec les probabilités. De la même manière que la mesure de Lebesgue peut être utilisée pour mesurer une longueur, une température ou un temps, rien n'empêche le formalisme des mesures finies d'être appliqué à la mesure de notions différentes.

Un *modèle de confiance* peut être créé sur tout sujet³ en définissant *un espace* correspondant aux paramètres dont on recherche les valeurs (fixées mais inconnues) puis *une mesure de confiance* sur cet espace. Cette dernière ne correspond pas à une donnée empirique. Elle ne peut pas être construite en suivant un algorithme précis comparable à la répétition d'une expérience. La manière dont elle sera définie dépendra des informations particulières que l'on prendra en compte et de la façon dont celles-ci seront traduites.

3. Le problème étudié n'est pas nécessairement en lien avec une expérience, répétée ou non. Une mesure de confiance peut être définie dès qu'une incertitude existe. En pratique, dès qu'il est envisageable d'effectuer un pari sur un sujet, une mesure de confiance peut être construite. S'il y a pari, il y a une issue incertaine. S'il y a une issue incertaine, il y a matière à détailler, décrire, mesurer cette incertitude. D'où une mesure. La quantité d'argent qu'un individu serait prêt à parier sur la valeur d'un paramètre inconnu traduit d'ailleurs essentiellement son degré d'incertitude.

Les points suivants présentent les grands étapes qui permettent de construire un modèle.

DÉFINITION DE L'ESPACE DE TRAVAIL DU MODÈLE

Un modèle de confiance est créé en listant tout d'abord l'intégralité des paramètres x liés à l'étude, que leur valeur soit connue ou non. Une variable $X \in \mathcal{D}_x$ est associée à chacun d'entre eux. L'ensemble de définition \mathcal{D}_x est celui des valeurs envisageables pour le paramètre x . Spécifiquement, $\mathcal{D}_x = \mathbb{R}$ s'il est certain que le paramètre est réel. Ces variables ne sont pas des fonctions. Pour les distinguer des variables aléatoires, elles seront appelées dans la suite des variables inconnues. Les valeurs de leurs paramètres associés sont fixes, mais pas nécessairement connues. L'ensemble de ces variables et de leur ensembles de définition fixe l'espace de travail.

Le choix de l'ensemble de définition \mathcal{D}_x est important. Il s'agit de l'une des *certitudes* insérées dans le modèle. Si un paramètre est supposé réel, *i.e.* si l'ensemble de définition de sa variable inconnue est fixé à l'ensemble \mathbb{R} , la valeur de confiance prédite sur des valeurs complexes sera à jamais nulle, et ce quelles que soient les informations ultérieures ajoutées. Plus généralement, l'annulation d'une densité de confiance sur un ensemble A quelconque est un effet irréversible. Quelle que soient les nouvelles informations considérées, la confiance que la valeur du paramètre étudié se situe dans A sera toujours nulle : rien ne pourra lui faire changer d'avis (voir section 1.1.3.3). En règle générale, ces absolutismes sont à manipuler avec précaution.

RESTRICTION DE L'ESPACE DE TRAVAIL DU MODÈLE

Une mesure de confiance est ensuite définie sur l'espace considéré en traduisant un lot d'informations. Le cas du joueur superstitieux précédent est une exception : la valeur de la mesure ne peut généralement pas être fixée aussi simplement. Chacune des informations considérées fournira plutôt une ou plusieurs contraintes sur la forme de la mesure. Une fois toutes les informations traduites, l'ensemble des contraintes obtenues devraient définir de manière univoque la mesure recherchée. Si ce n'est pas le cas, les informations considérées sont insuffisantes pour créer un modèle.

Les informations prises en compte peuvent être de natures ou d'origines très différentes. Tout d'abord, certains liens entre des paramètres peuvent être connus avec certitude. Dans le meilleur des cas, un paramètre peut être défini comme une fonction des autres. Par exemple, un paramètre x_3 peut être défini comme la somme de deux paramètres x_1 et x_2 . Pour traduire cette information, la mesure de confiance de l'espace \mathbb{R}^3 recherchée doit être localisée sur le sous-espace qui vérifie cette condition :

$$f_{\mathbb{C}}(X_1, X_2, X_3) = \delta(X_3 - (X_1 + X_2))f_{\mathbb{C}}(X_1, X_2) \tag{1.1}$$

L'indice \mathbb{C} permettra de distinguer dans la suite des densités de confiance de densités de probabilités $f_{\mathbb{P}}$.

En effet, cette équation se réécrit :

$$f_{\mathbb{C}}(X_3|X_1, X_2) = \delta(X_3 - (X_1 + X_2)) \tag{1.2}$$

La densité conditionnelle $f_{\mathbb{C}}(X_3|X_1, X_2)$ devra s'annuler en tout point excepté en $X_3 = X_1 + X_2$. Autrement dit, étant connues les valeurs de X_1 et de X_2 , la mesure devra traduire la certitude selon laquelle le paramètre X_3 ne peut prendre pour valeur que leur somme.

1.1 Notions de probabilité et de confiance

En particulier, quelles que soient les informations ajoutées par la suite, tout se passera comme si l'étude était effectuée dans \mathbb{R}^2 . L'espace de travail est alors restreint à celui des couples (X_1, X_2) des paramètres que l'on peut qualifier de libres x_1 et x_2 , la confiance sur la valeur du paramètre x_3 s'écrivant :

$$f_{\mathbb{C}}(X_3) = \int \delta(X_3 - (X_1 + X_2)) f_{\mathbb{C}}(X_1, X_2) dX_1 dX_2 \quad (1.3)$$

Plus généralement, si un paramètre x_3 est une fonction de $g(x_1, x_2)$, on utilise l'espace de travail \mathbb{R}^2 des couples (X_1, X_2) . La confiance sur les valeurs envisageables de x_3 est alors :

$$f_{\mathbb{C}}(X_3) = \int \delta(X_3 - g(X_1, X_2)) f_{\mathbb{C}}(X_1, X_2) dX_1 dX_2 \quad (1.4)$$

LIENS OU INDÉPENDANCE DES INFORMATIONS SUR LES PARAMÈTRES

Les informations sur les valeurs de certains paramètres peuvent à leur tour être liées. Une même source d'information peut fournir une estimation des valeurs de deux paramètres distincts. La valeur de l'un peut également être une indication de la valeur de l'autre. Lorsqu'aucun de ces cas n'est réalisé, les variables inconnues sont construites indépendantes : aucune information parmi celles considérées ne les relie. En terme de densités, si les informations collectées sur les paramètres x_1 et x_2 ne sont pas liées :

$$f_{\mathbb{C}}(X_1, X_2) = f_{\mathbb{C}}(X_1)f_{\mathbb{C}}(X_2) \quad (1.5)$$

Cette équation se réécrit à l'aide des densités conditionnelles et marginales :

$$f_{\mathbb{C}}(X_2|X_1) = f_{\mathbb{C}}(X_2) \quad (1.6)$$

En général, les informations sur les valeurs des paramètres seront liées. Cela se traduira sous la forme d'une densité conditionnelle $f_{\mathbb{C}}(X_2|X_1)$ variant avec X_1 . Cette dépendance peut être supposée et utilisée lors de la construction du modèle ou bien être déduite des autres informations présentes par le calcul. Dans le premier cas, une densité conditionnelle $f_{\mathbb{C}}(X_2|X_1)$, à une normalisation près, est appelée dans la littérature une *fonction de vraisemblance*.

PRIORS SUR LA VALEUR DE CERTAINS COEFFICIENTS

En dernier lieu, des informations peuvent également définir avant tout calcul une confiance sur les valeurs envisageables d'une partie des paramètres. Certains peuvent être considérés comme parfaitement connus auquel cas les lois de confiance sur leurs valeurs sont des lois de Dirac. D'autres peuvent être considérés comme partiellement connus. On modélise cette information en fixant la valeur d'une densité marginale $f_{\mathbb{C}}(X_1)$ particulière. Une telle distribution est usuellement appelée un *prior*.

1.1.3.3 Utilisation pratique d'un modèle de confiance

MODÈLE SIMPLE DE CONFIANCE À DEUX VARIABLE

En pratique, un observateur est souvent intéressé par la valeur d'un paramètre en particulier, par exemple x_1 . Les valeurs envisageables d'un paramètre secondaire x_2 peuvent être sensibles à la valeur de x_1 , à cause d'un lien théorique entre les deux par exemple.

Ce second paramètre joue alors le rôle d'une source indirecte d'information sur la valeur de x_1 . On construit un modèle de confiance pour décrire rigoureusement l'incertitude sur les valeurs de ces paramètres.

Les paramètres inconnus du problème sont x_1 et x_2 . Deux variables inconnues, réelles par exemple, leur sont associées : X_1 et X_2 . Elles définissent l'espace de travail.

Les valeurs de x_1 considérées comme possibles avant que toute information ne soit tirée de l'estimation de x_2 sont décrites à l'aide d'une densité $f_C(X_1)$. Les valeurs possibles de x_2 , étant supposée une valeur possible X_1 du paramètre de valeur x_1 exacte inconnue, sont décrites par une densité conditionnelle $f_C(X_2|X_1)$. Sa forme dépendra du problème étudié et du lien particulier supposé exister entre x_1 et x_2 . La mesure de confiance sur \mathbb{R}^2 est alors définie par la densité :

$$f_C(X_1, X_2) = f_C(X_2|X_1)f_C(X_1) \quad (1.7)$$

Ces éléments fixent le modèle de confiance.

ACTUALISATION DE LA CONFIANCE, PRIOR, POSTERIOR ET VRAISEMBLANCE

Avant que toute expérience de mesure du paramètre x_2 ne soit effectuée, sa valeur n'est pas connue. La mesure de confiance sur la valeur du premier paramètre est alors décrite en intégrant sur toutes les valeurs X_2 possibles :

$$\int f_C(X_1, X_2) dX_2 = f_C(X_1) \quad (1.8)$$

Il s'agit de la densité utilisée dans la construction de la mesure. Elle décrit proprement l'incertitude *a priori* sur la valeur de x_1 , *i.e.* avant que l'on ait cherché à collecter une nouvelle information au travers de x_2 . C'est pour cela qu'elle porte généralement le nom de *prior*.

Lorsque la valeur de x_2 est, pour une raison ou une autre, découverte, la densité $f_C(X_1)$ ne décrit plus l'incertitude de l'opérateur. Celui-ci possède une nouvelle information : la valeur du paramètre x_2 . Sa confiance est donc maintenant mesurée par $f_C(X_1|X_2 = x_2)$. Elle a pour valeur :

$$f_C(X_1|X_2) = \frac{f_C(X_2|X_1)}{f_C(X_2)} f_C(X_1) \quad (1.9)$$

En termes bayésiens, il est dit que la connaissance de x_2 "actualise" la densité sur X_1 . La densité conditionnelle $f_C(X_1|X_2)$ est d'ailleurs appelée un *posterior*. Bien entendu, la mesure de confiance sur \mathbb{R}^2 en elle-même n'a pas changée. La "densité sur X_1 ", quelle qu'elle soit, n'a pas non plus changé. Simplement, alors que l'incertitude sur x_1 était décrite par $f_C(X_1)$ avant la mesure de x_2 , elle est décrite par $f_C(X_1|X_2)$ après.

Plus généralement, pour un modèle de confiance quelconque, l'ajout d'une information se traduira le plus souvent par la restriction de l'espace de travail initial à un de ses sous-espace. La mesure de confiance décrivant le nouveau niveau d'incertitude est alors simplement la mesure héritée de la mesure générale.

Le rapport du membre de droite de l'équation (1.9) est appelé vraisemblance $\mathcal{V}(X_2|X_1)$ (*likelihood*, $\mathcal{L}(X_2|X_1)$). La valeur de X_2 étant fixée dans cette équation, le terme $f_C(X_2)^{-1}$ n'est qu'un terme de normalisation. On parle donc également de vraisemblance pour la densité conditionnelle $f_C(X_2|X_1)$.

L'équation (1.9) montre que si la confiance sur la valeur du paramètre x_1 s'annule sur un domaine $A \subset \mathcal{D}_x$, toute valeur actualisée de la confiance s'y annulera également.

1.1 Notions de probabilité et de confiance

L'annulation d'une densité sur un domaine traduit donc bel et bien une certitude. Quelles que soient les informations collectées par la suite, la confiance pour que x_1 appartienne à A sera à jamais nulle.

ACTUALISATION ET MODÈLE DE CONFIANCE

L'équation (1.9) s'obtient par deux utilisations consécutives de l'expression d'une densité conditionnelle :

$$f(a, b) = f(a|b)f(b) \quad (1.10)$$

Pour des raisons historiques, elle porte le nom de formule (ou théorème) de Bayes. Cette équation (et le modèle à deux paramètres inconnus précédent) est à la base des raisonnements dits bayésiens.

Il pourrait être tentant de chercher à ne pas s'embarasser d'une définition complète du modèle de confiance. En effet, en utilisant de manière répétée la formule de Bayes, on pourrait chercher à actualiser la mesure d'incertitude sur X_1 après toute nouvelle information indirecte I reçue :

$$\begin{aligned} & f_{\mathbb{C}}^{(0)}(X_1) \\ & \quad \downarrow \\ & f_{\mathbb{C}}^{(1)}(X_1) = \mathcal{V}(I_1|X_1)f_{\mathbb{C}}^{(0)}(X_1) \\ & \quad \downarrow \\ & f_{\mathbb{C}}^{(2)}(X_1) = \mathcal{V}(I_2|X_1)f_{\mathbb{C}}^{(1)}(X_1) \\ & \quad \downarrow \\ & \dots \end{aligned} \quad (1.11)$$

Cette approche n'est toutefois utilisable que lorsque la mesure de confiance est explicitement construite à partir d'un prior $f_{\mathbb{C}}^{(0)}(X_1)$ sur la valeur du paramètre recherché et de fonctions de vraisemblances $\mathcal{V}(I_n|X_1)$ successives, tous supposés donc connus. Ce n'est pas toujours le cas (cf par exemple le chapitre 2).

Cette manière d'aborder la mesure de confiance est également risquée. Raccourcir les notations peut en effet mener l'utilisateur à faire des erreurs. Il se pourrait par exemple que les informations I_1 et I_2 soient partiellement redondantes ou liées. Or, le passage de $f_{\mathbb{C}}^{(1)}(X_1)$ à $f_{\mathbb{C}}^{(2)}(X_1)$ requiert l'indépendance des informations successives :

$$\mathcal{V}(I_1, I_2|X_1) = \mathcal{V}(I_1|X_1)\mathcal{V}(I_2|X_1) \quad (1.12)$$

Reprenant le modèle à deux variables précédent, un groupe d'individus pourrait utiliser la mesure de x_2 pour actualiser leur incertitude sur x_1 et publier le posterior. Un autre groupe pourrait utiliser pour sa part la mesure de $x_3 = x_1 + x_2$ en utilisant le posterior précédent comme leur prior. La même fonction de vraisemblance serait alors appliquée à deux reprises à tort. L'utilisation d'un modèle propre, qui liste l'ensemble des données, informations et hypothèses employées, permet justement d'éviter ce genre de problèmes.

Ce dernier point est l'occasion d'effectuer une remarque générale sur l'appareil de la mesure de confiance. L'intérêt d'un modèle de confiance n'est pas tant de fournir une estimation de l'incertitude sur la valeur d'un paramètre recherché (ce qui, effectué grossièrement, est une tâche relativement simple) que de contrôler les informations utilisées dans celle-ci. Il ne s'agit pas d'obtenir une valeur de niveau de confiance mais plutôt

de connaître et *maîtriser* les différentes sources d'incertitude, leur imbrication et leurs contributions.

1.2 Confiance sur le résultat d'une expérience

1.2.1 Mesure de confiance et information

Au final, la mesure de confiance ne fait que dériver de manière rigoureuse les conséquences des informations qui lui sont fournies. La qualité de ces dernières définit donc l'intérêt des résultats obtenus. Il est par conséquent critique de contrôler les hypothèses utilisées dans la création d'une mesure particulière. Il ne s'agit pas d'un défaut de la mesure de confiance mais au contraire de l'un de ses intérêts principaux. La définition rigoureuse d'un modèle de confiance permet de lister précisément les informations prises en compte lors de l'étude d'une question. Comme présenté au chapitre 2, il permet même dans certains cas de mettre en évidence une incompatibilité entre certaines hypothèses.

En règle générale, il semble plus prudent de privilégier des modèles simples. Un modèle trop complexe est un modèle dont les informations fondatrices risquent de ne plus être contrôlées. Ses résultats ne peuvent alors plus être discutés et le modèle de confiance est inutilisable.

Un modèle de confiance peut en particulier traiter du résultat de la réalisation *unique* d'une expérience comme c'était le cas dans le jet de dé de la section précédente. Lorsqu'elle est connue, sa loi de probabilité est une donnée précieuse. Elle est objective comme peut l'être la donnée des dimensions d'un solide : cette information ne peut pas être remise en question. Les deux prochains paragraphes discutent de ce cas un peu plus en détails.

1.2.2 Mesures de confiance et de probabilité

Les notions de probabilité et de confiance sont fréquemment confondues. Elles s'expriment dans le même formalisme mathématique et il n'est pas nécessaire de faire la distinction entre les deux dans de nombreux cas : les deux mesures sont souvent numériquement identiques.

On se propose de construire une mesure de confiance sur le résultat *d'une réalisation particulière* du jet de dé de la section 1.1.1. Le seul paramètre inconnu est la valeur x obtenue dans ce lancé. On le modélise par une variable inconnue X définie sur $\{1, 2, 3, 4, 5, 6\}$. Supposons que le jet considéré remplisse les conditions de la définition de l'expérience générique de jet de dé présenté dans la section 1.1.2 :

Le dé est dans la main. Il est jeté sur la table.

Supposons également que la loi de probabilité \mathbb{P} de cette expérience générale soit connue. Une personne sensée choisirait alors la définition de la loi de confiance \mathbb{C} :

$$\mathbb{C}(X = i) = \mathbb{P}(\tilde{X} = i) \Leftrightarrow \mathbb{C}(X) = \mathbb{P}(\tilde{X} = X) \quad (1.13)$$

\tilde{X} est la variable aléatoire associée au résultat d'un jet générique et \mathbb{P} sa mesure de probabilité.

Ce choix n'est pas unique ! Un autre joueur peut avoir une confiance différente. Le joueur superstitieux précédent avait par exemple une confiance de 100% sur la sortie

1.2 Confiance sur le résultat d'une expérience

du six et de 0% sur les autres valeurs⁴. La mesure de confiance n'existe pas pour juger du meilleur choix à effectuer. Elle décrit une possibilité de confiance sur les issues de la question considérée.

La question qui doit donc nous intéresser est de savoir *pourquoi* l'on considère que le choix de l'équation (1.13) est le plus raisonnable. Après tout, la loi de probabilité correspond à la répétition de l'expérience un très grand nombre de fois alors que la mesure de confiance ne s'intéresse qu'à l'une de ses réalisations. Pourquoi ne pas définir la mesure de confiance différemment de la loi de probabilité ?

En fait, quelle que soit la question considérée (le résultat est-il supérieur à 3 ? Est-il pair ? Est-il égal à six ? ...) et quel que soit le gain considéré par réponse, il s'agit de la meilleure manière de définir la mesure de confiance *sur le long terme*. Autrement dit, si le joueur pouvait répéter ce pari une infinité de fois, il maximiserait son gain. Réciproquement, cette mesure est la seule à maximiser asymptotiquement le gain quels que soient la question étudiée et les gains possibles considérés. Une preuve de cette assertion est donnée dans le paragraphe suivant. Bien entendu, le joueur ne peut pas effectuer une infinité de paris. Il est possible qu'il soit perdant à chaque fois. Ce "meilleur" choix de mesure de confiance n'assure en aucun cas le succès.

Le fait que la loi de confiance optimale soit égale à la loi de probabilité de l'expérience considérée est en fait plus profond. La mesure de confiance doit être construite à l'aide des informations disponibles sur les issues possibles de la réalisation unique de l'expérience considérée. Or, on suppose ne pas posséder de donnée sur cette réalisation particulière. La seule information qui peut alors être utilisée est celle de la définition de l'expérience en elle-même. La loi de probabilité contient précisément des informations de l'expérience, sans information supplémentaire superflue ou non justifiée.

La confiance dans la valeur du résultat d'une expérience est un cas d'étude privilégié. En général, il n'existe pas de mesure de confiance meilleure que les autres. Seul l'étalement (ou l'entropie) des densités de confiance permettrait de définir une relation d'ordre entre elles mais puisque la validité des informations n'est pas toujours certaine, effectuer une comparaison entre différentes mesures est en fait un exercice beaucoup plus délicat.

CHOIX DE CONFIANCE LE PLUS RAISONNABLE

Supposons dans un premier temps que le pari ne porte que sur la valeur obtenue dans le jet de dé. Le joueur ne peut parier que sur un résultat à la fois. Le gain potentiel obtenu pour un pari réussi sur la sortie de la valeur i du dé est noté g_i . La mise initiale, toujours perdue, est notée m . Le gain obtenu lors d'un pari *unique* est une variable inconnue G_i où i est la valeur choisie par le joueur. Il s'écrit :

$$G_i = g_i \chi_{X=i} - m \quad (1.14)$$

4. La procédure d'apprentissage par actualisation de la mesure de confiance après plusieurs jets de dés ne change rien à ce problème. En effet, la valeur de la mesure de confiance va tendre numériquement vers celle de la mesure de probabilité après une infinité de jets quel que soit le prior (ne s'annulant nul part) si et seulement si les résultats des différents jets sont supposées non liés (dans les sous-espaces à mesure de probabilité connue, *i.e.* dans les sous-espaces où toute l'information tirée de la définition de l'expérience est connue). Cette hypothèse n'est clairement pas effectuée par un joueur superstitieux qui considère que le nombre six doit sortir au prochain lancer *parcequ'il* n'est pas sorti aux précédents. Ce joueur considère précisément que les résultats des jets précédents influent sur le résultat du jet suivant, quand bien même la loi de probabilité de l'expérience lui serait connue.

La fonction χ_A est la fonction caractéristique d'un ensemble A . Par extension, on note χ_b la fonction qui renvoie un lorsque la condition booléenne b est vraie et zéro dans le cas contraire. Le gain G_i est donc une variable inconnue qui modélise l'incertitude sur la valeur du gain réel, fixée et inconnue, et qui vaut $-m$ ou $g_i - m$.

Le joueur choisit d'effectuer son pari sur une valeur i particulière en fonction des informations dont il tient compte. Celles-ci lui font avoir une plus grande confiance dans ses possibilités de gain pour un choix de résultat plutôt qu'un autre. Plus précisément, il choisit la valeur i_M qui maximise l'espérance (pour la mesure de confiance) de son gain G_i :

$$\mathbb{E}_{\mathbb{C}}(G_{i_M}) = \max_i \{\mathbb{E}_{\mathbb{C}}(G_i)\} \quad (1.15)$$

Ces espérances s'écrivent en fonction des valeurs des confiances $c_i = \mathbb{C}(X = i)$:

$$\mathbb{E}_{\mathbb{C}}(G_i) = g_i c_i - m \quad (1.16)$$

La valeur du résultat parié i_M dépend donc des valeurs des gains $\{g_i\}_{i=1,\dots,6}$.

Supposons que le joueur réalise ce pari une infinité de fois. À chaque itération, le joueur joue sur le résultat i_M qui, selon les informations qu'il prend en compte pour déterminer sa confiance, lui donne la plus grande chance de gain. On suppose de plus que le joueur ne change pas d'une fois sur l'autre les informations qu'il utilise et qu'il connaisse la mesure de probabilité de l'expérience. Son choix de valeur i_M reste donc le même pour toutes les réalisations. On souhaite montrer que le choix de mesure de confiance (1.13) est l'unique cas qui maximise le gain total quel que soit le choix des g_i et qu'il s'avère donc être le choix le plus raisonnable.

Le gain total sur une infinité d'expériences a toutes les chances de tendre vers plus ou moins l'infini. Il est mieux décrit à la limite par le gain moyen, c'est à dire le rapport de la somme des gains individuels sur le nombre de jets du dé. Lorsque le nombre de réalisations de l'expérience tend vers l'infini, il s'agit précisément de la valeur de l'espérance de la variable aléatoire \tilde{G} correspondant au gain du joueur dans l'expérience répétée. On note \tilde{G}_i le gain qui serait obtenu si le joueur effectuait le choix i plutôt que le choix i_M . Ces gains s'écrivent :

$$\tilde{G}_i = g_i \chi_{\tilde{X}=i} - m \quad ; \quad \tilde{G} = \tilde{G}_{i_M} \quad (1.17)$$

Leurs espérances s'écrivent en fonction des probabilités $p_i = \mathbb{P}(\tilde{X} = i)$:

$$\mathbb{E}_{\mathbb{P}}(G_i) = g_i p_i - m \quad (1.18)$$

La relation que l'on cherche à démontrer s'écrit finalement :

$$\left(\forall \{g_i\}_{i=1,\dots,6}, \quad \mathbb{E}_{\mathbb{P}}(G_{i_M}) = \max_i \{\mathbb{E}_{\mathbb{P}}(G_i)\} \right) \Leftrightarrow (\forall i = 1, \dots, 6, \quad p_i = c_i) \quad (1.19)$$

Le sens réciproque est immédiat par définition (1.15) de i_M et c'est donc le sens direct qui nous intéresse dans la suite.

Supposons que l'espérance probabiliste est maximisée pour le gain \tilde{G}_{i_M} quel que soit le choix de paramètres $\{g_i\}$. En particulier, il est maximal pour les valeurs :

$$g_i = \frac{1}{p_i}, \quad \forall i \neq i_0, \quad g_{i_0} = \frac{1}{p_{i_0} - \epsilon} \quad (1.20)$$

1.2 Confiance sur le résultat d'une expérience

et ce quel que soit la valeur de ϵ strictement positive. On peut supposer que toutes les probabilités p_i sont non nulles, quitte à redéfinir l'expérience et imaginer que l'on utilise en fait un dé à cinq faces.

Par construction, l'espérance du gain \tilde{G}_{i_0} est maximale. Le paramètre i_0 est donc égal au paramètre i_M . L'espérance au sens de la mesure de confiance du gain G_{i_0} est donc à son tour le maximum des espérances des gains. Cela se réécrit :

$$c_{i_0} \frac{1}{p_{i_0} - \epsilon} - m \geq c_i \frac{1}{p_i} - m, \quad \forall i \neq i_0, \quad \forall \epsilon > 0 \quad (1.21)$$

À la limite :

$$\frac{c_{i_0}}{p_{i_0}} \geq \frac{c_i}{p_i}, \quad \forall i \neq i_0 \quad (1.22)$$

En redéfinissant les paramètres de gain $\{g_i\}$ on montre que cette relation est valable quelle que soit la valeur du paramètre i_0 . Les rapports sont donc tous égaux. Les mesures de probabilité et de confiance sont proportionnelles et donc égales par normalisation.

Le choix (1.13) de mesure est donc le seul candidat qui pourrait maximiser le gain quelle que soit la question (le résultat est-il pair?, ...) envisagée. En effet, il s'agit de la seule mesure qui maximise le gain dans le cas où la question est simplement celle de la valeur du résultat du jet de dé. Réciproquement, il est clair que la mesure de confiance définie comme étant numériquement égale à la mesure de probabilité maximisera le gain sur le long terme quelle que soit la question étudiée. La mesure de confiance (1.13) sur une unique réalisation d'une expérience est donc bel et bien le choix de mesure le plus raisonnable lorsque l'on connaît la mesure de probabilité associée.

1.2.3 Définition de l'expérience considérée

Dans l'exemple précédent, l'expérience était fixée et la loi de probabilité connue. En pratique, la loi de probabilité est souvent inconnue et l'expérience n'est pas nécessairement clairement définie.

Supposons que le joueur ait vu le dé tomber de sa main, la face 2 vers le haut, entendu ses deux rebonds, mais n'ait pas vu le résultat. Ce jet particulier du dé s'inscrit toujours dans la définition de l'expérience générique :

Le dé est dans la main. Il est jeté sur la table.

Si sa loi de probabilité est connue, le joueur peut décider d'utiliser la définition (1.13) de la mesure de confiance.

Un joueur plus avisé remarquerait que ce jet de dé s'inscrit également dans le cadre d'une expérience plus contrainte :

Le dé est dans la main, le numéro 2 face vers le haut. Il est jeté sur la table et rebondit précisément deux fois.

Cette expérience possède a priori une loi de probabilité différente de la première. Elle est plus contrainte et devrait être plus reproductible. L'étalement de sa loi de probabilité devrait être moins importante. Par conséquent, l'incertitude mesurée par la nouvelle loi de confiance définie par (1.13) devrait être plus faible également.

La totalité de l'information accessible peut être utilisée pour redéfinir l'expérience envisagée. Le problème de probabilité change donc à chaque fois. Par contre, le problème de confiance (la question posée, la réponse inconnue mais fixée) reste le même : il n'est

pas lié à l'expérience. Étant données certaines informations, l'expérience compatible la plus contrainte devrait définir la plus petite incertitude. Par exemple, le joueur pourrait idéalement avoir vu le résultat du dé. L'expérience la plus contrainte serait celle d'un jet de dé qui donne toujours ce même résultat. Sa loi de probabilité serait, par construction, une loi de Dirac. L'information utilisée dans le modèle de confiance (1.13) serait donc complète et l'incertitude nulle.

L'utilisation de la totalité de l'information accessible n'est pas nécessairement le meilleur choix qui puisse être effectué. Dans l'exemple précédent, le joueur pourrait connaître précisément la mesure de probabilité du jet de dé simple mais pas celle du jet avec la face du numéro deux dirigée vers le haut. Deux possibilités s'offrent à lui : négliger la nouvelle information ou tenter de l'ajouter sous la forme d'un modèle. Le premier choix est motivé par le principe général de créer des mesures simples aux informations contrôlées et d'origine certaine. Mais la deuxième approche reste attirante. Son succès dépendra expressément de la qualité du modèle utilisé. Pourtant, quelle que soit la dextérité du joueur à prendre en compte cette nouvelle information, la mesure de confiance ainsi construite restera ouverte à la controverse.

1.3 Confiance sur la valeur d'un paramètre inconnu

L'utilisation de la mesure de confiance (ou de probabilité *bayésienne*) est généralement présentée comme une alternative aux méthodes statistiques pour l'étude des problèmes d'inférence. Cette section s'attache donc à l'étude du rapport qu'il peut exister entre ces deux approches. Elle traite plus spécifiquement de l'exemple de l'estimation de la valeur inconnue d'un paramètre d'une loi de probabilité⁵.

Un exemple concret permet de fixer les idées. La suite de la section étudie le cas de l'estimation de la moyenne d'une loi normale. On note $f_{\mathbb{P}}^{(\theta)}$ la densité de probabilité sur la variable aléatoire \tilde{X} :

$$f_{\mathbb{P}}^{(\theta)}(\tilde{X}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\tilde{X}-\theta)^2}{\sigma^2}} \quad (1.23)$$

Le paramètre θ , fixé mais inconnu, est le paramètre recherché.

Ce problème n'est qu'un cas particulier d'estimation de l'incertitude sur la "valeur" de la mesure de probabilité d'une expérience. En effet, si la moyenne de la densité de probabilité n'est pas connue cela signifie que la loi de probabilité n'a pas été reconstruite empiriquement. De la même manière que le résultat d'un jet de dé est inconnu tant que ce dé n'a pas été lancé, une loi de probabilité ne peut pas être connue avant d'avoir été observée, c'est à dire avant d'avoir réalisé l'expérience associée un très grand nombre de fois. L'expression (1.23) n'est donc pas une information expérimentale, objective, mais une hypothèse.

Le problème est similaire à celui de l'exemple des sections précédentes. La mesure de probabilité est maintenant le paramètre inconnu. Une variable inconnue peut être définie sur l'ensemble des mesures finies. Une information supplémentaire est toutefois ajoutée ici pour simplifier l'étude. La mesure est supposée appartenir à une famille de mesures de densités $\{f_{\mathbb{P}}^{(\theta_1, \dots, \theta_n)}\}$. Cela restreint l'espace de travail à un simple espace numérique \mathbb{R}^n

5. Le problème des tests d'hypothèses, comme celle de l'existence du boson de Higgs, est très différent. L'approche fréquentiste, son interprétation et le traitement de ses résultats en terme de mesure de confiance est plus complexe dans ce cas que dans le cas de l'estimation de paramètre. Ce sujet n'est pas présenté ici.

1.3 Confiance sur la valeur d'un paramètre inconnu

de dimension finie correspondant au jeu de paramètres $(\theta_1, \dots, \theta_n)$ nécessaire pour fixer la valeur d'une mesure particulière. L'équation (1.23) correspond à un exemple de famille de lois de probabilité à $n = 1$ paramètre. Cette méthode d'estimation de la loi de probabilité est dite paramétrique.

Deux approches sont possibles à ce stade. La première est celle de l'inférence *bayésienne*. Une mesure de confiance est créée sur l'espace \mathbb{R}^n des paramètres inconnus. Pour cela, une ou plusieurs informations supplémentaires doivent être ajoutées sous la forme de *priors*. Cette approche est présentée à la sous-section 1.3.2. L'inférence *fréquentiste* (sous-section 1.3.1) utilise quant à elle la nature particulière de l'objet recherché, à savoir une mesure de probabilité. Cette méthode statistique permet d'obtenir un ou des encadrements sur les valeurs des paramètres sans utiliser cette fois-ci d'information complémentaire. Une certaine rigidité des résultats obtenus est toutefois le prix à payer.

Au bout du compte, l'approche la plus adaptée dépendra certainement du problème étudié, de la qualité et de l'objectivité des informations utilisables (paramétrisation de la loi de probabilité, éventuels priors) et de l'importance de la foi que l'utilisateur doit pouvoir porter sur le résultat obtenu. La sous-section 1.3.3 tente d'illustrer cette difficulté sous la forme de deux exemples.

1.3.1 Estimation statistique : approche "fréquentiste"

1.3.1.1 Expérience statistique

Dans la suite de cette section, j'appelle l'expérience dont on recherche la loi de probabilité (1.23) une *expérience élémentaire*. J'appellerai *expérience statistique* une répétition de N expériences élémentaires, N étant généralement large. Une variable aléatoire d'une expérience statistique est usuellement appelée une statistique.

Dans de nombreux cas, les N expériences élémentaires répétées sont identiques et leurs réalisations indépendantes. Une variable aléatoire \tilde{X} de l'expérience élémentaire définit naturellement une variable aléatoire $\tilde{S} = (\tilde{X}_1, \dots, \tilde{X}_N)$ sur l'expérience statistique. Les variables aléatoires \tilde{X}_i sont des statistiques à valeur dans \mathbb{R} , fonction de \tilde{S} . Notant e_i la matrice de taille $(N, 1)$ d'éléments $(e_i)_{n,1} = \delta_{in}$, elles s'écrivent :

$$\tilde{X}_i = e_i \cdot \tilde{S} \quad (1.24)$$

Les variables aléatoires \tilde{X}_i de l'expérience statistique sont dites indépendantes et identiquement distribuées (i.i.d.). On dit également dans ce cas que la statistique \tilde{S} est i.i.d.. Sa distribution de probabilité se déduit de celle de l'expérience élémentaire :

$$f_{\mathbb{P}}^{(\theta)}(\tilde{S}) = \prod_{i=1}^N f_{\mathbb{P}}^{(\theta)}(\tilde{X}_i) \quad \text{où} \quad f_{\mathbb{P}}^{(\theta)}(\tilde{X}_i) = f_{\mathbb{P}}^{(\theta)}(\tilde{X}) \quad \forall i \quad (1.25)$$

1.3.1.2 Estimation de paramètre

Un *estimateur* $\hat{\Theta}$ du paramètre inconnu θ de la densité élémentaire (1.23) est une variable aléatoire de l'expérience statistique, fonction de \tilde{S} , à valeur dans \mathbb{R} . C'est un cas particulier de statistique. On cherche généralement à ce qu'un estimateur soit *convergent*. En d'autres termes, on cherche à ce que sa loi de probabilité tende, dans la limite $N \rightarrow \infty$,

vers une loi de Dirac centrée en θ . La moyenne empirique $\hat{\Theta}$ d'une statistique i.i.d. de l'expérience élémentaire (1.23), définie par :

$$\hat{\Theta} = \frac{1}{N} \sum_{i=1}^N \tilde{X}_i = m \cdot \tilde{S} \quad (1.26)$$

est un estimateur convergent de θ . m est la matrice de taille $(N, 1)$ d'éléments $m_{n,1}$ égaux à $\frac{1}{N}$.

Une réalisation $\omega \in \Omega$ de l'expérience statistique fixe la valeur de toutes les variables aléatoires. On appelle $\tilde{S}(\omega)$ un *échantillon* et une réalisation $\hat{\Theta}(\omega) = m \cdot \tilde{S}(\omega)$ de l'estimateur une *estimation* de θ .

1.3.1.3 Intervalle de confiance statistique

CONSTRUCTION

Les méthodes d'estimation d'intervalles d'incertitude sur la valeur d'un paramètre définissent de nouvelles statistiques $\tilde{\Theta}_m^{(p)}$ et $\tilde{\Theta}_M^{(p)}$ de telle sorte que :

$$\mathbb{P}(\theta \in [\tilde{\Theta}_m^{(p)}; \tilde{\Theta}_M^{(p)}]) = p\% \quad (1.27)$$

Ces variables aléatoires sont des fonctions de \tilde{S} . Il existe de nombreux choix possibles de couples $(\tilde{\Theta}_m^{(p)}, \tilde{\Theta}_M^{(p)})$. Différents algorithmes (la construction de Neyman, celle de Feldman-Cousins, ...) permettent de définir un tel couple à partir d'un estimateur $\hat{\Theta}$ de θ .

Deux variables aléatoires données $\tilde{\Theta}_m^{(p)}$ et $\tilde{\Theta}_M^{(p)}$ prennent des valeurs $\tilde{\Theta}_m^{(p)}(\omega)$ et $\tilde{\Theta}_M^{(p)}(\omega)$ différentes à chaque réalisation de l'expérience statistique. L'équation (1.27) signifie que si celle-ci est répétée une infinité de fois, l'intervalle obtenu :

$$[\tilde{\Theta}_m^{(p)}(\omega); \tilde{\Theta}_M^{(p)}(\omega)] \quad (1.28)$$

contient le paramètre θ dans $p\%$ des cas.

INTERVALLE DE CONFIANCE STATISTIQUE

Considérons l'expérience :

Je réalise une expérience statistique dans le but de déterminer la valeur inconnue d'un paramètre θ de la loi de probabilité d'une expérience élémentaire. Je détermine deux statistiques $\tilde{\Theta}_m^{(90)}$ et $\tilde{\Theta}_M^{(90)}$ vérifiant l'équation (1.27) pour $p = 90$. J'en déduis un intervalle $[\tilde{\Theta}_m^{(90)}(\omega); \tilde{\Theta}_M^{(90)}(\omega)]$.

La variable aléatoire de cette expérience est une variable booléenne \tilde{Q} . Elle prend la valeur Vraie (v) si l'assertion :

$$\theta \in [\tilde{\Theta}_m^{(90)}(\omega); \tilde{\Theta}_M^{(90)}(\omega)] \quad (1.29)$$

est vraie, et la valeur Faux (f) sinon. Par construction, la loi de probabilité de cette variable aléatoire est :

$$\mathbb{P}(\tilde{Q} = v) = 0.9 \quad \mathbb{P}(\tilde{Q} = f) = 0.1 \quad (1.30)$$

Cette expérience est maintenant réalisée une unique fois. La valeur de l'assertion (1.29) est inconnue et définit une variable inconnue Q associée. Contrairement au cas du jet de dé

1.3 Confiance sur la valeur d'un paramètre inconnu

précédent, la valeur (fixée) du paramètre $\tilde{Q}(\omega)$ étudié n'est connue ni avant la réalisation de l'expérience, ni après.

Suivant les remarques de la section 1.2, un expérimentateur peut définir sa loi de confiance sur la valeur de Q par :

$$\mathbb{C}(Q) = \mathbb{P}(\tilde{Q} = Q) \quad (1.31)$$

Avant réalisation de l'expérience, un expérimentateur peut avoir confiance à 90% que la méthode statistique fournisse un intervalle contenant le paramètre θ recherché. De la même manière un expérimentateur peut avoir, après l'expérience, confiance à 90% que l'intervalle obtenu, et donc maintenant connu, contienne le paramètre. De manière générale, un intervalle (1.28) est donc appelé un intervalle de niveau de confiance $p\%$.

Cette approche ne donne pas pour autant une information directe sur le paramètre θ . Étant donnés deux réels quelconques $x, y \in \mathbb{R}$, elle ne permet pas d'associer une confiance à l'assertion $\theta \in [x; y]$. En effet, l'approche statistique est simplement un algorithme qui associe à chaque réalisation statistique $\omega \in \Omega$ une assertion (1.29), à raison dans $p\%$ des cas et à tort dans les $1-p\%$ cas restants. Elle définit une loi de probabilité sur une variable booléenne \tilde{Q} et non une mesure de confiance continue sur θ .

L'approche a pourtant l'avantage de se ramener au cas particulier où il *existe* une définition optimale de la mesure de confiance. Elle correspond donc l'approche la plus objective que l'on peut espérer.

Pour certains problèmes, cette approche peut malheureusement se révéler insuffisante. Une densité de confiance sur la valeur du paramètre peut être nécessaire. Il faut alors avoir recours à l'approche bayésienne présentée dans la section suivante.

VOCABULAIRE

D'après la construction précédente, une réalisation d'un intervalle $[\tilde{\Theta}_m^{(p)}(\omega); \tilde{\Theta}_M^{(p)}(\omega)]$ s'interprète bel et bien naturellement dans le cadre de la mesure de confiance, et non dans celui des probabilités. D'une part, cette dernière notion décrit la réalité. De l'autre, elle ne traite que du cas d'une répétition infinie d'une expérience. Or, un intervalle de niveau de confiance $p\%$ contient ou ne contient pas le paramètre θ recherché. Il n'y a pas de milieu : il ne le contient pas à $p\%$. De plus, cet intervalle est fixé. Il est obtenu pour une réalisation unique d'une expérience.

Il existe pourtant une difficulté de vocabulaire. La confiance est définie sur la variable booléenne Q et non sur l'intervalle lui-même. Le *niveau de confiance* de l'intervalle n'est pas une *confiance* au sens de la "taille" de l'intervalle vue au travers de la mesure de confiance. Il ne se rapporte même pas en fin de compte à l'intervalle lui-même. Un niveau de confiance correspond à la confiance d'une variable booléenne associée à une variable aléatoire du type de celle de l'équation (1.29). C'est donc un cas très particulier d'utilisation de la mesure de confiance.

Dans la suite, j'utiliserai l'expression *intervalle de confiance* pour désigner un intervalle de niveau de confiance (statistique) fixé. J'utiliserai les termes de *confiance d'un intervalle*, comme on peut parler de la longueur d'un intervalle ou de la probabilité d'un intervalle, pour désigner la mesure de cet intervalle à l'aide de la mesure de confiance, de la même manière que l'on peut utiliser la mesure de Lebesgue ou celle de probabilité. La confiance d'un intervalle s'écrira donc, dans le cas d'une mesure à densité $f_{\mathbb{C}}$:

$$\mathbb{C}([x; y]) = \int_x^y f_{\mathbb{C}}(z) dz \quad (1.32)$$

Au final, bien entendu, le résultat est le même : on a confiance à $p\%$ que le paramètre recherché appartienne à l'intervalle considéré⁶. Mais utiliser un vocabulaire légèrement différent permet de garder en mémoire la manière dont celui-ci et la mesure de confiance associée ont été construits. Cela peut jouer sur les possibilités de manipulation du résultat obtenu.

1.3.1.4 Composition d'intervalles de confiance statistiques

Pour être complet, et même si cela ne rentre pas directement dans le propos de ce chapitre, cette sous-section présente finalement la composition d'intervalles de confiance statistique. Cela fournit un exemple plus concret de la manipulation des mesures de confiance et de probabilité. L'ajout d'informations supplémentaires est également traitée dans la sous-section suivante.

COMPOSITION DES INTERVALLES AVANT RÉALISATION DE L'EXPÉRIENCE

Soit l'expérience statistique définie comme étant deux réalisations consécutives de l'expérience statistique précédente. Il s'agit en quelque sorte d'une expérience statistique de "niveau deux". N expériences élémentaires sont réalisées dans chaque expérience statistique de "niveau un". La variable aléatoire booléenne \tilde{Q} précédente définit la variable $(\tilde{Q}_1, \tilde{Q}_2)$ sur l'expérience statistique de niveau deux.

La loi de probabilité de la variable $(\tilde{Q}_1, \tilde{Q}_2)$ se déduit de celles des variables \tilde{Q} sur les expériences statistiques de niveau un :

$$\mathbb{P}((\tilde{Q}_1, \tilde{Q}_2)) = \mathbb{P}(\tilde{Q}_1)\mathbb{P}(\tilde{Q}_2) \quad \text{où} \quad \mathbb{P}(\tilde{Q}_i) = \mathbb{P}(\tilde{Q}) \quad (1.33)$$

Autrement dit :

$$\mathbb{P}((\tilde{Q}_1, \tilde{Q}_2) = (v, v)) = 0,81 \quad \mathbb{P}((\tilde{Q}_1, \tilde{Q}_2) = (f, f)) = 0,01 \quad (1.34)$$

$$\mathbb{P}((\tilde{Q}_1, \tilde{Q}_2) = (v, f)) = 0,09 \quad \mathbb{P}((\tilde{Q}_1, \tilde{Q}_2) = (f, v)) = 0,09 \quad (1.35)$$

L'expérience statistique de niveau deux va maintenant être réalisée une unique fois. Autrement dit, chacune des expériences statistiques de niveau un va être réalisée une fois. Le résultat correspondant à la variable aléatoire $(\tilde{Q}_1, \tilde{Q}_2)$, *i.e.* le paramètre de valeur inconnue $(\tilde{Q}_1(\omega), \tilde{Q}_2(\omega))$, définit une variable inconnue (Q_1, Q_2) de mesure de confiance naturelle :

$$\mathbb{C}((Q_1, Q_2) = (v, v)) = 0,81 \quad \mathbb{C}((Q_1, Q_2) = (f, f)) = 0,01 \quad (1.36)$$

$$\mathbb{C}((Q_1, Q_2) = (v, f)) = 0,09 \quad \mathbb{C}((Q_1, Q_2) = (f, v)) = 0,09 \quad (1.37)$$

En terme d'intervalles, le premier cas $(Q_1, Q_2) = (v, v)$ signifie par exemple que :

$$(\theta \in I_1) \wedge (\theta \in I_2) \Leftrightarrow (\theta \in I_1 \cap I_2) \quad (1.38)$$

Le symbole \wedge désigne le ET logique. Les autres cas se déduisent de manière similaire. Un expérimentateur peut donc avoir 81% de confiance que la prochaine réalisation de l'expérience statistique de niveau deux lui fournisse deux intervalles contenant le paramètre θ .

6. Encore une fois, ceci n'est vrai que dans les problèmes d'estimation de paramètres. Le cas des tests d'hypothèses est plus complexe. Les liens entre mesure de confiance et mesure de probabilité y sont plus délicats et l'expression de "niveau de confiance" (CL) ne correspond pas alors nécessairement à une utilisation de la mesure telle que définie ici.

1.3 Confiance sur la valeur d'un paramètre inconnu

COMPOSITION DES INTERVALLES APRÈS RÉALISATION DE L'EXPÉRIENCE

Après réalisation de l'expérience, l'expérimentateur possède une nouvelle information. En effet, les valeurs des intervalles de confiances sont maintenant utilisables. Lorsque les intervalles de confiance (fixés et connus) associés à Q_1 et Q_2 sont disjoints, les variables inconnues ne peuvent pas être simultanément vraies. En vocabulaire bayésien, la mesure de confiance utilisée est donc actualisée :

$$\mathbb{C}((Q_1, Q_2) | Q_1 \wedge Q_2 = f) = \frac{\mathbb{C}(Q_1 \wedge Q_2 = f | (Q_1, Q_2))}{\mathbb{C}(Q_1 \wedge Q_2 = f)} \mathbb{C}((Q_1, Q_2)) \quad (1.39)$$

Soit :

$$\mathbb{C}((v, v) | Q_1 \wedge Q_2 = f) = 0 \quad \mathbb{C}((f, f) | Q_1 \wedge Q_2 = f) = \frac{1}{19} \quad (1.40)$$

$$\mathbb{C}((v, f) | Q_1 \wedge Q_2 = f) = \frac{9}{19} \quad \mathbb{C}((f, v) | Q_1 \wedge Q_2 = f) = \frac{9}{19} \quad (1.41)$$

En terme d'intervalles, la confiance que l'on peut avoir pour que le paramètre θ appartienne à l'union des deux intervalles de confiance disjoints est de 18/19 et celle pour laquelle le paramètre n'appartienne à aucun des deux est de 1/19.

La nouvelle loi de confiance provient en fait d'une redéfinition de l'expérience statistique envisagée à l'aide des informations obtenues lors de la réalisation particulière considérée (cf section 1.2.3). Dans le cas du jet de dé, le résultat complet (la valeur donnée par le dé) était connue à l'issue de l'expérience. La nouvelle loi de probabilité était construite sur l'expérience de jet de dé qui donne toujours le même résultat. La confiance devenait donc une mesure de Dirac. Ici, le résultat complet de l'expérience, à savoir si θ appartient à l'un des intervalles de confiance, n'est pas disponible. Le fait que les intervalles soient disjoints restreint tout de même le nombre des réalisations que l'on peut considérer comme similaires à celle obtenue. Le jeu de confiances (1.40) provient donc de la loi de probabilité de la nouvelle expérience :

Deux expériences statistiques sont réalisées. Au plus l'une d'entre elles fournit un intervalle de confiance contenant le paramètre θ recherché.

En fait, l'expérience pourrait être définie de manière encore plus précise :

Deux expériences statistiques sont réalisées. Les intervalles de confiance fournis sont disjoints.

Le jeu de confiance sur la variable (Q_1, Q_2) en serait encore modifié mais il faudrait pour cela connaître la loi de probabilité associée. De la même manière, la loi de probabilité de l'expérience dans laquelle les deux intervalles obtenus ne sont pas disjoints n'a aucune raison d'être la même que celle de l'équation (1.34).

La composition d'un plus grand nombre d'intervalles de confiance statistique se déduit directement.

1.3.1.5 Intervalle de confiance et information supplémentaire

Une information d'origine non statistique peut être connue indépendamment des expériences considérées. Le cas le plus souvent cité est celui de la mesure d'une masse. En effet, un unique intervalle de confiance peut déborder sur les valeurs négatives de l'axe réel quand bien même la valeur de la masse est toujours positive.

Ce cas se traite de la même manière que dans le paragraphe précédent. Si l'on sait que par construction de la variable aléatoire \tilde{Q} l'intervalle de confiance construit contiendra dans $p\%$ des cas la valeur du paramètre recherché, il est généralement impossible de dire pour une réalisation particulière si l'intervalle obtenu contient réellement l'objet recherché ou non. Dans le cas où l'intervalle de confiance se situe dans une zone interdite, il est "par chance" possible de trancher : l'intervalle ne contient pas le paramètre.

Plus formellement, la confiance obtenue est modifiée en ajoutant une nouvelle variable inconnue booléenne H correspondant au domaine physique :

$$H : \theta \in \mathcal{D}_{phys} \tag{1.42}$$

De cette manière, la mesure (1.31) devient :

$$\mathbb{C}(Q|H = v) = \frac{\mathbb{C}(Q, H = v)}{\mathbb{C}(H = v)} = \mathcal{N} \cdot \mathbb{C}(Q, H = v) \tag{1.43}$$

\mathcal{N} est le facteur de normalisation. Si l'intervalle de confiance et le domaine physique sont de plus disjoints, l'information supplémentaire $Q \wedge H = f$ peut être utilisée :

$$\mathbb{C}(Q = v|H = v, Q \wedge H = f) = \mathcal{N} \cdot \mathbb{C}(Q = v, H = v|Q \wedge H = f) = 0 \tag{1.44}$$

$$\mathbb{C}(Q = f|H = v, Q \wedge H = f) = \mathcal{N} \cdot \mathbb{C}(Q = f, H = v|Q \wedge H = f) = 1 \tag{1.45}$$

Si les deux ensembles ne sont pas disjoints, l'équation (1.43) interprétée comme en (1.34), montre que l'intervalle de confiance est restreint à son intersection avec le domaine physique.

Bien entendu, la "chance" d'avoir à disposition une nouvelle information lorsque l'intervalle de confiance est en dehors du domaine physique est toute relative. Si la réalisation de l'expérience statistique a demandé beaucoup d'efforts, de temps et d'argent, il n'est pas agréable d'apprendre que l'on se situe précisément dans les $1 - p\%$ des cas où l'intervalle obtenu ne contient pas le paramètre recherché. D'autant plus que l'intervalle que l'on exclue était déjà inenvisageable puisque non physique. Lorsque l'intervalle est dans la zone physique, le niveau de confiance de l'intervalle obtenu peut être par contre plus élevé que les 90% initialement visés.

1.3.2 Mesure continue de confiance : approche "bayésienne"

1.3.2.1 Modèle de confiance

On modifie l'expérience considérée dans la construction du modèle de confiance de la section 1.3.1.3. On s'intéresse simplement à une réalisation unique $\omega_0 \in \Omega$ de l'expérience statistique. On ne cherche pas à définir un intervalle de confiance. Les paramètres de valeurs *a priori* inconnues sont les résultats :

$$s = (x_1, \dots, x_N) = \tilde{S}(\omega_0) \tag{1.46}$$

des différentes expériences élémentaires et le paramètre θ désignant la mesure de probabilité au sein de la famille des lois normales. On leur associe des variables inconnues $S = (X_1, \dots, X_N)$ et Θ puis on construit une mesure de confiance $f_{\mathbb{C}}(S, \Theta)$ sur l'espace correspondant.

1.3 Confiance sur la valeur d'un paramètre inconnu

Supposons un instant que la valeur θ du paramètre soit connue. Le problème est alors similaire à celui de la section 1.2. On construit la densité de confiance sur la variable inconnue S à l'aide de la densité de probabilité sur la variable aléatoire \tilde{S} :

$$f_{\mathbb{C}}(S|\Theta = \theta) = f_{\mathbb{P}}^{(\theta)}(\tilde{S} = S) = \prod_{i=1}^N \left[\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \theta)^2}{\sigma^2}} \right] \quad (1.47)$$

Plus généralement, on définit alors la densité conditionnelle par :

$$f_{\mathbb{C}}(S|\Theta) = \prod_{i=1}^N \left[\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - \Theta)^2}{\sigma^2}} \right] \quad (1.48)$$

De cette manière, il est certain que l'équation (1.47) est vérifiée, et ce quelle que puisse être la valeur réelle θ du paramètre. À une normalisation près, cette densité conditionnelle est habituellement trouvée sous le nom de *fonction de vraisemblance*.

La mesure de confiance sur l'espace total s'écrit :

$$f_{\mathbb{C}}(S, \Theta) = f_{\mathbb{C}}(S|\Theta)f_{\mathbb{C}}(\Theta) \quad (1.49)$$

Puisque la densité conditionnelle est déjà fixée, la définition de la distribution $f_{\mathbb{C}}(\Theta)$ complète la mesure de confiance. Cette densité s'appelle habituellement un *prior*. Malheureusement, aucune information totalement objective ne permet en général de fixer sa forme. On cherche alors une densité la plus neutre possible. Les deux sous-sections suivantes discutent succinctement ce problème.

Une fois le prior choisi, la mesure de confiance fournit une densité conditionnelle de confiance (aussi appelée *posterior*) sur la variable Θ , étant connu le résultat de l'expérience statistique :

$$f_{\mathbb{C}}(\Theta|S) = \frac{f_{\mathbb{C}}(S, \Theta)}{f_{\mathbb{C}}(S)} \quad (1.50)$$

La confiance (*a posteriori*) d'un intervalle quelconque $[x; y]$, $x, y \in \mathbb{R}$ sur la variable inconnue Θ est définie par :

$$\mathbb{C}(\Theta \in [x; y]|S) = \int_x^y f_{\mathbb{C}}(\Theta|S) d\Theta \quad (1.51)$$

1.3.2.2 Normalisation du prior. Information insuffisante

Avant la réalisation de l'expérience statistique, la valeur du paramètre θ est totalement inconnue. Le choix le plus naturel de prior ne privilégie aucune valeur de Θ :

$$\frac{f_{\mathbb{C}}(\Theta)}{f_{\mathbb{C}}(\Theta')} = 1 \quad \forall \Theta, \Theta' \in \mathbb{R} \quad (1.52)$$

Cette définition soulève deux difficultés. Tout d'abord, une fonction constante sur \mathbb{R} n'est d'intégrale finie que si elle est uniformément nulle. Son intégrale est alors elle-même égale à zéro. La densité du prior n'est donc pas normalisée et la mesure de confiance n'existe pas. En d'autres termes, l'information utilisée est insuffisante pour ne serait-ce que définir une mesure de confiance sur la valeur de Θ , *i.e.* fournir une estimation de la valeur de θ .

En pratique, il est possible de définir le prior comme la limite d'une mesure de confiance $f_{\mathbb{C},\epsilon}$ lorsqu'un paramètre ϵ tend vers zéro. On pose par exemple :

$$f_{\mathbb{C},\epsilon}(\Theta) = \frac{\epsilon}{2} \chi_{-\frac{1}{\epsilon} \leq \Theta \leq \frac{1}{\epsilon}} \quad (1.53)$$

où χ_A est la fonction caractéristique de l'ensemble A :

$$\chi_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases} \quad (1.54)$$

D'autres choix de densités $f_{\mathbb{C},\epsilon}$ que celui de l'expression (1.53) sont équivalents dès que l'équation (1.52) est vérifiée dans la limite $\epsilon \rightarrow 0$. Ce prior définit une mesure $f_{\mathbb{C},\epsilon}(S, \Theta)$ en tout ϵ non nul. En particulier, la densité $f_{\mathbb{C},\epsilon}(\Theta|S = s)$ est définie pour toute valeur de ϵ différente de zéro. À la limite, le prior et la mesure $f_{\mathbb{C}}(S, \Theta)$ s'annulent. Ce n'est pas nécessairement le cas de la densité conditionnelle $f_{\mathbb{C}}(\Theta|S = s)$. L'information est insuffisante pour fournir une estimation des valeurs du paramètre θ ou du couple (s, θ) lorsqu'aucune valeur de paramètre n'est connue. Elle peut par contre permettre de créer une estimation de la valeur du paramètre θ lorsque le résultat de l'expérience est connu.

1.3.2.3 Variable privilégiée

La densité $f_{\mathbb{C}}(\Theta)$ définie comme la limite de l'expression (1.53) s'annule uniformément. Elle ne contient pas suffisamment d'information pour estimer la valeur de θ . Elle n'en est pas pour autant dépourvue. En effet, le choix de prior (1.52) privilégie la valeur Θ du paramètre au détriment de toute autre fonction de Θ .

Considérons le nouveau prior défini en ϵ non nul par :

$$f_{\mathbb{C},\epsilon}(\Theta^2) = \epsilon \chi_{0 \leq \Theta^2 \leq \frac{1}{\epsilon}} \quad (1.55)$$

Les valeurs de Θ^2 sont considérées équivalentes. La densité sur Θ se déduit par changement de variable :

$$f_{\mathbb{C},\epsilon}(\Theta) = 2|\Theta| \epsilon \chi_{0 \leq \Theta^2 \leq \frac{1}{\epsilon}} \quad (1.56)$$

Les deux expressions possibles (1.56) et (1.55) du prior s'annulent dans la limite où ϵ tend vers zéro. L'information est insuffisante pour estimer le paramètre θ ou son carré. Par contre, quelle que soit la valeur de $\epsilon \leq \frac{1}{4}$:

$$\frac{f_{\mathbb{C},\epsilon}(\Theta = 2)}{f_{\mathbb{C},\epsilon}(\Theta = 1)} = 2 \quad (1.57)$$

La densité (1.56) ne traite pas de la même manière les valeurs $\Theta = 1$ et $\Theta = 2$ alors qu'elle traite symétriquement les valeurs $\Theta^2 = 1$ et $\Theta^2 = 4$: les choix de prior (1.52) et (1.56) ne sont pas équivalents. Ils fournissent donc tous deux une information mais celle-ci est différente dans les deux cas. Elle correspond à un choix de variable privilégiée : Θ ou Θ^2 . Les conséquences des priors sur le modèle en général et sur la densité conditionnelle $f_{\mathbb{C}}(\Theta|S)$ en particulier seront différentes.

De manière générale, il serait agréable d'utiliser un prior qui ne contienne pas d'information. Puisque celui-ci ne peut pas être défini depuis une distribution de probabilité, l'information qu'il porte s'apparente à une hypothèse. Le choix de cette densité est donc toujours ouvert à controverse. À défaut, on peut privilégier une variable inconnue fonction de Θ et lui associer une densité "uniformément nulle" telle que (1.52) ou (1.55). Ce choix pourra être guidé par des informations peu tangibles mais pour autant importantes telles que la nature (physique par exemple) du paramètre recherché.

1.3.3 Choix de méthode

1.3.3.1 Test répété

La prudence semblerait privilégier la méthode statistique d'estimation d'un paramètre inconnu. Moins d'information utilisée pour définir la mesure de confiance signifie qu'il y a moins de sources possibles d'erreur. Bien entendu, tant qu'aucune fausse certitude (annulation de la densité de confiance) n'aurait été utilisée dans la création du modèle, la prise en compte de suffisamment de résultats de l'expérience élémentaire, données objectives, corrigerait toute information erronée provenant d'un prior subjectif qui s'avèrerait trop éloigné de la réalité. Il semble toutefois préférable d'éviter d'avoir recours à des informations d'origine incertaine autant que possible. Tant que l'approche statistique est suffisante, il serait plus prudent de la privilégier. L'exemple imaginaire suivant pourrait entrer dans cette catégorie.

Une chaîne d'usinage traite des milliers de pièces par jour. Le centre de qualité est tenu de contrôler la conformité des pièces. Une cote particulière de la pièce doit se situer par exemple entre certaines limites inférieures et supérieures.

Supposons que la loi de probabilité de la valeur de cette cote soit une loi normale, de variance connue mais de moyenne inconnue. La moyenne pourrait dépendre de l'état d'usure de l'outillage de production. Le centre de qualité doit s'assurer que le nombre de rebus est inférieur à un certain seuil, cinq parts pour mille (5ppm) par exemple. Il vérifie pour cela la position de la moyenne de la loi.

Un échantillon statistique de N pièces est régulièrement extrait de la chaîne de production et testé. La méthode statistique fournit à chaque fois un encadrement de la moyenne à un niveau de confiance de 68%. La production est déclarée en conformité si toutes les valeurs de moyennes de cet intervalle prévoient moins de cinq parts pour mille de rebus.

Ce résultat peut être éventuellement considéré comme suffisant car les contrôles sont réguliers. En effet, il est important que la dérive de la valeur de la moyenne soit détectée mais il n'est peut être pas critique qu'elle le soit au moment précis où elle se produit. Or, il est "certain" que dans deux cas sur trois le centre de qualité a réellement vérifié que le nombre de rebus était inférieur aux 5ppm. Au pire des cas, la fréquence des contrôles peut être augmentée. L'ajout d'un prior n'est pas absolument nécessaire dans ce cas. On évite d'ajouter une information qui pourrait se révéler trop éloignée de la réalité.

1.3.3.2 Incertitude statistique et biais

En pratique, les résultats de l'approche statistique souffrent néanmoins d'une certaine rigidité. Cela limite sa portée. La moyenne de la loi normale peut par exemple ne pas être directement égale au paramètre cherché θ . En d'autres termes, l'expérience élémentaire peut être biaisée, d'un biais b de valeur fixe mais inconnue :

$$f_{\mathbb{P}}^{(\theta,b)}(\tilde{X}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\tilde{X}-(\theta+b))^2}{\sigma^2}} \quad (1.58)$$

La différence entre les approches statistique (*i.e.* *fréquentiste*) et de modèle de confiance (\simeq approche *bayésienne*) est claire dans ce cas. Les résultats (1.31) de la première sont difficilement généralisable à ce nouveau problème alors que la mesure de confiance s'étend sans aucune difficulté.

APPROCHE STATISTIQUE

Une expérience statistique est réalisée une unique fois. La réalisation $\omega \in \Omega$ est fixée. L'intervalle de confiance statistique (1.28) contient, à 90% de chance, la somme $\theta + b$:

$$\mathbb{C}(Q = v) = 90\% \quad \mathbb{C}(Q = f) = 10\% \quad (1.59)$$

Q est la variable inconnue booléenne associée à la valeur de l'assertion :

$$(\theta + b) \in [\tilde{\Theta}_m^{(90)}(\omega); \tilde{\Theta}_M^{(90)}(\omega)] \quad (1.60)$$

La valeur du biais b est inconnue. On définit une variable inconnue B associée. L'équation (1.59) se réécrit sous la forme :

$$\mathbb{C}(Q = v|B = b) = 90\% \quad \mathbb{C}(Q = f|B = b) = 10\% \quad (1.61)$$

Suivant le même raisonnement que dans la section 1.3.2.1, on pose de manière générale :

$$\mathbb{C}(Q = v|B) = 90\% \quad \mathbb{C}(Q = f|B) = 10\% \quad (1.62)$$

où maintenant :

$$Q \equiv [(\theta + B) \in [\tilde{\Theta}_m^{(90)}(\omega); \tilde{\Theta}_M^{(90)}(\omega)]] \quad (1.63)$$

Un prior $f_{\mathbb{C}}(B)$ achève enfin la définition de la mesure sur le couple (Q, B) . Il correspond à l'estimation du biais.

Cette mesure ne donne malheureusement pas d'information directe sur la valeur de θ . Elle ne fournit pas d'intervalle de confiance ou de densité de confiance sur les valeurs possibles du paramètre. En effet, supposons pour simplifier que le prior sur le biais soit simplement la somme pondérée de deux distributions de Dirac. Autrement dit, le biais ne peut prendre que deux valeurs : b_1 (confiance q) et b_2 . La mesure complète sur (Q, B) est donc :

$$\begin{aligned} \mathbb{C}((v, b_1)) &= 0,9q & \mathbb{C}((v, b_2)) &= 0,9(1-q) \\ \mathbb{C}((f, b_1)) &= 0,1q & \mathbb{C}((f, b_2)) &= 0,1(1-q) \end{aligned} \quad (1.64)$$

Le premier cas correspond par exemple à :

$$(\theta + b_1) \in [\tilde{\Theta}_m^{(90)}(\omega); \tilde{\Theta}_M^{(90)}(\omega)] \Leftrightarrow \theta \in [\tilde{\Theta}_m^{(90)}(\omega) - b_1; \tilde{\Theta}_M^{(90)}(\omega) - b_1] \quad (1.65)$$

et le second à :

$$(\theta + b_2) \in [\tilde{\Theta}_m^{(90)}(\omega); \tilde{\Theta}_M^{(90)}(\omega)] \Leftrightarrow \theta \in [\tilde{\Theta}_m^{(90)}(\omega) - b_2; \tilde{\Theta}_M^{(90)}(\omega) - b_2] \quad (1.66)$$

Supposons que les deux intervalles (1.65) et (1.66) ci-dessus possèdent une intersection non vide. On souhaiterait naturellement composer les confiances de la même manière que dans le cas de la section 1.3.1.4. Pourtant, les différents cas comme (v, b_1) et (v, b_2) ne sont pas ici compatibles : ils ne peuvent pas être réalisés simultanément. On ne peut pas les supposer indépendants. Ils ne correspondent même pas à des variables inconnues différentes. Le seul résultat valable est :

$$\mathbb{C}(\{(v, b_1), (v, b_2)\}) = 0,9q + 0,9(1-q) = 0,9 \quad (1.67)$$

Ce qui n'avance à rien.

Pour composer les "confiances" des intervalles (1.65) et (1.66) sur θ , il faudrait en fait fournir une répartition continue de cette confiance au sein même des segments. En d'autre termes, il faudrait fournir un prior sur θ . On retrouve l'approche "bayésienne". L'approche statistique de la section 1.3.1 ne permet donc pas de composer simplement l'incertitude statistique et un biais.

1.4 Applications : manipulations des incertitudes théoriques

MODÈLE DE CONFIANCE

L'approche de la section 1.3.2 définit une distribution de confiance sur le couple de variables inconnues (Θ, S) . La variable inconnue B représentant le biais b étend l'espace de définition de la mesure de confiance. Les équations (1.48) et (1.49), se réécrivent :

$$f_{\mathbb{C}}(S|\Theta, B) = \prod_{i=1}^N \left[\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i - (\Theta+B))^2}{\sigma^2}} \right] \quad (1.68)$$

$$f_{\mathbb{C}}(S, \Theta|B) = f_{\mathbb{C}}(S|\Theta, B)f_{\mathbb{C}}(\Theta|B) \quad (1.69)$$

Lorsque la connaissance de la valeur du biais b n'apporte pas d'indication sur la valeur du paramètre θ recherché, les variables inconnues Θ et B sont définies comme indépendantes. La densité conditionnelle sur Θ connaissant B est alors égale au prior sur Θ :

$$f_{\mathbb{C}}(\Theta|B) = f_{\mathbb{C}}(\Theta) \quad (1.70)$$

Un second prior modélisant l'incertitude sur le biais $f_{\mathbb{C}}(B)$ est nécessaire. La mesure de confiance sur l'ensemble des triplets (Θ, S, B) s'écrit finalement :

$$f_{\mathbb{C}}(S, \Theta, B) = f_{\mathbb{C}}(S|\Theta, B)f_{\mathbb{C}}(\Theta)f_{\mathbb{C}}(B) \quad (1.71)$$

La densité de confiance sur la valeur du paramètre θ , une fois l'expérience statistique réalisée, combine cette fois-ci les deux sources d'incertitudes :

$$f_{\mathbb{C}}(\Theta|S) = \int \frac{f_{\mathbb{C}}(S, \Theta, B)}{f_{\mathbb{C}}(S)} dB \quad (1.72)$$

1.4 Applications : manipulations des incertitudes théoriques

1.4.1 Incertitudes expérimentales et théoriques

Les méthodes statistiques sont également limitées lorsqu'il s'agit de combiner des informations de différentes origines. C'est le cas par exemple de la prise en compte simultanée des prédictions théoriques et expérimentales d'une même section efficace σ . Cette opération s'effectue naturellement dans un cadre purement bayésien.

Les prédictions théoriques ne sont pas toujours exactes. Elles peuvent être effectuées sous certaines approximations ou n'être que partielles. En physique des particules, de nombreux résultats se présentent sous forme de séries perturbatives. Seuls les premiers termes sont généralement connus car la complexité de leur calcul augmente rapidement avec l'ordre perturbatif. Le chapitre suivant étudie le problème de l'estimation des incertitudes théoriques dues à cette troncature. Le modèle proposé fournit une distribution de confiance $f_{\mathbb{C}}(\sigma|c_0, \dots, c_k)$ sur la valeur d'une section efficace σ étants connus les premiers termes (c_0, \dots, c_k) de la série. Ce paragraphe présente la combinaison de cette densité de confiance avec l'incertitude d'origine expérimentale.

VARIABLES INCONNUES

Les paramètres de valeur a priori inconnus sont la section efficace σ , sa mesure expérimentale σ_e et les coefficients c_n de la série perturbative définissant son expression théorique. Pour simplifier les notations, la distinction entre les paramètres et leur variables inconnues associée n'est plus effectuée dorénavant.

HYPOTHÈSE D'INDÉPENDANCE

Les informations sur les premiers coefficients c_0, \dots, c_k sont d'origine théorique alors que celles sur σ_e proviennent de l'expérience. La connaissance des termes de la série n'influence donc pas *directement* la confiance sur le résultat expérimental. Si la valeur de la section efficace était précisément connue, le calcul des coefficients de la série n'apporterait aucune information supplémentaire sur la valeur de σ_e :

$$f_{\mathbb{C}}(\sigma_e | \sigma, c_0, \dots, c_k) = f_{\mathbb{C}}(\sigma_e | \sigma) \quad (1.73)$$

Les variables inconnues σ_e et (c_0, \dots, c_k) sont donc indépendantes dans les sous-espaces à σ connu.

MESURE DE CONFIANCE ET INCERTITUDE SUR LA SECTION EFFICACE

La mesure de confiance complète s'écrit donc :

$$\begin{aligned} f_{\mathbb{C}}(\sigma, \sigma_e, c_0, \dots, c_k) &= f_{\mathbb{C}}(\sigma_e, c_0, \dots, c_k | \sigma) f_{\mathbb{C}}(\sigma) \\ &= f_{\mathbb{C}}(\sigma_e | \sigma) f_{\mathbb{C}}(c_0, \dots, c_k | \sigma) f_{\mathbb{C}}(\sigma) \\ &= f_{\mathbb{C}}(\sigma_e | \sigma) f_{\mathbb{C}}(\sigma | c_0, \dots, c_k) f_{\mathbb{C}}(c_0, \dots, c_k) \end{aligned} \quad (1.74)$$

La distribution de confiance sur la valeur de la section efficace, étant connus les premiers coefficients de son expression théorique et une mesure σ_e expérimentale, s'écrit :

$$f_{\mathbb{C}}(\sigma | \sigma_e, c_0, \dots, c_k) = \mathcal{N} \cdot f_{\mathbb{C}}(\sigma_e | \sigma) f_{\mathbb{C}}(\sigma | c_0, \dots, c_k) \quad (1.75)$$

L'élément \mathcal{N} défini par :

$$\mathcal{N} = \frac{f_{\mathbb{C}}(c_0, \dots, c_k)}{f_{\mathbb{C}}(\sigma_e, c_0, \dots, c_k)} \quad (1.76)$$

est un terme de normalisation. La densité $f_{\mathbb{C}}(\sigma_e | \sigma)$ étant par ailleurs connue, une mesure de confiance sur la valeur de la section efficace d'origine théorique $f_{\mathbb{C}}(\sigma | c_0, \dots, c_k)$ comme celle proposée au chapitre 2 complète l'estimation de l'incertitude totale sur la somme de la série σ .

1.4.2 Extraction de paramètre

Cette partie de ma thèse provient à l'origine de l'étude du problème d'extraction de la masse du quark top en production directe. C'est en recherchant une manière rigoureuse de combiner les incertitudes expérimentales et théoriques que toutes les réflexions de ce chapitre sont nées. La recherche d'une définition cohérente des incertitudes théoriques a donné à son tour le chapitre suivant. Pour conclure cette partie, je souhaite donc revenir sur ce sujet initial.

1.4.2.1 Construction du modèle de confiance

Les données expérimentales \mathcal{D} ne fournissent pas directement une estimation σ_e d'une section efficace σ . Elles doivent être traitées et interprétées. Cette étape nécessite entre autre la valeur réelle de certains paramètres physiques.

Ces paramètres ne sont en général pas totalement connus. La section efficace σ_e que l'on pensera avoir observé dépend du choix que l'on effectue sur leurs valeurs. En renversant

1.4 Applications : manipulations des incertitudes théoriques

le problème, il est donc possible d'utiliser la mesure de la section efficace pour estimer la valeur de ces paramètres. Le jeu de valeurs prédisant une section σ_e au plus proche de la valeur prédite par la théorie sera sélectionné.

On ne considère dans la suite qu'un seul paramètre pour simplifier : la masse m du quark top. La valeur extraite σ_e est donc une fonction des données expérimentales \mathcal{D} et de m .

$$\sigma_e = g(\mathcal{D}, m) \quad (1.77)$$

De la même manière, la prédiction théorique de la section efficace utilise des expressions \mathcal{T} analytiques. Une valeur m de la masse du quark doit être fixée pour calculer une valeur numérique des coefficients c_0, \dots, c_k de la série perturbative de la section efficace réelle.

$$(c_0, \dots, c_k) = h(\mathcal{T}, m) \quad (1.78)$$

VARIABLES INCONNUES

Ainsi, plus généralement qu'à la section précédente, les variables inconnues sont la section efficace σ , le jeu de données expérimentales \mathcal{D} , les expressions théoriques \mathcal{T} des coefficients et la masse m .

HYPOTHÈSE D'INDÉPENDANCE

On définit les variables inconnues \mathcal{T} et \mathcal{D} comme étant indépendantes lorsque σ et m sont connues :

$$f(\mathcal{D}, \mathcal{T} | \sigma, m) = f(\mathcal{D} | \sigma, m) f(\mathcal{T} | \sigma, m) \quad (1.79)$$

En effet, ces deux variables ne sont reliées qu'au travers des valeurs de la section efficace et de la masse. Une fois celles-ci connues, la connaissance de \mathcal{T} n'apporte par exemple aucune information supplémentaire sur les valeurs possibles de \mathcal{D} :

$$f(\mathcal{D} | \sigma, m, \mathcal{T}) = f(\mathcal{D} | \sigma, m) \quad (1.80)$$

Les incertitudes résiduelles, une fois σ et m connus, sont donc indépendantes.

EXPRESSION DE LA MESURE DE CONFIANCE

La mesure de confiance complète s'écrit dans ce cas :

$$f(\sigma, m, \mathcal{D}, \mathcal{T}) = f(\mathcal{T}, \mathcal{D} | \sigma, m) f(\sigma, m) \quad (1.81)$$

$$= f(\mathcal{T} | \sigma, m) f(\mathcal{D} | \sigma, m) f(\sigma, m) \quad (1.82)$$

$$= f(\sigma | \mathcal{T}, m) f(\sigma | \mathcal{D}, m) f(m) \left[\frac{f(\mathcal{T} | m) f(\mathcal{D} | m)}{f(\sigma | m)} \right] \quad (1.83)$$

La dernière ligne (1.83) est obtenue par application de la formule de Bayes sur les deux premières densités puis en utilisant celle des densités conditionnelles sur $f(\sigma, m)$.

1.4.2.2 Expressions des densités

DÉPENDANCE EN m

On fait l'hypothèse que la valeur de la masse m n'apporte d'information à la section σ ou aux autres données qu'au travers de \mathcal{T} ou \mathcal{D} . Autrement dit, on suppose que les densités $f(\sigma | m)$, $f(\mathcal{T} | m)$, $f(\mathcal{D} | m)$ sont indépendantes de m .

DENSITÉ $f(\sigma|\mathcal{T}, m)$

La densité $f(\sigma|\mathcal{T}, m)$ n'est rien d'autre que la densité de confiance $f(\sigma|c_0, \dots, c_k)$. En effet, celle-ci peut se réécrire :

$$f(\sigma|\mathcal{T}, m) = \int f(\sigma, c_0, \dots, c_k|\mathcal{T}, m) \prod dc_i \quad (1.84)$$

$$= \int f(\sigma|c_0, \dots, c_k|\mathcal{T}, m) f(c_0, \dots, c_k|\mathcal{T}, m) \prod dc_i \quad (1.85)$$

$$= \int f(\sigma|c_0, \dots, c_k) f(c_0, \dots, c_k|\mathcal{T}, m) \prod dc_i \quad (1.86)$$

$$= \int f(\sigma|c_0, \dots, c_k) \delta[(c_0, \dots, c_k) - h(\mathcal{T}, m)] \prod dc_i \quad (1.87)$$

$$= f(\sigma|(c_0, \dots, c_k) = h(\mathcal{T}, m)) \quad (1.88)$$

Le passage à la ligne (1.86) utilise deux points. La variable \mathcal{T} n'apporte de l'information sur la valeur de σ qu'au travers des coefficients. Suivant toujours le même raisonnement, une fois la valeur des coefficients connue, la donnée de \mathcal{T} ne modifie plus l'incertitude :

$$f(\sigma|c_0, \dots, c_k, \mathcal{T}, m) = f(\sigma|c_0, \dots, c_k, m) \quad (1.89)$$

D'autre part, on suppose comme au paragraphe précédent que la valeur de la masse m n'apporte pas d'information sur la section σ sans être combinée avec \mathcal{T} ou \mathcal{D} :

$$f(\sigma|c_0, \dots, c_k, m) = f(\sigma|c_0, \dots, c_k) \quad (1.90)$$

DENSITÉ $f(\sigma|\mathcal{D}, m)$

La densité $f(\sigma|\mathcal{D}, m)$ correspond de la même manière à la densité $f(\sigma|\sigma_e)$ estimant l'incertitude sur la section efficace étant données les données expérimentales seules.

CHOIX DE PRIOR

Le prior $f(m)$ peut être choisi uniformément nul en $\ln m$. En effet, la masse est supposée pouvoir être de tout ordre de grandeur. Le choix de $1 \pm 0,1$ GeV est *a priori* équivalent à celui de $1 \pm 0,1$ MeV. La densité doit donc être uniforme dans les ordres de grandeur, *i.e.* dans le logarithme de la masse.

1.4.2.3 Discussion du modèle

Si ce modèle peut sembler bien défini, ce n'est qu'en première apparence. Tout d'abord, il est difficile d'imaginer définir une réelle mesure sur l'espace des expressions théoriques \mathcal{T} des coefficients de la série perturbative. Cette notation semble donc plus symbolique que rigoureuse.

D'autre part, les densités $f(\mathcal{T}|m)$, $f(\mathcal{D}|m)$ et $f(\sigma|m)$ sont reliées :

$$f(\sigma|m) = \int f(\sigma|\mathcal{T}, m) f(\mathcal{T}|m) d\mathcal{T} \quad (1.91)$$

$$f(\sigma|m) = \int f(\sigma|\mathcal{D}, m) f(\mathcal{D}|m) d\mathcal{D} \quad (1.92)$$

Les densités $f(\sigma|\mathcal{T}, m)$ et $f(\sigma|\mathcal{D}, m)$ étant définies par ailleurs, il n'est pas du tout évident que l'on puisse "décider" que ces trois distributions sont indépendantes de la masse m .

1.4 Applications : manipulations des incertitudes théoriques

Cette mesure n'est donc pas proprement définie. Ses inconsistances semblent toutefois correspondre à des informations fines qui ne sont pas prises en compte. En première approximation, les résultats obtenus avec ce modèle devraient tout de même être exploitables.

1.4.2.4 Extraction de paramètre

La confiance sur la valeur de la masse, étant données les observations et les calculs théoriques, s'obtient finalement par la formule naturelle :

$$f(m|\mathcal{D}, \mathcal{T}) = \int f(\sigma, m|\mathcal{D}, \mathcal{T}) d\sigma \quad (1.93)$$

$$= \frac{1}{f(\mathcal{D}, \mathcal{T})} \int f(\sigma, m, \mathcal{D}, \mathcal{T}) d\sigma \quad (1.94)$$

$$= \mathcal{N} \left[\int f(\sigma|\mathcal{D}, m) f(\sigma|\mathcal{T}, m) d\sigma \right] f(m) \quad (1.95)$$

\mathcal{N} est un paramètre de normalisation. Le chemin emprunté semble très complexe pour arriver à une expression tellement simple. Il permet toutefois de vérifier les hypothèses utilisées et leur articulation. D'autant plus que le modèle ne semble pas réellement consistant.

1.5 Conclusion

Les mesures de confiance et de probabilité décrivent deux choses très différentes : l'incertitude d'un opérateur et les conséquences des informations dont il dispose d'une part et la non-reproductibilité d'une expérience de l'autre.

Dans le cas particulier où l'on s'intéresse à l'issue d'une réalisation particulière d'une expérience, les deux mesures existent simultanément. De plus, on peut définir une "définition optimale" du modèle de confiance. Celle-ci correspond à fixer la mesure de confiance de manière à ce qu'elle soit numériquement identique à celle de probabilité. Il n'y a alors pas nécessairement de raison de distinguer les deux mesures. Il est possible de confondre la probabilité d'un résultat avec la confiance que l'on puisse avoir qu'il soit réalisé.

Lorsque l'on cherche à tirer une information de la réalisation successive d'une expérience, *i.e.* lorsque l'on s'intéresse à un problème d'inférence, les deux notions de confiance et de probabilité vont encore une fois naturellement entrer en jeu. La première aura pour but d'estimer les conséquences des informations tirées des expériences répétées sur l'incertitude que peut avoir un opérateur. La seconde décrira les propriétés de la répétition de l'expérience en fonction des caractéristique d'une expérience unique.

La définition de la mesure de confiance dans un problème d'inférence, ou plus généralement du modèle de confiance, est plus complexe que dans le cas d'une réalisation unique d'une expérience. Il n'existe pas de définition optimale. Deux méthodes principales peuvent pourtant être mises en avant : les approches "fréquentiste" et "bayésienne". Les informations utilisées dans les deux cas diffèrent légèrement. L'approche fréquentiste est conservatrice. Elle évite l'introduction d'un prior qui décrirait une préconception sur le résultat. Ses résultats souffrent par contre d'une certaine rigidité. L'approche bayésienne au contraire introduit une information subjective au travers d'un prior. En contrepartie, elle définit un modèle de confiance modulable et aisément généralisable.

Il faut souligner encore une fois que les deux approches, fréquentiste et bayésienne, utilisent toutes deux les deux mesures de confiance et de probabilité. Elles correspondent à deux alternatives pour construire un modèle de confiance à l'aide des informations en présence.

Dans le cas où il n'y a pas d'expérience, il n'existe pas de mesure de probabilité. Seule la mesure de confiance peut être définie. A priori, il se pourrait qu'il existe encore une fois différentes manières de construire un modèle de confiance. Les informations utilisées et leur traduction dans le modèle considéré doivent donc être soigneusement étudiées.

Vocabulaire

Avant de conclure ce chapitre, je souhaite souligner une dernière fois la difficulté de vocabulaire qui existe actuellement. Pour cela, je désigne dans les paragraphes suivants par propension ce que j'ai appelé jusque là probabilité.

Il s'avère que les partisans de l'approche fréquentiste de l'inférence désignent par "probabilité" la propension, de la même manière que je l'ai fait jusque là. Au contraire, les partisans de l'approche bayésienne désignent par "probabilité" (ou "probabilité bayésienne") ce que j'ai appelé confiance. Un premier point d'incompréhension entre les deux groupes vient donc du fait qu'ils désignent par le même terme deux notions très différentes.

Un deuxième point d'incompréhension provient à mon sens de l'interprétation par de nombreux utilisateurs en terme d'incertitude de la mesure de propension, effectuée en

1.5 Conclusion

toute rigueur à tort.

D'un point de vue personnel, j'aimerais réserver l'appellation "probabilité" à la définition optimale de la mesure de confiance dans le cas particulier de l'étude du résultat d'une unique expérience, *i.e.* dans le cas où la mesure de confiance est numériquement égale à la mesure de propension. En effet, c'est un observateur qui va vouloir dire qu'un événement est "probable" ou non. Le terme correspond bel et bien à un degré de confiance d'une personne sur la possibilité que l'événement ait lieu. Pourtant, on parle bien de la "probabilité d'un événement", indépendamment de l'observateur considéré. Autrement dit, on considère qu'il s'agit d'une caractéristique de l'événement, une donnée essentiellement objective. Le terme de "probabilité" regroupe donc les propriétés des deux mesures de propension et de confiance.

Dans le cas où les mesures de confiance et de propension sont numériquement égales, elles peuvent formellement être confondues. Associer un nouveau terme, celui de "probabilité", à la mesure dans ce cas est donc intéressant. Celle-ci peut alors être interprétée en terme de confiance et de propension tour à tour. Cela correspond à ce qui est effectué en pratique. Les utilisateurs des méthodes fréquentistes en inférence parlent de "niveau de confiance" alors qu'ils n'utilisent que la mesure de propension. Il s'agit en fait du cas particulier où confiance et propension sont égales.

C'est également avec ce choix de vocabulaire que l'on peut dire que la mesure de confiance "étend" la notion de probabilité. On se rapproche en ce sens de l'assertion selon laquelle la notion de probabilité bayésienne "étend" la notion de probabilité au sens fréquentiste. En effet, selon la définition du terme proposée, la mesure de probabilité est un cas (très) particulier de mesure de confiance. Il s'agit du cas où elle est précisément égale à la valeur de la mesure de propension. Si mesures de propension et de confiance sont indistinguées sous la forme d'une mesure de probabilité, la notion de confiance est bel et bien une extension de celle de probabilité.

Finalement, alors que l'on peut définir une mesure de confiance différente d'une mesure de propension, que celle-ci existe ou non dans le problème étudié, la notion de propension peut *toujours* être utilisée pour définir un modèle de confiance. Une mesure de propension peut donc toujours être qualifiée de mesure de probabilité alors qu'il n'en est pas de même pour une mesure de confiance (toujours pour le choix de vocabulaire proposé). Ainsi, les partisans fréquentistes ont tous les droits de parler de "probabilité" en lieu et place de "propension", et ce même s'ils effectuent un raccourcis logique.

Ce choix de vocabulaire me semble donc correspondre à l'acceptation courante du mot "probabilité". Il me semble également utile pour expliquer les différentes assertions, qui semblent parfois incompatibles, des partisans bayésiens et fréquentistes. S'il n'en tenait qu'à moi, il s'agit donc des termes que j'utiliserais. Pourtant, le flou artistique qui règne sur les définitions des termes me fait dire qu'il serait certainement plus prudent de ne pas introduire un nouveau choix de termes. C'est pour cela que je me suis restreint au court de ce chapitre à l'utilisation du terme de probabilité pour la propension.

Bibliographie

- [1] G. D'Agostini, "Bayesian reasoning in data analysis, a critical introduction", World Scientific Publishing, 2003.
- [2] E.T. Jaynes, "Probability theory, the logic of science", Cambridge University Press, 2003.

Chapitre 2

Incertitudes théoriques

Sommaire

2.1	Incertitude théorique sur une série perturbative tronquée . .	55
2.1.1	Notations	55
2.1.2	Estimation conventionnelle de l'incertitude	55
2.1.3	Variation de l'échelle de renormalisation	56
2.1.4	Exemple du processus $e^+e^- \rightarrow$ hadrons	57
2.1.5	Ordre de grandeur des coefficients de la section	59
2.2	Modèle de confiance	60
2.2.1	Définition du modèle	60
2.2.2	Comportement du modèle	62
2.3	Discussion du modèle	65
2.3.1	Paramètres inconnus	65
2.3.2	Coefficients de même ordre de grandeur	66
2.3.3	Information partagée	67
2.3.4	Incertitude résiduelle	70
2.3.5	Nature du paramètre caché	73
2.4	Estimation de l'incertitude de troncature	74
2.4.1	Densité conditionnelle $f(\Delta_k c_0, \dots, c_k)$ approchée	74
2.4.2	Validité de l'approximation	78
2.4.3	Évaluation numérique de $f(\Delta_k c_0, \dots, c_k)$	80
2.4.4	Choix de densité $f(c_n \bar{c})$ non uniforme	81
2.5	Comparaison du modèle et de la méthode conventionnelle . .	81
2.5.1	Niveaux de confiance des intervalles conventionnels	81
2.5.2	Évolution des valeurs de confiance	84
2.5.3	Exemple du processus $e^+e^- \rightarrow$ hadrons	88
2.6	Approximation des ordres supérieurs	89
2.6.1	Ajustement du modèle	90
2.6.2	Comportement du modèle	92

Introduction

Le LHC (*Large Hadron Collider*) a finalement été lancé et, aucun trou noir n'ayant englouti la Terre, la course à la collecte de données et à leur analyse a maintenant commencé pour de bon. Tandis que l'objectif à court terme est de redécouvrir le Modèle Standard, celui à long terme est bien entendu de trouver des signaux de "nouvelle physique", qu'il s'agisse du boson de Higgs, de la supersymétrie ou de quelque chose d'autre, plus exotique, et potentiellement inattendu. Bien que chacun espère que des découvertes s'annonceront sous la forme de signaux clairs, il est concevable, et probablement également inévitable dans un premier temps, qu'elles se présentent plutôt cachées sous la forme d'un décalage subtile entre théorie et expérience. Si c'est le cas, un contrôle complet des incertitudes sur les prédictions théoriques devient naturellement d'une importance cruciale : comparant une mesure expérimentale à un calcul théorique, il doit être possible de dire s'ils sont en accord ou non, et avec quel degré de confiance une telle déclaration peut être effectuée. Ceci est impossible à réaliser sans que les mesures expérimentales et les prédictions théoriques ne soient fournies avec un degré, sensé et communément accepté, d'incertitude.

Bien que tout ce qui est présenté plus bas puisse s'appliquer à tout type de prédiction théorique en théorie perturbative, l'étude se placera dans le contexte de la Chromodynamique Quantique (*Quantum Chromodynamics*, QCD) : de nombreux processus et signaux de fonds au LHC dépendent du domaine de la QCD et, dû à la taille relativement large de la constante de couplage forte α_s et donc de la convergence plus lente des séries perturbatives, le problème de la précision théorique y est plus pressant. Les prédictions théoriques en QCD contiennent de multiples ingrédients, certaines contributions devant être en fin de compte extraites des données expérimentales, comme les fonctions de distribution partoniques (*Parton Distribution Functions*, PDFs) pour les collisions hadroniques et la valeur de α_s . Durant ces dernières années, de nombreux progrès ont été effectués sur α_s et les PDFs. L'incertitude avec laquelle est connue la constante de couplage est maintenant assez faible (voir par exemple [1]). De plus, différents groupes [2, 3, 4, 5, 6] ont extraits des jeux de PDFs avec des incertitudes associées d'origine expérimentales, et fournis des cadres formels pour les propager proprement à l'observable que l'on peut chercher à calculer. D'énormes progrès ont également été faits en effectuant des calculs aux ordres supérieurs pour un nombre important d'observables phénoménologiquement intéressantes [7], améliorant ainsi potentiellement la précision avec laquelle elles sont connues.

Un domaine où les progrès peuvent être discutés est celui de la compréhension des incertitudes théoriques résiduelles provenant des ordres supérieurs inconnus en théorie des perturbations. Cette incertitude est généralement estimée en variant des échelles non-physiques, notées ici collectivement μ , contenues dans un résultat perturbatif, comme celles de renormalisation et de factorisation, autour d'une valeur centrale μ_0 , habituellement choisie pour coïncider avec une échelle physique Q du processus étudié. Cette méthode, la gamme sur laquelle varier les échelles μ (typiquement $[\mu_0/2, 2\mu_0]$) et leur valeur centrale $\mu_0 = Q$ sont hautement conventionnels mais pourtant communément admises. Ils permettent à la communauté d'échanger efficacement une incertitude conventionnelle qui peut être facilement comparée entre différents calculs.

Parmi les limites de cette procédure, la plus flagrante est probablement qu'elle ne permet pas d'estimer la *confiance* qui peut être portée à l'intervalle d'incertitude obtenu. Plus précisément, il n'est pas possible d'associer un niveau de confiance, 68,3%, 95,5%, ou 99,7% par exemple, à la bande d'incertitude selon lequel elle contiendrait la somme

exacte¹ de la série. Ce manque de caractérisation claire de l'incertitude théorique perturbative signifie également que les procédures pour la combiner avec d'autres sources d'incertitudes, comme la valeur de α_s ou celle des PDFs, ne sont qu'au mieux ambiguës et discutables, comme illustré par la discussion récente [8] sur la manière d'estimer proprement l'incertitude totale sur la prédiction de la section efficace de production de Higgs en collision hadronique. Tout ceci rend potentiellement impossible d'évaluer complètement et rigoureusement la confiance avec laquelle un résultat expérimental peut concorder avec la théorie, rendant le pari sur la découverte d'une nouvelle physique une proposition entièrement non-scientifique et risquée.

L'objectif du travail présenté dans ce chapitre est précisément d'essayer de rendre un tel pari potentiellement sûr en utilisant le formalisme de la confiance (probabilités bayésiennes) présenté au chapitre 1 précédent. Pour cela, un modèle est construit et aboutit à une mesure de confiance bien définie pour une incertitude théorique perturbative, de telle manière que les niveaux de confiance peuvent être explicitement calculés. La section 2.1 revient dans un premier temps sur la méthode conventionnelle d'estimation des incertitudes théoriques par variation d'échelle. Elle discute également de la validité de l'approche et extrait l'hypothèse qui servira de pilier au modèle de confiance. La section 2.2 présente la mesure de confiance en elle-même, l'ensemble des informations utilisées dans sa construction et son comportement. La section 2.3 discute des hypothèses de bases du modèle et propose dans certain cas des variantes possibles. La mesure de confiance sur le reste de la section perturbative est finalement présentée dans la section 2.4 puis comparée aux estimations conventionnelles d'incertitudes dans la section 2.5. La section 2.6 propose enfin une extension du modèle au cas de coefficients d'ordres supérieurs partiellement connus. Une partie de ce chapitre a donné lieu à la publication [9].

Avant de clore cette introduction, le point suivant doit être souligné : ce travail ne cherche pas à améliorer la connaissance sur une prédiction perturbative en ajoutant une information non physique ou juste une spéculation sur sa forme, ou en recherchant une (improbable) information physique dans les mathématiques du formalisme. La seule information entrant dans le modèle, hormis les hypothèses généralement implicitement effectuées dans tous les cas (hypothèses motivées à la section 2.1, formalisées à la section 2.2 puis discutées en 2.3), sont les expressions des coefficients calculés. Les hypothèses utilisées ne sont pas nécessairement justes mais une fois qu'elles ont été effectuées, le chemin vers le calcul des niveaux de confiance est rigoureux.

1. On ne prend pas en compte ici le fait que les séries en QCD ne sont généralement pas convergentes mais simplement asymptotiques. Les prémisses d'un comportement asymptotique n'apparaissant habituellement qu'à des ordres assez élevés, cela n'affecte pas normalement les applications phénoménologiques réalistes. En pratique, la valeur de la "somme réelle de la série" peut être définie comme la valeur asymptotique de la série calculée à l'aide d'une prescription appropriée, ou même comme le résultat plus raffiné d'un ordre supérieur, bien que l'on continuera d'utiliser le terme de "vrai résultat" pour désigner le résultat désiré au-delà de ce qui a été réellement calculé.

2.1 Incertitude théorique sur une série perturbative tronquée

2.1.1 Notations

Dans la suite, la section efficace d'un processus ayant lieu à l'énergie Q est notée $\sigma(Q)$. Elle est obtenue sous forme la forme d'une série perturbative :

$$\sigma(Q) = \sum_{n=0}^{\infty} c_n(Q, \mu_R) \alpha_s^n(\mu_R) \quad (2.1)$$

μ_R est l'échelle de renormalisation, que l'on notera plus simplement μ dans la suite, et $\alpha_s(\mu)$ est la constante de couplage forte. La section efficace de production de hadrons dans les collisions e^+e^- est présentée comme exemple dans la section 2.1.4. Pour simplifier les notations, les dépendances des coefficients et de la constante de couplage ne sont pas écrites lorsqu'ils sont évalués en $\mu = Q$:

$$\sigma(Q) = \sum_{n=0}^{\infty} c_n(Q, Q) \alpha_s^n(Q) \equiv \sum_{n=0}^{\infty} c_n \alpha_s^n \quad (2.2)$$

La constante de couplage $\alpha_s(\mu)$ évolue avec l'échelle de renormalisation selon l'équation :

$$\frac{d\alpha_s}{d \ln \mu^2} = \beta(\alpha_s) = -\alpha_s^2 \sum_{n=0}^{\infty} \beta_n \alpha_s^n \quad (2.3)$$

$\beta(\alpha_s)$ est la fonction beta. À l'opposé, la section efficace totale ne dépend pas du choix de paramètre μ .

La somme partielle calculée jusqu'au dernier ordre connu, évaluée en une valeur μ quelconque de l'échelle de renormalisation ou en $\mu = Q$ fixé, est notée dans la suite :

$$\sigma_k(Q, \mu) \equiv \sum_{n=0}^k c_n(Q, \mu) \alpha_s^n(\mu) \quad \sigma_k \equiv \sum_{n=0}^k c_n \alpha_s^n \quad (2.4)$$

Le reste de la série est à son tour défini par :

$$\Delta_k(Q, \mu) \equiv \sum_{n=k+1}^{\infty} c_n(Q, \mu) \alpha_s^n(\mu) \quad \Delta_k \equiv \sum_{n=k+1}^{\infty} c_n \alpha_s^n \quad (2.5)$$

La somme totale de la série perturbative est donc la somme du résultat partiel σ_k , connu, et de Δ_k , inconnu. La contribution partielle est généralement utilisée comme estimation de $\sigma(Q)$. L'erreur commise est alors la valeur du terme Δ_k . Évaluer l'incertitude théorique correspond donc à évaluer les valeurs envisageables du reste de la série.

2.1.2 Estimation conventionnelle de l'incertitude

La dépendance explicite de $\sigma_k(Q, \mu)$ dans le paramètre μ de l'équation (2.4) rappelle qu'un calcul perturbatif tronqué à un ordre fini conserve une dépendance non physique dans le paramètre μ . Celle-ci est habituellement utilisée pour estimer l'erreur commise en assimilant la somme partielle à la somme totale. Intuitivement, la série tronquée n'a pas de raison d'être une meilleur estimation de la somme complète lorsqu'elle est estimée précisément en $\mu = Q$ plutôt que lorsqu'elle est calculée en des valeurs de l'échelle de

renormalisation suffisamment voisines de l'énergie du processus. De plus, la série complète ne dépend pas du paramètre μ non physique. La variation de la somme partielle avec l'échelle de renormalisation semblerait par conséquent être un témoin de sa différence avec $\sigma(Q)$.

Un résultat théorique est donc généralement cité sous la forme d'un intervalle d'incertitude $[\sigma_k^-, \sigma_k^+]$, défini à l'aide de la dépendance de la section partielle en fonction de μ et contenant la valeur de σ_k . Différentes possibilités existent :

1.

$$\sigma_k^- = \min\{\sigma_k(Q, Q/2), \sigma_k(Q, 2Q)\}, \sigma_k^+ = \max\{\sigma_k(Q, Q/2), \sigma_k(Q, 2Q)\} \quad (2.6)$$

2.

$$\sigma_k^- = \min_{\mu \in [Q/2, 2Q]} \{\sigma_k(Q, \mu)\}, \sigma_k^+ = \max_{\mu \in [Q/2, 2Q]} \{\sigma_k(Q, \mu)\} \quad (2.7)$$

3.

$$\sigma_k^\pm = \sigma_k \pm \frac{\delta_k}{2} \quad (2.8)$$

où :

$$\delta_k \equiv |\sigma_k(Q, Q/2) - \sigma_k(Q, 2Q)| \quad (2.9)$$

4. même forme que dans l'équation (2.8) mais pour

$$\delta_k \equiv \max_{\mu \in [Q/2, 2Q]} \{\sigma_k(Q, \mu)\} - \min_{\mu \in [Q/2, 2Q]} \{\sigma_k(Q, \mu)\} \quad (2.10)$$

Les paramètres σ_k^- et σ_k^+ sont donc créés à partir des valeurs extrémales de $\sigma_k(Q, \mu)$ sur un intervalle $[Q/2, 2Q]$ ou de ses valeurs sur les extrémités, de manière à former un segment centré sur σ_k ou non. Le choix de variation de μ d'un facteur 2 autour de l'énergie du processus Q est totalement conventionnel.

Cette approche est limitée. Quand bien même la largeur de l'intervalle d'incertitude obtenu, que l'on notera δ_k dans tous les cas, estimerait correctement l'ordre de grandeur du reste Δ_k de la série, il n'y a aucun moyen de décider jusqu'à quel degré cela est vrai.

2.1.3 Variation de l'échelle de renormalisation

En fait, il n'existe aucune raison objective simple pour laquelle la largeur de l'intervalle $[\sigma_k^-, \sigma_k^+]$ devrait correspondre à une bonne estimation du reste de la série Δ_k . La définition (2.5) montre que ce dernier est une fonction des coefficients inconnus c_n , $n > k$ évalués en $\mu = Q$. À l'opposé, la dépendance dans l'échelle de renormalisation de la section partielle est entièrement fixée par la valeur des coefficients connus c_n , $n \leq k$.

En effet, partant l'indépendance de la somme totale σ de la série perturbative par rapport à μ et l'équation d'évolution (2.3) de la constante de couplage, il est possible de reconstruire la dépendance complète de tout coefficient $c_n(Q, \mu)$ dans l'échelle de renormalisation en fonction de la valeur des coefficients c_l , $l \leq n$ d'ordre inférieur, évalués en $\mu = Q$:

$$\frac{d\sigma}{d \ln \mu^2} = 0 = \sum_{n=0}^{\infty} \left(\frac{dc_n}{d \ln \mu^2} \alpha_s^n - nc_n \alpha_s^{n-1} \sum_{l=0}^{\infty} \beta_l \alpha_s^{l+2} \right) \quad (2.11)$$

$$= \sum_{n=0}^{\infty} \alpha_s^n \left[\frac{dc_n}{d \ln \mu^2} - \sum_{l=0}^{n-1} l \beta_{n-1-l} c_l \right] \quad (2.12)$$

2.1 Incertitude théorique sur une série perturbative tronquée

En supposant que la séparation entre les coefficients et la constante de couplage ait été correctement effectué, *i.e.* en supposant que les coefficients ne dépendent pas de Λ_{QCD} :

$$\frac{dc_n}{d \ln \mu^2} = \sum_{l=0}^{n-1} j \beta_{n-1-l} c_l \quad \forall n \geq 0 \quad (2.13)$$

Les termes c_0 and c_1 sont donc toujours indépendants de μ . Par récurrence forte, les coefficients c_n , $n \geq 1$, sont des polynômes de degré inférieur ou égal à $n - 1$ en $\ln \frac{\mu^2}{Q^2}$:

$$c_n(Q, \mu) = \sum_{l=0}^{n-1} c_{n,l} \left(\ln \frac{\mu^2}{Q^2} \right)^l \quad \forall n \geq 1 \quad (2.14)$$

C'est bien le cas en $n = 1$ puisque c_1 est indépendant de μ . Supposant que c'est le cas pour tous les rangs jusqu'à n , l'équation (2.13) montre que $\frac{dc_{n+1}}{d \ln \mu^2}$ est un polynôme d'ordre inférieur ou égal à $n - 1$. $c_{n+1}(\mu)$ est donc bien un polynôme de degré inférieur ou égal à n . En réécrivant l'équation (2.13) ordre par ordre en $\ln \frac{\mu^2}{Q^2}$, on obtient finalement une relation de récurrence sur les éléments $c_{n,l}$:

$$c_{n,0} = c_n \quad c_{n,l} = \frac{1}{l} \sum_{i=0}^{n-1} i \beta_{n-1-i} c_{i,l-1} \quad \forall l \geq 1 \quad (2.15)$$

L'intégralité des coefficients $\{c_{n,l}\}_{l \leq n-1}$, et par suite la forme (2.14) de tout $c_n(Q, \mu)$, est donc bel et bien une fonction des coefficients $c_{i,0} = c_i$, $i \leq n$.

La dépendance (2.4) de $\sigma_k(Q, \mu)$ en μ et les intervalles $[\sigma_k^-, \sigma_k^+]$ sont par conséquent fixés par la valeur des coefficients calculés c_n , $n \leq k$. L'indépendance de la somme totale σ par rapport à μ n'impose aucun lien entre la valeur c_n en $\mu = Q$ des coefficients non calculés et la variation avec l'échelle de renormalisation de la somme partielle du résultat tronqué $\sigma_k(Q, \mu)$.

2.1.4 Exemple du processus $e^+e^- \rightarrow$ hadrons

La section totale $\sigma(e^+e^- \rightarrow$ hadrons) est l'une des observables les plus précisément connues en QCD perturbative. Ses coefficients ont été calculés de manière exacte jusqu'à l'ordre α_s^3 et une valeur approchée de l'ordre α_s^4 a été obtenue. Les valeurs précises sont référencées par exemple dans [1]. Ce processus est un outil idéal pour tester les résultats de cette étude.

On réutilise les notations introduites dans l'équation (9.7) de [1] :

$$\frac{\sigma(e^+e^- \rightarrow \text{hadrons}, Q)}{\sigma(e^+e^- \rightarrow \mu^+\mu^-, Q)} \equiv R(Q) = R_{EW}(Q) (1 + \delta_{QCD}(Q)) \quad (2.16)$$

$R_{EW}(Q)$ est une prédiction purement électrofaible. On s'intéresse au second facteur. Le paramètre $\delta_{QCD}(Q)$ se développe en série perturbative :

$$\delta_{QCD}(Q) = \sum_{n=1}^{\infty} c_n \alpha_s^n \quad (2.17)$$

L'utilisation du symbole δ dans un cas différent de celui de la section 2.1.2 est un accident malheureux. Deux choix de série perturbative $\sigma(Q)$ à étudier sont possibles :

$$\sigma(Q) \equiv 1 + \delta_{QCD}(Q) \quad (2.18)$$

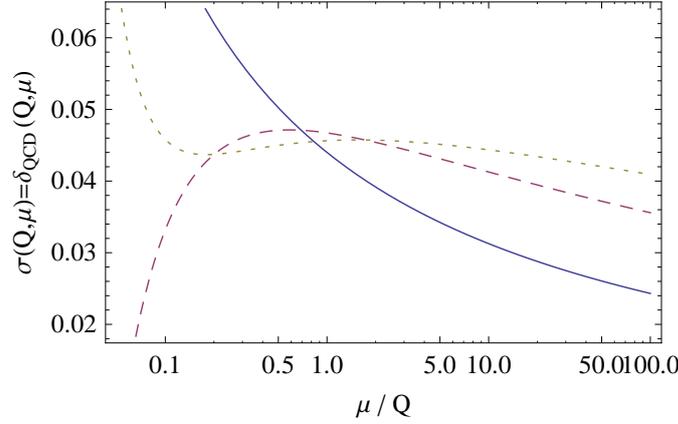


Figure 2.1 – Séries perturbatives partielles $\sigma_k(Q, \mu)$ de la section totale $\sigma(Q)$ définie par l'équation (2.18) pour $k = 1, 2$ et 3 , des tirets les plus espacés aux plus rapprochés.

ou :

$$\sigma(Q) \equiv \delta_{QCD}(Q) \quad (2.19)$$

En ce qui concerne l'étude de la variation de l'échelle de renormalisation, la seule différence entre les deux possibilités porte sur la valeur centrale des prédictions partielles, décalée d'une valeur $c_0 = 1$ d'un cas sur l'autre.

Les valeurs des premiers coefficients sont, pour $n_f = 5$ quarks non massifs :

$$c_1 = 0,31831 \quad c_2 = 0,142785 \quad c_3 = -0,412969 \quad c_4 \simeq -0,821356 \quad (2.20)$$

Pour le choix (2.19) de série perturbative et pour $\alpha_s(Q) = 0,118$, soit une échelle $Q \simeq 89,16$ GeV :

$$\sigma_1 = 0,0375606 \quad \sigma_2 = 0,0395487 \quad \sigma_3 = 0,0388702 \quad (2.21)$$

La dépendance complète en μ des séries tronquées $\sigma_k(Q, \mu)$, $k = 1, 2, 3$, se déduit de ces coefficients. La figure 2.1 présente ces trois séries partielles. On évalue les intervalles d'incertitude conventionnels aux différents ordres. Pour la définition (2.7) :

$$\begin{aligned} [\sigma_1^-; \sigma_1^+] &= [0,03401; 0,04197] \\ [\sigma_2^-; \sigma_2^+] &= [0,03871; 0,0398] \\ [\sigma_3^-; \sigma_3^+] &= [0,03855; 0,03893] \end{aligned} \quad (2.22)$$

La dépendance en μ des séries partielles $\sigma_k(Q, \mu)$ et les intervalles conventionnels précédents ont été calculés en utilisant une équation d'évolution pour α_s prenant en compte les coefficients de la fonction β jusqu'à β_2 , *i.e.* ce qui est nécessaire pour $\sigma_3(Q, \mu)$. Les valeurs des coefficients de la fonction β sont présentés par exemple dans [1]. Il est possible d'utiliser une précision plus faible pour $\sigma_1(Q, \mu)$ et $\sigma_2(Q, \mu)$ mais nous avons vérifié que ces changements ne modifient pas plus que le dernier chiffre significatif des valeurs approchées fournies plus haut.

2.1 Incertitude théorique sur une série perturbative tronquée

2.1.5 Ordre de grandeur des coefficients de la section

Malgré l'indépendance de Δ_k et de la variation des échelles, il s'avère que l'estimation conventionnelle de l'erreur présentée dans la section 2.1.2 donne des résultats compatibles avec l'expérience. D'une manière ou d'une autre, δ_k est du même ordre de grandeur que Δ_k dans de nombreux cas.

En première approximation, la largeur de l'intervalle d'incertitude conventionnel peut se réexprimer à l'aide de la dérivée de σ_k par rapport à μ en $\mu = Q$:

$$\delta_k \simeq \left| \frac{d\sigma_k}{d\mu} \right|_{\mu=Q} (2Q - Q/2) \simeq 3 \left| \frac{d\sigma_k}{d \ln \mu^2} \right|_{\mu=Q} \quad (2.23)$$

avec le changement de variable $d \ln \mu^2 = 2d\mu/\mu$. En utilisant les equations (2.12) et (2.13) :

$$\frac{d\sigma_k}{d \ln \mu^2} = \sum_{i=k+1}^{\infty} \alpha_s^i \sum_{j=0}^k j \beta_{i-1-j} c_j \quad (2.24)$$

La variation d'échelle δ_k est donc d'ordre α_s^{k+1} , soit un ordre supérieur que le dernier ordre calculé. Il est du même ordre en α_s que Δ_k .

Dans la gamme d'énergie où la constante de couplage forte α_s est très inférieure à l'unité :

$$\frac{d\sigma_k}{d \ln \mu^2} \simeq \alpha_s^{k+1} \sum_{j=0}^k j \beta_{k-j} c_j \quad (2.25)$$

Sous l'hypothèse que les coefficients c_j sont tous du même ordre de grandeur, et étant donné que les coefficients β_i connus vérifient la relation (cf table 2.1) :

$$i \beta_{k-i} |_{k \neq i} \ll i \beta_{k-i} |_{k=i} \quad (2.26)$$

la largeur des intervalles d'incertitude conventionnels δ_k prend la forme :

$$\delta_k \simeq 3k \beta_0 \cdot \alpha_s^{k+1} |c_k| \quad (2.27)$$

Sous la même hypothèse, l'expression (2.5) du reste devient :

$$\Delta_k \simeq \alpha_s^{k+1} c_{k+1} \quad (2.28)$$

Ainsi, s'il s'avère que les coefficients de la série perturbative sont du même ordre de grandeur, le reste de la série sera bel et bien du même ordre de grandeur que la largeur de l'intervalle d'incertitude conventionnel :

$$|\Delta_k| \simeq \delta_k \quad (2.29)$$

Cette hypothèse est donc une explication possible de l'apparent succès de la variation d'échelle à estimer la partie tronquée de la série.

L'hypothèse de coefficients du même ordre de grandeur n'est pas gratuite puisqu'elle est en fait à la base de l'approche perturbative. Les contributions des ordres inconnus sont implicitement considérés être des corrections à celles des ordres calculés, leur différence provenant de puissances d'un facteur d'expansion tel que α_s , inférieur à l'unité. Cette approche est plus généralement valide tant que la valeur des coefficients c_i n'augmente pas trop vite.

k	$i = 1$	$i = 2$	$i = 3$	$i = 4$
1	0.61	–	–	–
2	0.24	1.22	–	–
3	0.07	0.49	1.83	–
4	0.19	0.15	0.73	2.44

Tableau 2.1 – First values of $i\beta_{k-i}$, calculated with $n_f = 5$.

L'étude effectuée suppose que les coefficients de la série perturbative de la section efficace d'un processus donné partagent un même ordre de grandeur. Les premiers coefficients calculés permettent de construire une estimation de ce paramètre commun. Une fois extrait, celui-ci fournit une plage de valeurs envisageables pour chacun des coefficients inconnus. L'incertitude sur le reste Δ_k est alors obtenu par combinaison de celles sur les coefficients non calculés.

2.2 Modèle de confiance

2.2.1 Définition du modèle

Le modèle de confiance décrit ci-dessous est construit dans le but de fournir une densité de confiance conditionnelle $f_{\mathbb{C}}(\Delta_k | c_0, \dots, c_k)$ sur le reste d'une série perturbative étant connus ses premiers termes. Pour cela, une mesure de confiance générale, utilisable pour toute série perturbative, est créée. La mesure de probabilité n'apparaissant pas dans ce chapitre, toutes les densités présentées seront des densités de confiance $f_{\mathbb{C}}$. Pour simplifier les notations, les indices \mathbb{C} seront omis dans la suite.

PARAMÈTRES INCONNUS

Les paramètres mis en jeu sont la section totale σ , les sections partielles σ_k et les restes Δ_k pour tout $k \geq 0$ ainsi que les coefficients c_0, c_1, \dots de la série. La section totale, les sections partielles et les restes sont des fonctions des coefficients. La mesure est restreinte à l'ensemble des termes de la série.

En toute rigueur, la valeur la constante de couplage forte α_s , ou bien les coefficients de la fonction β et sa valeur $\alpha_s(M_Z)$ au pôle de la masse du boson Z , devrait également être ajoutée au modèle en tant que paramètre *a priori* inconnu. Pour une première approche, cette source d'incertitude n'est pas modélisée ici. La valeur du paramètre d'expansion est supposée précisément connue.

COEFFICIENTS DE MÊME ORDRE DE GRANDEUR

Les coefficients de la série perturbative sont supposés partager un même ordre de grandeur. Plus précisément ils sont supposés partager une forme de maximum commun \bar{c} , spécifique au processus étudié. Suivant le raisonnement présenté à la section 2.1.5 précédente, les premiers coefficients c_n , $n \leq k$, calculés donneront une information sur ce paramètre, restreignant les valeurs considérées comme possibles pour les termes inconnus c_n , $n > k$. L'ensemble des variables inconnues qui définissent l'espace sur lequel on crée

2.2 Modèle de confiance

la mesure de confiance est donc en fait constitué de ce paramètre \bar{c} et de l'ensemble des coefficients $\{c_i, i \geq 0\}$ a priori inconnus.

Trois hypothèses supplémentaires achèvent la définition de la mesure de confiance sur l'espace des variables $(\bar{c}, c_0, c_1, \dots)$. Celles-ci fixent donc également toute mesure héritée sur un sous-espace défini par la valeur des k premiers coefficients d'un processus particulier.

INFORMATION PARTAGÉE

Le paramètre \bar{c} traduit l'hypothèse d'une information, un ordre de grandeur commun, partagée par l'ensemble des coefficients. Aucun autre lien éventuel (indirect, cf section 2.3.3.4) n'est pris en compte. Lorsque \bar{c} est connu, les incertitudes résiduelles sur les coefficients sont alors indépendantes. Les variables inconnues correspondant aux coefficients sont donc définies comme des variables mutuellement indépendantes lorsque \bar{c} est connu :

$$f(\{c_i\}_I|\bar{c}) = \prod_{i \in I} f(c_i|\bar{c}) \quad \forall I \quad (2.30)$$

Dans ce modèle, la valeur de \bar{c} correspond à l'information totale que partagent les coefficients. Il s'agit de l'information maximale qui peut être extraite des valeurs des premiers coefficients c_0, \dots, c_k pour restreindre les valeurs envisageables des coefficients non calculés $c_n, n > k$.

INCERTITUDE RÉSIDUELLE

Connaissant *a priori* le coefficient \bar{c} , la densité de confiance résiduelle sur la valeur d'un coefficient inconnu est supposée uniforme :

$$f(c_n|\bar{c}) = \frac{1}{2\bar{c}} \begin{cases} 1 & \text{si } |c_n| \leq \bar{c} \\ 0 & \text{si } |c_n| > \bar{c} \end{cases} \equiv \frac{1}{2\bar{c}} \chi_{|c_n| \leq \bar{c}} \quad (2.31)$$

où χ_A représente la fonction caractéristique d'un ensemble A . Il serait possible (et certainement plus juste) d'utiliser une densité qui ne s'annule pas, comme une densité gaussienne, mais la forme (2.31) fournit des expressions analytiques plus simples. Cette expression est utilisée dans un premier temps pour étudier le comportement du modèle.

NATURE DU PARAMÈTRE CACHÉ

Le paramètre \bar{c} est supposé exister. Sa valeur n'est par contre pas connue *a priori*. Elle est cachée dans celles des coefficients c_n . Tant qu'aucun coefficient n'a été calculé, il est uniquement possible de dire qu'il s'agit d'un réel positif. Tous les ordres de grandeur du paramètre \bar{c} sont supposés également envisageables : la densité de confiance sur son logarithme est définie comme la limite d'une densité uniforme entre $-|\ln \epsilon|$ et $|\ln \epsilon|$ lorsque ϵ tend vers zéro.

$$f_\epsilon(\ln \bar{c}) = \frac{1}{2|\ln \epsilon|} \chi_{|\ln \bar{c}| \leq |\ln \epsilon|} \Leftrightarrow f_\epsilon(\bar{c}) = \frac{1}{2|\ln \epsilon|} \frac{1}{\bar{c}} \chi_{\epsilon \leq \bar{c} \leq 1/\epsilon} \quad (2.32)$$

Tous les calculs (analytiques et numériques) sont effectués à l'aide de cette densité en ϵ non nul. Les résultats finaux seront obtenus par passage à la limite en zéro. Une densité s'annule dans cette limite lorsque que les informations considérées sont insuffisantes pour restreindre les valeurs envisageables du paramètre étudié. Par exemple, la densité $f_\epsilon(\bar{c})$ tend vers une "densité uniformément nulle" car il n'y a pas assez d'information dans le modèle seul pour privilégier une valeur de \bar{c} . Comme le montre la section suivante, la densité $f(\bar{c}|c_0)$ est par contre bien définie.

2.2.2 Comportement du modèle

2.2.2.1 Densité $f_\epsilon(\bar{c}, c_0, c_1, \dots)$ de la mesure de confiance

La mesure sur l'espace formé par les variables inconnues \bar{c} et c_i , $i \geq 0$, est définie par la distribution $f_\epsilon(\bar{c}, c_0, c_1, \dots)$ lorsque ϵ est non nul. Son expression est obtenue à partir des hypothèses (2.30), (2.31) et (2.32). Plus précisément, l'expression de toute densité $f_\epsilon(\bar{c}, c_0, \dots, c_n)$ peut être obtenue, quelle que soit la valeur de l'entier n positif ou nul :

$$\begin{aligned} f_\epsilon(\bar{c}, c_0, \dots, c_n) &= f_\epsilon(c_0, \dots, c_n | \bar{c}) f_\epsilon(\bar{c}) \\ &= \left[\prod_{i=0}^n f(c_i | \bar{c}) \right] f_\epsilon(\bar{c}) \\ &= \left[\prod_{i=0}^n \frac{1}{2\bar{c}} \chi_{|c_i| \leq \bar{c}} \right] \frac{1}{2|\ln \epsilon|} \frac{1}{\bar{c}} \chi_{\epsilon \leq \bar{c} \leq \epsilon} \end{aligned} \quad (2.33)$$

Cette dernière expression se simplifie à l'aide de la propriété des fonctions caractéristiques :

$$\prod_{i \in I} \chi_{A_i} = \chi_{\cap_{i \in I} A_i} \quad (2.34)$$

En particulier, le produit des fonctions caractéristiques des ensembles définis par des inégalités $|c_i| \leq \bar{c}$ est égal à la fonction caractéristique de l'ensemble défini par : $\max(|c_0|, \dots, |c_n|) \leq \bar{c}$. La distribution $f_\epsilon(\bar{c}, c_0, \dots, c_n)$ se réécrit :

$$f_\epsilon(\bar{c}, c_0, \dots, c_n) = \frac{1}{2^{n+1}} \frac{1}{2|\ln \epsilon|} \frac{1}{\bar{c}^{n+2}} \chi_{\max(|c_0|, \dots, |c_n|, \epsilon) \leq \bar{c} \leq 1/\epsilon} \quad (2.35)$$

Lorsque le paramètre ϵ tend vers zéro, la densité tend vers une fonction uniformément nulle. Le modèle de confiance ne cherche pas à prédire un résultat. Les hypothèses (2.30), (2.31) et (2.32) à elles seules n'apportent pas suffisamment d'information pour privilégier certaines valeurs des différents paramètres.

2.2.2.2 Densité conditionnelle $f(\bar{c} | c_0, \dots, c_k)$

La densité $f(\bar{c} | c_0, \dots, c_k)$ correspond à la mesure de confiance sur \bar{c} dans le sous-espace où c_0, \dots, c_k sont fixés. Elle traduit la confiance qui peut être accordée, sous les hypothèses ayant construit le modèle, sur la valeur du paramètre \bar{c} lorsque les valeurs des premiers coefficients sont connus. Elle s'écrit en ϵ non nul, selon la formule de densité conditionnelle :

$$f_\epsilon(\bar{c} | c_0, \dots, c_k) = \frac{f_\epsilon(\bar{c}, c_0, \dots, c_k)}{f_\epsilon(c_0, \dots, c_k)} \quad (2.36)$$

La densité $f_\epsilon(c_0, \dots, c_k)$ est obtenue par marginalisation de la distribution $f_\epsilon(\bar{c}, c_0, \dots, c_k)$:

$$f_\epsilon(c_0, \dots, c_k) = \int f_\epsilon(\bar{c}, c_0, \dots, c_k) d\bar{c} \quad (2.37)$$

Après évaluation de l'intégrale (2.37), simplification de l'équation (2.36) et passage à la limite en ϵ nul, la densité conditionnelle sur la valeur du paramètre \bar{c} s'écrit :

$$f(\bar{c} | c_0, \dots, c_k) = (k+1) \frac{\bar{c}^{-k+1}}{\bar{c}^{k+2}} \chi_{\bar{c} > \bar{c}_{(k)}} \quad (2.38)$$

2.2 Modèle de confiance

où :

$$\bar{c}_{(k)} = \max(|c_0|, \dots, |c_k|) \quad (2.39)$$

L'information fournie par un unique coefficient, ajoutée aux hypothèses qui définissent le modèle, suffit à restreindre les valeurs envisageables du paramètre \bar{c} . La figure 2.2 présente la forme de la densité (2.38) pour différentes valeurs de k . Lorsqu'une infinité de coefficients sont connus ($k \rightarrow \infty$), celle-ci tend vers une distribution de Dirac centrée sur la valeur $\bar{c}_{(k)}$, fixé arbitrairement à un pour le bien du graphique.

Le paramètre $\bar{c}_{(k)}$ en k fini fournit une estimation la valeur de \bar{c} . Plus précisément, il fixe tout d'abord une borne inférieure sur ses valeurs envisageables. D'après la définition de la densité conditionnelle (2.31), tous les coefficients doivent être plus petits que \bar{c} . En particulier, les valeurs absolues des termes calculés $|c_0|, \dots, |c_k|$ doivent être inférieures à \bar{c} . Ce dernier ne peut donc être que supérieur à leur maximum $\bar{c}_{(k)}$.

D'autre part, le paramètre $\bar{c}_{(k)}$ donne un ordre de grandeur de l'incertitude sur \bar{c} . La densité (2.38) peut se réécrire :

$$f(x|c_0, \dots, c_k) = (k+1) \left(\frac{1}{1+x} \right)^{k+2} \chi_{x>0} \quad (2.40)$$

avec le changement de variable :

$$x = \frac{\bar{c} - \bar{c}_{(k)}}{\bar{c}_{(k)}} \quad (2.41)$$

Cette densité est indépendante des valeurs particulières des coefficients. Le paramètre $\bar{c}_{(k)}$ agit donc à la fois comme une origine sur la densité et comme un facteur d'échelle.

Enfin, il est tout à fait imaginable d'obtenir, selon la densité (2.31), une valeur de coefficient beaucoup plus petite que \bar{c} . Il est par contre beaucoup moins probable d'en obtenir cinq. Cela se traduit par l'évolution de la puissance de \bar{c} dans l'expression de la densité conditionnelle. L'estimation de ce paramètre par $\bar{c}_{(k)}$ est, à mesure que k augmente, de plus en plus certaine. Plus le nombre de coefficient connus d'un même processus est important, plus l'ordre de grandeur commun de l'ensemble des termes, qui est présupposé exister par construction du modèle, peut être considéré comme bien estimé.

2.2.2.3 Densité conditionnelle $f(c_n|c_0, \dots, c_k)$

La densité conditionnelle $f(c_n|c_0, \dots, c_k)$ de confiance sur un coefficient inconnu c_n se dérive de manière similaire. Pour tout n supérieur ou égal à $k+1$:

$$f_\epsilon(c_n|c_0, \dots, c_k) = \frac{f_\epsilon(c_0, \dots, c_k, c_n)}{f_\epsilon(c_0, \dots, c_k)} \quad (2.42)$$

Les coefficients jouent tous des rôles symétriques. La formule (2.35) s'applique aussi bien au jeu de coefficients $\{c_0, \dots, c_k, c_n\}$ qu'à celui des $\{c_0, \dots, c_k\}$. Étant donné que le maximum des valeurs absolues de variables c_i peut se réécrire :

$$\max(|c_0|, \dots, |c_k|, |c_n|) = \begin{cases} \max(|c_0|, \dots, |c_k|) & \text{si } |c_n| \leq \max(|c_0|, \dots, |c_k|) \\ |c_n| & \text{si } |c_n| > \max(|c_0|, \dots, |c_k|) \end{cases} \quad (2.43)$$

La densité conditionnelle se réécrit après intégration sur la variable \bar{c} , simplification de l'équation (2.42) et passage à la limite :

$$f(c_n|c_0, \dots, c_k) = \left(\frac{k+1}{k+2} \right) \frac{1}{2\bar{c}_{(k)}} \begin{cases} 1 & \text{si } |c_n| \leq \bar{c}_{(k)} \\ \frac{1}{(|c_n|/\bar{c}_{(k)})^{k+2}} & \text{si } |c_n| > \bar{c}_{(k)} \end{cases} \quad (2.44)$$

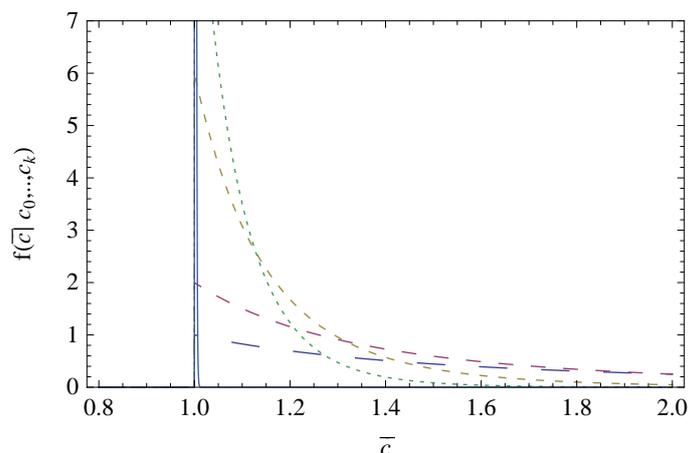


Figure 2.2 – Densité de confiance $f(\bar{c}|c_0, \dots, c_k)$ dans le cas $\bar{c}_{(k)} = 1$ pour, des tirets les plus larges au plus fins, $k = 0, 1, 5, 10$ and 10^3 (courbe continue).

La figure 2.3 présente le comportement de cette densité lorsque k augmente, pour une valeur de $\bar{c}_{(k)}$ toujours fixée à un. Ici aussi, un seul coefficient est suffisant pour obtenir une distribution de confiance non nulle. Lorsque l'entier k tend vers l'infini, la densité tend vers une densité uniforme entre $-\bar{c}_{(k)}$ et $\bar{c}_{(k)}$. Autrement dit, lorsqu'une infinité de coefficients sont connus, la densité de confiance sur les coefficients inconnus prend la forme (2.31) de la densité $f(c_n|\bar{c})$ où \bar{c} est fixé à $\bar{c}_{(k)}$.

Il s'agit du résultat attendu puisque, par construction, l'information maximale partagée par les coefficients est la valeur de \bar{c} . Lorsqu'une infinité de coefficients sont connus, la valeur de \bar{c} est considérée comme certaine, égale à $\bar{c}_{(k)}$. L'incertitude résiduelle est celle modélisée par la densité $f(c_n|\bar{c})$.

Pour une valeur finie de k , la densité de confiance $f(\bar{c}|c_0, \dots, c_k)$ est toujours plus étalée que la densité limite $f(c_n|\bar{c})$. L'incertitude est toujours plus importante lorsque seuls certains coefficients c_0, \dots, c_k sont connus plutôt que lorsque l'information maximale \bar{c} accessible ne l'est.

2.2.2.4 Comportement du modèle en pratique

Lorsque la section efficace d'un processus est évaluée, les coefficients de sa série perturbative sont calculés les uns après les autres. Le premier coefficient n'est pas en général le plus grand. Du moins, il n'y a aucune raison pour qu'il en soit ainsi. Le maximum des coefficients connus $\bar{c}_{(k)}$ change donc de valeur lorsque k évolue.

Par construction et pour le choix de densité (2.31) le paramètre $\bar{c}_{(k)}$ ne peut qu'augmenter. Par exemple, pour des valeurs imaginaires :

$$c_0 = 0,7 \quad c_1 = 0,85 \quad c_2 = -0,2 \quad c_3 = -0,9 \quad (2.45)$$

il prend successivement les valeurs :

$$\bar{c}_{(0)} = 0,7 \quad \bar{c}_{(1)} = 0,85 \quad \bar{c}_{(2)} = 0,85 \quad \bar{c}_{(3)} = 0,9 \quad (2.46)$$

Les densités de confiance $f(c_n|c_0, \dots, c_k)$ pour $k = 0, \dots, 3$ sur un coefficient inconnu c_n , $n \geq 4$, sont présentées sur la figure 2.4. Le plateau que présentent ces densités s'élargit avec

2.3 Discussion du modèle

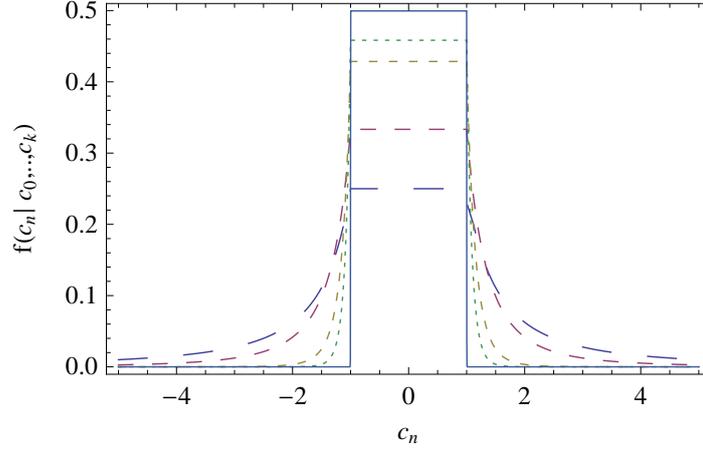


Figure 2.3 – Densité de confiance $f(c_n | c_0, \dots, c_k)$ dans le cas $\bar{c}_{(k)} = 1$ pour, des tirets les plus larges au plus fins, $k = 0, 1, 5, 10$ et 10^3 (courbe continue).

$\bar{c}_{(k)}$. Les densités continuent pourtant à se resserrer lorsque k augmente. En effet, les ailes présentes de part et d'autre de ce plateau prennent en compte la possibilité d'évolution de $\bar{c}_{(k)}$ avec k . Elles se résorbent peu à peu étant donné que $\bar{c}_{(k)}$ est considéré toujours mieux approximer \bar{c} . Sous les hypothèses du modèle, ce paramètre devrait varier plus lentement à mesure que k augmente.

2.3 Discussion du modèle

Des choix de paramètres inconnus à ceux des densités spécifiques utilisées, cette section revient sur les différentes hypothèses du modèle de la section précédente.

2.3.1 Paramètres inconnus

La section efficace totale ainsi que les sommes partielles et les restes de la série perturbative à différents ordres sont des paramètres précisément définis. La séparation entre le paramètre d'expansion et les coefficients de la série est par contre plus subjective. Les prédictions partielles de la section efficace totale du processus $e^+e^- \rightarrow$ hadrons sont par exemple généralement présentées sous la forme de séries tronquées en puissances de $\frac{\alpha_s}{\lambda}$.

De manière générale, le paramètre d'expansion peut être une fraction $\frac{\alpha_s}{\lambda}$ quelconque de α_s . La série perturbative se réécrit :

$$\sum c_n \alpha_s^n = \sum (\lambda^n c_n) \left(\frac{\alpha_s}{\lambda} \right)^n \quad (2.47)$$

Les nouveaux coefficients c'_n sont définis par :

$$c'_n = \lambda^n c_n \quad (2.48)$$

Ce sont ces termes qui sont désormais supposés être de même ordre de grandeur. L'expression de la densité conditionnelle de confiance (2.31) reste valable, les coefficients c'_n

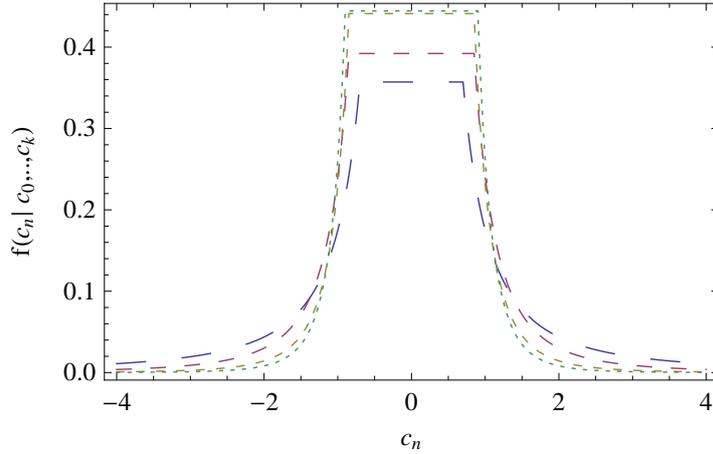


Figure 2.4 – Densité de confiance $f(c_n | c_0, \dots, c_k)$ lorsque $\bar{c}_{(k)}$ varie avec k . Les valeurs de ce paramètre sont données en (2.46).

étant écrits en lieu et place des c_n :

$$f(c'_n | \bar{c}) = \frac{1}{2\bar{c}} \chi_{|c'_n| \leq \bar{c}} \quad (2.49)$$

De la même manière, la mesure $f_\epsilon(\bar{c}, c'_0, c'_1, \dots)$ conserve son expression. Seule la densité conditionnelle $f(\Delta_k | c'_0, \dots, c'_k)$ présentée dans la section 2.4 suivante sera modifiée. En effet, en première approximation, le reste de la série Δ_k prendra l'expression :

$$\Delta_k \simeq c'_{k+1} \left(\frac{\alpha_s}{\lambda} \right)^{k+1} \quad (2.50)$$

Le changement de paramètre d'expansion permet de modifier la vitesse supposée de convergence de la série. Une valeur supérieure à un du paramètre λ suppose une série plus rapidement convergente. Une valeur plus faible que un ($\alpha_s < \lambda < 1$) traduit l'hypothèse d'une convergence plus lente.

2.3.2 Coefficients de même ordre de grandeur

En première approche, les coefficients ont été supposés de même ordre de grandeur. Le modèle peut pourtant s'adapter à d'autres cas. Le paramètre \bar{c} pourrait être modifié d'ordre en ordre pour modéliser la variation des coefficients avec n . Cette dépendance de l'ordre de grandeur des coefficients en fonction de l'ordre perturbatif pourrait provenir par exemple de l'évolution de la taille de l'espace des phases.

Par exemple, chaque coefficient c_n de la série pourrait être supposé, pour des valeurs faibles de n , d'ordre de grandeur $n! \bar{c}$. Le paramètre \bar{c} modéliserait toujours une information partagée par les différents termes de la série perturbative. La densité de confiance $f(c_n | \bar{c})$ serait alors fonction de l'ordre n .

Alternativement, la série perturbative pourrait de nouveau être réécrite, cette fois-ci sous la forme :

$$\sigma = \sum_{n=0}^{\infty} n! c'_n \alpha_s^n \quad (2.51)$$

2.3 Discussion du modèle

les nouveaux coefficients c'_n étant définis par :

$$c'_n = \frac{c_n}{n!} \quad (2.52)$$

De cette manière, les coefficients c'_n seraient encore une fois supposés de même ordre de grandeur et l'étude précédente pourrait leur être appliquée.

2.3.3 Information partagée

2.3.3.1 Informations incompatibles

Une section efficace σ est toujours positive. Cette information n'est pas prise en compte dans le modèle. Il pourrait être tentant de chercher à l'insérer directement dans la définition de la mesure en modifiant la densité de confiance conditionnelle sur la somme totale :

$$\begin{aligned} f_\epsilon(\sigma, \bar{c}, c_0, c_1, \dots) &= f(\sigma | \bar{c}, c_0, c_1, \dots) f_\epsilon(\bar{c}, c_0, c_1, \dots) \\ &= \delta\left(\sigma - \sum_0^\infty c_n \alpha_s^n\right) \chi_{\sigma > 0} \prod_{n=0}^\infty f(c_n | \bar{c}) f_\epsilon(\bar{c}) \end{aligned} \quad (2.53)$$

Mais par voie de conséquence, la densité de la mesure complète $f_\epsilon(\bar{c}, c_0, c_1, \dots)$ serait désormais :

$$\begin{aligned} f_\epsilon(\bar{c}, c_0, c_1, \dots) &= \int f_\epsilon(\sigma, \bar{c}, c_0, c_1, \dots) d\sigma \\ &= \chi_{\sum c_n \alpha_s^n > 0} \prod_{n=0}^\infty f(c_n | \bar{c}) f_\epsilon(\bar{c}) \\ &= \begin{cases} \prod_{n=0}^\infty f(c_n | \bar{c}) f_\epsilon(\bar{c}) & \text{si } \sum c_n \alpha_s^n > 0 \\ 0 & \text{sinon} \end{cases} \end{aligned} \quad (2.54)$$

Cette distribution est identique à la densité calculée précédemment, si ce n'est que son expression est maintenant tronquée. Avec les choix (2.31) et (2.32) de densité, elle n'est pas normalisée. L'ensemble des hypothèses du modèle d'une part et la nouvelle information du signe de la somme σ de l'autre ne sont pas directement compatibles.

Cette incompatibilité peut se comprendre plus facilement en repartant par exemple de l'équation (2.30) :

$$f(\{c_i\}_{i \in I} | \bar{c}) = \prod_{i \in I} f(c_i | \bar{c})$$

Elle implique que les informations sur les valeurs des coefficients sont indépendantes lorsque la valeur du paramètre \bar{c} est connue. En pratique, la valeur de c_0 donne pourtant une limite inférieure à la valeur de c_1 . Intuitivement, le somme $c_0 + \alpha_s c_1$ ne doit pas être trop négative de manière à ce que l'ajout des termes suivants prédissent bel et bien une section totale positive. Lorsque le signe de la somme de la série perturbative est considéré comme connu, les informations sur les valeurs des coefficients ne peuvent pas être supposées indépendantes.

Plus généralement, l'équation d'indépendance (2.30) n'est valide que lorsque la valeur du paramètre \bar{c} est la seule et unique information partagée (indirecte, cf section 2.3.3.4) considérée dans la construction du modèle. Cela ne signifie pas qu'il n'en existe pas en réalité d'autres mais seulement qu'elles ne sont pas prises en compte dans la mesure de confiance utilisée.

2.3.3.2 Premier modèle de prise en charge du signe de la section efficace

Différents modèles prenant en compte le signe de la somme peuvent être créés. En pratique, cette information supplémentaire modifierait essentiellement les densités de confiance $f(c_n|c_0, \dots, c_k)$ sur la demi-droite des réels négatifs. Au pire, lorsque seul le premier coefficient c_0 est connu, les changements s'effectueraient lorsque :

$$c_0 + \alpha_s c_1 < 0 \Leftrightarrow c_1 < -\frac{1}{\alpha_s} c_0 \quad (2.55)$$

Le paramètre $|c_0| = \bar{c}_{(0)}$ estime la valeur de \bar{c} et définit la largeur du plateau central de la densité de confiance sur le coefficient suivant. La prise en compte du signe de la série ne devrait donc modifier cette distribution de confiance qu'à une distance de plus de $10 \bar{c}_{(0)}$ de son extremum. Les conséquences devraient par conséquent être faibles, d'autant plus que l'ordre k du dernier coefficient calculé augmente et que la décroissance des densités de part et d'autre de $\bar{c}_{(k)}$ se fait plus rapide.

À titre d'exemple, le signe de la section efficace pourrait être pris en compte en supposant, en première approximation, que non seulement la somme totale σ mais aussi toutes les sommes partielles σ_k sont positives. Les expressions des densités conditionnelles $f(c_n|\bar{c}, c_0, \dots, c_{n-1})$ suivantes, avec la convention qu'une somme vide est nulle, réécritraient dans ce cas les anciennes hypothèses sur l'indépendance des coefficients, connaissant \bar{c} , et sur la forme de l'incertitude résiduelle :

$$\forall n \geq 0, \quad f_\epsilon(c_n|\bar{c}, c_0, \dots, c_{n-1}) = \mathcal{N} \chi_{|c_n| \leq \bar{c}} \chi_{c_n \geq -(\sum_{i=0}^{n-1} c_i \alpha_s^i)/a_s^n} \quad (2.56)$$

Le terme \mathcal{N} est un facteur de normalisation. L'ensemble des distributions en ϵ non nul $f_\epsilon(\bar{c}, c_0, \dots, c_n)$, définissant la mesure complète, se déduirait alors par récurrence :

$$f_\epsilon(\bar{c}, c_0, \dots, c_n) = f_\epsilon(c_n|\bar{c}, c_0, \dots, c_{n-1}) f_\epsilon(\bar{c}, c_0, \dots, c_{n-1}) \quad (2.57)$$

La figure 2.5 présente les densités de confiance $f(c_1|c_0)$ pour les deux modèles, avec ou sans la condition de positivité des sommes partielles. Les densités obtenues sont bien visuellement identiques autour de leur valeur extrémale. L'ajout de la nouvelle information n'est donc pas critique alors qu'elle complexifie les expressions des densités du modèle.

2.3.3.3 Deuxième modèle de prise en charge du signe de la section efficace

Un autre modèle, plus proche des manipulations habituelles des "probabilités bayésiennes" est possible. On introduit une variable inconnue booléenne sur le signe de la section efficace :

$$H : \sigma \in \mathbb{R}^+ \quad (2.58)$$

Les raisonnements de la section 2.3.3.1 montrent qu'il n'est pas consistant de supposer :

$$f(\{c_i\}_{i \in I}|\bar{c}, H) = \prod_{i \in I} f(c_i|\bar{c}, H) \quad (2.59)$$

Les informations sont incompatibles. Il est pourtant toujours possible de construire le modèle comme avant suivant l'hypothèse :

$$f(\{c_i\}_{i \in I}|\bar{c}) = \prod_{i \in I} f(c_i|\bar{c}) \quad (2.60)$$

2.3 Discussion du modèle

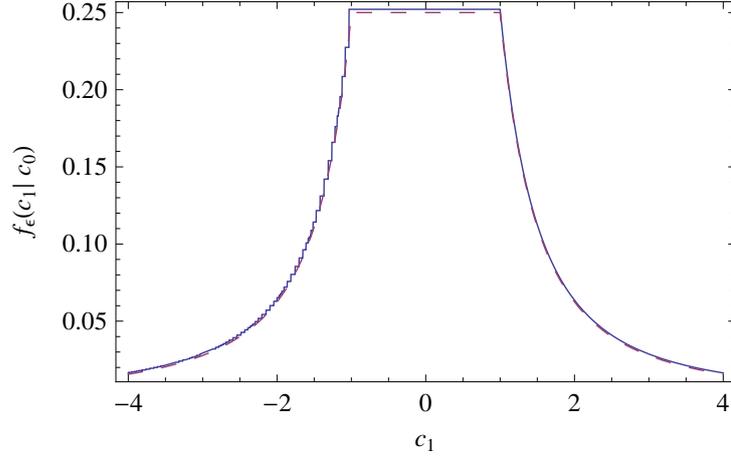


Figure 2.5 – Densités de confiance $f(c_1|c_0 = 1)$ obtenues dans les modèles imposant le signe des somme partielles (en trait continu, pour $\epsilon = 10^{-2}$) ou non (en pointillés).

et de n'utiliser la variable H qu'au tout dernier moment. Plus précisément, une fois la densité $f(\sigma|c_0, \dots, c_k)$ obtenue, il est possible d'ajouter la condition de signe par :

$$f(\sigma|c_0, \dots, c_k, H) = \frac{\mathbb{C}(H|\sigma, c_0, \dots, c_k)}{\mathbb{C}(H|c_0, \dots, c_k)} f(\sigma|c_0, \dots, c_k) \quad (2.61)$$

où l'on définit :

$$\mathbb{C}(H|\sigma, \dots) = \mathbb{C}(H|\sigma) = \delta_{\sigma \geq 0} \quad (2.62)$$

La mesure complète est en fait alors définie par :

$$f(\bar{c}, c_0, c_1, \dots, H = v) = \begin{cases} f(\bar{c}, c_0, c_1, \dots) & \text{si } \sum_0^\infty c_i \alpha_s^i \geq 0 \\ 0 & \text{si } \sum_0^\infty c_i \alpha_s^i < 0 \end{cases} \quad (2.63)$$

$$f(\bar{c}, c_0, c_1, \dots, H = f) = \begin{cases} 0 & \text{si } \sum_0^\infty c_i \alpha_s^i \geq 0 \\ f(\bar{c}, c_0, c_1, \dots) & \text{si } \sum_0^\infty c_i \alpha_s^i < 0 \end{cases}$$

L'hypothèse d'indépendance des informations sur les coefficients (2.60) est ici effectuée sur les densités marginalisées par rapport à la variable H . Les densités conditionnelles $f(c_i|\bar{c}, H)$ ne sont par contre bien entendu pas indépendantes.

2.3.3.4 Liens directs et indirects

Les informations sur l'ordre de grandeur commun \bar{c} et le signe positif de la section semblent donc devoir être traités de manières très différentes. La première hypothèse doit être intégrée directement dans la définition de la mesure alors que la seconde doit être utilisée *a posteriori* en restreignant l'espace d'étude au sous-espace vérifiant $\sigma \geq 0$.

La relation d'indépendance (2.60) de l'incertitude résiduelle est équivalente à l'hypothèse :

$$f(c_n|\bar{c}, \{c_i\}_{i \in I}) = f(c_n|\bar{c}) \quad \forall I, \forall n \notin I \quad (2.64)$$

Autrement dit, la connaissance de coefficients c_i ne peut apporter de l'information sur la valeur de c_n que par l'intermédiaire de \bar{c} . Lorsque ce dernier paramètre est totalement

connu, connaître de nouveaux coefficients ne peut rien ajouter. Le lien créé par l'ordre de grandeur commun \bar{c} entre les c_i est indirect. L'information transite par la valeur de \bar{c} .

Au contraire, la valeur du signe de la section efficace ne fait que relier les coefficients. Les valeurs des uns influencent toujours celles des autres. Le lien créé est direct. Étant connu le signe de la section, les valeurs considérées comme possibles pour le deuxième coefficient sont directement en lien avec la valeur du premier :

$$f(c_1|H, c_0 = 0, 5) \neq f(c_1|H, c_0 = 3) \quad (2.65)$$

On s'attend donc de manière générale à ce que :

$$f(c_1|H, c_0) \neq f(c_1|H) \quad (2.66)$$

Il semble que les liens directs et indirects ne doivent pas être pris en compte de la même manière dans la création d'un modèle de confiance. Les liens indirects sont ajoutés à la construction initiale sous la forme de relations d'indépendance. Supposons par exemple qu'il existe un deuxième lien indirect transitant par la connaissance de la valeur d'un second paramètre \bar{d} . Le modèle devrait être formé de manière à ce que, si les informations sont compatibles :

$$f(\{c_i\}|\bar{c}, \bar{d}) = \prod f(c_i|\bar{c}, \bar{d}) \quad (2.67)$$

Les liens directs semblent par contre devoir être ajoutés "dans un deuxième temps", à la méthode bayésienne. Plus précisément, ils semblent devoir être ajoutés en découpant l'espace de travail construit dans la première étape à la manière des équations (2.63).

2.3.4 Incertitude résiduelle

2.3.4.1 Confiance et certitude

La densité $f(c_n|\bar{c})$ définie par la formule (2.31) s'annule en dehors de $[-\bar{c}; \bar{c}]$. Elle traduit donc une certitude : connaissant la valeur du paramètre caché, il est "certain" que les coefficients inconnus sont plus petits en valeur absolue que \bar{c} .

Un modèle de confiance traduit des hypothèses. Ces hypothèses ne sont généralement ni vraies, ni totalement fausses. Elles relèvent de la croyance. Il est donc dangereux de chercher à traduire une certitude. En effet, il est toujours possible de se tromper en privilégiant une valeur possible d'un paramètre plutôt qu'une autre. Il est alors important de pouvoir changer d'avis lorsque suffisamment de preuves de l'erreur sont apportées. Les certitudes ne permettent pas ce revirement. Elles sont à banir ou, tout du moins, à manipuler avec grande prudence.

Numériquement, une fois que la densité de confiance sur un paramètre s'annule pour une valeur donnée, elle ne peut plus être modifiée. En effet, une nouvelle information "actualise" la mesure de confiance suivant la formule de Bayes sur les densités conditionnelles. Dans le modèle présent :

$$f(c_n|c_0, c_1) = f(c_n|c_0) \frac{f(c_1|c_0, c_n)}{f(c_1|c_0)} \quad (2.68)$$

Si la densité $f(c_n|c_0)$ s'annulait pour une valeur particulière de c_n , la densité actualisée $f(c_n|c_0, c_1)$ s'y annulerait également. Plus généralement, cette valeur serait toujours

2.3 Discussion du modèle

considérée comme impossible quelles que soient les éventuelles nouvelles informations qui pourraient être considérées.

L'annulation de la densité $f(c_n|\bar{c})$ n'est pas aussi critique. Cette densité représente la meilleure confiance que l'on puisse avoir sur la valeur du coefficient inconnu c_n . Dans le cadre du modèle, il est précisément impossible d'ajouter une nouvelle information à celle de la donnée cette valeur : il s'agit de l'information maximale disponible. De plus, la densité $f(c_n|\bar{c})$ est une limite qui n'est jamais atteinte. Les informations fournies sont les valeurs des coefficients calculés c_0, \dots, c_k . Malgré le choix de densité uniforme (2.31) pour $f(c_n|\bar{c})$, aucune des densités de confiance $f(c_n|c_0, \dots, c_k)$ ne s'annule sur la droite des réels.

2.3.4.2 Poids des coefficients calculés

La densité uniforme (2.31) présente un autre aspect critiquable. Elle mène à l'expression (2.44) de la densité de confiance $f(c_n|c_0, \dots, c_k)$ sur un coefficient inconnu c_n connaissant les coefficients c_0, \dots, c_k . Cette densité est fonction des valeurs des coefficients calculés au travers du maximum de leur valeur absolue $\bar{c}_{(k)}$.

Imaginons que cinq coefficients aient été calculés. Supposons que tous les coefficients soient du même ordre de grandeur. Pour fixer les idées, on les suppose de l'ordre de l'unité. le coefficient $\bar{c}_{(k)}$ est également de l'ordre de l'unité. Supposons que l'on calcule le sixième coefficient et qu'il soit de l'ordre de 10. Le paramètre $\bar{c}_{(k)}$ prend sa valeur. Supposons encore que l'on calcule trois coefficients de plus, tous de l'ordre de l'unité. Le paramètre $\bar{c}_{(k)}$ ne changerait pas et resterait de l'ordre de 10. Quand bien même tous les coefficients calculés par la suite seraient de l'ordre de l'unité, l'estimation de \bar{c} ne changerait pas. L'incertitude sur les coefficients calculés serait toujours dictée par le plus grand coefficient calculé.

Le "poids" du nombre de coefficients présents dans une zone n'est pas pris en compte. La densité uniforme (2.31) implique que l'information donnée par la valeur d'un coefficient est en quelque sorte binaire : *le coefficient c_n est plus petit que \bar{c}* est une proposition vraie tandis que *le coefficient est plus grand que \bar{c}* est faux. Il n'y a pas de notion de distance possible entre c_n et \bar{c} . L'information de la répartition particulière des coefficients calculés n'est pas utilisée.

De plus, la densité uniforme (2.31) s'annule en dehors de l'intervalle $[-\bar{c}; \bar{c}]$. Elle ne modélise pas le fait que l'on puisse s'attendre à un "accident", une valeur de coefficient anormalement élevée. Inversement, lorsque certains coefficients ont été calculés, la densité $f(c_n|\bar{c})$ ne permet pas d'interpréter la valeur anormalement élevée de l'un d'entre eux comme un événement rare, isolé. Tous les coefficients calculés sont utilisés de manière équivalente dans l'évaluation de la valeur de \bar{c} .

Imaginons une densité uniforme sur $[-\bar{c}; \bar{c}]$, rapidement décroissante de part et d'autre de cet intervalle et ne s'annulant jamais sur \mathbb{R} . Lorsque peu de coefficients auront été calculés, ceux-ci seront traités de manière similaire. Tout se passera comme si la densité $f(c_n|\bar{c})$ était la densité uniforme. Si un très grand coefficient est apparu, il dirigera l'estimation de \bar{c} . Pourtant, à mesure que le nombre de coefficients connus augmente, le coefficient anormalement grand sera de moins en moins utilisé dans la prédiction de l'ordre de grandeur commun. Celle-ci diminuera. Le grand coefficient sera bien considéré cette fois-ci comme un événement marginal.

Contrairement à ce que pourrait laisser penser les remarques du paragraphe 2.3.4.1

précédent, le choix de densité $f(c_n|\bar{c})$ uniforme apparaît donc comme une modélisation conservatrice de l'incertitude lorsque de nombreux coefficients ont été calculés. Concrètement, la différence avec le cas d'une densité qui ne s'annule pas sur la droite des réels ne devrait être importante que lorsque le nombre de coefficients calculés est élevé et que l'un d'eux ne partage pas l'ordre de grandeur commun. Pour des ordres plus bas, les valeurs exactes des limites des intervalles de valeur de confiance fixée devrait légèrement varier mais le comportement global du modèle devrait être identique.

2.3.4.3 Choix de variable privilégiée

Le densité uniforme (2.31) est définie sur le paramètre c_n . Elle traduit l'hypothèse qu'aucune valeur possible de c_n n'est privilégiée entre $-\bar{c}$ et \bar{c} . En d'autre terme, il y a "autant de chance" que la valeur de c_n se situe entre une certaine valeur x et $x + \delta x$ qu'entre x et $x - \delta x$, si tant est que l'intervalle $[x - \delta x; x + \delta x]$ soit inclus dans $[-\bar{c}; \bar{c}]$.

Il est possible de choisir alternativement de ne privilégier aucun ordre de grandeur du paramètre c_n . La densité $f(c_n|\bar{c})$ est alors uniforme en $\ln c_n$. Il est maintenant supposé qu'il y a "autant de chance" que c_n soit compris entre x et $x/\delta h$ qu'entre x et $x.\delta h$. De manière plus générale, la densité peut être définie comme uniforme sur une fonction quelconque de c_n . Cette expression particulière de c_n est alors privilégiée.

En pratique, seuls les deux cas de variables privilégiées c_n ou $\ln c_n$ ont été étudiés. La question du meilleur choix peut se poser. Un modèle créé sur une densité uniforme en $\ln c_n$ est plus complexe. Une densité uniforme sur c_n est naturellement centrée sur zéro. Un seul paramètre \bar{c} , contrôlant la largeur de la densité, est nécessaire. Une densité uniforme sur $\ln c_n$ est fixée par sa largeur $2 \ln h$ et la position $\ln \bar{c}$ de son centre : les coefficients sont compris entre \bar{c}/h et $\bar{c}.h$. Deux paramètres sont maintenant nécessaires. Par voie de conséquence, Il s'avère que ce modèle requiert la connaissance de plus d'un coefficient pour fournir des densités non nulles.

Un jeu de densités définissant le modèle construit sur le choix de variable privilégié $\ln c_n$ est par exemple :

$$f(c_n|\bar{c}, h) = \frac{1}{2 \ln h} \frac{1}{2|c_n|} \chi_{\bar{c}/h \leq |c_n| \leq \bar{c}.h} \quad (2.69)$$

$$f_{\epsilon, \epsilon'}(\bar{c}, h) = f_{\epsilon}(\bar{c}) f_{\epsilon'}(h) \quad (2.70)$$

$$f_{\epsilon}(\bar{c}) = \frac{1}{2|\ln \epsilon|} \frac{1}{\bar{c}} \chi_{\epsilon \leq \bar{c} \leq \frac{1}{\epsilon}} \quad (2.71)$$

$$f_{\epsilon'}(h) = \frac{1}{2|\ln \epsilon'|} \frac{1}{h} \chi_{\epsilon' \leq h \leq \frac{1}{\epsilon'}} \quad (2.72)$$

La condition d'indépendance (2.30), écrite maintenant pour \bar{c} et h connus, complète la mesure :

$$f(\{c_i, i \in I\}|\bar{c}, h) = \prod_I f(c_i|\bar{c}, h) \quad (2.73)$$

La densité de confiance générale a pour expression :

$$f_{\epsilon, \epsilon'}(c_0, \dots, c_n, \bar{c}, h) = \prod_{i=0}^n \left(\frac{1}{2 \ln h} \frac{1}{2|c_n|} \chi_{\bar{c}/h \leq |c_n| \leq \bar{c}.h} \right) \frac{1}{2|\ln \epsilon|} \frac{1}{\bar{c}} \chi_{\epsilon \leq \bar{c} \leq \frac{1}{\epsilon}} \frac{1}{2|\ln \epsilon'|} \frac{1}{h} \chi_{\epsilon' \leq h \leq \frac{1}{\epsilon'}} \quad (2.74)$$

2.3 Discussion du modèle

et la densité conditionnelle sur le couple $(\ln \bar{c}, \ln h)$ s'écrit par exemple, dès que k est supérieur ou égal à deux (c'est à dire pour au moins trois coefficients calculés) :

$$f(\ln \bar{c}, \ln h | c_0, \dots, c_k) = \frac{k-1}{k} \left[\frac{\ln h_{(k,+)}(\bar{c}) + \ln h_{(k,-)}(\bar{c})}{2 \ln h} \right]^{k+1} \frac{2 \chi_{h \geq \max\{h_{(k,\pm)}(\bar{c})\}}}{(\ln h_{(k,+)}(\bar{c}) + \ln h_{(k,-)}(\bar{c}))^2} \quad (2.75)$$

Les fonctions $h_{(k,\pm)}(\bar{c})$ estiment, pour une valeur centrale \bar{c} possible de l'ordre de grandeur des coefficients, l'étalement des coefficients à l'aide des plus grande et plus petite valeurs connues. Elles sont définies par :

$$\bar{c}/h_{(k,-)}(\bar{c}) = \min\{c_0, \dots, c_k\} \quad \bar{c} \cdot h_{(k,+)}(\bar{c}) = \max\{c_0, \dots, c_k\} \quad (2.76)$$

Le comportement de la densité est bien celui que l'on pouvait attendre. Elle est maximale pour les valeurs de (\bar{c}, h) où les coefficients sont répartis de part et d'autre de \bar{c} de manière à ce que les plus petit et plus grand termes soient précisément en \bar{c}/h et $\bar{c} \cdot h$. Encore une fois, l'utilisation de densités initiales uniformes crée une densité conditionnelle sur (\bar{c}, h) qui ne prend pas en compte le "poids" du nombre de coefficients dans une zone : la densité ci-dessus ne traduit pas le fait qu'il puisse y avoir plus de coefficients entre $\bar{c}/h_{(k,-)}$ et \bar{c} qu'entre \bar{c} et $\bar{c} \cdot h_{(k,+)}$.

Mais au delà du bon fonctionnement de la mesure ou de toute discussion sur le choix le plus approprié de variable privilégiée, ce modèle est inutilisable en pratique : il requiert plus de trois ordres connus. Il ne traduit donc pas la confiance généralement attribuée aux calculs effectués : la section de Born $\sigma_0 = c_0$ seule est habituellement considérée fournir une première estimation de la section totale.

2.3.5 Nature du paramètre caché

Le paramètre \bar{c} représente l'information cachée. Elle est intégrée au modèle à l'aide de la densité $f_\epsilon(\bar{c})$. De la même manière que dans le cas du problème du choix de variable privilégiée en c_n , le choix (2.32) traduit un a priori selon lequel les différents ordres de grandeur de \bar{c} seraient également envisageables en l'absence d'information sur la série perturbative. La densité est uniforme pour la variable $\ln \bar{c}$.

Il est encore une fois possible de privilégier une autre fonction de \bar{c} , telle que \bar{c} elle-même. Dans ce cas, ce sont les valeurs de \bar{c} qui sont supposées également envisageables. Du point de vue de la valeur de \bar{c} , le choix de densité uniforme sur son logarithme privilégie les petites valeurs du paramètre. On considère qu'il y a autant de chance de trouver \bar{c} entre 10^{-2} et 1 qu'entre 1 et 10^2 . À l'opposé, du point de vue de l'ordre de grandeur de \bar{c} , le choix de densité uniforme en \bar{c} privilégie les grands ordres de grandeur. On suppose alors qu'il y a autant de chance de trouver $\ln \bar{c}$ entre $\ln(1) = 0$ et $\ln(1,9) \simeq 0,6$ qu'entre $\ln(0,1) \simeq -2,3$ et $\ln(1) = 0$ alors que le premier intervalle est, du point de vue de $\ln \bar{c}$, beaucoup plus petit que le second.

Le choix de variable privilégiée ne peut venir que de la nature du paramètre étudié. Le paramètre \bar{c} est directement lié à la somme totale de la section. Il définit l'ordre de grandeur considéré comme possible des coefficients avant calcul. Il dicte donc l'ordre de grandeur considéré comme envisageable pour la somme de la série perturbative. Effectuer un choix de variable privilégiée sur \bar{c} revient en fait à faire un choix d'incertitude initiale sur la somme σ . A priori, avant tout calcul, il serait naturel de supposer que tous les ordres

de grandeurs de la section efficace sont équivalents. Une fois le calcul effectué, on considère d'ailleurs avoir la même précision si une section efficace est de l'ordre de nanobarn, connue à un pour cent près, que si elle est du picobarn, toujours connue à un pour cent près. On privilégie donc $\ln \bar{c}$.

D'un point de vue technique, si l'on privilégie une fonction \bar{c}^i , $i > 0$, de \bar{c} , les densités de confiance conditionnelles $f(c_n|c_0, \dots, c_k)$ et $f(\bar{c}|c_0, \dots, c_k)$ sont non uniformément nulles uniquement lorsque le nombre de coefficients calculés $k + 1$ est strictement plus grand que la puissance i . En effet, dans le cas contraire, les facteurs $\frac{1}{\bar{c}}$ de l'expression de la mesure complète (2.35) ne sont plus en nombre suffisant pour faire converger la densité (2.37) marginalisée en \bar{c} , densité qui correspond à la normalisation de la distribution conditionnelle $f(\bar{c}|c_0, \dots, c_k)$ (voir (2.36)). Comme précisé au paragraphe précédent, non seulement les mesures ne sont rapidement plus utilisables en pratique lorsque i augmente mais aucune d'elles ne permet plus non plus d'associer à la seule valeur de la section de Born une incertitude.

2.4 Estimation de l'incertitude de troncature

La valeur du reste Δ_k de la série perturbative correspond à l'erreur effectuée lorsque la somme totale est approchée par la somme partielle σ_k . Estimer son ordre de grandeur correspond donc à estimer l'incertitude de l'approximation.

Le reste, défini en (2.5) est une fonction des coefficients non calculés c_n . Sa densité de confiance s'écrit de manière générale :

$$f(\Delta_k|c_0, \dots, c_k) = \int \left[\delta\left(\Delta_k - \sum_{k+1}^{\infty} \alpha_s^n c_n\right) \right] f(c_{k+1}, c_{k+2}, \dots | c_0, \dots, c_k) dc_{k+1} dc_{k+2} \dots \quad (2.77)$$

Cette expression n'est pas utilisable en pratique car l'intégrale ne peut être réellement effectuée sur une infinité de coefficients. Il est nécessaire d'approximer la formule.

Cette section présente dans un premier temps la densité obtenue sur le reste dans l'approximation la plus simple où un seul terme de la définition (2.5) est conservé. Elle justifie ensuite le résultat obtenu en le comparant avec les distributions obtenues dans le cas d'approximations plus précises. Un dernier paragraphe présente enfin la forme de la densité de confiance lorsque la loi de confiance asymptotique sur les coefficients, *i.e.* la loi conditionnelle connaissant \bar{c} , n'est plus une loi uniforme mais une loi normale.

2.4.1 Densité conditionnelle $f(\Delta_k|c_0, \dots, c_k)$ approchée

Le paramètre α_s est généralement petit devant un. En première approximation, on peut supposer que le reste de la série perturbative à l'ordre k s'écrit :

$$\Delta_k = \sum_{k+1}^{\infty} \alpha_s^n c_n \simeq \alpha_s^{k+1} c_{k+1} \quad (2.78)$$

2.4 Estimation de l'incertitude de troncature

L'expression de la densité $f(\Delta_k|c_0, \dots, c_k)$ s'obtient alors à partir de l'équation (2.44) par changement de variable :

$$f(\Delta_k|c_0, \dots, c_k) \simeq \left(\frac{k+1}{k+2}\right) \frac{1}{2\alpha_s^{k+1}\bar{c}_{(k)}} \begin{cases} 1 & \text{si } |\Delta_k| \leq \alpha_s^{k+1}\bar{c}_{(k)} \\ \frac{1}{(|\Delta_k|/\alpha_s^{k+1}\bar{c}_{(k)})^{k+2}} & \text{si } |\Delta_k| > \alpha_s^{k+1}\bar{c}_{(k)} \end{cases} \quad (2.79)$$

La densité de confiance sur Δ_k dépend de la valeur des coefficients calculés au travers du paramètre $\bar{c}_{(k)}$.

L'équation (2.79) permet d'obtenir la forme exacte des plus petits intervalles de confiance $p\%$ sur Δ_k . Ce sont des intervalles centrés $[-d_k^{(p)}; d_k^{(p)}]$ où :

$$d_k^{(p)} = \alpha_s^{k+1}\bar{c}_{(k)} \begin{cases} [(k+2)(1-p\%)]^{-1/(k+1)} & \text{si } p\% > \frac{k+1}{k+2} \\ p\% \frac{k+2}{k+1} & \text{si } p\% \leq \frac{k+1}{k+2} \end{cases} \quad (2.80)$$

En effet, la densité est symétrique, continue, et décroissante de part et d'autre de zéro. Les intervalles minimaux sont donc centrés. Pour p suffisamment grand, $d_k^{(p)}$ sera supérieur à $\alpha_s^{k+1}\bar{c}_{(k)}$. La confiance à l'extérieur de l'intervalle est de $1-p\%$, la densité est symétrique et son expression sur $[d_k^{(p)}; \infty[$ est unique :

$$\frac{1-p\%}{2} = \int_{d_k^{(p)}}^{\infty} \left(\frac{k+1}{k+2}\right) \frac{1}{(\Delta_k/\alpha_s^{k+1}\bar{c}_{(k)})^{k+2}} \frac{d\Delta_k}{2\alpha_s^{k+1}\bar{c}_{(k)}} \quad (2.81)$$

La première partie de l'expression (2.80) se déduit de cette expression par intégration. Celle-ci est valide jusqu'à ce que :

$$d_k^{(p)} = \alpha_s^{k+1}\bar{c}_{(k)} \quad (2.82)$$

soit, en utilisant l'expression précédente :

$$p\% = \frac{k+1}{k+2} \quad (2.83)$$

Pour une valeur de p inférieure, la densité de confiance (2.79) est constante sur l'intervalle $[-d_k^{(p)}; d_k^{(p)}]$. On obtient l'expression de $d_k^{(p)}$ en écrivant que la confiance de l'intervalle est $p\%$.

Les parties gauche des figures 2.6 et 2.7 présentent les densités de confiance $f(\sigma|c_0, \dots, c_k)$ dans le cas $e^+e^- \rightarrow$ hadrons. Il s'agit des densités de confiance sur le reste Δ_k translatées de σ_k . Les barres horizontales correspondent aux intervalles dont la confiance est de 68,3% et 95,5%. Ces derniers sont reportés sur les graphiques de droite pour faciliter la comparaison entre les ordres.

Les deux figures 2.6 et 2.7 correspondent aux choix (2.18) et (2.19) de série perturbative $\sigma(Q)$. Dans le premier cas, $c_0 = 1$ est considéré comme un coefficient de la série. Cela suppose qu'il apporte une information sur l'ordre de grandeur des coefficients d'ordre supérieur. Dans l'autre, la série ne commence qu'à l'ordre $n = 1$. Une modification de la définition des coefficients c_n permet d'appliquer tous les résultats précédents. L'hypothèse selon laquelle tous les coefficients c_n sont du même ordre de grandeur s'applique tout aussi bien aux coefficients c'_n définis par :

$$c'_n \equiv \alpha_s c_{n+1} \quad (2.84)$$

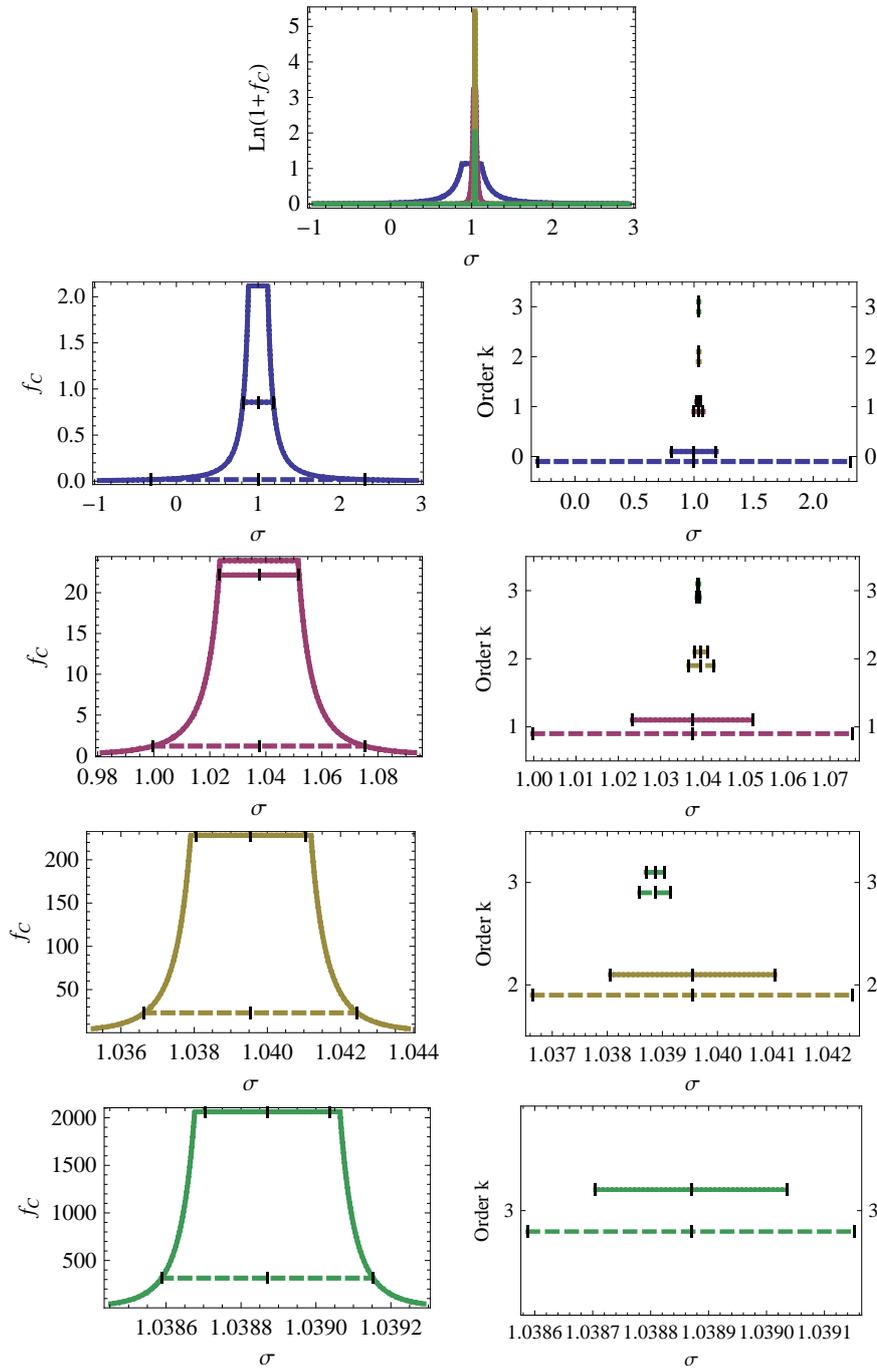


Figure 2.6 – À gauche, densités de confiance $f(\sigma|c_0, \dots, c_k)$, la somme σ étant définie par (2.18), pour k variant, de haut en bas, de 0 à 4. La densité $f(c_n|\bar{c})$ du modèle est uniforme. Les intervalles tracés sont de confiance 68, 3% (traits continus) et 95, 5% (pointillés). Ces intervalles sont reportés à droite, à chaque fois pour l'ordre k considéré et tous les ordres supérieurs. En haut, centré, l'ensemble des fonctions $\ln(1 + f)$ des densités à tous les ordres considérés.

2.4 Estimation de l'incertitude de troncature

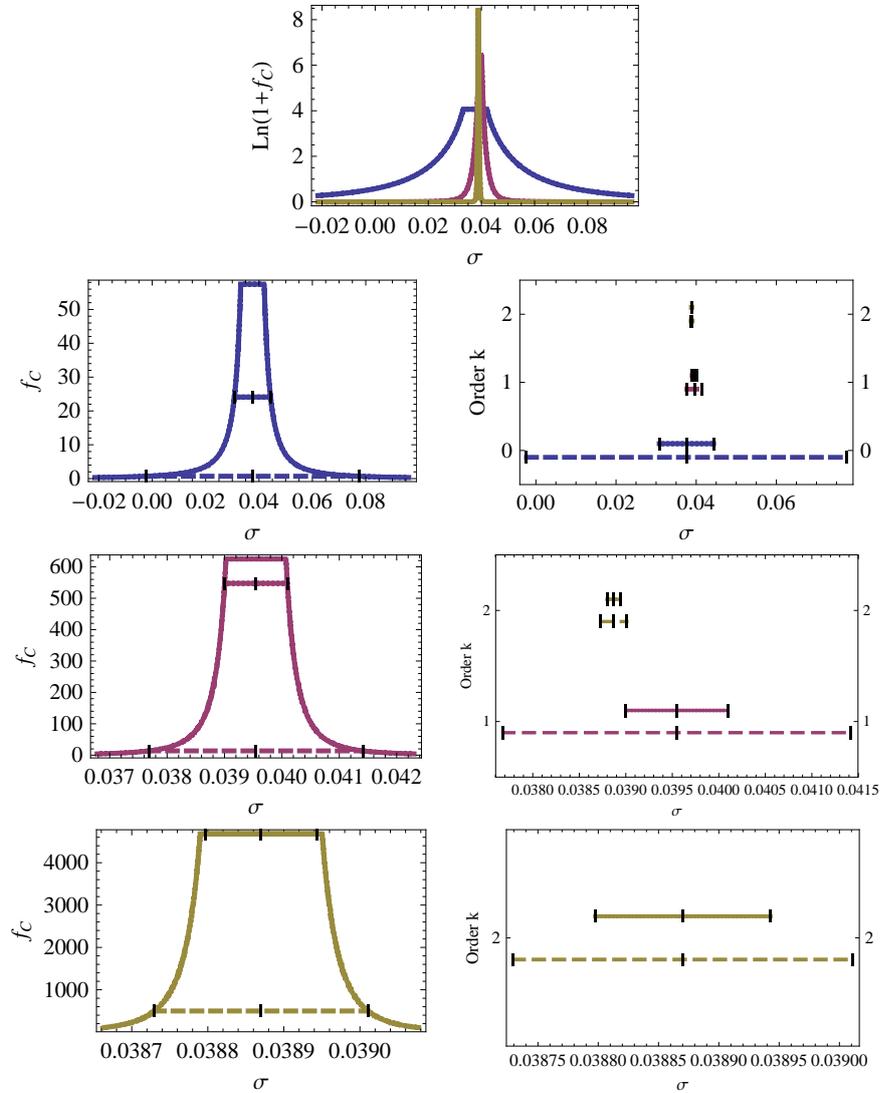


Figure 2.7 – À gauche, densités de confiance $f(\sigma|c'_0, \dots, c'_k)$, la somme σ étant définie par (2.19), pour k variant, de haut en bas, de 0 à 3. La densité $f(c'_n|\bar{c})$ du modèle est uniforme. Les intervalles tracés sont de confiance 68,3% (traits continus) et 95,5% (pointillés). Ces intervalles sont reportés à droite, à chaque fois pour l'ordre k considéré et tous les ordres supérieurs. En haut, centré, l'ensemble des fonctions $\ln(1+f)$ des densités à tous les ordres considérés.

La série perturbative $\sigma(Q)$ s'écrit en fonction des c'_n à partir de l'ordre $n = 0$:

$$\sigma(Q) = \sum_{n=1}^{\infty} (\alpha_s c_n) \alpha_s^{n-1} = \sum_{n=0}^{\infty} c'_n \alpha_s^n \quad (2.85)$$

Ce changement d'indice est possible car le modèle n'utilise que la valeur des coefficients en $\mu = Q$ fixé, pas leur variation.

Les deux choix de définition (2.18) et (2.19) peuvent être défendus. D'une part, les effets de l'interaction forte n'interviennent qu'à partir de l'ordre α_s . Le coefficient c_0 et les coefficients supérieurs ne font pas intervenir la même physique. D'un autre côté, les effets de l'interaction forte devraient tout de même constituer une correction de l'ordre 0. Les coefficients des puissances de α_s ne devraient pas être différents de la section de Born de plusieurs ordres de grandeur.

Dans le doute, l'ajout d'un coefficient ne peut que faire augmenter les valeurs des paramètres $\bar{c}_{(k)}$ aux différents ordres : le choix conservatif est de prendre en compte un nombre potentiellement trop grand de coefficient plutôt que de risquer de ne pas en prendre assez. La différence de puissances provenant du décalage des ordres est en effet rapidement négligeable. Dans le cas de la section efficace de production d'hadrons en collisions électrons-positrons, le coefficient c_0 est plus grand que tous les suivants. Il élargit les intervalles d'incertitudes.

2.4.2 Validité de l'approximation

De manière générale, l'approximation (2.78) peut être remplacée par différentes autres plus précises, d'expression :

$$\Delta_k = \sum_{k+1}^{\infty} \alpha_s^n c_n \simeq \sum_{k+1}^{k+N} \alpha_s^n c_n \quad (2.86)$$

N est le nombre d'ordres supérieurs désormais pris en compte. Le cas (2.78) précédent correspond en fait au choix $N = 1$. L'approximation (2.86) peut être substituée à la définition de Δ_k dans l'expression (2.77) de la densité conditionnelle. Les coefficients d'ordre supérieurs à $k+N+1$ s'intègrent et donnent un facteur 1. La densité conditionnelle de confiance sur Δ_k s'écrit alors :

$$f(\Delta_k | c_0, \dots, c_k) \simeq \int \left[\delta\left(\Delta_k - \sum_{k+1}^{k+N} \alpha_s^n c_n\right) f(c_{k+1}, \dots, c_{k+N} | c_0, \dots, c_k) \right] dc_{k+1} \dots dc_{k+N} \quad (2.87)$$

Cette nouvelle expression permet de comparer les estimations de la densité conditionnelle $f(\Delta_k | c_0, \dots, c_k)$ lorsque les incertitudes provenant de coefficients inconnus de plus en plus nombreux sont pris en compte.

La figure 2.8 présente les approximations (2.87) de la densité $f(\Delta_k | c_0, \dots, c_k)$ dans les cas $N = 1$, ce qui correspond à l'expression (2.79), et $N = 10$. Pour $\alpha_s = 0,12$, l'approximation (2.78) est suffisante. Ces choix donnent des résultats visuellement identiques. L'approximation se dégrade pour des valeurs plus importantes de la constante de couplage, comme $\alpha_s = 0,5$ (seconde colonne).

2.4 Estimation de l'incertitude de troncature

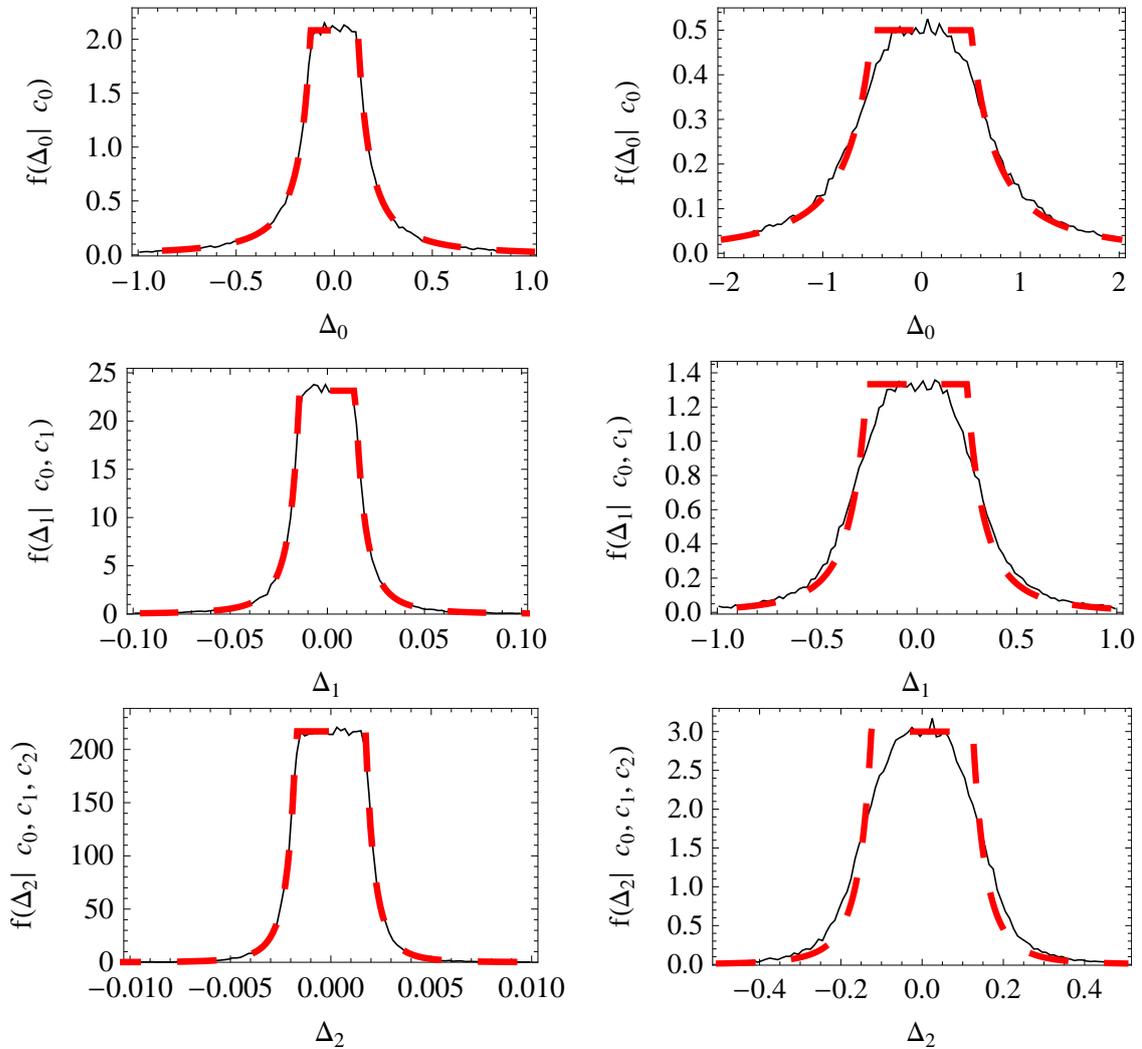


Figure 2.8 – Approximations numériques des densités de confiance $f(\Delta_k | c_0, \dots, c_k)$ pour $N = 10$ (traits pleins) et $N = 1$ (pointillés) dans le cas $\bar{c}_{(k)} = 1$ pour $k = 0$ (haut), $k = 1$ (milieu), et $k = 2$ (bas), pour $\alpha_s = 0,12$ (gauche) et $\alpha_s = 0,5$ (droite). Les densités de confiance (2.87) pour $N = 10$ ont été estimées à partir de 10^5 échantillons distribués sur des intervalles de largeur respective $\delta\Delta_0 = 2 \times 10^{-2}$, $\delta\Delta_1 = 2 \times 10^{-3}$ et $\delta\Delta_2 = 2 \times 10^{-4}$. La densité de confiance pour $N = 1$ a été calculée à l'aide de l'expression (2.79).

2.4.3 Évaluation numérique de $f(\Delta_k|c_0, \dots, c_k)$

Dans le cas $N = 10$, l'intégrale (2.87) est évaluée numériquement en ϵ non nul. Les mesures de confiance et de probabilité partagent le même formalisme. La densité conditionnelle peut être reconstruite sous forme d'histogramme, comme si elle était obtenue par répétition d'une expérience.

Plus précisément, la droite des réels est découpée en segments $[x_i; x_{i+1}]$ de même largeur $\delta\Delta_k$. Sur l'un de ces intervalles, en supposant le pas $\delta\Delta_k$ suffisamment petit :

$$f_\epsilon(\Delta_k = x_i|c_0, \dots, c_k) \delta\Delta_k \simeq \int_{x_i}^{x_{i+1}} f_\epsilon(\Delta_k|c_0, \dots, c_k) d\Delta_k \quad (2.88)$$

Étant donné que :

$$\int_A \delta(x - y) dy = \chi_{x \in A} \quad (2.89)$$

et que :

$$\begin{aligned} f_\epsilon(c_{k+1}, \dots, c_{k+N}|c_0, \dots, c_k) &= \int f_\epsilon(c_{k+1}, \dots, c_{k+N}|c_0, \dots, c_k, \bar{c}) f_\epsilon(\bar{c}|c_0, \dots, c_k) d\bar{c} \\ &= \int f_\epsilon(c_{k+1}, \dots, c_{k+N}|\bar{c}) f_\epsilon(\bar{c}|c_0, \dots, c_k) d\bar{c} \\ &= \int \left[\prod_{n=k+1}^{k+N} f_\epsilon(c_n|\bar{c}) \right] f_\epsilon(\bar{c}|c_0, \dots, c_k) d\bar{c} \end{aligned} \quad (2.90)$$

l'expression (2.88) se réécrit :

$$f_\epsilon(\Delta_k = x_i|c_0, \dots, c_k) \delta\Delta_k \simeq \int \chi_{\sum \alpha_s^n c_n \in [x_i; x_{i+1}]} \left[\prod_{k+1}^{k+N} f(c_n|\bar{c}) \right] f_\epsilon(\bar{c}|c_0, \dots, c_k) dc_{k+1} \dots dc_{k+N} d\bar{c} \quad (2.91)$$

Cette intégrale est estimée par méthode Monte Carlo en répétant les étapes suivantes un grand nombre M de fois :

1. une valeur de \bar{c} est tirée selon la densité $f_\epsilon(\bar{c}|c_0, \dots, c_k)$ (eq. 2.38) ;
2. pour tout n , $k+1 \leq n \leq k+N$, une valeur du coefficient c_i est tirée selon la densité $f(c_n|\bar{c})$;
3. les fonctions caractéristiques des intégrales (2.91) pour les différents intervalles $[x_i; x_{i+1}]$ de l'histogramme sont toutes évaluées parallèlement. Une seule sera non nulle : celle correspondant à l'intervalle $[x_i; x_{i+1}]$ contenant $\sum_{k+1}^{k+N} \alpha_s^n c_n$.

Les deux premières étapes peuvent s'interpréter comme la création d'une série perturbative imaginaire, qui partagerait les mêmes premiers coefficients c_0, \dots, c_k que la série réelle étudiée et qui vérifierait les hypothèses du modèle. Les résultats sur chacun des segments sont finalement renormalisés en divisant par le nombre d'échantillons M et le pas de l'histogramme $\delta\Delta_k$.

Cette méthode de calcul est utile. Elle permet d'évaluer simplement les densités des quantités étudiées. Il suffit de "simuler" un grand nombre de séries perturbatives, comme

2.5 Comparaison du modèle et de la méthode conventionnelle

si un grand nombre d'expériences étaient réalisées, et de calculer à chaque fois la valeur de la variable considérée. Ce point de vue ne correspond pas pour autant à la réalité. De la même manière que la méthode Monte Carlo, issue de raisonnements probabilistes, peut être employée pour calculer tout type d'intégrale, cette approche numérique n'est rien d'autre qu'une méthode de calcul d'une densité.

2.4.4 Choix de densité $f(c_n|\bar{c})$ non uniforme

La densité approximée (2.87) pour $N = 1$ peut également s'écrire, après intégration de la distribution de Dirac sur c_{k+1} , en utilisant (2.36) et la relation d'indépendance (2.30) :

$$f_\epsilon(\Delta_k|c_0, \dots, c_k) = \mathcal{N}_\epsilon \int f(c_0|\bar{c}) \dots f(c_k|\bar{c}) f(c_{k+1}) = \frac{\Delta_k}{\alpha_s^{k+1}}|\bar{c}) f_\epsilon(\bar{c}) d\bar{c} \quad (2.92)$$

Le facteur de normalisation \mathcal{N}_ϵ est donné par :

$$\mathcal{N}_\epsilon = \left[\alpha_s^{k+1} f_\epsilon(c_0, \dots, c_k) \right]^{-1} \quad (2.93)$$

L'intégrale (2.92) peut être évaluée numériquement pour tout choix de densité $f(c_n|\bar{c})$.

En pratique, il peut être préférable d'utiliser des densités $f(c_n|\bar{c})$ qui ne s'annulent jamais totalement. Dans le cas de la densité gaussienne :

$$f(c_n|\bar{c}) = \frac{1}{\sqrt{2\pi\bar{c}}} e^{-\frac{c_n^2}{2\bar{c}^2}} \quad (2.94)$$

et pour le processus $e^+e^- \rightarrow$ hadrons, on obtient les densités conditionnelles $f(\Delta_k|c_0, \dots, c_k)$ des figures 2.9 et 2.10 pour les choix (2.18) et (2.19) de série $\sigma(Q)$. Le comportement du modèle n'est pas lié au choix (2.31) de la densité $f(c_n|\bar{c})$ aux petits ordres. Les résultats généraux présentés dans la section 2.2.2 sont toujours valables. Les formules analytiques particulières sur la densité $f(\Delta_k|c_0, \dots, c_k)$ (eq. (2.79)) et sur les intervalles d'incertitudes à $p\%$ C.L. (eq. (2.80)) ne sont par contre plus utilisables. Leurs équivalents doivent être obtenus numériquement à partir de l'expression (2.92). Un package Mathematica a été développé pour effectuer ces calculs. Les figures 2.6, 2.7, 2.9 et 2.10 ont été obtenues à l'aide de cet outil.

2.5 Comparaison du modèle et de la méthode conventionnelle

2.5.1 Niveaux de confiance des intervalles conventionnels

La dérivation de la densité conditionnelle pour Δ_k ne nécessite aucune référence à la variation d'échelle de la somme partielle $\sigma_k(Q, \mu)$ avec l'échelle de renormalisation μ qui est utilisée dans l'estimation conventionnelle de l'incertitude sur la somme $\sigma(Q)$. La relation entre la mesure de confiance sur Δ_k et un intervalle $[\sigma_k^-; \sigma_k^+]$ tel que celui défini par (2.6) ou (2.7) est étudiée dans cette section pour évaluer l'étendue de la compatibilité des deux méthodes.

Étant donné une série perturbative spécifique et un jeu de coefficients (c_0, \dots, c_k) , la confiance d'un intervalle $[\Delta_k^-; \Delta_k^+]$ d'incertitude s'écrit :

$$\mathbb{C}(\Delta_k \in [\Delta_k^-; \Delta_k^+]|c_0, \dots, c_k) \equiv \int_{\Delta_k^-}^{\Delta_k^+} f(\Delta_k|c_0, \dots, c_k) d\Delta_k \quad (2.95)$$

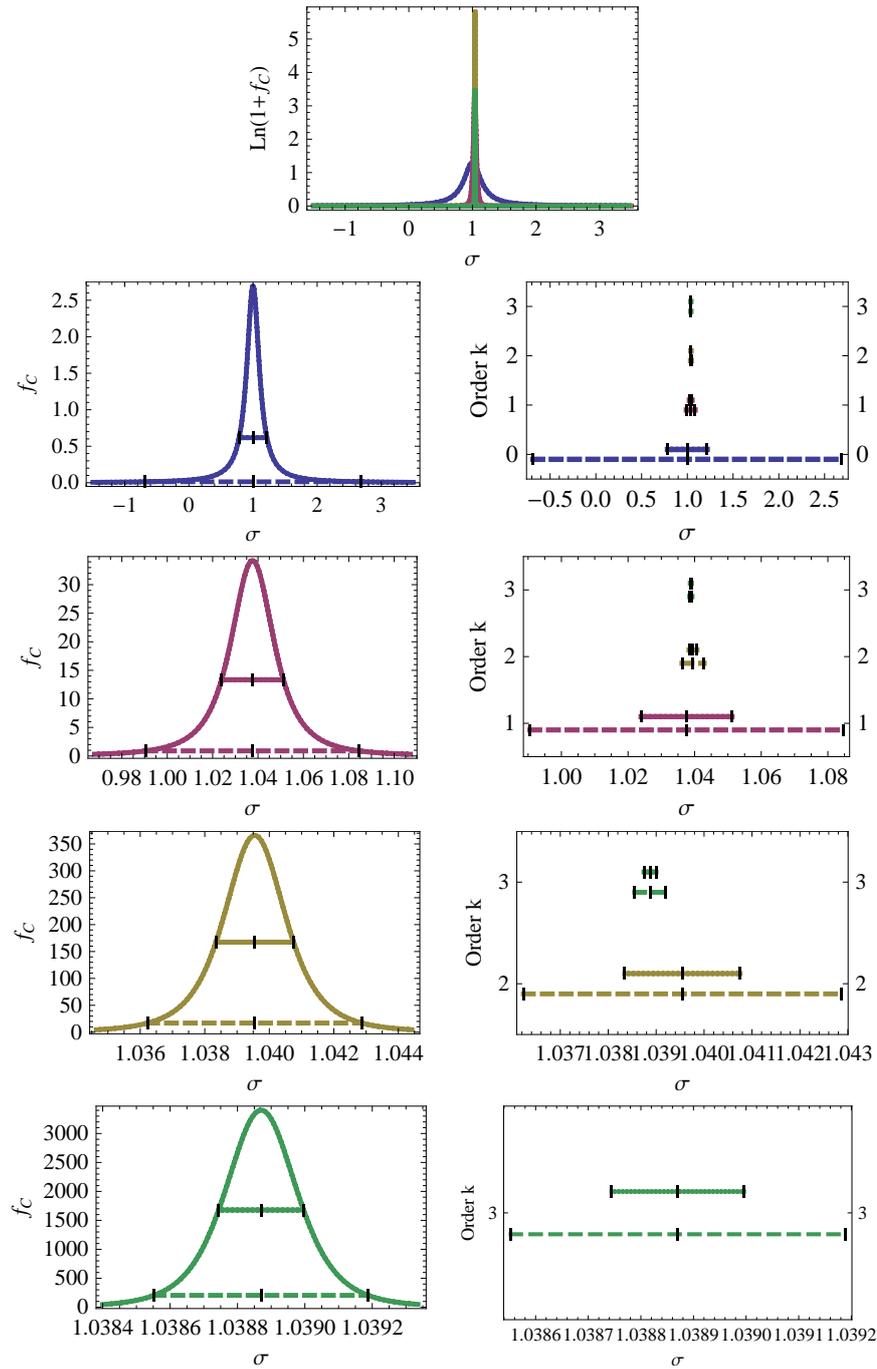


Figure 2.9 – À gauche, densités de confiance $f(\sigma|c_0, \dots, c_k)$, la somme σ étant définie par (2.18), pour k variant, de haut en bas, de 0 à 4. La densité $f(c_n|\bar{c})$ du modèle est gaussienne. Les intervalles tracés sont de confiance 68,3% (traits continus) et 95,5% (pointillés). Ces intervalles sont reportés à droite, à chaque fois pour l'ordre k considéré et tous les ordres supérieurs. En haut, centré, l'ensemble des fonctions $\ln(1 + f)$ des densités à tous les ordres considérés.

2.5 Comparaison du modèle et de la méthode conventionnelle

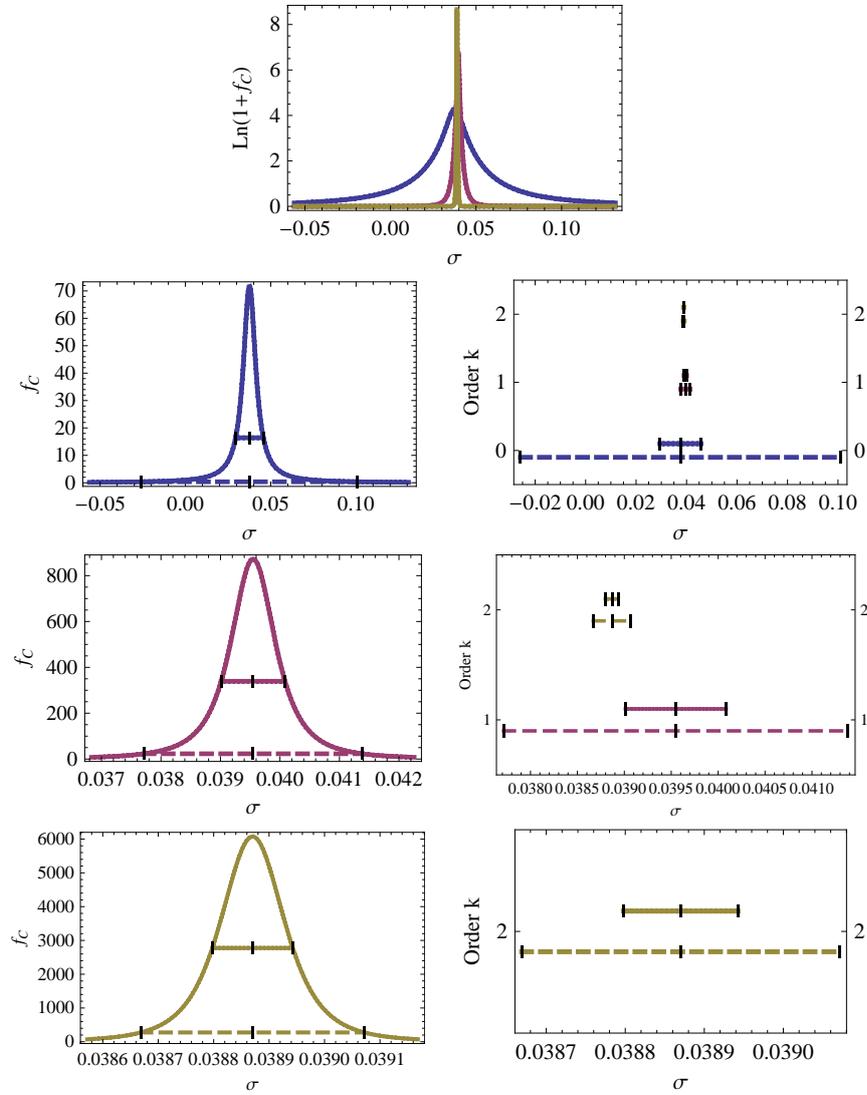


Figure 2.10 – À gauche, densités de confiance $f(\sigma|c'_0, \dots, c'_k)$, la somme σ étant définie par (2.19), pour k variant, de haut en bas, de 0 à 3. La densité $f(c'_n|\bar{c})$ du modèle est gaussienne. Les intervalles tracés sont de confiance 68,3% (traits continus) et 95,5% (pointillés). Ces intervalles sont reportés à droite, à chaque fois pour l'ordre k considéré et tous les ordres supérieurs. En haut, centré, l'ensemble des fonctions $\ln(1 + f)$ des densités à tous les ordres considérés.

Pour fixer les idées, l'intervalle d'incertitude conventionnel $[\sigma_k^-; \sigma_k^+]$ utilisé dans la suite est défini par l'équation (2.6). La densité de confiance est plus naturellement écrite en fonction du reste Δ_k de la somme. Les intervalles considérés sont translatés pour lui correspondre :

$$\Delta_k^- \equiv \min(\sigma_k(Q, Q/2), \sigma_k(Q, 2Q)) - \sigma_k = \sigma_k^- - \sigma_k \quad (2.96)$$

$$\Delta_k^+ \equiv \max(\sigma_k(Q, Q/2), \sigma_k(Q, 2Q)) - \sigma_k = \sigma_k^+ - \sigma_k \quad (2.97)$$

La variation de la somme partielle $\sigma_k(Q, \mu)$ en fonction de l'échelle de renormalisation μ dépend des valeurs des coefficients c_0, \dots, c_k . Il en est donc de même des valeurs des paramètres Δ_k^- et Δ_k^+ . La densité de confiance conditionnelle $f(\Delta_k | c_0, \dots, c_k)$ ne dépend quant à elle que de leur maximum $\bar{c}_{(k)}$. La confiance (2.95) de l'intervalle conventionnel estimé dans le modèle n'est donc pas une constante. Elle varie avec le processus étudié et la valeur des coefficients calculés. Celle-ci est donc une variable inconnue du modèle fonction de ces éléments :

$$x_k(c_0, \dots, c_k) \equiv \mathbb{C}(\Delta_k \in [\Delta_k^-; \Delta_k^+] | c_0, \dots, c_k) \quad (2.98)$$

Avant qu'un seul ordre de la série n'aie été calculé, la valeur de la confiance (2.95) est inconnue. Celle-ci ne sera obtenue que lorsque les coefficients c_0, \dots, c_k auront été évalués. Ses valeurs *a priori* envisageables sont notées x_k .

Les densités de confiance de ces variables inconnues x_k peuvent être évaluées numériquement en utilisant encore une fois le fait que les mesures de probabilités et de confiance partagent le même formalisme. Une grande quantité de configurations $(\bar{c}, c_0, \dots, c_k)$ est générée et la confiance (2.95) est à chaque fois évaluée. Les résultats sont rangés dans un histogramme. Plus précisément la densité $f(x_k | \bar{c})$ à \bar{c} fixé de la variable (2.98) est évaluée pour simplifier les calculs. Seuls les coefficients c_0, \dots, c_k sont tirés, et ce suivant la densité uniforme (2.31) $f(c_n | \bar{c})$ pour chacun d'entre eux. Cela permet en particulier de ne pas avoir recours à une valeur de ϵ non nul. La figure 2.11 présente les distributions des valeurs de confiance obtenues dans les cas $k = 1, 2$ et 3 pour $\bar{c} = 1$ et $7, 5$. Les densités ne dépendent pas de \bar{c} : elles correspondent donc également aux distributions $f(x_k)$.

La confiance (2.95) prédite par le modèle est en large partie non affectée par la valeur des coefficients : sa distribution prend une forme qui se rapproche de celle d'une distribution de Dirac centrée autour de 57%, 96% et 99% respectivement pour $k = 1, 2$ et 3 . Le pic en x_3 nul est un artéfact dû aux configurations où $\sigma_3(Q, 2Q)$ et $\sigma_3(Q, Q/2)$ sont accidentellement proches, conduisant à un intervalle $[\Delta_k^-; \Delta_k^+]$ anormalement petit. Il disparaît lorsque l'intervalle d'incertitude $[\sigma_k^-; \sigma_k^+]$ est défini selon une autre convention.

2.5.2 Évolution des valeurs de confiance

Certaines approximations permettent de simplifier l'expression de la variable inconnue x_k précédente et de prédire en partie le comportement observé sur la figure 2.11. La confiance peut être supposée essentiellement ne pas changer en première approche pour un intervalle $[\Delta_k^-; \Delta_k^+]$ centré :

$$\mathbb{C}(\Delta_k \in [\Delta_k^-; \Delta_k^+] | c_0, \dots, c_k) \simeq \mathbb{C}(\Delta_k \in [-\frac{\delta_k}{2}; \frac{\delta_k}{2}] | c_0, \dots, c_k) \quad (2.99)$$

2.5 Comparaison du modèle et de la méthode conventionnelle

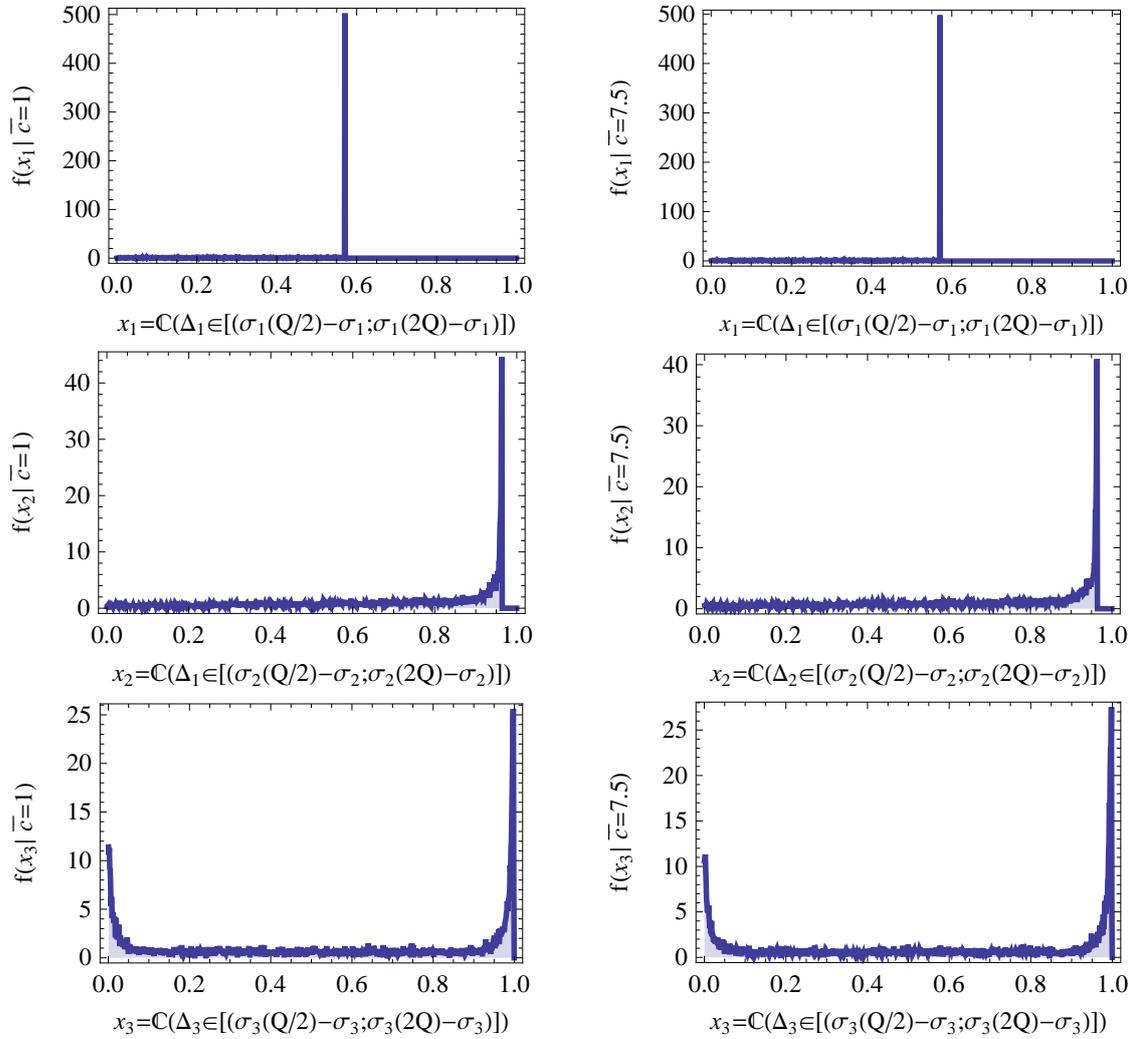


Figure 2.11 – Estimation numérique des densités de confiance exactes $f(x_k|\bar{c})$, où la variable inconnue x_k définie par (2.98) est la confiance de l'intervalle de variation des échelles, dans les cas $\bar{c} = 1$ (gauche) et 7.5 (droite) pour $k = 1$ (haut), $k = 2$ (milieu) et $k = 3$ (droite). Les figures ont été obtenus pour $N = 10^4$ échantillons distribués sur des intervalles de largeur 10^{-3} .

La densité (2.78) s'intègre facilement sur celui-ci. En utilisant l'expression approchée de la largeur de l'intervalle conventionnel (2.27) :

$$\mathbb{C}(\Delta_k \in [-\frac{\delta_k}{2}, \frac{\delta_k}{2}] | c_0, \dots, c_k) = \begin{cases} 1 - \frac{1}{k+2} \left[\frac{2}{3k\beta_0} \frac{\bar{c}_{(k)}}{|c_k|} \right]^{k+1} & \text{si } |c_k| \geq \frac{2}{3k\beta_0} \bar{c}_{(k)} \\ \frac{k+1}{k+2} \frac{3k\beta_0}{2} \frac{|c_k|}{\bar{c}_{(k)}} & \text{si } |c_k| < \frac{2}{3k\beta_0} \bar{c}_{(k)} \end{cases} \quad (2.100)$$

Cette expression est bel et bien indépendante de \bar{c} . De plus, lorsque le paramètre $\bar{c}_{(k)}$ est égal à $|c_k|$, la confiance de l'intervalle conventionnel est indépendante des valeurs des coefficients calculés. Étant donné que le facteur $\frac{2}{3.2\beta_0}$ est plus petit que l'unité, celle-ci s'écrit :

$$\mathbb{C}(\Delta_k \in [\Delta_k^-, \Delta_k^+] | c_0, \dots, c_k) \simeq \begin{cases} 1 - \frac{1}{k+2} \left[\frac{2}{3k\beta_0} \right]^{k+1} & \text{si } k \geq 2 \\ \frac{k+1}{k+2} \frac{3\beta_0}{2} & \text{si } k = 1 \end{cases} \quad (2.101)$$

Plus généralement, la valeur absolue du dernier coefficient calculé $|c_k|$ est toujours inférieure ou égale à $\bar{c}_{(k)}$. L'expression (2.101) correspond donc au maximum de la confiance de l'intervalle conventionnel calculée par le modèle.

Ces résultats peuvent être réexprimés en terme de densité de confiance sur la variable inconnue x_k de manière à faire le lien avec la section précédente. La densité conditionnelle $f(x_k | c_0, \dots, c_k, \bar{c})$ s'écrit, en supposant dans un premier temps que $|c_k|$ est toujours de l'ordre de $\bar{c}_{(k)}$:

$$f(x_k | c_0, \dots, c_k, \bar{c}) \simeq \begin{cases} \delta \left(x_k - \left[1 - \frac{1}{k+2} \left[\frac{2}{3k\beta_0} \right]^{k+1} \right] \right) & \text{si } k \geq 2 \\ \delta \left(x_k - \frac{k+1}{k+2} \frac{3\beta_0}{2} \right) & \text{si } k = 1 \end{cases} \quad (2.102)$$

La densité obtenue est indépendante de la valeur des coefficients et de \bar{c} : les densités $f(x_k)$ et $f(x_k | \bar{c})$ ont donc bel et bien les mêmes expressions. Pour des calculs au LO ($k = 1$), NLO ($k = 2$) et NNLO ($k = 3$), il s'agit de distributions de Dirac centrées sur des valeurs de confiance de 61%, 96% et 99.6% respectivement, pour $\beta_0 = 0,61$. Ces densités sont en bon accord avec celles obtenues depuis l'estimation numérique des densités exactes.

De manière plus générale, la densité doit s'annuler après ce pic car la confiance de l'intervalle d'incertitude conventionnel ne peut pas être plus élevée que (2.101). La distribution peut pas contre être non nulle sur des valeurs plus basses. Il s'agit bien de ce qui est observé sur la figure 2.11, notamment en $k = 2$ (graphiques du milieu).

La dépendance en k des résultats (2.100) et (2.102) montre que, sous le modèle proposé, la confiance accordée aux intervalles conventionnels $[\sigma_k^-, \sigma_k^+]$ n'est pas une constante mais dépend de l'ordre en perturbation étudié. Lorsque de nouveaux ordres supérieurs sont calculés dans une série perturbative, non seulement la taille de l'intervalle conventionnel diminue, mais sa confiance augmente.

SÉRIES DÉBUTANT À L'ORDRE l EN α_s

L'approximation analytique effectuée ci-dessus permet de considérer le cas de séries perturbatives dont le premier terme est à l'ordre l en α_s :

$$\sum_l^\infty c_n \alpha_s^n \quad (2.103)$$

2.5 Comparaison du modèle et de la méthode conventionnelle

La définition (2.19) de la série $\sigma(Q)$ pour le processus $e^+e^- \rightarrow$ hadrons correspond par exemple au cas $l = 1$.

La largeur de l'intervalle conventionnel d'incertitude est toujours :

$$\delta_k \simeq 3k\beta_0\alpha_s^{k+1}|c_k| \quad (2.104)$$

La densité de confiance (2.79) sur Δ_k conserve quant à elle la même expression pourvu qu'elle soit écrite en fonction des coefficients :

$$c'_n \equiv \alpha_s^l c_{n+l} \quad (2.105)$$

L'hypothèse selon laquelle les coefficients c_n sont tous du même ordre de grandeur se transmet directement aux c'_n . Cette définition est identique à la définition (2.84) déjà présentée pour le cas $l = 1$ dans le cas du processus $e^+e^- \rightarrow$ hadrons. En fonction des $k-l+1$ coefficients c_l, \dots, c_k calculés, la densité de confiance sur le reste de la série prend alors la forme :

$$f(\Delta_k | c_l, \dots, c_k) \simeq \left(\frac{k-l+1}{k-l+2} \right) \frac{1}{2\alpha_s^{k+1}\bar{c}_{(k)}} \begin{cases} 1 & \text{si } |\Delta_k| \leq \alpha_s^{k+1}\bar{c}_{(k)} \\ \frac{1}{(|\Delta_k|/\alpha_s^{k+1}\bar{c}_{(k)})^{k-l+2}} & \text{si } |\Delta_k| > \alpha_s^{k+1}\bar{c}_{(k)} \end{cases} \quad (2.106)$$

Le paramètre $\bar{c}_{(k)}$ est toujours défini comme le maximum des valeurs absolues des coefficients calculés.

La confiance de l'intervalle conventionnel (2.95), sous l'approximation (2.99), s'écrit dans ce cas :

$$\mathbb{C}(\Delta_k \in [-\frac{\delta_k}{2}, \frac{\delta_k}{2}] | c_l, \dots, c_k) = \begin{cases} 1 - \frac{1}{k-l+2} \left[\frac{2}{3k\beta_0} \frac{\bar{c}_{(k)}}{|c_k|} \right]^{k-l+1} & \text{si } |c_k| \geq \frac{2}{3k\beta_0}\bar{c}_{(k)} \\ \frac{k-l+1}{k-l+2} \frac{3k\beta_0}{2} \frac{|c_k|}{\bar{c}_{(k)}} & \text{si } |c_k| < \frac{2}{3k\beta_0}\bar{c}_{(k)} \end{cases} \quad (2.107)$$

et sa valeur maximale :

$$\mathbb{C}(\Delta_k \in [\Delta_k^-; \Delta_k^+] | c_l, \dots, c_k) \simeq \begin{cases} 1 - \frac{1}{k-l+2} \left[\frac{2}{3k\beta_0} \right]^{k-l+1} & \text{si } k \geq 2 \\ \frac{k-l+1}{k-l+2} \frac{3\beta_0}{2} & \text{si } k = 1 \end{cases} \quad (2.108)$$

Pour un processus débutant à l'ordre α_s^1 (*i.e.* $l = 1$), la valeur maximale de confiance de l'intervalle conventionnel d'incertitude est de 46% au LO ($k = 1$), 90% au NLO ($k = 2$) et 98,8% au NNLO ($k = 3$). Pour un processus commençant en α_s^2 (*i.e.* $l = 2$) cette équation prédit une confiance maximale de 73% au LO ($k=2$), 96% au NLO ($k=3$) et 99.5% au NNLO ($k=4$).

Les valeurs de confiance au NLO et NNLO en $k \geq 2$ sont plus importantes que celles pour $l = 0$ car les intervalles conventionnels sont proportionnellement plus larges. En effet, l'expression approchée (2.104) de δ_k s'écrit en fonction des coefficients (2.105), pour $k' = k - l$:

$$\delta_k \simeq 3(k' + l)\beta_0\alpha_s^{k'+1}|c'_{k'}| \quad (2.109)$$

Le *Leading Order* correspond à $k' = 0$, le NLO à $k' = 1$, et ainsi de suite. Le paramètre δ_k correspond en première approximation à la variation d'échelle du dernier ordre $\alpha_s^{k'} c'_{k'} =$

	Uniforme			Gaussienne				
$l = 0$	k'	1	2	3	k'	1	2	3
	C	19%	25%	61%	C	26%	36%	67%
$l = 1$	k'	0	1	2	k'	0	1	2
	C	46%	58%	77%	C	46%	61%	79%

Tableau 2.2 – Confiance des intervalles d’incertitude conventionnels du processus $e^+e^- \rightarrow$ hadrons calculées pour des densités initiales $f(c_n|\bar{c})$ uniformes ou gaussiennes, pour les choix (2.18) ($l = 0$) et (2.19) ($l = 1$) de définition de série perturbative.

$\alpha_s^{k'+l} c_k$ calculé. Plus précisément, il correspond à la variation de la puissance $k' + l$ de α_s en fonction de l’échelle de renormalisation. À k' fixé, lorsque la série perturbative débute à l’ordre l plutôt que 0, cette variation d’échelle est celle d’une puissance l plus grande de α_s . Comme $\alpha_s(\mu)$ est monotone décroissante, $(\alpha_s(\mu)/\alpha_s(Q))^{k'+l}$ décroît plus vite que $(\alpha_s(\mu)/\alpha_s(Q))^{k'}$. L’intervalle $[\alpha_s^{k'+l}(Q/2); \alpha_s^{k'+l}(2Q)]$ de variation en l non nul est donc plus grand que l’intervalle de variation $\alpha_s^l(Q)[\alpha_s^{k'}(Q/2); \alpha_s^{k'}(2Q)]$ du cas $l = 0$, rééchelonné à la même puissance de α_s .

2.5.3 Exemple du processus $e^+e^- \rightarrow$ hadrons

Il serait tentant de considérer les équations (2.101) et (2.108) comme les résultats principaux du modèle, permettant d’associer une confiance aux intervalles d’incertitude fournis par la méthode conventionnelle. La simplicité de ces équations et leur valeurs accidentellement proches des niveaux de confiance des variances gaussiennes, en font en apparence de bons candidats pour une telle identification. Il est pourtant important de garder en tête que ces équations dépendent du choix effectué pour la distribution (2.31), ainsi que des différentes approximations effectuées pour les obtenir. La figure 2.11 montre d’ailleurs que plus le dernier ordre k calculé est élevé, plus le nombre d’ “accidents”, *i.e.* de cas pour lesquels les valeurs maximales (2.101) et (2.108) de confiance ne s’appliquent pas, est grand.

Dans le cas du processus $e^+e^- \rightarrow$ hadrons, les confiances exactes des intervalles conventionnels (2.22), obtenues par intégration (2.95) de la densité de confiance sur le reste Δ_k , sont présentés dans le tableau 2.2. Elles ont été calculées dans les cas d’une densité $f(c_n|\bar{c})$ uniforme ou gaussienne, pour les choix de paramètre l nul ($\sigma(Q)$ défini par (2.18)) ou l égal à un ($\sigma(Q)$ défini par (2.19)).

Dans le cas où l est nul, et pour le choix de densité $f(c_n|\bar{c})$ uniforme, les confiances obtenues sont beaucoup plus faibles que les valeurs maximales de 60%, 96% et 99.6% prédites par la formule (2.101). Le premier coefficient c_0 est égal à un, les suivants sont les coefficients des équations (2.20). c_1 est par exemple beaucoup plus faible que c_0 . La méthode conventionnelle, au travers de la variation des échelle, fournit un intervalle d’incertitude de largeur δ_1 dicté par c_1 . La mesure de confiance estime à l’opposé l’incertitude par le maximum des coefficients calculés $\bar{c}_{(1)}$ qui s’avère être c_0 . L’intervalle conventionnel est donc proportionnellement beaucoup plus petit que la largeur de la densité sur le reste Δ_1 . À l’intégration, sa confiance sera bien plus faible que la valeur extrême de l’équation (2.101). Le raisonnement est identique aux ordres suivants. La méthode conventionnelle estime l’incertitude avec le dernier coefficient calculé. Le modèle de confiance utilise

2.6 Approximation des ordres supérieurs

la totalité des coefficients. Il privilégie le plus grand d'entre eux, à savoir, dans ce cas, c_0 .

Les expressions (2.100) et (2.107), plus complexes mais également plus complètes, permettent une meilleure estimation des confiances des intervalles conventionnels. Dans le cas où l est par exemple égal à un, celles-ci prédisent des valeurs de 46%, 55% et 98,8%. Ces résultats sont beaucoup plus proches des valeurs exactes prédites par le modèle et présentées dans le tableau 2.2. Seule la confiance du dernier ordre est réellement plus faible que la valeur fournie par la formule (2.107), l'intervalle d'incertitude conventionnel $[\sigma_3^-; \sigma_3^+]$ donné en (2.22) n'étant pas centré sur $\sigma_3(Q)$. Les valeurs exactes des confiances ne peuvent donc être obtenues, *in fine*, que par intégration de la densité de confiance sur le reste de la série Δ_k .

En pratique, la méthode proposée a pour objectif de se substituer à la méthode conventionnelle. Elle permet d'une part d'obtenir des formes de densité complètes, et ainsi d'évaluer la confiance d'intervalles donnés ou de combiner ses résultats avec d'autres sources d'incertitudes. Mais plus important encore, elle permet de contrôler les hypothèses utilisées dans la création de la mesure de confiance. Celles-ci ne traduiront certainement pas des informations vraies en toute circonstance. La mesure de confiance ne peut rien dire sur la véracité des hypothèses qu'elle traduit. Par contre, dans l'acte de création de la mesure, l'utilisateur est amené à expliciter les informations utilisées. Cela ouvre la voie à la critique des hypothèses sous-jacentes généralement admises, leur cadre d'application, ainsi qu'à la recherche d'un possible jeu d' "hypothèses minimales".

2.6 Approximation des ordres supérieurs

Il est possible de complexifier le modèle présenté jusque là. Les hypothèses peuvent être modifiées ou de nouvelles sources d'incertitude ou d'information peuvent être ajoutées. L'incertitude sur la valeur de $\alpha_s(Q)$ peut par exemple être prise en compte en définissant une variable α_s inconnue, indépendante de toutes les autres, et de densité $f(\alpha_s)$. Elle étend l'espace de travail mais ne modifie pas les densités de confiance sur les \bar{c}, c_0, \dots, c_k . La seule modification intervient dans la définition de la densité de confiance sur Δ_k qui s'écrit alors :

$$f(\Delta_k | c_0, \dots, c_k) = \int \delta(\Delta_k - \sum_{k+1}^{\infty} c_n \alpha_s^n) f(c_{k+1}, c_{k+2}, \dots | c_0, \dots, c_k) f(\alpha_s) dc_{k+1} dc_{k+2} \dots d\alpha_s \quad (2.110)$$

Le modèle étudié jusque là suppose d'autre part la connaissance parfaite de certains coefficients, jusque l'ordre k , et l'ignorance totale de ceux d'ordre supérieur. En pratique, il est souvent possible de connaître partiellement un coefficient d'ordre supérieur, typiquement calculé sous une certaine approximation ou obtenu comme une expansion d'une resommation à tous les ordres. Cette section présente succinctement l'adaptation du modèle à ce cas.

2.6.1 Ajustement du modèle

2.6.1.1 Informations supplémentaires

VARIABLES INCONNUES

Un ordre $k + 1$ peut par exemple avoir été évalué partiellement. Une nouvelle variable inconnue \tilde{c}_{k+1} est ajoutée. Celle-ci définit les valeurs *a priori* envisageables de l'approximation du coefficient c_{k+1} . L'espace de travail est maintenant celui des variables $(\tilde{c}, c_0, c_1, \dots, \tilde{c}_{k+1})$. Si plusieurs ordres sont connus approximativement, une nouvelle variable \tilde{c}_i est ajoutée pour chacun d'entre eux.

PARAMÈTRES D'APPROXIMATION

La variable \tilde{c}_{k+1} correspond à l'approximation de l'ordre c_{k+1} . Cette information sur la nature du paramètre doit être traduite :

Puisque \tilde{c}_{k+1} est une valeur approchée du coefficient d'ordre $k + 1$, sa valeur ne devrait pas fournir plus d'information que la valeur exacte de c_{k+1} elle-même. En d'autres termes, si la valeur de c_{k+1} est connue, la connaissance de la valeur approchée \tilde{c}_{k+1} ne doit rien changer. Pour un ensemble de coefficients $\{c_i\}$:

$$f(\{c_i\}|c_{k+1}, \tilde{c}_{k+1}) = f(\{c_i\}|c_{k+1}) \quad (2.111)$$

$$f(\tilde{c}|c_{k+1}, \tilde{c}_{k+1}) = f(\tilde{c}|c_{k+1}) \quad (2.112)$$

QUALITÉ D'APPROXIMATION

L'approximation de c_{k+1} par \tilde{c}_{k+1} peut être de plus ou moins bonne qualité. Une densité de confiance $f(\tilde{c}_{k+1}|c_{k+1})$ pour obtenir la valeur approchée \tilde{c}_{k+1} , étant donnée la valeur exacte c_{k+1} doit être introduite. Le choix de cette distribution dépend de la manière dont \tilde{c}_{k+1} a été obtenue. Le choix de la densité log-normale permet par exemple de fixer les idées et d'étudier le comportement du modèle modifié :

$$f(\tilde{c}_{k+1}|c_{k+1}) = \frac{1}{|\tilde{c}_{k+1}|} \frac{1}{\sqrt{2\pi \ln f}} \exp\left(-\frac{(\ln(\tilde{c}_{k+1}/c_{k+1}))^2}{2(\ln f)^2}\right) \quad (2.113)$$

Le paramètre f est fixé par l'utilisateur. L'élément \tilde{c}_{k+1} estime ici le coefficient c_{k+1} à un facteur f près.

2.6.1.2 Mesure de confiance

Les densités de confiance sur la valeur réelle du coefficient c_{k+1} , qui est approximativement connue, et sur un coefficient totalement inconnu c_n peuvent alors s'écrire, à des facteurs de normalisation collectivement appelés \mathcal{N} :

$$\begin{aligned} f(c_{k+1}|c_0, \dots, c_k, \tilde{c}_{k+1}) &= \mathcal{N} f(c_{k+1}|c_0, \dots, c_k) f(\tilde{c}_{k+1}|c_{k+1}) \\ f(c_n|c_0, \dots, c_k, \tilde{c}_{k+1}) &= \mathcal{N} \int f(c_n, c_{k+1}|c_0, \dots, c_k) f(\tilde{c}_{k+1}|c_{k+1}) \, dc_{k+1} \end{aligned} \quad (2.114)$$

Plus généralement, pour des ensembles arbitraires de coefficients connus $C_K = \{c_i\}_{i \in K}$, d'approximations $\tilde{C}_A = \{\tilde{c}_i\}_{i \in A}$ de coefficients $C_A = \{c_i\}_{i \in A}$ et de coefficients totalement inconnus $C_N = \{c_i\}_{i \in N}$ il est possible d'écrire :

$$f(C_N, C_A|C_K, \tilde{C}_A) = \mathcal{N} \cdot f(C_N, C_A|C_K) \cdot f(\tilde{C}_A|C_A) \quad (2.115)$$

2.6 Approximation des ordres supérieurs

La densité de confiance sur les coefficients inconnus C_N s'obtient par intégration de (2.115) sur les coefficients C_A . La densité de confiance conditionnelle sur le reste Δ_k de la série tronquée s'obtient en remplaçant la densité $f(C_N|C_K)$ par l'expression (2.115) dans sa définition générale (2.77).

2.6.1.3 Dérivation des densités (2.114)

La dérivation des densités $f(c_{k+1}|c_0, \dots, c_k, \tilde{c}_{k+1})$ et $f(c_n|c_0, \dots, c_k, \tilde{c}_{k+1})$ s'effectue toujours en ϵ non nul :

$$\begin{aligned} f_\epsilon(c_{k+1}|c_0, \dots, c_k, \tilde{c}_{k+1}) &= \frac{f_\epsilon(c_0, \dots, c_k, c_{k+1}, \tilde{c}_{k+1})}{\int f_\epsilon(c_0, \dots, c_k, c_{k+1}, \tilde{c}_{k+1}) \, dc_{k+1}} \end{aligned} \quad (2.116)$$

$$= \frac{f_\epsilon(c_0, \dots, c_k|c_{k+1}, \tilde{c}_{k+1})f_\epsilon(c_{k+1}, \tilde{c}_{k+1})}{\int f_\epsilon(c_0, \dots, c_k|c_{k+1}, \tilde{c}_{k+1})f_\epsilon(c_{k+1}, \tilde{c}_{k+1}) \, dc_{k+1}} \quad (2.117)$$

$$= \frac{f_\epsilon(c_0, \dots, c_k|c_{k+1})f_\epsilon(c_{k+1}, \tilde{c}_{k+1})}{\int f_\epsilon(c_0, \dots, c_k|c_{k+1})f_\epsilon(c_{k+1}, \tilde{c}_{k+1}) \, dc_{k+1}} \quad (2.118)$$

$$= \frac{[f_\epsilon(c_0, \dots, c_k, c_{k+1})/f_\epsilon(c_{k+1})][f_\epsilon(\tilde{c}_{k+1}|c_{k+1})f_\epsilon(c_{k+1})]}{\int [f_\epsilon(c_0, \dots, c_k, c_{k+1})/f_\epsilon(c_{k+1})][f_\epsilon(\tilde{c}_{k+1}|c_{k+1})f_\epsilon(c_{k+1})] \, dc_{k+1}} \quad (2.119)$$

$$= \frac{f_\epsilon(c_0, \dots, c_k, c_{k+1})f_\epsilon(\tilde{c}_{k+1}|c_{k+1})}{\int f_\epsilon(c_0, \dots, c_k, c_{k+1})f_\epsilon(\tilde{c}_{k+1}|c_{k+1}) \, dc_{k+1}} \quad (2.120)$$

$$= \mathcal{N}_\epsilon f_\epsilon(c_{k+1}|c_0, \dots, c_k) f(\tilde{c}_{k+1}|c_{k+1}) \quad (2.121)$$

Les lignes (2.116) et (2.117) utilisent l'expression des densités conditionnelle similaires à (2.36). En (2.118), le paramètre \tilde{c}_{k+1} disparaît en utilisant l'équation (2.111). L'expression des densités conditionnelles est de nouveau utilisée à la ligne (2.119), et la forme obtenue est simplifiée en (2.120). La relation des densités conditionnelles permet de ré-écrire une dernière fois la dernière ligne, la densité $f_\epsilon(c_0, \dots, c_k)$ étant incluse dans le facteur de normalisation. La densité $f(\tilde{c}_{k+1}|c_{k+1})$ ne dépend pas de ϵ . Dans la limite $\epsilon \rightarrow 0$:

$$f(c_{k+1}|c_0, \dots, c_k, \tilde{c}_{k+1}) = \mathcal{N} f(c_{k+1}|c_0, \dots, c_k) f(\tilde{c}_{k+1}|c_{k+1}) \quad (2.122)$$

La densité $f(c_n|c_0, \dots, c_k, \tilde{c}_{k+1})$, pour $n > k + 1$, s'écrit, avec le même type de manipulations :

$$\begin{aligned} f(c_n|c_0, \dots, c_k, \tilde{c}_{k+1}) &= \int f(c_n, c_{k+1}|c_0, \dots, c_k, \tilde{c}_{k+1}) \, dc_{k+1} \end{aligned} \quad (2.123)$$

$$= \int f(c_n|c_0, \dots, c_k, \tilde{c}_{k+1}, c_{k+1}) f(c_{k+1}|c_0, \dots, c_k, \tilde{c}_{k+1}) \, dc_{k+1} \quad (2.124)$$

$$= \mathcal{N} \int f(c_n|c_0, \dots, c_k, c_{k+1}) f(c_{k+1}|c_0, \dots, c_k) f(\tilde{c}_{k+1}|c_{k+1}) \, dc_{k+1} \quad (2.125)$$

$$= \mathcal{N} \int f(c_n, c_{k+1}|c_0, \dots, c_k) f(\tilde{c}_{k+1}|c_{k+1}) \, dc_{k+1} \quad (2.126)$$

À la ligne (2.124), on utilise l'expression d'une densité conditionnelle dans le sous-espace à $c_0, \dots, c_k, \tilde{c}_{k+1}$ fixés. Les deux dérivations s'étendent directement au cas d'ensembles C_A de coefficients approximatés et C_N de coefficients inconnus.

2.6.2 Comportement du modèle

Pour fixer les idées, ce paragraphe étudie par exemple la densité $f(\Delta_0|c_0, \tilde{c}_1)$ dans le cas d'un coefficient connu c_0 et d'un coefficient c_1 approximé par \tilde{c}_1 . Suivant la qualité supposée de l'approximation, l'incertitude sur c_1 ou c_2 va prédominer. Dans le cas général, l'expression de la densité doit être dérivée en conservant ces deux variables. L'approximation du reste de la série par ses deux premiers ordres est donc nécessaire :

$$\Delta_0 \simeq c_1\alpha_s + c_2\alpha_s^2 \quad (2.127)$$

Le nombre de paramètres α_s dans les expressions est réduit en s'intéressant plutôt :

$$f\left(\frac{\Delta_0}{\alpha_s}|c_0, \tilde{c}_1\right) \simeq f(c_1 + c_2\alpha_s = \frac{\Delta_0}{\alpha_s}|c_0, \tilde{c}_1) \quad (2.128)$$

Cette dernière densité est définie de manière similaire à celle sur Δ_0 dans l'équation (2.77). On utilise dans cette expression l'équation (2.115) puis on intègre la distribution de Dirac sur le coefficient approximé c_1 :

$$f(c_1 + c_2\alpha_s = x|c_0, \tilde{c}_1) = \int \delta(x - (c_1 + c_2\alpha_s))f(c_2, c_1|c_0, \tilde{c}_1) dc_1 dc_2 \quad (2.129)$$

$$= \mathcal{N} \int \delta(x - (c_1 + c_2\alpha_s))f(c_2, c_1|c_0)f(\tilde{c}_1|c_1) dc_1 dc_2 \quad (2.130)$$

$$= \mathcal{N} \int f(c_2, c_1^{(x, c_2)}|c_0)f(\tilde{c}_1|c_1^{(x, c_2)}) dc_2 \quad (2.131)$$

Le résultat de la fonction $c_1^{(x, c_2)} = x - c_2\alpha_s$ correspond à la valeur de c_1 fixée par intégration sur la mesure de Dirac. La densité $f(c_2, c_1|c_0)$ s'obtient de la même manière que $f(c_n|c_0, \dots, c_k)$:

$$f(c_2, c_1|c_0) = \lim_{\epsilon \rightarrow 0} \frac{f_\epsilon(c_0, c_1, c_2)}{f_\epsilon(c_0)} \propto \left(\frac{1}{\bar{c}_{(2)}}\right)^3 \quad (2.132)$$

L'expression de la densité recherchée prend finalement la forme :

$$f(c_1 + c_2\alpha_s = x|c_0, \tilde{c}_1) = \mathcal{N} \int \left(\frac{1}{\bar{c}_{(2)}^{(x, c_2)}}\right)^3 f(\tilde{c}_1|c_1^{(x, c_2)}) dc_2 \quad (2.133)$$

où $\bar{c}_{(2)}^{(x, c_2)} = \max(c_0, c_1^{(x, c_2)}, c_2)$.

La figure 2.12 présente l'évolution de la densité de confiance (2.133), pour $f(\tilde{c}_1|c_1)$ défini par (2.113), à mesure que la qualité supposée de l'approximation de c_1 par \tilde{c}_1 augmente. Les paramètres c_0 et \tilde{c}_1 sont fixés arbitrairement à 0,9 et 0,83 respectivement.

Sur le graphique en haut à gauche, l'approximation de c_1 par \tilde{c}_1 est supposée être très peu précise ($f = 50$). La seule information utile venant de \tilde{c}_1 est son signe (positif). L'incertitude sur $\frac{\Delta_0}{\alpha_s}$ est dominée par celle sur c_1 . La densité $f(\frac{\Delta_0}{\alpha_s}|c_0, \tilde{c}_1)$ (en trait continu rouge) coïncide donc avec $f(c_1|c_0, \tilde{c}_1)$ (en pointillés marrons). Ces deux densités sont de plus essentiellement identiques à $f(c_1|c_0)$. Plus précisément, puisque \tilde{c}_1 indique que c_1 est positif, $f(c_1|c_0, \tilde{c}_1)$ s'annule sur le demi axe négatif des abscisses et coïncide sur l'autre demi axe avec le double de la densité $f(c_1|c_0)$ (courbe la plus étalée en tirets noirs). Les différences provenant de l'incertitude sur c_2 et la faible information fournie par la valeur de \tilde{c}_1 sont presque négligeables et ne se font ressentir qu'autour de zéro.

2.6 Approximation des ordres supérieurs

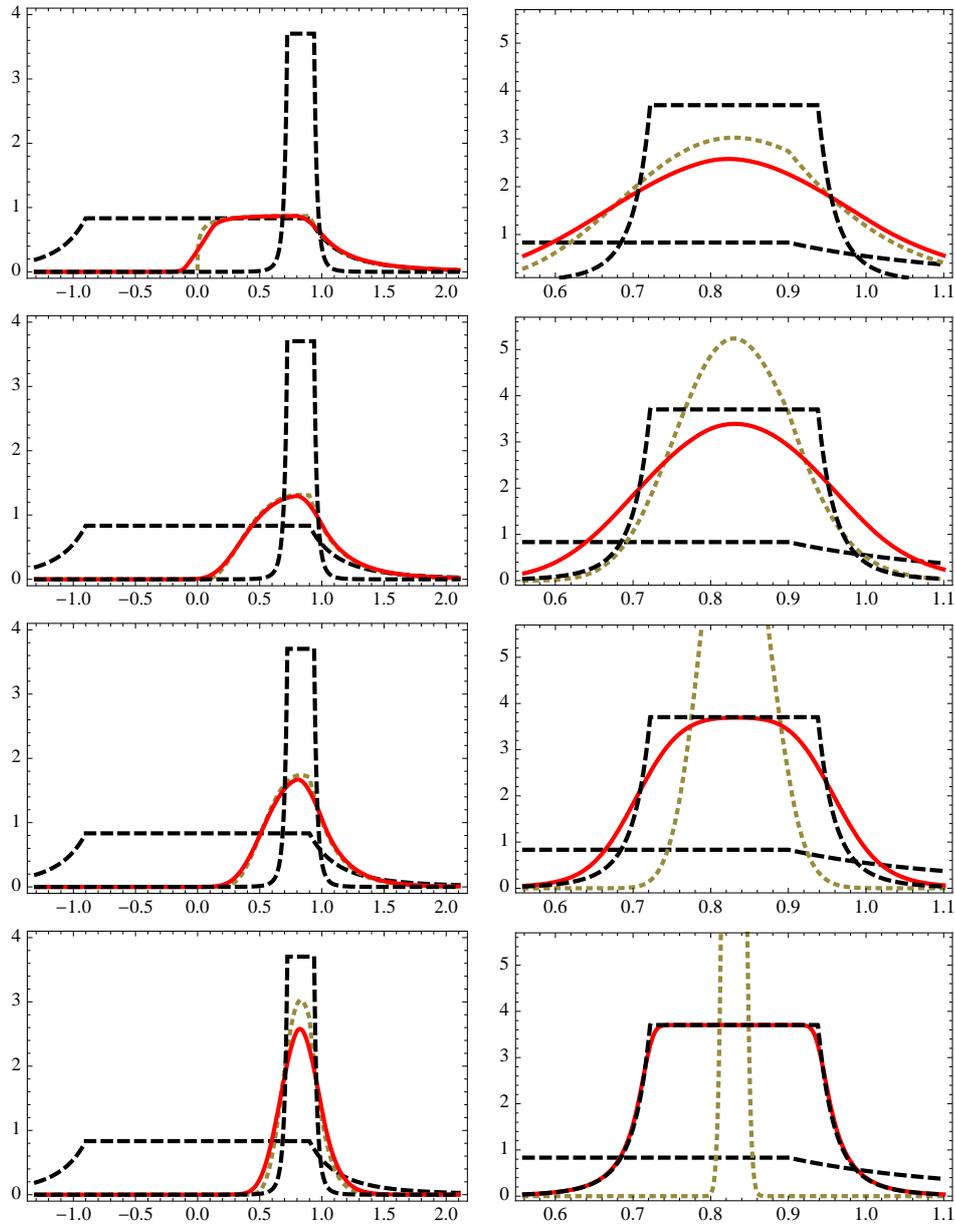


Figure 2.12 – Densité $f(c_1 + \alpha_s c_2 | c_0, \tilde{c}_1) \simeq f(\frac{\Delta_0}{\alpha_s} | c_0, \tilde{c}_1)$ (trait continu rouge) pour $c_0 = 0.9$, $\tilde{c}_1 = 0.83$, $\alpha_s = 0.12$ et une densité $f(\tilde{c}_1 | c_1)$ log-normale de paramètre de largeur $\ln f$ avec $f = 50; 2; 1,5; 1,2$, de haut en bas sur la gauche, et $f = 1,2; 1,1; 1,05; 1,01$, de haut en bas sur la droite. Les deux courbes en tirets larges noirs représentent respectivement $2f(c_1 | c_0) \simeq 2f(\frac{\Delta_0}{\alpha_s} | c_0)$ (courbe la plus étalée) et $f(c_1 + \alpha_s c_2 | c_0, c_1 = \tilde{c}_1) \simeq f(\frac{\Delta_0}{\alpha_s} | c_0, c_1 = \tilde{c}_1)$ (courbe la plus piquée). Elles correspondent aux limites de la densité étudiée dans lesquelles l'approximation \tilde{c}_1 est supposée de très mauvaise ou de très bonne qualité. La dernière courbe, en pointillés marrons, représente $f(c_1 | c_0, \tilde{c}_1)$. Elle coïncide avec $f(\frac{\Delta_0}{\alpha_s} | c_0, \tilde{c}_1)$ tant que l'incertitude sur c_1 est limitante.

Lorsque l'approximation de c_1 par \tilde{c}_1 est supposée de meilleure qualité, l'incertitude totale diminue. La confiance selon laquelle $\frac{\Delta_0}{\alpha_s}$ prendrait des valeurs autour de \tilde{c}_1 augmente (deux graphiques de gauche, au milieu). L'incertitude sur $\frac{\Delta_0}{\alpha_s}$ est toujours dominée par celle sur c_1 mais l'information fournie par \tilde{c}_1 n'est plus négligeable devant celle fournie par c_0 . Les densités $f(\frac{\Delta_0}{\alpha_s}|c_0, \tilde{c}_1)$ et $f(c_1|c_0, \tilde{c}_1)$ sont toujours indiscernables mais elles ne coïncident plus avec $f(c_1|c_0)$.

Pour f suffisamment petit, l'incertitude sur c_1 rejoint celle sur $\alpha_s c_2$ (à gauche en bas et à droite et en haut. Ce sont les mêmes graphiques pour des échelles différentes). La densité $f(\frac{\Delta_0}{\alpha_s}|c_0, \tilde{c}_1)$ se détache de $f(c_1|c_0, \tilde{c}_1)$. Sa largeur provient des incertitudes sur c_1 et $\alpha_s c_2$.

Pour une très bonne approximation \tilde{c}_1 de c_1 (graphique à droite en bas), l'incertitude sur c_2 domine. La différence entre \tilde{c}_1 et c_1 est négligeable. La densité totale prend la forme de $f(c_1 + \alpha_s c_2|c_0, c_1 = \tilde{c}_1)$ où le coefficient c_1 est supposé connu, de valeur \tilde{c}_1 .

Le modèle a donc exactement le comportement attendu. Il permet de modéliser la transition continue entre la méconnaissance d'un coefficient et sa connaissance complète. À l'utilisation, la difficulté provient bien entendu du choix de densité (2.113) et de modélisation/estimation de la qualité de l'approximation.

Bibliographie

- [1] S. Bethke, Eur. Phys. J. **C64** (2009) 689-703. [arXiv :0908.1135 [hep-ph]].
- [2] A. D. Martin, W. J. Stirling, R. S. Thorne, G. Watt, Eur. Phys. J. **C63** (2009) 189-285. [arXiv :0901.0002 [hep-ph]].
- [3] H. -L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, C. -P. Yuan, Phys. Rev. **D82** (2010) 074024. [arXiv :1007.2241 [hep-ph]].
- [4] R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, J. Rojo, M. Ubiali, Nucl. Phys. **B838** (2010) 136-206. [arXiv :1002.4407 [hep-ph]].
- [5] F. D. Aaron *et al.* [H1 and ZEUS Collaboration], JHEP **1001** (2010) 109 [arXiv :0911.0884 [hep-ex]].
- [6] S. Alekhin, J. Blumlein, S. Klein, S. Moch, Phys. Rev. **D81** (2010) 014032. [arXiv :0908.2766 [hep-ph]].
- [7] J. R. Andersen *et al.*, “The SM and NLO Multileg Working Group : Summary report”, Proceedings of the 6th Les Houches Workshop : Physics at TeV Colliders, Les Houches, France, 8-26 Jun 2009, arXiv :1003.1241 [hep-ph].
- [8] J. Baglio, A. Djouadi, JHEP **1010** (2010) 064. [arXiv :1003.4266 [hep-ph], arXiv :1009.1363 [hep-ph]].
- [9] M. Cacciari, N. Houdeau, JHEP **1109** (2011) 039. [arXiv :1105.5152 [hep-ph]].
- [10] Bruno de Finetti, Theory of Probability : A Critical Introductory Treatment, John Wiley, New York, 1974
- [11] O. Berger, Statistical Decision Theory and Bayesian Analysis, Springer-Verlag, 1985
- [12] M. Rubin, G. P. Salam, S. Sapeta, JHEP **1009** (2010) 084. [arXiv :1006.2144 [hep-ph]].
- [13] K. Nakamura *et al.* [Particle Data Group Collaboration], J. Phys. G **G37** (2010) 075021.

Conclusion

De manière générale, le formalisme des mesures finies peut être utilisé pour décrire des notions très différentes. Effectuer une distinction entre elles est bien souvent inutile car elles s'approchent généralement de près ou de loin à une notion, vague, de probabilité. Ce terme est donc employé pour désigner toute mesure finie en rapport avec une étude physique.

Dans certains cas, il faut pourtant séparer proprement l'outil mathématique de l'objet mesuré. Dans le cas du traitement d'incertitudes théorique, la distinction est indispensable car deux notions non équivalentes entrent en jeu. J'ai donc fait le choix d'utiliser des vocabulaires différents pour les mesures de confiance et de fréquence de résultat d'expérience, réservant au dernier cas le terme de probabilité pour me conformer à l'acceptation qui me semble la plus répandue du terme.

La clarification entre les notions étant faite, la construction du modèle de confiance sur l'erreur de troncature se déroule naturellement. Pourtant, il est important que ce dernier ne soit bien interprété que comme une première proposition de cadre formel de description pour cette source d'incertitude théorique. L'intérêt d'une mesure de confiance se mesurant à la qualité des informations utilisées dans sa construction, le modèle proposé peut être vu comme une coquille vide prête à accueillir des hypothèses qui restent encore à définir à l'aide d'arguments physiques.

Ainsi le modèle lui-même ne définit pas quels sont les coefficients sensés être du même ordre de grandeur. Il ne précise pas non plus de "véritable" coefficient d'expansion de la série perturbative. Les choix génériques qui ont été effectués ne sont motivés que par l'apparent succès de la méthode conventionnelle de la variation d'échelle. Il s'agit plus d'un exemple académique, utile pour présenter le formalisme, que d'un lot d'hypothèses abouties. En pratique, les choix à effectuer devront certainement être réévalués au cas par cas, de la même manière que les sources d'incertitude systématiques d'une expérience doivent être étudiées les unes après les autres.

L'architecture globale de la mesure devra elle-même sans aucun doute être retouchée plus ou moins en profondeur pour l'interfacer avec les estimations d'incertitudes des fonctions de distributions partoniques ou de fragmentation. Ce modèle apparaît donc finalement comme un nouveau jalon vers une description propre et contrôlée des incertitudes sur les prédictions théoriques en théorie quantique des champs.

Deuxième partie

Phénoménologie au LHC

Chapitre 3

Production de quarks lourds au LHC

Sommaire

3.1	Production et évolution de quarks lourds	104
3.1.1	Production d'un quark lourd. Factorisation	104
3.1.2	Production de quarks massifs	107
3.1.3	Fragmentation non perturbative	112
3.1.4	Désintégration	112
3.2	Prédictions théoriques FONLL pour le LHC	113
3.2.1	Paramètres utilisés	113
3.2.2	Incertitudes théoriques	114
3.2.3	Production de quarks lourds au LHC, 7 TeV	115

Introduction

La physique fondamentale a pour objectif de développer la compréhension du comportement du monde aux limites de la connaissance actuelle. Des outils théoriques sont développés et leurs prédictions sont testées expérimentalement. En physique des particules, les expériences considérées sont des collisions à hautes énergies. Les produits observés, leur nombre, leurs propriétés et leur fréquence d'apparition sont comparés aux prédictions proposées par les différents modèles ou théories.

Pour cela, les prédictions théoriques doivent être calculées. Alors que le nouveau collisionneur, le LHC, est enfin entré en activité, celles-ci doivent être estimées pour tous les modèles existants. C'est l'occasion de tester les nouveaux modèles proposés, mais aussi de mesurer plus finement les paramètres rentrant en jeu dans les anciens modèles et de repousser leur limite d'application.

Ce chapitre s'intéresse aux prédictions du Modèle Standard de production de quarks lourds dans des collisions à 7TeV. Il présente succinctement en section 3.1 le principe des différentes étapes de calcul des sections efficaces observées puis les résultats obtenus à l'aide de l'outil numérique FONLL dans la section 3.2.

3.1 Production et évolution de quarks lourds

3.1.1 Production d'un quark lourd. Factorisation

3.1.1.1 Quarks légers et quarks lourds

Un quark est une particule élémentaire fermionique (spin demi-entier) du Modèle Standard, sensible à l'interaction forte. Il en existe six, ordonnés en trois générations¹ (cf table 3.1). Les particules élémentaires fermioniques du Modèle Standard qui ne sont pas sensibles à l'interaction forte sont appelés des leptons.

u $m_u = 2,5_{-0,8}^{+0,6}$ MeV	c $m_c = 1,29_{-0,11}^{+0,05}$ GeV	t $m_t = 172,9 \pm 0,6 \pm 0,9$ GeV
d $m_d = 5,_{-0,9}^{+0,7}$ MeV	s $m_s = 100_{-20}^{+30}$ MeV	b $m_b = 4,19_{-0,06}^{+0,18}$ GeV

Tableau 3.1 – Quarks du Modèle Standard et leur masses (cf [1]).

L'interaction forte, décrite par la chromo-dynamique quantique ou *Quantum ChromoDynamics* (QCD), est portée par huit bosons vecteurs (spin entier) : les gluons. Elle présente la propriété de liberté asymptotique : à petites échelles, sa constante de couplage est faible. Le comportement des particules interagissantes est alors proche de celui de particules libres et il est bien décrit par un traitement perturbatif. À l'opposé, la constante de couplage diverge à grandes échelles. Sur les grands temps et les grandes distances ces quarks et gluons se regroupent jusqu'à former des systèmes liés, les hadrons, essentiellement insensibles à l'interaction forte. Ceux-ci sont classés en deux familles : les baryons, contenant trois quarks, et les mésons, contenant un quark et un anti-quark.

Les quarks et les gluons ne sont donc pas observables aux échelles macroscopiques en tant que particules libres. Ils n'interviennent que dans des processus rapides sur de petites distances comme les collisions entre hadrons (collisions *hadroniques*) ou entre leptons (collisions *leptoniques*). On les appelle collectivement des partons.

La position de la transition entre les petites échelles et les grandes échelles est estimée par le paramètre $\Lambda_{QCD} \sim 200$ MeV. Au premier ordre en perturbation, il est défini implicitement à l'aide de l'équation d'évolution de la constante de couplage forte α_s par :

$$\alpha_s(\mu) = \frac{2\pi}{\beta_0 \ln(\mu/\Lambda_{QCD})}, \quad \forall \mu \gg \Lambda_{QCD} \quad (3.1)$$

où β_0 est le premier terme du développement de la fonction beta en série de puissances de α_s :

$$\frac{d\alpha_s}{d \ln \mu^2} = \beta(\mu) \quad \beta(\mu) = - \sum_{n=0}^{\infty} \beta_n \alpha_s^n(\mu) \quad (3.2)$$

Les quarks u , d et s , de masse inférieure au paramètre Λ_{QCD} , sont appelés des quarks légers. L'effet de leur masse ne se fait ressentir qu'à des échelles où les quarks sont confinés. Décrits dans un domaine d'échelles perturbatives, ces particules ont le même comportement que si elles étaient de masse nulle. Les quarks c , b et t , de masse supérieure à Λ_{QCD} ,

1. Les masses des quarks présentées sont les masses $\overline{\text{MS}}$, excepté pour le quark top dont la masse est celle au pôle.

3.1 Production et évolution de quarks lourds

sont appelés des quarks lourds. L'effet de leur masse doit cette fois être pris en compte dans une étude perturbative de leur interactions. Le quark top a la particularité de se désintégrer en un quark plus léger avant d'avoir le temps de s'hadroniser. On s'intéresse dans ce chapitre plus particulièrement aux quarks lourds c et b qui forment bien des mésons observables lorsqu'ils sont produits dans des collisions.

3.1.1.2 Théorème de factorisation

PRINCIPE DE LA FACTORISATION

Depuis la création d'un quark lourd dans une collision jusqu'à l'observation des hadrons produits dans une collision, différents processus entrent en jeu sur différentes échelles de temps et d'énergie. Dans les premiers instants suivant la collision, les processus agissant sur des grands échelles de temps sont gelés. À l'inverse, au bout d'un temps long, les processus plus rapides sont moyennés. Les caractéristiques principales de l'évolution du système à un instant t sur un intervalle de temps dt sont dictées par les processus aux échelles² d'énergies $\frac{1}{t}$.

Ce principe devrait permettre de séparer l'étude de l'évolution du système après la collision en plusieurs étapes successives. Chacune d'entre elle correspondrait à des échelles et donc des phénomènes physiques différents. Les outils les décrivant pourraient ainsi être adaptés au cas par cas. Une partie de cette factorisation des étapes a été démontrée. Le reste du "théorème de factorisation" est en fait une hypothèse couramment utilisée.

La création de la particule et les premiers instants de son évolution sont décrits de manière perturbative. Il s'agit du processus dur. La collision des particules élémentaires incidentes crée le quark recherché, éventuellement virtuel. Le système se détend dans ce cas en émettant un ou plusieurs gluons jusqu'à atteindre sa couche de masse.

Sur les temps plus longs, les processus dictant l'évolution du système ne sont plus décrits par l'approche perturbative. La puissance de l'interaction forte sur les grandes distances confine les quarks et les gluons. Les particules émises dans la première phase se réordonnent pour former des mésons ou des baryons. Cette étape, dite d'hadronisation, doit être modélisée ou paramétrisée puis ajustée sur les résultats expérimentaux.

Les hadrons formés se désintègrent finalement et ce sont généralement les particules produites à ce stade que l'on observe.

"THÉORÈME" DE FACTORISATION

Les résultats obtenus se présentent sous la forme d'un jeu de probabilités suivant lequel une collision peut produire tel ou tel autre élément. Ces probabilités s'expriment sous la forme de sections efficaces notées génériquement σ . Le travail théorique vise à prédire leurs valeurs tandis que celui expérimental a pour objet de les observer. La validation de la théorie dépend de l'accord entre leurs résultats.

Lorsqu'elle est calculée théoriquement, la section efficace de production des produits finaux de désintégration est factorisée suivant les différents processus successifs mis en jeu. Elle s'écrit génériquement et symboliquement dans une collision leptonique :

$$\sigma \simeq \bar{\sigma} \otimes D \tag{3.3}$$

2. On se place dans un système d'unités où les constantes c et \hbar sont égales à un.

Le produit de convolution de l'équation (3.3) est défini par :

$$(f \otimes g)(x) = \int_0^\infty f(z) g\left(\frac{x}{z}\right) \frac{dz}{z} \quad (3.4)$$

Il est effectué sur la fraction x d'énergie de la collision emportée par le parton (gluon ou quark) considéré. Plus précisément, on utilise dans la suite la variable de Björken :

$$x = \frac{2p \cdot q}{q^2} \quad (3.5)$$

où p est la quadri-impulsion du parton considéré et q la quadri-impulsion du centre de masse de la collision.

Le terme $\bar{\sigma}$ décrit la collision elle-même et les tous premiers instants du système créé à ce moment. Il s'agit par exemple de la section efficace partonique, éventuellement corrigée (par exemple dans le schéma de factorisation $\overline{\text{MS}}$) auquel cas elle est notée $\hat{\sigma}$. La fonction de fragmentation D modélise des processus ultérieurs dont l'hadronisation. Une échelle μ_F modélise la séparation entre les temps courts, décrits par la section efficace du processus dur $\bar{\sigma}$, et les temps intermédiaires décrits par la fonction de fragmentation perturbative. Elle est appelée échelle de factorisation.

Le résultat peut enfin être multiplié par un rapport de branchement Br , ou convolué avec une fonction mesurant la fraction des hadrons produits qui se désintègrent dans les particules observées.

Dans le cas d'une collision hadronique, le processus dur ne s'effectue pas entre deux leptons mais entre deux partons provenant chacun de l'un des deux hadrons incidents. Les probabilités qu'il s'agisse de partons de telles ou telles autres natures et propriétés cinématiques doivent être prises en compte. Celles-ci sont comptabilisées à l'aide d'une convolution adéquate avec des fonctions F dites de distributions partonique. Le théorème de factorisation s'écrit alors symboliquement :

$$\sigma \simeq F \otimes \bar{\sigma} \otimes D \quad (3.6)$$

Plus précisément, la section efficace de production d'un hadron lourd H (*i.e.* contenant un quark lourd) portant une fraction x de l'énergie du faisceau dans la collision à l'énergie S de deux hadrons initiaux A et B d'impulsions P_A et P_B :

$$A(P_A) + B(P_B) \rightarrow H + X \quad (3.7)$$

s'écrit :

$$\frac{d\sigma}{dx}(x) = \sum_{i,j,k} \int dz_i dz_j \frac{dz}{z} \quad (3.8)$$

$$F_{i/A}(z_i; \mu_F) F_{j/B}(z_j; \mu_F) \left. \frac{d\bar{\sigma}}{dz} \right|_{ij \rightarrow kX} (z_i, z_j, z; \mu_F) D_{H/k}\left(\frac{x}{z}; \mu_F\right)$$

La somme porte d'une part sur les partons i et j provenant des hadrons A et B et d'impulsion $p_i = z_i P_A$ et $p_j = z_j P_B$ qui vont participer à la collision partonique à l'énergie $s = z_i z_j S$. Elle porte également sur les partons finaux k qui vont potentiellement donner un hadron H après une étape d'hadronisation pendant laquelle ils vont perdre une fractions x/z de leur énergie.

3.1 Production et évolution de quarks lourds



Figure 3.1 – Diagrammes de Feynman de l'ordre de Born (gauche) et d'émission virtuelle (droite) contribuant au premier ordre en α_s de la section partonique du processus $e^+e^- \rightarrow q\bar{q}(g)$.



Figure 3.2 – Diagrammes de Feynman d'émission réelle contribuant au premier ordre en α_s de la section partonique du processus $e^+e^- \rightarrow q\bar{q}(g)$.

Le processus dur dépend du détail de la collision entre les partons incidents. Les fonctions de fragmentation, de distribution partonique et les rapports de branchement devraient par contre en être indépendantes. Les processus physiques qu'ils décrivent ont lieu sur des échelles différentes, à des instants différents. Il s'agit du principe d'universalité. En pratique, les fonctions de fragmentation et les rapports de branchements sont par exemple ajustés dans certaines collisions bien connues, comme les collisions leptoniques, puis utilisés pour la prédictions des sections efficaces en collisions hadroniques.

3.1.2 Production de quarks massifs

3.1.2.1 Prédiction à ordre fixe en théorie des perturbations

SECTION EFFICACE DE PRODUCTION DE QUARKS MASSIFS

Un quark lourd Q peut être créé en collision leptonique, typiquement une collision électron/positron, ou hadronique (proton/proton ou proton/anti-proton,...). Les figures 3.1 et 3.2 présentent par exemple les diagrammes de Feynman à l'ordre un en perturbations en puissance de α_s dans le premier cas.

La section efficace du processus dur $\bar{\sigma} = \sigma_Q$ peut être calculée à tous les ordres. Elle est aujourd'hui connue jusqu'à l'ordre α_s^2 . La section efficace différentielle première en la variable de Björken x normalisée correspondant aux diagrammes des figures 3.1 et 3.2 est par exemple :

$$\begin{aligned} \frac{1}{\sigma_Q} \frac{d\sigma_Q}{dx} = & \delta(1-x) + \frac{\alpha_s(\hat{Q}^2)}{2\pi} \left(C_F + C_F \left[\ln \frac{\hat{Q}^2}{m^2} \left(\frac{1+x^2}{1-x} \right)_+ \right. \right. \\ & + 2 \frac{1+x^2}{1-x} \ln x - \left. \left. \left(\frac{\ln(1-x)}{1-x} \right)_+ (1+x^2) \right. \right. \\ & \left. \left. + \frac{1}{2} \left(\frac{1}{1-x} \right)_+ (x^2 - 6x - 2) + \left(\frac{2}{3}\pi^2 - \frac{5}{2} \right) \delta(1-x) \right] \right) + \mathcal{O}\left(\frac{m}{\hat{Q}}\right) \end{aligned} \quad (3.9)$$

L'expression exacte est connue mais les corrections en puissances de m/\hat{Q} ne sont pas détaillées ici pour simplifier l'expression. \hat{Q} est l'énergie du centre de masse de la collision au niveau du processus dur, m est la masse du quark lourd et $C_F = 4/3$ est le facteur de couleur. La distribution "plus" associée à une fonction f est définie sur l'ensemble des fonctions test g par la relation :

$$\int_0^1 [f(x)]_+ g(x) dx = \int_0^1 f(x) [g(x) - g(1)] dx \quad (3.10)$$

et par :

$$\forall x / 0 \leq x < 1, \quad [f(x)]_+ = f(x) \quad (3.11)$$

La section efficace complète de production de hadrons σ_H s'écrit comme le produit de la section efficace du processus dur et d'une fonction de fragmentation non perturbative décrivant entre autre l'hadronisation :

$$\sigma_H = \sigma_Q \otimes D_{NP} \quad (3.12)$$

Dans le cas d'une collision hadronique, une publication de Collins, Soper et Sterman [2] justifie plus précisément la factorisation des fonctions de distributions partoniques à des corrections en puissance de (Λ_{QCD}/m) près :

$$\sigma_{AB \rightarrow QX} = F_{i/A} \otimes F_{j/B} \otimes \sigma_{ij \rightarrow QX} + \mathcal{O} \left(\frac{\Lambda_{QCD}}{m} \right)^p \quad (3.13)$$

En notant σ_Q la section de production $\sigma_{ij \rightarrow QX}$ d'un quark lourd Q depuis les partons i et j , la section totale après hadronisation s'écrit symboliquement :

$$\sigma_H = F \otimes \sigma_Q \otimes D_{NP} \quad (3.14)$$

DOMAINE DE VALIDITÉ DE LA PRÉDICTION

Lorsque la collision partonique s'effectue à une énergie \hat{Q} proche de la valeur de la masse du quark produit, l'expérience montre que la section efficace approchée (3.9) décrit bien le processus dur. Par contre, lorsque l'énergie du centre de masse \hat{Q} est très supérieure à m , la présence du logarithme $\ln(\hat{Q}^2/m^2)$ montre que ce n'est plus le cas. Ce facteur provient de l'émission d'un gluon quasi-collinéaire au quark. Aux ordres supérieurs, des émissions collinéaires multiples fournissent des facteurs $\alpha_s^n \ln(\hat{Q}/m)^m$ qui peuvent être de l'ordre de l'unité ou plus grands. Leurs contributions ne sont plus négligeables et l'approximation de la section totale par les premiers ordres perturbatifs ne fonctionne plus.

La présence de ces logarithmes collinéaires signifie que lorsqu'il est produit dans une collision à haute énergie, et donc créé à haute virtualité, le quark lourd va émettre un très grand nombre de gluons pour se désexciter et atteindre finalement sa couche de masse. La description perturbative, où l'on considère l'émission d'un gluon comme une simple correction, ne peut pas fonctionner. L'émission de gluons est indispensable pour rendre compte fidèlement du comportement du système.

De la même manière, certains termes de l'expression (3.9) divergent dans la limite où la variable x tend vers un. Ceux-ci correspondent à l'émission de gluons de faible énergie que l'on appelle des gluons mous. Par conséquent, lorsque le quark considéré rassemble

3.1 Production et évolution de quarks lourds

la majorité de l'énergie de la collision, leur émission en grand nombre devient une part importante de la dynamique du système.

Des techniques ont été développées pour prendre en compte les contributions des émissions multiples, qu'elles soient collinéaires [3, 16, 5] ou molles [5, 6]. Dans chacun des deux cas, la somme de ces éléments se traduit par la somme d'une famille de facteurs $\alpha_s^n \ln()^m$, les logarithmes étant respectivement ceux de rapports \hat{Q}/m ou ceux de fonctions de x , ce qui vaut aux méthodes le nom de resommations des logarithmes. Dans le premier cas par exemple, la prise en compte des termes $\alpha_s^n \ln(\hat{Q}/m)^n$ (tous les gluons émis sont collinéaires) s'appelle la resommation aux logarithmes collinéaires dominants (*Leading Log*, LL). La resommation des termes en $\alpha_s^n \ln(\hat{Q}/m)^{n-1}$ (un gluon émis n'est pas collinéaire) est celle des *Next-to-Leading Log* (NLL) et ainsi de suite. L'objectif est de resommer à la fois les logarithmes collinéaires et les logarithmes mous de manière à décrire au mieux la dynamique du système.

La section suivante présente par exemple rapidement le principe de la resommation des logarithmes collinéaires à l'aide du formalisme de la fragmentation perturbative qui est implémentée dans l'outil de calcul de sections efficaces FONLL [7, 8].

3.1.2.2 Processus dur de production de quark non massif et fragmentation perturbative

QUARKS LOURDS PRODUITS À HAUTE ÉNERGIE ET QUARKS LÉGERS

La description perturbative de la production de quarks présentée dans le paragraphe précédent est spécifique aux quarks lourds Q . Lorsque l'énergie du centre de masse de la collision partonique est trop élevée et que cette approche ne fonctionne plus, il est toujours possible de se tourner vers la méthode utilisée dans le cas des quarks légers q .

En effet, la masse d'un quark léger est plus basse que la limite Λ_{QCD} du domaine perturbatif de l'interaction forte. Or les collisions considérées se situent généralement à des énergies \hat{Q} qui en sont bien au-delà. L'énergie du centre de masse de la collision partonique est donc toujours plus grande que la masse des quarks légers produits. Le problème est donc similaire à la production de quarks lourds dans des collisions à très haute énergie.

Lorsque l'énergie du centre de masse \hat{Q} est bien plus élevée que la masse m des quarks produits, le temps caractéristique des processus de la collision $1/\hat{Q}$ est beaucoup plus court que celui de masse des quarks $1/m$. Les effets de masse ne se font pas sentir à ces échelles courtes : ils sont gelés. Tout se passe comme s'ils n'existaient pas, comme si leur temps caractéristique $1/m$ était infini. Autrement, dit, tout se passe comme si la masse des quarks était nulle.

Ainsi, les premiers instants après la création de l'état virtuel dans le processus dur, la dynamique du système peut être étudiée comme si le quark considéré était de masse nulle, qu'il soit en fait lourd et produit à très haute énergie ou simplement léger.

SECTION EFFICACE DE PRODUCTION DE QUARKS NON MASSIFS

L'approche perturbative sur les temps de l'ordre de $1/\hat{Q}$ est valide. Le système nouvellement créé n'a pas encore eu le temps d'émettre un grand nombre de gluons. La section efficace du processus dur σ_q de production d'un quark léger q , considéré de masse nulle, peut être calculée à un ordre fini. Elle devrait représenter correctement les premiers instants après la collision.

Par contre, ce résultat ne peut pas décrire les processus physiques sur des temps plus longs. D'une part, les effets des masses réelles des quarks lourds vont se faire ressentir sur des temps de l'ordre de $1/m$. De l'autre, le calcul de la section précédente montre qu'un grand nombre de gluons va être émis à mesure que le quark virtuel se désexcite : la description perturbative à un ordre faible ne peut pas fonctionner.

Le calcul des éléments de matrice contribuant à la section efficace s'effectue pourtant sur des temps allant jusqu'à l'infini. La section σ_q contient donc une description incorrecte des temps longs. Un contre-terme est alors utilisé pour retirer de la section efficace σ_q , pour factoriser, cette partie. La section efficace du processus dur est alors la section corrigée (ou "soustraite") $\bar{\sigma} = \hat{\sigma}$. Plusieurs méthodes, ou schémas, de factorisation sont possibles. Le plus courant est le schéma de factorisation $\overline{\text{MS}}$.

Dans le cas de la production de quarks non massifs dans une collision électron-positron, la section efficace partonique différentielle première en la variable x de Björken de production de quarks non massifs s'écrit à l'ordre α_s comme la somme de deux contributions, dites virtuelle (provenant des diagrammes 3.1) et réelle (diagrammes 3.2) :

$$\frac{d\sigma_q}{dx} = \frac{d\sigma_q^{(r)}}{dx} + \frac{d\sigma_q^{(v)}}{dx} \quad (3.15)$$

Celle-ci diverge (cf chapitre 4). Elle ne peut donc pas, en effet, être une bonne description de la production de quarks. La section différentielle corrigée par le contre-terme :

$$\frac{d\hat{\sigma}}{dx} = \frac{d\sigma_q^{(r)}}{dx} + \frac{d\sigma_q^{(v)}}{dx} + \frac{d\hat{\sigma}^{(c)}}{dx} \quad (3.16)$$

est par contre finie. Les expressions de ces contributions sont présentées au chapitre 4.

FONCTION DE FRAGMENTATION PERTURBATIVE

Dans cette approche, la bonne description des temps jusqu'à $1/m$ est rajoutée sous la forme d'une nouvelle fonction de fragmentation dite perturbative :

$$\sigma_H = \underbrace{\hat{\sigma} \otimes D_P}_{\simeq \sigma'_Q} \otimes D'_{NP} \quad (3.17)$$

La factorisation de la production de quark lourd, à des corrections en puissances près, en une section efficace corrigée de production de quarks non massifs et en une fonction de fragmentation universelle :

$$\sigma'_Q = \hat{\sigma} \otimes D_P + \mathcal{O}\left(\frac{m}{\hat{Q}}\right) \quad (3.18)$$

a cette fois-ci été prouvée. Il s'agit du théorème de factorisation des singularités de masse [5]. Les corrections en puissances de m/\hat{Q} ne sont généralement pas obtenues par cette approche et elles sont négligées en pratique (eq. 3.17).

La factorisation (3.18) est effectuée à une échelle μ_F proche de l'échelle de collision du processus dur \hat{Q} . De cette manière, la section efficace $\hat{\sigma}$ n'est utilisée que pour décrire que des temps jusqu'à $1/\mu_F \sim 1/\hat{Q}$ après la collision. Elle relègue le reste du comportement du domaine perturbatif, à savoir l'évolution du système des temps de l'ordre de $1/\mu_F$ à $1/m$, à la fonction D_P d'où son nom.

La fonction de fragmentation perturbative doit alors décrire les deux effets physiques présentés plus tôt et ayant lieu dans ces temps. Elle doit prendre en compte l'émission

3.1 Production et évolution de quarks lourds

multiple de gluons provenant de la désexcitation de l'état virtuel initialement créé, de l'échelle μ_F proche de l'énergie du centre de masse de collision \hat{Q} à une échelle μ_{F_0} , supérieure mais proche de la masse du quark. En second lieu, la fonction de fragmentation doit également contenir l'effet de la masse m du quark elle-même. En effet, tant que l'échelle étudiée est grande devant m , *i.e.* supérieure à μ_{F_0} , le quark peut être considéré comme étant de masse nulle. Son émission de gluons de désexcitation est importante. Celle-ci s'estompe entre μ_{F_0} et m lorsque l'état virtuel se rapproche de la couche de masse : la masse du quark agit comme un *cutoff* sur l'émission de gluons.

La "resommation des logarithmes collinéaires", traduction calculatoire de la prise en compte de l'émission multiple de gluons collinéaires de désexcitation, doit donc être effectuée dans l'expression de la fonction de fragmentation perturbative. Puisque cette dernière est universelle, le calcul peut être réalisé une seule fois. La fonction de fragmentation peut par la suite être réutilisée dans l'étude de nombreux processus différents. Il suffit d'utiliser à chaque fois la section efficace corrigée de production de quark non massif $\hat{\sigma}$ associée au processus étudié.

La resommation est effectuée à l'aide des équations d'évolution d'Altarelli-Parisi. Elle permet d'exprimer la fonction de fragmentation perturbative à l'échelle μ_F en fonction de celle à l'échelle μ_{F_0} et un terme comprenant l'ensemble des émissions multiples. La "condition initiale" en μ_{F_0} est ensuite obtenue en comparant les expressions de la section efficace de production de quarks non massifs σ_q et massifs σ_Q . Puisque les échelles comprises entre μ_F et m sont plus grandes que Λ_{QCD} , toutes ces manipulations peuvent être effectuées analytiquement à l'aide d'expressions perturbatives [5].

3.1.2.3 Comparaison des deux approches

L'approche perturbative directe de la production de quarks lourds de la section 3.1.2.1 et l'approche par la fonction de fragmentation perturbative de la section 3.1.2.2 sont valides pour différentes valeurs d'énergie de collision partonique \hat{Q} . Leurs expressions :

$$\sigma_H = \sigma_Q \otimes D_{NP} \quad \sigma_H = \sigma'_Q \otimes D'_{NP} \quad (3.19)$$

sont différentes car les approximations effectuées ne sont pas les mêmes dans les deux cas. En particulier, lorsque l'énergie du centre de masse de la collision partonique est proche de la masse d'un quark, la première méthode décrit proprement la collision dure. L'expression de la section efficace de production de quark lourd σ'_Q est incomplète : il lui manque des contributions en puissances de m/\hat{Q} . Pour des ordres de grandeur de \hat{Q} plus élevés, la seconde méthode est mieux adaptée. L'émission multiple de gluons est une part importante de la dynamique du système et elle n'est pas décrite par le calcul à ordre fixe σ_Q .

Dans les deux cas, le processus physique étudié est pourtant le même : la création d'un état virtuel et sa désexcitation jusqu'à la couche de masse du quark. Il serait plus agréable d'avoir à disposition une unique expression de la section efficace du processus dur valable sur toute la gamme des énergies de collision \hat{Q} , par exemple en extrapolant les résultats (3.19) sur la zone intermédiaire.

En pratique, une expression générale est en fait nécessaire pour étudier les collisions hadroniques. En effet, l'énergie du centre de masse \hat{Q} du processus dur est exactement connue dans une collision électron-positron car elle vient directement des valeurs des quadri-impulsions des deux particules incidentes. En collision hadronique, ce n'est plus

le cas. Les particules intervenant dans le processus dur ne sont pas les hadrons (protons, anti-protons) eux-mêmes mais des partons les constituants. Leurs quadri-impulsions ne sont pas déterminées. Le calcul de la section efficace fait alors intervenir toutes les expressions des processus durs pour toutes les valeurs possibles de \hat{Q} via le produit de convolution de l'équation (3.6) avec les fonctions de distribution partonique F . Les deux expressions, dans les cas de grande ou de petite énergie de collision partonique \hat{Q} , doivent être obligatoirement recollées pour des valeurs intermédiaires de manière à obtenir une expression valide sur toute la gamme d'énergies de collision partonique. Cette opération s'appelle le *matching* [7].

3.1.3 Fragmentation non perturbative

3.1.3.1 Fonction de fragmentation non perturbative

Les processus physiques ayant lieu sur des temps caractéristiques d'au plus $1/m$ étant décrits perturbativement, la fonction de fragmentation non perturbative représente les effets sur des temps plus longs. Ils sont supposés universels puis extraits des collisions les plus nettes : les collisions électrons-positrons.

Pour simplifier cette extraction, la fonction de fragmentation peut être paramétrée. La forme historiquement proposée par Kartvelishvili *et al.* [9] est utilisée dans la suite :

$$D_{NP}(x; \alpha) = (\alpha + 1)(\alpha + 2)x^\alpha(1 - x) \quad (3.20)$$

3.1.3.2 Extraction dans l'espace de Mellin

L'extraction s'effectue dans l'espace de la transformée de Mellin. Celle-ci est définie comme étant l'unique transformée intégrale représentant le produit de convolution de l'équation 3.4 par le produit simple des transformées (cf chapitre 4) :

$$\mathcal{M}[f \otimes g] = \mathcal{M}[f] \cdot \mathcal{M}[g] \quad (3.21)$$

Ces transformées sont des fonctions de la variable complexe N .

Ainsi, la transformée de Mellin de la section efficace complète de production de hadrons dans des collisions hadronique peut s'écrire comme le produit de transformées de Mellin de la section du processus dur et de la fonction de fragmentation non perturbative :

$$\mathcal{M}[\sigma_H] = \mathcal{M}[\sigma_Q] \mathcal{M}[D_{NP}] \quad (3.22)$$

Le membre de gauche est mesuré expérimentalement et le premier terme de droite est obtenu théoriquement. Le jeu de paramètres de la fonction D_{NP} est alors ajusté sur ces données [10].

3.1.4 Désintégration

Les hadrons lourds créés ne sont pas stables. Les éléments observés peuvent être des produits de leur désintégration. La probabilité qu'un hadron lourd se désintègre en un jeu précis de particule est appelée rapport de branchement (Br). Ceux-ci sont obtenus expérimentalement.

3.2 Prédictions théoriques FONLL pour le LHC

m (GeV)	4,5	4,75	5
α	26,7	24,2	22,2

Tableau 3.2 – Valeurs du paramètre α de la fonction de fragmentation non-perturbative D_{NP} pour les différentes valeurs de masse du quark b utilisées.

3.2 Prédictions théoriques FONLL pour le LHC

L’outil FONLL a été appliqué avec succès pour fournir des prédictions de production de quarks b [10, 11] et c [12] au Tevatron (collisions $p\bar{p}$ à 1TeV) et de quarks lourds à RHIC [13]. La section efficace à ordre fixe du processus dur utilisée pour les prédictions à basse énergie de collision partonique est déterminée à l’ordre NLO. Les logarithmes collinéaires sont resommés à l’ordre NLL.

La section efficace étudiée dans le cas du LHC est la section différentielle première en l’impulsion transverse p_T des mésons lourds D (contenant un quark c) et B (contenant un quark b). Les paramètres utilisés pour la prédiction sont les suivants.

3.2.1 Paramètres utilisés

La prédiction centrale de FONLL est calculée en la valeur centrale de la masse du quark considéré et du jeu de PDF utilisé, ainsi qu’en fixant les échelles de renormalisation et de factorisation égales à la masse transverse, $\mu_{R,F} = \mu_0 = \sqrt{p_T^2 + m^2}$, où m est la masse du quark lourd considéré.

3.2.1.1 Mésons B

Pour la production de quark b , la masse $m_b = 4,75$ GeV est utilisée. Les cas $m_b = 4,5$ GeV et $m_b = 5$ GeV sont également explorés pour estimer l’erreur possible provenant de l’incertitude sur la valeur exacte de la masse du quark.

La fonction de fragmentation non perturbative D_{NP} est de la forme (3.20). Les valeurs du paramètre α pour les différents choix de masse sont présentées dans le du tableau 3.2. Elles ont été extraites dans [14] par ajustement sur les données en collisions leptonique du LEP [15, 16] dans l’espace de Mellin en $N = 5$. Ces données correspondent à la production de mésons B, tous types confondus. Les données séparées pour les cas B^0 et B^\pm ne sont pas disponibles et l’on est forcé de supposer des fonctions de fragmentations similaires dans tous ces cas.

La fonction (3.20) est normalisée à l’unité. Utilisée pour un état spécifique, le ratio de branchement approprié (*e.g.* $b \rightarrow B^+$) est ajouté.

3.2.1.2 Mésons D

Dans le cas de la production de quarks c , la situation est plus complexe. D’une part, les données expérimentales pour les mésons individuels (D^* , D^\pm , D^0 et \bar{D}^0 , D_s) sont disponibles. De l’autre, les différences et similarités de la fragmentations d’un quark lourd en un méson pseudoscalaire (D) ou vecteur (D^*) sont au moins partiellement théoriquement comprises [17]. Cette compréhension peut donc être exploitée dans le but de minimiser le nombre de paramètres à extraire des données expérimentales.

Dans ce but, dans la référence [12], les fonctions de fragmentation non perturbatives dans différents mésons D ont été construits exclusivement en fonction de la fragmentation $c \rightarrow D^*$, dont le paramètre est extrait des données e^+e^- d'ALEPH [18]. La désintégration de D^* en états D a été modélisée théoriquement, et les différents ratios de branchement ont été extraits des données. La production primaire de mésons D depuis la fragmentation de quarks c a été décrite en terme du même paramètre non perturbatif ajusté aux données $c \rightarrow D^*$, bien qu'une fonction différente, obtenue en [17], a été utilisée. Les résultats détaillés pour les fonctions de fragmentation non perturbatives $c \rightarrow D^*$, $c \rightarrow D^+$ et $c \rightarrow D^0$ sont données dans les équations (10), (9) et (5) de [12] respectivement. Elles dépendent d'un seul paramètre non perturbatif r déterminé depuis les données.

Les valeurs $r = 0,1$ pour $m_c = 1,5$ GeV, $r = 0,06$ pour $m_c = 1,3$ GeV et $r = 0,135$ pour $m_c = 1,7$ GeV ont été utilisées. Elles correspondent comme dans le cas des mésons B à un ajustement dans l'espace de Mellin en $N = 5$.

3.2.2 Incertitudes théoriques

La méthode d'évaluation des incertitudes utilisée dans ce chapitre n'est pas en lien avec le travail présenté en première partie. Les incertitudes théoriques sont estimées à l'aide de la variation combinée des deux échelles non physiques μ_R et μ_F , de la variation des masses des quarks lourds, et à l'aide de l'incertitude associée aux valeurs des fonctions de distribution partoniques. Ces trois incertitudes sont combinées en quadrature :

$$d\sigma_{-\Delta-}^{+\Delta+} \quad (3.23)$$

où :

$$\Delta_{\pm} = \sqrt{\Delta_{\pm,echelles}^2 + \Delta_{\pm,masse}^2 + \Delta_{\pm,PDF}^2} \quad (3.24)$$

Les incertitudes sur la fragmentation non perturbative ou sur la désintégration ne sont pas prises ici en compte. Certaines évaluations ont montré qu'elles n'étaient pas significatives.

INCERTITUDE D'ÉCHELLES

Pour éviter une compensation accidentelle entre les dépendances en μ_F et μ_R de la section efficace qui pourrait apparaître en gardant les deux échelles égales, on calcule l'incertitude d'échelles non physiques en les variant indépendamment l'une de l'autre sur l'intervalle $0,5 < \xi_{R,F} < 2$, avec la contrainte $0,5 < \xi_R/\xi_F < 2$, où $\xi_{R,F} = \mu_{R,F}/\mu_0$. En pratique, la section efficace est évaluée (en utilisant les valeurs centrales de la masse et des PDFs) aux sept points :

$$(\xi_R, \xi_F) = \{(1; 1), (0,5; 0,5), (2; 2), (0,5; 1), (1; 0,5), (2; 1), (1; 2)\} \quad (3.25)$$

et l'enveloppe des résultats est prise. Celle-ci définit, en tout point de la distribution considérée, les deux extrêmes :

$$d\sigma_{-\Delta-,echelles}^{+\Delta+,echelles} \quad (3.26)$$

INCERTITUDE DE MASSE

Les trois valeurs de masse $m_c = 1,5; 1,3; 1,7$ GeV sont utilisées pour le quark c et $m_b = 4,75; 4,5; 5$ GeV pour b . Les paramètres non perturbatifs sont ajustés pour chaque masse à leur valeur appropriée et la section est évaluée en $\xi_{R,F} = 1$ avec le jeu central de PDFs. L'enveloppe est prise, définissant :

$$d\sigma_{-\Delta-,masse}^{+\Delta+,masse} \quad (3.27)$$

3.2 Prédictions théoriques FONLL pour le LHC

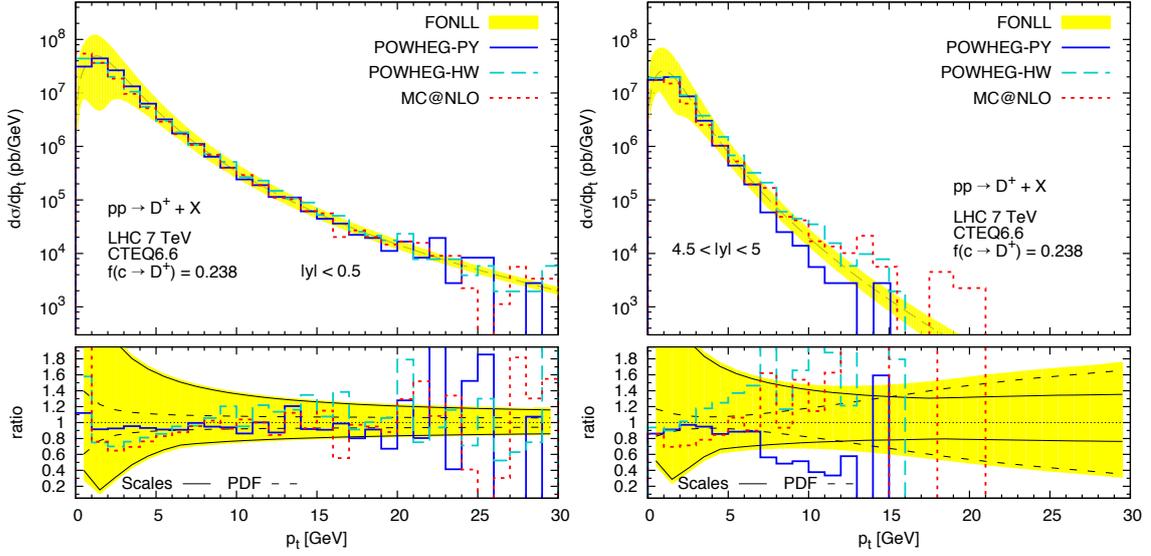


Figure 3.3 – Distribution du moment transverse de D^+ produits. Graphique de gauche : région de rapidité centrale $|y| < 0,5$. Graphique de droite : région de grande rapidité $4,5 < |y| < 5$.

INCERTITUDE DE FONCTION DE DISTRIBUTION PARTONIQUE

L'incertitude des PDFs, lorsqu'elle est disponible, est évaluée de la manière suggérée par le jeu spécifique utilisé :

$$d\sigma_{-\Delta-,PDF}^{+\Delta+,PDF} \quad (3.28)$$

Le jeu utilisé par défaut est CTEQ 6.6 [19].

3.2.3 Production de quarks lourds au LHC, 7 TeV

3.2.3.1 Production de mésons D

La production de mésons D^+ est considérée sur la figure 3.3. Une fraction quark $c \rightarrow D^+$ de fragmentation de 0,238 est utilisée. La prédiction FONLL est représentée en tirets au centre de sa bande d'incertitude jaune. Les prédictions fournies par trois outils Monte Carlo, POWHEG avec PYTHIA et POWHEG avec HERWIG [20, 21, 22] ainsi que MC@NLO [23] sont également représentées. Toutes les prédictions sont largement compatibles dans la région de rapidité centrale $|y| < 0,5$ alors que quelques différences apparaissent dans la région de grande rapidité $4,5 < |y| < 5$. L'incertitude FONLL est élargie dans cette zone, en grande partie à cause d'une contribution beaucoup plus importante des PDFs.

La production de méson D^0 , présentée en 3.4 est très similaire.

3.2.3.2 Production de mésons B

La production de méson B^+ est présentée sur la figure 3.5. La qualité de la compatibilité entre les différentes prédictions est comparable au cas de la production de D^+ . Les

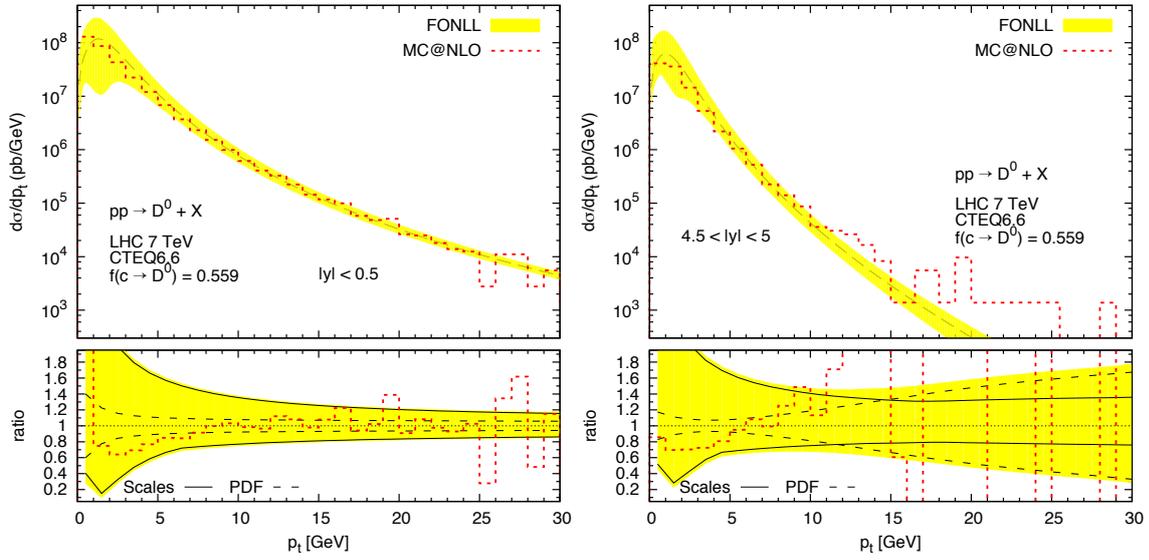


Figure 3.4 – Distribution du moment transverse de D^0 produits. Graphique de gauche : région de rapidité centrale $|y| < 0,5$. Graphique de droite : région de grande rapidité $4,5 < |y| < 5$.

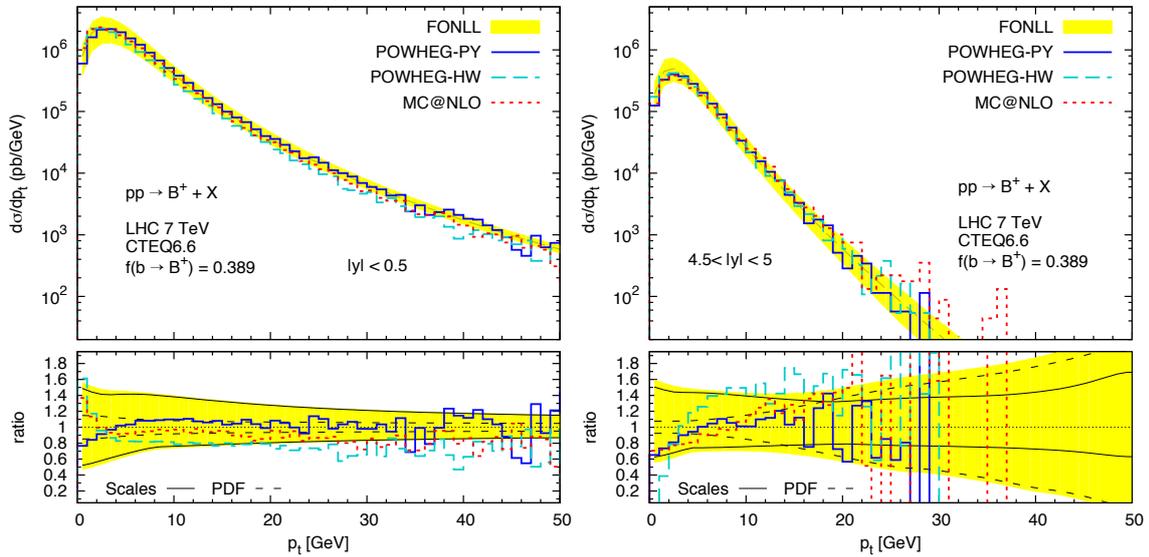


Figure 3.5 – Distribution du moment transverse de B^+ produits. Graphique de gauche : région de rapidité centrale $|y| < 0,5$. Graphique de droite : région de grande rapidité $4,5 < |y| < 5$.

3.2 Prédictions théoriques FONLL pour le LHC

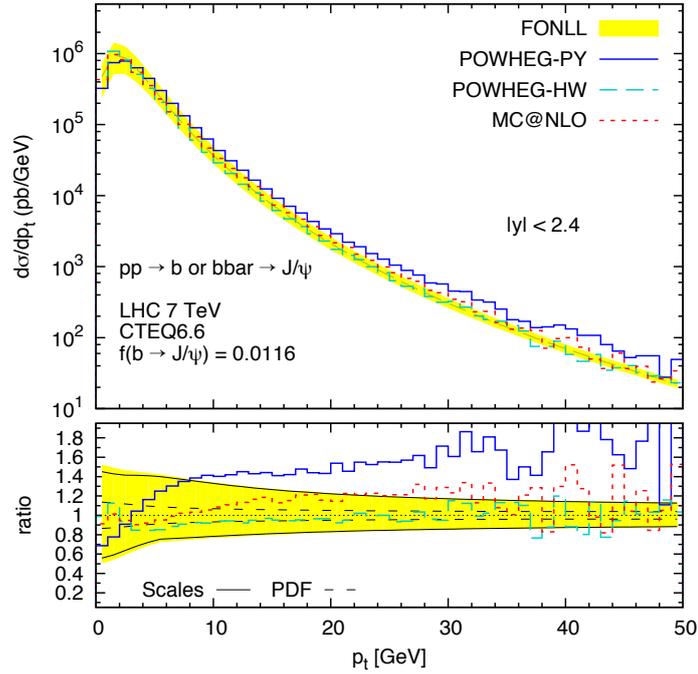


Figure 3.6 – Distribution du moment transverse de J/Ψ produits dans la région de rapidité centrale $|y| < 2,4$.

autres caractéristiques, comme l’élargissement de l’incertitude à large rapidité en conséquence d’une incertitude plus large provenant des PDFs, sont similaires.

3.2.3.3 Production $B \rightarrow J/\Psi$

La figure 3.6 compare les prédictions pour la distribution de moment transverse pour la production “non prompt” de J/Ψ (*i.e.* venant de la désintégration d’un méson contenant un quark b) dans la région de rapidité centrale $|y| < 2,4$. Dans ce cas, la prédiction de l’implémentation POWHEG-PY diffère de manière importante des autres et les données expérimentales pourraient être discriminantes. Des mesures expérimentales à 7 TeV de CMS, LHCb et ATLAS ont été comparées à FONLL et ont été trouvées être en bon accord, bien qu’une inspection précise des différentes prédictions dans le but de voir si quelques unes d’entre elles sont peut être défavorisées n’a pas encore été effectuée.

Bibliographie

- [1] K. Nakamura *et al.* (Particle Data Group), *J. Phys. G* **37**, 075021 (2010)
- [2] J. C. Collins, D. E. Soper and G. F. Sterman, *Nucl. Phys. B* **263** (1986) 37.
- [3] B. Mele, P. Nason, *Phys. Lett.* **B245** (1990) 635-639.
- [4] B. Mele, P. Nason, *Nucl. Phys.* **B361** (1991) 626-644.
- [5] M. Cacciari and S. Catani, *Nucl. Phys. B* **617** (2001) 253 [arXiv :hep-ph/0107138].
- [6] M. Cacciari and E. Gardi, *Nucl. Phys. B* **664** (2003) 299 [arXiv :hep-ph/0301047].
- [7] M. Cacciari, M. Greco and P. Nason, *JHEP* **9805** (1998) 007 [arXiv :hep-ph/9803400].
- [8] M. Cacciari, S. Frixione and P. Nason, *JHEP* **0103** (2001) 006 [arXiv :hep-ph/0102134].
- [9] V. G. Kartvelishvili, A. K. Likhoded, V. A. Petrov, *Phys. Lett.* **B78** (1978) 615.
- [10] M. Cacciari, P. Nason, *Phys. Rev. Lett.* **89** (2002) 122003. [hep-ph/0204025].
- [11] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, *JHEP* **0407** (2004) 033 [arXiv :hep-ph/0312132].
- [12] M. Cacciari and P. Nason, *JHEP* **0309** (2003) 006 [arXiv :hep-ph/0306212].
- [13] M. Cacciari, P. Nason and R. Vogt, *Phys. Rev. Lett.* **95** (2005) 122001 [arXiv :hep-ph/0502203].
- [14] M. Cacciari, P. Nason and C. Oleari, *JHEP* **0604** (2006) 006 [arXiv :hep-ph/0510032].
- [15] A. Heister *et al.* [ALEPH Collaboration], *Phys. Lett. B* **512** (2001) 30 [arXiv :hep-ex/0106051].
- [16] G. Abbiendi *et al.* [OPAL Collaboration], *Eur. Phys. J. C* **29** (2003) 463 [arXiv :hep-ex/0210031].
- [17] E. Braaten, K. m. Cheung, S. Fleming and T. C. Yuan, *Phys. Rev. D* **51** (1995) 4819 [arXiv :hep-ph/9409316].
- [18] R. Barate *et al.* [ALEPH Collaboration], *Eur. Phys. J. C* **16** (2000) 597 [arXiv :hep-ex/9909032].
- [19] P. M. Nadolsky *et al.*, *Phys. Rev. D* **78** (2008) 013004 [arXiv :0802.0007 [hep-ph]].
- [20] P. Nason, *JHEP* **0411** (2004) 040 [arXiv :hep-ph/0409146].
- [21] S. Frixione, P. Nason and C. Oleari, *JHEP* **0711** (2007) 070 [arXiv :0709.2092 [hep-ph]].
- [22] S. Alioli, P. Nason, C. Oleari and E. Re, *JHEP* **1006** (2010) 043 [arXiv :1002.2581 [hep-ph]].
- [23] S. Frixione and B. R. Webber, *JHEP* **0206** (2002) 029 [arXiv :hep-ph/0204244].

Chapitre 4

Annulation numérique des divergences infrarouges et collinaires

Sommaire

4.1	Annulation numérique des divergences infrarouges et collinaires	124
4.1.1	Principe de l'annulation	124
4.1.2	Cas du processus $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$	124
4.1.3	Manipulation des distributions	128
4.2	Annulation numérique des divergences dans l'espace de Mellin	129
4.2.1	Transformée de Mellin	129
4.2.2	Cas du processus $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$	130
4.3	Limitations et utilité de l'approche	136
4.3.1	Dépendance explicite en ϵ	136
4.3.2	Prolongement analytique dans l'espace conjugué	136
4.3.3	Transformée de Mellin en k dimensions et position des pôles . . .	137
4.3.4	Domaine d'application et utilité de la méthode	137

Introduction

Le calcul des prédictions théoriques souffre de deux types de difficultés : la complexité mathématique des expressions rencontrées et l'augmentation factorielle du nombre de diagrammes de Feynman à évaluer. Du fait du second point il est aujourd'hui indispensable d'avoir recours à l'outil informatique. Malheureusement, certaines étapes de calcul comme les prolongements analytiques résistent au passage au numérique. Un travail particulier est donc nécessaire pour adapter les calculs et fournir des algorithmes implémentables. Pour une revue des difficultés rencontrées dans l'automatisation des calculs de sections efficaces, voir *e.g.* [1].

Deux difficultés mathématiques particulières limitent l'automatisation des calculs des sections efficaces. D'une part, les intégrales présentes dans les expressions des amplitudes des diagrammes de Feynman comportant des boucles sont complexes et non référencées. Différentes méthodes telles que l'algorithme de Passarino-Veltman [2] ou les méthodes unitaires [3, 4] ont été développées pour automatiser l'évaluation des amplitudes à une boucle. Un panel d'outils tels que les identités d'intégration par partie [5] ou la transformée de Mellin-Barnes [6] sont également disponibles pour le calcul des amplitudes à deux boucles ou plus.

D'autre part, l'intégration sur l'espace des phases des amplitudes des différentes contributions d'une section efficace se heurte à l'émergence de divergences d'origine collinéaire et infrarouge. Le calcul ne peut donc pas être effectué par une simple évaluation Monte Carlo. Deux méthodes en particulier, la méthode de découpe de l'espace des phases [7, 8, 9] et la méthode de soustraction [10, 11, 12, 13, 14], permettent de traiter les divergences avant d'effectuer l'intégrale sur l'espace des phases. La deuxième approche permet par exemple d'ajouter à chaque contribution un terme retirant les comportements singuliers. Ainsi, les intégrales obtenues étant bien définies, elles peuvent être évaluées numériquement. Les termes singuliers sont de plus universels, si bien que leur génération peut être automatisée [15].

Si ces dernières méthodes permettent d'effectuer l'intégration sur l'espace des phases au premier ordre en perturbation, leur application devient plus difficile aux ordres supérieurs. Ce chapitre présente une tentative de numériser intégralement le traitement des divergences infrarouges et collinéaires en utilisant les propriétés de la transformée de Mellin, ce qui aurait pu être une voie vers le traitement des ordres supérieurs. La première section présente le principe de l'annulation numérique des divergences, dans l'espace classique et la seconde section dans l'espace de Mellin. Une discussion des limites de l'approche est présentée dans la dernière section.

4.1 Annulation numérique des divergences infrarouges et collinéaires

4.1.1 Principe de l'annulation

Les différentes contributions d'une section efficace divergent lorsque le paramètre de régularisation, ϵ en régularisation dimensionnelle, tend vers zéro. Pour des valeurs faibles mais non nulles, elles prennent des valeurs finies presque partout. La section efficace complète, éventuellement exclusive, est donc également finie en ϵ non nul. Elle peut y être évaluée numériquement.

Contrairement à celui des différentes contributions, le développement de la section efficace complète en puissances de ϵ autour de zéro est un développement de Taylor : il ne présente pas de pôles. La section totale prend des valeurs finies en ϵ non nul et tend vers une valeur également finie lorsque ϵ tend vers zéro. Cette dernière peut donc être extrapolée à partir des valeurs de la section en ϵ non nul.

Cette extrapolation peut s'effectuer de manière numérique. La section différentielle complète est évaluée en N valeurs de ϵ non nulles, ces valeurs étant proche de zéro. La dépendance en ϵ est ensuite ajustée par un polynôme de degré $N - 1$. La valeur réelle de la section est enfin approximée par celle du polynôme en ϵ nul. Dans le cas particulier où un seul point est utilisé, la section efficace complète est simplement approchée par sa valeur en un ϵ non nul, proche de zéro.

Cette approche est totalement numérique. Les développements en séries de Laurent des différentes contributions ne sont plus nécessaires. La méthode s'applique à toutes les sections efficaces, quelles que soient les expressions des sections partielles. Il suffit que ces expressions puissent être évaluées numériquement.

4.1.2 Cas du processus $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$

CONTRIBUTIONS À LA SECTION PARTONIQUE DIFFÉRENTIELLE

La toute première partie de la thèse a été consacrée au calcul de la section efficace partonique différentielle première du processus $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ à l'ordre α_s en la variable de Björken définie par :

$$x = \frac{2p \cdot q}{q^2} \quad (4.1)$$

où p est la quadri-impulsion de l'un des quarks réels émis et q la quadri-impulsion du centre de masse partonique. Le calcul est effectué pour des quarks non massifs.

La section efficace se compose de trois parties : une composante virtuelle, une composante réelle et un contre-terme de factorisation des divergences collinéaires. La composante virtuelle est issue des diagrammes de Feynman de la figure 3.1. En régularisation dimensionnelle à $D = 4 - 2\epsilon$ dimensions, son expression est :

$$\frac{d\sigma_\epsilon^{(v)}}{dx} = \left(\frac{4 - 2\epsilon}{3 - 2\epsilon} - 2\epsilon \right) \frac{\pi\alpha^2}{\hat{Q}^2} \left[\left(\frac{\hat{Q}^2}{4\pi\mu^2} \right)^{-\epsilon} 2^{2\epsilon} \frac{\sqrt{\pi}/2}{\Gamma(\frac{3}{2} - 2\epsilon)} \right] \left(1 - \left(\frac{\hat{Q}^2}{4\pi\mu^2} \right)^{-\epsilon} \frac{C_F\alpha_S}{\pi} \cos(\epsilon\pi) F_\epsilon \right) \delta(1 - x) \quad (4.2)$$

4.1 Annulation numérique des divergences infrarouges et collinéaires

\hat{Q} est l'énergie du centre de masse de la collision au niveau partonique. C_F est le Casimir quadratique de $SU(3)$. Il est égal à $\frac{4}{3}$. Enfin, le facteur F_ϵ est défini par :

$$F_\epsilon = 1 + \Gamma(-\epsilon) \frac{\Gamma(1+\epsilon)}{\Gamma(2-2\epsilon)} \left(\Gamma(-\epsilon) + \frac{1}{2}\Gamma(1-\epsilon) \right) \quad (4.3)$$

La composante réelle provient des diagrammes de Feynman de la figure 3.2. Toujours en régularisation dimensionnelle, son expression différentielle seconde en (x_1, x_2) , les variables de Björken associées aux deux quarks réels émis, est :

$$\begin{aligned} \frac{d^2\sigma_\epsilon^{(r)}}{dx_1 dx_2} = & \frac{4 - 4\epsilon C_F \alpha^2 \alpha_S}{3 - 2\epsilon} \frac{1}{\hat{Q}^2} \frac{1}{2} \left[\left(\frac{\hat{Q}^2}{4\pi\mu^2} \right)^{-2\epsilon} \frac{1}{\Gamma(2-2\epsilon)} \right] ((1-x_1)(1-x_2)(x_1+x_2-1))^{-\epsilon} \\ & (1-\epsilon) \left((1-\epsilon) \frac{x_1^2 + x_2^2}{(1-x_1)(1-x_2)} - 2\epsilon \left(2 - \frac{x_1 x_2}{(1-x_1)(1-x_2)} \right) \right) \end{aligned} \quad (4.4)$$

Après intégration sur l'une des deux variables, l'expression différentielle première de la contribution partonique est :

$$\begin{aligned} \frac{d\sigma_\epsilon^{(r)}}{dx} = & \frac{4 - 4\epsilon C_F \alpha^2 \alpha_S}{3 - 2\epsilon} \frac{1}{\hat{Q}^2} \frac{1}{2} \left[\left(\frac{\hat{Q}^2}{4\pi\mu^2} \right)^{-2\epsilon} \frac{\Gamma(1-\epsilon)\Gamma(2-\epsilon)}{\Gamma(2-2\epsilon)^2} \right] x^{-2\epsilon} \\ & \frac{1}{-\epsilon} \frac{(2\epsilon^2 - 3\epsilon + 1) + 2\epsilon(2-\epsilon)x + (\frac{1}{2}\epsilon^2 - \frac{7}{2}\epsilon + 1)x^2}{[1-x]^{1+\epsilon}} \end{aligned} \quad (4.5)$$

Enfin, le contre-terme de régularisation des divergences collinéaires dûes à l'approximation de masse nulle est :

$$\frac{d\hat{\sigma}_\epsilon^{(c)}}{dx} = \left(\frac{4-2\epsilon}{3-2\epsilon} - 2\epsilon \right) \frac{C_F \alpha^2 \alpha_S}{\hat{Q}^2} \frac{1}{2} \left[\left(\frac{\hat{Q}^2}{4\pi\mu^2} \right)^{-\epsilon} 2^{2\epsilon} \frac{\sqrt{\pi}/2}{\Gamma(\frac{3}{2}-2\epsilon)} \frac{1}{\Gamma(1-\epsilon)} \right] \frac{1}{\epsilon} \left[\frac{1+x^2}{1-x} \right]_+ \quad (4.6)$$

Les distributions "plus" sont définies par les prescriptions présentées au chapitre précédent en (3.10).

La section totale différentielle première est la limite en ϵ nul de la somme des trois contributions :

$$\frac{d\hat{\sigma}}{dx} = \lim_{\epsilon \rightarrow 0} \left(\frac{d\sigma_\epsilon^{(v)}}{dx} + \frac{d\sigma_\epsilon^{(r)}}{dx} + \frac{d\hat{\sigma}_\epsilon^{(c)}}{dx} \right) \quad (4.7)$$

En tout x différent de zéro ou de un, l'expression limite est régulière et s'écrit :

$$\begin{aligned} \frac{d\hat{\sigma}}{dx} = & \frac{4 C_F \alpha^2 \alpha_S}{3} \frac{1}{\hat{Q}^2} \frac{1}{2} \\ & \left(\left(\ln \frac{\hat{Q}^2}{\mu^2} + \ln(1-x) + 2 \ln x - \frac{3}{4} \right) \frac{1+x^2}{1-x} + \frac{1}{4}(9+5x) - \frac{7}{2}(1+x) + 3 \right) \end{aligned} \quad (4.8)$$

COMPORTEMENTS DES DIFFÉRENTES CONTRIBUTIONS SUR]0; 1[

La contribution virtuelle (4.2) à la section efficace différentielle première du processus partonique $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ est nulle pour toute valeur de la variable de Björken x dans]0; 1[. La valeur de la section différentielle complète sur cet intervalle est la somme de

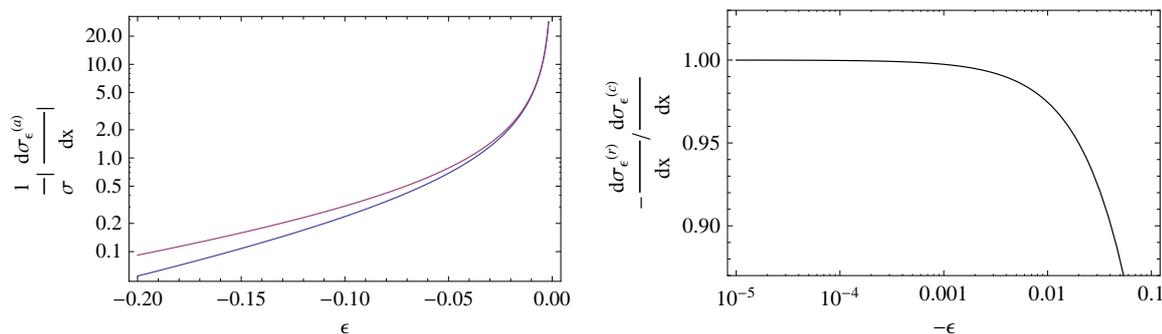


Figure 4.1 – À gauche : contribution réelle (4.5) (bleu) et valeur absolue du contre-terme collinéaire (4.6) (rouge) à la section efficace différentielle du processus partonique $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ en la variable de Björken x , normalisé par la section totale σ , en $x = 0,5$ fixé, en fonction du paramètre de régularisation dimensionnel ϵ . À droite : rapport de la contribution réelle et de l’opposé du contre-terme collinéaire en fonction de ϵ et en $x = 0,5$ fixé.

la contribution réelle (4.5) et du contre-terme (4.6). Chacun de ces termes diverge lorsque le paramètre de régularisation ϵ tend vers zéro. La figure 4.1 montre les comportements de ces deux contributions en $x = 0,5$ en fonction de ϵ . Les divergences se compensent au voisinage de zéro.

ANNULATION NUMÉRIQUE

La valeur exacte de la section efficace différentielle partonique peut être estimée sur $]0; 1[$ par sa valeur en ϵ non nul, proche de zéro. La précision de l’approximation augmente lorsque la valeur de ϵ utilisée se rapproche de zéro. La figure 4.2 présente le rapport de la valeurs de la section différentielle complète en $\epsilon = -10^{-2}$ et de sa valeur exacte calculée depuis l’expression (4.8). La précision atteinte est de plus de 7% sur la plus grande partie de l’intervalle.

Le paramètre ϵ ne peut pas être choisi arbitrairement petit. Lorsque sa valeur s’approche de zéro, les différentes contributions divergent et leur estimation numérique devient instable. La figure 4.3 présente une fois encore le rapport de la valeur de la section efficace différentielle partonique à la valeur exacte (4.8) atteinte lorsque le paramètre de régularisation ϵ est nul. La variable de Björken x est maintenant fixée à 0,5 et le rapport est calculé en fonction de la valeur du paramètre ϵ . Le calcul numérique devient instable pour des valeurs de ϵ inférieures à 10^{-10} .

AMÉLIORATION DE LA PRÉCISION

L’approximation de la valeur exacte de la section différentielle partonique par la valeur de cette section en ϵ non nul peut être vue comme le cas particulier d’une extrapolation polynômiale par un polynôme de degré nul. De manière plus générale, les valeurs de la section en N valeurs de ϵ distinctes permettent de définir un polynôme d’extrapolation de degré $N - 1$ (extrapolation de Lagrange). La valeur exacte de la section différentielle est approximée par la valeur de ce polynôme en ϵ nul.

4.1 Annulation numérique des divergences infrarouges et collinéaires

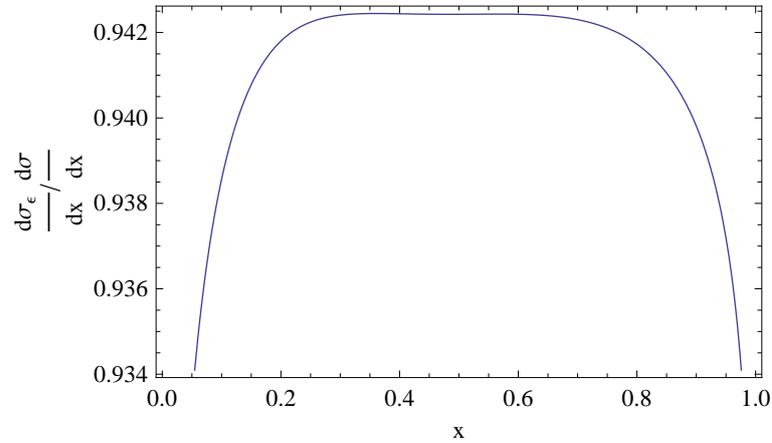


Figure 4.2 – Évaluation numérique du rapport de la valeur de la section efficace différentielle du processus partonique $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ en la variable de Björken x en $\epsilon = -10^{-2}$ et de sa valeur exacte (*i.e.* en $\epsilon = 0$) (4.8) sur l'intervalle $]0; 1[$.

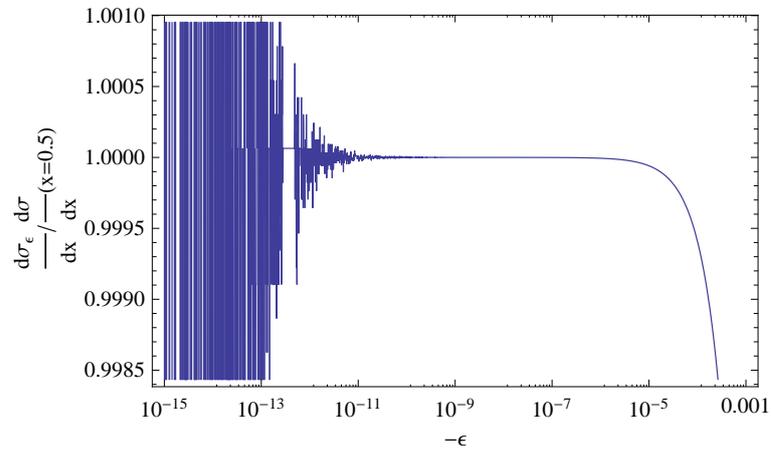


Figure 4.3 – Évaluation numérique du rapport de la valeur de la section efficace différentielle du processus partonique $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ en la variable de Björken x et de sa valeur exacte (*i.e.* en $\epsilon = 0$) (4.8), en $x = 0,5$ fixé, en fonction du paramètre de régularisation ϵ .

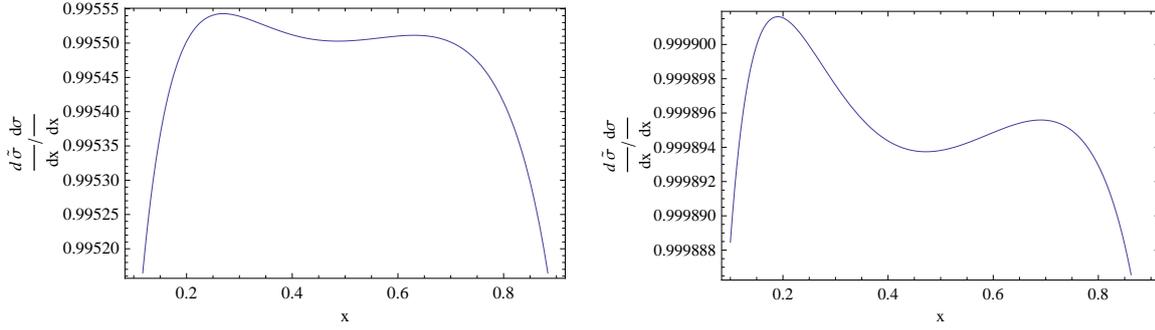


Figure 4.4 – Évaluations numériques des rapports de la valeur de la section efficace différentielle du processus partonique $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ en la variable de Björken x extrapolée en $\epsilon = 0$ et de sa valeur exacte (*i.e.* en $\epsilon = 0$) (4.8) sur l'intervalle $]0; 1[$. L'extrapolation est effectuée par un polynôme de Lagrange de degré un (gauche) ou deux (droite) à partir des points d'abscisse $\epsilon \{-3.10^{-2}; -10^{-2}\}$ et $\{-3.10^{-2}; -1, 5.10^{-2}; -10^{-2}\}$ respectivement.

La précision de l'extrapolation s'améliore lorsque le degré du polynôme utilisé augmente. La figure 4.4 présente les rapport de l'estimation de la section et de sa valeur exacte lorsque le polynôme d'extrapolation est de degré un ou deux. Les valeurs des abscisses ϵ des points utilisés pour définir les polynômes d'extrapolation sont respectivement $\{-3.10^{-2}; -10^{-2}\}$ et $\{-3.10^{-2}; -1, 5.10^{-2}; -10^{-2}\}$.

Le temps de calcul de l'approximation augmente avec le nombre de points utilisés. Pour chaque degré ajouté au polynôme, les contributions partielles doivent notamment être évaluées en une valeur de ϵ supplémentaire. On cherche à limiter le degré utilisé pour la précision demandée. La précision requise peut être atteinte en suivant l'algorithme schématique suivant :

- Diminuer la valeur du paramètre ϵ tant que la précision recherchée n'est pas atteinte ;
- S'assurer de la stabilité du calcul numérique des différentes contributions à chaque modification de ϵ ;
- Si le calcul devient instable avant que la précision requise ne soit atteinte, augmenter le nombre de points d'extrapolation utilisé.

4.1.3 Manipulation des distributions

Cette méthode s'applique lorsque les expressions des différentes contributions de la section efficace peuvent être évaluées numériquement. Elle ne s'applique pas lorsque ces contributions sont des distributions. Plus précisément, la méthode ne permet d'évaluer la section que sur le support non singulier de ses contributions, c'est à dire aux points où celles-ci coïncident avec des fonctions régulières.

Cette limite est contraignante. Les sections efficaces physiques sont obtenues par convolution des sections partoniques avec des fonctions de distribution partonique et de fragmentation. Les sections efficaces partoniques différentielles dans lesquelles l'annulation des divergences doit être effectuée sont généralement elles-mêmes des distributions.

Les différentes contributions à la section efficace différentielle première du processus partonique $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ sont par exemple singulières aux extrémités de leurs intervalles de définition, en $x = 0$ et $x = 1$. La section différentielle partonique com-

4.2 Annulation numérique des divergences dans l'espace de Mellin

plète (4.7) est une distribution. De plus, les sections efficaces physiques, obtenues par convolution avec des fonctions de fragmentations, dépendent du comportement singulier de cette distribution. La section efficace intégrée dépend elle-même de la contribution virtuelle alors que celle-ci est purement singulière puisque proportionnelle à une distribution de Dirac. Le comportement singulier de la section différentielle ne peut pas être négligé.

Les distributions présentes dans les expressions des différentes contributions de la section efficace différentielle du processus $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ sont tempérées à support compact. Leurs transformées de Fourier sont des fonctions régulières. La méthode proposée d'annulation numérique des divergences peut donc être appliquée dans l'espace conjugué. Puisque la connaissance de la transformée de Fourier d'une distribution tempérée est équivalente à la connaissance de la distribution elle-même, la transformée de Fourier d'une section efficace différentielle, après annulation des divergences, contient l'information de son comportement en tant que distribution. On utilise en pratique la transformée de Mellin en lieu et place de la transformée de Fourier puisqu'elle est adaptée à la procédure de factorisation.

4.2 Annulation numérique des divergences dans l'espace de Mellin

4.2.1 Transformée de Mellin

4.2.1.1 Définition

La transformée de Mellin est usuellement définie sur les fonctions $f(x)$, $x \in \mathbb{R}^+$. Elle leur associe des fonctions \tilde{f} de la variable complexe définies par :

$$\mathcal{M}[f](n) = \tilde{f}(n) = \int_0^\infty f(x)x^{n-1} dx \quad (4.9)$$

En particulier, la transformée de Mellin en $n = 1$ de la fonction f est, si elle existe, son intégrale sur \mathbb{R}^+ . L'intégrale (4.9) est convergente sur une bande de l'espace complexe délimitée par deux droites parallèles à l'axe imaginaire :

$$\mathcal{B}_D = \{n \in \mathbb{C} / n = n_R + i\omega, n_R \in]n_m; n_M[\} \quad (4.10)$$

Cette bande est appelé la bande de définition de la transformée de f . La transformée $\mathcal{M}[f](n)$ y est holomorphe. Les distributions considérées dans la suite sont à support compact, inclu dans $[0; 1]$. La définition (4.9) de leur transformée de Mellin se réécrit :

$$\mathcal{M}[f](n) = \tilde{f}(n) = \int_0^1 f(x)x^{n-1} dx \quad (4.11)$$

La transformée de Mellin inverse est définie par :

$$\mathcal{M}^{-1}[\tilde{f}](x) = f(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} x^{-n} \tilde{f}(n) dn \quad (4.12)$$

Le paramètre c est un réel inclu dans la bande de définition de la transformée de f . Les transformées de Mellin de deux fonctions différentes peuvent avoir la même expression analytique. Leurs bandes de définition sont alors disjointes. Suivant que le paramètre c de l'équation (4.12) appartient à l'une ou l'autre des bandes, la transformée de Mellin inverse aura pour image l'une des deux fonctions plutôt que l'autre.

4.2.1.2 Convolution de Mellin

Le produit de convolution de Mellin est défini comme au chapitre précédent par :

$$(f \otimes g)(x) = \int_0^\infty f(z) g\left(\frac{x}{z}\right) \frac{dz}{z} \quad (4.13)$$

La transformée de Mellin est l'unique transformée intégrale vérifiant la propriété :

$$\mathcal{M}[f \otimes g] = \mathcal{M}[f]\mathcal{M}[g] \quad (4.14)$$

Le produit de convolution des distributions à support inclus dans $[0; 1]$ s'écrit :

$$(f \otimes g)(x) = \int_0^1 f(z) g\left(\frac{x}{z}\right) \frac{dz}{z} \quad (4.15)$$

4.2.2 Cas du processus $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$

4.2.2.1 Transformée de Mellin

CONTRIBUTION RÉELLE

La transformée de la contribution réelle (4.5) de la section efficace différentielle s'écrit sur sa bande de définition :

$$\sigma_\epsilon^{(r)}(n) = \int_0^1 \frac{d\sigma_\epsilon^{(r)}}{dx} x^{n-1} dx \quad (4.16)$$

L'intégrale converge pour toute valeur de n réelle strictement supérieure à 2ϵ . La bande de définition de la transformée de Mellin de la contribution réelle a donc pour trace sur l'axe réel l'intervalle $]2\epsilon; +\infty[$. L'expression de la transformée est holomorphe :

$$\sigma_\epsilon^{(r)}(n) = \left(\frac{2-2\epsilon}{3-2\epsilon}\right) \frac{C_F \alpha^2 \alpha_S}{\hat{Q}^2} \left[\left(\frac{\hat{Q}^2}{4\pi\mu^2}\right)^{-2\epsilon} \frac{\Gamma(2-\epsilon)}{\Gamma(2-2\epsilon)^2} \frac{\Gamma(-\epsilon)^2 \Gamma(n-2\epsilon)}{\Gamma(2+n-3\epsilon)} \right. \\ \left. \left[2n(1+n) - 5\left(1 + \frac{n}{2}(5+n)\right)\epsilon + \left(21 + \frac{n}{2}(25+n)\right)\epsilon^3 + 8\epsilon^4 \right] \right] \quad (4.17)$$

Les pôles de cette fonction de la variable complexe n sont situés sur l'axe réel négatif, en $n = -k - 2 + 3\epsilon$ et $n = -k + 2\epsilon$, $k \in \mathbb{N}$.

CONTRIBUTION VIRTUELLE

La contribution virtuelle (4.2) est proportionnelle à une distribution de Dirac. La forme intégrale de sa transformée de Mellin est définie pour tout n complexe :

$$\sigma_\epsilon^{(v)}(n) = \int_0^1 \frac{d\sigma_\epsilon^{(v)}}{dx} x^{n-1} dx \quad (4.18)$$

Sa bande de définition est l'espace complexe tout entier. La transformée de Mellin de la contribution virtuelle est le coefficient de la distribution de Dirac de son expression (4.2) :

$$\sigma_\epsilon^{(v)}(n) = \left(\frac{4-2\epsilon}{3-2\epsilon} - 2\epsilon\right) \frac{\pi\alpha^2}{\hat{Q}^2} \left[\left(\frac{\hat{Q}^2}{4\pi\mu^2}\right)^{-\epsilon} 2^{2\epsilon} \frac{\sqrt{\pi}/2}{\Gamma(\frac{3}{2}-2\epsilon)} \right] \\ \left[1 - \left(\frac{\hat{Q}^2}{4\pi\mu^2}\right)^{-\epsilon} \frac{C_F \alpha_S}{\pi} \cos(\epsilon\pi) F_\epsilon \right] \quad (4.19)$$

4.2 Annulation numérique des divergences dans l'espace de Mellin

F_ϵ est l'expression définie par l'équation (4.3). La transformée de Mellin de la contribution virtuelle est une fonction constante de la variable complexe n .

CONTRE-TERME DE FACTORISATION

L'intégrale définissant la transformée de Mellin du contre-terme (4.6) :

$$\sigma_\epsilon^{(c)}(n) = \int_0^1 \frac{d\sigma_\epsilon^{(c)}}{dx} x^{n-1} dx \quad (4.20)$$

est définie sur le demi-plan complexe de partie réelle positive. L'expression de la transformée après intégration est définie dans le plan complexe entier :

$$\sigma_\epsilon^{(c)}(n) = \left(\frac{4-2\epsilon}{3-2\epsilon} - 2\epsilon \right) \frac{C_F \alpha^2 \alpha_S}{2\hat{Q}^2} \left[\left(\frac{\hat{Q}^2}{4\pi\mu^2} \right)^{-\epsilon} 2^{2\epsilon} \frac{\sqrt{\pi}/2}{\Gamma(3/2-2\epsilon)} \right] \frac{1}{\epsilon\Gamma(1-\epsilon)} \left[-2 \sum_{i=1}^{n+1} \frac{1}{i} + \frac{2n+1}{n(n+1)} + \frac{3}{2} \right] \quad (4.21)$$

Cette fonction de la variable complexe n est également holomorphe. Elle présente deux pôles en $n = 0$ et $n = -1$.

SECTION TOTALE

La méthode conventionnelle d'annulation des divergences par développement en série fournit l'expression de la transformée de Mellin de la section totale réelle, *i.e.* en ϵ nul. Celle-ci permet dans la suite d'estimer la qualité de l'annulation numérique proposée. L'expression de la section totale dans l'espace de Mellin n'est pas reproduite ici mais elle peut être trouvée dans [16].

4.2.2.2 Annulation des divergences

COMPORTEMENT DES DIFFÉRENTES CONTRIBUTIONS

Les transformées de Mellin des trois contributions en ϵ non nul sont des fonctions numériques sur le plan complexe privé des positions des pôles. Chacune des trois composantes diverge lorsque le paramètre de régularisation s'approche de zéro. La figure 4.5 présente les comportements de la somme de la contribution réelle et du contre-terme d'une part et de l'opposé de la contribution virtuelle de l'autre dans l'espace de Mellin, normalisés dans les deux cas par la valeur de la section intégrée σ , en tant que fonctions de ϵ et en $n = 1$. Les divergences des trois contributions se compensent.

ANNULATION DES DIVERGENCES

L'annulation numérique des divergences dans l'espace de Mellin est similaire à celle effectuée dans l'espace initial. Les différentes contributions sont évaluées en N valeurs de ϵ non nulles en un point n fixé du plan complexe. Les résultats sont sommés :

$$\forall \epsilon, \quad \sigma_\epsilon(n) = \sigma_\epsilon^{(r)}(n) + \sigma_\epsilon^{(v)}(n) + \sigma_\epsilon^{(c)}(n) \quad (4.22)$$

Les valeurs de la section totale $\sigma_\epsilon(n)$ sont utilisées pour ajuster sa dépendance en ϵ aux alentours de zéro. La valeur de la transformée de Mellin de la section totale $\sigma(n)$ est approchée par la valeur du polynôme extrapolateur en ϵ nul.

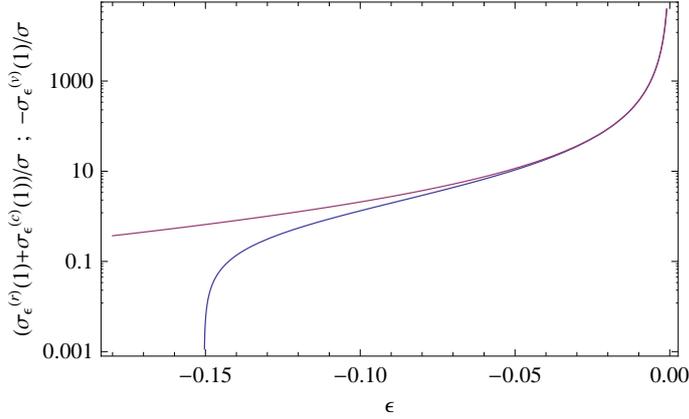


Figure 4.5 – Somme des transformées de Mellin de la contribution réelle et du contre-terme collinéaire (bleu) et transformée de l’opposé de la contribution virtuelle (rouge) du processus $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ dans le voisinage de $\epsilon = 0$ à $n = 1$ fixé. Les quantités tracées sont normalisées par la section efficace intégrée σ .

En première approximation, pour $N = 1$, la transformée de la section efficace différentielle $\sigma(n)$ est égale à la transformée de la section en ϵ non nul, mais proche de zéro :

$$\sigma(n) \simeq \sigma_\epsilon(n) \quad (4.23)$$

La partie gauche de la figure 4.6 présente la dépendance en ϵ du rapport de ces deux quantités en $n = 1, 10, 100$ et 1000 . De la même manière que dans l’espace direct, le calcul de la transformée $\sigma_\epsilon(n)$ est instable lorsque le paramètre ϵ prend des valeurs trop petites. La partie droite de la figure présente l’estimation numérique du même rapport en $n = 1$ pour des valeurs très faibles de ϵ .

EXTRAPOLATION POLYNÔMIALE D’ORDRE SUPÉRIEUR

La précision de l’approximation (4.23) obtenue en $n = 1$ avant d’atteindre l’instabilité est largement suffisante pour les applications phénoménologiques. Ce ne serait pas nécessairement vrai dans le cas d’un autre processus. Pour illustrer l’utilisation d’une extrapolation d’ordre supérieur, les intégrales des transformées de Mellin (4.17) et (4.20) de la contribution réelle et du contre-terme sont estimées numériquement. La figure 4.7 présente de nouveau les sommes des transformées de la contribution réelle et du contre-terme collinéaire d’une part et l’opposé de la transformée de la contribution virtuelle de l’autre. La compensation des divergences échoue pour des valeurs de ϵ beaucoup plus éloignées de zéro.

Les figures 4.8 et 4.9 présentent les sommes des transformées numériques des différentes contributions, normalisées par leur valeur exacte fournie par [16], ainsi que les extrapolations d’ordre un et deux en fonction de ϵ pour différentes valeurs entières de n . Les approximations (4.23) d’ordre zéro sont éloignées de 15 à 22% de la valeur exacte. Les polynômes d’ordre deux fournissent à l’opposé des valeurs à moins de 2% d’erreur.

4.2 Annulation numérique des divergences dans l'espace de Mellin

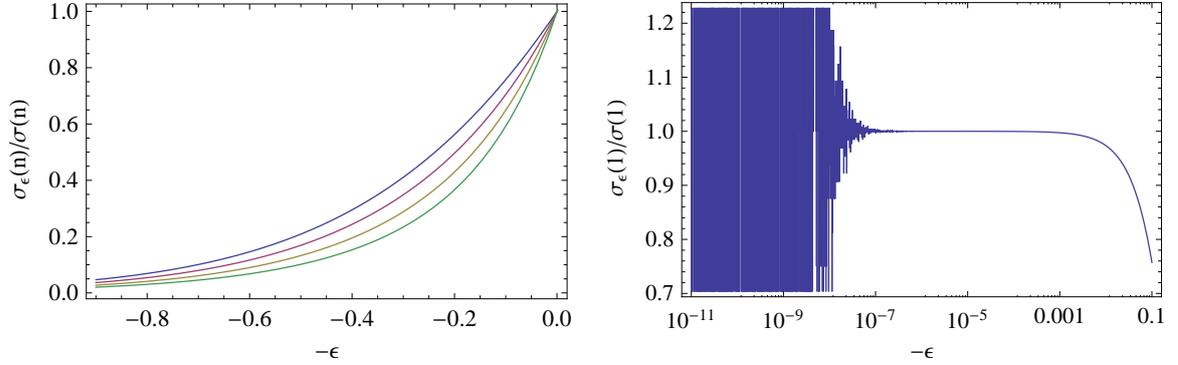


Figure 4.6 – Évaluation numérique du rapport des valeurs des transformées de Mellin de la section efficace différentielle du processus partonique $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ en la variable de Björken x en ϵ non nul et de sa valeur exacte (*i.e.* en $\epsilon = 0$) au voisinage de $\epsilon = 0$. À gauche, les transformées de Mellin sont évaluées en $n = 1, 10, 100, 1000$ de haut en bas. À droite, elle le sont en $n = 1$.

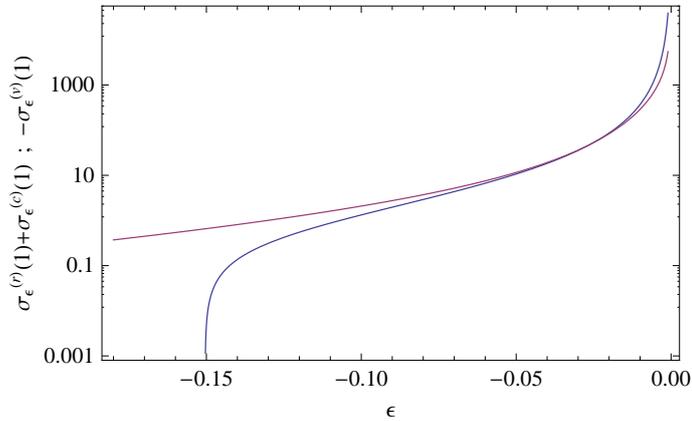


Figure 4.7 – Somme des transformées de Mellin de la contribution réelle et du contre-terme collinéaire (bleu) et transformée de l'opposé de la contribution virtuelle (rouge) du processus $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ dans le voisinage de $\epsilon = 0$ à $n = 1$ fixé. Les quantités tracées sont normalisées par la section efficace intégrée σ . Les transformées de Mellin de la contribution réelle et du contre-terme collinéaire sont effectuées numériquement.

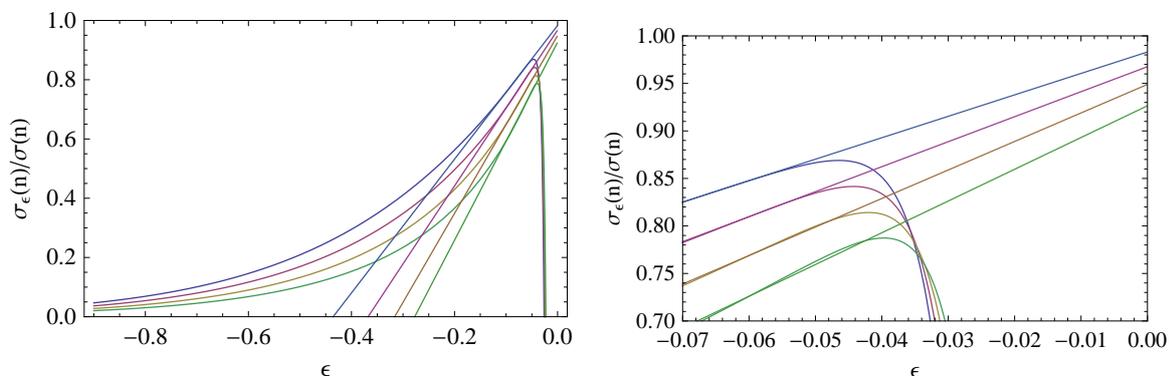


Figure 4.8 – Évaluation numérique du rapport des valeurs des transformées de Mellin de la section efficace différentielle du processus partonique $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ en la variable de Björken x en ϵ non nul et de sa valeur exacte (*i.e.* en $\epsilon = 0$) au voisinage de $\epsilon = 0$. Les transformées de Mellin sont évaluées en $n = 1, 10, 100, 1000$ de haut en bas. Les droites représentent les extrapolations d'ordre un de ces rapports.

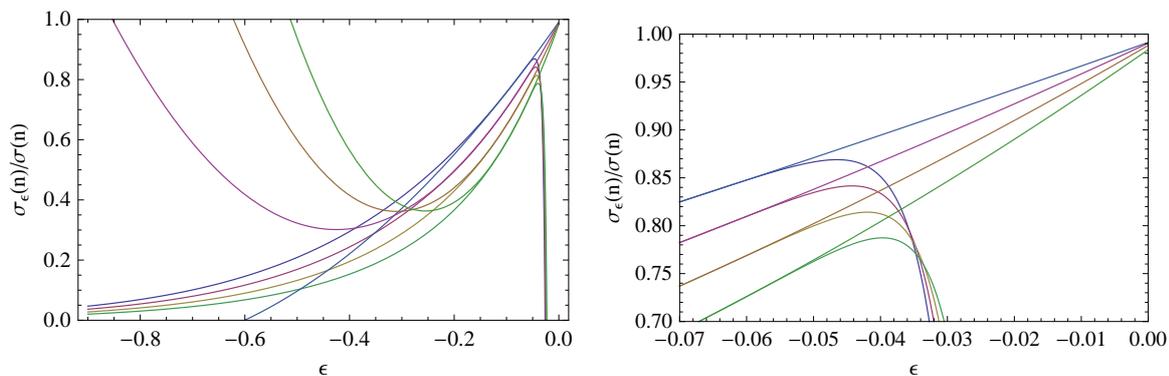


Figure 4.9 – Évaluation numérique du rapport des valeurs des transformées de Mellin de la section efficace différentielle du processus partonique $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ en la variable de Björken x en ϵ non nul et de sa valeur exacte (*i.e.* en $\epsilon = 0$) au voisinage de $\epsilon = 0$. Les transformées de Mellin sont évaluées en $n = 1, 10, 100, 1000$ de haut en bas. Les paraboles représentent les extrapolations d'ordre deux de ces rapports.

4.2 Annulation numérique des divergences dans l'espace de Mellin

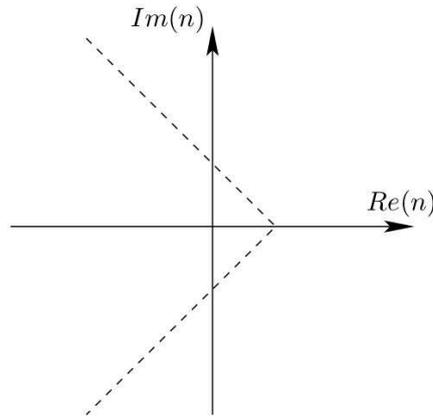


Figure 4.10 – Contour d'intégration schématisé utilisé pour l'évaluation numérique de la transformée de Mellin inverse de l'expression de la section efficace du processus $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ dans l'espace conjugué.

4.2.2.3 Transformée de Mellin inverse

La transformée de Mellin inverse (4.12) est effectuée numériquement. La bande de définition de la transformée de Mellin de la section totale est le demi-plan complexe de partie réelle positive. Le contour d'intégration est une droite parallèle à l'axe imaginaire de ce demi-plan, d'équation $n = c + ix, x \in \mathbb{R}, c \in \mathbb{R}^{+*}$.

L'intégrale complexe ainsi définie converge très lentement. Comme les pôles de la transformée de Mellin $\sigma(n)$ sont situés sur l'axe réel négatif, le contour d'intégration peut être déformé suivant la figure 4.10 vers le demi-plan complexe de partie réelle négative du moment qu'il coupe l'axe réel sur son demi-axe positif. L'intégrand x^{-n} accélère alors la convergence.

La figure 4.11 présente les valeurs de la section efficace différentielle de son support non singulier obtenues lorsque la transformée de Mellin des différentes contributions est effectuée analytiquement, l'annulation des divergences est effectuée dans l'espace conjugué (pour $N = 1$ en $\epsilon = 10^{-4}$ à gauche et pour $N = 3$ en $\epsilon = \{-0, 6.10^{-1}; -0, 7.10^{-1}; -10^{-1}\}$ à gauche) et la transformée de Mellin inverse est effectuée numériquement. Ces résultats sont normalisés par les valeurs exactes fournies par l'équation (4.8). L'erreur accumulée est de moins de 2% dans le cas d'une extrapolation quadratique à partir d'estimations en des valeurs de ϵ de l'ordre de -10^{-1} . De plus, le résultat obtenu dans l'espace de Mellin contient maintenant les contributions singulières à la section efficace différentielle. Les figures 4.6, 4.8 et 4.9 montrent que la section intégrée ($n = 1$) est bien reproduite : la contribution virtuelle, proportionnelle à une distribution de Dirac, est correctement prise en compte.

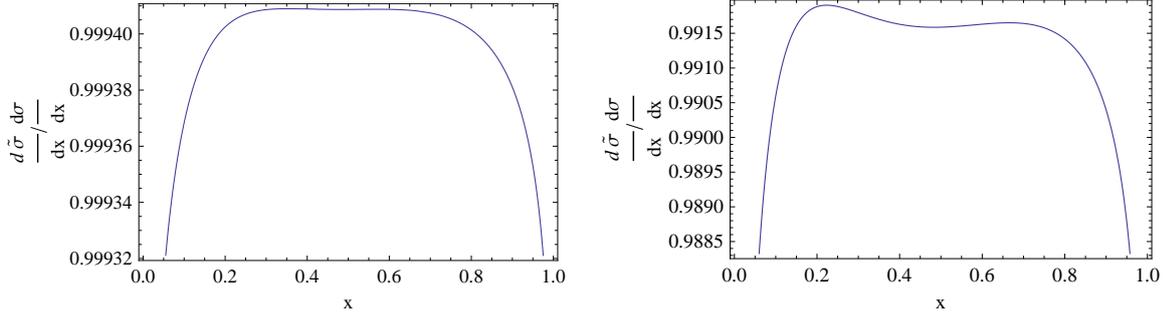


Figure 4.11 – Résultat de l'évaluation de la section efficace différentielle du processus $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$ après annulation numérique des divergences dans l'espace de Mellin. À gauche, l'annulation numérique des divergence est faite pour $N = 1$ en $\epsilon = 10^{-4}$. À droite, elle est effectuée pour $N = 3$ en $\epsilon = \{-0, 6.10^{-1}; -0, 7.10^{-1}; -10^{-1}\}$.

4.3 Limitations et utilité de l'approche

4.3.1 Dépendance explicite en ϵ

L'annulation numérique proposée des divergences collinéaires et infra-rouges utilise explicitement la dépendance des différentes contributions de la section efficace étudiée en fonction du paramètre de régularisation ϵ . La contribution réelle, par exemple, provient d'une intégration sur l'espace des phases. Son expression en $\epsilon = \frac{D-4}{2}$ non nul correspond au prolongement du résultat de cette intégration effectuée en une dimension D entière quelconque.

Les intégrations sur l'espace des phases ne peuvent être effectuées numériquement qu'en certaines valeurs entières de la dimension D . Le prolongement de ces résultats ne peut pas être calculé numériquement. Pour obtenir la dépendance en ϵ nécessaire à l'application de la méthode présentée, les intégrations sur l'espace des phases doivent être effectuées, au moins en partie, analytiquement.

4.3.2 Prolongement analytique dans l'espace conjugué

Le calcul de la transformée de Mellin inverse (4.12) s'effectue par intégration sur un contour de la bande de définition de la transformée de la section efficace dans l'espace conjugué. Sur ce contour, le calcul des transformées des différentes contributions de la section peut être effectué numériquement. En effet, par définition de la bande de définition, les intégrales (4.17), (4.18) et (4.20) sont convergentes.

En pratique, on s'attend à ce que la convergence de l'intégrale (4.12) de la transformée inverse soit très lente. De la même manière que dans le cas du processus $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$, le contour doit être déformé vers le demi-plan complexe de partie réelle négative. Le calcul numérique de la transformée inverse est alors plus stable. En contrepartie, une majorité du contour d'intégration se situe en dehors de la bande de définition de la transformée de Mellin de la section efficace. Au moins une des intégrales définissant les transformées des différentes contributions n'y est pas convergente. Sur ces parties du contour, le prolongement analytique de l'expression de la transformée de Mellin de la contribution en cause est indispensable. Ainsi, s'il s'avère que la bande de définition de la transformée de

4.3 Limitations et utilité de l'approche

la section totale n'est pas le plan complexe en entier, au moins une des transformées de Mellin de ses contributions doit être calculée analytiquement et son expression doit être étendue à tout le plan complexe.

4.3.3 Transformée de Mellin en k dimensions et position des pôles

Le recours au calcul numérique n'est pas nécessaire dans des cas aussi simple que le calcul de la section efficace différentielle première $e^+e^- \rightarrow \gamma^* \rightarrow q\bar{q}(g)$. Il pourrait par contre être utile pour l'étude de sections efficaces différentielles en k variables. L'annulation numérique des divergences collinéaires et infra-rouges s'effectuerait de la même manière que dans le cas d'une section différentielle première. Une transformée de Mellin serait effectuée pour chaque variable de chaque contribution. Les expressions totalement numériques obtenues pourraient être sommées dans l'espace conjugué \mathbb{C}^k en ϵ non nul.

Le retour à l'espace direct s'effectuerait par une transformée de Mellin inverse pour chaque paramètre complexe n_1, \dots, n_k . Cela correspondrait à effectuer une intégrale sur une surface de dimension réelle k dans \mathbb{C}^k . La convergence rapide de cette intégrale nécessiterait la déformation de la surface d'intégration vers la partie de l'espace complexe \mathbb{C}^k où les différents paramètres n_i sont de partie réelle négative. Cette déformation ne devrait traverser aucun pôle. Quand bien même l'expression analytique de la section dans l'espace de Mellin et la position de ses pôles serait connue, les questions de l'existence d'une telle surface déformée et de son identification deviendrait rapidement inextricable à haute dimension k .

4.3.4 Domaine d'application et utilité de la méthode

Au final, d'après les remarques de la section 4.3.1, une dépendance des différentes contributions à la section efficace étudiée dans le paramètre de régularisation ϵ est nécessaire pour appliquer la méthode proposée d'annulation des divergences collinéaires et infra-rouges. Les intégrales sur l'espace des phases ne peuvent pas être calculées numériquement en une dimension non entière. Les expressions des contributions doivent pour cette raison être calculées, au moins en partie, analytiquement. En fait, suivant le raisonnement de la section 4.3.2, les expressions doivent même être connues entièrement analytiquement. Il doit être possible de décider si leurs transformées de Mellin peuvent être calculées numériquement et, dans le cas négatif, d'obtenir leur forme analytique. Enfin, et comme discuté en section 4.3.3, la généralisation de la méthode d'annulation à des sections différentielles en plusieurs variables semble difficile à réaliser et encore plus à automatiser.

La méthode proposée ne peut finalement être utilisée de manière générale que sur une section efficace différentielle de premier ordre dont les expressions analytiques des différentes contributions, après intégration sur l'espace des phases, sont connues. L'annulation restante des divergences collinéaires et infra-rouges par développement en série de Laurent des contributions est simple dans ces cas. En effet, les fonctions présentes dans leurs expressions sont connues et leur développements en séries tabulés. L'annulation numérique des divergences ne pourrait être utile que si de nouvelles expressions apparaissaient. Mais la complexité et la lourdeur des transformations de Mellin directes et réciproques, devant être répétées à chaque fois que le résultat doit être calculé, semblent cher payé lorsqu'il suffirait de calculer les développements en séries des nouvelles fonctions apparues une fois

Annulation numérique des divergences infrarouges et collinéaires

pour toute. D'autant plus que la connaissance de ces développements pourrait bénéficier à de nombreux domaines.

Bibliographie

- [1] S. Weinzierl, PoS A **CAT** (2007) 005 [arXiv :0707.3342 [hep-ph]].
- [2] G. Passarino and M. J. G. Veltman, Nucl. Phys. B **160**, 151 (1979).
- [3] Z. Bern, L. J. Dixon, D. C. Dunbar and D. A. Kosower, Nucl. Phys. B **425**, 217 (1994) [arXiv :hep-ph/9403226].
- [4] Z. Bern, L. J. Dixon, D. C. Dunbar and D. A. Kosower, Nucl. Phys. B **435**, 59 (1995) [arXiv :hep-ph/9409265].
- [5] K. G. Chetyrkin and F. V. Tkachov, Nucl. Phys. B **192**, 159 (1981).
- [6] V. A. Smirnov and O. L. Veretin, Nucl. Phys. B **566**, 469 (2000) [arXiv :hep-ph/9907385].
- [7] W. T. Giele and E. W. N. Glover, Phys. Rev. D **46**, 1980 (1992).
- [8] W. T. Giele, E. W. N. Glover and D. A. Kosower, Nucl. Phys. B **403**, 633 (1993) [arXiv :hep-ph/9302225].
- [9] S. Keller and E. Laenen, Phys. Rev. D **59**, 114004 (1999) [arXiv :hep-ph/9812415].
- [10] S. Frixione, Z. Kunszt and A. Signer, Nucl. Phys. B **467**, 399 (1996) [arXiv :hep-ph/9512328].
- [11] S. Catani and M. H. Seymour, Nucl. Phys. B **485**, 291 (1997) [Erratum-ibid. B **510**, 503 (1998)] [arXiv :hep-ph/9605323].
- [12] S. Dittmaier, Nucl. Phys. B **565**, 69 (2000) [arXiv :hep-ph/9904440].
- [13] L. Phaf and S. Weinzierl, JHEP **0104**, 006 (2001) [arXiv :hep-ph/0102207].
- [14] S. Catani, S. Dittmaier, M. H. Seymour and Z. Trocsanyi, Nucl. Phys. B **627**, 189 (2002) [arXiv :hep-ph/0201036].
- [15] S. Weinzierl, Eur. Phys. J. C **45**, 745 (2006) [arXiv :hep-ph/0510157].
- [16] B. Mele, P. Nason, Nucl. Phys. **B361** (1991) 626-644.