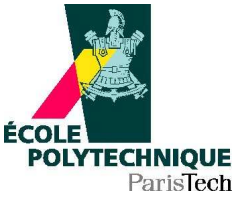
Mathematical analysis of light propagation in optical fibers with randomly varying birefringence

Maxime Gazeau



Thèse pour l'obtention du titre de docteur de l'École Polytechnique
spécialité : **Mathématiques Appliquées**
présentée par

Maxime Gazeau

Analyse de modèles mathématiques pour la propagation de la lumière dans les fibres optiques en présence de biréfringence aléatoire

Composition du jury :

Mme	Anne de Bouard	Directrice
M.	Josselin Garnier	Rapporteur
Mme	Erika Hausenblas	Rapporteur
M.	Arnaud Debussche	Examinateur
Mme	Annie Millet	Examinatrice
M.	François Alouges	Examinateur

Soutenue le 19 Octobre 2012 à l'École Polytechnique.

Centre de Mathématiques Appliquées de l'École Polytechnique – UMR 7641

Remerciements

Débuter une thèse c'est un peu, enfin je l'imagine, comme diriger une expédition. Au début, il faut se mettre en quête de financements, car partir dans une expédition au long cours coûte cher. Je remercie ici le conseil régional d'Ile-de-France qui accepta ce projet et le finança pendant trois ans. Partir seul dans une aventure d'une telle ampleur est voué à l'échec. Il faut de l'aide, du soutien et du réconfort que l'on trouve auprès d'un général d'expérience, compétent et charismatique mais également auprès d'autres corps expéditionnaires, partis dans des galères similaires.

Souvent, on me demande ce qui, précisemment, m'a conduit à débuter cette aventure. Le désir d'apprendre, l'inconnu, la curiosité ! Probablement un peu de tout cela mais encore plus sur un coup de folie car "ce [les coups de folie] sont les seules choses que l'on ne regrette pas[(1)]". Si je ne regrette pas ce choix, je le dois avant tout à ma directrice de thèse Anne de Bouard qui a été un général formidable. Merci d'avoir accepté de me superviser et de me faire confiance il y a trois ans déjà. Merci de m'avoir initié à la recherche en me guidant patiemment. Enfin, merci pour ce sujet passionnant et mathématiquement très riche.

Je souhaite également exprimer ma profonde gratitude à tous ceux qui m'ont permis de mener à bien ce travail de thèse. Je remercie sincèrement les professeurs Josselin Garnier et Erika Hausenblas pour avoir accepté de rapporter cette thèse ainsi que pour leurs lectures attentives du manuscrit. C'est un grand honneur pour moi que d'avoir Josselin dans mon jury car nombreux de ses articles ont été sur mon bureau pendant ces dernières années. J'espère que nos chemins se recroiseront dans le futur. Je tiens également à exprimer toute ma gratitude à Annie Millet, qui m'a fait l'honneur de présider mon jury de thèse, ainsi qu'à Arnaud Debussche et François Alouges pour avoir accepté d'être examinateurs dans ce jury.

Je salue tous les membres du département de Mathématiques appliquées de l'Ecole Polytechnique au sein duquel il y règne une ambiance chaleureuse et conviviale. Je souhaite en particulier remercier Grégoire Allaire et Frédéric Coquel pour leur formidable cours de M2 mais aussi, tout simplement, pour leur grande gentillesse. Enfin, un grand merci à Christophe Besse et Stephan de Bièvre pour avoir accepté de m'encadrer pour un post-doctorat au sein du Laboratoire Paul Painlevé à Lille.

Je tiens également à remercier les nombreux doctorants du CMAP et en particulier ceux du bureau 2011. Je sais que parmi eux certains attachent de l'importance à voir leurs noms apparaîtrent dans la longue liste des remerciements. C'est avec un grand plaisir que je remercie ceux qui m'ont accompagné tout au long de ma thèse : Camille, Laurent, Xavier, Xiaolu, Michael, Chao, Hao, Khalid, Sylvie, Emilie, Florent, Soledad, Nicolas, Franscico, Abdul ainsi que la horde de nouveaux doctorants contribuant au renouvellement des générations. Merci à tous mes amis, matheux et non matheux, que j'ai toujours grand plaisir à voir : Julien, Guillaume,

(1) Oscar Wilde, *Le portrait de Dorian Gray*, 1891

Antonin, Marco, Khaled, Nafi, Emmanuel, Cécile, Julie, Sébastien, Kenza, Laure, l'autre Laure, Cédric, Marion, Eric, ... sans oublier mes "nouveaux" amis architectes qui se reconnaitront.

Je remercie mes amis Versaillais triathlètes (et non triathlètes) pour les longues sorties en vélo du dimanche matin dans la vallée de chevreuse. J'espère que notre déménagement ne changera en rien ces bonnes habitudes. Je remercie également les adjudants de la section Judo de l'X grâce à qui j'ai pu continuer à pratiquer ce beau sport.

Il serait incorrect et malpoli de ma part de ne pas remercier les secrétaires du CMAP pour leur incroyable efficacité, leur gentillesse et leur sourire en toutes circonstances. Si il règne une ambiance agréable au laboratoire, elles n'y sont pas étrangères.

Je repense également à mes parents et à leurs regards ébahis lorsque je leur ai annoncé que je débutais une thèse de mathématiques. Merci de m'avoir toujours soutenu dans mes choix en m'accordant votre confiance. J'adresse également un vif remerciement à mes trois frangins. Ce n'est pas souvent que l'on peut exprimer toute l'affection que l'on porte à ses petits frères alors je saisis l'occasion qui m'est donnée ici pour le leur dire : Benoit, Adrien, Thomas, je suis content de vous avoir. Enfin, je souhaiterais dédier cette thèse à ma femme Kristel. Merci d'avoir été présente à mes côtés et de te refuser à comprendre les mathématiques. En d'autres temps cela aurait agacé Henri Poincaré, mais je suis sûr qu'il aurait changé d'avis en te rencontrant.

Table des matières

Chapitre 1

Introduction générale

La “communication”, au sens premier du terme, est un échange entre un émetteur et un récepteur. A l'époque de l'empire romain, les voies romaines servaient de “canaux” de communications pour le transport des personnes et des marchandises. Aujourd'hui, les moyens de communications ont pris une forme bien différente et on parle de “télécommunication[1]” lorsque l'information est transmise sur des distances importantes via un système électrique ou optique.

Maxwell démontra dans les années 1860, grâce à sa théorie de l'électromagnétisme, que la lumière était une onde électromagnétique et que par conséquent elle pouvait se propager le long d'un guide d'onde au même titre que d'autres ondes électromagnétiques. Avec les fibres optiques, l'information est transmise via la lumière et non plus sous forme électrique comme c'est le cas avec les câbles coaxiaux. Lorsque l'on parle de fibres optiques, on pense immédiatement au très haut débit pour les systèmes de télécommunications. Pourtant, leur première utilisation, dans les années 1950, fut en médecine avec l'invention du fibroscope. L'utilisation de la fibre optique dans les télécommunications ne fut pas immédiate et fut conditionnée par plusieurs avancées technologiques (laser, augmentation de la qualité des fibres $\cdots$) qui ont permis d'envisager le transport de l'information via ce canal de transmission.

Depuis, les transmissions optiques ont considérablement évolué notamment avec l'avènement d'Internet, dans la seconde moitié du 20^{e} siècle, et surtout du Web qui a permis l'ouverture du réseau internet au grand public. Ce mode de transmission constitue un réel enjeu économique et stratégique pour les acteurs industriels du secteur des télécommunications. La demande croissante de débits fait face à la capacité insuffisante des réseaux et il faut sans cesse trouver de nouvelles innovations permettant d'augmenter les débits. Une des dernières en date est le multiplexage en longueur d'onde (Wavelength Division Multiplexing en anglais) qui permet de faire passer plusieurs signaux de longueurs d'onde différentes dans une seule fibre optique. Récemment, la compagnie Japonaise NEC a réussi à établir une transmission optique autorisant pour chaque longueur d'onde un débit de 100 Gigabits par seconde pour un total de 88 longueurs d'onde en utilisant le réseau de fibre optique Japonais déjà existant[2] :

[1] Terme introduit par E. Estaunié, “Traité pratique de télécommunication électrique”, 1904.

[2] T. Kato and J. Jasper, “NEC Contributes to Japan's First 100Gbps-DWDM Fiber Cable

> "NEC Corporation announced today [June 21, 2012] its participation in Japan's first DWDM[(3)] transmission of 100Gbps per wavelength [...] connecting the 710 km between Tokyo, Nagoya and Osaka[...]. Tests were conducted using existing commercial cable [...]. These actual operating conditions present a variety of operational challenges, including [...] variability in optical fiber characteristics (Chromatic Dispersion, Polarization-Mode Dispersion,[(4)])[...]. In recent years, an urgent need to strengthen network infrastructure emerged as transmission traffic rapidly increased from the growing use of smartphones, video sites and SNS[...]. This [100Gbps-DWDM transmission system] enables telecommunications carriers to efficiently maintain network infrastructure that boasts high speed, high capacity and high reliability."

Des technologies toujours plus performantes ont ainsi permis à la compagnie NEC d'accroître les capacités de transmission par fibres optiques au Japon. Néanmoins, les ingénieurs sont toujours confrontés aux mêmes effets dispersifs (dispersion chromatique, effet Kerr optique, dispersion modale de polarisation (PMD)) qui dégradent les performances des télécommunications et dont l'effet cumulé reste assez mal compris. La dispersion modale de polarisation est un phénomène de dispersion qui résulte du changement de l'état de polarisation du champ électrique au cours de sa propagation dans la fibre. La PMD se manifeste par une différence de vitesse de groupe entre les deux composantes du champ électrique, ainsi que par un échange d'énergie entre les deux modes. Depuis une dizaine d'années, de nombreux physiciens ont entrepris l'étude de l'évolution de la lumière dans les fibres optiques en présence de la dispersion modale de polarisation. Ces études ont débouché sur la dérivation d'une équation permettant de modéliser l'évolution de l'enveloppe lentement variable du champ électrique [76, 77, 78, 110] et qui a donné suite à une série de travaux permettant d'étudier la dynamique de la PMD ainsi que l'impact des différents effets dispersifs lorsque ceux ci interagissent [39, 42, 43, 60, 72, 73, 75, 107, 108, 109].

Mon travail de thèse s'inscrit dans la continuité de ces recherches et l'objectif est d'étudier, d'un point de vue à la fois théorique et numérique, la propagation de la lumière dans les systèmes de transmission optique, en présence de dispersion modale de polarisation. Le deuxième chapitre est écrit en français, tandis que les trois autres le sont en anglais. On utilise les notations physiques pour les chapitres 2 et 5, c'est-à-dire que z correspond à la variable d'évolution et t au temps retardé. Dans les chapitres 3 et 4 on utilise les notations mathématiques : t désigne l'évolution le long de la fibre et x la variable d'espace. Les quatre chapitres sont divisés comme suit

Chapitre 2. Ce chapitre est introductif et ne comporte aucun résultat mathématique nouveau. Il présente les principes de la propagation optique et sa modélisation

Transmission Using Commercial Fiber Cable between Tokyo, Nagoya and Osaka (710km)", mis en ligne le 21 juin 2012. URL : http://www.nec.com/en/press/201206/global_20120621_01.html.

(3) Dense Wavelength Division Multiplexing (DWDM). DWDM is transmission technology that utilizes multiplexing optical fiber to simultaneously transmit numerous optical signals of different wavelengths. This is a high density version of Wavelength Division Multiplexing (WDM).

(4) Polarization-Mode Dispersion (PMD). PMD is a phenomenon where a propagation time difference arises among components of orthogonal polarization mode within optical fiber.

mathématique. L'objectif est d'abord d'introduire les notions physiques (biréfringence, dispersion modale de polarisation, différence de vitesse de groupe etc.) utilisées au cours de la thèse. On présente ensuite, à partir des équations de Maxwell, la dérivation d'un système couplé d'équations de Schrödinger non-linéaires permettant de modéliser la dispersion modale de polarisation. Dans le repère local des axes de biréfringence, l'équation s'écrit

$$i\frac{\partial \Phi}{\partial z} + \Sigma\left(\theta_z, z\right)\Phi + ib'\sigma_3\frac{\partial \Phi}{\partial t} + \frac{d_0}{2}\frac{\partial^2 \Phi}{\partial t^2} + \frac{5}{6}\left|\Phi\right|^2\Phi + \frac{1}{6}\left(\Phi^*\sigma_3\Phi\right)\sigma_3\Phi + \frac{1}{3}\begin{pmatrix}\overline{\Phi_1}\Phi_2^2 e^{-4ibz}\\ \\ \overline{\Phi_2}\Phi_1^2 e^{4ibz}\end{pmatrix} = 0, \tag{1.0.1}$$

et sera le point de départ de l'analyse mathématique du chapitre 3. Dans ce modèle, la biréfringence, à l'origine de la dispersion modale de polarisation, est paramétrée par les variables b et θ. Le paramètre d_0 correspond à la dispersion de vitesse de groupe et b' la dérivée de b par rapport à la fréquence angulaire. La description de l'évolution du champ électrique, soumis aux variations aléatoires de la biréfringence, est beaucoup plus compliquée que dans le cas usuel où l'équation de Schrödinger non linéaire s'applique, notamment à cause du couplage linéaire et non linéaire. Lorsque la biréfringence est nulle et que l'une des composantes est nulle, on retrouve l'équation de NLS standard.

Chapitre 3. La modélisation physique de ce problème fait intervenir plusieurs échelles caractéristiques. L'échelle la plus petite est la longueur de battement L_B correspondant à la distance nécessaire pour que le champ électrique retourne à son état de polarisation initial. La longueur caractéristique suivante est la longueur de corrélation de la PMD l_c. Enfin, les échelles suivantes sont relatives aux effets dispersifs (effet Kerr optique l_{nl} et dispersion chromatique l_d) et sont du même ordre de grandeur que la longueur de la fibre l.

L'objectif de ce chapitre est d'exhiber une dynamique asymptotique pour l'enveloppe lentement variable du champ électrique, décrite par (1.0.1), en utilisant des techniques de séparation d'échelles et la théorie de l'approximation-diffusion [9, 38, 40, 57, 70, 73, 87, 88]. Lorsque les différentes échelles physiques d'un problème peuvent être séparées, les méthodes d'approximation-diffusion permettent de prouver des théorèmes limites, au sens de la convergence faible de mesures de probabilités, pour des processus stochastiques, solution d'équations différentielles stochastiques et d'équations aux dérivées partielles stochastiques. Ces résultats constituent une généralisation du théorème suivant qui établit que le mouvement Brownien est une limite universelle [6].

Théorème 1.0.1 (Théorème central limite fonctionnel)**.** *On considère une suite* $\left(\xi_k\right)_{k\in\mathbb{Z}}$ *de variables aléatoires stationnaire, ergodique et de moments du second ordre finis tel que* $\sum_{n=1}^{+\infty}\mathbb{E}\left(\left(\mathbb{E}\left(\xi_n\middle|\mathcal{F}_0\right)\right)^2\right)^{1/2} < +\infty$*. Alors,* $\mathbb{E}\left(\xi_n\right) = 0$ *et*

$$\sigma^2 = \mathbb{E}\left(\xi_0^2\right) + 2\sum_{k=1}^{+\infty}\mathbb{E}\left(\xi_0\xi_k\right)$$

converge absolument. De plus, si $\sigma > 0$ *et* $S_n = \frac{1}{\sqrt{n}} \sum_{k=1}^{[nt]} \xi_k$*, alors* S_n *converge en loi vers* σW*, où* W *est un mouvement Brownien, dans les espaces de Skorohod.*

On démontre facilement à partir du résultat précédent une version continue de ce théorème. Soit $\epsilon > 0$ un petit paramètre adimensionné. Soit $(m(t))_{t\in\mathbb{R}_+}$ un processus de Markov, centré, homogène, stationnaire et ergodique tel que $0 < \sigma^2 = 2\int_0^{+\infty} \mathbb{E}\left(m(0)m(t)\right) dt < +\infty$. Alors,

$$\frac{1}{\epsilon}\int_0^t m\left(\frac{s}{\epsilon^2}\right) ds \underset{\epsilon\to 0}{\Longrightarrow} \sigma W \qquad \text{dans } C([0,T],\mathbb{R}),$$

où $\Longrightarrow$ indique la convergence en loi. L'exemple le plus classique est le processus de Langevin de paramètre λ dont les fonctions de transition sont des mesures gaussiennes de moyenne $m(0)\exp(-\lambda t)$ et de variance $\sigma(t) = (1-\exp(-2\lambda t))/(2\lambda)$. La mesure invariante est donnée par $\sqrt{\lambda/\pi}\exp(-\lambda x^2)$. Les théorèmes d'approximation-diffusion sont une généralisation de ce résultat pour des processus de Markov solution d'un système dynamique perturbé. En dimension finie, la théorie est assez bien développée et la dynamique asymptotique est donnée par une équation différentielle stochastique. En dimension infinie, on a l'intuition que ce type de résultat reste vrai pour des EDP perturbées aléatoirement. Considérons par exemple l'équation de transport stochastique suivante

$$\partial_t X_\epsilon = \frac{1}{\epsilon} m\left(\frac{t}{\epsilon^2}\right) \partial_x X_\epsilon, \qquad t \geqslant 0, \quad x \in \mathbb{R},$$

avec pour donnée initiale $X_\epsilon(0) = X(0) = X_0 \in L^2(\mathbb{R})$. Dans le domaine fréquentiel, la solution de cette équation est donnée par $\widehat{X}_\epsilon(t) = \exp\left(i\xi\int_0^t \frac{1}{\epsilon} m\left(\frac{u}{\epsilon^2}\right) du\right)\widehat{X}_0$. La version continue du théorème central limite fonctionnel permet de conclure que $\widehat{X}_\epsilon$ converge en loi, lorsque ϵ tend vers zéro, vers $\widehat{X}(t) = \exp\left(i\xi\sigma W(t)\right)\widehat{X}_0$ dans $C\left([0,T],L^2\right)$. L'équation limite obtenue est

$$dX = \sigma\partial_x X \circ dW(t),$$

où $\circ$ désigne l'intégral stochastique au sens de Stratonovich. Cet exemple est trivial car la solution de l'EDP perturbée, ainsi que la solution de l'équation limite, est une fonction continue du bruit, dont l'expression est explicite. Récemment, des résultats d'approximation-diffusion ont été obtenus pour certaines EDP (à coefficients aléatoires) dispersives linéaires [43] et non linéaires [21, 28]. Le résultat obtenu par Garnier et Marty [43] est relatif à l'équation de Manakov-PMD linéaire. Pour démontrer ce résultat ils utilisent également la transformée de Fourier, de fortes hypothèses de régularité sur la donnée initiale ainsi que la méthode de la fonction test perturbée. L'équation limite obtenue dans ce cas est une EDP linéaire stochastique dirigée par un mouvement brownien 3d. Pour les EDP dispersives non linéaires [21, 28], la preuve utilise la continuité de la solution par rapport au processus directeur et la version continue du théorème central limite fonctionnel. Ces méthodes ne sont pas applicables dans notre situation parce que les dynamiques asymptotiques ne sont pas des applications continues du bruit et que l'utilisation de la transformée de Fourier n'est pas possible à cause de la non-linéarité.

L'obtention de la dynamique asymptotique pour l'équation (1.0.1) s'effectue en deux étapes. Dans un premier temps, on s'intéresse à l'évolution du champ sur l'échelle correspondante à l_c dont la dynamique adimensionnée est donnée par

$$i\frac{\partial \Phi_\varepsilon}{\partial t} + \frac{1}{\varepsilon}\Sigma\left(\alpha_\varepsilon(t), \frac{t}{\varepsilon^2}\right)\Phi_\varepsilon + ib'\sigma_3\frac{\partial \Phi_\varepsilon}{\partial x} + \frac{d_0}{2}\frac{\partial^2 \Phi_\varepsilon}{\partial x^2} + \frac{5}{6}\left|\Phi_\varepsilon\right|^2\Phi_\varepsilon + \frac{1}{6}\left(\Phi_\varepsilon^*\sigma_3\Phi_\varepsilon\right)\sigma_3\Phi_\varepsilon + \frac{1}{3}\begin{pmatrix}\overline{\Phi_{1,\varepsilon}}\Phi_{2,\varepsilon}^2 e^{-4ibt/\varepsilon^2} \\ \overline{\Phi_{2,\varepsilon}}\Phi_{1,\varepsilon}^2 e^{4ibt/\varepsilon^2}\end{pmatrix} = 0, \tag{1.0.2}$$

avec $\varepsilon = \sqrt{L_B/l_c}$ et $\alpha_\varepsilon(t) = \alpha(t/\varepsilon^2)$. Soit $W = (W_1, W_2)$ un mouvement Brownien de dimension 2. On introduit l'équation aux dérivées partielles stochastique

$$id\Psi(t) + \left\{ib'\sigma_3\frac{\partial \Psi}{\partial x} + \frac{d_0}{2}\frac{\partial^2 \Psi}{\partial x^2} + \frac{5}{6}\left|\Psi\right|^2\Psi + \frac{1}{6}\left(\Psi^*\sigma_3\Psi\right)\sigma_3\Psi\right\}dt + \gamma_s\sigma_3\Psi dt + i\gamma_c\Psi dt - \sqrt{\gamma_c}\left(\sigma_1\Psi dW_1(t) + \sigma_2\Psi dW_2(t)\right) = 0. \tag{1.0.3}$$

Le résultat suivant établit qu'à l'échelle de la longueur de corrélation, la dynamique asymptotique est donnée par l'Equation (1.0.3).

Théorème 1.0.2. *Soit α un processus de Feller homogène, centré, stationnaire et ergodique et dont le générateur infinitésimal $\mathscr{L}_\alpha$ satisfait l'alternative de Fredholm. Soit $T > 0$ et $\Phi_\varepsilon(0) = X_0 = v \in \mathbb{H}^3(\mathbb{R})$, alors l'unique solution globale Φ_ε de (1.0.2) converge en loi vers l'unique solution globale Ψ de (1.0.3) dans l'espace des fonctions continues sur $[0,T]$ à valeurs dans $\mathbb{H}^1$.*

Dans la seconde étape, on s'intéresse aux variations du champ électrique sur les échelles plus longues c'est-à-dire sur les échelles de la dispersion chromatique et de l'effet Kerr. On introduit donc un petit paramètre $\epsilon = \sqrt{l_c/l_d}$ et les notations $\Psi_\epsilon(t) = \frac{1}{\epsilon}\Psi(\frac{t}{\epsilon^2}, \frac{x}{\epsilon})$ et $X_\epsilon(t) = \frac{1}{\epsilon}X(\frac{t}{\epsilon^2}, \frac{x}{\epsilon})$. Un changement de variable $\Psi_\epsilon(t) = Z_\epsilon(t)X_\epsilon(t)$ permet de décrire l'évolution du champ électrique dans le référentiel où, pour un champ linéaire, le vecteur de Stokes est constant le long de la fibre à la fréquence nulle. Dans ce référentiel, l'évolution de X_ϵ est donnée par une équation que nous appellerons équation de Manakov PMD

$$i\frac{\partial X_\epsilon(t)}{\partial t} + \frac{ib'}{\epsilon}\boldsymbol{\sigma}\left(\nu_\epsilon(t)\right)\frac{\partial X_\epsilon(t)}{\partial x} + \frac{d_0}{2}\frac{\partial^2 X_\epsilon(t)}{\partial x^2} + F_{\nu_\epsilon(t)}(X_\epsilon(t)) = 0, \tag{1.0.4}$$

où ν est un processus de diffusion évoluant sur la sphère $\mathbb{S}^3$ de $\mathbb{R}^4$. De plus, ce processus admet une unique mesure invariante sous laquelle il est ergodique. Soit $W = (W_1, W_2, W_3)$ un mouvement Brownien de dimension 3. L'équation limite, lorsque ϵ tend vers zéro, est donnée par

$$idX(t) + \left(\frac{d_0}{2}\frac{\partial^2 X(t)}{\partial x^2} + F(X(t))\right)dt + i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k\frac{\partial X(t)}{\partial x}\circ dW_k(t) = 0, \tag{1.0.5}$$

où les matrices σ_k, $k = 1, 2, 3$ sont les matrices de Pauli. Le résultat de convergence est décrit dans le théorème suivant

Théorème 1.0.3. *Soit $\epsilon > 0$ et $X_\epsilon(0) = X(0) = v \in \mathbb{H}^3$. Soit $X_\epsilon \in C(\mathbb{R}_+, \mathbb{H}^3(\mathbb{R}))$ l'unique solution de (1.0.4) et $X \in C([0, \tau^*), \mathbb{H}^3(\mathbb{R}))$ l'unique solution locale de (1.0.5), τ^* étant le temps d'arrêt maximal d'existence. Alors X_ϵ converge en loi vers X dans l'espace des fonctions continues, pouvant exploser en temps fini, à valeurs dans $\mathbb{H}^1$.*

Dans ce chapitre, on démontre d'abord le Théorème 1.0.3 puis le Théorème 1.0.2, qui se démontre avec les mêmes arguments. La démonstration du Théorème 1.0.3 se fait en trois étapes : dans un premier temps, on montre le résultat d'existence suivant.

Théorème 1.0.4. *Soit $\epsilon > 0$ et $X_\epsilon(0) = v \in \mathbb{L}^2(\mathbb{R})$, alors l'équation (1.0.4) admet une unique solution globale X_ϵ dans $C(\mathbb{R}_+, \mathbb{L}^2) \cap \mathbb{L}^8_{loc}(\mathbb{R}_+, \mathbb{L}^4)$. De plus, la norme $\mathbb{L}^2$ de la solution X_ϵ est constante au cours du temps, i.e. pour tout t appartenant à $\mathbb{R}_+$, $\|X_\epsilon(t)\|_{\mathbb{L}^2} = \|v\|_{\mathbb{L}^2}$. Si on suppose plus de régularité pour la donnée initiale, i.e. $X_\epsilon(0) = v \in \mathbb{H}^1$ (resp. $\mathbb{H}^2$, resp. $\mathbb{H}^3$), alors l'unique solution correspondante appartient à $C(\mathbb{R}_+, \mathbb{H}^1)$ (resp. $C(\mathbb{R}_+, \mathbb{H}^2)$, resp. $C(\mathbb{R}_+, \mathbb{H}^3)$).*

Il n'est pas possible de prouver des estimations de Strichartz ([13, 101]) directement sur le système (1.0.4) à cause de la dépendance temporelle de la matrice $\boldsymbol{\sigma}$. Ainsi, on démontre d'abord l'existence d'une unique solution globale pour l'équation (1.0.3) et on en déduit un résultat d'existence pour l'équation (1.0.4).

La seconde étape consiste à étudier le caractère bien posé de l'équation limite (1.0.5) et on prouve le Théorème suivant.

Théorème 1.0.5. *Soit $X_0 = v \in \mathbb{H}^1(\mathbb{R})$, alors il existe un temps d'arrêt maximal $\tau^*(v, \omega)$ et une unique solution forte X (au sens probabiliste) à l'équation (1.0.5), tel que $X \in C([0, \tau^*), \mathbb{H}^1(\mathbb{R}))$ presque sûrement. De plus, la norme $\mathbb{L}^2$ est presque sûrement constante, i.e. $\forall t \in [0, \tau^*), \|X(t)\|_{\mathbb{L}^2} = \|v\|_{\mathbb{L}^2}$ et le temps d'arrêt maximal d'existence satisfait l'alternative suivante*

$$\tau^*(v, \omega) = +\infty \quad ou \quad \limsup_{t \nearrow \tau^*(v,\omega)} \|X(t)\|_{\mathbb{H}^1} = +\infty.$$

De plus si $v \in \mathbb{H}^2$, alors $X \in C([0, \tau^*), \mathbb{H}^2(\mathbb{R}))$ et τ^* satisfait

$$\tau^*(v, \omega) = +\infty \quad ou \quad \lim_{t \nearrow \tau^*(v,\omega)} \|X(t)\|_{\mathbb{H}^1} = +\infty. \tag{1.0.6}$$

La difficulté qui se présente ici est la même que précédemment, à la différence que cette fois-ci, il n'existe pas de transformations permettant d'obtenir des estimations de Strichartz. La raison vient du fait que l'équation est dirigée par un mouvement brownien multidimensionnel ([32, 102, 112]). On obtient l'existence d'une unique solution locale en utilisant cette fois une méthode de compacité et le problème martingale.

On montre enfin le théorème limite 1.0.3 (au sens de la convergence de mesure de probabilité [6, 34, 88]). Le schéma de la preuve est classique : on montre d'abord la tension de la suite de mesure de probabilité $\mathcal{L}_\epsilon(X_\epsilon)$ dans un espace métrisable

séparable et complet, dans lequel la tension est équivalente à la relative compacité séquentielle, et dont les ensembles compacts pourront être caractérisés par le théorème d'Ascoli et de Banach-Alaoglu. L'unicité de la limite est obtenue grâce à la convergence des problèmes martingales. Cette étape permet en fait de caractériser de manière unique les distributions fini-dimensionnelles. On a utilisé la méthode de la fonction test perturbée introduite par Papanicolaou-Stroock-Varadhan [88] et Blankenship-Papanicolaou [9], qui permet d'identifier la limite éventuelle, et surtout d'obtenir des bornes uniformes en ϵ qui permettront de démontrer la tension et la convergence des problèmes martingales. En effet, des estimations a priori sur l'équation (1.0.4) ne permettent pas d'obtenir les bornes uniformes nécessaires sur les modules de continuité, à cause du terme en $1/\epsilon$. Jusqu'à présent, cette méthode n'avait été utilisée que dans le cadre des EDS et de la dimension finie. En dimension infinie, la situation est beaucoup plus technique car il faut travailler dans les bons espaces pour obtenir la tension.

Chapitre 4. Dans le chapitre précédent, on a montré que la dynamique asymptotique de la solution de l'équation de Manakov PMD (1.0.4) est donnée par l'équation de Manakov stochastique (1.0.5). Dans ce chapitre, l'objectif est de construire un schéma pour l'équation (1.0.5) qui préserve les invariants de cette équation c'est-à-dire la norme $\mathbb{L}^2$. Il existe deux approches différentes basées sur le fait que l'intégrale de Stratonovich est égale à l'intégrale d'Itô plus une correction. Dans de nombreuses situations, l'approche "Itô" est avantageuse car la discrétisation explicite du bruit réduit le temps de calcul lors des simulations numériques et simplifie les calculs pour l'obtention d'un ordre de convergence. Néanmoins dans notre situation, cette approche se révèle désastreuse. Non seulement, le schéma ainsi construit ne préserve pas la norme $\mathbb{L}^2$ discrete, mais en plus, il est presque sûrement instable. En effet, comme dans le cas déterministe, un schéma contenant une discrétisation explicite du gradient n'est stable en norme $\mathbb{L}^2$ que sous une condition CFL (Courant-Friedrichs-Levy) dépendant de la vitesse. Ici, une condition CFL déterministe n'est presque sûrement jamais satisfaite car les vitesses sont données par des variables Gaussiennes et imposer une condition CFL aléatoire, qui dépendrait de toute la trajectoire brownienne, serait beaucoup trop restrictif. On propose donc une approche alternative basée sur une discrétisation semi-implicite de l'intégrale de Stratonovich qui permet au schéma de conserver la norme $\mathbb{L}^2$ discrète, ce qui par ailleurs assure la stabilité inconditionnelle du schéma dans la même norme. On considère un schéma semi-discret de type Crank Nicolson, basé sur celui existant pour l'équation déterministe, qui s'écrit

$$\begin{cases} X_N^{n+1} - X_N^n + H_{\Delta t,n} X_N^{n+1/2} - iF\left(X_N^n, X_N^{n+1}\right)\Delta t = 0 \\ F\left(X_N^n, X_N^{n+1}\right) = \frac{4}{9}\left(\left|X_N^n\right|^2 + \left|X_N^{n+1}\right|^2\right) X_N^{n+1/2}, \end{cases} \tag{1.0.7}$$

où $\Delta t > 0$ est le pas de temps et $H_{\Delta t,n}$, un opérateur matriciel, donné par

$$H_{\Delta t,n} = -i\Delta t I_2 \partial_x^2 + \sqrt{\gamma \Delta t} \sum_{k=1}^{3} \sigma_k \chi_k^n \partial_x. \tag{1.0.8}$$

Le schéma est stable et on montre qu'il est précis à l'ordre 1/2 en probabilité. On note $\widetilde{X}^n$ la solution exacte de l'équation (1.0.5) évaluée au temps t_n. Le résultat principal de ce chapitre est le suivant.

Théorème 1.0.6. *Soit $X_0 \in \mathbb{H}^6$, alors pour tout temps d'arrêt τ tel que presque sûrement $\tau < \tau^*$*

$$\lim_{C\to+\infty} \mathbb{P}\left(\max_{n=0,\dots,N_\tau} \left\|X^n - \widetilde{X}^n\right\|_{\mathbb{H}^1} \geqslant C\Delta t^{1/2}\right) = 0,$$

uniformément en Δt. On dit alors que le schéma est d'ordre $1/2$ en probabilité. Si de plus $\delta < \frac{1}{2}$, alors il existe une variable aléatoire K_δ tel que

$$\max_{n=0,\dots,N_\tau} \left\|X^n - \widetilde{X}^n\right\|_{\mathbb{H}^1} \leqslant K_\delta(T,R,\omega)\Delta t^\delta.$$

Pour prouver ce résultat, on est confronté à plusieurs difficultés. La première est due à la discrétisation semi-implicite de l'intégrale de Stratonovich, qui impose de travailler avec des propagateurs aléatoires, dépendants de n, et des processus non-adaptés. La seconde difficulté, déjà rencontrée au chapitre 3, est due au fait que le propagateur aléatoire, solution de l'équation linéaire associée à (1.0.5), n'a pas de formulation explicite. En effet, l'équation est dirigée par un mouvement brownien multidimensionnel ne satisfaisant pas l'hypothèse de commutativité. Enfin la partie dérive est fortement non linéaire et on obtient un ordre fort en probabilité de 1/2. Ce type de résultat est similaire à ceux obtenus pour une équation de Schrödinger non linéaire par A. de Bouard et A. Debussche dans [20] et pour des équations paraboliques par J. Printems dans [95].

Chapitre 5. Ce dernier chapitre est consacré à la simulation numérique de l'équation (1.0.5). L'objectif est d'étudier numériquement l'évolution de la dispersion modale de polarisation et la stabilité des solitons de Manakov soumis à la PMD. Dans un premier temps, on introduit les schémas numériques qui seront implémentés dans ce chapitre c'est-à-dire le schéma de Crank-Nicolson (1.0.7), un schéma de Relaxation ainsi qu'un schéma de splitting. La partie suivante constitue une validation de l'implémentation numérique, d'abord dans un cadre où la biréfringence est homogène, puis dans le cadre d'une fibre optique aléatoirement biréfringente en l'absence de dispersion chromatique et d'effet Kerr. Dans ces deux situations, il existe des résultats théoriques permettant de valider la pertinence des simulations numériques. Enfin, dans la dernière partie, on étudie numériquement la propagation et la collision de solitons de Manakov lorsque ceux-ci sont soumis à la PMD. Si la propagation de ces solitons est similaire à ceux de NLS (forme inchangée et translation du profil), la collision entre deux solitons est beaucoup plus compliquée, puisqu'elle fait intervenir la notion de polarisation des deux composantes. Lorsque l'on rajoute en plus les changements de polarisation induits par la biréfringence, la collision de deux solitons de Manakov devient alors extrêmement difficile à analyser. Dans une seconde étude, on s'intéresse à l'évolution des statistiques de la PMD. On effectue des simulations de Monte-Carlo et on introduit pour cela une méthode de réduction de variance qui constitue une généralisation de la méthode de Romberg-Statistique, utilisée par Kebaier [66] dans le cadre de processus de diffusion. Le principe de la méthode, assez

simple à mettre en œuvre, est basé sur la méthode de variable de contrôle et s'adapte facilement à d'autres équations aux dérivées partielles stochastiques. On note X_j^n une approximation de $X(z_n, t_j)$ par un schéma de différences finies sur une grille fine et X_l^m une approximation de $X(z_m, t_l)$ sur une grille grossière. La méthode consiste à estimer $\mathbb{E}\left(f\left(\widetilde{X}_j^n\right)\right)$ par la méthode de Monte-Carlo en introduisant la variable

$$Q = f\left(X_{j,k}^n\right) - f\left(X_{l,k}^m\right) + \mathbb{E}\left(f\left(X_l^m\right)\right),$$

où $\left(X_{j,k}^n\right)_{1\leqslant k\leqslant K}$ est un échantillon i.i.d de K variables X_j^n. L'approximation de la solution de (1.0.5) sur la grille grossière sert donc de variable de contrôle. Souvent, $\mathbb{E}\left(f\left(X_l^m\right)\right)$ n'est pas connue et on doit aussi l'estimer par la méthode de Monte-Carlo. L'estimateur par variables de contrôle, sans biais et fortement consistant, est donné par

$$\widehat{\mu}_l^n = \frac{1}{N_n}\sum_{k=1}^{N_n}\left[f\left(X_{j,k}^n\right) - f\left(X_{l,k}^m\right)\right] + \frac{1}{N_m}\sum_{k=1}^{N_m} f\left(\widehat{X}_{l,k}^m\right),$$

où N_n et N_m sont le nombre de trajectoires respectivement sur la grille fine et la grille grossière. Les variables aléatoires $X_{j,k}^n$ et $X_{l,k}^m$ sont calculées en utilisant les mêmes trajectoires Browniennes qui doivent être indépendantes de celles utilisées pour le calcul de $\widehat{X}_{l,k}^m$. Toute la difficulté de la méthode repose sur le choix de paramètres optimaux pour minimiser la complexité de l'algorithme, et qui est un arbitrage entre l'erreur de discrétisation du schéma et l'erreur statistique de l'approximation de l'espérance.

Chapitre 2

Modélisation de la propagation de la lumière dans les systèmes de transmission optique en présence de biréfringence aléatoire

Dans ce premier chapitre, on introduit les notions physiques qui ont motivé l'étude mathématique des chapitres suivants ainsi que la problématique qui s'est posée aux ingénieurs et physiciens lorsque ceux-ci ont étudié l'évolution de la lumière dans les fibres optiques longues distances. Outre les phénomènes de dispersion habituels, ils ont observé un nouveau phénomène plus complexe, appelé Dispersion Modale de Polarisation (Polarization Mode Dispersion (PMD) en anglais) qui a son origine dans la biréfringence aléatoire. Ce phénomène est dû au caractère vectoriel de la lumière et à la façon dont celle-ci se propage dans la fibre optique. L'équation de Schrödinger non linéaire ne permettant pas de modéliser un phénomène aussi complexe, Menyuk et al. ont dérivé dans [76, 77, 78] un système couplé d'équations de Schrödinger valide dans les fibres optiques biréfringentes. Ce travail a été le point de départ d'une série de travaux aboutissant à l'équation de Manakov PMD [75, 107, 108, 109, 110] qui tient compte de la biréfringence (linéaire) aléatoire. On présente dans ce chapitre les notions de base concernant la propagation de la lumière dans les fibres optiques [2, 14, 39, 61]. On présente ensuite la dérivation formelle de l'équation de Manakov PMD, à partir de [76, 77, 78], et en particulier, on mettra en évidence les différences permettant d'aboutir à cette équation plutôt qu'à l'équation de Schrödinger. Depuis que l'hypothèse selon laquelle la biréfringence varie aléatoirement le long de la fibre a été admise, les physiciens ont essayé de déterminer les statistiques (loi, moments, ...) de la PMD afin de mieux comprendre son comportement [39, 42]. L'estimation de ces statistiques fera l'objet des simulations numériques du chapitre 5.

2.1 Les fibres optiques

La fibre optique est un guide d'onde diélectrique transparent[(1)] (généralement fabriqué en silice SiO_2) et à symétrie de révolution constitué d'un cœur, où circule la lumière, d'une gaine optique et d'une gaine plastique pour la protéger. La figure 2.1 représente la vue d'une coupe transversale d'une fibre optique. Depuis qu'elle est

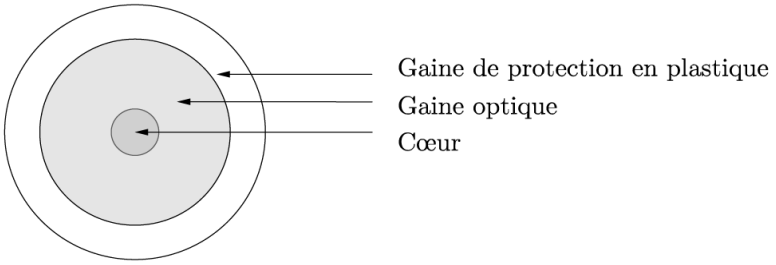


FIG. 2.1: Schéma de la coupe transversale d'une fibre optique.

utilisée comme canal de communications, la fibre optique est devenue incontournable dans les télécommunications en remplaçant le fil électrique pour le téléphone puis pour internet. L'information numérique (constituée des 0 et 1 du code binaire) est transformée en impulsions lumineuses par un laser, puis injectée au cœur de la fibre via une lentille. On appelle modulation l'opération transformant le code binaire en impulsions lumineuses et démodulation l'opération inverse. L'onde transportant le signal s'appelle l'onde porteuse. La lumière peut se propager le long de la fibre car celle-ci joue le rôle de guide d'onde grâce à la loi de Snell-Descartes et au principe de réflexion totale. En effet, les indices de réfraction du cœur (n_c) et de la gaine (n_g) sont choisis tel que $n_c > n_g$ et $n_c - n_g \sim 10^{-3}$. On parle alors d'ondes faiblement guidées. Dans une fibre parfaite, l'indice de réfraction du cœur (n_c) ne dépend que de la distance par rapport à l'axe de symétrie de la fibre. Lorsqu'il est constant dans tout le cœur, on parle de fibres à saut d'indice. Il existe également des fibres optiques plus complexes, appelées fibres à gradient d'indice, pour lesquelles l'indice de réfraction décroit graduellement à l'intérieur du cœur du centre de la fibre jusqu'à la gaine. Dans ce cas, les rayons sont courbes. Lorsque la lumière incidente est dans un cône d'acceptance, le brusque changement de l'indice de réfraction à l'interface cœur-gaine entraîne une réflexion totale de la lumière (principe de la réflexion interne totale) et il n'y a pas de réfraction [1, 2] (voir figure 2.2).

[(1)] permettant la propagation de la lumière

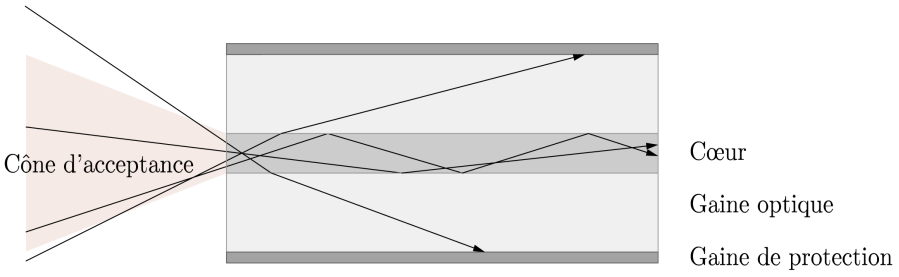


FIG. 2.2: La fibre optique à saut d'indice comme un guide d'onde pour la lumière.

Il existe deux familles de fibres optiques pour différentes fonctionnalités : la fibre multimode et la fibre monomode (Single Mode Fiber). Les fibres multimodes sont apparues les premières car plus faciles et moins chères à fabriquer et à connecter entre elles. Elles sont principalement utilisées sur de courtes distances (dans un immeuble ou un campus par exemple) tandis que la fibre monomode, apparue dans les années 1980, est utilisée sur de longues distances (transatlantique[(2)]) pour des réseaux optiques à très haut débit. La principale différence entre ces deux types de fibres réside dans le diamètre du cœur. Il est de l'ordre de 25 μm pour la fibre multimode tandis qu'il est inférieur à 10 μm pour la fibre monomode. Une fibre optique multimode accepte alors plusieurs modes de propagation (comme le montre la figure 2.2) tandis que la fibre monomode (figure 2.3) n'en supporte qu'un dont la direction de propagation est donnée par l'axe horizontal de la fibre (d'où son nom). Un mode (de propagation) correspond en fait à un vecteur propre, associé à un nombre d'onde, pour une équation de Maxwell stationnaire dont l'expression dépend des propriétés du guide d'onde (comme on le verra dans la section suivante). Lorsque le guide d'onde est une fibre optique parfaite, les modes propres sont les solutions de l'équation de Helmholtz. Cette équation peut être résolue en utilisant la symétrie de révolution de la fibre autour de l'axe de propagation z et les coordonnées cylindriques [2]. Il s'avère que le nombre de fonctions propres dépend du rayon du cœur de la fibre et de la différence des indices de réfraction à l'interface cœur-gaine [2, 11]. Pour les fibres monomodes, celui-ci est suffisament petit pour qu'il n'existe qu'un seul mode de propagation. Néanmoins, comme on le verra dans la suite, la fibre optique n'est en général pas parfaite, c'est-à-dire qu'elle présente des asymétries et que sa réponse linéaire est anisotrope. Les modes propres sont alors solutions d'une équation beaucoup plus compliquée dont l'expression est donnée dans la section 2.4 [68, 71, 77, 78]. Dans cette thèse, on considérera des communications longues distances et donc des fibres monomodes.

[(2)] Des exemples de connexion sous marines peuvent être trouvés à l'adresse suivante : http://igm.univ-mlv.fr/~dr/XPOSE2009/Transmission_sur_fibre_optique/cablesousmarin.html

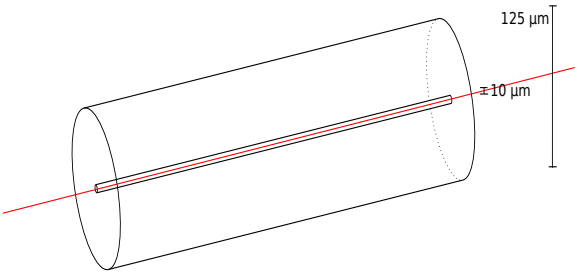


FIG. 2.3: Fibre optique monomode.

L'utilisation de la lumière comme vecteur d'information via les fibres optiques n'a été possible que grâce à l'invention du laser en 1960 et à la conception de fibres optiques à faible taux d'atténuation (affaiblissement du signal au cours de la propagation). Le laser étant modulable à des fréquences élevées et caractérisé par une forte intensité, il permit la transmission de l'information sur de grandes distances. C'est en 1966 qu'a eu lieu la première expérimentation de transmission de données par fibre optique. Quelques années plus tard, en 1970, le design de fibres à faible absorption a permis l'utilisation de la fibre optique dans les réseaux de télécommunications. L'énorme avantage des fibres optiques, combiné à l'utilisation du laser, est que la fréquence des impulsions lumineuses est bien plus élevée que celles des ondes électromagnétiques dans les câbles coaxiaux (dont le cœur est en métal), ce qui a permis d'atteindre des débits de communications jusqu'alors inenvisageable. Dans les fibres optiques très récentes, ce débit peut dépasser le térabit/seconde en comparaison des 10 Gbit/seconde pour les câbles coaxiaux. Néanmoins, bien que la fibre optique soit un support à faible atténuation, c'est aussi un milieu dispersif. C'est ainsi la cause de plusieurs phénomènes (linéaires et non linéaires) entraînant une déformation du signal. L'impact de ces phénomènes sur la dégradation du signal est d'autant plus important que la distance est élevée. Ils limitent donc de fait la distance maximale de transmission et les débits.

2.2 La polarisation de la lumière

La lumière est une onde électromagnétique transversale (i.e. qui évolue dans un plan perpendiculaire à l'axe de propagation). En optique, on ne s'intéresse qu'à l'évolution du champ électrique ; les équations de Maxwell permettant de retrouver la donnée du champ magnétique (car la magnétisation dans une fibre est nulle). On considérera dans la suite que l'axe de propagation est donné par la direction z. Le champ électrique $\mathbf{E}$ associé à la lumière est alors donné par le champ de vecteur $\mathbf{E} = (E_x, E_y, 0)$.

L'état de polarisation d'une onde décrit la façon dont évolue, au cours du temps,

l'extrémité du vecteur champ électrique autour de l'axe de propagation z [61]. On dit que la lumière est naturelle ou non polarisée lorsqu'elle se propage de façon désordonnée dans l'espace. On dit qu'elle est cohérente ou polarisée (laser) lorsque l'évolution de la direction du vecteur champ électrique décrit une ellipse autour de z qui peut, selon les cas, dégénérer en une droite ou un cercle. L'avantage de telles ondes est que l'étude de l'évolution du champ électromagnétique associée est prévisible. Afin de décrire la polarisation d'une onde électromagnétique, il existe plusieurs formalismes complémentaires, dont celui de Jones et de Stokes-Mueller. Nous utiliserons surtout ce dernier, qui a l'avantage de pouvoir être utilisé dans des milieux dépolarisants. Une onde est dite partiellement polarisée lorsque sa polarisation ne reste pas constante et change le long de la fibre. En revanche, l'inconvénient de cette représentation est qu'elle ne tient pas compte de la phase de l'onde (contrairement au formalisme de Jones qui tient compte de la phase mais ne permet de traiter que des ondes totalement polarisées). Pour une description plus complète de l'état de polarisation d'une onde plane, on pourra se référer à l'ouvrage [61].

2.2.1 Le formalisme de Jones

L'onde plane monochromatique joue un rôle essentiel dans l'étude de la polarisation des ondes lumineuses. En effet, toute solution des équations de Maxwell (i.e. tout champ électromagnétique) peut s'écrire à l'aide de la transformée de Fourier comme une superposition d'ondes planes monochromatiques

$$\mathbf{E}\left(\mathbf{r},t\right)=\frac{1}{(2\pi)^4}\int_{\mathbb{R}^3}\int_{\mathbb{R}}\widehat{\mathbf{E}}\left(\mathbf{k},\omega\right)e^{-i(\omega t-\mathbf{kr})}d\mathbf{k}d\omega.$$

On présentera donc la polarisation d'une onde lumineuse dans le contexte des ondes planes monochromatiques. Considérons donc une onde plane monochromatique se propageant dans la direction z

$$\begin{aligned}\mathbf{E}(z,t) &= 2\mathcal{R}\mathrm{e}\left\{u_x e^{i(k_x z-\omega t+\varphi_x)}\widehat{\mathbf{e}}_x+u_y e^{i(k_y z-\omega t+\varphi_y)}\widehat{\mathbf{e}}_y\right\}\\ &= u_x\cos\left(k_x z-\omega t+\varphi_x\right)\widehat{\mathbf{e}}_x+u_y\cos\left(k_y z-\omega t+\varphi_y\right)\widehat{\mathbf{e}}_y\end{aligned}$$

où $\widehat{\mathbf{e}}_x$ et $\widehat{\mathbf{e}}_y$ sont des vecteurs unitaires définissant le plan d'onde, perpendiculaires à l'axe de propagation $\widehat{\mathbf{e}}_z$, des deux modes. Pour simplifier, on considère que $(\widehat{\mathbf{e}}_x,\widehat{\mathbf{e}}_y,\widehat{\mathbf{e}}_z)$ est la base canonique de $\mathbb{R}^3$ et constitue le repère de travail (selon lequel sera étudié la polarisation). De plus u_x et u_y sont des constantes réelles positives définissant l'amplitude de l'onde ; φ_x et φ_y correspondent à des translations de phases de chacune des composantes et dont les valeurs varient entre $-\pi$ et π. On obtient en $z=0$ l'équation d'une ellipse

$$\left(\frac{E_x}{u_x}\right)^2+\left(\frac{E_y}{u_y}\right)^2-2\frac{E_xE_y}{u_xu_y}\cos\left(\varphi\right)=\sin^2\left(\varphi\right),\tag{2.2.1}$$

où $\varphi=\varphi_y-\varphi_x$ est le déphasage entre les deux polarisations orthogonales E_x et E_y. L'évolution de l'extrémité du vecteur champ électrique $\mathbf{E}$ dans un plan orthogonal à sa direction de propagation est donc une ellipse qui décrit l'état de polarisation de la lumière. Il est clair que suivant les valeurs du déphasage φ (c'est-à-dire du

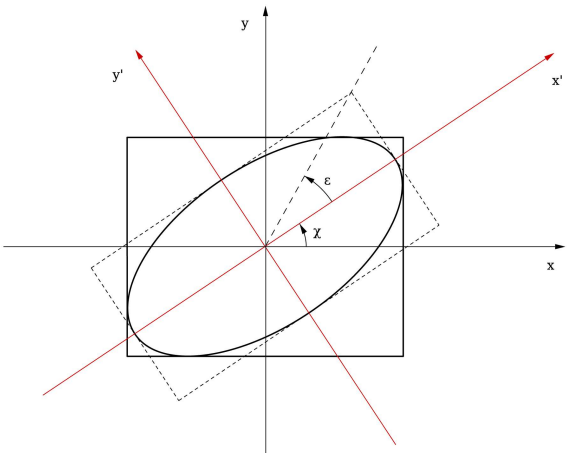


FIG. 2.4: Représentation de l'évolution du vecteur **E** dans le cas d'une polarisation elliptique. Les angles χ et ε correspondent respectivement à l'inclinaison de l'ellipse et à son ellipticité.

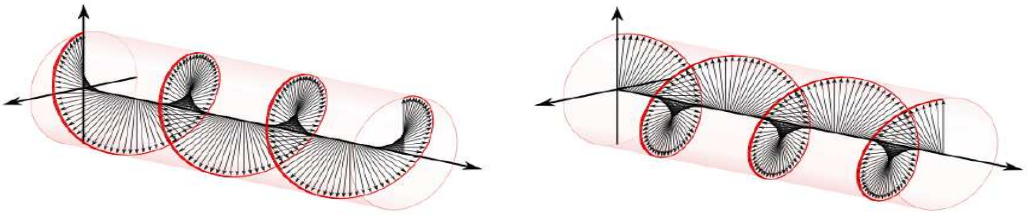


FIG. 2.5: Polarisation circulaire gauche et droite.

retard d'une des composantes par rapport à l'autre), l'ellipse aura une forme différente. Lorsque $\varphi = 0$ ou est un multiple de π, l'ellipse dégénère en une droite. Les deux composantes oscillent en même temps et on parle de polarisation rectiligne ou linéaire. Lorsque le déphasage est multiple de $\pi/2$, l'équation (2.2.1) est celle d'un cercle et on parle alors de polarisation circulaire. Sinon, on parle de polarisation elliptique. La figure 2.4 décrit un état de polarisation elliptique en fonction des paramètres de l'ellipse. Le signe de φ détermine le sens de rotation de la polarisation [61]. La rotation est dite directe ou droite (sens des aiguilles d'une montre) si φ est compris entre 0 et π et rétrograde ou gauche (sens inverse des aiguilles d'une montre) si φ est compris entre $-\pi$ et 0. La figure 2.5 illustre l'évolution du champ électrique de polarisation circulaire gauche et droite. Le vecteur complexe de Jones **J** donné par

$$\mathbf{J} = \left(u_x e^{-i\varphi_x}\widehat{\mathbf{e}}_x + u_y e^{-i\varphi_y}\widehat{\mathbf{e}}_y\right)$$

permet donc de décrire complètement l'état de polarisation d'une onde plane monochromatique (et donc dans des milieux non dépolarisant). En normalisant les vecteurs de Jones **J**, un état de polarisation peut donc être représenté par un point sur la sphère $\mathbb{S}^3 \subset \mathbb{R}^4$. Néanmoins, cette représentation fait intervenir les amplitudes des champs complexes et la phase (φ_x, φ_y) qui sont des quantités non mesurables en pratique. On préfèrera alors utiliser la représentation de Stokes qui a l'avantage de

ne s'écrire qu'en fonction des intensités des deux composantes.

2.2.2 Le formalisme de Stokes et la représentation de Poincaré

Le formalisme de Stokes est particulièrement utile pour décrire l'état de polarisation d'une onde plane monochromatique car tout état de polarisation peut être représenté de façon unique par un point sur la sphère unité $\mathbb{S}^2 \subset \mathbb{R}^3$, la sphère de Poincaré. Dans ce cas, le vecteur de Stokes est défini à l'aide du vecteur de Jones par la relation [61]

$$S = \mathbf{J}^* \overrightarrow{\sigma} \mathbf{J},$$

où $\overrightarrow{\sigma} = (\sigma_0, \sigma_3, \sigma_1, \sigma_2)$ est le vecteur des matrices de Pauli qui est donné par

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Les composantes de ce vecteur se réecrivent

$$\begin{pmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} |u_x|^2 + |u_y|^2 \\ |u_x|^2 - |u_y|^2 \\ 2u_x u_y \cos(\varphi) \\ 2u_x u_y \sin(\varphi) \end{pmatrix}.$$

Lorsque l'on normalise le vecteur de Stokes celui-ci évolue alors sur la sphère $\mathbb{S}^2 \subset \mathbb{R}^3$ car $S_0^2 = S_1^2 + S_2^2 + S_3^2$. La composante S_1 décrit le pourcentage de polarisation linéaire horizontale ou verticale, S_2 le pourcentage de polarisation linéaire à $\pm 45°$ et S_3 le pourcentage de polarisation circulaire. Ainsi, les polarisations rectilignes sont situées sur l'équateur de la sphère, les polarisations circulaires gauches et droites sont situées au pôles sud et nord de la sphère. Les polarisations elliptiques droites dans l'hémisphère nord et elliptiques gauches dans l'hémisphère sud. La figure 2.6 illustre ces propos. Deux points diamétralement opposés sur la sphère de Poincaré sont dits orthogonalement polarisés. Au contraire, lorsque deux points sont proches sur la sphère de Poincaré, alors les polarisations sont similaires. En coordonnées sphériques, les coefficients de Stokes se réécrivent facilement en fonction des angles d'inclinaison χ et d'ellipticité ε

$$\begin{pmatrix} S_1/S_0 \\ S_2/S_0 \\ S_3/S_0 \end{pmatrix} = \begin{pmatrix} \cos(2\varepsilon)\cos(2\chi) \\ \cos(2\varepsilon)\sin(2\chi) \\ \sin(2\varepsilon) \end{pmatrix}.$$

La figure 2.7 montre une représentation alternative de l'état de polarisation de la lumière en fonction des paramètres de l'ellipse. Pour la lumière naturelle (ou par-

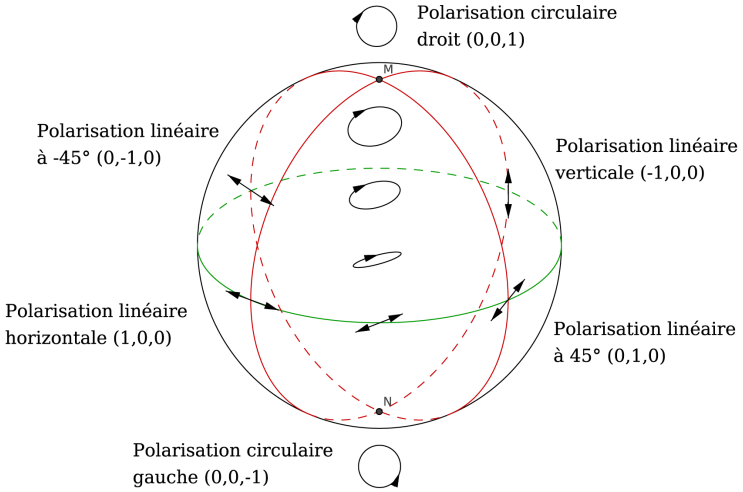


FIG. 2.6: Représentation des différents états de polarisation sur la sphère de Poincaré.

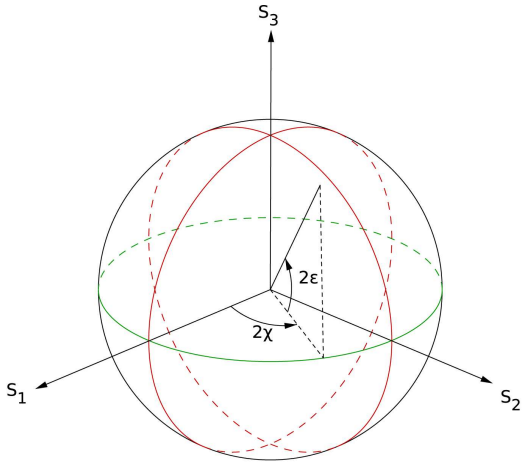


FIG. 2.7: Représentation des coefficients de Stokes sur la sphère de Poincaré.

tiellement polarisée) le vecteur de Stokes $S = (S_0, S_1, S_2, S_3)$ est défini un peu différemment. Il est cette fois donné par

$$S(z) = \langle \mathbf{J}, \overrightarrow{\sigma}\mathbf{J} \rangle (z) = \frac{1}{2T} \int_{-T}^{T} \mathbf{J}^* \overrightarrow{\sigma}\mathbf{J}(z,t)dt.$$

Le degré de polarisation (Degree of Polarization (DOP)) est utile pour mesurer la dépolarisation d'une onde. Il est défini par [61]

$$P_d(z) = \frac{\sqrt{S_1^2(z) + S_2^2(z) + S_3^2(z)}}{S_0},$$

et prend ses valeurs entre 0 et 1. Quand $P_d = 1$, cela signifie que l'onde a une polarisation constante et on a l'équivalence avec la formulation précédente. Au contraire, lorsque $P_d = 0$, l'onde est complètement dépolarisée (c'est le cas de la lumière naturelle). Dans le cas de la lumière partiellement polarisée, il sera également utile d'exprimer les coefficients de Stokes de la transformée de Fourier $\widehat{\mathbf{E}}$ du champ électrique $\mathbf{E}$ [39, 42]

$$\begin{aligned}
\widehat{s}_1(\omega, z) &= \widehat{\mathbf{E}}^* \sigma_3 \widehat{\mathbf{E}}(\omega, z) = \left|\widehat{E}_x\right|^2 - \left|\widehat{E}_y\right|^2 \\
\widehat{s}_2(\omega, z) &= \widehat{\mathbf{E}}^* \sigma_1 \widehat{\mathbf{E}}(\omega, z) = 2\mathcal{R}\mathrm{e}\left(\widehat{E}_x^* \widehat{E}_y\right) \\
\widehat{s}_3(\omega, z) &= \widehat{\mathbf{E}}^* \sigma_2 \widehat{\mathbf{E}}(\omega, z) = 2\mathcal{I}\mathrm{m}\left(\widehat{E}_x^* \widehat{E}_y\right).
\end{aligned}$$

Dans le cas où les effets de la dispersion chromatique et les effets non linéaires sont négligés, le vecteur $(\widehat{s}_1/\widehat{s}_0, \widehat{s}_2/\widehat{s}_0, \widehat{s}_3/\widehat{s}_0)$ évolue lui aussi sur la sphère $\mathbb{S}^2$ de $\mathbb{R}^3$. Une courbe sur cette sphère permet alors de décrire l'évolution de la polarisation de la lumière le long de la fibre à une fréquence fixée.

Remarque 2.2.1. *Le passage d'une représentation à l'autre sera particulièrement utile lors des simulations numériques. Dans le cas où les effets de la dispersion chromatique sont négligeables, l'évolution du vecteur de Stokes est décrite, à une fréquence angulaire fixée, par une équation différentielle stochastique. Il est alors aisé de simuler numériquement l'évolution des vecteurs de Stokes. En revanche, dans le cas non linéaire, l'équation décrivant l'évolution de la PMD est beaucoup plus compliquée. On préférera donc dans un premier temps simuler le vecteur de Jones pour ensuite obtenir les coefficients de Stokes. Les figures 5.5 et 5.8 du Chapitre 5 illustrent les cas où l'état de polarisation varie respectivement de façon uniforme et aléatoire.*

2.3 Phénomènes affectant la propagation de la lumière dans une fibre optique

Bien que les avancées technologiques aient permis de construire des fibres optiques avec une atténuation extrêmement faible (jusqu'à 0.15 dB.km^{-1} pour les fibres en silice combinées à une longueur d'onde de 1.55 μm [2]), il existe de nombreux éléments de perturbations du signal dont certains peuvent conduire à des phénomènes de dispersion très complexes.

2.3.1 La dispersion

Lors de la progression d'une impulsion lumineuse dans une fibre optique, on observe que le signal s'étale. Ce phénomène est appelé dispersion et apparaît lorsque les vitesses de phases v_φ des ondes composant la lumière dépendent de leurs fréquences angulaires ω (i.e $v_\varphi(\omega)$). Les différentes ondes formant le paquet d'ondes ne voyagent donc pas à la même vitesse et le paquet d'ondes se disperse. Si la dispersion est trop importante, la détection du signal à la sortie de la fibre devient compliquée et l'information initiale est mal décodée.

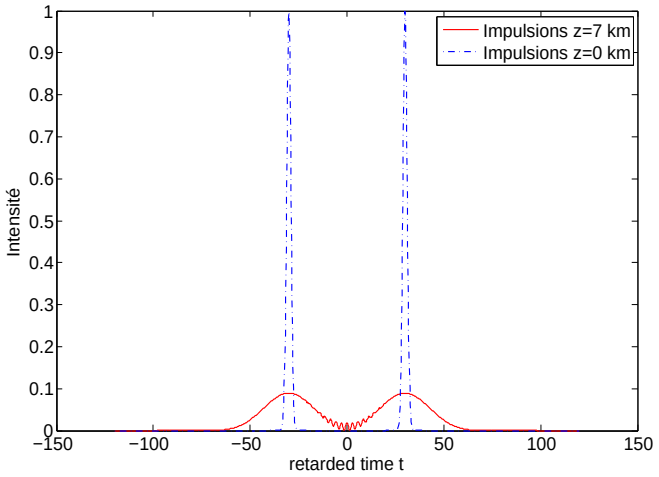


FIG. 2.8: Effets de la dispersion chromatique sur un train d'ondes. Les deux impulsions ont commencé à interagir.

Pour éviter ces interférences, il faut soit augmenter l'intervalle entre les impulsions (et donc réduire le débit), soit augmenter la distance de dispersion[(3)] à l'aide d'équipements additionnels (les répéteurs).

2.3.2 La dispersion chromatique et la dispersion modale

La dispersion chromatique résulte du fait que l'on veut générer des impulsions courtes pour augmenter le débit. Or d'après la théorie de Fourier, plus l'impulsion est courte et plus le spectre est large. Ainsi si le milieu est dispersif, toutes les longueurs d'onde ne se propageront pas à la même vitesse. Cependant, plusieurs avancées technologiques ont permis de limiter les effets de la dispersion chromatique : la fabrication de lasers et l'utilisation de fibres plus performantes (fibres à gestion de dispersion, fibres à dispersion décalée, etc.). En effet, les lasers sont capables d'émettre une lumière cohérente, caractérisée par un très haut degré de monochromaticité, une grande directionnalité et une forte intensité, rendant possible la transmission par solitons. Les fibres à gestion de dispersion consistent en une concaténation périodique

[(3)] distance de dispersion acceptable pour laquelle l'information pourra être décodée.

de tronçons de dispersion normale et anormale. Ce type de fibre a permis de réduire la dispersion chromatique pour des longueurs d'onde autour de 1.3 μm. En revanche pour ces longueurs d'onde, l'atténuation est plus forte qu'à 1.55 μm. Les fibres à dispersion décalée permettent d'atteindre une dispersion chromatique quasi-nulle pour des longueurs d'onde d'atténuation minimale c'est à dire 1.55 μm [2]. Néanmoins, pour les transmissions longues distances un juste équilibre devra être trouvé entre la dispersion chromatique et les autres effets dispersifs afin de diminuer la dispersion totale. D'un point de vue mathématique, la dispersion chromatique est décrite via un développement de Taylor du nombre d'onde $k(\omega)$ autour de la fréquence angulaire ω_0 de l'onde porteuse i.e.

$$k(\omega) := n(\omega)\frac{\omega}{c} = k(\omega_0) + k'(\omega_0)(\omega - \omega_0) + \frac{1}{2}k''(\omega_0)(\omega - \omega_0)^2 + \cdots$$

avec les notations $k'(\omega_0) = \frac{dk(\omega)}{d\omega}|_{\omega=\omega_0}$ et $k''(\omega_0) = \frac{d^2k(\omega)}{d\omega^2}|_{\omega=\omega_0}$. Les vitesses de phase et de groupe sont respectivement données par $v_\varphi = \omega/k = c/n(\omega)$ et

$$v_g = \frac{d\omega}{dk} = \frac{c}{n} - \frac{ck}{n^2}\frac{dn(\omega)}{dk}.$$

De plus,

$$\frac{dk(\omega)}{d\omega} = \frac{n}{c} + \frac{\omega}{c}\frac{dn(\omega)}{d\omega},$$

où k' est égal à l'inverse de la vitesse de groupe v_g tandis que k'' représente la dispersion de la vitesse de groupe (Group Velocity Dispersion (GVD) en anglais). Il correspond à l'étalement de l'impulsion causé par la dépendance de la vitesse de groupe à la fréquence. L'ordre supérieur k''' correspond à une distorsion d'impulsion ultra-courte et on la négligera dans la suite. On notera également l_d la distance de dispersion, correspondant à la distance limite à une dispersion chromatique tolérable.

On parle de dispersion modale lorsque les rayons lumineux composant le signal peuvent parcourir la fibre selon des modes (ou directions) différents. Bien sûr, cette notion de dispersion fait intervenir la notion de fibres multimodes et monomodes. Dans les fibres multimodes, il y a dispersion modale à cause du large diamètre du cœur (voir figure 2.2). Les rayons lumineux ne se propageant pas dans la même direction, ils n'arriveront pas en même temps à la sortie de la fibre. Les fibres optiques multimodes ont donc une capacité de transmission limitée et c'est la raison pour laquelle elles ne sont utilisées que sur de courtes distances. Théoriquement, il n'y a pas de dispersion modale dans les fibres monomodes car les rayons lumineux ne peuvent se propager que dans une direction : l'axe horizontal de la fibre. On verra dans la section suivante qu'en raison de la présence de dispersion modale de polarisation une fibre monomode est plutôt bi-modale.

2.3.3 La biréfringence et la dispersion modale de polarisation

La biréfringence, aussi connue sous le nom de "double réfraction", est un phénomène qui apparaît lorsque l'on tient compte de l'aspect vectoriel de la lumière. A l'entrée de la fibre, l'onde lumineuse se divise en deux composantes (ou modes) dont

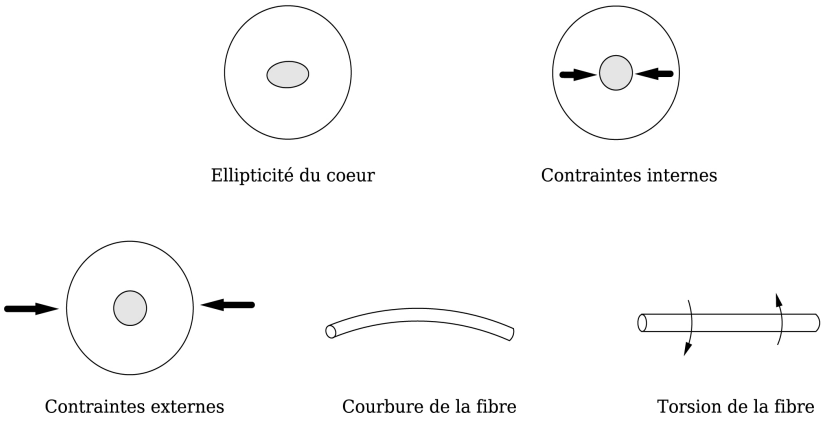


FIG. 2.9: Causes de la biréfringence dans la fibre.

les polarisations sont orthogonales (i.e. diamétralement opposées sur la sphère de Poincaré). On verra que selon le type de biréfringence que l'on considère, ces deux modes ont des polarisations différentes. Lorsque ces deux modes n'ont pas les mêmes caractéristiques, parce que le milieu n'est pas parfaitement isotrope, les impulsions s'élargissent et le signal est dégradé. C'est ce qu'on appelle la dispersion modale de polarisation (Polarization Mode Dispersion (PMD) en anglais).

Les fibres optiques monomodes sont donc plutôt bi-modales car elles permettent en fait la propagation de ces deux modes polarisés orthogonalement. Dans une fibre optique monomode idéale, la section transverse possède une parfaite symétrie de révolution. Dans ce cas, les deux modes se propagent avec les mêmes caractéristiques : vitesse de phase, vitesse de groupe, dispersion chromatique etc. On dit alors qu'ils sont dégénérés. Cependant, au cours du processus de fabrication, la fibre peut présenter certains défauts (voir Fig. 2.9) tels qu'une ellipticité du cœur ou alors des contraintes internes (lors du processus de fabrications) ou externes (liées à l'environnement extérieur par exemple des variations de température ou le vieillissement du câble). Tous ces phénomènes détruisent la dégénérescence des deux modes et induisent de la biréfringence dans la fibre optique, c'est-à-dire une différence de vitesse de phase entre les deux modes. Cette différence de vitesse de phase (même petite) crée un déphasage φ entre les deux polarisations qui se traduit alors par une évolution des états de polarisation du champ électrique le long de la fibre. Il apparaît alors localement un axe rapide et un axe lent de propagation dont les indices de réfraction sont respectivement donnés par n_1 et n_2. La différence des indices de réfraction effectifs est donnée par $\Delta n = n_1 - n_2$. Dans les configurations réalistes, il est de l'ordre de $10^{-6} - 10^{-7}$, ce qui est beaucoup plus petit que la différence des indices de réfraction entre le cœur de la fibre et la gaine optique. La force de biréfringence mesure la différence entre les nombres d'ondes des deux composantes

$$b(z, \omega) = k_1(z, \omega) - k_2(z, \omega).$$

Néanmoins dans de nombreuses applications, la force de biréfringence est évaluée en

la fréquence de l'onde porteuse ω_0 c'est à dire que $b(z) = b(z, \omega_0)$ et donc

$$b(z) = \frac{\omega_0 n_1(z)}{c} - \frac{\omega_0 n_2(z)}{c} = \frac{\omega_0 \Delta n(z)}{c} = k_0 \Delta n(z) = \frac{2\pi}{\lambda_0} \Delta n(z),$$

où $k_0 = k(\omega_0)/n$ et $\lambda_0 = n\lambda$ sont respectivement le nombre d'onde et la longueur d'onde de la lumière dans le vide. On définit également la longueur de battement L_B (ou beat length en Anglais)

$$L_B = \frac{\lambda_0}{\Delta n} = \frac{2\pi}{b}.$$

Dans les fibres optiques monomodes, la biréfringence est généralement considérée comme importante (b grand) et la longueur de battement associée est très courte, de l'ordre du mètre, ce qui est beaucoup plus petit que la longueur de la fibre.

Lorsque l'anisotropie est homogène et en l'absence d'effets non linéaires, il existe deux axes tels que si l'onde incidente est polarisée selon l'un des deux, alors sa polarisation reste inchangée au cours de sa propagation. Ces deux états de polarisations sont appelés états propres de polarisation. Les axes associés sont l'axe lent et rapide et constituent une base de décomposition pour l'ensemble des états de polarisation d'une onde plane monochromatique. Dans tous les autres cas, lorsque la polarisation de l'onde incidente n'est pas orientée selon l'un de ces axes, les deux modes échangent leurs énergies de façon périodique au cours de leur évolution le long de la fibre sur une période donnée par L_B. Autrement dit, si ceux-ci sont initialement en phase, alors ils le sont de nouveau après une distance donnée par L_B. Sur la sphère de Poincaré, cela signifie que le vecteur de Stokes revient à sa position initiale après une période L_B après avoir dessiné un cercle sur la sphère (voir Fig. 5.5). Si l'on considère que les axes propres sont donnés par la base canonique $(\widehat{\mathbf{e}}_x, \widehat{\mathbf{e}}_y, \widehat{\mathbf{e}}_z)$ de $\mathbb{R}^3$, alors les états propres de polarisation sont les polarisations linéaires horizontales et verticales (dans le cas de la biréfringence linéaire). Ainsi, dans une fibre optique linéairement biréfringente, si l'onde incidente est linéairement polarisée à 45° par rapport aux axes propres, alors sa polarisation va varier de façon cyclique selon la figure 2.10.

Les perturbations précédentes n'ont pas toutes les mêmes effets sur la fibre et n'entraînent pas le même type de biréfringence. La biréfringence linéaire résulte des contraintes internes ou externes tandis que la biréfringence elliptique résulte de torsions de la fibre. Lorsqu'on parle de biréfringence, on sous-entend en général biréfringence linéaire, c'est-à-dire qu'en l'absence d'effet Kerr, seule une onde plane polarisée rectilignement peut se propager sans déformation. Par analogie, on utilise l'expression biréfringence circulaire (ou activité optique) pour désigner un milieu dans lequel les états propres de polarisation sont les polarisations circulaires droite et gauche. Dans un milieu circulairement biréfringent, les ondes de polarisation circulaire se propagent sans se déformer. On parle de biréfringence elliptique lorsque les deux phénomènes sont présents. On décompose cette fois le champ électrique dans la base orthonormale complexe (pour le produit scalaire hermitien)

$$\widehat{\mathbf{e}}_1 = \frac{\widehat{\mathbf{e}}_x + ir\widehat{\mathbf{e}}_y}{\sqrt{1+r^2}} \qquad \text{et} \qquad \widehat{\mathbf{e}}_2 = \frac{r\widehat{\mathbf{e}}_x - i\widehat{\mathbf{e}}_y}{\sqrt{1+r^2}}$$

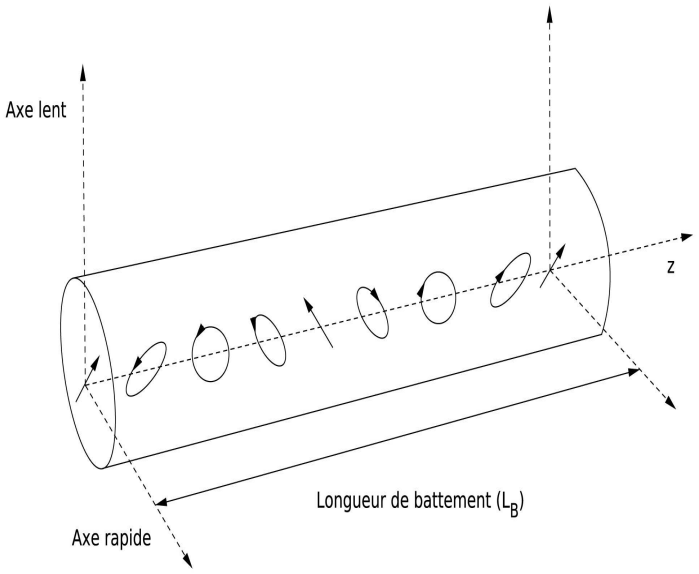


FIG. 2.10: Evolution périodique de la polarisation induite par de la biréfringence linéaire.

où $r = \tan(\phi/2)$ et l'angle ϕ caractérise l'ellipticité des axes propres de polarisation. La biréfringence linéaire correspond au cas où $\phi = 0$, la biréfringence circulaire au cas où $\phi = \pi/2$. Les expériences réalisées jusque-là sur les fibres optiques monomodes tendent à confirmer que la biréfringence est principalement linéaire, on considérera donc que $\phi = 0$.

Le fait que les indices de réfraction des deux modes dépendent de la fréquence est donc à l'origine de la dispersion modale de polarisation, qui se traduit par une différence de vitesse de groupe entre les deux modes (voir Fig. 2.11) induisant un étalement du signal à la sortie de la fibre lorsque les deux modes ne font plus qu'un. On note $\Delta\tau$ cette différence de vitesse de groupe (Differential Group Delay (DGD)) entre le mode lent et le mode rapide qui se définit en ps.km^{-1} par

$$\frac{\Delta\tau}{l} = \frac{d}{d\omega}\left(\frac{\omega\Delta n}{c}\right) = \frac{\Delta n}{c} + \frac{\omega}{c}\frac{d\Delta n}{d\omega}$$

avec l la longueur de la fibre. Néanmoins, la PMD ne se manifeste pas seulement via cette différence de vitesse de groupe. On a déjà mentionné que les causes de la biréfringence étaient variées et que par conséquent elle ne pouvait pas être considérée comme uniforme. On appelle angle de biréfringence θ l'angle de rotation des axes locaux de biréfringence par rapport aux axes propres initiaux. On considère donc que les paramètres de la biréfringence (b, θ) peuvent varier aléatoirement le long de la fibre, ce qui implique que la polarisation du champ électrique évolue elle aussi de manière aléatoire (voir Fig. 5.8) en fonction des propriétés locales de la biréfringence. Ces variations entraînent un échange d'énergie entre les deux modes qu'on appelle couplage des modes (Polarization Mode Coupling). Contrairement à la dispersion chromatique, la PMD évolue de façon dynamique et aléatoire le long de la fibre, ce qui la rend difficile à contrôler. L'effet dévastateur est dû à l'accumulation de la

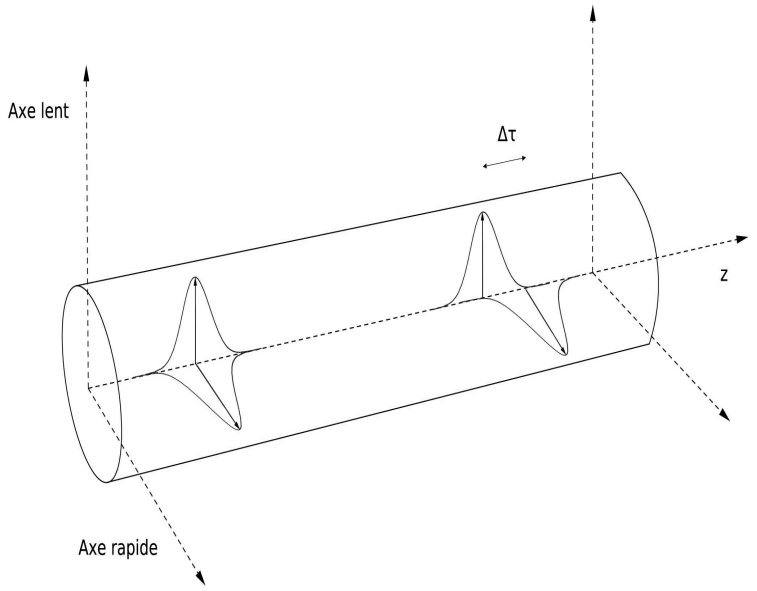


FIG. 2.11: Effet de la PMD sur une impulsion linéairement polarisée à 45° par rapport aux axes de biréfringence sur un tronçon de fibre optique biréfringente.

PMD et non à sa valeur locale.

Ces variations sont caractérisées par la longueur de corrélation l_c (ou longueur de couplage). Cette longueur correspond à la distance à partir de laquelle l'état de polarisation de la lumière ne dépend plus de l'état de polarisation initial. Généralement, elle est de l'ordre de plusieurs dizaines de mètres. Lorsque $l \ll l_c$, la longueur de la fibre est trop petite pour que le couplage des modes ait lieu. La description précédente de la PMD, induite par la biréfringence uniforme, est alors valide et $\Delta\tau$ varie linéairement. En revanche, lorsque $l_c \ll l$, la fibre est considérée longue distance et la valeur moyenne de $\Delta\tau$ varie comme la racine carrée de la longueur de la fibre.

2.3.4 Effets non linéaires

On parle de fibres optiques linéaires quand le matériel réagit linéairement à l'intensité du champ électrique. L'optique non linéaire concerne les phénomènes physiques apparaissant lorsqu'un milieu matériel est soumis à un faisceau lumineux suffisamment intense (le laser par exemple) pour modifier la réponse du milieu au champ électromagnétique. De nouveaux processus vont intervenir, donnant lieu à une grande richesse de phénomènes comme par exemple l'automodulation de phase (Self-Phase Modulation (SPM) en anglais), la modulation de phase croisée (Cross-Phase Modulation (XPM)) et le mélange à quatre ondes dégénéré (degenerate Four wave mixing (FWM)). La première démonstration expérimentale de l'existence de ces phénomènes non linéaires est apparue en 1961 grâce à P.Franken.

On parle d'effet Kerr lorsque l'indice de réfraction du matériau varie proportionnellement à l'intensité du champ électrique E. Dans le cas de fibres optiques

linéairement biréfringentes, les variations des indices de réfraction n_x et n_y des deux modes sont données par

$$\begin{cases} \Delta n_x &= n_2 \left(|E_x|^2 + \frac{2}{3} |E_y|^2 + \frac{1}{3} E_x^* E_y^2 \right) \\ \Delta n_y &= n_2 \left(|E_y|^2 + \frac{2}{3} |E_x|^2 + \frac{1}{3} E_y^* E_x^2 \right) \end{cases} \tag{2.3.1}$$

où n_2 est l'indice de réfraction non linéaire du matériau indiquant la proportion dans laquelle l'indice varie en fonction de l'intensité du champ. Le premier terme correspond à l'automodulation de phase, le second à la modulation de phase croisée et le dernier terme correspond au mélange à quatre ondes dégénéré. L'indice de réfraction n_j est alors déterminé par

$$n_j = n_L + \Delta n_j, \qquad j = x, y.$$

2.4 Dérivation de l'équation de Manakov PMD

Dans cette partie, on dérive l'équation de Manakov PMD à partir des équations de Maxwell et on donne les principaux arguments permettant d'obtenir ce système. Pour une dérivation complète du système de Manakov PMD, on pourra consulter les articles [77, 78, 75, 76, 108, 109, 110] ainsi que les monographies [2, 55, 101]. Dans un premier temps, on insiste sur les hypothèses amenant à ce système d'équation plutôt qu'à l'équation de Schrödinger non linéaire. On fait les hypothèses suivantes :

Hypothèses

1) On tient compte du caractère vectoriel de la lumière. La dérivation de l'équation de Schrödinger suppose que l'état de polarisation de la lumière incidente reste identique lors de sa propagation le long de la fibre. Les deux modes propres ont alors exactement les mêmes caractéristiques (vitesses de phase, vitesse de groupe, etc.) et voyagent indépendamment (pas d'échanges d'énergies). On peut alors considérer que le champ électrique est scalaire (la donnée d'un mode propre suffit à retrouver l'autre). Ce n'est plus le cas si on s'intéresse aux fibres optiques biréfringentes.
2) On suppose que la fibre optique est monomode. Il n'existe donc qu'un axe de propagation possible (l'axe de la fibre) dont la coordonnée sera notée z.
3) On suppose que la lumière incidente est un laser (lumière quasi-monochromatique) et donc que le spectre en fréquence est localisé autour de la fréquence de l'onde porteuse ω_0.
4) Lors de la dérivation, on fera une approximation paraxiale (perturbation d'une superposition d'ondes planes monochromatiques) des équations de Maxwell. La prise en compte des variations transverses du champ électrique dans la fibre ne change pas la forme finale de l'équation et on pourra se placer dans le cadre d'une approximation par ondes planes, dans les variables transverses, du champ électrique.
5) On supposera également que l'amplitude est une fonction lentement variable en temps et dans la direction de propagation z.

6) On suppose que la réponse linéaire de la fibre a une composante anisotrope et de fait que la fibre est biréfringente dans la direction de propagation z. L'anisotropie signifie que les propriétés du milieu ne sont pas identiques selon les modes de propagation. Cette composante ne sera pas négligée et conduira à la biréfringence qu'on supposera importante (par rapport à la dispersion chromatique et aux effets non linéaires).

7) En revanche, on suppose que la réponse non linéaire est essentiellement isotrope.

8) On se place dans une configuration réaliste lors de l'étude des fibres optiques monomodes, à savoir que la longueur de battement L_B est de l'ordre du mètre tandis que la longueur de corrélation l_c est de l'ordre d'une centaine de mètres. Les longueurs l_d et l_{nl} relatives à la dispersion chromatique et à l'effet Kerr sont beaucoup plus grandes et sont de l'ordre d'une centaine de kilomètres. Ainsi, on est donc dans un régime où $L_B \ll l_c \ll l_d \sim l_{nl}$.

Dans cette section uniquement, on utilisera la définition suivante de la transformée de Fourier

$$\widehat{f}(\omega) = \int_{\mathbb{R}} f(t) e^{i\omega t} dt$$

et de la transformée de Fourier inverse

$$f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \widehat{f}(\omega) e^{-i\omega t} d\omega.$$

Sur l'échelle la plus fine, correspondant au diamètre du cœur et à la longueur d'onde (de l'ordre du μm), l'évolution d'une onde électromagnétique dans un milieu diélectrique est décrite par les équations de Maxwell. Sur une échelle plus longue, correspondant à la longueur de battement L_B et à la longueur de corrélation l_c, une approximation des équations de Maxwell permet d'obtenir une description de l'enveloppe lentement variable du champ électrique par un système couplé d'équations de Schrödinger. Enfin si on regarde l'évolution du champ électrique sur des distances encore plus longues, relatives à l_d et l_{nl}, on obtient, après moyennisation du système d'équations de Schrödinger, le système de Manakov PMD [39, 77, 78]. Dans l'analyse et la dérivation du modèle on introduira donc trois paramètres $1 \gg \delta_1 \gg \delta_2 \gg \delta_3$ proportionnels respectivement à l'inverse de L_B, l_c et $l_d \sim l_{nl}$.

2.4.1 Un système couplé d'équations de Schrödinger non linéaires

Le point de départ pour l'établissement du système de Schrödinger non linéaire et de l'équation de Manakov PMD est donc l'ensemble des équations de Maxwell décrivant la propagation des ondes électromagnétiques dans un milieu matériel en l'absence de courant et de charges électriques (car le milieu est diélectrique). On pose $\mathbf{E} = \mathbf{E}(\mathbf{r}, t)$, $\mathbf{H} = \mathbf{H}(\mathbf{r}, t)$, $\mathbf{D} = \mathbf{D}(\mathbf{r}, t)$ et $\mathbf{B} = \mathbf{B}(\mathbf{r}, t)$ qui définissent respectivement le champ électrique, le champ magnétique, l'induction électrique et l'induction magnétique de l'onde. Par ailleurs, $\mathbf{r} = (x, y, z) \in \mathbb{R}^3$ et $t \in \mathbb{R}_+$ sont respectivement

les variables spatiales et temporelle. Les équations de Maxwell prennent la forme :

$$\operatorname{div} \mathbf{B} = \operatorname{div} \mathbf{D} = 0 \tag{2.4.1}$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} \tag{2.4.2}$$

$$\frac{\partial \mathbf{D}}{\partial t} = \nabla \times \mathbf{H}. \tag{2.4.3}$$

Le système précédent n'étant pas fermé, on lui adjoint les relations constitutives suivantes :

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \qquad \text{et} \qquad \mathbf{B} = \frac{1}{\epsilon_0 c^2} \mathbf{H} + \mathbf{M} \tag{2.4.4}$$

avec c la vitesse de la lumière dans le vide et la constante ϵ_0 est la permittivité du vide. La polarisation magnétique $\mathbf{M}$ est nulle car le verre n'est pas un matériau magnétique. La polarisation électrique $\mathbf{P}$ est la réponse de la fibre à l'intensité du champ électrique $\mathbf{E}$ (à ne pas confondre avec la polarisation de la lumière). On considère que la fibre a une réponse à la fois linéaire et non linéaire, c'est-à dire-que $\mathbf{P} = \mathbf{P}_L + \mathbf{P}_{NL}$.

La polarisation linéaire $\mathbf{P}_L$ et le champ électrique $\mathbf{E}$ sont liés par la susceptibilité linéaire $\chi^{(1)}$. De la même façon, les propriétés non linéaires d'un matériau sont caractérisées par un certain nombre de susceptibilités optiques non linéaires $\chi^{(n)}$ d'ordre $n \geqslant 2$. Bien évidemment, selon les propriétés du milieu matériel, les équations que l'on obtiendra seront différentes. Par exemple, lorsque l'induction électrique est proportionnelle au champ électrique, l'équation (2.4.8) est l'équation des ondes. Dans un premier temps, on considère un milieu transparent et centrosymétrique. Un milieu matériel dit centrosymétrique (tel que les liquides, les verres ou les gaz) est invariant par symétrie d'inversion. En d'autres termes, un changement de sens du champ électrique $\mathbf{E}$ induit le même changement dans la polarisation (i.e. si $\mathbf{E}$ donne $\mathbf{P}$ alors $-\mathbf{E}$ donne $-\mathbf{P}$). Ce qui signifie que la susceptibilité d'ordre $\chi^{(2)}$ est nulle, si bien que ces matériaux ne peuvent pas produire de processus de second ordre non linéaire.

On définit z comme la distance le long de la fibre, c'est-à-dire la distance de propagation, $r_\perp = (x, y)$ correspondant donc aux directions transverses. On suppose que le tenseur linéaire $\chi^{(1)}$ varie lentement dans la variable z sur une échelle donnée par δ_2 i.e. $\chi^{(1)}(\mathbf{r}, t) = \chi^{(1)}(\delta_2 z, \mathbf{r}_\perp, t)$. La polarisation linéaire est ainsi donnée par

$$\mathbf{P}_L(\mathbf{r}, t) = \epsilon_0 \int_{\mathbb{R}} \chi^{(1)}(\delta_2 z, \mathbf{r}_\perp, t - t_1) \cdot \mathbf{E}(\mathbf{r}, t_1) dt_1. \tag{2.4.5}$$

On suppose que le tenseur diélectrique linéaire s'écrit

$$\chi^{(1)}(\delta_2 z, \mathbf{r}_\perp, t) = \chi_L(\delta_2 z, \mathbf{r}_\perp, t)\, Id + \delta_2 \Delta\chi_L(\delta_2 z, \mathbf{r}_\perp, t), \tag{2.4.6}$$

où $\Delta\chi_L$ contient les effets de la biréfringence due à l'anisotropie et l'asymétrie présentes dans la fibre qui n'est pas parfaitement circulaire. En général $\chi^{(1)}$ est à valeurs complexes, la partie réelle correspondant à l'indice de réfraction linéaire et la partie imaginaire au coefficient d'absorption. On suppose que l'absorption dans la fibre est négligeable et donc que $\chi^{(1)}$ est à valeurs réelles. Les phénomènes non linéaires du

second ordre étant nuls, les plus importants sont ceux du troisième ordre. Dans le cadre d'un régime faiblement non linéaire, la réponse non linéaire s'écrit donc pour tout $i \in \{1,2,3\}$

$$\mathbf{P}_{NL,i}(\mathbf{r},s) = \\ \delta_3\epsilon_0 \sum_{j,k,l=1}^{3} \int_{\mathbb{R}^3} \chi^{(3)}_{ijkl}(\delta_2 z, \mathbf{r}_\perp, s-t_1, s-t_2, s-t_3) E_j(\mathbf{r},t_1) E_k(\mathbf{r},t_2) E_l(\mathbf{r},t_3) dt \quad (2.4.7)$$

où $t = (t_1,t_2,t_3)$. La polarisation au temps t ne dépend de $\mathbf{E}$ que pour des temps antérieurs à t, ce qui revient à supposer que $\chi^{(1)}(t-t') = 0$ pour tout $t' > t$. De même $\chi^{(3)}$ est nul lorsqu'une de ces composantes temporelles est négative.

On revient maintenant aux équations de Maxwell. En supposant tous les champs réguliers et en utilisant l'équation (2.4.2) avec l'équation (2.4.3), on obtient

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r},t) + \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \mathbf{D}}{\partial t^2}(\mathbf{r},t) = 0. \quad (2.4.8)$$

Remarquons d'abord l'égalité

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r},t) = -\partial_z^2 \mathbf{E}_\perp(\mathbf{r},t) + \widehat{\mathbf{e}}_z \partial_z \nabla_\perp \cdot \mathbf{E}_\perp(\mathbf{r},t) + \partial_z \nabla_\perp \mathbf{E}_z + \nabla_\perp \times \nabla_\perp \times \mathbf{E}(\mathbf{r},t),$$

où $\nabla_\perp = (\partial_x, \partial_y, 0)^t$ et $\mathbf{E}_\perp = (E_x, E_y, 0)^t$. On se place dans le cadre d'une approximation paraxiale des équations de Maxwell en supposant que la solution est une perturbation d'une onde plane de fréquence ω_0 et de nombres d'ondes $k(\omega_0)$. On suppose également que l'enveloppe du champ électrique varie lentement en temps et dans la direction z par rapport à ses variations dans les variables transverses. On cherche donc des solutions aux équations de Maxwell sous la forme

$$\mathbf{E}(\mathbf{r},t) = \mathbf{F}(\delta_1 z, \mathbf{r}_\perp, \delta_1 t, \omega_0) \exp(ik(\omega_0)z - i\omega_0 t) + \text{complexe conjugué}. \quad (2.4.9)$$

On considère l'évolution de l'enveloppe $\mathbf{F}$ du champ électrique sur la plus petite longueur caractéristique δ_1. A cette échelle, le champ électrique s'écrit

$$\mathbf{E}(\mathbf{r},t) = \mathbf{F}(\mathbf{r}_\perp, \omega_0) \exp(ik(\omega_0)z - i\omega_0 t) + \text{complexe conjugué}. \quad (2.4.10)$$

En accord avec les hypothèses selon lesquelles la dispersion chromatique, la biréfringence et les effets non linéaires varient sur des échelles beaucoup plus grande (on les néglige donc pour le moment), en insérant l'expression (2.4.10) et en ne gardant que les termes d'ordre zéro, on obtient [68, 77, 78]

$$\begin{aligned} &k^2(\omega_0)\mathbf{F}_\perp(\mathbf{r}_\perp,\omega_0) + ik(\omega_0)\vec{\mathbf{e}}_z \nabla_\perp \cdot \mathbf{F}_\perp(\mathbf{r}_\perp,\omega_0) \\ &+ik(\omega_0)\nabla_\perp F_z(\mathbf{r}_\perp,\omega_0) + \nabla_\perp \times \nabla_\perp \times \mathbf{F}(\mathbf{r}_\perp,\omega_0) \\ &-\frac{\omega_0^2}{c^2}[1+\widehat{\chi}_L(\mathbf{r}_\perp,\omega_0)]\mathbf{F}(\mathbf{r}_\perp,\omega_0) = 0. \end{aligned} \quad (2.4.11)$$

Remarque 2.4.1. *Lorsque la fibre est parfaitement homogène, la susceptibilité linéaire ne dépend pas des variables spatiales. On obtient, à partir de* (2.4.1) *et* (2.4.4)*, que* $\nabla(div\ \boldsymbol{E}) = 0$ *et que* $\nabla \times \nabla \times \boldsymbol{E} = -\Delta \boldsymbol{E}$*. L'équation* (2.4.11) *se réduit alors à l'équation de Helmholtz.*

Le problème aux valeurs propres (2.4.11) détermine le nombre de modes se propageant dans la fibre et doit être résolu sur un domaine cylindrique, correspondant à la fibre optique, avec les conditions aux limites $\mathbf{F}(\rho, \omega_0) \to 0$ lorsque ρ tends vers $+\infty$ [77]. Lorsque le rayon du cœur est suffisamment petit, il n'existe que deux modes de propagation $\mathbf{R}_x$ et $\mathbf{R}_y$ qui constituent les deux modes propres [68, 77]. Ces deux modes vérifient la relation $\widehat{\mathbf{e}}_z \times \mathbf{R}_x = \mathbf{R}_y$ et sont normalisés tel que

$$\int_0^{2\pi} \int_{\mathbb{R}_+} \rho \left|\mathbf{R}_{x,\perp}(\mathbf{r}_\perp, \omega_0)\right|^2 d\rho d\theta = \int_0^{2\pi} \int_{\mathbb{R}_+} \rho \left|\mathbf{R}_{y,\perp}(\mathbf{r}_\perp, \omega_0)\right|^2 d\rho d\theta = 1$$

avec $\mathbf{R}_{x,\perp}$ et $\mathbf{R}_{x,\perp}$ les composantes transverses de ces deux modes. Dans ce cas $\mathbf{F}$ s'écrit comme la superposition de ces deux modes [68, 77]. En incluant les variations sur les échelles plus longues, $\mathbf{F}$ s'écrit encore comme la superposition des deux modes $\mathbf{R}_x$ et $\mathbf{R}_y$ [68, 77]

$$\begin{aligned}&\mathbf{F}(\delta_1 z, \mathbf{r}_\perp, \delta_1 t, \omega_0) = \\ &\left(\frac{\omega_0}{2\epsilon_0 c^2 k(\omega_0)}\right)^{1/2} \left[U_{\delta_1,x}(z,t)\mathbf{R}_x(\mathbf{r}_\perp, \omega_0) + U_{\delta_1,y}(z,t)\mathbf{R}_y(\mathbf{r}_\perp, \omega_0)\right] + \text{c.c}\end{aligned} \tag{2.4.12}$$

où $U_{\delta_1,j}(z,t) = U_j(\delta_1 z, \delta_1 t)$ pour $j = x, y$. Pour obtenir rigoureusement l'évolution de $\mathbf{F}$ sur les échelles plus longues, on écrit l'évolution de l'enveloppe (2.4.12), à partir de l'équation (2.4.8), puis on moyenne sur les variables transverses [77]. La moyennisation sur les variables transverses donne les coefficients de la dispersion chromatique et de l'effet Kerr optique. Une approche simplifiée consiste à supposer que le champ électrique $\mathbf{E}$ est une onde plane dans ses variables transverses [76, 78]. Cette approche ne donne pas les bons coefficients mais aboutit à la bonne équation. Dans ce cas, la relation précédente se simplifie et devient [76, 78]

$$\begin{aligned}\mathbf{E}(z,t) &= \left(\frac{\omega_0}{2\epsilon_0 c^2 k(\omega_0)}\right)^{1/2} \left[U_x(\delta_1 z, \delta_1 t)\widehat{\mathbf{e}}_x + U_y(\delta_1 z, \delta_1 t)\widehat{\mathbf{e}}_y\right] \exp\left(ik(\omega_0)z - i\omega_0 t\right) + \text{c.c} \\ &= C(\omega_0)\mathbf{U}_{\delta_1}(z,t) \exp\left(ik(\omega_0)z - i\omega_0 t\right) + \text{c.c}\end{aligned} \tag{2.4.13}$$

où $\widehat{\mathbf{e}}_x$ et $\widehat{\mathbf{e}}_y$ sont deux vecteurs unitaires orthogonaux perpendiculaires à la direction de propagation satisfaisant la relation $\widehat{\mathbf{e}}_z \times \widehat{\mathbf{e}}_x = \widehat{\mathbf{e}}_y$. Ils définissent les axes de polarisation des deux modes propres et sont adimensionnés. Les équations de Maxwell (2.4.1) se simplifient

$$\partial_z^2 \mathbf{E}(z,t) - \frac{1}{\epsilon_0 c^2} \partial_t^2 \left[\epsilon_0 \mathbf{E}(z,t) + \mathbf{P}(z,t)\right] = 0. \tag{2.4.14}$$

En utilisant la transformée de Fourier, par changement de variables et utilisant la relation (2.4.6), l'expression pour la polarisation linéaire (2.4.5) devient

$$\begin{aligned}\mathbf{P}_L(z,t) &= \frac{\epsilon_0}{2\pi} \int_{\mathbb{R}} \widehat{\chi}^{(1)}(\delta_2 z, \omega) \cdot \widehat{\mathbf{E}}(z,\omega) \exp\left(-i\omega t\right) d\omega + c.c \\ &= \frac{\epsilon_0}{2\pi} \int_{\mathbb{R}} \widehat{\chi}_L(\omega + \omega_0) \widehat{\mathbf{U}}_{\delta_1}(z,\omega) \exp\left(ik(\omega_0)z - i(\omega + \omega_0)t\right) d\omega \\ &\quad + \frac{\epsilon_0}{2\pi} \int_{\mathbb{R}} \delta_2 \Delta\widehat{\chi}_L(\omega + \omega_0) \cdot \widehat{\mathbf{U}}_{\delta_1}(z,\omega) \exp\left(ik(\omega_0)z - i(\omega + \omega_0)t\right) d\omega + \text{c.c}\end{aligned}$$

On suppose que $\mathbf{E}$ (et donc $\mathbf{U}$) est à support compact en fréquence avec un support proche de $\pm\omega_0$. Ainsi, par un développement de Taylor sur la susceptibilité linéaire (qu'on a supposé régulière) autour de la pulsation ω_0, on obtient que

$$\widehat{\chi}_L(\omega+\omega_0) = \widehat{\chi}_L(\omega_0) + \widehat{\chi}'_L(\omega_0)\omega + \frac{1}{2}\widehat{\chi}''_L(\omega_0)\omega^2 + o\left(\omega^2\right)$$

et

$$\Delta\widehat{\chi}_L(\omega+\omega_0) = \Delta\widehat{\chi}_L(\omega_0) + \Delta\widehat{\chi}'_L(\omega_0)\omega + o\left(\omega\right).$$

Ainsi,

$$\mathbf{P}_L(z,t) = \epsilon_0 \mathbf{Q}_L \mathbf{U}_{\delta_1}(t,z)\exp\left(ik(\omega_0)z - i\omega_0 t\right) + \text{c.c}, \tag{2.4.15}$$

où l'opérateur matriciel $\mathbf{Q}_L$ est donné par

$$\mathbf{Q}_L = \widehat{\chi}_L(\omega_0) + i\widehat{\chi}'_L(\omega_0)\partial_t - \frac{1}{2}\widehat{\chi}''_L(\omega_0)\partial_t^2 + \delta_2\Delta\widehat{\chi}_L(\omega_0) + i\delta_2\Delta\widehat{\chi}'_L(\omega_0)\partial_t.$$

Il n'est pas nécessaire de tenir compte des ordres supérieurs qui sont négligeables lorsque l'on considère de longues distances de propagation. On a donc négligé les termes d'ordre $\delta_1^3 \gg \delta_2\delta_1^2$ dans le développement de Taylor. L'ordre 1 gouverne la vitesse de l'impulsion lumineuse et l'ordre 2 la dispersion chromatique. On s'intéresse maintenant à la contribution non linéaire de la polarisation. L'hypothèse que le champ électrique $\mathbf{E}$ est localisé en fréquence autour de $\pm\omega_0$ et que le guide d'onde est faiblement anisotrope entraîne que la contribution de $\mathbf{P}_{NL}$ autour des fréquences $\pm 3\omega_0$ peut être négligée. Ainsi les termes non linéaires que l'on prend en compte sont tous du type $\mathbf{U}^2\mathbf{U}^*$. On suppose de plus que le milieu est isotrope et à symétrie de rotation, ce qui signifie que parmi les éléments du tenseur $\chi^{(3)}$, seulement trois éléments sont indépendants les uns des autres, de sorte que pour $i,j,k,l \in \{x,y\}$

$$\chi^{(3)}_{ijkl} = \chi^{(3)}_{xxyy}\delta_{ij}\delta_{kl} + \chi^{(3)}_{xyxy}\delta_{ik}\delta_{jl} + \chi^{(3)}_{xyyx}\delta_{il}\delta_{jk},$$

où δ_{ij} est le symbole de Kronecker. Grâce aux hypothèses précédentes et en utilisant la transformée de Fourier, ainsi que la symétrie du tenseur $\chi^{(3)}_{ijkl}(\tau_1,\tau_2,\tau_3)$ dans ses variables τ_1 et τ_2, l'expression de la polarisation non linéaire peut être simplifiée de la façon suivante [2, 76, 78]

$$\mathbf{P}_{NL}(z,t) = \delta_3\epsilon_0\mathbf{Q}_{NL}(\delta_1 z, \delta_1 t)\exp(ik(\omega_0)z - i\omega_0 t) + \text{c.c} \tag{2.4.16}$$

avec

$$\begin{aligned}
\mathbf{Q}_{NL}(\delta_1 z, \delta_1 t) = {} & \frac{1}{(2\pi)^3}\times \\
& \Bigg[\int_{\mathbb{R}^3} 2\widehat{\chi}^{(3)}_{xxxx}(\omega_1-\omega_0,\omega_2+\omega_0,\omega_3+\omega_0)\left[\widehat{\mathbf{U}}^*_{\delta 1}(\omega_1)\cdot\widehat{\mathbf{U}}_{\delta 1}(\omega_2)\right]\widehat{\mathbf{U}}_{\delta 1}(\omega_3)\exp(i\omega t)d\omega \\
& + \int_{\mathbb{R}^3}\widehat{\chi}^{(3)}_{xxxx}(\omega_1+\omega_0,\omega_2+\omega_0,\omega_3-\omega_0)\left[\widehat{\mathbf{U}}_{\delta 1}(\omega_1)\cdot\widehat{\mathbf{U}}_{\delta 1}(\omega_2)\right]\widehat{\mathbf{U}}^*_{\delta 1}(\omega_3)\exp(i\omega t)d\omega\Bigg]
\end{aligned}$$

où $\omega = (\omega_1, \omega_2, \omega_3)$. Sous l'hypothèse de transparence du matériau, on a la relation suivante $\widehat{\chi}^{(3)}_{xxxx}(\omega_0, -\omega_0, \omega_0) = \widehat{\chi}^{(3)}_{xxxx}(\omega_0, \omega_0, -\omega_0) := \widehat{\chi}^{(3)}_{NL}$. De plus, par un développement de Taylor autour de ω_0 et en ne conservant que les termes d'ordre δ_3, on obtient

$$\begin{aligned}\mathbf{Q}_{NL}(\delta_1 z, \delta_1 t) = \Big\{ & 2\widehat{\chi}^{(3)}_{NL}\Big[\mathbf{U}^*_{\delta 1}(z,t)\cdot\mathbf{U}_{\delta_1}(z,t)\Big]\mathbf{U}_{\delta 1}(z,t) \\ & + \widehat{\chi}^{(3)}_{NL}\Big[\mathbf{U}_{\delta 1}(z,t)\cdot\mathbf{U}_{\delta 1}(z,t)\Big]\mathbf{U}^*_{\delta 1}(z,t)\Big\}.\end{aligned}$$

Les termes suivants contiennent des dérivées et sont d'ordre supérieur et ne sont pas gardés dans l'approximation. On définit $\epsilon(\omega_0) = 1 + \widehat{\chi}_L(\omega_0) = n^2(\omega_0)$ où n est l'indice de réfraction linéaire. Dans la configuration que l'on a considérée et qui est physiquement raisonnable, on a $\delta_1^2 \gg \delta_1\delta_2 \gg \delta_3 \gg \delta_1^3$. En substituant les expressions (2.4.13), (2.4.15) et (2.4.16) dans (2.4.14) et en ne gardant que les termes au plus d'ordre δ_3 on obtient l'équation suivante

$$\begin{aligned}&-k^2(\omega_0)\mathbf{U}_{\delta_1}(z,t) + 2ik(\omega_0)\partial_z\mathbf{U}_{\delta_1}(z,t) + \partial_z^2\mathbf{U}_{\delta_1}(z,t) + \frac{\omega_0^2}{c^2}\epsilon(\omega_0)\mathbf{U}_{\delta_1}(z,t) \\ &+i\left[\frac{\omega_0^2}{c^2}\epsilon'(\omega_0) + 2\frac{\omega_0}{c^2}\epsilon(\omega_0)\right]\partial_t\mathbf{U}_{\delta_1}(z,t) - \left[\frac{\omega_0^2}{2c^2}\epsilon''(\omega_0) + 2\frac{\omega_0}{c^2}\epsilon'(\omega_0) + \frac{1}{c^2}\epsilon(\omega_0)\right]\partial_t^2\mathbf{U}_{\delta_1}(z,t) \\ &+\frac{\omega_0^2}{c^2}\delta_2\Delta\widehat{\chi}_L(\omega_0)\cdot\mathbf{U}_{\delta_1}(z,t) + i\delta_2\left[\frac{\omega_0^2}{c^2}\Delta\widehat{\chi}'_L(\omega_0) + \frac{2\omega_0}{c^2}\Delta\widehat{\chi}_L(\omega_0)\right]\cdot\partial_t\mathbf{U}_{\delta_1}(z,t) \\ &+\frac{3\omega_0^3}{2c^4k(\omega_0)}\delta_3\widehat{\chi}^{(3)}_{NL}\left\{\left(|U_{\delta_1,1}|^2 + \frac{2}{3}|U_{\delta_1,2}|^2\right)\mathbf{U}_{\delta_1}(z,t) + \frac{1}{3}\begin{pmatrix}U_{\delta_1,2}^2\overline{U}_{\delta_1,1}\\ \\ U_{\delta_1,1}^2\overline{U}_{\delta_1,2}\end{pmatrix}\right\} = 0. \quad (2.4.17)\end{aligned}$$

Il n'y a pas de dérivée en temps dans les termes non linéaires car ceux-ci font apparaître des termes d'ordre supérieurs $\delta_1\delta_3$. La seule contribution de la dérivée seconde est donc l'ajout du coefficient $-\omega_0^2$ devant $\mathbf{P}_{NL}$. En identifiant les puissances de δ_1 dans l'équation, on obtient à l'ordre 0 la relation de dispersion

$$k^2(\omega_0) - \frac{\omega_0^2}{c^2}\epsilon(\omega_0) = 0. \quad (2.4.18)$$

En dérivant la relation de dispersion par rapport à ω_0, on obtient les relations suivantes

$$\begin{cases} 2k(\omega_0)k'(\omega_0) = \left[\dfrac{\omega_0^2}{c^2}\epsilon(\omega_0)\right]' = \dfrac{\omega_0^2}{c^2}\epsilon'(\omega_0) + 2\dfrac{\omega_0}{c^2}\epsilon(\omega_0) \\ \\ k(\omega_0)k''(\omega_0) = \dfrac{\epsilon(\omega_0)}{c^2} + \dfrac{2\omega_0}{c^2}\epsilon'(\omega_0) + \dfrac{\omega_0^2}{2c^2}\epsilon''(\omega_0) - (k'(\omega_0))^2. \end{cases} \quad (2.4.19)$$

On identifie maintenant les termes d'ordre δ_1 et δ_2 dans (2.4.17) et on obtient l'équation de transport suivante

$$2ik(\omega_0)\partial_z\mathbf{U}_{\delta_1}(z,t) + i\left[\frac{\omega_0^2}{c^2}\epsilon(\omega_0)\right]'\partial_t\mathbf{U}_{\delta_1}(z,t) + \frac{\omega_0^2}{c^2}\delta_2\Delta\widehat{\chi}_L(\omega_0)\cdot\mathbf{U}_{\delta_1}(z,t) = 0 \quad (2.4.20)$$

que l'on peut réécrire, en incorporant la relation de dispersion,

$$i\partial_z \mathbf{U}_{\delta_1}(z,t) + ik'(\omega_0)\partial_t \mathbf{U}_{\delta_1}(z,t) + \delta_2 \Delta B(\omega_0)\cdot \mathbf{U}_{\delta_1}(z,t) = 0$$

où $\Delta B(\omega_0) = \left[\omega_0^2/2c^2k(\omega_0)\right]\Delta\widehat{\chi}_L(\omega_0)$. En prenant la dérivée par rapport à z dans l'équation précédente puis en réinjectant l'équation (2.4.20), on obtient l'expression suivante pour la dérivée seconde

$$\begin{aligned}\partial_z^2 \mathbf{U}_{\delta_1}(z,t) = {} & \left(k'(\omega_0)\right)^2 \partial_t^2 \mathbf{U}_{\delta_1}(z,t) - 2i\delta_2 k'(\omega_0)\Delta B(\omega_0)\cdot \partial_t \mathbf{U}_{\delta_1}(z,t) \\ & - \left(\delta_2\right)^2 \left[\Delta B(\omega_0)\right]^2 \cdot \mathbf{U}_{\delta_1}(z,t). \end{aligned} \tag{2.4.21}$$

On injecte les équations (2.4.21) et (2.4.19) dans (2.4.17) et en supprimant les paramètres d'ordre, le système se réécrit

$$\begin{aligned} i\partial_z \mathbf{U}(z,t) + & \left[\Delta B(\omega_0) - \frac{\left(\Delta B(\omega_0)\right)^2}{2k(\omega_0)}\right]\cdot \mathbf{U}(z,t) \\ & - \frac{k''}{2}\partial_t^2 \mathbf{U}(z,t) + i\left[k'(\omega_0) + \left(\Delta B(\omega_0)\right)'\right]\cdot \partial_t \mathbf{U}(z,t) \\ & + \gamma\left\{\frac{5}{6}\left|\mathbf{U}\right|^2 \mathbf{U}(z,t) + \frac{1}{6}\left(\mathbf{U}^*\sigma_3\mathbf{U}\right)\sigma_3\mathbf{U}(z,t) + \frac{1}{3}\begin{pmatrix} U_y^2\overline{U}_x \\ U_x^2\overline{U}_y \end{pmatrix}\right\} = 0 \end{aligned}$$

où $\gamma = 3\omega_0^3\widehat{\chi}_{NL}^{(3)}/(4c^4k^2(\omega_0))$ est le coefficient Kerr et

$$\left(\Delta B(\omega_0)\right)' = \frac{1}{2k(\omega_0)}\left(\frac{\omega_0^2}{c^2}\Delta\widehat{\chi}_L'(\omega_0) + \frac{2\omega_0}{c^2}\Delta\widehat{\chi}_L(\omega_0)\right) - \frac{k'(\omega_0)}{k(\omega_0)}\Delta B(\omega_0).$$

Remarque 2.4.2. *On retrouve ici le coefficient k'' devant $\partial_t^2 \boldsymbol{U}(z,t)$, défini comme le coefficient de dispersion de la vitesse de groupe. Le signe de k'' dans l'équation précédente est important : on parle de régime de dispersion anormale lorsqu'il est négatif et de dispersion normale lorsqu'il est positif. Dans le cas du système de Manakov en régime de dispersion anormale, il existe certaines ondes pour lesquelles l'automodulation de phase et la modulation de phase croisée compensent exactement la dispersion chromatique donnant lieu à une impulsion qui se propage sans déformations à la vitesse de groupe. On les appelle ondes solitaires. On parle de solitons lorsque celles-ci ont en plus la propriété de rester stable après une collision.*

On suppose que la biréfringence est linéaire c'est à dire que

$$\Delta B(z,\omega_0) = b(z,\omega_0)\left[\sin\left(\theta(z)\right)\sigma_1 + \cos\left(\theta(z)\right)\sigma_3\right] = b(z,\omega_0)\Sigma\left(\theta(z)\right)$$

où θ est l'angle de rotation des axes de biréfringence (par rapport aux axes initiaux de polarisation) autour de l'axe de propagation et ne dépend pas de la fréquence. On rappelle que les matrices de Pauli sont données par

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

En négligeant le terme $\frac{(\Delta B(\omega_0))^2}{2k(\omega_0)}$ et en définissant le temps retardé $t \to t - k'(\omega_0)z$, l'équation précédente peut alors se réécrire [77, 78]

$$i\frac{\partial \mathbf{U}}{\partial z} + b(z,\omega_0)\Sigma\left(\theta(z)\right)\mathbf{U} + ib'(z,\omega_0)\Sigma\left(\theta(z)\right)\frac{\partial \mathbf{U}}{\partial t} - \frac{k''(\omega_0)}{2}\frac{\partial^2 \mathbf{U}}{\partial t^2} + \gamma\left\{\frac{5}{6}\left|\mathbf{U}\right|^2\mathbf{U} + \frac{1}{6}\left(\mathbf{U}^*\sigma_3\mathbf{U}\right)\sigma_3\mathbf{U} + \frac{1}{3}\begin{pmatrix} U_x^* U_y^2 \\ \\ U_y^* U_x^2 \end{pmatrix}\right\} = 0, \tag{2.4.22}$$

où t est maintenant le temps retardé. Jusqu'à présent, on a regardé l'évolution de l'enveloppe lentement variable du champ électrique sur des échelles correspondant à la longueur de battement L_B et à la longueur de corrélation l_c et on a établi que cette évolution était décrite par un système couplé d'équations de Schrödinger.

Remarque 2.4.3. *Lorsque la biréfringence est supposée constante (et linéaire) le long de la fibre, on peut choisir sans pertes de généralités $\theta = 0$ et l'équation* (2.4.22) *se réduit au système couplé*

$$i\frac{\partial \boldsymbol{U}}{\partial z} + b\sigma_3\boldsymbol{U} + ib'\sigma_3\frac{\partial \boldsymbol{U}}{\partial t} - \frac{k''}{2}\frac{\partial^2 \boldsymbol{U}}{\partial t^2} + \gamma\left\{\frac{5}{6}\left|\boldsymbol{U}\right|^2\boldsymbol{U} + \frac{1}{6}\left(\boldsymbol{U}^*\sigma_3\boldsymbol{U}\right)\sigma_3\boldsymbol{U} + \frac{1}{3}\begin{pmatrix} U_x^* U_y^2 \\ \\ U_y^* U_x^2 \end{pmatrix}\right\} = 0.$$

Remarque 2.4.4. *Dans le cas où la biréfringence est nulle ($b = 0$), ou lorsque le champ électrique est capable de maintenir son état de polarisation (dans les fibres à maintien de polarisation par exemple) et que la lumière est initialement polarisée selon un des deux modes propres, il n'y a pas d'interactions non linéaires croisées. L'équation de Schrödinger non linéaire scalaire est alors une bonne approximation de l'évolution de l'enveloppe lentement variable du champ électrique.*

2.4.2 Biréfringence aléatoire

Pour le moment, on n'a pas fait d'hypothèses sur la façon dont la biréfringence évoluait. On a vu dans les parties précédentes que la biréfringence avait des causes variées et devait être considérée comme aléatoire. Dans [110], Menyuk et Wai ont proposé deux modèles de biréfringence aléatoire pour les fibres optiques monomodes. Dans cette thèse, on considère le premier modèle dans lequel il est supposé que la force de biréfringence b est constante mais que θ varie aléatoirement. Les variations de θ sont modélisées par un bruit blanc $\dot{\xi}$ en z

$$\frac{d\theta(z)}{dz} = \theta_z = \dot{\xi}(z)$$

de variance σ^2 et de donnée initiale $\theta(0) = 0$. Puisque $\Sigma^2(\theta(z)) = I_2$ et que $\mathbb{E}\left(\cos\left(\theta(z)\right)\right) = \exp\left(-\frac{\sigma^2}{2}z\right)$, la longueur de corrélation de la PMD est donnée par

$$l_c = \int_0^{+\infty} \mathbb{E}\left(\cos(\theta(0))\cos(\theta(z))\right) dz = 2/\sigma^2.$$

Dans la suite de la dérivation, on veut comprendre les effets couplés entre la biréfringence linéaire et les effets non linéaires et en particulier l'impact sur les coefficients de la modulation de phase. Bien que l'effet Kerr ne soit pas difficile à modéliser, il est difficile de comprendre les effets engendrés par l'interaction entre l'effet Kerr et la PMD. Cette difficulté se retrouvera dans l'étude mathématique de ces équations pour lesquelles peu de résultats sont encore valables. On suit la rotation d'angle θ des axes de biréfringence autour de l'axe de propagation z, la description de l'évolution du champ électrique s'en trouvera simplifiée. On transforme donc $\mathbf{U}$ en un nouveau champ de vecteur $\mathbf{A}$ par la rotation suivante

$$\mathbf{U} = \begin{pmatrix} \cos(\theta(z)/2) & -\sin(\theta(z)/2) \\ \sin(\theta(z)/2) & \cos(\theta(z)/2) \end{pmatrix} \mathbf{A}.$$

L'équation décrivant l'évolution de A, dans le repère local des axes de biréfringence, est donnée par [110, 107]

$$\begin{aligned} i\frac{\partial \mathbf{A}}{\partial z} + \widetilde{\Sigma}(\theta(z))\mathbf{A} + ib'\sigma_3\frac{\partial \mathbf{A}}{\partial t} - \frac{k''}{2}\frac{\partial^2 \mathbf{A}}{\partial t^2} \\ + \frac{5}{6}|\mathbf{A}|^2\mathbf{A} + \frac{1}{6}(\mathbf{A}^*\sigma_3\mathbf{A})\sigma_3\mathbf{A} + \frac{1}{3}N(\mathbf{A}) = 0, \end{aligned} \tag{2.4.23}$$

où la matrice $\widetilde{\Sigma}(\theta(z))$ et le terme non linéaire $N(\mathbf{A})$ sont donnés par

$$\widetilde{\Sigma}(\theta(z)) = \begin{pmatrix} b & -\frac{i}{2}\theta_z \\ \frac{i}{2}\theta_z & -b \end{pmatrix} \qquad \text{et} \qquad N(\mathbf{A}) = \begin{pmatrix} \overline{A_1}A_2^2 \\ \overline{A_2}A_1^2 \end{pmatrix}$$

Le changement de jauge $\mathbf{\Phi} = \exp(-ibz\sigma_3)\mathbf{A}$ permet de décrire l'évolution de $\mathbf{\Phi}$ de la façon suivante

$$\begin{aligned} i\frac{\partial \mathbf{\Phi}}{\partial z} + \Sigma(\theta_z, z)\mathbf{\Phi} + ib'\sigma_3\frac{\partial \mathbf{\Phi}}{\partial t} + \frac{d_0}{2}\frac{\partial^2 \mathbf{\Phi}}{\partial t^2} \\ + \frac{5}{6}|\mathbf{\Phi}|^2\mathbf{\Phi} + \frac{1}{6}(\mathbf{\Phi}^*\sigma_3\mathbf{\Phi})\sigma_3\mathbf{A} + \frac{1}{3}N(\mathbf{\Phi}, z) = 0, \end{aligned} \tag{2.4.24}$$

avec

$$\Sigma(\theta_z, z) = \begin{pmatrix} 0 & -\frac{i}{2}\theta_z e^{-2ibz} \\ \frac{i}{2}\theta_z e^{2ibz} & 0 \end{pmatrix} \qquad \text{et} \qquad N(\mathbf{\Phi}, z) = \begin{pmatrix} \overline{\Phi_1}\Phi_2^2 e^{-4ibz} \\ \overline{\Phi_2}\Phi_1^2 e^{4ibz} \end{pmatrix}.$$

Dans la suite le coefficient de dispersion de la vitesse de groupe sera noté $d_0 = -k''$.

Chapter 3

Limit theorems for the Manakov PMD equation

3.1 Introduction

The Manakov PMD equation has been introduced by Wai and Menyuk in [110] to study light propagation over long distance in random birefringent optical fibers. Due to the various length scales present in this problem, a small parameter ϵ appears in the rescaled equation. Our aim in this paper is to prove a diffusion limit theorem for this equation for which we will have to generalize the perturbed test function method [9, 70, 88] to the case of infinite dimension. In [43, 73], a limit theorem is proved for the linear part of the Manakov PMD equation using the Fourier transform and the theory of diffusion approximation for random ODE. Obviously the method in [43, 73] does not work for a nonlinear PDE. In [21, 73], a limit theorem is proved for a non linear scalar PDE driven by a one dimensional noise. The proof relies on the fact that the solution processes are continuous functions of the noise. These methods are no longer applicable to the limit equation that we will consider which is driven by a three dimensional noise, because the solution cannot be written as a continuous function of the noise. Indeed in a general setting a strong solution of a stochastic equation is only a measurable function of the initial data and the Brownian Motion driving the equation. However in the case of a one dimensional noise, Doss [32] and Sussman [102] proved that the solution of such an equation can be written as a continuous function of the Brownian motion. This result has been extended by Yamato [112] to multidimensional Brownian Motions when the Lie algebra generated by the vector fields of the equation is nilpopent of step p. He actually proves the equivalence between the nilpotent hypothesis and the fact that the solution can be written as a continuous function of iterated Stratonovich integrals. In our case the vector fields driving the Manakov PMD Equation are functions of the Pauli matrices and the nilpotent hypothesis of Yamato is not satisfied. This motivates the use of the perturbed test function method. Note that the method has been used for a linear PDE in [29] and a PDE with bounded diffusion coefficients in [89].

We are also interested in the mathematical analysis of both the Manakov PMD and the limit equations. Using a unitary transformation, we are able to establish

Strichartz estimates for the transformed equation, that are not available for the Manakov PMD equation. This result will then enable us to prove global existence of solutions. The limiting equation is also studied. We use a compactness method to study the existence and uniqueness of solutions of this latter equation, due to the lack of nilpotent hypothesis and to the absence of unitary transformation similar to the Manakov PMD case.

3.1.1 Presentation of the model

Optical fibers are thin, transparent and flexible fibers along which the light propagates to transmit information over long distances and so are of huge interest in modern communications. In a perfect fiber, the two transverse components of the electric field are degenerate in the sense that they propagate with the same characteristics : group velocity, chromatic dispersion, refractive indices ($n_1 = n_2$), etc. However during the fabrication process the fiber may present defects like an ellipticity of the core or suffer from mechanical distortions like stress constraints or twisting [1, 2]. These phenomena induce modal birefringence ($n_1 \neq n_2$) characterized by an orientation angle θ and an amplitude b. If $n_1 > n_2$, we then define a slow axe and a rapid axe corresponding respectively to the mode indices n_1 and n_2. The orientation angle θ describes the rotation of the local polarization axes with respect to the initial axes. The birefringence strength (or degree of modal birefringence) is given by $b = |n_1 - n_2|\, k_0 = k_1 - k_2$ where k_1, k_2 are the components of the wave vector and k_0 the wavenumber of the incident light in Vacuum. The beat length $L_B = \frac{2\pi}{k_1 - k_2}$ indicates the length required for the polarization to return to its initial state. There exist several types of birefringence that do not have the same effect on the electric field. Usually linear birefringence is studied (in the absence of Kerr effect, a linearly polarized light remains linearly polarized), although it has been shown that the birefringence could also be elliptic (occurring in case of twisting, see Menyuk [76]). In case of a uniform anisotropy along the fiber, the birefringence parameters (θ, b) are constant. However in realistic configurations, the anisotropy is not uniform along the fiber. We assume, as in [107, 108, 109, 110], that the birefringence is randomly varying, implying Polarization Mode Dispersion (PMD). The difference of velocity of the two modes, due to random change of the birefringence (and so of the refractive indices), induces coupling between the two polarized modes and pulse spreading : PMD is one of the limiting factors of high bit rate transmission.

In [110], Wai and Menyuk assumed that there is no polarization-dependent loss and considered that communication fibers are nearly linearly birefringent. We here use one of the models introduced in [110] for which the local axes of birefringence are bended with an angle θ randomly varying along the propagation axe and that b and b' (the frequency derivative of b) are constant along this axe. Let us recall that the Pauli matrices are defined by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and let us consider the coupled nonlinear Schrödinger equation transformed into the

frame of the local axes of birefringence ([72, 110])

$$i\frac{\partial\Phi}{\partial t}+\Sigma(\theta_t,t)\Phi+ib'\sigma_3\frac{\partial\Phi}{\partial x}+\frac{d_0}{2}\frac{\partial^2\Phi}{\partial x^2}$$
$$+\frac{5}{6}\left|\Phi\right|^2\Phi+\frac{1}{6}\left(\Phi^*\sigma_3\Phi\right)\sigma_3\Phi+\frac{1}{3}N\left(\Phi,t\right)=0, \tag{3.1.1}$$

where d_0 is the group velocity dispersion parameter and

$$\Sigma(\theta_t,t)=\begin{pmatrix}0 & -\frac{i}{2}\frac{d\theta(t)}{dt}e^{-2ibt}\\ \frac{i}{2}\frac{d\theta(t)}{dt}e^{2ibt} & 0\end{pmatrix}\qquad\text{and}\qquad N\left(\Phi,t\right)=\begin{pmatrix}\overline{\Phi_1}\Phi_2^2e^{-4ibt}\\ \overline{\Phi_2}\Phi_1^2e^{4ibt}\end{pmatrix}.$$

We recall that in the context of fiber optics, x corresponds to the retarded time while t corresponds to the distance along the fiber. Following Wai and Menyuk [72, 108, 109, 110] we denote by l the fiber length. We also denote by l_d the dispersion length scale and l_{nl} the nonlinear length scale related to Kerr effect. The fiber autocorrelation length l_c is the length over which two polarization components remain correlated. We consider, as in [110], a typical configuration where $l\sim l_d\sim l_{nl}\gg l_c\gg L_B$ i.e we consider "relatively small" propagation distances. Assuming, as in [43, 73, 109], that the correlation length of $d\theta/dt$ is much shorter than the birefringence beat length and that $|d\theta/dt|\ll b$, we set $d\theta/dt=2\varepsilon\alpha(t)$, where $\varepsilon=\sqrt{L_B/l_c}$ is a small dimensionless parameter and α a Markov process with good ergodic properties. Thus introducing the re-scaled process $\alpha_\varepsilon(t)=\alpha\left(t/\varepsilon^2\right)$, the dynamic of Φ_ε, at a length scale corresponding to the correlation length l_c, is given by

$$i\frac{\partial\Phi_\varepsilon}{\partial t}+\frac{1}{\varepsilon}\Sigma\left(\alpha_\varepsilon(t),\frac{t}{\varepsilon^2}\right)\Phi_\varepsilon+ib'\sigma_3\frac{\partial\Phi_\varepsilon}{\partial x}+\frac{d_0}{2}\frac{\partial^2\Phi_\varepsilon}{\partial x^2}$$
$$+\frac{5}{6}\left|\Phi_\varepsilon\right|^2\Phi_\varepsilon+\frac{1}{6}\left(\Phi_\varepsilon^*\sigma_3\Phi_\varepsilon\right)\sigma_3\Phi_\varepsilon+\frac{1}{3}N\left(\Phi_\varepsilon,\frac{t}{\varepsilon^2}\right)=0. \tag{3.1.2}$$

It will be proved that the asymptotic dynamic of Φ_ε is given by the stochastic semilinear dynamic (see Theorem 3.1.4)

$$id\Psi(t)+\left\{ib'\sigma_3\frac{\partial\Psi}{\partial x}+\frac{d_0}{2}\frac{\partial^2\Psi}{\partial x^2}+\frac{5}{6}\left|\Psi\right|^2\Psi+\frac{1}{6}\left(\Psi^*\sigma_3\Psi\right)\sigma_3\Psi\right\}dt$$
$$+\gamma_s\sigma_3\Psi dt-\sqrt{\gamma_c}\left(\sigma_1\Psi\circ dW_1(t)+\sigma_2\Psi\circ dW_2(t)\right)=0, \tag{3.1.3}$$

where $W=(W_1,W_2)$ is a $2d$ real valued Brownian motion and $\circ$ denotes the Stratonovich product. As in [72, 110], we introduce a unitary matrix

$$Z(t)=\begin{pmatrix}\nu_1(t) & \overline{\nu_2}(t)\\ -\nu_2(t) & \overline{\nu_1}(t)\end{pmatrix}, \tag{3.1.4}$$

where the process ν is solution of the stochastic equation

$$\begin{aligned}d\nu(t) &= i\sqrt{\gamma_c}\left(\sigma_1\nu(t)\circ dW_1(t)+\sigma_2\nu(t)\circ dW_2(t)\right)+i\gamma_s\sigma_3\nu(t)dt \qquad (3.1.5)\\ &= i\sqrt{\gamma_c}\left(\sigma_1\nu(t)dW_1(t)+\sigma_2\nu(t)dW_2(t)\right)+i\gamma_s\sigma_3\nu(t)dt-\gamma_c\nu(t)dt,\end{aligned}$$

where $|\nu_1(0)|^2+|\nu_2(0)|^2=1$. The second equation is the corresponding Ito equation. In addition γ_c, γ_s are two constants determined by α and given by

$$\gamma_c = \int_0^\infty \cos(2bt)\mathbb{E}\left(\alpha(0)\alpha(t)\right)dt \text{ and } \gamma_s = \int_0^\infty \sin(2bt)\mathbb{E}\left(\alpha(0)\alpha(t)\right)dt.$$

We also consider, for $t \in \mathbb{R}_+$, the matrix :

$$\begin{aligned}\boldsymbol{\sigma}\left(\nu(t)\right) &= \begin{pmatrix} |\nu_1|^2 - |\nu_2|^2 & 2\overline{\nu_1\nu_2} \\ 2\nu_1\nu_2 & |\nu_2|^2 - |\nu_1|^2 \end{pmatrix} = \begin{pmatrix} m_3 & m_1 - im_2 \\ m_1 + im_2 & -m_3 \end{pmatrix} \\ &= \sigma_1 m_1(t) + \sigma_2 m_2(t) + \sigma_3 m_3(t), \end{aligned} \tag{3.1.6}$$

which characterizes the linear birefringence and where m_1, m_2, m_3 are real valued processes. Then $\nu(t) \in \mathbb{S}^3$ a.s, the unit sphere in $\mathbb{C}^2 \sim \mathbb{R}^4$. We denote by Λ the unique invariant probability measure of ν (see Section 3.5) and by $\mathbb{E}_\Lambda(.)$ the expectation with respect to Λ. Then we can remove the rapid variation of the state of polarization in the evolution of Φ using the change of variable $\Phi(t) = Z(t)X(t)$. The evolution of the electric field envelope $X = (X_1, X_2)^t$ is now given by (see Lemma 3.2.1)

$$i\frac{\partial X}{\partial t} + \frac{d_0}{2}\frac{\partial^2 X}{\partial x^2} + \frac{8}{9}|X|^2 X = -ib'\boldsymbol{\sigma}\left(\nu(t)\right)\frac{\partial X}{\partial x} - \frac{1}{6}\left(N_\nu(X) - \mathbb{E}_\Lambda\left(N_\nu(X)\right)\right), \tag{3.1.7}$$

where $N_\nu(X) = \left(N_{1,\nu}(X), N_{2,\nu}(X)\right)^t$ satisfy

$$\begin{aligned} N_{1,\nu}(X) &= \left(m_1^2 + m_2^2\right)\left(2|X_2|^2 - |X_1|^2\right)X_1 + \left(m_1 - im_2\right)m_3(2|X_1|^2 - |X_2|^2)X_2 \\ &\quad + \left(m_1 - im_2\right)^2 X_2^2\overline{X_1} + \left(m_1 + im_2\right)m_3 X_1^2\overline{X_2} \\ N_{2,\nu}(X) &= \left(m_1^2 + m_2^2\right)\left(2|X_1|^2 - |X_2|^2\right)X_2 - \left(m_1 + im_2\right)m_3(2|X_2|^2 - |X_1|^2)X_1 \\ &\quad - \left(m_1 - im_2\right)m_3 X_2^2\overline{X_1} + \left(m_1 + im_2\right)^2 X_1^2\overline{X_2}. \end{aligned}$$

The process $m = (m_1, m_2, m_3)$ satisfying $m = (g_1(\nu), g_2(\nu), g_3(\nu))$, it can be proved (see Section 3.5) that

$$\mathbb{E}_\Lambda\left(N_{1,\nu}(X)\right) = \frac{2}{3}\left(2|X_2|^2 - |X_1|^2\right)X_1, \qquad \mathbb{E}_\Lambda\left(N_{2,\nu}(X)\right) = \frac{2}{3}\left(2|X_1|^2 - |X_2|^2\right)X_2.$$

We set

$$F_{\nu(t)}(X(t)) = \frac{8}{9}|X|^2 X - \frac{1}{6}\left(N_\nu(X) - \mathbb{E}_\Lambda\left(N_\nu(X)\right)\right). \tag{3.1.8}$$

Equation (3.1.7) is of great interest for the study of dispersion because the main effects leading to signal distortions (Kerr effect, chromatic dispersion, PMD) can be easily identified : on the left hand side, the first term describes the evolution of the pulse along the fiber. The second one corresponds to the chromatic dispersion and the last term to the Kerr effect averaged on the Poincaré sphere. On the right hand side of the equation, the first term describes the linear PMD effect and the second term describes nonlinear PMD.

The Manakov PMD equation (3.1.7) is written in dimensionless form. According to the length scales we consider, we set $X_\epsilon(t,x) = \frac{1}{\epsilon} X\left(\frac{t}{\epsilon^2}, \frac{x}{\epsilon}\right)$ and $\nu_\epsilon(t) = \nu\left(\frac{t}{\epsilon^2}\right)$ where ν is solution of (3.1.5); then the electric field X_ϵ has the following evolution

$$i\frac{\partial X_\epsilon(t)}{\partial t} + \frac{ib'}{\epsilon}\boldsymbol{\sigma}\left(\nu_\epsilon(t)\right)\frac{\partial X_\epsilon(t)}{\partial x} + \frac{d_0}{2}\frac{\partial^2 X_\epsilon(t)}{\partial x^2} + F_{\nu_\epsilon(t)}(X_\epsilon(t)) = 0, \tag{3.1.9}$$

where the term $F_{\nu_\epsilon(t)}(X_\epsilon(t))$ is given by (3.1.8).

In various physical situations, the long time behavior of a phenomenon subject to random perturbations requires to take care of the different characteristic length scales of the problem. In this context Papanicolaou-Stroock-Varadhan [88] and Blankenship-Papanicolaou [9] introduced the approximation diffusion theory for random ordinary differential Equations. This method has been used to study wave propagation in random media [38] and in particular in randomly birefringent fibers [43, 73] but only few results exist on limit theorems for random PDEs. In the latter, the authors studied the evolution, in an optical fiber, of the linear field envelope X_ϵ given by

$$i\frac{\partial X_\epsilon(t)}{\partial t} + \frac{ib'}{\epsilon}\boldsymbol{\sigma}\left(\nu_\epsilon(t)\right)\frac{\partial X_\epsilon(t)}{\partial x} + \frac{d_0}{2}\frac{\partial^2 X_\epsilon(t)}{\partial x^2} = 0,$$

and proved that the asymptotic dynamics, when ϵ goes to zero, is given by

$$idX(t) + \left(\frac{d_0}{2}\frac{\partial^2 X(t)}{\partial x^2}\right)dt + i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k\frac{\partial X(t)}{\partial x}\circ dW_k(t) = 0,$$

where $W = (W_1, W_2, W_3)$ is a 3d Brownian motion, and $\gamma = (b')^2/6\gamma_c$. Note that the linear PMD effect reduces to one single parameter γ in front of the three Brownian motions. Generalizing the perturbed test function method, we will prove that the asymptotic dynamic of (3.1.9) is given by the stochastic nonlinear evolution :

$$idX(t) + \left(\frac{d_0}{2}\frac{\partial^2 X(t)}{\partial x^2} + F(X(t))\right)dt + i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k\frac{\partial X(t)}{\partial x}\circ dW_k(t) = 0, \tag{3.1.10}$$

where the nonlinear function F reduces to $F(X(t)) = \frac{8}{9}\left|X(t)\right|^2 X(t)$ that is simply the expectation, with respect to the invariant measure Λ, of $F_{\nu_\epsilon(t)}(X_\epsilon(t))$. We will also make use of the following equivalent Ito formulation :

$$idX(t) + \left(\left(\frac{d_0}{2} - \frac{3i\gamma}{2}\right)\frac{\partial^2 X(t)}{\partial x^2} + F(X)(t)\right)dt + i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k\frac{\partial X(t)}{\partial x}dW_k(t) = 0. \tag{3.1.11}$$

Note that a different regime concerned with long propagation distances and corresponding to $l \gg l_{nl} \sim l_d$ is of physical interest; however this regime would lead to another asymptotic analysis which is beyond the scope of this paper.

This paper is organized as follows : in Section 3.1.2 we give notations that will be used along the paper and state the main results. Section 3.2 is devoted to the proof of well-posedness for the Manakov PMD equation. In Section 3.3 we study the local

well-posedness of the limiting Equation (3.1.10). Finally in Section 3.4 we prove the convergence in law of X_ϵ to X as ϵ goes to zero. In Section 3.5 we recall some results obtained in [43, 73] about the driving process ν and in Section 3.6 proofs of technical results used in Section 3.4 are gathered. This paper ends with Section 3.7 where we apply the method of Section 3.4 to prove the convergence of Φ_ε to Ψ.

3.1.2 Notations and main results

Before stating the main results of this article, let us give some definitions and notations.

For all $p \geqslant 1$, we define $\mathbb{L}^p(\mathbb{R}) = \left(L^p(\mathbb{R};\mathbb{C})\right)^2$ the Lebesgue spaces of functions with values in $\mathbb{C}^2$. Identifying $\mathbb{C}$ with $\mathbb{R}^2$, we define a scalar product on $\mathbb{L}^2(\mathbb{R})$ by

$$(u,v)_{\mathbb{L}^2} = \sum_{i=1}^{2} \mathcal{R}\mathrm{e}\left\{\int_{\mathbb{R}} u_i \overline{v_i} dx\right\}.$$

We denote by $\mathbb{W}^{m,p}, m \in \mathbb{N}^*, p \in \mathbb{N}^*$ the space of functions in $\mathbb{L}^p$ such that their m first derivatives are in $\mathbb{L}^p$. If $p = 2$, then we denote $\mathbb{H}^m(\mathbb{R}) = \mathbb{W}^{m,2}(\mathbb{R})$, $m \in \mathbb{N}$. We will also use $\mathbb{H}^{-m}$ the topological dual space of $\mathbb{H}^m$ and denote $\langle .,. \rangle$ the paring between $\mathbb{H}^m$ and $\mathbb{H}^{-m}$. The Fourier transform of a tempered distribution $v \in \mathcal{S}'(\mathbb{R})$ is either denoted by $\widehat{v}$ or $\mathcal{F}v$. If $s \in \mathbb{R}$ then $\mathbb{H}^s$ is the fractional Sobolev space of tempered distributions $v \in \mathcal{S}'(\mathbb{R})$ such that $(1+|\xi|^2)^{s/2}\widehat{v} \in \mathbb{L}^2$. Let $(E, \|.\|_E)$ and $(F, \|.\|_F)$ be two Banach spaces. We denote by $\mathcal{L}(E,F)$ the space of linear continuous functions from E into F, endowed with its natural norm. If I is an interval of $\mathbb{R}$ and $1 \leqslant p \leqslant +\infty$, then $L^p(I;E)$ is the space of strongly Lebesgue measurable functions f from I into E such that $t \mapsto \|f(t)\|_E$ is in $L^p(I)$. The space $L^p(\Omega, E)$ is defined similarly where $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. We denote by $L^p_w(I,E)$ the space $L^p(I,E)$ endowed with the weak (or weak star) topology. For a real number $0 < \alpha < 1$ and $p \geqslant 1$, we denote by $W^{\alpha,p}([0,T],E)$ the fractional Sobolev space of functions u in $L^p(0,T;E)$ satisfying

$$\int_0^T \int_0^T \frac{\|u(t)-u(s)\|_E^p}{|t-s|^{\alpha p+1}} dsdt < +\infty.$$

The space $C^\beta([0,T];E)$ is the space of Hölder continuous functions of order $\beta > 0$ with values in E and we denote by $\mathcal{M}(E)$ the set of probability measures on E, endowed with the topology of the weak convergence $\sigma(\mathcal{M}(E), C_b(E))$.

We will use the space

$$\mathcal{K} = \left(C\left([0,T], \mathbb{H}^1_{loc}\right) \cap C_w\left([0,T], \mathbb{H}^1\right) \cap L^\infty_w\left(0,T;\mathbb{H}^2\right)\right) \times C([0,T],\mathbb{R}),$$

where $C_w([0,T],\mathbb{H}^m), m \in \mathbb{Z}$ is the space of functions f in $L^\infty(0,T;\mathbb{H}^m)$, weakly continuous from $[0,T]$ into $\mathbb{H}^m$. As the solution of our limit equation will not necessary be global in time, we need to introduce a space of exploding paths, as in [3], by adding a point Δ, which acts as a cemetery point, at infinity in $\mathbb{H}^1$; then

$$\mathcal{E}\left(\mathbb{H}^1\right) = \left\{f \in C\left([0,T], \mathbb{H}^1 \cup \{\Delta\}\right), f(t_0) = \Delta \text{ for } t_0 \in [0,T] \Rightarrow f(t) = \Delta \text{ for } t \in [t_0,T]\right\}.$$

We define a topology on $\mathbb{H}^1 \cup \{\Delta\}$ such that the open sets of $\mathbb{H}^1 \cup \{\Delta\}$ are the open sets of $\mathbb{H}^1$ and the complementary in $\mathbb{H}^1 \cup \{\Delta\}$ of the closed bounded sets in $\mathbb{H}^1$. For any $f \in C\left([0,T], \mathbb{H}^1 \cup \{\Delta\}\right)$ we denote the blowing-up time $\tau(f)$ by

$$\tau(f) = \inf\left\{t \in [0,T], f(t) = \Delta\right\},$$

with the convention $\tau(f) = +\infty$ if $f(t) \neq \Delta$ for all $t \in [0,T]$. We endow the space $\mathcal{E}\left(\mathbb{H}^1\right)$ with the topology induced by the uniform convergence in $\mathbb{H}^1$ on every compact set of $[0, \tau(f))$.

Let $(\mathcal{A}, \mathcal{G}, \mathbb{Q})$ be a probability space endowed with the complete filtration $(\mathcal{G}_t)_{t \geqslant 0}$ generated by a two dimensional Brownian Motion $W = (W_1, W_2)$ which is driving the diffusion process ν given by (3.1.5). We first state an existence and uniqueness result for Equation (3.1.9).

Theorem 3.1.1. *Let $\epsilon > 0$ and suppose that $X_\epsilon(0) = v \in \mathbb{L}^2(\mathbb{R})$, then there exists a unique global solution X_ϵ to Equation* (3.1.9) *such that, $\mathbb{Q}$-almost surely,*

$$X_\epsilon \in C\left(\mathbb{R}_+, \mathbb{L}^2\right) \cap C^1\left(\mathbb{R}_+, \mathbb{H}^{-2}\right) \cap \mathbb{L}^8_{loc}\left(\mathbb{R}_+, \mathbb{L}^4\right).$$

Moreover Equation (3.1.9) *preserves the $\mathbb{L}^2$ norm i.e for all $t \in \mathbb{R}_+$:*

$$\|X_\epsilon(t)\|_{\mathbb{L}^2} = \|v\|_{\mathbb{L}^2}.$$

If in addition $X_\epsilon(0) = v \in \mathbb{H}^1$ (resp. $\mathbb{H}^2$, resp. $\mathbb{H}^3$) then corresponding solution is in $C\left(\mathbb{R}_+, \mathbb{H}^1\right)$ (resp. $C\left(\mathbb{R}_+, \mathbb{H}^2\right)$, resp. $C\left(\mathbb{R}_+, \mathbb{H}^3\right)$).

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which is defined a 3-dimensional real valued Brownian motion $W = (W_1, W_2, W_3)$. We denote by $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ the complete filtration generated by W. The next theorem gives existence and uniqueness of local solution for (3.1.10)

Theorem 3.1.2. *Let $X_0 = v \in \mathbb{H}^1(\mathbb{R})$ then there exists a maximal stopping time $\tau^*(v,\omega)$ and a unique strong solution X (in the probabilistic sense) to* (3.1.10)*, such that $X \in C\left([0,\tau^*), \mathbb{H}^1(\mathbb{R})\right)$ $\mathbb{P}-a.s.$ Furthermore the $\mathbb{L}^2$ norm is almost surely preserved, i.e, $\forall t \in [0,\tau^*), \|X(t)\|_{\mathbb{L}^2} = \|v\|_{\mathbb{L}^2}$ and the following alternative holds for the maximal existence time of the solution :*

$$\tau^*(v,\omega) = +\infty \quad or \quad \limsup_{t \nearrow \tau^*(v,\omega)} \|X(t)\|_{\mathbb{H}^1} = +\infty.$$

Moreover if $v \in \mathbb{H}^2$, then $X \in C\left([0,\tau^*), \mathbb{H}^2(\mathbb{R})\right)$ and τ^* satisfies

$$\tau^*(v,\omega) = +\infty \quad or \quad \lim_{t \nearrow \tau^*(v,\omega)} \|X(t)\|_{\mathbb{H}^1} = +\infty. \tag{3.1.12}$$

Note that we do not obtain global existence for Equation (3.1.10), due to the lack of control of the evolution of the $\mathbb{H}^1$ norm (see Remark 3.3.1).

Using these existence theorems, we are able to prove a diffusion approximation result for the nonlinear system of PDEs (3.1.9).

Theorem 3.1.3. *Let $X_\epsilon(0) = X_0 = v$ be in $\mathbb{H}^3(\mathbb{R})$, then the solution X_ϵ of (3.1.9) given by Theorem 3.1.1 converges in law to the solution X of (3.1.10) in $\mathcal{E}(\mathbb{H}^1)$ i.e for all functions f in $C_b(\mathcal{E}(\mathbb{H}^1))$,*

$$\lim_{\epsilon \to 0} \mathcal{L}(X_\epsilon)(f) = \mathcal{L}(X)(f).$$

Note that we consider here the Manakov PMD equation (3.1.9), but the method may be carried out to other nonlinear Schrödinger equations. Let us first emphasize the key points that allow us to prove Theorem 3.1.3.

The first point is that the noise term is a linear function of the unknown X_ϵ. This particular structure leads to a stochastic partial differential equation for the limiting equation. The second point is the fact that the Pauli matrices are hermitian. This is important to obtain the conservation of the $\mathbb{L}^2$ norm for both equations. Finally we use that the driving process ν is a homogeneous Markov ergodic process defined on a compact state space such that $\mathbb{E}_\Lambda(\boldsymbol{\sigma}(y)) = 0$. The hypothesis on the driving noise may be weakened as in the case of random ordinary differential equation assuming good mixing properties (for example exponential decay of the covariance function). The boundedness of $\boldsymbol{\sigma}(\nu_\epsilon(t))$ seems to be necessary. It is used to prove uniform bounds in Lemma 3.4.5 for tightness. On the other hand, the lack of Strichartz estimates for the limiting equation (3.1.10) is a negative aspect. Thus we use that $F(v)$ is locally lipschitz in $\mathbb{H}^1(\mathbb{R})$ to prove existence and uniqueness of a local solution to Equation (3.1.10). But if $\boldsymbol{\sigma}(\nu_\epsilon(t))$ were a one dimensional process, larger dimension and larger power in the nonlinear term could be considered.

Other types of nonlinear Schrödinger equations may be considered, replacing for example $i\frac{\partial X_\epsilon}{\partial x}$ by X_ϵ and assuming that the matrices σ_k are real valued and symmetric. This latter equation is simpler to handle using Strichartz estimates for the fundamental solution and because $\boldsymbol{\sigma}(\nu_\epsilon(t))X_\epsilon(t)$ can be treated as a perturbation as far as we are concerned with existence of solutions. An excellent exemple of such generalization is given in the next result. Let $\left(\widetilde{\mathcal{A}}, \widetilde{\mathcal{G}}, \widetilde{\mathcal{G}}_t, \widetilde{\mathbb{Q}}\right)$ be a filtered probability space.

Theorem 3.1.4. *Let $\varepsilon > 0$ and suppose that α is $\widetilde{\mathbb{Q}}$ almost surely bounded and that $\Phi_\varepsilon(0) = v \in \mathbb{L}^2(\mathbb{R})$, then there almost surely exists a unique global solution $\Phi_\varepsilon \in C(\mathbb{R}_+, \mathbb{L}^2) \cap \mathbb{L}^8_{loc}(\mathbb{R}_+, \mathbb{L}^4)$ to Equation (3.1.2). Moreover Equation (3.1.2) preserves the $\mathbb{L}^2$ norm and if in addition $\Phi_\varepsilon(0) = v \in \mathbb{H}^m, m = 1, 2, 3$ then the corresponding solution is in $C(\mathbb{R}_+, \mathbb{H}^m), m = 1, 2, 3$.*

Suppose that the process α in (3.1.2) is a Feller process, with separable state space S, with an unique invariant probability measure μ under which it is ergodic. Let us assume that α satisfies the null mass condition $\mathbb{E}_\mu(\alpha) = 0$ and that its infinitesimal generator $\mathscr{L}_\alpha$ satisfies the Fredholm alternative. Suppose that $\Phi_\varepsilon(0) = X_0 = v \in \mathbb{H}^3(\mathbb{R})$, then the solution Φ_ε of (3.1.2) converges in law in $C([0,T]; \mathbb{H}^1)$ to the solution Ψ of (3.1.3), which exists and is unique thanks to Proposition 3.2.4, i.e. for all functions f in $C_b(C([0,T]; \mathbb{H}^1))$,

$$\lim_{\varepsilon \to 0} \mathcal{L}(\Phi_\varepsilon)(f) = \mathcal{L}(\Psi)(f).$$

These results establish a first approach for the Diffusion-Approximation theory in infinite dimensions. It would be interesting (and challenging) to generalize these results in dimension 2 and 3 or for different types of noises (space dependent noises for exemple). There are a lot of contexts where this type of equations arise. A first exemple would be to consider Bose Einstein condenstates and the Gross-Pitaevskii equation.

3.2 The Manakov PMD equation : proof of Theorem 3.1.1

The point here is that no Strichartz estimates are available for (3.1.9) because of the lack of commutativity of the matrix $\boldsymbol{\sigma}$ at different time : $\boldsymbol{\sigma}(\nu(t))\boldsymbol{\sigma}(\nu(s)) \neq \boldsymbol{\sigma}(\nu(s))\boldsymbol{\sigma}(\nu(t))$. Consequently only local existence and uniqueness for initial data in $\mathbb{H}^1$ can be easily proved directly on Equation (3.1.9). The idea of the proof is then to find a unitary transformation such that Strichartz estimates are available for the transformed equation. This change of unknown is given in the next result :

Lemma 3.2.1. *Let us denote for $t \in \mathbb{R}_+$*

$$Z_\epsilon(t) = \begin{pmatrix} \nu_{1,\epsilon}(t) & \overline{\nu_{2,\epsilon}}(t) \\ -\nu_{2,\epsilon}(t) & \overline{\nu_{1,\epsilon}}(t) \end{pmatrix},$$

where $\nu_\epsilon = \nu(t/\epsilon^2)$, ν given by (3.1.5). Assuming that $X_\epsilon \in C([0,T],\mathbb{L}^2)$, we set $\Psi_\epsilon(t) = Z_\epsilon(t)X_\epsilon(t)$; then the evolution of the electric field Ψ_ϵ is given by the stochastic Itô equation

$$\begin{aligned} id\Psi_\epsilon(t) + \left\{ \frac{ib'}{\epsilon}\sigma_3\frac{\partial\Psi_\epsilon}{\partial x} + \frac{d_0}{2}\frac{\partial^2\Psi_\epsilon}{\partial x^2} + \frac{5}{6}|\Psi_\epsilon|^2\Psi_\epsilon + \frac{1}{6}\left(\Psi_\epsilon^*\sigma_3\Psi_\epsilon\right)\sigma_3\Psi_\epsilon \right\} dt \\ + \frac{\gamma_s}{\epsilon^2}\sigma_3\Psi_\epsilon dt + \frac{i\gamma_c}{\epsilon^2}\Psi_\epsilon dt - \frac{\sqrt{\gamma_c}}{\epsilon}\left(\sigma_1\Psi_\epsilon d\widetilde{W}_1(t) + \sigma_2\Psi_\epsilon d\widetilde{W}_2(t)\right) = 0, \end{aligned} \tag{3.2.1}$$

where $\widetilde{W}_j(t) = \epsilon W_j(t/\epsilon^2), j = 1,2$, and with initial conditions

$$\Psi_\epsilon(0) = \begin{pmatrix} \nu_{1,\epsilon}(0)v_1 + \overline{\nu_{2,\epsilon}}(0)v_2 \\ \\ -\nu_{2,\epsilon}(0)v_1 + \overline{\nu_{1,\epsilon}}(0)v_2 \end{pmatrix} = \psi_0.$$

Proof. Using the equation satisfied by ν_ϵ and because $|\nu_{1,\epsilon}(t)|^2 + |\nu_{2,\epsilon}(t)|^2 = 1$ for any $t \geqslant 0$, we obtain :

$$idZ_\epsilon(t)Z_\epsilon^{-1}\Psi_\epsilon(t) = -\frac{\gamma_s}{\epsilon^2}\sigma_3\Psi_\epsilon dt - \frac{i\gamma_c}{\epsilon^2}\Psi_\epsilon dt + \frac{\sqrt{\gamma_c}}{\epsilon}\sigma_1\Psi_\epsilon d\widetilde{W}_1(t) + \frac{\sqrt{\gamma_c}}{\epsilon}\sigma_2\Psi_\epsilon d\widetilde{W}_2(t).$$

The nonlinear part of Equation (3.2.1) is obtained as in the derivation given in [110]. □

We first investigate the behavior of the linear equation :

$$i\frac{\partial \Psi_\epsilon}{\partial t} + \frac{1}{\epsilon} ib' \sigma_3 \frac{\partial \Psi_\epsilon}{\partial x} + \frac{d_0}{2} \frac{\partial^2 \Psi_\epsilon}{\partial x^2} = 0, \tag{3.2.2}$$

with initial condition $\Psi_\epsilon(0) = \psi_0 \in \mathbb{L}^2$.

Proposition 3.2.1. *The unbounded matrix operator* $H_\epsilon = \frac{id_0}{2} I_2 \frac{\partial^2}{\partial x^2} - \frac{b'}{\epsilon} \sigma_3 \frac{\partial}{\partial x}$ *defined on* $\mathscr{D}(H_\epsilon) = \mathbb{H}^2$ *is the infinitesimal generator of a unique strongly continuous unitary group* $U_\epsilon(t)$ *on* $\mathbb{L}^2$. *Moreover* $U_\epsilon(t)$ *may be expressed as a convolution kernel i.e for* $\psi_0 \in \mathcal{S}(\mathbb{R})$

$$U_\epsilon(t)\psi_0 = A_\epsilon(t) \star \psi_0 = \frac{1}{\sqrt{2\pi i d_0 t}} \begin{pmatrix} \exp\left\{ \frac{i}{2} \frac{(x - b't/\epsilon)^2}{d_0 t} \right\} & 0 \\ 0 & \exp\left\{ \frac{i}{2} \frac{(x + b't/\epsilon)^2}{d_0 t} \right\} \end{pmatrix} \star \psi_0.$$

Proof of Proposition 3.2.1. Assuming $\psi_0 \in \mathcal{S}(\mathbb{R})$ and taking the Fourier transform in space of Equation (3.2.2) we obtain readily

$$\frac{\partial \widehat{\Psi}_\epsilon}{\partial t} = -\frac{1}{\epsilon} ib' \sigma_3 \xi \widehat{\Psi}_\epsilon - i\frac{d_0 \xi^2}{2} \widehat{\Psi}_\epsilon.$$

Since σ_3 does not depend on time, we obtain

$$\widehat{\Psi}_\epsilon(t) = R_\epsilon(t)\widehat{\psi_0} = \begin{pmatrix} \exp\left\{ -\frac{id_0}{2} \xi^2 t - i\frac{b'}{\epsilon} \xi t \right\} & 0 \\ 0 & \exp\left\{ -\frac{id_0}{2} \xi^2 t + i\frac{b'}{\epsilon} \xi t \right\} \end{pmatrix} \widehat{\psi_0}.$$

The statement of Proposition 3.2.1 follows then in a classical way, setting $A_\epsilon(t) = \mathcal{F}^{-1}(R_\epsilon(t))$. □

The explicit formulation of the kernel given in Proposition 3.2.1 allows immediately to get the following dispersive estimates : if $p \geqslant 2$, $t \neq 0$, then $U_\epsilon \in \mathcal{L}\left(\mathbb{L}^{p'}, \mathbb{L}^p\right)$ where p' is such that $\frac{1}{p} + \frac{1}{p'} = 1$ and for all $\psi_0 \in \mathbb{L}^{p'}$,

$$\|U_\epsilon(t)\psi_0\|_{\mathbb{L}^p} \leqslant (2\pi |d_0| |t|)^{-1/2+1/p} \|\psi_0\|_{\mathbb{L}^{p'}}. \tag{3.2.3}$$

Using then classical arguments (see [13, 44]) one may prove Strichartz inequalities for $U_\epsilon(t)$.

Proposition 3.2.2. *The following properties hold :*

1. *For every* $\psi_0 \in \mathbb{L}^2(\mathbb{R})$, $U_\epsilon(.)\psi_0 \in L^8(\mathbb{R}; \mathbb{L}^4) \cap C(\mathbb{R}; \mathbb{L}^2)$. *Furthermore, there exists a constant* C *such that*

$$\|U_\epsilon(.)\psi_0\|_{L^8(\mathbb{R};\mathbb{L}^4)} \leqslant C \|\psi_0\|_{\mathbb{L}^2} \quad \textit{for every } \psi_0 \in \mathbb{L}^2.$$

2. *Let* I *be an interval of* $\mathbb{R}$ *and* $t_0 \in I$. *Let* $f \in L^{8/7}\left(I, \mathbb{L}^{4/3}\right)$ *then the function*

$$t \mapsto \int_{t_0}^{t} U_\epsilon(t-s) f(s) ds,$$

belongs to $L^8(I,\mathbb{L}^4)\cap C(I,\mathbb{L}^2)$. *Furthermore, there exists a constant* C *independent of* I *such that for every* $f\in L^{8/7}\left(I,\mathbb{L}^{4/3}\right)$

$$\left\|\int_{t_0}^{\cdot}U_\epsilon(.-s)f(s)ds\right\|_{L^8(I,\mathbb{L}^4)\cap L^\infty(I,\mathbb{L}^2)}\leqslant C\left\|f\right\|_{L^{8/7}\left(I,\mathbb{L}^{4/3}\right)}.$$

We now turn to the study of the nonlinear problem. We will use, as is classical, a cut-off argument on the nonlinear term which is not lipschitz. The cut off we consider here is of the same form as the one considered in [17]. We first prove an existence and uniqueness result for this truncated equation, then deduce from this result the existence of a unique solution for Equation (3.2.1). We denote :

$$f(\Psi_\epsilon)=\frac{5}{6}\left|\Psi_\epsilon\right|^2\Psi_\epsilon+\frac{1}{6}\left(\Psi_\epsilon^*\sigma_3\Psi_\epsilon\right)\sigma_3\Psi_\epsilon.$$

Let $\Theta\in C_c^\infty(\mathbb{R})$ with $\mathrm{supp}\Theta\subset[-2;2]$ such that $\Theta(x)=1$ for $|x|\leqslant 1$ and $0\leqslant\Theta(x)\leqslant 1$ for $x\in\mathbb{R}$. Let $R>0$ and $\Theta_R(x)=\Theta(x/R)$. We then consider the following equation

$$\begin{aligned}\Psi_\epsilon^R(t) &= U_\epsilon(t)\psi_0+\frac{i\gamma_s}{\epsilon^2}\int_0^t U_\epsilon(t-s)\sigma_3\Psi_\epsilon^R(s)ds-\frac{\gamma_c}{\epsilon^2}\int_0^t U_\epsilon(t-s)\Psi_\epsilon^R(s)ds\\ &+i\int_0^t U_\epsilon(t-s)\Theta_R\left(\left\|\Psi_\epsilon^R\right\|_{L^8(0,s;\mathbb{L}^4)}\right)f\left(\Psi_\epsilon^R(s)\right)ds \qquad (3.2.4)\\ &-\frac{i\sqrt{\gamma_c}}{\epsilon}\int_0^t U_\epsilon(t-s)\sigma_1\Psi_\epsilon^R(s)d\widetilde{W}_1(s)-\frac{i\sqrt{\gamma_c}}{\epsilon}\int_0^t U_\epsilon(t-s)\sigma_2\Psi_\epsilon^R(s)d\widetilde{W}_2(s),\end{aligned}$$

which is the mild form of the Ito equation :

$$\begin{aligned}&id\Psi_\epsilon^R(t)+\left\{\frac{ib'}{\epsilon}\sigma_3\frac{\partial\Psi_\epsilon^R(t)}{\partial x}+\frac{d_0}{2}\frac{\partial^2\Psi_\epsilon^R(t)}{\partial x^2}+\frac{\gamma_s}{\epsilon^2}\sigma_3\Psi_\epsilon^R(t)+\frac{i}{\epsilon^2}\gamma_c\Psi_\epsilon^R(t)\right\}dt \qquad (3.2.5)\\ &-\frac{\sqrt{\gamma_c}}{\epsilon}\sigma_1\Psi_\epsilon^R d\widetilde{W}_1(t)-\frac{\sqrt{\gamma_c}}{\epsilon}\sigma_2\Psi_\epsilon^R d\widetilde{W}_2(t)+\Theta_R\left(\left\|\Psi_\epsilon^R\right\|_{L^8(0,t;\mathbb{L}^4)}\right)f\left(\Psi_\epsilon^R(t)\right)dt=0,\end{aligned}$$

with initial condition $\Psi_\epsilon^R(0)=\psi_0$.

Proposition 3.2.3. *Let* $\Psi_\epsilon^R(0)=\psi_0\in\mathbb{L}^2(\mathbb{R})$. *Let* $T>0$ *and* $\mathcal{U}_c^T=C([0,T];\mathbb{L}^2)\cap L^8(0,T;\mathbb{L}^4)$; *then Equation* (3.2.4) *has a unique strong adapted solution* $\Psi_\epsilon^R\in L^8\left(\mathcal{A};\mathcal{U}_c^T\right)$, *for any* $T>0$.

Proof of Proposition 3.2.3. We use a fixed point argument in the Banach space $L^8\left(\mathcal{A};\mathcal{U}_c^T\right)$ for sufficiently small time T depending on R. We first need to establish estimates on the stochastic integrals

$$J_{j,\epsilon}\Psi_\epsilon(t)=\int_0^t U_\epsilon(t-s)\sigma_j\Psi_\epsilon(s)d\widetilde{W}_j(s),\quad j=1,2.$$

Lemma 3.2.2. *Let* $T>0$ *; then for each adapted process* $\Psi_\epsilon\in L^8\left(\mathcal{A};\mathcal{U}_c^T\right)$ *and for* $j=1,2$ *the stochastic integral* $J_{j,\epsilon}\Psi_\epsilon$ *belongs to* $L^8\left(\mathcal{A};\mathcal{U}_c^T\right)$. *Moreover for any* $T>0$ *and* t *in* $[0,T]$ *we have the estimates*

$$\mathbb{E}\left(\left\|J_{j,\epsilon}\Psi_\epsilon\right\|^8_{L^8(0,T;\mathbb{L}^4)\cap L^\infty(0,T;\mathbb{L}^2)}\right)\leqslant CT^4\mathbb{E}\left(\left\|\Psi_\epsilon\right\|^8_{L^\infty(0,T;\mathbb{L}^2)}\right).$$

Proof of Lemma 3.2.2. Since $\Psi_\epsilon \in L^8\left(\mathcal{A};\mathcal{U}_c^T\right)$ and is adapted, we may apply the Burkholder Davis Gundy inequality in the Banach space $\mathbb{L}^4(\mathbb{R})$ (which is UMD space [10]) :

$$\begin{aligned}\mathbb{E}\left(\|J_{j,\epsilon}\Psi_\epsilon\|^8_{L^8(0,T;\mathbb{L}^4)}\right) &= \mathbb{E}\left(\int_0^T\left\|\int_0^t U_\epsilon(t-s)\sigma_j\Psi_\epsilon(s)d\widetilde{W}_j(s)\right\|^8_{\mathbb{L}^4}dt\right)\\ &\leqslant \int_0^T\mathbb{E}\left(\sup_{0\leqslant u\leqslant t}\left\|\int_0^u U_\epsilon(t-s)\sigma_j\Psi_\epsilon(s)d\widetilde{W}_j(s)\right\|^8_{\mathbb{L}^4}\right)dt\\ &\leqslant C\mathbb{E}\left(\int_0^T\left(\int_0^t\|U_\epsilon(t-s)\sigma_j\Psi_\epsilon(s)\|^2_{\mathbb{L}^4}\,ds\right)^4dt\right).\end{aligned}$$

Using Hölder inequality in time, Fubini and a change of variable :

$$\mathbb{E}\left(\int_0^T\left(\int_0^t\|U_\epsilon(t-s)\sigma_j\Psi_\epsilon(s)\|^2_{\mathbb{L}^4}\,ds\right)^4dt\right)\leqslant T^3\mathbb{E}\left(\int_0^T\|U_\epsilon(.)\sigma_j\Psi_\epsilon(s)\|^8_{L^8(0,T;\mathbb{L}^4)}\,ds\right).$$

On the other hand, by Proposition 3.2.2,

$$\begin{aligned}\mathbb{E}\left(\int_0^T\|U_\epsilon(.)\sigma_j\Psi_\epsilon(s)\|^8_{L^8(0,T;\mathbb{L}^4)}\,ds\right) &\leqslant C\mathbb{E}\left(\int_0^T\|\Psi_\epsilon(s)\|^8_{\mathbb{L}^2}\,ds\right)\\ &\leqslant CT\mathbb{E}\left(\|\Psi_\epsilon\|^8_{L^\infty(0,T;\mathbb{L}^2)}\right).\end{aligned}$$

Combining these inequalities leads to the estimate in $L^8(0,T;\mathbb{L}^4)$. The other estimate is proved using Burkholder inequality in Hilbert space and the unitary property of the group U_ϵ. Finally $U_\epsilon(t)$ being a unitary semigroup in $\mathbb{L}^2$, Theorem 6.10 in [94] tells us that, provided $\Psi_\epsilon \in L^8\left(\mathcal{A}, L^2(0,T;\mathbb{L}^2)\right)$, then $J_{j,\epsilon}\Psi_\epsilon(.)$ has continuous modification with values in $\mathbb{L}^2(\mathbb{R})$. □

Given $\Psi^R_\epsilon \in L^8\left(\mathcal{A};\mathcal{U}_c^T\right)$, we denote by $\mathcal{T}\Psi^R_\epsilon(t)$ the right hand side of (3.2.4). Since the group $U_\epsilon(.)$ maps $\mathbb{L}^2(\mathbb{R})$ into $C\left(\mathbb{R},\mathbb{L}^2(\mathbb{R})\right)$, Proposition 3.2.2 and Lemma 3.2.2 easily imply that the mapping $\mathcal{T}$ maps $L^8\left(\mathcal{A};\mathcal{U}_c^T\right)$ into itself. Let now Ψ^R_ϵ and Φ^R_ϵ being adapted processes with values in $L^8\left(\mathcal{A};\mathcal{U}_c^T\right)$, then using Proposition 3.2.2, the same arguments as in [17] for the cut-off and Lemma 3.2.2 applied to $J_{j,\epsilon}\left(\Phi^R_\epsilon(t)-\Psi^R_\epsilon(t)\right)$, we get

$$\mathbb{E}\left(\left\|\mathcal{T}\Psi^R_\epsilon-\mathcal{T}\Phi^R_\epsilon\right\|^8_{\mathcal{U}_c^T}\right)^{1/8}\leqslant\left(\frac{CT}{\epsilon^2}+\frac{CT^{1/2}}{\epsilon}+C(R)T^{1/2}\right)\mathbb{E}\left(\left\|\Psi^R_\epsilon-\Phi^R_\epsilon\right\|^8_{\mathcal{U}_c^T}\right)^{1/8}.$$

We conclude that $\mathcal{T}$ is a contraction mapping if T is chosen such that $CT/\epsilon^2 + CT^{1/2}/\epsilon + C(R)T^{1/2} < 1$. As usual, iterating the procedure, we deduce the existence of a unique solution of Equation (3.2.4) in $L^8\left(\mathcal{A};\mathcal{U}_c^T\right)$ for all $T>0$. □

Our aim is now to get global existence for the process Ψ_ϵ, solution of Equation (3.2.1) which may be constructed from the above results. Let us set

$$\kappa^R_\epsilon(\psi_0,\omega)=\inf\left\{t\geqslant 0,\left\|\Psi^R_\epsilon\right\|_{L^8(0,t;\mathbb{L}^4)}\geqslant R\right\},$$

which is a $\mathcal{G}_\epsilon(t)$ stopping time. It can be proved using Strichartz estimates and the integral formulation (3.2.4) (see [17, 18]) that κ_ϵ^R is nondecreasing with R and that $\Psi_\epsilon^R = \Psi_\epsilon^{R'}$ on $[0, \kappa_\epsilon^R]$ for $R < R'$. Thus we are able to define a local solution Ψ_ϵ to Equation (3.2.1) on the random interval $[0, \kappa_\epsilon^*(\psi_0))$, where $\kappa_\epsilon^*(\psi_0) = \lim_{R\to+\infty} \kappa_\epsilon^R$, by setting $\Psi_\epsilon(t) = \Psi_\epsilon^R(t)$ on $[0, \kappa_\epsilon^R]$. It remains to prove that $\kappa_\epsilon^* = +\infty$ almost surely. From the construction of the stopping time κ_ϵ^* it is clear that a.s,

$$\text{if } \kappa_\epsilon^*(\psi_0) < +\infty \text{ then } \lim_{t \nearrow \kappa_\epsilon^*(\psi_0)} \left\|\Psi_\epsilon^R\right\|_{L^8(0,t;\mathbb{L}^4)} = +\infty. \tag{3.2.6}$$

The arguments are adapted from [17]. We first prove the following lemma :

Lemma 3.2.3. *Let $\Psi_\epsilon(0) = \psi_0$ be as in Proposition 3.2.3 and Ψ_ϵ^R be the corresponding solution of (3.2.5); then for any $t < T$*

$$\left\|\Psi_\epsilon^R(t)\right\|_{\mathbb{L}^2} = \|\psi_0\|_{\mathbb{L}^2} \ \ a.s,$$

and there is a constant $M_\epsilon > 0$, depending on T and $\|\psi_0\|_{\mathbb{L}^2}$, but independent of R, such that

$$\mathbb{E}\left(\left\|\Psi_\epsilon^R\right\|_{L^8(0,T;\mathbb{L}^4)}\right) \leqslant M_\epsilon(T). \tag{3.2.7}$$

Proof. To prove that the $\mathbb{L}^2$ norm of the solution Ψ_ϵ^R of (3.2.5) is constant in time, we apply formally the Ito formula to $\frac{1}{2}\left\|\Psi_\epsilon^R(t)\right\|_{\mathbb{L}^2}^2$ and notice that by integration by parts

$$\left(b'\sigma_3\frac{\partial\Psi_\epsilon^R}{\partial x}, \Psi_\epsilon^R\right)_{\mathbb{L}^2} = -\left(\Psi_\epsilon^R, b'\sigma_3\frac{\partial\Psi_\epsilon^R}{\partial x}\right)_{\mathbb{L}^2} = 0.$$

Since $\sigma_j^* = \sigma_j,\ j = 1, 2, 3$, where $*$ stands for the conjuguate transpose, we get

$$\left(\Psi_\epsilon^R(t), i\sigma_j\Psi_\epsilon^R(t)\right)_{\mathbb{L}^2} = 0 \qquad \text{for } j = 1, 2, 3.$$

Moreover because the Itô corrections cancel with the damping term $-\frac{\gamma_c}{\epsilon^2}\Psi_\epsilon^R$ of Equation (3.2.5), we get $\left\|\Psi_\epsilon^R(t)\right\|_{\mathbb{L}^2} = \|\psi_0\|_{\mathbb{L}^2}, \forall t \leqslant T$. The computations can be made rigorous by a regularization procedure.

In order to prove (3.2.7), we follow the procedure in [17, 18]. Using the integral formulation (3.2.4), the conservation of the $\mathbb{L}^2$-norm and Proposition 3.2.2, we obtain for a.e $\omega \in \Omega$ and for all time T_1 such that $T \geqslant T_1 > 0$:

$$\left\|\Psi_\epsilon^R\right\|_{L^8(0,T_1;\mathbb{L}^4)} \leqslant K_\epsilon(\omega) + CT_1^{1/2}\left\|\Psi_\epsilon^R\right\|_{L^8(0,T_1;\mathbb{L}^4)}^3, \tag{3.2.8}$$

where

$$K_\epsilon(\omega) = C\left(1 + \frac{T}{\epsilon^2}\right)\|\psi_0\|_{\mathbb{L}^2} + \frac{1}{\epsilon}\sum_{j=1}^{2}\left\|J_{j,\epsilon}\Psi_\epsilon^R\right\|_{L^8(0,T;\mathbb{L}^4)}.$$

From inequality (3.2.8) it follows that $\left\|\Psi_\epsilon^R\right\|_{L^8(0,T_1;\mathbb{L}^4)} \leqslant 2K_\epsilon(\omega)$ if T_1 is chosen for example such that $T_1(\omega) = \inf\left(T, 2^{-6}\left(C^{1/2}K_\epsilon\right)^{-4}\right)$. If $T_1 < T$ we can reiterate the process on small time intervals $[lT_1, (l+1)T_1] \subset [0, T]$ (keeping R fixed and

varying l) to get $\left\|\Psi_\epsilon^R\right\|_{L^8(lT_1,(l+1)T_1;\mathbb{L}^4)} \leqslant 2K_\epsilon(\omega)$. Summing these estimates, using $T_1 = 2^{-6}C^{-2}\left(K_\epsilon\right)^{-4}$ and Young inequality, we obtain

$$\left\|\Psi_\epsilon^R\right\|_{L^8(0,T;\mathbb{L}^4)} \leqslant C(T)\left(K_\epsilon(\omega)\right)^5.$$

Taking the expectation in the above inequality, using Holder inequality and Lemma 3.2.2 we get the following estimate

$$\mathbb{E}\left(\left\|\Psi_\epsilon^R\right\|_{L^8(0,T;\mathbb{L}^4)}\right) \leqslant C(T)\left(\left(1+\frac{T}{\epsilon^2}\right)^5\left\|\psi_0\right\|_{\mathbb{L}^2}^5 + \frac{CT^{5/2}}{\epsilon^5}\left\|\psi_0\right\|_{\mathbb{L}^2}^5\right), \tag{3.2.9}$$

from which (3.2.7) follows. □

We easily deduce from Lemma 3.2.3 and (3.2.6) that $\kappa_\epsilon^* = +\infty$ a.s. and as in [17] the existence and uniqueness of a solution Ψ_ϵ of (3.2.1), a.s. in $\mathcal{U}_c^T$ for any $T > 0$.

To end the proof of Theorem 3.1.1 we have to extend those results to the process X_ϵ. For a.e ω in $\mathcal{A}$ and for each $t \geqslant 0$ we set $X_\epsilon(t) = Z_\epsilon^{-1}(t)\Psi_\epsilon(t)$. By definition of the process $Z_\epsilon^{-1}(t)$ (which in particular is measurable with respect to $\mathcal{G}_\epsilon(t)$) and properties of Ψ_ϵ, we easily deduce that $X_\epsilon(t)$ is adapted and continuous with values in $\mathbb{L}^2$, and satisfy (3.1.9), hence is C^1 with values in $\mathbb{H}^{-2}$. By unitarity of Z_ϵ we also deduce that for all $t \geqslant 0$

$$\left\|\Psi_\epsilon(t)\right\|_{\mathbb{L}^2}^2 = \left(X_\epsilon(t), Z_\epsilon^{-1}(t)Z_\epsilon(t)X_\epsilon(t)\right)_{\mathbb{L}^2} = \left\|X_\epsilon(t)\right\|_{\mathbb{L}^2}^2,$$

and since the coefficients of $Z_\epsilon^{-1}(t)$ are a.s uniformly bounded, $X_\epsilon \in L^8_{loc}\left(\mathbb{R}_+, \mathbb{L}^4\right)$ a.s; Theorem 3.1.1 is proved. □

We now extend the previous global existence results to more regular initial data. T being fixed, we denote

$$\mathcal{V}^T = L^\infty\left(0,T;\mathbb{H}^1\right) \cap L^8\left(0,T;\mathbb{W}^{1,4}\right) \quad \text{and} \quad \mathcal{V}_c^T = C\left(0,T;\mathbb{H}^1\right) \cap L^8\left(0,T;\mathbb{W}^{1,4}\right).$$

Proposition 3.2.4. *Let $\Psi_\epsilon(0) = \psi_0 \in \mathbb{H}^1$ and let $T > 0$; then Equation (3.2.1) has a unique strong solution Ψ_ϵ with trajectories in $C\left(0,T;\mathbb{H}^1\right)$.*

Proof of Proposition 3.2.4. Let ψ_0 be in $\mathbb{H}^1$. Given $\Psi_\epsilon^R \in L^8\left(\mathcal{A};\mathcal{V}^T\right)$ we denote by $\mathcal{T}\Psi_\epsilon^R(t)$ the right hand side of (3.2.4) and $\mathcal{U}^T = L^\infty\left(0,T;\mathbb{L}^2\right) \cap L^8\left(0,T;\mathbb{L}^4\right)$. By Proposition 3.2.2, Lemma 3.2.2 applied to $\partial_x\Psi_\epsilon^R$ and Hölder inequality we deduce that

$$\left\|\mathcal{T}\partial_x\Psi_\epsilon^R\right\|_{L^8(\mathcal{U}^T)} \leqslant C\left\|\partial_x\psi_0\right\|_{\mathbb{L}^2} + \left(\frac{CT}{\epsilon^2} + \frac{CT^{1/2}}{\epsilon} + CT^{1/2}4R^2\right)\left\|\partial_x\Psi_\epsilon^R\right\|_{L^8(\mathcal{U}^T)}.$$

Therefore we conclude that choosing $R_0 = 2C\left\|\Psi_0\right\|_{\mathbb{H}^1}$, $\mathcal{T}$ maps the closed ball of $L^8\left(\mathcal{A};\mathcal{V}^T\right)$ with radius R_0 into itself, provided T is small enough depending only on R and ϵ, but not on R_0. Combining with the fact that $\mathcal{T}$ is a contraction in $L^8\left(\mathcal{A};\mathcal{U}^T\right)$ and that the balls of $L^8\left(\mathcal{A};\mathcal{V}^T\right)$ are closed for the norm in $L^8\left(\mathcal{A};\mathcal{U}^T\right)$, we conclude to the existence of a unique fixed point $\Psi_\epsilon^R \in L^8\left(\mathcal{A};\mathcal{V}^T\right)$. Using Proposition 3.2.2 and Lemma 3.2.2, we get continuity of the solution in $\mathbb{H}^1$. Since the cut-off only depends on the $L^8\left(0,T,\mathbb{L}^4\left(\mathbb{R}\right)\right)$ norm, we deduce that there is a unique global solution Ψ_ϵ to (3.2.1) with paths in $C\left(0,T;\mathbb{H}^1\right)$. Since the transformation Z_ϵ does not depend on x, we conclude that these results still hold true for X_ϵ. □

Proposition 3.2.5. *Let $\Psi_\epsilon(0) = \psi_0 \in \mathbb{H}^m$, $m = 2, 3$. Let $T > 0$; then Equation (3.2.1) has a unique strong solution Ψ_ϵ with paths in $C(0, T; \mathbb{H}^m)$, $m = 2, 3$.*

Proof. We consider Equation (3.2.5) but with $\Theta_R(\|\Psi_\epsilon^R\|_{L^8(0,t;\mathbb{L}^4)})$ replaced by $\Theta_R(\|\Psi_\epsilon^R(t)\|^2_{\mathbb{H}^1})$. Given Ψ_ϵ^R in $L^8(\mathcal{A}; L^\infty(0, T; \mathbb{H}^2(\mathbb{R})))$ we denote by $\mathcal{T}\Psi_\epsilon^R(t)$ the right hand side of the integral formulation of this equation. We easily prove that $\mathcal{T}$ maps the closed ball of $L^8(\mathcal{A}; L^\infty(0, T; \mathbb{H}^2(\mathbb{R})))$ with radius R_0 into itself, for $R_0 = 2C\|\Psi_0\|_{\mathbb{H}^2}$, provided that T is small enough, depending only on R and ϵ, but not on R_0. Using that this ball is closed for the norm in $L^8(\mathcal{A}; L^\infty(0, T; \mathbb{H}^1(\mathbb{R})))$ and that $\mathcal{T}$ is a contraction for the norm in $L^8(\mathcal{A}; L^\infty(0, T; \mathbb{H}^1(\mathbb{R})))$, we deduce that there exists a unique solution Ψ_ϵ with paths in $C(0, T; \mathbb{H}^2(\mathbb{R}))$ a.s, which is global since the solution is global in $\mathbb{H}^1$. Existence and uniqueness in $\mathbb{H}^3$ can be proved by the same arguments. Again those results are easily extended to X_ϵ and this conclude the proof of Theorem 3.1.1. □

3.3 The limiting equation : proof of Theorem 3.1.2

In order to prove a local existence and uniqueness result for the system (3.1.10) we use a compactness approach (see for example [36]) motivated by the fact that we do not know if Strichartz estimates are available for (3.1.10). Indeed no transformation similar to the Manakov PMD case seems to be available, as the equation $dX(t) = -\sqrt{\gamma}\sum_{k=1}^3 \sigma_k \frac{\partial X(t)}{\partial x} dW_k(t)$ can not be solved in a simple way. We first prove existence of a unique solution in $\mathbb{H}^1$ for the linear part of the equation, defining then a random propagator, and then consider the nonlinear part as a perturbation. We will strongly use the fact that the nonlinearity is locally lipschitz in $\mathbb{H}^1$. The regularity in $\mathbb{H}^2$ will follow with the same arguments as for Equation (3.2.1). Let us consider the linear part of Equation (3.1.10)

$$\begin{aligned} dX(t) &= \left(i\frac{d_0}{2}\frac{\partial^2 X}{\partial x^2}\right)dt - \sqrt{\gamma}\sum_{k=1}^3 \sigma_k \frac{\partial X(t)}{\partial x} \circ dW_k(t) \\ &= \left(i\frac{d_0}{2} + \frac{3\gamma}{2}\right)\frac{\partial^2 X}{\partial x^2}dt - \sqrt{\gamma}\sum_{k=1}^3 \sigma_k \frac{\partial X(t)}{\partial x} dW_k(t), \end{aligned} \tag{3.3.1}$$

with initial data $X(0) = v \in \mathbb{H}^2$. We introduce, for $\eta > 0$, the mollifier $J_\eta = \left(I - \eta\frac{\partial^2}{\partial x^2}\right)^{-1}$. We denote by X_η the solution of the regularized Ito equation

$$dX_\eta(t) = \left(i\frac{d_0}{2} + \frac{3\gamma}{2}\right)\frac{\partial^2 J_\eta^2 X_\eta}{\partial x^2}dt - \sqrt{\gamma}\sum_{k=1}^3 \sigma_k \frac{\partial J_\eta X_\eta(t)}{\partial x} dW_k(t), \tag{3.3.2}$$

and $X_\eta(0) = v \in \mathbb{H}^2$. Since the operators $\partial_x^2 J_\eta^2$ and $\partial_x J_\eta$ are bounded from $\mathbb{H}^1$ into $\mathbb{H}^1$ (with constants depending on η), we easily get, thanks to the Doob inequality, Fubini theorem, the Ito isometry and the independence of $(W_k)_{k=1,2,3}$, the existence and uniqueness of a solution X_η to (3.3.2) with paths in $C([0, T], \mathbb{H}^2)$ for any $T > 0$.

Moreover it is easy to see that the $\mathbb{H}^2$ norm of X_η is conserved since the Pauli matrices are hermitian. Consequently the process

$$M_\eta(t) = -X_\eta(t) + X_\eta(0) + \int_0^t \left(\frac{id_0}{2} + \frac{3\gamma}{2}\right) \frac{\partial^2 J_\eta^2 X_\eta}{\partial x^2} ds,$$

is a $\mathcal{F}_t$ martingale with paths in $C\left([0,T], \mathbb{L}^2\right)$. Let us compute the quadratic variation. Let $a = (a_1, a_2)^t$ and $b = (b_1, b_2)^t$ be in $\mathbb{L}^2$ and $T \geqslant t \geqslant s \geqslant 0$; then

$$\begin{aligned} &\mathbb{E}\left((a, M_\eta(t))_{\mathbb{L}^2} (b, M_\eta(t))_{\mathbb{L}^2} - (a, M_\eta(s))_{\mathbb{L}^2} (b, M_\eta(s))_{\mathbb{L}^2} \middle| \mathcal{F}_s \right) \\ &= \gamma \sum_{k=1}^{3} \mathbb{E}\left(\int_s^t \left(a, \sigma_k \frac{\partial J_\eta X_\eta}{\partial x}\right)_{\mathbb{L}^2} \left(b, \sigma_k \frac{\partial J_\eta X_\eta}{\partial x}\right)_{\mathbb{L}^2} du \middle| \mathcal{F}_s \right). \end{aligned}$$

We deduce that the quadratic variation of $M_\eta(t)$ is given by :

$$(b, \ll M_\eta(t) \gg a)_{\mathbb{L}^2} = \gamma \sum_{k=1}^{3} \int_0^t \left(a, \sigma_k \frac{\partial J_\eta X_\eta}{\partial x}\right)_{\mathbb{L}^2} \left(b, \sigma_k \frac{\partial J_\eta X_\eta}{\partial x}\right)_{\mathbb{L}^2} du. \tag{3.3.3}$$

Using the conservation of the $\mathbb{H}^2$ norm and Equation (3.3.2) we get for all $0 \leqslant \alpha < \frac{1}{2}$

$$\mathbb{E}\left(\|X_\eta\|_{C^\alpha([0,T];\mathbb{L}^2)} \right) \leqslant C_\alpha(T), \tag{3.3.4}$$

where $C_\alpha(T)$ is a constant independent of η. Using Ascoli-Arzela and Banach Alaoglu theorems, Markov inequality and inequality (3.3.4), we get that the sequence $\left(\mathcal{L}\left(X_\eta\right)\right)_{\eta>0}$ is tight on $C_w\left([0,T], \mathbb{H}^1\left(\mathbb{R}\right)\right) \cap L_w^\infty\left(0,T,\mathbb{H}^2\right)$. The Skorokhod theorem ([6],[34]) implies that on some probability space $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathcal{F}_t}, \widetilde{\mathbb{P}}\right)$, there exist a sequence of stochastic processes $\left(\widetilde{X}_\eta\right)_{\eta>0}$, and a process $\widetilde{X}$, such that :

$$\mathcal{L}\left(\widetilde{X}_\eta\right) = \mathcal{L}\left(X_\eta\right), \qquad \mathcal{L}\left(\widetilde{X}\right) = \mathcal{L}\left(X\right),$$

and $\lim_{\eta\to 0} \widetilde{X}_\eta = \widetilde{X}$, $\widetilde{\mathbb{P}} - a.s$ in $C_w\left([0,T], \mathbb{H}^1\right) \cap L_w^\infty\left(0,T,\mathbb{H}^2\right)$. For all $\eta > 0$ and $t \in [0,T]$ we define the process

$$\widetilde{M_\eta}(t) = -\widetilde{X}_\eta(t) + \widetilde{X}_\eta(0) + \int_0^t \left(\frac{id_0}{2} + \frac{3\gamma}{2}\right) \frac{\partial^2 J_\eta^2 \widetilde{X}_\eta}{\partial x^2}(s) ds.$$

We deduce from the above laws equality that $\widetilde{M_\eta}(t)$ is a square integrable continuous martingale with values in $\mathbb{L}^2$ with respect to the filtration $\widetilde{\mathcal{F}}_t$ and that the quadratic variation $\ll \widetilde{M_\eta}(t) \gg$ is given by formula (3.3.3) replacing X_η by $\widetilde{X}_\eta$. Let $a \in \mathbb{H}^1$, then by the above martingale property we get for all $s \leqslant t$:

$$\mathbb{E}\left(\left(a, \widetilde{M_\eta}(t) - \widetilde{M_\eta}(s)\right)_{\mathbb{L}^2} \middle| \widetilde{\mathcal{F}_s} \right) = 0.$$

Using the almost sure convergence in $C_w\left([0,T],\mathbb{H}^1(\mathbb{R})\right)$ of X_η, the boundedness in $\mathbb{H}^{-1}$ of the operator J_η and the conservation of the $\mathbb{H}^1$ norm, we get the almost sure convergence in $C_w\left([0,T],\mathbb{H}^{-1}(\mathbb{R})\right)$ of $\widetilde{M}_\eta$ to $\widetilde{M}$, where

$$\widetilde{M}(t)=\widetilde{X}(t)-\widetilde{X}(0)-\int_0^t\left(\frac{id_0}{2}+\frac{3\gamma}{2}\right)\frac{\partial^2\widetilde{X}}{\partial x^2}(s)ds.$$

Hence $\widetilde{M}$ is a weakly continuous martingale with values in $\mathbb{H}^{-1}$. Moreover using the a.s convergence in $C_w\left([0,T],\mathbb{H}^1(\mathbb{R})\right)$ and dominated convergence theorem, we get for all $t,s\in[0,T],t\geqslant s$ and for any $a,b\in\mathbb{H}^1$,

$$\lim_{\eta\to0}\mathbb{E}\left(\left\langle b,\ll\widetilde{M}_\eta(t)\gg a\right\rangle\middle|\widetilde{\mathcal{F}}_s\right)=\gamma\sum_{k=1}^3\mathbb{E}\left(\int_0^t\left\langle a,\sigma_k\frac{\partial\widetilde{X}}{\partial x}(u)\right\rangle\left\langle b,\sigma_k\frac{\partial\widetilde{X}}{\partial x}(u)\right\rangle du\middle|\widetilde{\mathcal{F}}_s\right).$$

Thus the quadratic variation $\left\langle b,\ll\widetilde{M}(t)\gg a\right\rangle$ is given, for all $t\in[0,T]$, by

$$\left\langle b,\ll\widetilde{M}(t)\gg a\right\rangle=\gamma\sum_{k=1}^3\int_0^t\left\langle a,\sigma_k\frac{\partial\widetilde{X}}{\partial x}(u)\right\rangle\left\langle b,\sigma_k\frac{\partial\widetilde{X}}{\partial x}(u)\right\rangle du. \tag{3.3.5}$$

Noticing that $\widetilde{M}(0)=0$ and using the representation theorem for continuous square integrable martingales we obtain that, on a possibly enlarged space $\left(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathcal{F}}_t,\widetilde{\mathbb{P}}\right)$, one can find a Brownian motion $\widetilde{W}=\left(\widetilde{W}_1,\widetilde{W}_2,\widetilde{W}_3\right)$ such that

$$\left\langle a,\widetilde{M}(t)\right\rangle=\sqrt{\gamma}\int_0^t\sum_{k=1}^3\left\langle a,\sigma_k\frac{\partial\widetilde{X}}{\partial x}(s)\right\rangle d\widetilde{W}_k(s).$$

Thus we deduce that $\left(\widetilde{X},\widetilde{W}\right)$ is a weak solution of Equation (3.3.1) on $\left(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathcal{F}}_t,\widetilde{\mathbb{P}}\right)$ with values in $C_w\left([0,T],\mathbb{H}^1(\mathbb{R})\right)\cap L^\infty(0,T,\mathbb{H}^2)$. To conclude the proof we have to prove pathwise uniqueness of the solution and strong continuity in $\mathbb{H}^1$. Since $\widetilde{X}\in L^\infty(0,T,\mathbb{H}^2)$ is solution of (3.3.1), we easily deduce that $\widetilde{X}\in C^\alpha([0,T],\mathbb{L}^2)$ for any $\alpha\in[0,1/2)$. By interpolation we obtain that $\widetilde{X}\in C([0,T],\mathbb{H}^1)$. It follows, using Ito formula, that pathwise uniqueness holds for Equation (3.3.1) in $C([0,T],\mathbb{H}^1)$. This implies, by the Yamada-Watanabe theorem, that the solution exists in the strong sense. Thus we can define a random unitary propagator $U(t,s)$ which is strongly continuous from $\mathbb{H}^2$ into $\mathbb{H}^1$. This random propagator can be extended to a random propagator from $\mathbb{H}^1$ into $\mathbb{H}^1$ using the continuity of X in $\mathbb{H}^1$, the density of $\mathbb{H}^2$ into $\mathbb{H}^1$ and the isometry property of $U(t,s)$ in $\mathbb{H}^1$.

The local existence of the non linear problem (3.1.10) in $\mathbb{H}^1$ follows from the construction of the random propagator U: we consider a cut-off function $\Theta\in C_c^\infty(\mathbb{R})$, $\Theta\geqslant0$ satisfying

$$\Theta_R\left(\|X(t)\|^2_{\mathbb{H}^1}\right)=\begin{cases}1 & \text{if } \|X(t)\|^2_{\mathbb{H}^1}\leqslant R\\ 0 & \text{if } \|X(t)\|^2_{\mathbb{H}^1}\geqslant 2R.\end{cases}$$

and first construct a solution X^R of the cut-off equation :

$$idX^R(t)+\left(\frac{d_0}{2}\frac{\partial^2 X^R}{\partial x^2}+\Theta_R\left(\left\|X^R(t)\right\|^2_{\mathbb{H}^1}\right)F\big(X^R\big)(t)\right)dt+i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k\frac{\partial X^R(t)}{\partial x}\circ dW_k(t)=0, \tag{3.3.6}$$

with initial data $X^R(0)=v\in\mathbb{H}^1$ and whose integral formulation is given a.e by

$$X^R(t)=U(t,0)v+i\int_0^t\Theta_R\left(\left\|X^R(s)\right\|^2_{\mathbb{H}^1}\right)U(t,s)F\big(X^R(s)\big)ds. \tag{3.3.7}$$

The existence and uniqueness of $X^R\in L^\rho\left(\Omega;C\left(0,T;\mathbb{H}^1\right)\right)$, solution of (3.3.7), is easily obtained by a fixed point argument since the nonlinear term is globally lipschitz. Introducing the nondecreasing stopping time

$$\tau^R=\inf\left\{t\geqslant 0,\left\|X^R(t)\right\|^2_{\mathbb{H}^1}\geqslant R\right\},$$

we may then define a local solution X to Equation (3.1.10) on a random interval $[0,\tau^*(v))$, where $\tau^*(v)=\lim_{R\to+\infty}\tau^R$ almost surely, by setting $X(t)=X^R(t)$ on $[0,\tau^R]$. Then for any stopping time $\tau<\tau^*$ we have constructed a unique local solution with paths a.s in $C\left([0,\tau],\mathbb{H}^1\right)$. It follows from the construction of the stopping time τ^* that if $\tau^*<+\infty$ then $\limsup_{t\to\tau^*}\|X(t)\|_{\mathbb{H}^1}=+\infty$. Let us now prove that if $v\in\mathbb{H}^2$ then the maximal stopping time satisfies the following alternative

$$\tau^*=+\infty\text{ or }\lim_{t\to\tau^*}\|X(t)\|_{\mathbb{H}^1}=+\infty. \tag{3.3.8}$$

We note that the random propagator commutes with derivation. Hence if $v\in\mathbb{H}^2$, then $U(.,0)v\in C\left([0,T],\mathbb{H}^2\right)$. We easily deduce, using Equation (3.3.1) and interpolating $\mathbb{H}^1$ between $\mathbb{H}^2$ and $\mathbb{L}^2$, that $U(.,0)v\in C^\beta\left([0,T],\mathbb{H}^1\right)$ for $\beta\in[0,1/4)$. By a fixed point argument in $\mathbb{H}^2$ and Equation (3.3.7), we conclude that $X\in C^\beta\left([0,\tau],\mathbb{H}^1\right)$ for any stopping time $\tau<\tau^*$ and for the same maximal time existence τ^*. Hence using the condition on τ^* and uniform continuity of X in $\mathbb{H}^1$, we get that (3.3.8) holds.

Remark 3.3.1. *We were not able to prove the global wellposedness for Equation* (3.1.10)*. Due to the lack of Strichartz estimates, we cannot control the evolution of the* $\mathbb{H}^1$ *norm. Even though the deterministic energy provides a control on the* $\mathbb{H}^1$ *norm because we are in the subcritical case, its evolution for a solution of Equation* (3.1.10)*, which is given in the next Lemma, involves terms which are not well controlled. However, we cannot really conclude to the real occurence of blow up or not in this model. It is clear that on a physical point of view such a phenomenon should not occur.*

Lemma 3.3.1. *Let the functional* H *be defined for* $u\in\mathbb{H}^1(\mathbb{R})$ *by*

$$H(u)=\frac{d_0}{4}\int_{\mathbb{R}}\left|\frac{\partial u}{\partial x}\right|^2dx-\frac{2}{9}\int_{\mathbb{R}}|u|^4\,dx.$$

Then for any stopping time τ such that $\tau < \tau^*$, we have

$$\begin{aligned} H\left(X(\tau)\right) &= H\left(X_0\right) + \sqrt{\gamma}\frac{8}{9}\sum_{k=1}^{3}\int_0^\tau \left\langle |X|^2 X, \sigma_k \frac{\partial X}{\partial x}\right\rangle \circ dW_k(s) \\ &= H\left(X_0\right) + \sqrt{\gamma}\frac{8}{9}\sum_{k=1}^{3}\int_0^\tau \left\langle |X|^2 X, \sigma_k \frac{\partial X}{\partial x}\right\rangle dW_k(s) \\ &+ \frac{2\gamma}{9}\int_0^\tau \int_{\mathbb{R}} \left(\partial_x |X_1|^2 + \partial_x |X_2|^2\right)^2 dxds - \frac{4}{9}\gamma\int_0^\tau\int_{\mathbb{R}} \left| X_1 \frac{\partial X_2}{\partial x} - \frac{\partial X_1}{\partial x} X_2\right|^2 dxds \\ &+ \frac{12}{9}\gamma\int_0^\tau\int_{\mathbb{R}} \partial_x |X_1|^2 \partial_x |X_2|^2 dxds. \end{aligned}$$

Proof. The first equality follows by Stratonovich differential calculus applied to the functional H and because the process X is solution of (3.1.10). The calculation can be made rigorous by localization (H is C^2 but not bounded) and regularization through convolution. The second equality is obtained writing the evolution of H in its Ito formulation, that is

$$\begin{aligned} H\left(X(\tau)\right) &= H\left(X_0\right) + \sqrt{\gamma}\frac{8}{9}\sum_{k=1}^{3}\int_0^\tau \left\langle |X|^2 X, \sigma_k \frac{\partial X}{\partial x}\right\rangle dW_k(s) \\ &+ \frac{24}{9}\gamma\int_0^\tau \left\langle X, \partial_x X \mathcal{R}\mathrm{e}\left(X.\partial_x\overline{X}\right)\right\rangle ds - \frac{8}{9}\gamma\sum_{k=1}^{3}\int_0^\tau \left\langle X, \sigma_k\partial_x X \mathcal{R}\mathrm{e}\left(X.\overline{\sigma_k}\partial_x\overline{X}\right)\right\rangle ds, \end{aligned}$$

where we used the unitary of the Pauli matrices and $\sigma_k = \sigma_k^*$, for $k = 1, 2, 3$. Easy calculations lead to the expression given above. □

3.4 Diffusion limit of the Manakov-PMD equation : Proof of Theorem 3.1.3

The aim of this part is the proof of the convergence result given in Theorem 3.1.3. For this purpose we have to cut-off Equation (3.1.9) in order to get uniform bounds, with respect to ϵ, of high order moments of the $\mathbb{H}^2$ norm of the solution. Let us denote by X_ϵ^R the solution of the cut-off equation

$$\begin{cases} i\dfrac{\partial X_\epsilon^R(t)}{\partial t} + \dfrac{ib'}{\epsilon}\boldsymbol{\sigma}\left(\nu_\epsilon(t)\right)\dfrac{\partial X_\epsilon^R}{\partial x} + \dfrac{d_0}{2}\dfrac{\partial^2 X_\epsilon^R}{\partial x^2} + \Theta_R\left(\left\|X_\epsilon^R(t)\right\|_{\mathbb{H}^1}^2\right) F_{\nu_\epsilon(t)}\left(X_\epsilon^R\right) = 0 \\ X_0 = v \in \mathbb{H}^3(\mathbb{R}). \end{cases} \tag{3.4.1}$$

The proof will consist of the following steps :

3.4.1 We prove uniform bounds on the solution X_ϵ^R of (3.4.1). These bounds will enable us to prove tightness on $\mathcal{K}$.

3.4.2 We use the perturbed test function method to get convergence of the generators in some sense [38, 70, 88]. This method formally gives a candidate for the limit process.

3.4.3 Setting $Z_\epsilon^R = \left(X_\epsilon^R, \left\|X_\epsilon^R(.)\right\|_{\mathbb{H}^1}^2\right)$, we then prove that the family of laws $\mathcal{L}\left(Z_\epsilon^R\right) = \mathbb{P} \circ \left(Z_\epsilon^R\right)^{-1}$ is tight on $\mathcal{K}$ and we deduce that the process Z_ϵ^R converges in law, up to a subsequence.

3.4.4 Combining the previous steps and using the martingale problem formulation, we identify the limit and conclude to the weak convergence of the whole sequence X_ϵ^R.

3.4.5 Finally we get rid of the cut off and we conclude that the sequence $(X_\epsilon)_{\epsilon>0}$ converges in law to X in $\mathcal{E}\left(\mathbb{H}^1\right)$ using the Skorokhod Theorem.

3.4.1 Uniform bounds on X_ϵ^R

Recall that a unique solution $\Psi_\epsilon^R \in C\left(\mathbb{R}_+, \mathbb{H}^3\right)$ of the following equation exists (see Section 3.2).

$$id\Psi_\epsilon^R(t) + \left\{ \frac{ib'}{\epsilon}\sigma_3 \frac{\partial \Psi_\epsilon^R(t)}{\partial x} + \frac{d_0}{2}\frac{\partial^2 \Psi_\epsilon^R(t)}{\partial x^2} + \frac{\gamma_s}{\epsilon^2}\sigma_3 \Psi_\epsilon^R(t) + \frac{i}{\epsilon^2}\gamma_c \Psi_\epsilon^R(t) \right\} dt \qquad (3.4.2)$$
$$- \frac{\sqrt{\gamma_c}}{\epsilon}\sigma_1 \Psi_\epsilon^R d\widetilde{W}_1(t) - \frac{\sqrt{\gamma_c}}{\epsilon}\sigma_2 \Psi_\epsilon^R d\widetilde{W}_2(t) + \Theta_R\left(\left\|\Psi_\epsilon^R(t)\right\|_{\mathbb{H}^1}^2\right) f\left(\Psi_\epsilon^R(t)\right) dt = 0.$$

A solution X_ϵ^R to (3.4.1) is then easily deduced from $X_\epsilon^R(t) = Z_\epsilon^{-1}(t)\Psi_\epsilon^R(t)$.

Lemma 3.4.1. *Let $\psi_0 \in \mathbb{H}^3$ and Ψ_ϵ^R be the solution of (3.4.2); then for all $T > 0$ there exists a positive constant $C(R,T)$ independent of ϵ, such that, a.s for every t in $[0,T]$,*

$$\left\|\Psi_\epsilon^R(t)\right\|_{\mathbb{H}^3} \leqslant C(R,T).$$

Similar bounds hold for $X_\epsilon^R(t) = Z_\epsilon^{-1}(t)\Psi_\epsilon^R(t)$ for any $t \in [0,T]$ since Z_ϵ^{-1} is almost surely bounded.

Proof of Lemma 3.4.1. The bounds on the $\mathbb{H}^3$ norm are obtained using an energy method. Using a regularization procedure, Ito formula applied to $\left\|\partial_x \Psi_\epsilon^R(t)\right\|_{\mathbb{L}^2}^2$ and Equation (3.4.2), we obtain for all $t \in [0,T]$

$$\left\|\partial_x \Psi_\epsilon^R(t)\right\|_{\mathbb{L}^2}^2 = \left\|\partial_x \psi_0\right\|_{\mathbb{L}^2}^2 + 2\int_0^t \left\langle \partial_x \Psi_\epsilon^R(s), d\partial_x \Psi_\epsilon^R(s)\right\rangle + \frac{2\gamma_c}{\epsilon^2}\int_0^t \left\|\partial_x \Psi_\epsilon^R(s)\right\|_{\mathbb{L}^2}^2 ds,$$

hence

$$\begin{aligned}\left\|\partial_x \Psi_\epsilon^R(t)\right\|_{\mathbb{L}^2}^2 &\leqslant \left\|\partial_x \psi_0\right\|_{\mathbb{L}^2}^2 + 2\int_0^t \Theta_R\left(\left\|\Psi_\epsilon^R(s)\right\|_{\mathbb{H}^1}^2\right)\left\|\partial_x f\left(\Psi_\epsilon^R(s)\right)\right\|_{\mathbb{L}^2}\left\|\partial_x \Psi_\epsilon^R(s)\right\|_{\mathbb{L}^2} ds \\ &\leqslant \left\|\partial_x \psi_0\right\|_{\mathbb{L}^2}^2 + C(R)\int_0^t \left\|\partial_x \Psi_\epsilon^R(s)\right\|_{\mathbb{L}^2}^2 ds.\end{aligned}$$

By Gronwall Lemma we deduce that

$$\left\|\partial_x \Psi_\epsilon^R(t)\right\|_{\mathbb{L}^2}^2 \leqslant \|\partial_x \psi_0\|_{\mathbb{L}^2}^2 \exp\left(C(R)T\right).$$

Using the same procedure for $\left\|\partial_x^2 X_\epsilon^R\right\|_{\mathbb{L}^2}^2$, Gagliardo-Nirenberg and Young inequalities,

$$\begin{aligned}
&\left\|\partial_x^2 \Psi_\epsilon^R(t)\right\|_{\mathbb{L}^2}^2 - \left\|\partial_x^2 \psi_0\right\|_{\mathbb{L}^2}^2 \\
&\qquad \leqslant C \int_0^t \Theta_R\left(\left\|\Psi_\epsilon^R(s)\right\|_{\mathbb{H}^1}^2\right)\left(\left\|\Psi_\epsilon^R(s)\right\|_{\mathbb{L}^\infty}^2 + 1\right)\left\|\partial_x^2 \Psi_\epsilon^R(s)\right\|_{\mathbb{L}^2}^2 ds \\
&\qquad\quad + C \int_0^t \Theta_R\left(\left\|\Psi_\epsilon^R(s)\right\|_{\mathbb{H}^1}^2\right)\left\|\Psi_\epsilon^R(s)\right\|_{\mathbb{L}^\infty}^4 \left\|\partial_x \Psi_\epsilon^R(s)\right\|_{\mathbb{L}^2}^6 ds.
\end{aligned}$$

By Sobolev embeddings, properties of the cut off function and again Gronwall Lemma, we conclude

$$\left\|\partial_x^2 \Psi_\epsilon^R(t)\right\|_{\mathbb{L}^2}^2 \leqslant \left\|\partial_x^2 \psi_0\right\|_{\mathbb{L}^2}^2 C(R,T).$$

A bound on $\left\|\partial_x^3 X_\epsilon^R\right\|_{\mathbb{L}^2}^2$ may be obtained similarly using the previous estimates and Gronwall Lemma. □

Remark 3.4.1. *To prove the convergence result we need an initial data in* $\mathbb{H}^3(\mathbb{R})$. *We will explain later where exactly we need this extra regularity but this is mainly due to the fact that we prove tightness in* $C\left([0,T], \mathbb{H}^1\right)$.

Remark 3.4.2. *Note that we first prove convergence in law for the couple of random variables* $\left(X_\epsilon^R, \left\|X_\epsilon^R(.)\right\|_{\mathbb{H}^1}^2\right)$. *This is due to the fact that the cut off is not continuous for the weak topology in* $\mathbb{H}^1$ *neither for the strong topology in* $\mathbb{H}^1_{loc}$. *These arguments have already been used in [19].*

3.4.2 The perturbed test function method

Note that the process X_ϵ^R is not Markov due to the presence of ν_ϵ. However $(X_\epsilon^R, \nu_\epsilon)$ is Markov, by construction of ν. We denote by $\mathscr{L}_\epsilon^R$ its infinitesimal generator. Let us compute $\mathscr{L}_\epsilon^R f$ for f sufficiently smooth such that f maps $\mathbb{H}^{-1} \times \mathbb{S}^3$ into $\mathbb{R}$ and is of class C_b^2. Let $\langle ., . \rangle$ be the duality product between $\mathbb{H}^1$ and $\mathbb{H}^{-1}$. Then, for $\epsilon > 0$ and for X_ϵ^R solution of the Manakov-PMD Equation (3.4.1),

$$\begin{aligned}
f\left(X_\epsilon^R(t), \nu_\epsilon(t)\right) - f(v,y) =& f\left(X_\epsilon^R(t), \nu_\epsilon(t)\right) - f\left(v, \nu_\epsilon(t)\right) + f\left(v, \nu_\epsilon(t)\right) - f(v,y) \\
=& \left\langle D_v f\left(v, \nu_\epsilon(t)\right), X_\epsilon^R(t) - v \right\rangle + R\left(X_\epsilon^R(t), v\right) \\
&+ f\left(v, \nu_\epsilon(t)\right) - f(v,y),
\end{aligned}$$

where

$$R\left(X_\epsilon^R(t), v\right) = \int_0^1 (1-\theta)\left\langle D_v^2 f\left(v + \theta\left(X_\epsilon^R(t) - v\right)\right)\left(X_\epsilon^R(t) - v\right), X_\epsilon^R(t) - v \right\rangle d\theta,$$

and $D_v^2 f(v) \in \mathcal{L}(\mathbb{H}^{-1}, \mathbb{H}^1)$. Thus

$$\begin{aligned}
&\frac{1}{t}\mathbb{E}\Big(f\left(X_\epsilon^R(t), \nu_\epsilon(t)\right) - f(v,y)\big| \left(X(0), \nu(0)\right) = (v,y) \Big) \\
&= \mathbb{E}\left(\left\langle D_v f\left(v, \nu_\epsilon(t)\right), \frac{X_\epsilon^R(t) - v}{t} \right\rangle \middle| \left(X(0), \nu(0)\right) = (v,y) \right) \\
&\quad + \mathbb{E}\left(\frac{R\left(X_\epsilon^R(t), v\right)}{t} \middle| X(0) = v \right) + \mathbb{E}\left(\frac{f\left(v, \nu_\epsilon(t)\right) - f\left(v,y\right)}{t} \middle| \nu(0) = y \right).
\end{aligned}$$

We know by Theorem 3.1.1 that if $v \in \mathbb{H}^3$ then $X_\epsilon^R \in C^1\left([0,T], \mathbb{H}^1\right)$. Thus by the mean value Theorem, Equation (3.4.1), the almost sure boundedness of ν, Lemma 3.4.1 and the conservation of the $\mathbb{L}^2$ norm :

$$\begin{aligned}
\frac{1}{t}\left\|X_\epsilon^R(t) - v\right\|_{\mathbb{L}^2} &\leqslant \sup_{s\in[0,t]} \left\|\partial_s X_\epsilon^R(s)\right\|_{\mathbb{L}^2} \\
&\leqslant \sup_{s\in[0,t]} \left(\left\| \frac{b'}{\epsilon}\boldsymbol{\sigma}\left(\nu_\epsilon(t)\right)\partial_x X_\epsilon^R(s) \right\|_{\mathbb{L}^2} + \left\| \frac{d_0}{2}\partial_x^2 X_\epsilon^R(s) \right\|_{\mathbb{L}^2} \right. \\
&\qquad \left. + \left\| \Theta_R\left(\left\|X_\epsilon^R(s)\right\|_{\mathbb{H}^1}^2 \right) F_{\nu_\epsilon(s)}\left(X_\epsilon^R(s)\right) \right\|_{\mathbb{L}^2} \right) \\
&\leqslant \left(\frac{b'}{\epsilon} + \frac{d_0}{2} \right) C(R,T) + 2RC\left\|v\right\|_{\mathbb{L}^2}.
\end{aligned}$$

Thus, by the boundedness of $D_v^2 f$, the continuity of $t \mapsto X_\epsilon^R(t)$ in $\mathbb{L}^2$ and the previous bounds, we conclude that

$$\frac{R\left(X_\epsilon^R(t), v\right)}{t} \leqslant C\left(R, T, \epsilon\right) \sup_{w\in\mathbb{H}^1} \left\|D_v^2 f(w)\right\|_{\mathcal{L}(\mathbb{H}^{-1},\mathbb{H}^1)} \left(1 + \left\|v\right\|_{\mathbb{L}^2}\right) \left\|X_\epsilon(t) - v\right\|_{\mathbb{L}^2}$$

and the right hand side above tends to zero as t goes to zero. Now, we perform the change of variables $t' = t/\epsilon^2$, to get

$$\frac{1}{t}\mathbb{E}\left(f\left(v, \nu_\epsilon(t)\right) - f(v,y)\right| \nu(0) = y) = \frac{1}{\epsilon^2 t'}\mathbb{E}\left(f\left(v, \nu(t')\right) - f(v,y)\right| \nu(0) = y).$$

Thus, using the Markov property of the process ν, and using Equation (3.4.1) again, we get an expression of the infinitesimal generator $\mathscr{L}_\epsilon^R$ of the Markov process $\left(X_\epsilon^R, \nu_\epsilon\right)$:

$$\begin{aligned}
\mathscr{L}_\epsilon^R f(v,y) &= \lim_{t\to 0}\frac{1}{t}\left(\mathbb{E}\Big(f\left(X_\epsilon^R(t), \nu_\epsilon(t)\right) - f(v,y)\big| \left(X(0), \nu(0)\right) = (v,y) \Big) \right) \\
&= \left\langle D_v f(v,y), \partial_t X_\epsilon^R(t)\big|_{t=0} \right\rangle + \frac{1}{\epsilon^2}\mathscr{L}_\nu f(v,y) \\
&= \left\langle D_v f(v,y), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\left\|v\right\|_{\mathbb{H}^1}^2\right) F_y\left(v\right) \right\rangle \qquad (3.4.3) \\
&\quad - \frac{1}{\epsilon}\left\langle D_v f(v,y), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x} \right\rangle + \frac{1}{\epsilon^2}\mathscr{L}_\nu f(v,y),
\end{aligned}$$

where $\mathscr{L}_\nu$ is the infinitesimal generator of ν and $\mathscr{D}_\nu$ its domain. The perturbed test function method gives (by identifying its infinitesimal generator) an idea of the limit law of the sequence $\left(X_\epsilon^R\right)_{\epsilon>0}$. It provides in addition convergences that are useful to prove the weak convergence of the sequence of measures $\left(\mathcal{L}\left(X_\epsilon^R\right)\right)_{\epsilon>0}$.

Proposition 3.4.1 (Perturbed test function method)**.** *There exists a limiting infinitesimal generator $\left(\mathscr{L}^R, \mathscr{D}^R\right)$ such that for all sufficiently smooth and real valued functions $f \in \mathscr{D}^R$ and for all positive ϵ, there exists a test function f_ϵ and positive constants $C_1(K)$ and $C_2(K)$ satisfying*

$$\sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}} |f_\epsilon(v,y) - f(v)| \leqslant \epsilon C_1(K) \tag{3.4.4}$$

$$\sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}} \left|\mathscr{L}_\epsilon^R f_\epsilon(v,y) - \mathscr{L}^R f(v)\right| \leqslant \epsilon C_2(K), \tag{3.4.5}$$

where $\mathcal{B}(K)$ denotes the closed ball of $\mathbb{H}^3(\mathbb{R})$ with radius K.

Proof. The idea is to prove that for all suitable test function f, one can find a function f_ϵ of the form

$$f_\epsilon(v,y) = f(v) + \epsilon f^1(v,y) + \epsilon^2 f^2(v,y), \tag{3.4.6}$$

such that Proposition 3.4.1 holds. We plug this expression of f_ϵ into (3.4.3) and formally compute the expression of $\mathscr{L}_\epsilon^R f_\epsilon$:

$$\begin{aligned}
\mathscr{L}_\epsilon^R f_\epsilon(v,y) = &\left\langle D_v f(v), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|_{\mathbb{H}^1}^2\right) F_y(v)\right\rangle - \left\langle D_v f^1(v,y), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle \\
&+ \mathscr{L}_\nu f^2(v,y) + \frac{1}{\epsilon}\mathscr{L}_\nu f^1(v,y) - \frac{1}{\epsilon}\left\langle D_v f(v), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle \\
&+ \epsilon\left\langle D_v f^1(v,y), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|_{\mathbb{H}^1}^2\right) F_y(v)\right\rangle \\
&- \epsilon\left\langle D_v f^2(v,y), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle \\
&+ \epsilon^2\left\langle D_v f^2(v,y), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|_{\mathbb{H}^1}^2\right) F_y(v)\right\rangle,
\end{aligned} \tag{3.4.7}$$

and we notice that $\mathscr{L}_\nu f(v)$ is identically zero because f does not depend on $\nu = (\nu_1, \nu_2)$. The aim is to wisely choose the functions f^1 and f^2 and the regularity of f so that $\mathscr{L}_\epsilon^R f_\epsilon$ is well defined and that f_ϵ and $\mathscr{L}_\epsilon^R f_\epsilon$ converge in the sense of Proposition 3.4.1. In particular, we need to cancel the terms with a factor $1/\epsilon$ and we need the terms with factors ϵ or ϵ^2 to be $\mathcal{O}(\epsilon)$ on bounded sets. In order to cancel the $1/\epsilon$ terms, we look for a function f^1 solution of the Poisson equation :

$$\mathscr{L}_\nu f^1(v,y) = \left\langle D_v f(v), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle. \tag{3.4.8}$$

By Corollary 3.5.1, we know that

$$\mathbb{E}_\Lambda\left(g_j(\nu)\right) = 0 \qquad \forall j = 1,2,3.$$

We deduce that $\left\langle D_v f(v), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle$, which is a linear combination of $m_j = g_j(y)$ (see (3.1.6)), is of null mass with respect to the invariant measure Λ. Hence $\left\langle D_v f(v), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle$ is a function of $y \in \mathbb{S}^3$, which satisfies the assumptions of Proposition 3.5.1, provided that f is sufficiently smooth i.e $f \in C^1\left(\mathbb{H}^{-1}\right)$ and $v \in \mathbb{L}^2$. It follows that the solution f^1 of the Poisson Equation (3.4.8) can be written as :

$$\begin{aligned} f^1(v,y) &= \mathscr{L}_\nu^{-1}\left(\left\langle D_v f(v), b'\boldsymbol{\sigma}(.)\frac{\partial v}{\partial x}\right\rangle\right)(y) \\ &= -\left\langle D_v f(v), b'\widetilde{\boldsymbol{\sigma}}(y)\frac{\partial v}{\partial x}\right\rangle, \end{aligned} \tag{3.4.9}$$

where

$$\widetilde{\boldsymbol{\sigma}}(y) = \int_0^{+\infty} \mathbb{E}\left(\boldsymbol{\sigma}\left(\nu(t)\right)|\, \nu(0) = y\right) dt. \tag{3.4.10}$$

By Proposition 3.5.1, there is a positive constant M such that

$$\left\| |\widetilde{\boldsymbol{\sigma}}(y)| \right\|_\infty \leqslant M, \quad \forall y \in \mathbb{S}^3, \tag{3.4.11}$$

and $f^1(v,y)$ is a continuous bounded function of y for $v \in \mathbb{L}^2$. We now have to choose the function f^2, but we cannot choose $\mathscr{L}_\nu f^2$ cancelling the terms

$$\left\langle D_v f(v), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right) F_y(v)\right\rangle - \left\langle D_v f^1(v,y), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle,$$

because they do not satisfy the null mass condition with respect to Λ. Hence we look for a solution f^2 of the Poisson equation :

$$\begin{aligned} \mathscr{L}_\nu f^2(v,y) = &-\left\langle D_v f(v), i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right) F_y(v)\right\rangle + \left\langle D_v f(v), i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right) F(v)\right\rangle \\ &+ \left\langle D_v f^1(v,y), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle - \mathbb{E}_\Lambda\left(\left\langle D_v f^1(v,y), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle\right), \end{aligned} \tag{3.4.12}$$

where, due to (3.4.9),

$$\begin{aligned} &\left\langle D_v f^1(v,y), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle \\ &\quad = -(b')^2\left\langle D_v^2 f(v)\widetilde{\boldsymbol{\sigma}}(y)\frac{\partial v}{\partial x}, \boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle - (b')^2\left\langle D_v f(v), \widetilde{\boldsymbol{\sigma}}(y)\boldsymbol{\sigma}(y)\frac{\partial^2 v}{\partial x^2}\right\rangle. \end{aligned} \tag{3.4.13}$$

Moreover thanks to expression (3.4.13), Fubini Theorem and Corollary 3.5.1

$$\begin{aligned} &-\mathbb{E}_\Lambda\left(\left\langle D_v f^1(v,y), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle\right) \\ &= (b')^2\sum_{j,k=1}^3\left\langle D_v^2 f(v)\sigma_k\frac{\partial v}{\partial x}, \sigma_j\frac{\partial v}{\partial x}\right\rangle\int_0^{+\infty}\mathbb{E}_\Lambda\left(g_k\left(\nu(t)\right)g_j\left(\nu(0)\right)\right)dt \\ &\quad + (b')^2\sum_{j,k=1}^3\left\langle D_v f(v), \sigma_k\sigma_j\frac{\partial^2 v}{\partial x^2}\right\rangle\int_0^{+\infty}\mathbb{E}_\Lambda\left(g_k\left(\nu(t)\right)g_j\left(\nu(0)\right)\right)dt \\ &= \frac{\gamma}{2}\sum_{k=1}^3\left\langle D_v^2 f(v)\sigma_k\frac{\partial v}{\partial x}, \sigma_k\frac{\partial v}{\partial x}\right\rangle + \frac{3\gamma}{2}\left\langle D_v f(v), \frac{\partial^2 v}{\partial x^2}\right\rangle, \end{aligned} \tag{3.4.14}$$

where $\gamma = (b')^2/6\gamma_c$. Provided that f is of class $C^2(\mathbb{H}^{-1})$ and $v \in \mathbb{H}^1$ and because $f^1(v,.)$ is of class $C_b^2(\mathbb{S}^3)$ for any $v \in \mathbb{H}^1$, we can now define, by Proposition 3.5.1, a unique solution, up to a constant, to the Poisson Equation (3.4.12). This solution f^2 is expressed as :

$$\begin{aligned} f^2(v,y) &= \mathscr{L}_\nu^{-1}\left(\left\langle D_v f(v), i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right)\left(F_y(v) - F(v)\right)\right\rangle\right) \\ &\quad -\mathscr{L}_\nu^{-1}\left(\left\langle D_v f^1(v,y), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle - \mathbb{E}_\Lambda\left(\left\langle D_v f^1(v,y), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle\right)\right) \\ &= \left\langle D_v f(v), i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right)\widetilde{F}(v,y)\right\rangle \qquad (3.4.15) \\ &\quad -(b')^2\sum_{k,l=1}^{3}\left\langle D_v^2 f(v)\sigma_k\frac{\partial v}{\partial x}, \sigma_l\frac{\partial v}{\partial x}\right\rangle\widetilde{\widetilde{g}}_{k,l}(y) - \left\langle D_v f(v), (b')^2\widetilde{\widetilde{\boldsymbol{\sigma}}}(y)\frac{\partial^2 v}{\partial x^2}\right\rangle, \end{aligned}$$

where

$$\widetilde{F}(v,y) = \int_0^{+\infty}\mathbb{E}\left(F_{\nu(t)}(v) - F(v)\middle|\,\nu(0) = y\right)dt,$$

and

$$\widetilde{\widetilde{g}}_{k,l}(y) = \int_0^{+\infty}\left(\int_t^{+\infty}\mathbb{E}\left(g_k(\nu(s))\,g_l(\nu(t))\middle|\,\nu(0) = y\right)ds - \frac{\gamma}{2(b')^2}\delta_{kl}\right)dt,$$

and

$$\widetilde{\widetilde{\boldsymbol{\sigma}}}(y) = \int_0^{+\infty}\left(\int_t^{+\infty}\mathbb{E}\left(\boldsymbol{\sigma}(\nu(s))\,\boldsymbol{\sigma}(\nu(t))\middle|\,\nu(0) = y\right)ds - \frac{3\gamma}{2(b')^2}\right)dt.$$

Replacing $\mathscr{L}_\nu f^1$ and $\mathscr{L}_\nu f^2$ in Equation (3.4.7), respectively by the right hand side of (3.4.8) and (3.4.12) and using expression (3.4.14) we get :

$$\begin{aligned} \mathscr{L}_\epsilon^R f_\epsilon(v,y) &= \left\langle D_v f(v), \left(\frac{id_0}{2} + \frac{3\gamma}{2}\right)\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right)F(v)\right\rangle \qquad (3.4.16) \\ &\quad +\frac{\gamma}{2}\sum_{k=1}^{3}\left\langle D_v^2 f(v)\sigma_k\frac{\partial v}{\partial x}, \sigma_k\frac{\partial v}{\partial x}\right\rangle \\ &\quad +\epsilon\left\langle D_v f^1(v,y), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right)F_y(v)\right\rangle \\ &\quad -\epsilon\left\langle D_v f^2(v,y), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle \\ &\quad +\epsilon^2\left\langle D_v f^2(v,y), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right)F_y(v)\right\rangle, \end{aligned}$$

and we define the limiting operator by :

$$\begin{aligned} \mathscr{L}^R f(v) &= \left\langle D_v f(v), \left(\frac{id_0}{2} + \frac{3\gamma}{2}\right)\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right)F(v)\right\rangle \qquad (3.4.17) \\ &\quad +\frac{\gamma}{2}\sum_{k=1}^{3}\left\langle D_v^2 f(v)\sigma_k\frac{\partial v}{\partial x}, \sigma_k\frac{\partial v}{\partial x}\right\rangle. \end{aligned}$$

Hence if we define $\mathscr{D}^R$ as the space of functions which are the restriction to $\mathbb{H}^3$ of functions f from $\mathbb{H}^{-1}$ into $\mathbb{R}$ of class $C^3(\mathbb{H}^{-1})$ and such that f and its first three derivatives are bounded on bounded sets of $\mathbb{H}^{-1}$, then the functions f^1 and f^2 are well defined for $f \in \mathscr{D}^R$. Moreover if $f \in \mathscr{D}^R$ then $\mathscr{L}_\epsilon^R f_\epsilon$ is well defined for $v \in \mathbb{H}^3$.

We now write that

$$\sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}} |f_\epsilon(v,y) - f(v)| \leqslant \epsilon \sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}} \left|f^1(v,y)\right| + \epsilon^2 \sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}} \left|f^2(v,y)\right|,$$

and use the following result, which is proved in Section 3.6.

Lemma 3.4.2. *Let $f \in \mathscr{D}^R$ and f^1 and f^2 be respectively solution of Equation* (3.4.8) *and* (3.4.12). *Then*

$$\sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}} \left|f^1(v,y)\right| \leqslant C_1(K) \qquad \text{and} \qquad \sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}} \left|f^2(v,y)\right| \leqslant C_2(K).$$

This proves the first convergence of Proposition 3.4.1. With $\mathscr{L}^R f(v)$ given by (3.4.17), the second convergence (3.4.5) in Proposition 3.4.1 follows from (3.4.16) and the next Lemma, which is proved in Section 3.6. □

Lemma 3.4.3. *Let $f \in \mathscr{D}^R$ and f^1, f^2 be respectively solutions of Equation* (3.4.8) *and* (3.4.12). *Then*

$$\sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}} \left|\left\langle D_v f^1(v,y), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right) F_y(v)\right\rangle\right| \leqslant C_1(K),$$

$$\sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}} \left|\left\langle D_v f^2(v,y), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle\right| \leqslant C_2(K),$$

$$\sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}} \left|\left\langle D_v f^2(v,y), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right) F_y(v)\right\rangle\right| \leqslant C_3(K).$$

3.4.3 Tightness of the family of probability measures $\left(\mathcal{L}\left(Z_\epsilon^R\right)\right)_{\epsilon>0}$

To prove tightness on $\mathcal{K}$ of the sequence of probability measure $\mathcal{L}\left(Z_\epsilon^R\right) = \mathbb{P}\circ\left(Z_\epsilon^R\right)^{-1}$, we need to obtain uniform bounds in ϵ on Z_ϵ^R in the space

$$\left(C\left([0,T],\mathbb{H}^2\right) \cap C^\alpha\left([0,T],\mathbb{H}^{-1}\right)\right) \times C^\delta\left([0,T],\mathbb{R}\right),$$

for suitable $\alpha, \delta > 0$. Note that uniform bound of X_ϵ^R in $C\left([0,T],\mathbb{H}^2\right)$ are given by Lemma 3.4.1. The perturbed test function method will enable us to get uniform bound in $C^\alpha\left([0,T],\mathbb{H}^{-1}\right)$. Such bounds can not be directly obtained using Equation (3.4.1) because of the $1/\epsilon$ term. In order to obtain such bounds we use again the perturbed test function method for convenient test functions. Let $(\widetilde{e_j})_{j\in\mathbb{N}^*}$ be a complete orthonormal system in $\mathbb{L}^2$. Recall that $\langle .,. \rangle$ is the duality product between

$\mathbb{H}^1$-$\mathbb{H}^{-1}$ and $(.,.)_{\mathbb{L}^2}$ the inner product in $\mathbb{L}^2$. By definition of $\mathbb{H}^s, s \in \mathbb{R}$ we can define a complete orthonormal system $(e_j)_{j\in\mathbb{N}^*}$ on $\mathbb{H}^1$ from $(\widetilde{e_j})_{j\in\mathbb{N}^*}$

$$\|v\|^2_{\mathbb{H}^{-1}} = \left\|\left(1+\xi^2\right)^{-1/2}\widehat{v}\right\|^2_{\mathbb{L}^2} = \sum_{j=1}^{+\infty}\left(\left(1+\xi^2\right)^{-1/2}\widehat{v},\widehat{\widetilde{e_j}}\right)^2_{\mathbb{L}^2} = \sum_{j=1}^{+\infty}\langle e_j, v\rangle^2,$$

where $e_j = \mathcal{F}^{-1}\left(\left(1+\xi^2\right)^{-1/2}\widehat{\widetilde{e_j}}\right)$ for any $j \in \mathbb{N}^*$. We denote by $(f_j)_{j\in\mathbb{N}^*}$ the family of test functions in $\mathscr{D}^R$ defined by

$$\begin{aligned} f_j : \mathbb{H}^{-1} &\to \mathbb{R} \\ v &\mapsto f_j(v) = \langle e_j, v\rangle. \end{aligned}$$

For $v \in \mathbb{H}^3$, we also consider particular perturbed test functions $f_{j,\epsilon}$ of the form

$$f_{j,\epsilon}(v,y) = f_j(v) + \epsilon f_j^1(v,y), \tag{3.4.18}$$

where, for all j in $\mathbb{N}^*$, $f_j^1(v,y) = \langle e_j, \varphi^1(v,y)\rangle$ for a given function φ^1 with values in $\mathbb{H}^2$. We now choose φ^1 as a solution of the Poisson equation in y :

$$\mathscr{L}_\nu\varphi^1(v,y) - b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x} = 0, \tag{3.4.19}$$

whose explicit formulation is given by (see Proposition 3.5.1):

$$\varphi^1(v,y) = -b'\widetilde{\boldsymbol{\sigma}}(y)\frac{\partial v}{\partial x}, \tag{3.4.20}$$

where $\widetilde{\boldsymbol{\sigma}}(y)$ is given by (3.4.10). We point out that φ^1 behaves in its first variable like $\frac{\partial}{\partial x}$ and is linear in v. Consequently for all j in $\mathbb{N}^*$:

$$\begin{aligned} &\mathscr{L}^R_\epsilon f_{j,\epsilon}\left(X^R_\epsilon(t),\nu_\epsilon(t)\right) \\ &= \left\langle e_j, \frac{id_0}{2}\frac{\partial^2 X^R_\epsilon(t)}{\partial x^2} + i\Theta_R\left(\left\|X^R_\epsilon(t)\right\|^2_{\mathbb{H}^1}\right)F_{\nu_\epsilon(t)}\left(X^R_\epsilon(t)\right)\right\rangle \\ &+ \left\langle e_j, (b')^2\widetilde{\boldsymbol{\sigma}}(\nu_\epsilon(t))\boldsymbol{\sigma}(\nu_\epsilon(t))\frac{\partial^2 X^R_\epsilon(t)}{\partial x^2}\right\rangle \\ &- \epsilon\left\langle e_j, b'\widetilde{\boldsymbol{\sigma}}(\nu_\epsilon(t))\frac{\partial}{\partial x}\left(\frac{id_0}{2}\frac{\partial^2 X^R_\epsilon(t)}{\partial x^2} + i\Theta_R\left(\left\|X^R_\epsilon(t)\right\|^2_{\mathbb{H}^1}\right)F_{\nu_\epsilon(t)}\left(X^R_\epsilon(t)\right)\right)\right\rangle. \end{aligned} \tag{3.4.21}$$

For all $t \in [0,T]$, we define the process M^R_ϵ with values in $\mathbb{H}^{-1}$ given for any j in $\mathbb{N}^*$ by :

$$\begin{aligned} &\left\langle e_j, M^R_\epsilon(t)\right\rangle \\ =\ & f_{j,\epsilon}\left(X^R_\epsilon,\nu_\epsilon\right)(t) - f_{j,\epsilon}(v,y) - \int_0^t \mathscr{L}^R_\epsilon f_{j,\epsilon}\left(X^R_\epsilon(s),\nu_\epsilon(s)\right)ds \\ =\ & \left\langle e_j, X^R_\epsilon - v\right\rangle + \epsilon\left\langle e_j, \varphi^1(X^R_\epsilon,\nu_\epsilon) - \varphi^1(v,y)\right\rangle - \int_0^t \mathscr{L}^R_\epsilon f_{j,\epsilon}\left(X^R_\epsilon(s),\nu_\epsilon(s)\right)ds. \end{aligned}$$

Given the fact that $\mathscr{L}_\epsilon^R$ is the infinitesimal generator of the continuous Markov process $\left(X_\epsilon^R, \nu_\epsilon\right)$ and $\mathscr{L}_\epsilon^R f_{j,\epsilon}$ is well defined because $f_j \in \mathscr{D}^R$, then $\left\langle e_j, M_\epsilon^R(t)\right\rangle$ is a real valued continuous martingale. Moreover it is a square integrable martingale, as follows from the bounds on the $\mathbb{H}^3$ norm of X_ϵ^R obtained in Lemma 3.4.1. To prove tightness of the family of probability measures $\mathcal{L}\left(Z_\epsilon^R\right)$ on $\mathcal{K}$ we need estimates of moments on the processes X_ϵ^R and $\left\|X_\epsilon^R(.)\right\|_{\mathbb{H}^1}^2$. Before proving these estimates we introduce a process Y_ϵ^R close in probability to X_ϵ^R for which it will be easier to get those estimates, using in particular the Kolmogorov criterion. The idea is to use Lemma 3.4.4 below to get tightness of the family $\mathcal{L}\left(Z_\epsilon^R\right)$ from convergence in law of a subsequence of Y_ϵ^R.

Lemma 3.4.4. *Let us define the process Y_ϵ^R as*

$$X_\epsilon^R(t) - Y_\epsilon^R(t) = \epsilon\left(\varphi^1(v,y) - \varphi^1\left(X_\epsilon^R(t), \nu_\epsilon(t)\right)\right), \quad \forall t \in [0,T]; \tag{3.4.22}$$

then for all $\delta > 0$:

$$\mathbb{P}\left(\left\|X_\epsilon^R - Y_\epsilon^R\right\|_{C([0,T],\mathbb{H}^1)} > \delta\right) \leqslant \frac{\epsilon}{\delta} C_1(T,R).$$

Proof of Lemma 3.4.4. Using Markov inequality and Lemma 3.4.1 we get for all $\delta > 0$,

$$\begin{aligned}
&\mathbb{P}\left(\sup_{t\in[0,T]} \left\|X_\epsilon^R(t) - Y_\epsilon^R(t)\right\|_{\mathbb{H}^1} > \delta\right) \\
&\leqslant \frac{\epsilon}{\delta}\mathbb{E}\left(\sup_{t\in[0,T]} \left\|\varphi^1\left(X_\epsilon^R(t), \nu_\epsilon(t)\right) - \varphi^1(v,y)\right\|_{\mathbb{H}^1}\right) \\
&\leqslant \frac{\epsilon}{\delta}\mathbb{E}\left(\sup_{t\in[0,T]} \left\|b'\widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(t)\right)\frac{\partial X_\epsilon^R(t)}{\partial x} - b'\widetilde{\boldsymbol{\sigma}}(y)\frac{\partial v}{\partial x}\right\|_{\mathbb{H}^1}\right) \\
&\leqslant \epsilon\frac{2M}{\delta}C(T,R),
\end{aligned}$$

where M is given by (3.4.11). □

Note that the process Y_ϵ^R is also defined by the identity, for all j in $\mathbb{N}^*$ and all t in $[0,T]$:

$$\begin{aligned}
\left\langle e_j, Y_\epsilon^R(t)\right\rangle &= \left\langle e_j, X_\epsilon^R(t)\right\rangle - \epsilon\left\langle e_j, \varphi^1(v,y) - \varphi^1\left(X_\epsilon^R(t), \nu_\epsilon(t)\right)\right\rangle \\
&= \left\langle e_j, M_\epsilon^R(t)\right\rangle + \langle e_j, v\rangle + \int_0^t \mathscr{L}_\epsilon^R f_{j,\epsilon}\left(X_\epsilon^R(s), \nu_\epsilon(s)\right) ds.
\end{aligned} \tag{3.4.23}$$

Lemma 3.4.5. *For all $1 \geqslant \epsilon > 0$, there exist three positive constants $C_1(T,R)$, $C_2(T,R)$ and $C_3(T,R)$ depending on final time T and on the cut-off radius R, but independent of ϵ, such that*

$$\mathbb{E}\left(\left\|Y_\epsilon^R\right\|_{C([0,T],\mathbb{H}^2)}^4\right) \leqslant C_1(T,R), \tag{3.4.24}$$

$$\mathbb{E}\left(\left\|Y_\epsilon^R\right\|_{C^\alpha([0,T],\mathbb{H}^{-1})}\right) \leqslant C_2(T,R), \tag{3.4.25}$$

$$\mathbb{E}\left(\left\|\left\|Y_\epsilon^R\right\|_{\mathbb{H}^1}^2\right\|_{C^\delta([0,T],\mathbb{R})}\right) \leqslant C_3(T,R), \tag{3.4.26}$$

where $0 < \alpha < \frac{1}{2}$ *and* $\delta = \alpha/3 > 0$.

Proof of Lemma 3.4.5. Thanks to Lemma 3.4.1 we know that the solution X_ϵ^R of Equation (3.4.1) is uniformly bounded, for all ϵ, in $\mathbb{H}^3$ by a constant C depending on R and T. We conclude, using the explicit formulation of φ_1 given by (3.4.20) and Equation (3.4.22), that (3.4.24) holds.

To prove inequality (3.4.25), we first need an intermediate estimate that will be proved in Section 3.6:

Lemma 3.4.6. *There exists a positive constant* $C(R,T)$ *such that for all* $t, s \in [0,T]$

$$\mathbb{E}\left(\left\|Y_\epsilon^R(t) - Y_\epsilon^R(s)\right\|_{\mathbb{H}^{-1}}^4\right) \leqslant C(R,T)(t-s)^2.$$

Then we deduce from Lemma 3.4.6 :

$$\mathbb{E}\Big(\left\|Y_\epsilon^R\right\|_{\mathbb{W}^{\gamma,4}([0,T],\mathbb{H}^{-1})}^4\Big) \leqslant C(R,T),$$

for any $\gamma < 1/2$. We use the Sobolev embedding $\mathbb{W}^{\gamma,4}\left([0,T],\mathbb{H}^{-1}\right) \hookrightarrow C^\alpha\left([0,T],\mathbb{H}^{-1}\right)$ for $\gamma - \alpha > 1/4$ and $\gamma < 1/2$, which implies $\alpha < 1/4$. Thus we deduce the second inequality (3.4.25).

It remains to prove the last bound (3.4.26). Note that for $t, s \in [0,T]$

$$\begin{aligned}
&\left|\left\|Y_\epsilon^R(t)\right\|_{\mathbb{H}^1}^2 - \left\|Y_\epsilon^R(s)\right\|_{\mathbb{H}^1}^2\right| \\
\leqslant\ & C \sup_{r\in[0,T]} \left\|Y_\epsilon^R(r)\right\|_{\mathbb{H}^1} \left\|Y_\epsilon^R(t) - Y_\epsilon^R(s)\right\|_{\mathbb{H}^1} \\
\leqslant\ & C \sup_{r\in[0,T]} \left\|Y_\epsilon^R(r)\right\|_{\mathbb{H}^1} \sup_{r\in[0,T]} \left\|Y_\epsilon^R(r)\right\|_{\mathbb{H}^2}^{2/3} \left\|Y_\epsilon^R(t) - Y_\epsilon^R(s)\right\|_{\mathbb{H}^{-1}}^{1/3}.
\end{aligned}$$

It follows that if $\delta = \alpha/3$,

$$\left\|\left\|Y_\epsilon^R(.)\right\|_{\mathbb{H}^1}^2\right\|_{C^\delta([0,T],\mathbb{R})} \leqslant C \sup_{r\in[0,T]} \left\|Y_\epsilon^R(r)\right\|_{\mathbb{H}^2}^{5/3} \left\|Y_\epsilon^R\right\|_{C^\alpha([0,T],\mathbb{H}^{-1})}^{1/3}.$$

Inequality (3.4.26) is then implied by Hölder inequality, (3.4.24) and (3.4.25). □

Remark 3.4.3. *The extra* $\mathbb{H}^3$ *regularity is needed precisely in the first step of the above proof in order to estimate the* $\mathbb{H}^2$ *norm of* Y_ϵ^R*, which involves the gradient of* X_ϵ^R.

Proposition 3.4.2. *The family of laws* $\left(\mathcal{L}\left(Z_\epsilon^R\right)\right)_{\epsilon>0}$ *is tight on* $\mathcal{K}$.

Proof of Proposition 3.4.2. We set $\widetilde{Z}_\epsilon^R = \left(Y_\epsilon^R, \left\|Y_\epsilon^R(.)\right\|_{\mathbb{H}^1}^2\right)$. Denoting by $\mathcal{B}(K)$ the closed ball of $\left(C\left([0,T];\mathbb{H}^2\left(\mathbb{R}\right)\right) \cap C^\alpha\left([0,T];\mathbb{H}^{-1}\left(\mathbb{R}\right)\right)\right) \times C^\delta\left([0,T];\mathbb{R}\right)$ with radius K, for α and δ as in Lemma 3.4.5, we deduce using Ascoli-Arzela and Banach-Alaoglu theorems that $\mathcal{B}(K)$ is compact in $\mathcal{K}$. Using Markov inequality and Lemma 3.4.5, we get

$$\begin{aligned}
\mathbb{P}\left(\widetilde{Z}_\epsilon^R \notin \mathcal{B}(K)\right) &\leqslant \frac{1}{K}\mathbb{E}\left(\max\left\{\left\|Y_\epsilon^R\right\|_{C([0,T];\mathbb{H}^2)}, \left\|Y_\epsilon^R\right\|_{C^\alpha([0,T];\mathbb{H}^{-1})}, \left\|\left\|Y_\epsilon^R\right\|_{\mathbb{H}^1}^2\right\|_{C^\delta([0,T])}\right\}\right) \\
&\leqslant \frac{1}{K}\max\left(C_1^{1/4}(T,R), C_2(T,R), C_3(T,R)\right).
\end{aligned}$$

We conclude that the family of laws $\left(\mathcal{L}\left(\widetilde{Z}_\epsilon^R\right)\right)_{\epsilon>0}$ is tight on $\mathcal{K}$ and by the Prokhorov theorem we obtain the relative compactness of the sequence of laws $\left(\mathcal{L}\left(\widetilde{Z}_\epsilon^R\right)\right)_{\epsilon>0}$ i.e, up to a subsequence, the sequence $\mathcal{L}\left(\widetilde{Z}_\epsilon^R\right)$ weakly converges to a probability measure $\mathcal{L}\left(\widehat{Z}^R\right)$ where $\widehat{Z}^R = \left(\widehat{X}^R, \gamma^R\right)$. We may now use Lemma 3.4.4 to prove that the family of laws $\mathcal{L}\left(Z_\epsilon^R\right)$ is tight. Indeed it easily follows from Lemma 3.4.4 and the above convergence in law that for all $g \in C_b(\mathcal{K})$

$$\lim_{\epsilon\to 0} \mathbb{E}\left(g\left(Z_\epsilon^R\right)\right) = \mathbb{E}\left(g\left(\widehat{Z}^R\right)\right).$$

□

3.4.4 Convergence in law of the process X_ϵ^R

In order to get the convergence in law of the whole sequence $\left(X_\epsilon^R\right)_{\epsilon>0}$, it remains to characterize the limit, i.e. to prove that $\widehat{X}^R = X^R$, the solution of Equation (3.3.6), and that $\gamma^R(t) = \left\|X^R(t)\right\|^2_{\mathbb{H}^1}$ for any $t \in [0,T]$. The tool here will be the use of the martingale problem formulation introduced by Stroock-Varadhan in [100].

Proposition 3.4.3. *The whole sequence X_ϵ^R converges in law to X^R in $C([0,T],\mathbb{H}^1)$.*

Proof of the Proposition 3.4.3. In order to prove that any subsequence of X_ϵ^R converges to the same limit X^R, solution of Equation (3.3.6), we will prove the convergence of the martingale problem for suitable test functions $f \in \mathscr{D}^R$. To this purpose let us define, for $a \in \mathbb{H}^1$ with compact support, the particular test function $f_a(.) = \langle a, . \rangle$, so that $f_a \in \mathscr{D}^R$. From this particular choice, we construct a perturbed test function $f_{a,\epsilon}$

$$f_{a,\epsilon}(v,y) = f_a(v) + \epsilon f_a^1(v,y) + \epsilon^2 f_a^2(v,y),$$

obtained thanks to Proposition 3.4.1. The correctors f_a^1 and f_a^2 are chosen to be solution of the Poisson equations (3.4.8) and (3.4.12) for f_a. Let us denote by Z_ϵ^R a subsequence converging to $\widehat{Z}^R$ and define the $\mathbb{H}^{-1}$ valued process $\mathbf{N}_\epsilon^R\left(Z_\epsilon^R(t)\right)$, associated to Equation (3.4.1)

$$\left\langle a, \mathbf{N}_\epsilon^R\left(Z_\epsilon^R(t)\right)\right\rangle = f_{a,\epsilon}(X_\epsilon^R(t), \nu_\epsilon(t)) - f_{a,\epsilon}(v,y) - \int_0^t \mathscr{L}_\epsilon^R f_{a,\epsilon}\left(X_\epsilon^R(s), \nu_\epsilon(s)\right) ds,$$

where $\mathscr{L}_\epsilon^R$ is given by (3.4.7). We also define the process $\mathbf{N}^R\left(Z_\epsilon^R(t)\right)$

$$\left\langle a, \mathbf{N}^R\left(Z_\epsilon^R(t)\right)\right\rangle = f_a\left(X_\epsilon^R(t)\right) - f_a(v) - \int_0^t \mathscr{L}^R f_a\left(X_\epsilon^R(s)\right) ds,$$

where $\mathscr{L}^R$ is given by expression (3.4.17). Moreover we denote by $\mathscr{L}_{\gamma^R}^R$ the operator whose expression is given by (3.4.17) replacing $\left\|\widehat{X}^R(t)\right\|_{\mathbb{H}^1}$ by $\gamma^R(t)$ in the cut-off function. Let us now define $\left\langle a, \mathbf{N}^R\left(\widehat{Z}^R(t)\right)\right\rangle$ by

$$\left\langle a, \mathbf{N}^R\left(\widehat{Z}^R(t)\right)\right\rangle = f_a\left(\widehat{X}^R(t)\right) - f_a(v) - \int_0^t \mathscr{L}_{\gamma^R}^R f_a\left(\widehat{X}^R(s)\right) ds. \qquad (3.4.27)$$

The process $\left\langle a, \mathbf{N}_\epsilon^R\left(Z_\epsilon^R(t)\right)\right\rangle$ is a real continuous martingale because $\left(X_\epsilon^R, \nu_\epsilon\right)$ is a Markov process and because $\mathscr{L}_\epsilon^R f_{a,\epsilon}$ is well defined since $X_\epsilon^R(t) \in \mathbb{H}^3$. Moreover it is a square integrable martingale, as follows from the bounds on the $\mathbb{H}^3$ norm of X_ϵ^R obtained in Lemma 3.4.1. The above martingale property implies that for all $t, s \in [0,T], t \geqslant s$,

$$\mathbb{E}\left[\left\langle a, \mathbf{N}_\epsilon^R\left(Z_\epsilon^R(t)\right) - \mathbf{N}_\epsilon^R\left(Z_\epsilon^R(s)\right)\right\rangle \middle| \sigma\left(Z_\epsilon^R(u), \nu_\epsilon(u)\right), u \leqslant s\right] = 0.$$

It follows in particular that for all test functions $h_1, \ldots, h_m \in C_b\left(\mathbb{H}^1_{loc} \times \mathbb{R}\right)$ and $0 \leqslant t_1 \ldots < t_m \leqslant s \leqslant t$

$$\mathbb{E}\left[\left\langle a, \mathbf{N}_\epsilon^R\left(Z_\epsilon^R(t)\right) - \mathbf{N}_\epsilon^R\left(Z_\epsilon^R(s)\right)\right\rangle \prod_{j=1}^m h_j\left(Z_\epsilon^R(t_j)\right)\right] = 0.$$

Using Proposition 3.4.1, Lemma 3.4.1 and the boundedness of the functions h_j, we get

$$\mathbb{E}\left(\left\langle a, \mathbf{N}_\epsilon^R\left(Z_\epsilon^R(t)\right) - \mathbf{N}^R\left(Z_\epsilon^R(t)\right) - \mathbf{N}_\epsilon^R\left(Z_\epsilon^R(s)\right) + \mathbf{N}^R\left(Z_\epsilon^R(s)\right)\right\rangle \prod_{j=1}^m h_j(Z_\epsilon^R(t_j))\right)$$
$$\leqslant \epsilon C(R,T).$$

Let us consider a cut off function $\chi_{R_0} \in C_c^\infty(\mathcal{K})$ satisfying

$$\chi_{R_0}(u) = \begin{cases} 1 & \text{if } u \in \mathcal{B}_{\mathcal{K}}(R_0) \\ 0 & \text{if } u \notin \mathcal{B}_{\mathcal{K}}(2R_0) \end{cases}$$

where $\mathcal{B}_{\mathcal{K}}(R_0)$ denotes the closed ball of radius R_0 of the space $\mathcal{K}$ and R_0 is chosen such that $X_\epsilon^R \in \mathcal{B}_{\mathcal{K}}(R_0)$ a.s (see Lemma 3.4.1). Note that, by continuity of the functions χ_{R_0} and $\{h_j\}_{j\in\{1,\ldots,m\}}$ respectively in $\mathcal{K}$ and $\mathbb{H}^1_{loc} \times \mathbb{R}$, by continuity of $f_a(.)$ for the weak topology in $\mathbb{H}^1$, by continuity and boundedness of Θ_R in $C([0,T];\mathbb{R})$, by continuity of F from $\mathbb{H}^1$ to $\mathbb{H}^{-1}$ and the bounds on $F\left(X_\epsilon^R(t)\right)$ obtained thanks to Lemma 3.4.1, the function

$$\left\langle a, \mathbf{N}^R\left(Z_\epsilon^R(t)\right)\right\rangle \chi_{R_0}\left(Z_\epsilon^R\right) \prod_{j=1}^m h_j\left(Z_\epsilon^R(t_j)\right)$$

is a bounded and continuous function of Z_ϵ^R from $\mathcal{K}$ into $\mathbb{R}$. We deduce by convergence in law of Z_ϵ^R to $\widehat{Z}^R$ in $\mathcal{K}$, since the test function a is compactly supported, that for all $t, s \in [0,T], t \geqslant s$

$$\mathbb{E}\left(\left\langle a, \mathbf{N}^R\left(\widehat{Z}^R(t)\right) - \mathbf{N}^R\left(\widehat{Z}^R(s)\right)\right\rangle \chi_{R_0}\left(\widehat{Z}^R\right) \prod_{j=1}^m h_j(\widehat{Z}^R(t_j))\right) = 0. \qquad (3.4.28)$$

Since, almost surely, X_ϵ^R belongs to the closed ball $\mathcal{B}_{\mathcal{K}}(R_0)$, we deduce that almost surely $\widehat{X}^R \in \mathcal{B}_{\mathcal{K}}(R_0)$. Thus we conclude from (3.4.28) that $\left\langle a, \mathbf{N}^R\left(\widehat{Z}^R(.)\right)\right\rangle$

is a continuous square integrable martingale with respect to the filtration $\mathcal{G}_t = \sigma\left(\widehat{Z}^R(s), s \leqslant t\right)$ and this holds for any $a \in \mathbb{H}^1$ with compact support.

In order to identify the equation satisfied by $\widehat{X}^R$, we consider, for $a, b \in \mathbb{H}^1$ with compact support, the function $g_{a,b}(v) = f_a(v) f_b(v) \in \mathscr{D}^R$ and the perturbed test function $g_{a,b,\epsilon}$

$$g_{a,b,\epsilon}(v,y) = g_{a,b}(v) + \epsilon g^1_{a,b}(v,y) + \epsilon^2 g^2_{a,b}(v,y),$$

obtained thanks to Proposition 3.4.1. Thus functions $g^1_{a,b}(v,y)$ and $g^2_{a,b}(v,y)$ are chosen to be solution of the Poisson Equations (3.4.8) and (3.4.12) for $g_{a,b}$. Let us now define the real valued continuous martingale

$$\mathbf{H}^R_{a,b,\epsilon}\left(Z^R_\epsilon(t)\right) = g_{a,b,\epsilon}(X^R_\epsilon(t), \nu_\epsilon(t)) - g_{a,b,\epsilon}(v,y) - \int_0^t \mathscr{L}^R_\epsilon g_{a,b,\epsilon}\left(X^R_\epsilon(s), \nu_\epsilon(s)\right) ds.$$

Using the same arguments as before, we may prove that

$$\begin{aligned}
&\lim_{\epsilon \to 0} \mathbb{E}\left(\left(\mathbf{H}^R_{a,b,\epsilon}\left(Z^R_\epsilon(t)\right) - \mathbf{H}^R_{a,b,\epsilon}\left(Z^R_\epsilon(s)\right)\right) \chi_{R_0}\left(Z^R_\epsilon\right) \prod_{j=1}^m h_j(Z^R_\epsilon(t_j))\right) \\
&\quad = \quad \mathbb{E}\left(\left(\mathbf{H}^R_{a,b}\left(\widehat{Z}^R(t)\right) - \mathbf{H}^R_{a,b}\left(\widehat{Z}^R(s)\right)\right) \chi_{R_0}\left(\widehat{Z}^R\right) \prod_{j=1}^m h_j(\widehat{Z}^R(t_j))\right),
\end{aligned}$$

where

$$\mathbf{H}^R_{a,b}\left(\widehat{Z}(t)\right) = g_{a,b}(\widehat{X}^R(t)) - g_{a,b}(v) - \int_0^t \mathscr{L}^R_{\gamma^R} g_{a,b}\left(\widehat{X}^R(s)\right) ds.$$

From the above convergence and the martingale property of $\mathbf{H}^R_{a,b,\epsilon}\left(Z^R_\epsilon(t)\right)$, we deduce that $\mathbf{H}^R_{a,b}\left(\widehat{Z}^R(.)\right)$ is a continuous real valued martingale. A classical computation then shows that the quadratic variation of the martingale $\mathbf{N}^R\left(\widehat{Z}^R(t)\right)$ defined in (3.4.27) is given by

$$\begin{aligned}
&\left\langle b, \ll \mathbf{N}^R\left(\widehat{Z}^R(t)\right) \gg a \right\rangle \\
&= \int_0^t \mathscr{L}^R_{\gamma^R}\left(f_a\left(\widehat{Z}^R(s)\right) f_b\left(\widehat{Z}^R(s)\right)\right) - f_a\left(\widehat{Z}^R(s)\right) \mathscr{L}^R_{\gamma^R} f_b\left(\widehat{Z}^R(s)\right) \\
&- f_b\left(\widehat{Z}^R(s)\right) \mathscr{L}^R_{\gamma^R} f_a\left(\widehat{Z}^R(s)\right) ds.
\end{aligned}$$

Applying the operator $\mathscr{L}^R_{\gamma^R}$ respectively to the test functions f_a and $g_{a,b}$, we obtain that

$$\mathscr{L}^R_{\gamma^R} f_a\left(\widehat{Z}^R(t)\right) = \left\langle a, \left(\frac{id_0}{2} + \frac{3\gamma}{2}\right) \frac{\partial^2 \widehat{X}^R}{\partial x^2} + i\Theta_R\left(\left\|\gamma^R(t)\right\|^2_{\mathbb{H}^1}\right) F\left(\widehat{X}^R\right)\right\rangle,$$

and

$$\begin{aligned}
&\mathscr{L}^R_{\gamma^R} g_{a,b}\left(\widehat{Z}^R(t)\right)\\
= \;& f_b\left(\widehat{X}^R(t)\right)\left\langle a, \left(\frac{id_0}{2}+\frac{3\gamma}{2}\right)\frac{\partial^2 \widehat{X}^R(t)}{\partial x^2} + i\Theta_R\left(\gamma^R(t)\right) F\left(\widehat{X}^R(t)\right)\right\rangle\\
&+ f_a\left(\widehat{X}^R(t)\right)\left\langle b, \left(\frac{id_0}{2}+\frac{3\gamma}{2}\right)\frac{\partial^2 \widehat{X}^R(t)}{\partial x^2} + i\Theta_R\left(\gamma^R(t)\right) F\left(\widehat{X}^R(t)\right)\right\rangle\\
&+\gamma\sum_{k=1}^{3}\left\langle a, \sigma_k\frac{\partial \widehat{X}^R(t)}{\partial x}\right\rangle\left\langle b, \sigma_k\frac{\partial \widehat{X}^R(t)}{\partial x}\right\rangle.
\end{aligned}$$

We deduce that the quadratic variation is given by formula (3.3.5) wiht $\widetilde{X}$ replaced by $\widehat{X}^R$. Thus, using the martingale representation theorem, we can write the $\mathcal{G}_t$-martingale $\mathbf{N}^R\left(\widehat{Z}^R(t)\right)$ as the stochastic integral

$$\left\langle a, \mathbf{N}^R\left(\widehat{Z}^R(t)\right)\right\rangle = \sqrt{\gamma}\int_0^t \sum_{k=1}^{3}\left\langle a, \sigma_k\frac{\partial \widehat{X}^R(s)}{\partial x}\right\rangle dW_k(s),$$

where $W=(W_1,W_2,W_3)$ is a real valued Brownian motion on a possibly enlarged space $(\Omega,\mathcal{G},\mathcal{G}_t,\mathbb{P})$. We deduce that $\left(\widehat{X}^R, W\right)$ is a weak solution in $C\left([0,T];\mathbb{H}^1_{loc}\right)\cap C_w\left([0,T];\mathbb{H}^1\right)\cap L^\infty_w\left(0,T;\mathbb{H}^2\right)$ of the equation

$$\begin{aligned}
id\widehat{X}^R(t) + &\left(\frac{d_0}{2}\partial_x^2\widehat{X}^R(t) + \Theta_R\left(\gamma^R(t)\right)F\Big(\widehat{X}^R(t)\Big)\right)dt\\
&+ i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k\partial_x\widehat{X}^R(t)\circ dW_k(t) = 0, \qquad (3.4.29)
\end{aligned}$$

with initial condition $X_0 = v\in\mathbb{H}^3$. The next step consists in proving that almost surely $\gamma^R(t) = \left\|\widehat{X}^R(t)\right\|^2_{\mathbb{H}^1}$. Using the Skorokhod representation theorem, we can construct new random variables (that we still denote Z^R_ϵ, $\widehat{Z}^R$) on a new common probability space $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbb{P})$ with respectively $\mathcal{L}\left(Z^R_\epsilon\right)$ and $\mathcal{L}\left(\widehat{Z}^R\right)$ as probability measure and with values in $\mathcal{K}$ such that

$$\lim_{\epsilon\to 0} Z^R_\epsilon = \widehat{Z}^R \quad \mathbb{P}\ a.s \text{ in } \mathcal{K}.$$

Since $\widehat{X}^R\in L^\infty\left(0,T;\mathbb{H}^2\right)$, we deduce using Equation (3.4.29) that $\widehat{X}^R\in C\left([0,T];\mathbb{L}^2\right)$. Hence applying Itô formula, it is easy to see, since Θ_R is a real valued function, that almost surely

$$\left\|\widehat{X}^R(t)\right\|_{\mathbb{L}^2} = \|v\|_{\mathbb{L}^2} = \left\|X^R_\epsilon(t)\right\|_{\mathbb{L}^2}, \qquad \forall t\in[0,T],\ \forall\epsilon>0.$$

Thus we deduce the strong convergence of $X^R_\epsilon(t)$ to $\widehat{X}^R(t)$ in $\mathbb{L}^2$, a.s for each $t\in[0,T]$. Since X^R_ϵ converges to $\widehat{X}^R$ in $L^\infty_w\left(0,T;\mathbb{H}^2\right)$, we get using Lemma 3.4.1 that

$$\left\|\widehat{X}^R\right\|_{L^\infty(0,T;\mathbb{H}^2)} \leqslant \liminf_{\epsilon\to 0}\left\|X^R_\epsilon\right\|_{L^\infty(0,T;\mathbb{H}^2)} \leqslant C(R,T) \quad \mathbb{P}-a.s.$$

Interpolating $\mathbb{H}^1$ between $\mathbb{L}^2$ and $\mathbb{H}^2$, we conclude that

$$\lim_{\epsilon\to 0}\left\|X_\epsilon^R(t)-\widehat{X}^R(t)\right\|_{\mathbb{H}^1}=0,\quad \forall t\in[0,T],\ \mathbb{P}-a.s, \tag{3.4.30}$$

and $\widehat{X}^R\in C\left([0,T];\mathbb{H}^1\right)$; it follows that, almost surely for all t in $[0,T]$, $\gamma^R(t)=\left\|\widehat{X}^R(t)\right\|^2_{\mathbb{H}^1}$ and $\widehat{X}^R$ is a solution of (3.3.6). Thus the limit in law of X_ϵ^R is unique and is given by the solution X^R of Equation (3.3.6).

The final step consists in recovering the convergence in law in $C\left([0,T],\mathbb{H}^1\right)$. Since Y_ϵ^R is uniformly bounded in ϵ in $C^\alpha\left([0,T],\mathbb{H}^{-1}\right)\cap C\left([0,T];\mathbb{H}^2\right)$ with $0\leqslant\alpha<1/2$, we deduce that it is a.s uniformly bounded in ϵ in $C^\beta\left([0,T],\mathbb{H}^1\right)$ with $\beta=\alpha/3$. Moreover using pointwise convergence (3.4.30), expression (3.4.22) and uniform bounds (3.4.1), we get pointwise convergence in $\mathbb{H}^1$ of Y_ϵ^R to X^R. We conclude that Y_ϵ^R converges in law to X^R in $C\left([0,T],\mathbb{H}^1(\mathbb{R})\right)$ and by Lemma 3.4.4, the convergence in law of X_ϵ^R to X^R in $C\left([0,T],\mathbb{H}^1(\mathbb{R})\right)$ follows. □

Remark 3.4.4. *Using Arzela-Ascoli and Banach-Alaoglu theorems, Lemma 3.4.5 and Tychonov Theorem, we deduce that $\left(\mathcal{L}\left(X_\epsilon^R\right)\right)_{R\in\mathbb{N}}$ is tight on $\mathcal{K}^{\mathbb{N}}$. Thus the same arguments as above lead to the convergence in law of $\left(X_\epsilon^R\right)_{R\in\mathbb{N}}$ to $\left(X^R\right)_{R\in\mathbb{N}}$ (see [19]).*

3.4.5 Convergence of $(X_\epsilon)_{\epsilon>0}$ to X

Using the Skorokhod Theorem, we can construct new random variables $\widetilde{X}_\epsilon^R$, $\widetilde{X}^R$ on a common probability space $\left(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathcal{F}_t},\widetilde{\mathbb{P}}\right)$ and with values in $C\left([0,T],\mathbb{H}^1\right)$ such that for any $R>0$,

$$\begin{cases}\widetilde{\mu}_\epsilon^R=\mu_\epsilon^R\\ \widetilde{\mu}^R=\mu^R\end{cases}\qquad\text{and } \widetilde{X}_\epsilon^R\underset{\epsilon\to 0}{\longrightarrow}\widetilde{X}^R\quad\widetilde{\mathbb{P}}\ a.s \text{ in } C\left([0,T],\mathbb{H}^1\right).$$

We define the escape times $\widetilde{\tau}^R$ and $\widetilde{\tau}_\epsilon^R$ associated to the cut-off :

$$\widetilde{\tau}^R=\inf\left\{t\in[0,T],\left\|\widetilde{X}^R(t)\right\|_{\mathbb{H}^1}>R\right\}\quad\text{and}\quad\widetilde{\tau}_\epsilon^R=\inf\left\{t\in[0,T],\left\|\widetilde{X}_\epsilon^R(t)\right\|_{\mathbb{H}^1}>R\right\}.$$

Let $\widetilde{X}_\epsilon$ and $\widetilde{X}$ be the processes, with values in $\mathcal{E}\left(\mathbb{H}^1\right)$, defined respectively by $\widetilde{X}_\epsilon(t)=\widetilde{X}_\epsilon^R(t)$ for $t<\widetilde{\tau}_\epsilon^R$ and $\widetilde{X}(t)=\widetilde{X}^R(t)$ for $t<\widetilde{\tau}^R$, $\widetilde{X}(t)=\Delta$ for $t\geqslant\tau^*=\lim_{R\to+\infty}\widetilde{\tau}^R$. Then if $\tau<\tau^*$ a.s is a stopping time, the process $\widetilde{X}_\epsilon$ converges to $\widetilde{X}$ a.s in $C\left([0,\tau],\mathbb{H}^1(\mathbb{R})\right)$. Hence the convergence in law in $\mathcal{E}\left(\mathbb{H}^1\right)$ follows. □

3.5 Study of the driving process ν

In this appendix, we briefly explain the set of hypotheses on the driving process α stated in Theorem 3.1.4. We also recall some results obtained in [43, 73] about the driving process ν.

As it is classical in the homogenization theory, the theory of Diffusion-Approximation uses the method of correctors (known as the perturbed test functions method). These correctors are defined as a solution of Poisson equations. The hypotheses stated in Theorem 3.1.4 on the process α garantee the existence of a unique (up to an additive constant) solution to this equation. These hypotheses are usual in Diffusion-Approximation theory [38, 41, 43, 73]. However, some of them may be relaxed and in particular similar results may be proved for ergodic diffusion in $\mathbb{R}^d$ [90, 91, 92].

Let α be a homogeneous Markov process, with separable state space S and transition functions P_t. For any bounded measurable function f on S, we define the operator T_t [97, 98]

$$T_t f(x) = \mathbb{E}\left(f(\alpha(t)) \,|\alpha_0 = x\right).$$

The Chapman-Kolmogorov equation, satisfied by the transition functions P_t, immediately implies that the bounded operator T_t has the algebraic semigroup property $T_t T_s = T_{t+s}$. A Markov process is said to be a Feller process if $(T_t)_{t\geqslant 0}$ is a strongly continuous contraction semigroup from the space of bounded continuous functions $C_b(S)$ into itself. Moreover a Feller process has a modification such that its sample paths are cad-lag (right continous with left hand limits) [34, 69]. This allows to define the infinitesimal generator $\mathscr{L}_\alpha$ from its domain $\mathscr{D}$ into $C_b(S)$ as the limit in $C_b(S)$ of

$$\mathscr{L}_\alpha f = \lim_{t\to 0^+} \frac{T_t f - f}{t}.$$

Poisson processes and diffusion processes with lipschitz coefficients are example of Feller processes. In the latter case the infinitesimal generator is a second order partial differential operator. From the Chapman-kolmogorov equation, the finite-dimensional distributions of a Feller process are entirely determined by its initial distribution ν and the semigroup $(T_t)_{t\geqslant 0}$ [34]. Hence, they are in turn determined by its generator $(\mathscr{L}_\alpha, \mathscr{D})$ or by a sufficiently large subset of $\mathscr{D}$ [34]. The easiest way to caracterize a Feller process is thought the martingale problem introduced by Stroock and Varadhan [100].

The Fredholm alternative is the statement that either $\mathscr{L}_\alpha u = -f$ admits a unique solution or has a solution if and only if f satisfies the null mass condition $\mathbb{E}_\mu(f(\alpha)) = 0$, where μ is the unique invariant probability measure of α. This solution is not unique because the operator $\mathscr{L}_\alpha$ is not injective since its null space contains constant functions. Actually, homogeneous boundary conditions are missing to ensure the injectivity of this operator. The null mass condition is necessary. Indeed, consider for exemple that S is a subset of the euclidian space and that the invariant probability measure μ is absolutely continuous with respect to the Lebesgue measure with density probability function p. Then, by integration by parts and by definition of an invariant measure

$$\mathbb{E}_\mu(\mathscr{L}_\alpha u) = \int_S \mathscr{L}_\alpha u(x) p(x) dx = \int_S u(x) \mathscr{L}_\alpha^* p(x) dx = 0.$$

Similar condition is obtained for the Laplace problem with homogeneous Neuman boundary conditions.

A Feller process is ergodic if and only if its T_t-invariant functions are constants, i.e. $T_t f = f$ is equivalent to say that f is constant [69]. Therefore, ergodicity insures that the solution of the homogeneous equation $\mathscr{L}_\alpha u = 0$ has only constant solutions. Ergodicity also states that the process forgets its initial distribution as t goes to infinity

$$T_t f(x) \xrightarrow[t\to+\infty]{} \int_S f(y)\mu(dy).$$

Thus, a particular solution of the inhomogeneous Poisson equation $\mathscr{L}_\alpha u = -f$ is given by

$$u(x) = \int_0^\infty T_s f(x) ds < +\infty.$$

For Feller processes, ergodicity is related to the notion of recurrence and transience [69]. In some case, when the state space is compact, it is possible to check that the Fredholm alternative is satisfied. The next result was used in [43, 73] to prove Proposition 3.5.1.

Theorem 3.5.1. *Assume that a Feller process α, with compact state space S, has transition functions P_t that are absolutely continuous with respect to a borel measure on S. Then, it is recurrent (in the sense of Harris) and has a unique invariant probability measure μ under which the process is ergodic. Moreover, if $f \in L^\infty(S)$ satisfies $\mathbb{E}_\mu(f(\alpha)) = 0$, then*

$$u(x) = \int_0^\infty T_s f(x) ds < +\infty$$

and u is a continuous bounded function on S. If u belongs to $\mathscr{D}$, the domain of $\mathscr{L}_\alpha$, then u is a solution of the Poisson equation $\mathscr{L}_\alpha u = -f$.

Let us now recall the results obtained in [43, 73] about the driving process ν.

Proposition 3.5.1. *The process $\nu = (\nu_1, \nu_2)^t$ is a Feller process that evolves on the unit sphere $\mathbb{S}^3$ of $\mathbb{C}^2 \sim \mathbb{R}^4$. Furthermore it admits a unique invariant measure Λ, which is the uniform measure on $\mathbb{S}^3$, under which it is ergodic. For all $f \in C_b^2(\mathbb{S}^3)$ satisfying the null mass condition $\mathbb{E}_\Lambda(f(\nu)) = \int_{\mathbb{S}^3} f(y)\Lambda(dy) = 0$, the Poisson equation $\mathscr{L}_\nu u(y) + f(y) = 0$ admits a unique solution of class $C_b^2(\mathbb{S}^3)$, up to a constant, which can be written as $u(y) = \int_0^{+\infty} \mathbb{E}[f(\nu(t))|\nu_0 = y]\, dt$.*

Let us recall that $\boldsymbol{\sigma}(\nu(t)) = \sigma_1 m_1 + \sigma_2 m_2 + \sigma_3 m_3$ where $m_j(t) = g_j(\nu(t))$. We now state a result related to the effect of the random PMD on the pulse evolution.

Corollary 3.5.1.

1. *The process $m = (m_1, m_2, m_3) \in \mathbb{S}^3$ is a Feller process with a unique invariant measure $\Lambda \circ g^{-1}$ under which it is ergodic.*

2. *For $j = 1, 2, 3$: $\mathbb{E}_\Lambda(g_j(\nu)) = \mathbb{E}_{\Lambda \circ g_j^{-1}}(m) = 0$ and $\mathbb{E}_\Lambda(g_j(\nu(t))g_k(\nu(t))) = \delta_{jk}/3$. As a consequence,*

$$\mathbb{E}_\Lambda(N_{1,\nu}(X)) = \frac{2}{3}\left(2|X_2|^2 - |X_1|^2\right)X_1, \quad \mathbb{E}_\Lambda(N_{2,\nu}(X)) = \frac{2}{3}\left(2|X_1|^2 - |X_2|^2\right)X_2.$$

3. For $j,k=1,2,3$:

$$\int_0^{+\infty}\mathbb{E}_\Lambda\left[g_j\left(\nu(0)\right)g_k\left(\nu(t)\right)\right]dt=\begin{cases}\dfrac{1}{12\gamma_c} & \text{if } j=k\\ 0 & \text{if } j\neq k,\end{cases}$$

where γ_c is the constant appearing in (3.1.5).

3.6 Proof of technical Lemmas

Proof of Lemma 3.4.2. Let v be in $\mathbb{H}^3$. Using the explicit representation (3.4.9) of f^1, we obtain, since $D_vf(v)\in\mathbb{H}^1(\mathbb{R})$, that

$$\begin{aligned}\left|f^1(v,y)\right| &= \left|\left\langle D_vf(v),b'\widetilde{\boldsymbol{\sigma}}(y)\frac{\partial v}{\partial x}\right\rangle\right|\\ &\leqslant b'\left\|D_vf(v)\right\|_{\mathbb{H}^1}\left\|\frac{\partial v}{\partial x}\right\|_{\mathbb{H}^{-1}}\sum_{j=1}^{3}\left|\int_0^{+\infty}\mathbb{E}\left(g_j\left(\nu(t)\right)|\,\nu(0)=y\right)dt\right|.\end{aligned}$$

Moreover by Proposition 3.5.1 the integral $\int_0^{+\infty}\mathbb{E}\left(g_j\left(\nu(t)\right)|\,\nu(0)=y\right)dt$ converges because g_j is a bounded function of $\nu\in\mathbb{S}^3$. Since $v\mapsto D_vf(v)$ is a continuous function which is bounded on bounded sets of $\mathbb{H}^{-1}$, we deduce that :

$$\sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}}\left|f_\epsilon^1(v,y)\right|\leqslant b'C(K).$$

The function f^2 given by (3.4.15) may be bounded using the same arguments. Indeed

$$\left\langle D_vf(v),i\Theta_R\left(\|v\|_{\mathbb{H}^1}^2\right)\widetilde{F}(v,y)\right\rangle\leqslant\left\|D_vf(v)\right\|_{\mathbb{H}^1}\left\|\widetilde{F}(v,y)\right\|_{\mathbb{H}^{-1}}.$$

Since for all $v\in\mathbb{H}^3$, $y\mapsto F_y(v)-F(v)$ is a function of class C_b^2 on $\mathbb{S}^3$, with values in $\mathbb{H}^{-1}$, satisfying the null mass condition of Proposition 3.5.1, the term $\widetilde{F}(v,y)$ is bounded. Moreover $v\mapsto F_y(v)-F(v)$ is bounded in $\mathbb{H}^{-1}$ on bounded sets of $\mathbb{H}^1$ by the continuous embeddings $\mathbb{H}^1(\mathbb{R})\hookrightarrow\mathbb{L}^4(\mathbb{R})$ and $\mathbb{L}^{4/3}(\mathbb{R})\hookrightarrow\mathbb{H}^{-1}(\mathbb{R})$. In addition

$$\begin{aligned}&\left|(b')^2\sum_{k,l=1}^{3}\left\langle D_v^2f(v)\sigma_k\frac{\partial v}{\partial x},\sigma_l\frac{\partial v}{\partial x}\right\rangle\widetilde{\widetilde{g}}_{k,l}(y)+\left\langle D_vf(v),(b')^2\,\widetilde{\widetilde{\boldsymbol{\sigma}}}(y)\frac{\partial^2v}{\partial x^2}\right\rangle\right|\\ &\leqslant C\sum_{k,l=1}^{3}\left(\left|\left\langle D_v^2f(v)\sigma_k\frac{\partial v}{\partial x},\sigma_l\frac{\partial v}{\partial x}\right\rangle\right|+\left|\left\langle D_vf(v),\sigma_k\sigma_l\frac{\partial^2v}{\partial x^2}\right\rangle\right|\right)\\ &\leqslant C\left(\left\|D_v^2f(v)\right\|_{\mathscr{L}(\mathbb{H}^{-1},\mathbb{H}^1)}\left\|\frac{\partial v}{\partial x}\right\|_{\mathbb{H}^{-1}}^2+\left\|D_vf(v)\right\|_{\mathbb{H}^1}\left\|\frac{\partial^2v}{\partial x^2}\right\|_{\mathbb{H}^{-1}}\right).\end{aligned}$$

Since $v\mapsto D_vf(v)$ and $v\mapsto D_v^2f(v)$ are bounded on bounded sets of $\mathbb{H}^{-1}(\mathbb{R})$, we conclude the proof of the lemma. □

Proof of Lemma 3.4.3. Replacing f^1 by its expression (3.4.9) we get

$$\begin{aligned}
&\left\langle D_v f^1(v,y), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right) F_y(v)\right\rangle \\
&= -\left\langle D_v^2 f(v) b'\widetilde{\boldsymbol{\sigma}}(y)\frac{\partial v}{\partial x}, \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right) F_y(v)\right\rangle \\
&\quad -\left\langle D_v f(v), b'\widetilde{\boldsymbol{\sigma}}(y)\frac{id_0}{2}\frac{\partial^3 v}{\partial x^3} + ib'\widetilde{\boldsymbol{\sigma}}(y)\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right)\partial_x F_y(v)\right\rangle.
\end{aligned}$$

By the assumptions on f, $v \mapsto D_v f(v)$ and $v \mapsto D_v^2 f(v)$ are continuous bounded functions on bounded sets of $\mathbb{H}^{-1}(\mathbb{R})$. Moreover $D_v^2 f(v) \in \mathcal{L}(\mathbb{H}^{-1}, \mathbb{H}^1)$, $D_v f(v) \in \mathbb{H}^1$ and $\frac{\partial^3 v}{\partial x^3} \in \mathbb{L}^2$. Using the bound (3.4.11), we deduce that

$$\sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}}\left|\left\langle D_v f^1(v,y), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right) F_y(v)\right\rangle\right| \leqslant C(K).$$

Let us now compute the first derivative of f^2 using expression (3.4.15); for all h in $\mathbb{H}^1$ and v in $\mathbb{H}^3$:

$$\begin{aligned}
&\left\langle D_v f^2(v,y), h\right\rangle \\
&= \left\langle D_v^2 f(v)h, i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right)\widetilde{F}(v,y)\right\rangle + \left\langle D_v f(v), 2i\Theta_R'\left(\|v\|^2_{\mathbb{H}^1}\right)(v,h)_{\mathbb{H}^1}\widetilde{F}(v,y)\right\rangle \\
&+ \left\langle D_v f(v), i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right) D_v\widetilde{F}(v,y).h\right\rangle - (b')^2\sum_{k,l=1}^3 D_v^3 f(v).\left(\sigma_k\frac{\partial v}{\partial x}, \sigma_l\frac{\partial v}{\partial x}, h\right)\widetilde{\widetilde{g}}_{k,l}(y) \\
&- 2(b')^2\sum_{k,l=1}^3\left\langle D_v^2 f(v)\sigma_k\frac{\partial h}{\partial x}, \sigma_l\frac{\partial v}{\partial x}\right\rangle\widetilde{\widetilde{g}}_{k,l}(y) \\
&- \left\langle D_v^2 f(v)h, (b')^2\widetilde{\widetilde{\boldsymbol{\sigma}}}(y)\frac{\partial^2 v}{\partial x^2}\right\rangle - \left\langle D_v f(v), (b')^2\widetilde{\widetilde{\boldsymbol{\sigma}}}(y)\frac{\partial^2 h}{\partial x^2}\right\rangle.
\end{aligned}$$

Taking respectively $h = \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right)F_y(v)$ and $h = b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}$, we conclude

$$\sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}}\left|\left\langle D_v f^2(v,y), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} + i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right)F_y(v)\right\rangle\right| \leqslant C(K),$$

and

$$\sup_{\substack{v\in\mathcal{B}(K)\\ y\in\mathbb{S}^3}}\left|\left\langle D_v f^2(v,y), b'\boldsymbol{\sigma}(y)\frac{\partial v}{\partial x}\right\rangle\right| \leqslant C(K),$$

since $v \mapsto D_v^3 f(v)$ is bounded on bounded set of $\mathbb{H}^{-1}(\mathbb{R})$ with values in $\mathcal{L}_3(\mathbb{H}^{-1}, \mathbb{R})$ and $\frac{\partial^4 v}{\partial x^4} \in \mathbb{H}^{-1}$. □

Proof of lemma 3.4.6. Let us recall that the family $\{e_i\}_{i\in\mathbb{N}^*}$ denotes a complete orthonormal system of $\mathbb{H}^1$ constructed from a complete orthonormal system $\{\widetilde{e_i}\}_{i\in\mathbb{N}^*}$ in $\mathbb{L}^2$ and $\langle .,. \rangle$ is the duality product between $\mathbb{H}^1 - \mathbb{H}^{-1}$. Then

$$\left\|Y_\epsilon^R(t) - Y_\epsilon^R(s)\right\|^4_{\mathbb{H}^{-1}} = \left\{\sum_{i=1}^{+\infty}\left\langle e_i, Y_\epsilon^R(t) - Y_\epsilon^R(s)\right\rangle^2\right\}^2.$$

Using twice the Young inequality and the expression of Y_ϵ^R given by (3.4.23) and (3.4.21), we obtain :

$$
\begin{aligned}
&\left\|Y_\epsilon^R(t)-Y_\epsilon^R(s)\right\|_{\mathbb{H}^{-1}}^4 \\
&\leqslant C\left\|\frac{d_0}{2}\int_s^t \partial_x^2 X_\epsilon^R(t')dt'\right\|_{\mathbb{H}^{-1}}^4 + C\left\|\int_s^t \Theta_R\left(\left\|X_\epsilon^R(t')\right\|_{\mathbb{H}^1}^2\right) F_{\nu_\epsilon(t')}\left(X_\epsilon^R(t')\right)dt'\right\|_{\mathbb{H}^{-1}}^4 \\
&+ C\left\|\int_s^t (b')^2\,\widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(t')\right)\boldsymbol{\sigma}\left(\nu_\epsilon(t')\right)\partial_x^2 X_\epsilon^R(t')dt'\right\|_{\mathbb{H}^{-1}}^4 + C\left[\sum_{i=1}^{+\infty}\left\langle e_i, M_\epsilon^R(t)-M_\epsilon^R(s)\right\rangle^2\right]^2 \\
&+ C\epsilon^4\left\|\int_s^t b'\widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(t')\right)\partial_x\left(\frac{d_0}{2}\partial_x^2 X_\epsilon^R(t') + \Theta_R\left(\left\|X_\epsilon(t')\right\|_{\mathbb{H}^1}^2\right) F_{\nu_\epsilon(t')}\left(X_\epsilon^R(t')\right)\right)dt'\right\|_{\mathbb{H}^{-1}}^4.
\end{aligned}
$$

We bound each terms separately. Using Lemma 3.4.1,

$$
\left\|\int_s^t \frac{d_0}{2}\frac{\partial^2 X_\epsilon^R(t')}{\partial x^2}dt'\right\|_{\mathbb{H}^{-1}}^4 \leqslant C(R,T)(t-s)^4.
$$

Using that F is cubic and Lemma 3.4.1,

$$
\left\|\int_s^t \Theta_R\left(\left\|X_\epsilon(t')\right\|_{\mathbb{H}^1}^2\right) F_{\nu_\epsilon(t')}\left(X_\epsilon^R(t')\right)dt'\right\|_{\mathbb{H}^{-1}}^4 \leqslant C(R,T)(t-s)^4,
$$

and using, Lemma 3.4.1 and the bound (3.4.11)

$$
\left\|\int_s^t (b')^2\,\widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(t')\right)\boldsymbol{\sigma}\left(\nu_\epsilon(t')\right)\frac{\partial^2 X_\epsilon^R(t')}{\partial x^2}dt'\right\|_{\mathbb{H}^{-1}}^4 \leqslant C(R,T)(t-s)^4.
$$

Finally, we bound the ϵ^4 term that is well defined because X_ϵ^R has values in $\mathbb{H}^3$. Using the Cauchy Schwarz inequality, Lemma 3.4.1 and the bound (3.4.11), we get for all $\epsilon < 1$:

$$
\begin{aligned}
&\epsilon^4\left\|\int_s^t b'\widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(t')\right)\frac{\partial}{\partial x}\left(\frac{d_0}{2}\frac{\partial^2 X_\epsilon^R(t')}{\partial x^2} + \Theta_R\left(\left\|X_\epsilon\right\|_{\mathbb{H}^1}^2\right) F_{\nu_\epsilon(t')}\left(X_\epsilon^R(t')\right)\right)dt'\right\|_{\mathbb{H}^{-1}}^4 \\
&\quad\leqslant \epsilon^4(b')^4M^4\left\|\int_s^t \frac{d_0}{2}\frac{\partial^3 X_\epsilon^R(t')}{\partial x^3} + \frac{\partial}{\partial x}\left(\Theta_R\left(\left\|X_\epsilon\right\|_{\mathbb{H}^1}^2\right) F_{\nu_\epsilon(t')}\left(X_\epsilon^R(t')\right)\right)dt'\right\|_{\mathbb{H}^{-1}}^4 \\
&\quad\leqslant C(R,T)(t-s)^4.
\end{aligned}
$$

Taking the expectation and adding the previous estimates, we deduce that :

$$
\begin{aligned}
&\mathbb{E}\left(\left\|Y_\epsilon^R(t)-Y_\epsilon^R(s)\right\|_{\mathbb{H}^{-1}}^4\right) \\
&\quad\leqslant C(R,T)\,(t-s)^4 + C\mathbb{E}\left(\left[\sum_{j\in\mathbb{N}^*}\left\langle e_j, M_\epsilon^R(t)-M_\epsilon^R(s)\right\rangle^2\right]^2\right).
\end{aligned}
$$

In order to prove a uniform bound, with respect to ϵ, of the second term, we will use the Burkholder-Davis-Gundy inequality and consequently we have to compute the quadratic variation $\ll M_\epsilon^R(t) \gg$ of $M_\epsilon^R(t)$ defined, for all $j \in \mathbb{N}^*$, by

$$
\left\langle e_j, M_\epsilon^R(t)\right\rangle = f_{j,\epsilon}\left(X_\epsilon^R(t),\nu_\epsilon(t)\right) - f_{j,\epsilon}\left(v,y\right) - \int_0^t \mathscr{L}_\epsilon^R f_{j,\epsilon}\left(X_\epsilon^R(s),\nu_\epsilon(s)\right)ds,
$$

where $\mathscr{L}_\epsilon^R f_{j,\epsilon}\left(X_\epsilon^R(s),\nu_\epsilon(s)\right)$ is given by (3.4.21). The next Lemma states that the process
$\ll M_\epsilon^R(t) \gg$ can be expressed only in terms of the infinitesimal generator $\mathscr{L}_\nu$ of the Markov process ν.

Lemma 3.6.1. *For all j in $\mathbb{N}^*$*

$$\left\langle e_j, \ll M_\epsilon^R(t) \gg e_j \right\rangle = (b')^2 \int_0^t \mathscr{L}_\nu \left(\left\langle e_j, \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right) \frac{\partial X_\epsilon^R}{\partial x}(s) \right\rangle^2 \right) ds$$
$$- 2\,(b')^2 \int_0^t \left\langle e_j, \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right) \frac{\partial X_\epsilon^R}{\partial x}(s) \right\rangle \left\langle e_j, \mathscr{L}_\nu \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right) \frac{\partial X_\epsilon^R}{\partial x}(s) \right\rangle ds.$$

Thus using the Burkhölder-Davis-Gundy inequality

$$\mathbb{E}\left(\left(\sum_{j=1}^{+\infty} \left\langle e_j, M_\epsilon^R(t) - M_\epsilon^R(s) \right\rangle^2 \right)^2 \right) \leqslant C(R,T)\,|t-s|^2,$$

thanks to Lemma 3.4.1 and Proposition 3.5.1. Adding the previous estimates,

$$\mathbb{E}\Big(\left\| Y_\epsilon^R(t) - Y_\epsilon^R(s) \right\|_{\mathbb{H}^{-1}}^4 \Big) \leqslant C\,(R,T)\,|t-s|^2,$$

and Lemma 3.4.6 is proved. □

Proof of Lemma 3.6.1. A classical computation shows that for all $j \in \mathbb{N}^*$:

$$\left\langle e_j, \ll M_\epsilon^R(t) \gg e_j \right\rangle = \int_0^t \mathscr{L}_\epsilon^R \left(f_{j,\epsilon}\left(X_\epsilon^R(s),\nu_\epsilon(s)\right)^2 \right) ds$$
$$- 2 \int_0^t f_{j,\epsilon}\left(X_\epsilon^R(s),\nu_\epsilon(s)\right) \mathscr{L}_\epsilon^R f_{j,\epsilon}\left(X_\epsilon^R(s),\nu_\epsilon(s)\right) ds.$$

Now, for all j in $\mathbb{N}^*$,

$$\begin{aligned} \left(f_{j,\epsilon}\left(X_\epsilon^R(s),\nu_\epsilon(s)\right)\right)^2 &= \left\langle e_j, X_\epsilon^R(s) \right\rangle^2 - 2b'\epsilon \left\langle e_j, X_\epsilon^R(s) \right\rangle \left\langle e_j, \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right) \frac{\partial X_\epsilon^R}{\partial x}(s) \right\rangle \\ &\quad + (b')^2 \epsilon^2 \left\langle e_j, \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right) \frac{\partial X_\epsilon^R}{\partial x}(s) \right\rangle^2. \end{aligned}$$

Thus we get

$$
\begin{aligned}
&\mathscr{L}_\epsilon^R \left(f_{j,\epsilon}\left(X_\epsilon^R(s),\nu_\epsilon(s)\right)\right)^2 \\
&= 2\left\langle e_j, X_\epsilon^R(s)\right\rangle \left\langle e_j, \frac{id_0}{2}\frac{\partial^2 X_\epsilon^R(s)}{\partial x^2} + i\Theta_R\left(\left\|X_\epsilon^R(s)\right\|_{\mathbb{H}^1}^2\right) F_{\nu_\epsilon(s)}\left(X_\epsilon^R(s)\right)\right\rangle \\
&- 2b'\epsilon \left\langle e_j, \frac{id_0}{2}\frac{\partial^2 X_\epsilon^R(s)}{\partial x^2} + i\Theta_R\left(\left\|X_\epsilon^R(s)\right\|_{\mathbb{H}^1}^2\right) F_{\nu_\epsilon(s)}\left(X_\epsilon^R(s)\right)\right\rangle \left\langle e_j, \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right)\frac{\partial X_\epsilon^R}{\partial x}(s)\right\rangle \\
&- 2b'\epsilon \left\langle e_j, X_\epsilon^R(s)\right\rangle \left\langle e_j, \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right)\frac{\partial}{\partial x}\left(\frac{id_0}{2}\frac{\partial^2 X_\epsilon^R(s)}{\partial x^2} + i\Theta_R\left(\left\|X_\epsilon^R(s)\right\|_{\mathbb{H}^1}^2\right) F_{\nu_\epsilon(s)}\left(X_\epsilon^R(s)\right)\right)\right\rangle \\
&- 2\left(b'\right)^2 \left\langle e_j, \mathscr{L}_\nu\widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right)\frac{\partial X_\epsilon^R}{\partial x}(s)\right\rangle \left\langle e_j, \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right)\frac{\partial X_\epsilon^R}{\partial x}(s)\right\rangle \\
&+ \left(b'\right)^2 \mathscr{L}_\nu\left(\left\langle e_j, \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right)\frac{\partial X_\epsilon^R}{\partial x}(s)\right\rangle^2\right) \\
&+ 2b'\left\langle e_j, \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right)\frac{\partial}{\partial x}\left(b'\boldsymbol{\sigma}\left(\nu_\epsilon(s)\right)\frac{\partial X_\epsilon^R(s)}{\partial x}\right)\right\rangle \left\langle e_j, X_\epsilon^R(s)\right\rangle \\
&+ 2\left(b'\right)^2\epsilon^2 \left\langle e_j, \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right)\frac{\partial X_\epsilon^R}{\partial x}(s)\right\rangle \left\langle e_j, \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right)\frac{id_0}{2}\frac{\partial^3 X_\epsilon^R(s)}{\partial x^3}\right\rangle \\
&+ 2\left(b'\right)^2\epsilon^2 \left\langle e_j, \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right)\frac{\partial X_\epsilon^R}{\partial x}(s)\right\rangle \left\langle e_j, i\widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right)\Theta_R\left(\left\|X_\epsilon^R(s)\right\|_{\mathbb{H}^1}^2\right)\frac{\partial}{\partial x}F_{\nu_\epsilon(s)}\left(X_\epsilon^R(s)\right)\right\rangle \\
&- 2\left(b'\right)^2\epsilon \left\langle e_j, \widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right)\frac{\partial X_\epsilon^R}{\partial x}(s)\right\rangle \left\langle e_j, b'\widetilde{\boldsymbol{\sigma}}\left(\nu_\epsilon(s)\right)\boldsymbol{\sigma}\left(\nu_\epsilon(s)\right)\frac{\partial^2 X_\epsilon^R(s)}{\partial x^2}\right\rangle.
\end{aligned}
$$

The same kind of computations for the term $2f_{j,\epsilon}\mathscr{L}_\epsilon^R f_{j,\epsilon}$ lead to the result. □

3.7 An application of the Diffusion-Approximation theory in infinite dimension : Proof of Theorem 3.1.4

In this section, Theorem 3.1.4 is proved using similar arguments as those used to prove theorem 3.1.1 and 3.1.3. However, the proof is simpler because the solutions of (3.1.2) and (3.1.3) are global in time. Therefore, we only give the main arguments and computations leading to the final result.

Existence and uniqueness. The proof of the existence part is done by a Banach fixed-point procedure. Since the process α is assumed to be almost surely bounded, there is no need to introduce a cut-off function and bounds are obtained almost surely. Following the proof of Theorem 3.1.1 the unbounded operator

$$H = ib'\sigma_3\partial_x + \frac{d_0}{2}\partial_x^2$$

with domain $\mathbb{H}^2$ generates a strongly continuous group of unitary operators $U(t)$ on $\mathbb{L}^2$ satisfying Strichartz estimates. We set

$$G\left(\Phi_\varepsilon, t/\varepsilon^2\right) = \frac{5}{6}\left|\Phi_\varepsilon\right|^2\Phi_\varepsilon + \frac{1}{6}\left(\Phi_\varepsilon^*\sigma_3\Phi_\varepsilon\right)\sigma_3\Phi_\varepsilon + \frac{1}{3}N\left(\Phi_\varepsilon, t/\varepsilon^2\right).$$

and consider the integral formulation

$$\Phi_\varepsilon(t) = U(t)\Phi_0 + \frac{i}{\varepsilon}\int_0^t U(t-s)\Sigma\left(\alpha_\varepsilon(s), \frac{s}{\varepsilon^2}\right)\Phi_\varepsilon(s)ds + \int_0^t U(t-s)G\left(\Phi_\varepsilon(s), \frac{s}{\varepsilon^2}\right)ds, \tag{3.7.1}$$

which is the mild formulation of Equation (3.1.3). Let us denote by $\mathcal{B}_R$ the closed ball of $\mathcal{U}_c^T$, of radius R with center 0. Given $\Phi_\varepsilon \in \mathcal{U}_c^T$, we denote by $\mathcal{T}\Phi_\varepsilon(t)$ the right hand side of Equation (3.7.1). Applying Proposition 3.2.2, using similar arguments as in Lemma 3.2.2 and Hölder inequality, we obtain that for any $T > 0$ and $t \in [0, T]$, there exists a positive constant C such that almost surely

$$\left\|\mathcal{T}\Phi_\varepsilon\right\|_{\mathcal{U}_c^T} \leqslant C\left\|\Phi_0\right\|_{\mathbb{L}^2} + \frac{CT}{\varepsilon}\sup_{t\in[0,T]}\left|\alpha_\varepsilon(t)\right|\left\|\Phi_\varepsilon\right\|_{\mathcal{U}_c^T} + CT^{1/2}\left\|\Phi_\varepsilon\right\|_{\mathcal{U}_c^T}^3.$$

We deduce that almost surely $\mathcal{T}$ maps $\mathcal{B}_R$ into itself choosing $R = 2C\left\|\Psi_0\right\|_{\mathbb{H}^2}$ and provided that T is such that $CT/\varepsilon M + CT^{1/2}R^2 < 1/2$, where $M = \sup_{t\in[0,T]}\left|\alpha_\varepsilon(t)\right|$. Using similar arguments, we also prove that $\mathcal{T}$ is a contraction mapping on $\mathcal{B}_R$, if T is chosen small enough depending again on R, ε and M. From classical arguments for the NLS equation [13, 44, 65, 101], we construct a maximal solution on $[0, t_\varepsilon^*)$ and the following alternative on the blow-up time holds :

$$t_\varepsilon^* = +\infty \qquad \text{or} \qquad \lim_{t\to t_\varepsilon^*}\left\|\Phi_\varepsilon(t)\right\|_{\mathbb{L}^2} = +\infty.$$

Since the process α is real valued, the matrix $\Sigma\left(\alpha_\varepsilon, t/\varepsilon^2\right)$ is hermitian and the conservation of the $\mathbb{L}^2$ norm follows. Let us now denote by $\mathcal{B}_R$ the closed ball of radius R of $\mathcal{V}^T$. The existence and uniqueness of a local solution of Equation (3.1.2) in $\mathbb{H}^1$ is obtained combining the fact that $\mathcal{T}$ is a contraction mapping in $\mathcal{U}^T$, that $\mathcal{B}_R$ is a complete metric space endowed with the norm of $\mathcal{U}^T$ and $\mathcal{T}\mathcal{B}_R \subset \mathcal{B}_R$. The continuity of the paths of the solution is obtained afterwards by Proposition 3.2.2. From the blow-up alternative, a bound on the $\mathbb{L}^2$ norm of the gradient will ensure the existence of a global solution. Usually, this bound is obtained thanks to the conservation of the energy. However, the energy H, associated to the deterministic part of Equation (3.1.2)

$$\begin{aligned} H(\Phi_\varepsilon) = \frac{d_0}{2}\int_{\mathbb{R}}\left|\partial_x\Phi_\varepsilon\right|^2 dx - \frac{5}{24}\int_{\mathbb{R}}\left|\Phi_\varepsilon\right|^4 dx - \frac{1}{24}\int_{\mathbb{R}}\left(\left|\Phi_{1,\varepsilon}\right|^2 - \left|\Phi_{2,\varepsilon}\right|^2\right)^2 dx \\ - \frac{1}{12}\int_{\mathbb{R}}\mathcal{R}\mathrm{e}\left\{\overline{\Phi_{1,\varepsilon}}^2\Phi_{2,\varepsilon}^2 e^{-4ibt/\varepsilon^2}\right\}dx \end{aligned}$$

is not preserved. Nevertheless, a priori estimates may be obtained on the $\mathbb{H}^1$ norm of the solution. We know, from the $\mathbb{L}^2$ existence, that $\left\|\Phi_\varepsilon(t)\right\|_{\mathbb{L}^2} = \left\|\Phi_0\right\|_{\mathbb{L}^2}, \forall t \in \mathbb{R}_+$. Using the integral formulation (3.7.1), Proposition 3.2.2 and Hölder inequality,

$$\begin{aligned} \left\|\partial_x\Phi_\varepsilon\right\|_{L^\infty(0,T;\mathbb{L}^2)} \leqslant \left\|\partial_x\Phi_0\right\|_{\mathbb{L}^2} + \frac{CMT}{\varepsilon}\left\|\partial_x\Phi_\varepsilon\right\|_{L^\infty(0,T;\mathbb{L}^2)} \\ + CT^{1/2}\left\|\Phi_\varepsilon\right\|_{L^4(0,T;\mathbb{L}^\infty)}^2\left\|\partial_x\Phi_\varepsilon\right\|_{L^\infty(0,T;\mathbb{L}^2)}. \end{aligned}$$

Using again formulation (3.7.1), Strichartz estimates for the admissible pair $(4,\infty)$, the conservation of the $\mathbb{L}^2$ norm and Hölder inequality, we obtain that there exists a positive constant C such that

$$\|\Phi_\varepsilon\|_{L^4(0,T;\mathbb{L}^\infty)} \leqslant C\left(\left(1+\frac{TM}{\varepsilon}\right)\|\Phi_0\|_{\mathbb{L}^2} + T^{1/2}\|\Phi_\varepsilon\|^3_{L^8(0,T;\mathbb{L}^4)}\right).$$

Using the same arguments as in Lemma 3.2.3, we obtain that the norm $\|\Phi_\varepsilon\|_{L^8(0,T;\mathbb{L}^4)}$ is almost surely bounded by the $\mathbb{L}^2$ norm of the initial data Φ_0, i.e.

$$\|\Phi_\varepsilon\|_{L^8(0,T;\mathbb{L}^4)} \leqslant C\left(T, \|\Phi_0\|_{\mathbb{L}^2}, \varepsilon\right).$$

Combining the above estimates leads to the following bound on the $\mathbb{L}^2$ norm of the gradient

$$\begin{aligned}\|\partial_x\Phi_\varepsilon\|_{L^\infty(0,T;\mathbb{L}^2)} \quad \leqslant \quad & \|\partial_x\Phi_0\|_{\mathbb{L}^2} + \frac{CMT}{\varepsilon}\|\partial_x\Phi_\varepsilon\|_{L^\infty(0,T;\mathbb{L}^2)} \\ & + CT^{1/2}\left(\|\Phi_0\|_{\mathbb{L}^2} + T^{1/2}C\left(T,\|\Phi_0\|_{\mathbb{L}^2},\varepsilon\right)^3\right)^2\|\partial_x\Phi_\varepsilon\|_{L^\infty(0,T;\mathbb{L}^2)}.\end{aligned}$$

We conclude that there exists a unique global solution Φ_ε in $C(\mathbb{R}_+;\mathbb{H}^1)$ for an initial data in $\mathbb{H}^1$. However, this bound is not uniform in ε and may not be used to prove a tighness criterion.

Limit theorem for probability measures. In this paragraph, we prove the second part of Theorem 3.1.4. Let us observe that by an application of the Banach fixed point theorem, we may easily prove the continuity of the mapping

$$\begin{aligned}\Phi : C([0,T]) &\to C([0,T],\mathbb{H}^1) \\ \alpha &\mapsto \Phi(\alpha)\end{aligned}$$

for initial data $\Phi_0 \in \mathbb{H}^1$. However, the limit process Ψ, solution of (3.1.3), is not a continuous mapping from $C([0,T])$ into $C([0,T];\mathbb{H}^1)$ of the two dimensional Brownian motion (W_1,W_2). Hence, we may not proceed as in [21, 28, 73]. Accordingly, we prove the limit theorem using the tools developed in the previous sections. In order to work with homogeneous Markov process, and define the infinitesimal generator in the usual way, we follow the approach given in [38, 41], defining the period $Z_0 = \pi/b$ and the mapping

$$\begin{aligned}\tau : \mathbb{R}_+ &\to \mathbb{R}_+/Z_0\mathbb{N} \\ t &\mapsto t \bmod Z_0.\end{aligned}$$

We denote by Φ^R_ε the solution of a cut-off version of Equation (3.1.2), i.e. the cut-off function $\Theta_R\left(\left\|\Phi^R_\varepsilon(t)\right\|^2_{\mathbb{H}^1}\right)$ is added in front of the nonlinear term $G\left(\Phi_\varepsilon, t/\varepsilon^2\right)$. Therefore, the enlarged process $\left(\Phi^R_\varepsilon, \alpha_\varepsilon, \tau\left(t/\varepsilon^2\right)\right)$, defined on the state space $\mathbb{H}^3 \times S \times [0,Z_0)$, is Markov and we denote $\mathscr{L}^R_\varepsilon$ its infinitesimal generator. Thus, for all

test functions $f : \mathbb{H}^{-1} \times S \times [0, Z_0) \to \mathbb{R}$ of class C_b^2 the infinitesimal generator is given by

$$\begin{aligned}\mathscr{L}_\varepsilon^R f(v, \alpha_0, \tau) = &\left\langle D_v f(v, \alpha_0, \tau), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} - b'\sigma_3 \frac{\partial v}{\partial x} + i\Theta_R\left(\|v\|_{\mathbb{H}^1}^2\right) G(v, \tau)\right\rangle \\ &+ \frac{1}{\varepsilon}\left\langle D_v f(v, \alpha_0, \tau), i\Sigma(\alpha_0, \tau)\, v\right\rangle + \frac{1}{\varepsilon^2}\mathcal{Q} f(v, \alpha_0, \tau),\end{aligned}$$

where $\mathcal{Q} = \mathscr{L}_\alpha + \partial_\tau$. We now introduce a perturbed test function of the form

$$f_\varepsilon(v, \alpha_0, \tau) = f(v) + \varepsilon f^1(v, \alpha_0, \tau) + \varepsilon^2 f^2(v, \alpha_0, \tau)$$

and formally compute the expression $\mathscr{L}_\varepsilon^R f_\varepsilon$

$$\begin{aligned}\mathscr{L}_\varepsilon^R f_\varepsilon(v, \alpha_0, \tau) = &\left\langle D_v f(v), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} - b'\sigma_3 \frac{\partial v}{\partial x} + i\Theta_R\left(\|v\|_{\mathbb{H}^1}^2\right) G(v, \tau)\right\rangle \\ &+ \frac{1}{\varepsilon}\mathcal{Q} f^1(v, \alpha_0, \tau) + \frac{1}{\varepsilon}\left\langle D_v f(v), i\Sigma(\alpha_0, \tau)\, v\right\rangle + \mathcal{Q} f^2(v, \alpha_0, \tau) \\ &+ \left\langle D_v f^1(v, \alpha_0, \tau), i\Sigma(\alpha_0, \tau)\, v\right\rangle + \mathcal{O}(\varepsilon),\end{aligned} \tag{3.7.2}$$

where

$$\begin{aligned}\mathcal{O}(\varepsilon) = &\varepsilon\left\langle D_v f^1(v, \alpha_0, \tau), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} - b'\sigma_3 \frac{\partial v}{\partial x} + i\Theta_R\left(\|v\|_{\mathbb{H}^1}^2\right) G(v, \tau)\right\rangle \\ &+ \varepsilon\left\langle D_v f^2(v, \alpha_0, \tau), i\Sigma(\alpha_0, \tau)\, v\right\rangle \\ &+ \varepsilon^2\left\langle D_v f^2(v, \alpha_0, \tau), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} - b'\sigma_3 \frac{\partial v}{\partial x} + i\Theta_R\left(\|v\|_{\mathbb{H}^1}^2\right) G(v, \tau)\right\rangle.\end{aligned}$$

The first corrector is chosen to cancel the $1/\varepsilon$ terms i.e.

$$\mathcal{Q} f^1(v, \alpha_0, \tau) = -\left\langle D_v f(v), i\Sigma(\alpha_0, \tau)\, v\right\rangle. \tag{3.7.3}$$

Obviously, the joint markov process (α, τ) has an invariant measure that is the product of the invariant probability measure μ of α and the lebesgue measure on $[0, Z_0]$, which is normalized to get the uniform probability measure. Moreover, the joint process (α, τ) is ergodic under this invariant measure. Indeed, the semigroup associated to the order one partial derivative ∂_τ is the shift operator S_τ, i.e. for any continuous bounded functions f

$$S_\tau f(x) = f(x + \tau).$$

Thus, the S_τ-invariant functions satisfy $f(x + \tau) = f(x)$, for any $\tau \in [0, Z_0)$. Therefore, f is constant and τ is ergodic. Since $\Sigma(\alpha_0, \tau)$ satisfies the null mass condition, with respect to the joint invariant measure, a particular solution of the Poisson equation is given by

$$\begin{aligned}f^1(v, \alpha_0, \tau) &= \int_0^{+\infty} T_t S_t \left\langle D_v f(v), i\Sigma(\alpha_0, \tau)\, v\right\rangle dt \\ &= \int_0^{+\infty} \mathbb{E}\left(\left\langle D_v f(v), i\Sigma(\alpha(t), \tau + t)\, v\right\rangle \middle| \alpha(0) = \alpha_0\right) dt.\end{aligned} \tag{3.7.4}$$

We denote

$$g(v,\alpha_0,\tau)=\left\langle D_vf(v),\frac{id_0}{2}\frac{\partial^2 v}{\partial x^2}-b'\sigma_3\frac{\partial v}{\partial x}+i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right)G\left(v,\tau\right)\right\rangle \\ +\left\langle D_vf^1(v,\alpha_0,\tau),i\Sigma\left(\alpha_0,\tau\right)v\right\rangle$$

and the test function f^2 is chosen as a solution of the Poisson equation

$$\mathcal{Q}f^2(v,\alpha_0,\tau)=-g(v,\alpha_0,\tau)+\mathbb{E}_\mu\left(g(v,\alpha_0,\tau)\right) \tag{3.7.5}$$

Hence, for this choice of perturbed test functions, the expression of the infinitesimal generator $\mathscr{L}_\varepsilon^R$ is given by

$$\mathscr{L}_\varepsilon^R f_\varepsilon(v,\alpha_0,\tau)=\mathbb{E}_\mu\left(g(v,\alpha_0,\tau)\right)+\mathcal{O}(\varepsilon).$$

Writing that $\Sigma\left(\alpha_0,\tau\right)=\alpha_0\cos(2b\tau)\sigma_2-\alpha_0\sin(2b\tau)\sigma_1$, the expectation may be written as

$$\mathbb{E}_\mu\left(\left\langle D_vf^1(v,\alpha_0,\tau),i\Sigma\left(\alpha_0,\tau\right)v\right\rangle\right) \\ =\sum_{i,j=1}^{2}\left\langle D_v^2f(v)\sigma_iv,\sigma_jv\right\rangle C_{ij}-\sum_{i,j=1}^{2}\left\langle D_vf(v),\sigma_i\sigma_jv\right\rangle C_{ij},$$

where the covariance coefficients C_{ij} are given by

$$\begin{cases} C_{11}=C_{22}=\dfrac{1}{2}\gamma_c:=\dfrac{1}{2}\displaystyle\int_0^{+\infty}\mathbb{E}\left(\alpha_0\alpha(t)\right)\cos(2bt)dt \\ C_{12}=-C_{21}=\dfrac{1}{2}\gamma_s:=\dfrac{1}{2}\displaystyle\int_0^{+\infty}\mathbb{E}\left(\alpha_0\alpha(t)\right)\sin(2bt)dt. \end{cases}$$

Therefore, using properties of the pauli matrices stated in 4.1.1, we deduce that

$$\mathbb{E}_\mu\left(\left\langle D_vf^1(v,\alpha_0,\tau),i\Sigma\left(\alpha_0,\tau\right)v\right\rangle\right) \\ =\frac{\gamma_c}{2}\sum_{j=1}^{2}\left\langle D_v^2f(v)\sigma_jv,\sigma_jv\right\rangle-\gamma_c\left\langle D_vf(v),v\right\rangle+\gamma_s\left\langle D_vf(v),i\sigma_3v\right\rangle \tag{3.7.6}$$

Replacing $\mathcal{Q}f^1$ and $\mathcal{Q}f^2$ in Equation (3.7.2), by respectively the right hand side of (3.7.3) and (3.7.5) and using the expression (3.7.6), we get

$$\mathscr{L}_\varepsilon^R f_\varepsilon(v,\alpha_0,\tau) \\ =\left\langle D_vf(v),\frac{id_0}{2}\frac{\partial^2 v}{\partial x^2}-b'\sigma_3\frac{\partial v}{\partial x}+i\Theta_R\left(\|v\|^2_{\mathbb{H}^1}\right)F\left(v\right)\right\rangle \\ +\frac{\gamma_c}{2}\sum_{j=1}^{2}\left\langle D_v^2f(v)\sigma_jv,\sigma_jv\right\rangle-\gamma_c\left\langle D_vf(v),v\right\rangle+\gamma_s\left\langle D_vf(v),i\sigma_3v\right\rangle+\mathcal{O}(\varepsilon)$$

where the nonlinear mapping $F(v)=\frac{5}{6}\left|v\right|^2v+\frac{1}{6}\left(v^*\sigma_3v\right)\sigma_3v$ is simply the averaging of the mapping $F(v,\tau)$ over one period Z_0. The formal infinitesimal limit generator

is thus given by

$$\mathscr{L}^R f(v) = \left\langle D_v f(v), \frac{id_0}{2}\frac{\partial^2 v}{\partial x^2} - b'\sigma_3 \frac{\partial v}{\partial x} + i\Theta_R\left(\|v\|^2_{\mathrm{H}^1}\right) F(v) \right\rangle$$
$$+ \frac{\gamma_c}{2}\sum_{j=1}^{2} \left\langle D_v^2 f(v)\sigma_j v, \sigma_j v\right\rangle - \gamma_c \left\langle D_v f(v), v\right\rangle + \gamma_s \left\langle D_v f(v), i\sigma_3 v\right\rangle .$$

The end of the proof follows exactly the same lines as in the proof of Theorem 3.1.3.

Chapter 4

Strong order of convergence of a semidiscrete scheme for the stochastic Manakov equation

4.1 Introduction

In this chapter, we consider a semidiscrete version of the Crank Nicolson scheme for the stochastic Manakov equation (4.1.1), that we adapt to our stochastic setting. Our aim is to analyse the order of the error for this scheme and we prove that the strong order is 1/2. The stochastic Manakov equation is a stochastic perturbation, in the stratonovich sense, of the Manakov equation. This equation appears as the asymptotic dynamic of the Manakov PMD equation (3.1.9) and is given by (see Chapter 3)

$$idX(t)+\left(\frac{d_0}{2}\frac{\partial^2 X(t)}{\partial x^2}+F\left(X(t)\right)\right)dt+i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k\frac{\partial X(t)}{\partial x}\circ dW_k(t)=0,\quad t\geqslant 0, x\in\mathbb{R}, \tag{4.1.1}$$

where γ is a small positive parameter given by the physics of the problem, d_0 is the group velocity dispersion (GVD) and $W=(W_1,W_2,W_3)$ is a 3d real valued Brownian motion. The matrices $\sigma_1,\sigma_2,\sigma_3$ are the Pauli matrices and the nonlinear term is given by $F\left(X(t)\right)=\frac{8}{9}\left|X\right|^2X(t)$. The equivalent Itô formulation is given by

$$idX(t)+\left(\left(\frac{d_0}{2}-\frac{3i\gamma}{2}\right)\frac{\partial^2 X(t)}{\partial x^2}+F(X)(t)\right)dt+i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k\frac{\partial X(t)}{\partial x}dW_k(t)=0. \tag{4.1.2}$$

For simplicity, we will take in the sequel $d_0/2=1$ and we set $C_\gamma=i+\frac{3\gamma}{2}$. In the deterministic case (i.e. when $\gamma=0$), when one considers the Manakov Equation, both the mass and the Hamiltonian H are preserved. This is not the case for the stochastic equation that preserves only the mass, the Hamiltonian structure being destroyed by the noise (see Chapter 3). Several numerical approximations have been proposed to simulate the solution of the deterministic equation, such as the Crank-Nicolson scheme [31, 33], the Relaxation scheme [4] and Fourier split-step schemes [5, 64, 103, 111]. These schemes are known to be conservative. Indeed,

both the Crank-Nicolson and Relaxation schemes have very good properties since the discrete mass and the discrete Hamiltonian are preserved. On the contrary, the splitting scheme fails in preserving exactly H. The design of a numerical scheme for (4.1.1) is based on the fact that it must preserve the discrete mass since it is an invariant quantity in the continuous case. Numerical schemes for (4.1.1) are constructed from the above mentioned schemes. In this chapter, we are interested in the numerical analysis of a semi-discrete Crank-Nicolson scheme given by

$$\begin{cases} X_N^{n+1} - X_N^n + H_{\Delta t,n} X_N^{n+1/2} - iF\left(X_N^n, X_N^{n+1}\right)\Delta t = 0 \\ F\left(X_N^n, X_N^{n+1}\right) = \frac{4}{9}\left(\left|X_N^n\right|^2 + \left|X_N^{n+1}\right|^2\right) X_N^{n+1/2}, \end{cases}$$

where $X_N^{n+1/2} = \left(X_N^{n+1} + X_N^n\right)/2$, Δt is the time step and $\sqrt{\Delta t}\chi_k^n = W_k\left((n+1)\Delta t\right) - W_k\left(n\Delta t\right)$ is the noise increment. The random matrix operator $H_{\Delta t,n}$ is defined by

$$H_{\Delta t,n} = -i\Delta t I_2 \partial_x^2 + \sqrt{\gamma \Delta t}\sum_{k=1}^{3} \sigma_k \chi_k^n \partial_x,$$

The question that needs to be addressed is the discretization of the noise term. There is actually two different approaches based on the fact that, in the continuous case, Equation (4.1.1) and Equation (4.1.2) are equivalent. Hence, one may either propose an implicit discretization of the Stratonovich integral, using the midpoint rule, or an explicit discretization of the Itô integral. However, in the discrete setting, the two formulations are not equivalent. Indeed, the discrete $\mathbb{L}^2$ norm is not preserved when considering the Itô approach, while the implicit discretization of the Stratonovich integral allows preservation of the mass. Note that the conservation of the discrete mass immediately leads to the unconditional $\mathbb{L}^2$ stability of the scheme.

There is actually a more profound reason that keeps us from using a numerical scheme based on the Itô equation and this reason lies in the fact that the noise term contains a one order derivative. It is well known from the deterministic literature, that explicit schemes for the advection equation require a stability criterion to converge, while implicit schemes are stable. Indeed, the basic explicit scheme for this equation is stable in $\mathbb{L}^2$ only under the CFL condition $|V|\,\Delta t \leqslant \Delta x$, where Δt is the time step, Δx is the space step and V the constant velocity. When considering the Itô approach, the discretization of the stochastic integral has to be explicit in order to be consistent with the equation, since an implicit discretization converges to the backward Itô integral. Therefore, the Itô approach leads to a CFL condition that depends on Gaussian random variables. Since they are not bounded, this random stability condition may be very restrictive. For similar reasons, it may happen that implicit schemes are not stable [83], but this is not the case for the stochastic Manakov PMD equation.

Numerical simulations of stochastic differential equations are used in practice to solve complicated equations and to lighten some hidden behaviours such as large deviations [12]. In optics, numerical simulations of the stochastic Manakov equation (4.1.1) may help to understand the impact of the PMD on the pulse spreading. Depending on the problem, one may not be interested in the same quantities. On one

hand, one may be interested in the computation of path samples (related to strong solutions) to emphasize, for exemple, the relation between various parameters in the dynamics. On the other hand, if the quantity under interest depends only on the law of the dynamics, one will focus on weak approximations. It is important to understand that the two types of approximations may lead to very different numerical schemes. The pathwise error analysis of numerical schemes for SDE has been intensively studied [35, 67, 82, 104], whereas the weak error analysis started later with the work of Milstein [80, 81] and Talay [105], who used the Kolmogorov equation associated to the SDE to obtain a weak order of convergence. Usually, for Euler schemes, the strong order is 1/2. More sophisticated schemes exist to increase the pathwise order but their numerical implementation requires to compute multiple iterated integrals, which may be difficult if the dynamics is driven by a multi-dimensional Brownian motion.

The numerical analysis of SPDEs combines stochastic analysis together with PDEs numerical approximation. Most of the results are concerned with the analysis of pathwise convergence for solutions of semi-linear and quasi-linear parabolic equations (for a non exhaustive list, see [16, 48, 49, 50, 51, 52, 58, 79, 95]). There is some recent literature on dispersive equations for stochastic nonlinear Schrödinger equations [19, 20, 74] and for a stochastic Korteweg-de-Vries equation [25, 26]. Weak order for SPDEs has been considered later [23, 27, 59]; the proof consists then in using the Kolmogorov equation which is now a PDE with an infinite number of variables. In our case, the difficult point lies in the fact that the noise term contains a one order derivative that cannot be treated as a perturbation (see Chapter 3). Then, it requires a lot of regularity on the initial data to carry the computations. Moreover, as mentioned above, an implicit discretization of the noise has to be considered. The delicate point, in order to obtain the strong error, is to deal with random matrices. Another difficulty is that the stochastic Manakov equation (4.1.1) may have blowing-up solutions (see Chapter 3). Therefore, we have to define a discrete solution of the scheme up to a discrete blow-up time as in [20].

In a first section, we introduce some notations and the main result of this chapter. Then, following the approach of Chapter 3 for the continuous equation, we construct a discrete random propagator associated to the linear equation. We study the linear Euler equation with implicit discretization of the noise and prove that the strong order is 1/2. In a third section, we give a result on the strong order of convergence for the nonlinear equation.

4.1.1 Notation and main result

For all $p \geqslant 1$, we define $\mathbb{L}^p(\mathbb{R}) = \left(L^p(\mathbb{R};\mathbb{C})\right)^2$ the Lebesgue spaces of functions with values in $\mathbb{C}^2$. Identifying $\mathbb{C}$ with $\mathbb{R}^2$, we define a scalar product on $\mathbb{L}^2(\mathbb{R})$ by

$$(u,v)_{\mathbb{L}^2} = \sum_{i=1}^{2} \mathcal{R}\mathrm{e}\left\{\int_{\mathbb{R}} u_i \overline{v_i} dx\right\}.$$

We denote by $\mathbb{H}^m(\mathbb{R}), m \in \mathbb{N}$ the space of functions in $\mathbb{L}^2$ such that their m first derivatives are in $\mathbb{L}^2$. We will also use $\mathbb{H}^{-m}$ the topological dual space of $\mathbb{H}^m$ and

denote $\langle .,. \rangle$ the paring between $\mathbb{H}^m$ and $\mathbb{H}^{-m}$. The Fourier transform of a tempered distribution $v \in \mathcal{S}'(\mathbb{R})$ is either denoted by $\widehat{v}$ or $\mathcal{F}v$. If $s \in \mathbb{R}$ then $\mathbb{H}^s$ is the fractional Sobolev space of tempered distributions $v \in \mathcal{S}'(\mathbb{R})$ such that $(1+|\xi|^2)^{s/2}\widehat{v} \in \mathbb{L}^2$. Let $(E, \|.\|_E)$ and $(F, \|.\|_F)$ be two Banach spaces. We denote by $\mathcal{L}(E,F)$ the space of linear continuous functions from E into F, endowed with its natural norm. If I is an interval of $\mathbb{R}$ and $1 \leqslant p \leqslant +\infty$, then $L^p(I;E)$ is the space of strongly Lebesgue measurable functions f from I into E such that $t \mapsto \|f(t)\|_E$ is in $L^p(I)$. The space $L^p(\Omega, E)$ is defined similarly where $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space.

We assume that the set $\{\Delta t\} = \{\Delta t_n\}_{n \in \mathbb{N}}$ is a discrete sequence converging to 0. We define a final time $T > 0$ and an interval $[0,T]$ on which we will consider the approximation of the solution of (4.1.1). Moreover $N_T = [T/\Delta t]$, the integer part of $T/\Delta t$. Similarly for any stopping time τ, $N_\tau = [\tau/\Delta t]$. Moreover we write $t_n = n\Delta t$ for any $n \in [\![0, N]\!]$ where N is either N_T or N_τ according to the situation. We denote by $L^\infty(0,T;\mathbb{H}^m)$ the space of all bounded sequences for $n = 0, \cdots, N$ with values in $\mathbb{H}^m$ endowed with the supremum norm

$$\|X_N\|_{L^\infty(0,T;\mathbb{H}^m)} = \sup_{\substack{n \in \mathbb{N} \\ n\Delta t \leqslant T}} \|X_N^n\|_{\mathbb{H}^m}.$$

Moreover for a $n \times n$ matrix $A = \{a_{ij}\}$, the uniform norm is defined by

$$|||A|||_\infty = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}|$$

and the spectral norm of A is defined by

$$|||A|||_2 = \sqrt{\rho(A^*A)}$$

where A^* is the conjugate transpose and ρ is the spectral radius. Finally we denote $dW_0(u) = du$ and we introduce the notations for $j,k \in [\![0,3]\!]$

$$W_j^{n,s}(f) = \int_{t_n}^{s} f(u) dW_j(u)$$
$$W_{j,k}^{n,n+1}(f) = \int_{t_n}^{t_{n+1}} \int_{t_n}^{s} f(u) dW_j(u) dW_k(s).$$

We recall that the Pauli matrices have the following properties

Property 4.1.1. *Let $j,k \in [\![0,3]\!]$, then*

- *Commutation relations : $[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k$.*
- *Anticommutation relations : $\sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij} \cdot I_2$ and $\sigma_i = \sigma_i^*$,*

where $\varepsilon_{ijk} = (i-j)(j-k)(k-i)/2$ is the Levi-Civita symbol.

Our aim is to design a scheme which preserves the properties of the continuous solution and is consistent with the stochastic Manakov equation (4.1.1). By definition, the Stratonovich integral is defined as

$$\int_0^T Y(s) \circ dW(s) = \lim_{|\pi| \to 0} \sum_{i=1}^{n} \frac{Y(t_i) + Y(t_{i-1})}{2} \Delta W^i,$$

where $\pi = \{0 = t_1 < \cdots < t_n = T\}$ is a homogeneous partition of $[0, T]$ and $\Delta W^i = (W(t_i) - W(t_{i-1}))$. The limit, if it exists, in the sense of convergence in probability, is called the Stratonovich integral. Let us describe our numerical scheme. We denote by X_N^n a numerical approximation of the solution X, at time t_n, of Equation (4.1.1). We set $X_N^{n+1/2} = \frac{1}{2}\left(X_N^{n+1} + X_N^n\right)$ and for $0 \leqslant n \leqslant N-1$

$$\chi_k^n = \frac{W_k\left((n+1)\Delta t\right) - W_k\left(n\Delta t\right)}{\sqrt{\Delta t}} \sim \mathcal{N}(0,1).$$

The scheme reads as follows

$$\begin{cases} X_N^{n+1} - X_N^n + H_{\Delta t,n} X_N^{n+1/2} - iF\left(X_N^n, X_N^{n+1}\right)\Delta t = 0 \\ F\left(X_N^n, X_N^{n+1}\right) = \frac{4}{9}\left(\left|X_N^n\right|^2 + \left|X_N^{n+1}\right|^2\right) X_N^{n+1/2}, \end{cases} \tag{4.1.3}$$

where $H_{\Delta t,n}$ is the random matrix operator defined by

$$H_{\Delta t,n} = -i\Delta t I_2 \partial_x^2 + \sqrt{\gamma \Delta t} \sum_{k=1}^{3} \sigma_k \chi_k^n \partial_x$$

with domain $\mathscr{D}(H_{\Delta t,n}) = \mathbb{H}^2(\mathbb{R}) \subset \mathbb{L}^2(\mathbb{R})$ independent of n and I_2 is the 2×2 identity matrix. This operator is easily described thanks to the Fourier transform. Indeed, for any $\xi \in \mathbb{R}$

$$\widehat{H}_{\Delta t,n}(\xi) = \begin{pmatrix} i\Delta t\left|\xi\right|^2 + i\sqrt{\gamma\Delta t}\chi_3^n \xi & i\sqrt{\gamma\Delta t}\left(\chi_1^n - i\chi_2^n\right)\xi \\ i\sqrt{\gamma\Delta t}\left(\chi_1^n + i\chi_2^n\right)\xi & i\Delta t\left|\xi\right|^2 - i\sqrt{\gamma\Delta t}\chi_3^n \xi \end{pmatrix}. \tag{4.1.4}$$

Moreover, we set

$$\begin{aligned} U_{\Delta t,n} &= (\mathrm{Id} + \frac{1}{2} H_{\Delta t,n})^{-1} (\mathrm{Id} - \frac{1}{2} H_{\Delta t,n}) \\ T_{\Delta t,n} &= (\mathrm{Id} + \frac{1}{2} H_{\Delta t,n}), \end{aligned}$$

where Id is the identity mapping in $\mathbb{L}^2$. We denote by $\widetilde{X}^n = X(t_n)$ the solution of Equation (4.1.1), evaluated at the point t_n. Let us now give the main result of this chapter stating that the approximation of Equation (4.1.1) by the scheme (4.1.3) has an order $1/2$ in probability.

Theorem 4.1.1. *Assume that $X_0 \in \mathbb{H}^6$, then for any stopping time $\tau < \tau^*$ almost surely we have*

$$\lim_{C\to+\infty} \mathbb{P}\left(\max_{n=0,\dots,N_\tau} \left\|X^n - \widetilde{X}^n\right\|_{\mathbb{H}^1} \geqslant C\Delta t^{1/2}\right) = 0,$$

uniformly in Δt. Then we say, according to [95], that the scheme has an order $1/2$ in probability. Moreover, for any $\delta < \frac{1}{2}$, there exists a random variable K_δ such that

$$\max_{n=0,\dots,N_\tau} \left\|X^n - \widetilde{X}^n\right\|_{\mathbb{H}^1} \leqslant K_\delta(T,R,\omega)\Delta t^\delta.$$

The order of convergence is obtained writing the Stratonovich integral as an Itô integral plus a correction

$$\sum_{i=1}^{n} \frac{Y(t_i)+Y(t_{i-1})}{2}\Delta W^i = \sum_{i=1}^{n} Y(t_{i-1})\Delta W^i + \sum_{i=1}^{n} \frac{Y(t_i)-Y(t_{i-1})}{2}\Delta W^i$$

and comparing it with the continuous Itô Equation.

Remark 4.1.1. *The time centering method used for this scheme allows it to be conservative (at least for the mass). In the deterministic case, it can be shown that this discretization has some pathologies : there is no local smoothing effect and no Strichartz estimates are available ([62, 63]). Instead, if we considered an implicit discretization of the second order derivative, the scheme would have good dispersive estimates as in the continuous case but would be dissipative.*

4.2 The linear equation.

In this section, we study the approximation of the solution of the linear equation

$$idX(t) + \frac{\partial^2 X(t)}{\partial x^2}dt + i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k \frac{\partial X(t)}{\partial x}\circ dW_k(t) = 0, \quad t \geqslant 0, x \in \mathbb{R} \tag{4.2.1}$$

given by the semidiscrete implicit approximation

$$X_N^{n+1} - X_N^n + H_{\Delta t,n} X_N^{n+1/2} = 0. \tag{4.2.2}$$

To lighten the notation, we will remove the dependence in N of the unknown X_N. The aim of this section is to give an existence result of an adapted solution for the scheme (4.2.2) and to give an estimate of the discretization error. The results are stated in Propositions 4.2.1 and 4.2.2 below.

4.2.1 Existence and stability

The next proposition states that the solution of the scheme (4.2.2) is uniquely defined and adapted, and that the mass is preserved.

Proposition 4.2.1. *Given $X_0 \in \mathbb{H}^m$ for $m \in \mathbb{N}$, then there exists a unique adapted discrete solution $(X^n)_{n=0,\cdots,N}$ to (4.2.2) that belongs to $L^\infty(0,T;\mathbb{H}^m)$. Moreover the $\mathbb{H}^m$ norm of the solution X^n of (4.2.2) is constant i.e. for all $n \in [\![0,N]\!]$*

$$\|X^n\|_{\mathbb{H}^m} = \|X_0\|_{\mathbb{H}^m}. \tag{4.2.3}$$

Proof of Proposition 4.2.1. Assume that X^n is a $\mathcal{F}_{n\Delta t}-$measurable random variable with values in $\mathbb{H}^m$. We have to prove that the operator $T_{\Delta t,n}$ is invertible for a.e. $\omega \in \Omega$ and for all n. Let us first note that $A_{\Delta t} = \Delta t I_2 \partial_x^2$ is selfadjoint in $\mathbb{L}^2$ with domain $\mathbb{H}^2$ and that $B_{\Delta t,n} = i\sqrt{\gamma\Delta t}\sum_{k=1}^3 \sigma_k \chi_k^n \partial_x v$ is a symmetric operator from $\mathbb{H}^1$ into $\mathbb{L}^2$ for a.e. $\omega \in \Omega$. Using Property 4.1.1 of the Pauli matrices, Cauchy-Schwarz and Young inequalities, we may prove that $B_{\Delta t,n}$ is almost surely $A_{\Delta t}$-bounded with relative bound smaller than 1. Indeed, for any $u \in \mathbb{H}^2$

$$\begin{aligned}
\left\| \sqrt{\gamma\Delta t}\sum_{k=1}^3 \sigma_k \chi_k^n \partial_x v \right\|_{\mathbb{L}^2}^2 &= -\left\langle \Delta t \partial_x^2 v, 3\gamma \left(\chi_k^n\right)^2 v \right\rangle \\
&\leqslant C(\gamma,\omega) \left\| (\Delta t)\partial_x^2 v \right\|_{\mathbb{L}^2} \|v\|_{\mathbb{L}^2} \\
&\leqslant \frac{1}{2}\left\| (\Delta t)\partial_x^2 v \right\|_{\mathbb{L}^2}^2 + \frac{C(\gamma,\omega)^2}{2}\|v\|_{\mathbb{L}^2}^2,
\end{aligned}$$

where $C(\gamma,\omega) = 3\gamma\left(\chi_k^n\right)^2$. Since $|C(\gamma,\omega)| < +\infty$ a.s., we deduce thanks to Kato-Rellich Theorem that $iH_{\Delta t,n}$ is selfadjoint in $\mathbb{L}^2$ with domain $\mathbb{H}^2$ and it follows that $T_{\Delta t,n}$ is invertible from $\mathbb{H}^2$ into $\mathbb{L}^2$. Hence, the unique $\mathcal{F}_{n+1}$-adapted solution is given by $X^{n+1} = U_{\Delta t,n} X^n$ a.s. The conservation of the $\mathbb{L}^2$ norm follows because $H_{\Delta t,n}$ is skew symmetric and that $U_{\Delta t,n}^* U_{\Delta t,n} = \mathrm{Id}$. □

In our case, the operator $T_{\Delta t,n}$ is invertible for every Δt. Thus, the implementation of the scheme (4.2.2) does not require to use a truncation of the noise term as in [83] to insure stability.

4.2.2 Strong order of convergence

Let us now be concerned with the order of convergence of the scheme (4.2.2). To this purpose, we denote by $\widetilde{X}^n = X(t_n)$ the solution of (4.2.1), evaluated at the point t_n, and define the vector error $e^n = \widetilde{X}^n - X^n$.

Proposition 4.2.2. *If $X_0 \in \mathbb{H}^{m+5}$, $m \in \mathbb{N}$, then the scheme (4.2.2) is convergent and for any $p \geqslant 1$*

$$\mathbb{E}\left(\max_{n\in[\![0,N]\!]} \|e^n\|_{\mathbb{H}^m}^{2p}\right) \leqslant C(T,\gamma,p,\|X_0\|_{\mathbb{H}^{m+5}})\Delta t^p. \tag{4.2.4}$$

Moreover, for all $p > 1$ and $\delta < \frac{1}{2} - \frac{1}{2p}$, there exists a positive constant $C_\delta(T,\gamma,\omega,p)$, depending on $\|X_0\|_{\mathbb{H}^{m+5}}$, such that

$$\max_{n=0,\cdots,N} \|e^n\|_{\mathbb{H}^m} \leqslant C_\delta(T,\gamma,\omega,p)\Delta t^\delta \quad a.s. \tag{4.2.5}$$

It may be surprising to require so much regularity on the initial data to prove a $L^p(\Omega)$ order of a linear equation. Usually, the order is obtained using the explicit expression of the group $S(t)$, solution of the free Schrödinger equation (that is $\gamma = 0$ in Equation (4.2.1)), and the mild form of the Itô equation. In our case, we cannot proceed similarly because of the implicit discretization of the noise and the presence of a differential operator in this term.

Before proving this result, we state a Lemma giving an estimate of the growth of the solution $X(s)$ of (4.2.1) starting at $\widetilde{X}^n$. This estimate will be used along the proof to obtain the order of convergence. We denote by $\widetilde{e}^n(s)$ the difference $\widetilde{e}^n(s) = X(s) - \widetilde{X}^n$, for all $s \in [t_n, t_{n+1}]$.

Lemma 4.2.1. *For any $p \geqslant 1$ and $m \in \mathbb{N}$, if $X_0 \in \mathbb{H}^{m+1}$ then*

$$\mathbb{E}\left(\sup_{t_n \leqslant s \leqslant t_{n+1}} \|\widetilde{e}^n(s)\|_{\mathbb{H}^m}^{2p}\right) \leqslant C_p(\gamma)\, \|X_0\|_{\mathbb{H}^{m+1}}^{2p}\, \Delta t^p \qquad \forall n = 0, \cdots, N.$$

Proof. Using the group $S(t)$, solution of the free Schrödinger equation (that is $\gamma = 0$ in Equation (4.2.1)), and writing the Itô formulation of Equation (4.2.1) under its mild form, we get

$$X(t) - \widetilde{X}^n = (S(t - t_n) - \mathrm{Id})\, \widetilde{X}^n + i\sqrt{\gamma} \sum_{k=1}^{3} \int_{t_n}^{t} S(t-u)\sigma_k \partial_x X(u) dW_k(u).$$

Using the Fourier transform, it can easily be shown that

$$\|(S(t) - \mathrm{Id})\, f\|_{\mathbb{H}^m} \leqslant C t^{1/2}\, \|f\|_{\mathbb{H}^{m+1}}, \qquad \forall f \in \mathbb{H}^{m+1},$$

from which we deduce, together with (4.2.3), that

$$\mathbb{E}\left(\sup_{t_n \leqslant s \leqslant t_{n+1}} \left\|(S(s - t_n) - \mathrm{Id})\, \widetilde{X}^n\right\|_{\mathbb{H}^m}^{2p}\right) \leqslant C_p\, \|X_0\|_{\mathbb{H}^{m+1}}^{2p}\, \Delta t^p.$$

Moreover, since X is adapted and belongs to $L^{2p}(\Omega, C([0,T], \mathbb{H}^m))$, we may apply the Burkholder-Davis-Gundy inequality to the stochastic convolution. Using the contraction property of the semigroup $S(t)$ and (4.2.3), we obtain the estimate

$$\mathbb{E}\left(\sup_{t_n \leqslant s \leqslant t_{n+1}} \left\|\sqrt{\gamma} \sum_{k=1}^{3} \int_{t_n}^{s} S(s-u)\sigma_k \partial_x X(u) dW_k(u)\right\|_{\mathbb{H}^m}^{2p}\right) \leqslant C_p\, \|X_0\|_{\mathbb{H}^{m+1}}^{2p}\, \gamma^p \Delta t^p.$$

This concludes the proof of the Lemma. □

Proof of Proposition 4.2.2. The proof is done for $m = 0$, but is easily extended to the general setting. Using the Itô formulation of Equation (4.2.1) and evaluating its solution on the time interval $[t_n, t_{n+1}]$, we obtain for all $x \in \mathbb{R}$,

$$\widetilde{X}^{n+1} = \widetilde{X}^n + C_\gamma W_0^{n,n+1} \left(\partial_x^2 X\right) - \sqrt{\gamma} \sum_{k=1}^{3} \sigma_k W_k^{n,n+1} \left(\partial_x X\right) \tag{4.2.6}$$

$$= \widetilde{X}^n - H_{\Delta t,n} \widetilde{X}^{n+1/2} + \epsilon_1^n + \epsilon_2^n, \tag{4.2.7}$$

where $C_\gamma = i + \frac{3\gamma}{2}$ and the processes ϵ_1^n and ϵ_2^n are given by

$$\begin{cases} \epsilon_1^n = iW_0^{n,n+1}\left(\partial_x^2 X - \partial_x^2 \widetilde{X}^{n+1/2}\right) \\ \epsilon_2^n = \sqrt{\gamma}\sum_{k=1}^3 \sigma_k \left(\partial_x \widetilde{X}^{n+1/2}\Delta W_k^n - W_k^{n,n+1}\left(\partial_x X\right)\right) + \frac{3\gamma}{2} W_0^{n,n+1}\left(\partial_x^2 X\right). \end{cases} \tag{4.2.8}$$

It follows that the local error at time t_{n+1} is given by

$$e^{n+1} = U_{\Delta t,n} e^n + T_{\Delta t,n}^{-1}\left(\epsilon_1^n + \epsilon_2^n\right).$$

Denoting

$$\mathcal{U}_{\Delta t}^{n,l} = \begin{cases} U_{\Delta t,n-1}\cdots U_{\Delta t,1}U_{\Delta t,0} & \text{for } l = 0 \\ U_{\Delta t,n-1}\cdots U_{\Delta t,l}T_{\Delta t,l-1}^{-1} & \text{for } l \in [\![1, n-1]\!] \\ T_{\Delta t,n-1}^{-1} & \text{for } l = n, \end{cases}$$

we get by induction the recursive formula for the global error

$$e^n = \mathcal{U}_{\Delta t}^{n,0} e^0 + \sum_{l=1}^n \mathcal{U}_{\Delta t}^{n,l}\left(\epsilon_1^{l-1} + \epsilon_2^{l-1}\right).$$

We are now interested in finding an estimate for the two quantities

$$\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^n \mathcal{U}_{\Delta t}^{n,l}\epsilon_1^{l-1}\right\|_{\mathbb{L}^2}^{2p}\right) \qquad \text{and} \qquad \mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^n \mathcal{U}_{\Delta t}^{n,l}\epsilon_2^{l-1}\right\|_{\mathbb{L}^2}^{2p}\right). \tag{4.2.9}$$

Since $e^0 = 0$, an estimate of the global error e^n in $L^{2p}\left(\Omega; L^\infty\left(0,T;\mathbb{L}^2\right)\right)$ of the scheme will follow. The proof combines arguments from mean-square approximation of stochastic differential equations and approximation of partial differential equations. Here, the main difference lies in the implicit discretization of the noise and the fact that the operator $\mathcal{U}_{\Delta t}^{n,l}$ is not $\mathcal{F}_{t_{l-1}}$ adapted, since it depends on the Brownian increments after time t_{l-1}. Let us write the remainder term ϵ_1^{l-1}, given in (4.2.8), as the sum of two terms $\epsilon_{1,1}^{l-1}$ and $\epsilon_{1,2}^{l-1}$. Writing that

$$\widetilde{X}^{l-1/2} = \widetilde{X}^{l-1} + \frac{\widetilde{X}^l - \widetilde{X}^{l-1}}{2} \tag{4.2.10}$$

and using Equation (4.2.6), we get that the expressions of $\epsilon_{1,1}^{l-1}$ and $\epsilon_{1,2}^{l-1}$ are given respectively by

$$\epsilon_{1,1}^{l-1} = iC_\gamma\left(W_{0,0}^{l-1,l}\left(\partial_x^4 X\right) - \frac{1}{2}W_0^{l-1,l}\left(\partial_x^4 X\right)\Delta t\right) \tag{4.2.11}$$

and

$$\epsilon_{1,2}^{l-1} = -i\sqrt{\gamma}\sum_{k=1}^3\left(W_{k,0}^{l-1,l}\left(\sigma_k\partial_x^3 X\right) - \frac{1}{2}W_k^{l-1,l}\left(\sigma_k\partial_x^3 X\right)\Delta t\right). \tag{4.2.12}$$

We proceed similarly for the term ϵ_2^{l-1} writing it as a sum of three terms $\epsilon_2^{l-1} = \epsilon_{2,1}^{l-1} + \epsilon_{2,2}^{l-1} + \epsilon_{2,3}^{l-1}$. Using again (4.2.10) and Equation (4.2.6), we obtain that

$$\begin{aligned}
&\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k\partial_x\widetilde{X}^{l-1/2}\Delta W_k^{l-1} \\
&\quad= \sqrt{\gamma}\sum_{k=1}^{3}W_k^{l-1}\left(\sigma_k\partial_x\widetilde{X}^{l-1}\right) + \frac{\sqrt{\gamma}}{2}C_\gamma\sum_{k=1}^{3}\sigma_k W_0^{l-1,l}\left(\partial_x^3 X\right)\Delta W_k^{l-1} \\
&\qquad- \frac{\gamma}{2}\sum_{j,k=1}^{3}\sigma_j\sigma_k W_j^{l-1,l}\left(\partial_x^2 X\right)\Delta W_k^{l-1}.
\end{aligned}$$

Therefore, the truncation error ϵ_2^{l-1}, given in Expression (4.2.8), can now be expressed thanks to

$$\begin{cases}
\epsilon_{2,1}^{l-1} = -\sqrt{\gamma}\sum_{k=1}^{3}W_k^{l-1,l}\left(\sigma_k\partial_x\widetilde{e}^{l-1}\right) \\
\epsilon_{2,2}^{l-1} = \frac{3\gamma}{2}W_0^{l-1,l}\left(\partial_x^2 X\right) - \frac{\gamma}{2}\sum_{j,k=1}^{3}\sigma_j\sigma_k W_j^{l-1,l}\left(\partial_x^2 X\right)\Delta W_k^{l-1} \\
\epsilon_{2,3}^{l-1} = \frac{\sqrt{\gamma}}{2}C_\gamma\sum_{k=1}^{3}\sigma_k W_0^{l-1,l}\left(\partial_x^3 X\right)\Delta W_k^{l-1}.
\end{cases} \tag{4.2.13}$$

To obtain estimates on the two quantities (4.2.9) , we will have to bound expectations of the form

$$\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\gamma^p\left\|\sum_{l=1}^{n}\mathcal{U}_{\Delta t}^{n,l}\sum_{k=1}^{3}W_{k,0}^{l-1,l}\left(\sigma_k\partial_x^3 X\right)\right\|_{\mathbb{L}^2}^{2p}\right).$$

We wish to write that

$$\mathcal{U}_{\Delta t}^{n,l}\sum_{k=1}^{3}W_{k,0}^{l-1,l}\left(\sigma_k\partial_x^3 X\right) = \sum_{k=1}^{3}W_{k,0}^{l-1,l}\left(\mathcal{U}_{\Delta t}^{n,l}\sigma_k\partial_x^3 X\right)$$

and apply the Burkholder-Davis-Gundy inequality to the stochastic integrals. Unfortunately, since $\mathcal{U}_{\Delta t}^{n,l}$ is not $\mathcal{F}_{l-1}$ adapted, the above equality does not hold. Therefore, we introduce the following process

$$\mathcal{V}_{\Delta t}^{l-1} = \begin{cases}
\mathrm{Id} & \text{for } l = 0,1 \\
T_{\Delta t,0}U_{\Delta t,1}^{-1}T_{\Delta t,1}^{-1} & \text{for } l = 2 \\
T_{\Delta t,0}U_{\Delta t,1}^{-1}\cdots U_{\Delta t,l-1}^{-1}T_{\Delta t,l-1}^{-1} & \text{for } l \in [\![3,n]\!],
\end{cases}$$

and separating the adapted part from the non-adapted part, we write

$$\mathcal{U}_{\Delta t}^{n,l} = \mathcal{U}_{\Delta t}^{n,1}\left(\mathcal{V}_{\Delta t}^{l-1} - \mathcal{V}_{\Delta t}^{l-2} + \mathcal{V}_{\Delta t}^{l-2}\right).$$

Now, using the contraction property of $\mathcal{U}_{\Delta t}^{n,1}$ in $\mathbb{L}^2$, we may write, for $q = 1, 2$,

$$\begin{aligned}
&\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\mathcal{U}_{\Delta t}^{n,l}\epsilon_q^{l-1}\right\|_{\mathbb{L}^2}^{2p}\right)\\
&\quad=\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\mathcal{U}_{\Delta t}^{n,1}\sum_{l=1}^{n}\left(\mathcal{V}_{\Delta t}^{l-1}-\mathcal{V}_{\Delta t}^{l-2}+\mathcal{V}_{\Delta t}^{l-2}\right)\epsilon_q^{l-1}\right\|_{\mathbb{L}^2}^{2p}\right)\\
&\quad\leqslant C_p\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\left(\mathcal{V}_{\Delta t}^{l-1}-\mathcal{V}_{\Delta t}^{l-2}\right)\epsilon_q^{l-1}\right\|_{\mathbb{L}^2}^{2p}\right)+C_p\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\mathcal{V}_{\Delta t}^{l-2}\epsilon_q^{l-1}\right\|_{\mathbb{L}^2}^{2p}\right).
\end{aligned}$$

Since $\mathcal{V}_{\Delta t}^{l-2}$ is $\mathcal{F}_{l-1}$ adapted, we are allowed to use the Burkholder-Davis-Gundy inequality. The next Lemma gives useful estimates to conclude the first part of the proof of Proposition 4.2.2.

Lemma 4.2.2. *For all* $\left(f^l\right)_{l\in[\![1,N]\!]}\in\left(\mathbb{H}^1\left(\mathbb{R}\right)\right)^N$ *and for all* $p>1$, *there exists a positive constant* $C(\gamma,T,p)$, *independent of* N, *such that*

$$\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\left(\mathcal{V}_{\Delta t}^{l-1}-\mathcal{V}_{\Delta t}^{l-2}\right)f^l\right\|_{\mathbb{L}^2}^{2p}\right)\leqslant C(\gamma,T,p)N^p\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\|f^n\|_{\mathbb{H}^1}^{4p}\right). \quad (4.2.14)$$

Moreover, if for any $l\in[\![1,N]\!]$, $f^l=\epsilon_q^{l-1}, q=1,2$, *then there exists two positive constants* C *and* C_1, *independent of* N, *such that*

$$\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\left(\mathcal{V}_{\Delta t}^{l-1}-\mathcal{V}_{\Delta t}^{l-2}\right)f^l\right\|_{\mathbb{L}^2}^{2p}\right)\leqslant C(\gamma,T,p,q,\|X_0\|_{\mathbb{H}^5})\Delta t^p. \quad (4.2.15)$$

and

$$\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\mathcal{V}_{\Delta t}^{l-2}f^l\right\|_{\mathbb{L}^2}^{2p}\right)\leqslant C_1(\gamma,T,p,q,\|X_0\|_{\mathbb{H}^4})\Delta t^p. \quad (4.2.16)$$

Proof of Lemma 4.2.2. The proof of this lemma is divided into two parts. In a first step, we prove inequality (4.2.14). The second step consists in proving estimate (4.2.16). By similar arguments, the bound (4.2.15) is easily obtained from (4.2.14).

Proof of estimate (4.2.14)**.** We begin this proof with a lemma stating that $\mathcal{V}_{\Delta t}^l$ is almost surely a bounded operator in $\mathbb{L}^2$ with a random continuity constant.

Lemma 4.2.3. *The random matrix operator* $\mathcal{V}_{\Delta t}^l$ *is almost surely a bounded operator in* $\mathbb{L}^2$ *and for any* $l=0,\cdots,n$,

$$\left\|\mathcal{V}_{\Delta t}^l f\right\|_{\mathbb{L}^2}\leqslant C_{0,l}\left(\omega\right)\|f\|_{\mathbb{L}^2},$$

where the random variable $C_{0,l}\left(\omega\right)$ *is integrable at any order such that its moments do not depend on* l.

Proof. By unitarity property of the matrices $U_{\Delta t,l}$, for any $l = 0, \cdots, n$, we know that $\left\|\left|U_{\Delta t,l}^{-1}\right|\right\|_2 = 1$. Therefore, by Plancherel theorem and Hölder inequality

$$\begin{aligned}\left\|\mathcal{V}_{\Delta t}^l f\right\|_{\mathbb{L}^2} &\leqslant \sup_{\xi\in\mathbb{R}} \left\|\left|\widehat{\mathcal{V}}_{\Delta t}^l(\xi)\right|\right\|_2 \|f\|_{\mathbb{L}^2} \\ &\leqslant \sup_{\xi\in\mathbb{R}} \left\{\left\|\left|\widehat{T}_{\Delta t,0}(\xi)\right|\right\|_\infty \left\|\left|\widehat{T}_{\Delta t,l}^{-1}(\xi)\right|\right\|_\infty\right\} \|f\|_{\mathbb{L}^2}.\end{aligned}$$

We claim that the random variable $\left\|\left|\widehat{T}_{\Delta t,0}\right|\right\|_\infty \left\|\left|\widehat{T}_{\Delta t,l}^{-1}\right|\right\|_\infty$ is almost surely bounded by a constant $C_{0,l}(\omega)$, independent of ξ, that is integrable at any order. In Fourier space, the inverse of the operator $T_{\Delta t,l}$ is given by

$$\widehat{T}_{\Delta t,l}^{-1}(\xi) = \frac{1}{2\det(l,\xi)} \begin{pmatrix} 2 + i\Delta t\,|\xi|^2 - i\sqrt{\gamma\Delta t}\chi_3^l\xi & -i\sqrt{\gamma\Delta t}\left(\chi_1^l - i\chi_2^l\right)\xi \\ -i\sqrt{\gamma\Delta t}\left(\chi_1^l + i\chi_2^l\right)\xi & 2 + i\Delta t\,|\xi|^2 + i\sqrt{\gamma\Delta t}\chi_3^l\xi \end{pmatrix},$$

where $\det(l,\xi) = 1 + \frac{\gamma\Delta t}{4}\sum_{k=1}^3 \left(\chi_k^l\right)^2 |\xi|^2 - \frac{\Delta t^2}{4}|\xi|^4 - i\Delta t\,|\xi|^2$. Therefore,

$$\begin{aligned}&\left\|\left|\widehat{T}_{\Delta t,0}(\xi)\right|\right\|_\infty \left\|\left|\widehat{T}_{\Delta t,l}^{-1}(\xi)\right|\right\|_\infty \\ &\quad= \frac{1}{4\left|\det(l,\xi)\right|}\left(\left|2 + i\Delta t\,|\xi|^2 - i\sqrt{\gamma\Delta t}\chi_3^l\xi\right| + \left|\sqrt{\gamma\Delta t}\left(\chi_1^l - i\chi_2^l\right)\xi\right|\right) \\ &\quad\times\left(\left|2 + i\Delta t\,|\xi|^2 + i\sqrt{\gamma\Delta t}\chi_3^0\xi\right| + \left|\sqrt{\gamma\Delta t}\left(\chi_1^0 - i\chi_2^0\right)\xi\right|\right).\end{aligned}$$

Applying the triangular inequality, we obtain the next bound

$$\begin{aligned}&\left\|\left|\widehat{T}_{\Delta t,0}(\xi)\right|\right\|_\infty \left\|\left|\widehat{T}_{\Delta t,l}^{-1}(\xi)\right|\right\|_\infty \\ &\leqslant \frac{1}{4\left|\det(l,\xi)\right|}\left(4 + 4\Delta t\,|\xi|^2 + \Delta t^2|\xi|^4 + 2\sqrt{\gamma\Delta t}\,|\xi|\sum_{k=1}^3\left(\left|\chi_k^0\right| + \left|\chi_k^l\right|\right)\right) \\ &+ \frac{1}{4\left|\det(l,\xi)\right|}\left(\sqrt{\gamma\Delta t}\Delta t\,|\xi|^3\sum_{k=1}^3\left(\left|\chi_k^0\right| + \left|\chi_k^l\right|\right) + \gamma\Delta t\,|\xi|^2\sum_{k=1}^3\left|\chi_k^0\right|\sum_{k=1}^3\left|\chi_k^l\right|\right).\end{aligned} \tag{4.2.17}$$

Denoting $x = \Delta t^{1/2}|\xi|$ and $y = \sum_{k=1}^3\left(\chi_k^l\right)^2$, we define the mapping f from $\mathbb{R}_+^2$ into $\mathbb{R}_+$

$$f(x,y) = \frac{x^4/4}{\left(\left(1 + \frac{\gamma x^2}{4}y - \frac{x^4}{4}\right)^2 + x^4\right)^{1/2}}.$$

We can prove that there exists a positive constant C, such that for any $y \in \mathbb{R}_+$ and any $x \in \mathbb{R}_+$

$$f(x,y) < \sqrt{C}\max\left(\frac{\gamma y}{4}, 1\right).$$

We may proceed similarly for the other terms in the right hand side in (4.2.17). Therefore, the right hand side in (4.2.17) is uniformly bounded in ξ by a polynomial function of y, $|\chi_k^0|$ and $\left|\chi_k^l\right|$. Therefore, its moments do not depend on l. □

We now state a Lemma giving an estimate of the local error between the unbounded random operator $T_{\Delta t,n-1}U^{-1}_{\Delta t,n}T^{-1}_{\Delta t,n}$ and the identity mapping.

Lemma 4.2.4. *For any n and $p \in \mathbb{N}$, there exists a positive random constant $C_{n-1,n}(\omega) < +\infty$ a.s. belonging to $L^{2p}(\Omega)$ for any $f \in \mathbb{H}^1$*

$$\left\| \left[T_{\Delta t,n-1}U^{-1}_{\Delta t,n}T^{-1}_{\Delta t,n} - Id \right] f \right\|_{\mathbb{L}^2} \leqslant C_{n-1,n}(\omega)\sqrt{\Delta t}\, \|f\|_{\mathbb{H}^1} \qquad a.s.$$

Moreover, the moments of the random variable $C_{n-1,n}$ are independent of n.

Proof. Let us recall that the determinant of $\left(\mathrm{Id} - \frac{1}{2}\widehat{H}_{\Delta t,n}(\xi)\right)$ is given by

$$\det(n,\xi) = 1 + \frac{\gamma\Delta t}{4}\sum_{k=1}^{3}(\chi^n_k)^2\,|\xi|^2 - \frac{\Delta t^2}{4}\,|\xi|^4 - i\Delta t\,|\xi|^2$$

and that the following relation holds

$$\widehat{T}_{\Delta t,n-1}\widehat{U}^{-1}_{\Delta t,n}\widehat{T}^{-1}_{\Delta t,n} - \mathrm{Id} = \left(\frac{1}{2}\widehat{H}_{\Delta t,n-1} + \frac{1}{2}\widehat{H}_{\Delta t,n}\right)\left(\mathrm{Id} - \frac{1}{2}\widehat{H}_{\Delta t,n}(\xi)\right)^{-1}.$$

A computation gives

$$\widehat{H}_{\Delta t,n}(\xi)\left(\mathrm{Id} - \frac{1}{2}\widehat{H}_{\Delta t,n}(\xi)\right)^{-1} = \frac{1}{\det(n,\xi)}\begin{pmatrix} h_+(\Delta t,\xi,\omega) & g_-(\Delta t,\xi,\omega) \\ g_+(\Delta t,\xi,\omega) & h_-(\Delta t,\xi,\omega)\end{pmatrix}$$

where

$$\begin{aligned} h_\pm(\Delta t,\xi,\omega) &= i\Delta t\xi^2 - \frac{\gamma\Delta t}{2}\sum_{k=1}^{3}(\chi^n_k)^2\,\xi^2 + \frac{\Delta t^2}{2}\xi^4 \pm i\sqrt{\gamma\Delta t}\chi^n_3\xi \\ g_\pm(\Delta t,\xi,\omega) &= i\sqrt{\gamma\Delta t}\,(\chi^n_1 \pm i\chi^n_2)\,\xi. \end{aligned}$$

From the proof of Lemma 4.2.3, we obtain that there exists a random variable $C_n(\omega)$, integrable at any order, such that

$$\sup_{\xi\in\mathbb{R}}\frac{\Delta t^{3/2}\,|\xi|^3}{|\det(n,\xi)|} \leqslant C_n(\omega).$$

We easily deduce, from the above estimate, that

$$\sup_{\xi\in\mathbb{R}}\frac{1}{\left(1+|\xi|^2\right)^{1/2}}\frac{\Delta t^2\,|\xi|^4}{|\det(n,\xi)|} \leqslant C_n(\omega)\sqrt{\Delta t}.$$

Hence,

$$\sup_{\xi\in\mathbb{R}}\frac{1}{\left(1+|\xi|^2\right)^{1/2}}\left\|\left|\widehat{T}_{\Delta t,n-1}\widehat{U}^{-1}_{\Delta t,n}\widehat{T}^{-1}_{\Delta t,n} - \mathrm{Id}\right|\right\|_\infty \leqslant C_{n-1,n}(\omega)\sqrt{\Delta t},$$

where $C_{n-1,n}(\omega)$ is a polynomial function of $\sum_{k=1}^{3}(\chi^n_k)^2$, $\left|\chi^{n-1}_k\right|$ and $|\chi^n_k|$. Hence, the moments of $C_{n-1,n}(\omega)$ are independent of n. □

Now, we prove estimate (4.2.14). For any $l \geqslant 2$

$$\begin{aligned}\mathcal{V}_{\Delta t}^{l-1} - \mathcal{V}_{\Delta t}^{l-2} &= \mathcal{V}_{\Delta t}^{l-2}\left[T_{\Delta t,l-2}U_{\Delta t,l-1}^{-1}T_{\Delta t,l-1}^{-1} - \mathrm{Id}\right] \\ &= \mathcal{V}_{\Delta t}^{l-2}\left[\left(\mathrm{Id} + \frac{1}{2}H_{\Delta t,l-2}\right)\left(\mathrm{Id} - \frac{1}{2}H_{\Delta t,l-1}\right)^{-1} - \mathrm{Id}\right].\end{aligned}$$

Therefore, applying Lemma 4.2.3,

$$\begin{aligned}&\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\left(\mathcal{V}_{\Delta t}^{l-1} - \mathcal{V}_{\Delta t}^{l-2}\right)f^l\right\|_{\mathbb{L}^2}^{2p}\right) \\ &\leqslant N^{2p}\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\mathcal{V}_{\Delta t}^{n-2}\left[T_{\Delta t,n-2}U_{\Delta t,n-1}^{-1}T_{\Delta t,n-1}^{-1} - \mathrm{Id}\right]f^n\right\|_{\mathbb{L}^2}^{2p}\right) \\ &\leqslant N^{2p}\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left(C_{0,n-2}\right)^{2p}\left\|\left[T_{\Delta t,n-2}U_{\Delta t,n-1}^{-1}T_{\Delta t,n-1}^{-1} - \mathrm{Id}\right]f^n\right\|_{\mathbb{L}^2}^{2p}\right).\end{aligned}$$

From Lemma 4.2.4, we deduce that,

$$\begin{aligned}&\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\left(\mathcal{V}_{\Delta t}^{l-1} - \mathcal{V}_{\Delta t}^{l-2}\right)f^l\right\|_{\mathbb{L}^2}^{2p}\right) \\ &\leqslant N^{2p}\left(\Delta t\right)^p\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left(C_{0,n-2}\right)^{2p}\left(C_{n-2,n-1}\right)^{2p}\left\|f^n\right\|_{\mathbb{H}^1}^{2p}\right).\end{aligned}$$

We obtain, using the Cauchy-Schwarz inequality, that

$$\begin{aligned}&\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\left(\mathcal{V}_{\Delta t}^{l-1} - \mathcal{V}_{\Delta t}^{l-2}\right)f^l\right\|_{\mathbb{L}^2}^{2p}\right) \\ &\leqslant T^pN^p\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left(C_{0,n-2}\right)^{4p}\left(C_{n-2,n-1}\right)^{4p}\right)^{1/2}\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|f^n\right\|_{\mathbb{H}^1}^{4p}\right)^{1/2}.\end{aligned}$$

Applying again the Cauchy-Schwarz inequality and the Burkholder-Davis-Gundy inequality and using the fact that the moments of $C_{0,n-2}$ and $C_{n-2,n-1}$ are independent of n, we obtain that $\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left(C_{0,n-2}\right)^{4p}\left(C_{n-2,n-1}\right)^{4p}\right)$ is independent of n. Thus, the following inequality holds

$$\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\left(\mathcal{V}_{\Delta t}^{l-1} - \mathcal{V}_{\Delta t}^{l-2}\right)f^l\right\|_{\mathbb{L}^2}^{2p}\right) \leqslant C(\gamma,T,p)N^p\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|f^n\right\|_{\mathbb{H}^1}^{4p}\right)^{1/2}.$$

Proof of estimate (4.2.16) **for** $q = 1$**.** Writing $\widetilde{X}(s) = \widetilde{X}^{l-1} + \widetilde{e}^{l-1}$, we rewrite $\epsilon_{1,1}^{l-1}$, given in (4.2.11), as follows

$$\epsilon_{1,1}^{l-1} = iC_\gamma\left(W_{0,0}^{l-1,l}\left(\partial_x^4\,\widetilde{e}^{l-1}\right) - \frac{1}{2}W_0^{l-1,l}\left(\partial_x^4\,\widetilde{e}^{l-1}\right)\Delta t\right).$$

We focus on the first term in the above expression, the other term being bounded in a similar way. Using the Minkowski inequality, the contraction property of $\mathcal{U}_{\Delta t}^{n,l}$ in $\mathbb{L}^2(\mathbb{R})$ for every $l \in [\![1,n]\!]$ and the conservation of the $\mathbb{L}^2$ norm, we get

$$\begin{aligned}
&\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\mathcal{U}_{\Delta t}^{n,l}W_{0,0}^{l-1,l}(\partial_x^4\widetilde{e}^{l-1})\right\|_{\mathbb{L}^2}^{2p}\right)\\
&\leqslant \mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left(\sum_{l=1}^{n}\left\|W_{0,0}^{l-1,l}\left(\partial_x^4\widetilde{e}^{l-1}\right)\right\|_{\mathbb{L}^2}\right)^{2p}\right)\\
&\leqslant \mathbb{E}\left(\left(\sum_{l=1}^{N}\sup_{t_{l-1}\leqslant u\leqslant t_l}\left\|\partial_x^4\widetilde{e}^{l-1}(u)\right\|_{\mathbb{L}^2}\frac{\Delta t^2}{2}\right)^{2p}\right)\\
&\leqslant C\left(\|X_0\|_{\mathbb{H}^4}^{2p}\right)T^{2p}\Delta t^{2p}\left(1+\gamma^p\right).
\end{aligned}$$

For the term $\epsilon_{1,2}^{l-1}$, whose expression is given in (4.2.12), let us notice that after integration by part, we have the following equality

$$\begin{aligned}
W_{k,0}^{l-1,l}(1) &= W_k(t_l)t_l - W_k(t_{l-1})t_{l-1} - W_{0,k}^{l-1,l}(1) - \Delta W_k^{l-1}t_{l-1} - W_k(t_{l-1})\Delta t\\
&= \Delta W_k^{l-1}\Delta t - W_{0,k}^{l-1,l}(1).
\end{aligned}$$

Hence, the second term $\epsilon_{1,2}^{l-1}$ can be written as

$$\begin{aligned}
\epsilon_{1,2}^{l-1} &= -i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k\partial_x^3\widetilde{X}^{l-1}\left(\frac{1}{2}\Delta W_k^{l-1}\Delta t - W_{0,k}^{l-1,l}(1)\right) \qquad (4.2.18)\\
&\quad -i\sqrt{\gamma}\sum_{k=1}^{3}\left(W_{0,k}^{l-1,l}\left(\sigma_k\partial_x^3\widetilde{e}^{l-1}\right) - \frac{1}{2}W_k^{l-1,l}\left(\sigma_k\partial_x^3\widetilde{e}^{l-1}\right)\Delta t\right).
\end{aligned}$$

We recall that we wish to estimate the expectation $\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\mathcal{V}_{\Delta t}^{l-2}\epsilon_{1,2}^{l-1}\right\|_{\mathbb{L}^2}^{2p}\right)$. In the above expression (4.2.18), all the terms may be bounded using similar arguments. Thus, we only do the computation for

$$\mathcal{V}_{\Delta t}^{l-2}\sigma_k\partial_x^3\widetilde{X}^{l-1}W_{0,k}^{l-1,l}(1) = W_{0,k}^{l-1,l}\left(\mathcal{V}_{\Delta t}^{l-2}\sigma_k\partial_x^3\widetilde{X}^{l-1}\right),$$

since $\mathcal{V}_{\Delta t}^{l-2}\sigma_k\partial_x^3\widetilde{X}^{l-1}$ is $\mathcal{F}_{l-1}$ adapted. By orthogonality of the increments of the three dimensional Brownian Motion,

$$\mathbb{E}\left(W_{0,k}^{l-1,l}\left(\mathcal{V}_{\Delta t}^{l-2}\sigma_k\partial_x^3\widetilde{X}^{l-1}\right)W_{0,j}^{l'-1,l'}\left(\mathcal{V}_{\Delta t}^{l'-2}\sigma_j\partial_x^3\widetilde{X}^{l'-1}\right)\right) = 0 \qquad \text{if } l\neq l' \text{ or } k\neq j.$$

Hence, we obtain that

$$\ll\sum_{l=1}^{n}\sum_{k=1}^{3}W_{0,k}^{l-1,l}\left(\mathcal{V}_{\Delta t}^{l-2}\sigma_k\partial_x^3\widetilde{X}^{l-1}\right)\gg=\sum_{l=1}^{n}\sum_{k=1}^{3}\ll W_{0,k}^{l-1,l}\left(\mathcal{V}_{\Delta t}^{l-2}\sigma_k\partial_x^3\widetilde{X}^{l-1}\right)\gg,$$

where $\ll . \gg$ denotes the quadratic variation process. Thanks to the conservation of the $\mathbb{H}^1$ norm, the solution X of Equation (4.2.1) has all its moments bounded in $\mathbb{H}^1$ and the stochastic integral is a true martingale. Thus, applying the Burkholder-Davis-Gundy inequality and Cauchy-Schwarz inequality

$$\begin{aligned}
&\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\gamma^p\left\|\sum_{l=1}^{n}\sum_{k=1}^{3}W_{0,k}^{l-1,l}\left(\mathcal{V}_{\Delta t}^{l-2}\sigma_k\partial_x^3\widetilde{X}^{l-1}\right)\right\|_{\mathbb{L}^2}^{2p}\right)\\
&\leqslant C\mathbb{E}\left(\gamma^p\left(\sum_{l=1}^{N}\sum_{k=1}^{3}\int_{t_{l-1}}^{t_l}\left\|W_0^{l-1,s}\left(\mathcal{V}_{\Delta t}^{l-2}\sigma_k\partial_x^3\widetilde{X}^{l-1}\right)\right\|_{\mathbb{L}^2}^2 ds\right)^p\right)\\
&\leqslant CT^p\gamma^p\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\sup_{s\in[t_{n-1},t_n]}\sum_{k=1}^{3}\left\|W_0^{n-1,s}\left(\mathcal{V}_{\Delta t}^{n-2}\sigma_k\partial_x^3\widetilde{X}^{n-1}\right)\right\|_{\mathbb{L}^2}^{2p}\right).
\end{aligned}$$

Applying Lemma 4.2.3 and again the Cauchy-Schwarz inequality, we get

$$\begin{aligned}
&\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\sup_{s\in[t_{n-1},t_n]}\sum_{k=1}^{3}\left\|W_0^{n-1,s}\left(\mathcal{V}_{\Delta t}^{n-2}\sigma_k\partial_x^3\widetilde{X}^{n-1}\right)\right\|_{\mathbb{L}^2}^{2p}\right)\\
&\leqslant \Delta t^{2p}\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left[C_{0,n-2}\right]^{2p}\left\|\partial_x^3\widetilde{X}^{n-1}\right\|_{\mathbb{L}^2}^{2p}\right)\\
&\leqslant \Delta t^{2p}\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left[C_{0,n-2}\right]^{4p}\right)^{1/2}\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\partial_x^3\widetilde{X}^{n-1}\right\|_{\mathbb{L}^2}^{4p}\right)^{1/2}.
\end{aligned}$$

Hence, from the conservation of the $\mathbb{H}^1$ norm, it follows that

$$\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\gamma^p\left\|\sum_{l=1}^{n}\sum_{k=1}^{3}W_{0,k}^{l-1,l}\left(\mathcal{V}_{\Delta t}^{l-2}\sigma_k\partial_x^3\widetilde{X}^{l-1}\right)\right\|_{\mathbb{L}^2}^{2p}\right)\leqslant C\left(\gamma,T,\|X_0\|_{\mathbb{H}^3}\right)\Delta t^{2p}.$$

Collecting the above estimates leads to estimate (4.2.16) for $q=1$.

Proof of estimate (4.2.16) **for** $q=2$**.** The first and second terms $\epsilon_{2,1}^{l-1}$ and $\epsilon_{2,2}^{l-1}$ in (4.2.13) will give the order of convergence of the scheme. The third one $\epsilon_{2,3}^{l-1}$ may be bounded similarly as in the previous step. To bound $\epsilon_{2,1}^{l-1}$, we may use again the independence of the increments of the Brownian Motion, Burkholder-Davis-Gundy inequality, Lemma 4.2.1 and 4.2.3 and Cauchy-Schwarz inequality

$$\begin{aligned}
&\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\gamma^p\left\|\sum_{l=1}^{n}\sum_{k=1}^{3}W_{k}^{l-1,l}\left(\mathcal{V}_{\Delta t}^{l-2}\sigma_k\partial_x\widetilde{e}^{l-1}\right)\right\|_{\mathbb{L}^2}^{2p}\right)\\
&\leqslant C\gamma^p\mathbb{E}\left(\left(\sum_{l=1}^{N}\sum_{k=1}^{3}W_0^{l-1,l}\left(\left\|\mathcal{V}_{\Delta t}^{l-2}\sigma_k\partial_x\widetilde{e}^{l-1}\right\|_{\mathbb{L}^2}^2\right)\right)^p\right)\\
&\leqslant C\left\|X_0\right\|_{\mathbb{H}^2}^{2p}\gamma^p\left(1+\gamma^p\right)T^p\Delta t^p.
\end{aligned}$$

We conclude the proof obtaining an estimate for $\epsilon_{2,2}^{l-1}$. Writing again $\widetilde{X}(s) = \widetilde{X}^{l-1} + \widetilde{e}^{l-1}(s)$ and using Equation (4.2.6), we obtain

$$\begin{aligned}\epsilon_{2,2}^{l-1} &= \frac{3\gamma}{2}\frac{\partial^2 \widetilde{X}^{l-1}}{\partial x^2}\Delta t - \frac{\gamma}{2}\sum_{j,k=1}^{3}\sigma_j\sigma_k\frac{\partial^2 \widetilde{X}^{l-1}}{\partial x^2}\Delta W_j^{l-1}\Delta W_k^{l-1} \\ &\quad + \frac{3\gamma}{2}W_0^{l-1,l}\left(\partial_x^2\widetilde{e}^{l-1}\right) - \frac{\gamma}{2}\sum_{j,k=1}^{3}W_j^{l-1,l}\left(\sigma_j\sigma_k\partial_x^2\widetilde{e}^{l-1}\right)\Delta W_k^{l-1}.\end{aligned}$$

By Property 4.1.1 of the Pauli matrices, the following identity holds

$$\frac{\gamma}{2}\sum_{j,k=1}^{3}\sigma_j\sigma_k\Delta W_j^{l-1}\Delta W_k^{l-1} = \frac{\gamma}{2}\sum_{k=1}^{3}\left(\Delta W_k^{l-1}\right)^2. \tag{4.2.19}$$

Moreover, for any $t_{l-1} \leqslant s \leqslant t$,

$$\mathbb{E}\left(\left(W_k^{l-1,t}(1)\right)^4\middle|\mathcal{F}_s\right) = \left(W_k^{l-1,s}(1)\right)^4 + 3(t-s)^2 + 6\left(W_k^{l-1,s}(1)\right)^2(t-s).$$

It follows that

$$M_t = \left(W_k^{l-1,t}(1)\right)^4 - 6(t-t_{l-1})\left(W_k^{l-1,t}(1)\right)^2 + 3(t-t_{l-1})^2$$

is a $\mathcal{F}_t$ martingale. Therefore, the quadratic variation of the $\mathcal{F}_t$ martingale $\left(W_k^{l-1,t}(1)\right)^2 - (t-t_{l-1})$ is given by

$$\begin{aligned}\ll \left(W_k^{l-1,t}(1)\right)^2 - (t-t_{l-1}) \gg &= \left(\left(W_k^{l-1,t}(1)\right)^2 - (t-t_{l-1})\right)^2 - M_t \\ &= 4(t-t_{l-1})\left(W_k^{l-1,t}(1)\right)^2 - 2(t-t_{l-1})^2.\end{aligned}$$

Consequently, taking $t = t_l$ leads to

$$\ll \left(W_k^{l-1,l}(1)\right)^2 - \Delta t \gg = 4\Delta t\left(W_k^{l-1,l}(1)\right)^2 - 2\Delta t^2.$$

We deduce, applying the Burkholder-Davis-Gundy inequality and using equality (4.2.19), that

$$\begin{aligned}&\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\mathcal{V}_{\Delta t}^{l-2}\left(\frac{3\gamma}{2}\frac{\partial^2 \widetilde{X}^{l-1}}{\partial x^2}\Delta t - \frac{\gamma}{2}\sum_{j,k=1}^{3}\sigma_j\sigma_k\frac{\partial^2 \widetilde{X}^{l-1}}{\partial x^2}\Delta W_j^{l-1}\Delta W_k^{l-1}\right)\right\|_{\mathbb{L}^2}^{2p}\right) \\ &\leqslant \frac{\gamma^{2p}}{4^p}\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\sum_{k=1}^{3}\mathcal{V}_{\Delta t}^{l-2}\frac{\partial^2 \widetilde{X}^{l-1}}{\partial x^2}\left(\Delta t - \left(\Delta W_k^{l-1}\right)^2\right)\right\|_{\mathbb{L}^2}^{2p}\right) \\ &\leqslant C\gamma^{2p}\mathbb{E}\left(\left(\sum_{l=1}^{N}\sum_{k=1}^{3}\left\|\mathcal{V}_{\Delta t}^{l-2}\frac{\partial^2 \widetilde{X}^{l-1}}{\partial x^2}\left(4\Delta t\left(\Delta W_k^{l-1}\right)^2 - 2\Delta t^2\right)\right\|_{\mathbb{L}^2}\right)^p\right).\end{aligned}$$

Moreover, applying Lemma 4.2.3 and using the conservation of the $\mathbb{H}^1$ norm, we obtain that

$$\begin{aligned}
&\mathbb{E}\left(\left(\sum_{l=1}^{N}\sum_{k=1}^{3}\left\|\mathcal{V}_{\Delta t}^{l-2}\frac{\partial^2\widetilde{X}^{l-1}}{\partial x^2}\left(4\Delta t\left(\Delta W_k^{l-1}\right)^2-2\Delta t^2\right)\right\|_{\mathbb{L}^2}\right)^p\right)\\
&\quad\leqslant \|X_0\|_{\mathbb{H}^2}^{2p}\,\mathbb{E}\left(\left(\sum_{l=1}^{N}\sum_{k=1}^{3}C_{0,l-2}\left|4\Delta t\left(\Delta W_k^{l-1}\right)^2-2\Delta t^2\right|\right)^p\right)\\
&\quad\leqslant \|X_0\|_{\mathbb{H}^2}^{2p}\,N^{p-1}\sum_{l=1}^{N}\sum_{k=1}^{3}\mathbb{E}\left(\left(C_{0,l-2}\right)^p\left|4\Delta t\left(\Delta W_k^{l-1}\right)^2-2\Delta t^2\right|^p\right).
\end{aligned}$$

Finally, we conclude applying the Cauchy-Schwarz inequality that

$$\begin{aligned}
&\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}\mathcal{V}_{\Delta t}^{l-2}\left(\frac{3\gamma}{2}\frac{\partial^2\widetilde{X}^{l-1}}{\partial x^2}\Delta t-\frac{\gamma}{2}\sum_{j,k=1}^{3}\sigma_j\sigma_k\frac{\partial^2\widetilde{X}^{l-1}}{\partial x^2}\Delta W_j^{l-1}\Delta W_k^{l-1}\right)\right\|_{\mathbb{L}^2}^{2p}\right)\\
&\quad\leqslant C\,\|X_0\|_{\mathbb{H}^2}^{2p}\,\gamma^{2p}T^p\Delta t^p.
\end{aligned}$$

The last term $\epsilon_{2,3}^n$ in (4.2.13) may be bounded similarly as $\epsilon_{1,2}^n$. □

Collecting the previous bounds, we obtain that the scheme has a global strong rate of convergence of order $\sqrt{\Delta t}$. We now aim at proving an almost sure order using the previous result. Let $\delta>0$ and set

$$A_N=\left\{\max_{n\in[\![1,N]\!]}N^{\delta}\left\|X^n-\widetilde{X}^n\right\|_{\mathbb{L}^2}>1\right\}$$

Using the $L^{2p}(\Omega)$ order given by (4.2.4) and the Tchebychev inequality, we get for any $p\geqslant 1$

$$\mathbb{P}(A_N)\leqslant N^{2p\delta}C\left(\|X_0\|_{\mathbb{H}^5},T,p,\gamma\right)\Delta t^p.$$

The series of general terms $N^{p(2\delta-1)}$ converges if $\delta<\frac{1}{2}-\frac{1}{2p}$. The Borel-Cantelli lemma then says that

$$\mathbb{P}\left(\limsup_{N\to+\infty}A_N\right)=0.$$

Therefore, there almost surely exists a N_0 such that for all $N\geqslant N_0(\delta)$

$$\max_{n\in[\![1,N]\!]}N^{\delta}\left\|\widetilde{e}^n\right\|_{\mathbb{L}^2}\leqslant 1.$$

Hence, there exists a constant M, depending on p,T,γ,δ,ω and $\|X_0\|_{\mathbb{H}^5}$, such that

$$\max_{n\in[\![1,N]\!]}N^{\delta}\left\|\widetilde{e}^n\right\|_{\mathbb{L}^2}\leqslant M\left(p,T,\gamma,\delta,\|X_0\|_{\mathbb{H}^5},\omega\right)\qquad\text{a.s for all } N.$$

□

4.3 The Crank-Nicolson scheme

In this section, we investigate the order in probability and the almost sure order for the Crank-Nicolson scheme (4.1.3) as an approximation of Equation (4.1.1). The notion of order in probability is not usual in the context of numerical analysis of stochastic equation. This notion is weaker than the strong order in time and is used here because of the nonlinear drift. Let us denote by $U(t,s), t \geqslant s, t, s \in \mathbb{R}_+$ the random unitary propagator defined as the unique solution of the linear equation of (4.1.1)

$$idX(t) + \frac{\partial^2 X(t)}{\partial x^2}dt + i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k \frac{\partial X(t)}{\partial x} \circ dW_k(t) = 0, \quad t \geqslant 0, x \in \mathbb{R},$$

that exists (see Chapter 3). Then, Equation (4.1.1), with initial condition $X_0 = v$, can be written in its mild form

$$X(t) = U(t,0)v + i\int_0^t U(t,s)F(X(s))ds. \tag{4.3.1}$$

We introduce a cut-off function $\Theta \in C_c^\infty(\mathbb{R})$, $\Theta \geqslant 0$ satisfying $\Theta(x) = 1$ for $x \in [0,1]$ and $\Theta(x) = 0$ for $x \geqslant 2$. We then define $\Theta_R(.) = \Theta(\|.\|_{\mathbb{H}^1}/R)$ for any $R \in \mathbb{N}^*$ and introduce the cut-off equation

$$X_R(t) = U(t,0)v + i\int_0^t \Theta_R^2(X_R(s))\, U(t,s)F(X_R(s))ds, \tag{4.3.2}$$

which is the mild formulation of the equation

$$idX_R(t) + \left(\frac{\partial^2 X_R(t)}{\partial x^2} + \Theta_R^2(X_R(t))\, F(X_R(t))\right)dt + i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k\frac{\partial X_R(t)}{\partial x}\circ dW_k(t) = 0. \tag{4.3.3}$$

This section is organized in two parts. We first define a cut-off scheme, as an approximation of the continuous cut-off equation (4.3.3), and prove existence and uniqueness of a global solution to this scheme. We also prove that the strong mean-square rate of convergence of this approximation to the continuous cut-off equation is 1/2. From this solution, we construct a discrete solution to the Crank Nicolson scheme (4.1.3) and define a discrete blow-up time. Using the time order for the cut off scheme, we obtain a probability order and a.s. order for the discrete scheme (4.1.3).

4.3.1 A truncated scheme : existence and strong order

Let us consider a cut-off version of the scheme (4.1.3)

$$X_R^{n+1} = X_R^n - H_{\Delta t,n}X_R^{n+1/2} + i\Delta t\Theta_{X_R}^{n,n+1}F\left(X_R^n, X_R^{n+1}\right) \tag{4.3.4}$$

where $\Theta_{X_R}^{n,n+1} = \Theta_R(X_R^n)\,\Theta_R\left(X_R^{n+1}\right)$. Recall that the nonlinear function F is given by

$$F\left(X_R^n, X_R^{n+1}\right) = \frac{4}{9}\left(|X_R^n|^2 + \left|X_R^{n+1}\right|^2\right)X_R^{n+1/2}.$$

This discretization of the nonlinear term is consistent with the continuous equation (4.3.3). Indeed replacing X_R^n and X_R^{n+1} by the continuous solution X_R in the discrete nonlinear term $\Theta_{X_R}^{n,n+1} F\left(X_R^n, X_R^{n+1}\right)$, we obtain the term $\Theta_R^2\left(X_R\right) F\left(X_R\right)$ appearing in (4.3.3). Now, we state in the next Proposition an existence and convergence result for the scheme (4.3.4). This will be useful to define a solution, up to the blow-up time, for (4.1.3) and a rate of convergence in a sense that should be specified.

Proposition 4.3.1. *Let $X_0 \in \mathbb{H}^1$. Then there exists a unique adapted discrete solution $X_R^N = (X_R^n)_{n=0,\cdots,N}$ to (4.3.4) that belongs to $L^\infty\left(0,T;\mathbb{H}^1\right)$. Furthermore for any $n \in \mathbb{N}$ such that $n \leqslant N$, the $\mathbb{L}^2$ norm is almost surely preserved i.e $\|X_R^n\|_{\mathbb{L}^2} = \|X_0\|_{\mathbb{L}^2}$.*

Proof. Assume that $X_0 \in \mathbb{H}^1$, the integral formulation of the cut-off scheme (4.3.4) is then given by

$$X_R^n = \mathcal{U}_{\Delta t}^{n,0} X_0 + i\Delta t \sum_{l=1}^{n} \mathcal{U}_{\Delta t}^{n,l} \Theta_{X_R}^{l-1,l} F\left(X_R^{l-1}, X_R^l\right), \tag{4.3.5}$$

where $\mathcal{U}_{\Delta t}^{n,l}$ is the discrete random propagator solution of the linear equation (4.2.2). We denote by $\mathcal{T} X_R^n$ the right hand side of the previous equation and we will prove that $\mathcal{T}$ is a contraction mapping in $L^\infty\left(0,T;\mathbb{H}^1\right)$. The proof is similar to the one in the continuous case [17, 22] . Using that $\mathbb{H}^1(\mathbb{R})$ is an algebra and the fact that $\mathcal{U}_{\Delta t}^{n,0}$ maps $\mathbb{H}^1$ into $L^\infty\left(0,T;\mathbb{H}^1\right)$, we easily get that $\mathcal{T}$ maps $L^\infty\left(0,T;\mathbb{H}^1\right)$ into itself a.s. Let us now assume that Y_R^N and X_R^N belong to $L^\infty\left(0,T;\mathbb{H}^1\right)$. We set as in [17, 22]

$$\tau_R^1 = \inf\left\{n\Delta t \leqslant T, \|X_R^n\|_{\mathbb{H}^1} \geqslant 2R\right\}.$$

We define in the same way the stopping time τ_R^2 associated with Y_R^N. Using the isometry property of the operator $\mathcal{U}_{\Delta t}^{n,0}$ for $n = 0, \cdots, N$, we get

$$\begin{aligned}
&\left\|\mathcal{T} X_R^N - \mathcal{T} Y_R^N\right\|_{L^\infty(0,T;\mathbb{H}^1)} \\
&\quad\leqslant \Delta t \left\| \sum_{l=1}^{n} \mathcal{U}_{\Delta t}^{n,l} \left(\Theta_{X_R}^{l-1,l} F\left(X_R^{l-1}, X_R^l\right) - \Theta_{Y_R}^{l-1,l} F\left(Y_R^{l-1}, Y_R^l\right)\right)\right\|_{L^\infty(0,T;\mathbb{H}^1)} \\
&\quad\leqslant T \sup_{\substack{n\in\mathbb{N}^* \\ n\Delta t\leqslant T}} \left\|\Theta_{X_R}^{n-1,n} F\left(X_R^{n-1}, X_R^n\right) - \Theta_{Y_R}^{n-1,n} F\left(Y_R^{n-1}, Y_R^n\right)\right\|_{\mathbb{H}^1}.
\end{aligned}$$

Assuming then that $\tau_R^1 \leqslant \tau_R^2$ and writing $[0,T] = [0,\tau_R^1) \cup [\tau_R^1, \tau_R^2) \cup [\tau_R^2, T]$, we get

$$\begin{aligned}
\left\|\mathcal{T} X_R^N - \mathcal{T} Y_R^N\right\|_{L^\infty(0,T;\mathbb{H}^1)} \leqslant\ & T \sup_{\substack{n\in\mathbb{N}^* \\ n\Delta t<\tau_R^1}} \left\|\left(\Theta_{X_R}^{n-1,n} - \Theta_{Y_R}^{n-1,n}\right) F\left(X_R^{n-1}, X_R^n\right)\right\|_{\mathbb{H}^1} \\
&+T \sup_{\substack{n\in\mathbb{N}^* \\ n\Delta t<\tau_R^1}} \left\|\Theta_{Y_R}^{n-1,n}\left(F\left(X_R^{n-1}, X_R^n\right) - F\left(Y_R^{n-1}, Y_R^n\right)\right)\right\|_{\mathbb{H}^1} \\
&+T \sup_{\substack{n\in\mathbb{N}^* \\ \tau_R^1\leqslant n\Delta t<\tau_R^2}} \left\|\Theta_{Y_R}^{n-1,n} F\left(Y_R^{n-1}, Y_R^n\right)\right\|_{\mathbb{H}^1}.
\end{aligned}$$

Moreover,

$$\partial_x F\left(X_R^{n-1}, X_R^n\right) = \frac{4}{9}\left(2\mathcal{R}\mathrm{e}\left\{X_R^{n-1}\partial_x \overline{X}_R^{n-1}\right\} + 2\mathcal{R}\mathrm{e}\left\{X_R^n \partial_x \overline{X}_R^n\right\}\right) X_R^{n-1/2} + \frac{4}{9}\left(\left|X_R^{n-1}\right|^2 + |X_R^n|^2\right)\partial_x X_R^{n-1/2}.$$

Hence, using the mean value theorem, the fact that $\Theta_R = \Theta\left(\frac{\|.\|_{\mathbb{H}^1}}{R}\right)$, the Sobolev embedding $\mathbb{H}^1 \hookrightarrow L^\infty$ and the inequality $\|X_R^n\|_{\mathbb{H}^1} \leqslant 2R$ on $[0, \tau_R^1)$, we deduce that

$$\sup_{\substack{n\in\mathbb{N}^* \\ n\Delta t<\tau_R^1}} \left\|\left(\Theta_{X_R}^{n-1,n} - \Theta_{Y_R}^{n-1,n}\right) F\left(X_R^{n-1}, X_R^n\right)\right\|_{\mathbb{H}^1} \leqslant CR^2 \left\|\Theta'\right\|_{L^\infty(\mathbb{R})} \sup_{\substack{n\in\mathbb{N}^* \\ n\Delta t<\tau_R^1}} \left(\left\|X_R^{n-1} - Y_R^{n-1}\right\|_{\mathbb{H}^1} + \left\|X_R^n - Y_R^n\right\|_{\mathbb{H}^1}\right).$$

Using again that $\mathbb{H}^1(\mathbb{R})$ is an algebra

$$\sup_{\substack{n\in\mathbb{N}^* \\ n\Delta t\leqslant\tau_R^1}} \left\|\Theta_{Y_R}^{n-1,n}\left(F\left(X_R^{n-1}, X_R^n\right) - F\left(Y_R^{n-1}, Y_R^n\right)\right)\right\|_{\mathbb{H}^1} \leqslant CR^2 \sup_{\substack{n\in\mathbb{N}^* \\ n\Delta t\leqslant\tau_R^1}} \left(\left\|X_R^{n-1} - Y_R^{n-1}\right\|_{\mathbb{H}^1} + \left\|X_R^n - Y_R^n\right\|_{\mathbb{H}^1}\right).$$

Noticing that $\Theta_{X_R}^{n-1,n} = 0$ on $[\tau_R^1, \tau_R^2]$ and proceeding as for the first term, we obtain a similar estimate for the third term. Collecting the previous results leads to

$$\left\|\mathcal{T}X_R^N - \mathcal{T}Y_R^N\right\|_{L^\infty(0,T;\mathbb{H}^1)} \leqslant CTR^2 \left\|X_R^N - Y_R^N\right\|_{L^\infty(0,T;\mathbb{H}^1)},$$

where C is a constant independent of N. We deduce that $\mathcal{T}$ is a contraction mapping from $L^\infty(0,T;\mathbb{H}^1)$ into itself as soon as T is chosen such that $T < R^{-2}/C$. Since T only depends on R the unique solution may be extended to the whole interval. Moreover, we easily obtain that the scheme (4.1.3) preserves the $\mathbb{L}^2$ norm taking the scalar product in $\mathbb{L}^2$ of Equation (4.3.4) with $\left(\overline{X}_R^{n+1/2}\right)^t$ since $\mathcal{U}_{\Delta t}^{n,0}$ is an isometry in $\mathbb{L}^2$. □

Let us set $e_R^n = X_R^n - \widetilde{X}_R^n$, where X_R^n is the solution of (4.3.4) and $\widetilde{X}_R^n$ is the solution of (4.3.3) evaluated at time t_n.

Proposition 4.3.2. *For any $T \geqslant 0$ and $p \geqslant 1$, there exists a positive constant C, depending on R, T and p, and the $\mathbb{H}^6$ norm of the initial data, such that*

$$\mathbb{E}\left(\max_{n=0,\cdots,N} \|e_R^n\|_{\mathbb{H}^1}^{2p}\right) \leqslant C(R, T, p, \gamma, \|X_0\|_{\mathbb{H}^6})\Delta t^p. \tag{4.3.6}$$

Before proving Proposition 4.3.2, let us state a Lemma which gives uniform bounds for the solution X_R of the cut-off equation (4.3.3).

Lemma 4.3.1. *Let $X_0 \in \mathbb{H}^6$ and X_R be the solution of (4.3.3); then for all $T > 0$ there exists a positive constant $C(R,T)$, such that, a.s for every t in $[0,T]$,*

$$\|X_R(t)\|_{\mathbb{H}^6} \leqslant C\left(R, T, \|X_0\|_{\mathbb{H}^6}\right).$$

Proof. The proof is similar to the one of Lemma 4.1 in Chapter 3. □

Proof of Proposition 4.3.2. The proof follows the same lines as in [15] and strongly uses the fact that $\Theta_R^2 F$ is globally lipschitz on $\mathbb{H}^1$. Using the Duhamel formulation (4.3.2) for the continuous cut off equation and the discrete Duhamel equation (4.3.5), we get for any $p \geqslant 1$

$$\begin{aligned}\mathbb{E}\left(\max_{n=0,\cdots,N}\|e_R^n\|_{\mathbb{H}^1}^{2p}\right) \leqslant & C_p\mathbb{E}\left(\max_{n=0,\cdots,N}\left\|\left(U\left(t_n,0\right)-\mathcal{U}_{\Delta t}^{n,0}\right)X_0\right\|_{\mathbb{H}^1}^{2p}\right)\\ &+C_p\mathbb{E}\left(\max_{n=0,\cdots,N}\left\|\sum_{l=1}^{n}A_1^{l-1,l}+A_2^{l-1,l}+A_3^{l-1,l}\right\|_{\mathbb{H}^1}^{2p}\right),\end{aligned}$$

where

$$\left\{\begin{array}{l} A_1^{l-1,l}=\displaystyle\int_{t_{l-1}}^{t_l}\left(\Theta_R^2\left(X_R(s)\right)-\Theta_{X_R}^{l-1,l}\right)U\left(t_n,s\right)F\left(X_R(s)\right)ds\\ A_2^{l-1,l}=\displaystyle\int_{t_{l-1}}^{t_l}\left(U\left(t_n,s\right)-\mathcal{U}_{\Delta t}^{n,l}\right)\Theta_{X_R}^{l-1,l}F\left(X_R(s)\right)ds\\ A_3^{l-1,l}=\displaystyle\int_{t_{l-1}}^{t_l}\mathcal{U}_{\Delta t}^{n,l}\Theta_{X_R}^{l-1,l}\left(F\left(X_R(s)\right)-F\left(X_R^{l-1},X_R^{l}\right)\right)ds.\end{array}\right. \tag{4.3.7}$$

From Proposition 4.2.2, we already know that for any $p \geqslant 1$

$$\mathbb{E}\left(\max_{n=0,\cdots,N}\left\|\left(U\left(t_n,0\right)-\mathcal{U}_{\Delta t}^{n,0}\right)X_0\right\|_{\mathbb{H}^1}^{2p}\right) \leqslant C(T,\gamma)\,\|X_0\|_{\mathbb{H}^6}^{2p}\,\Delta t^p.$$

Let us now denote by $\widetilde{e}_R^n(s)$ the difference $\widetilde{e}_R^n(s) = X_R(s) - \widetilde{X}_R^n$ for all $s \in [t_n, t_{n+1}]$ and state an intermediate result which gives a local estimate on $\widetilde{e}_R^n(s)$.

Lemma 4.3.2. *For any $p \geqslant 1$ and $m \in \mathbb{N}$, if $X_0 \in \mathbb{H}^{m+1}$ then there exists a positive constant $C_1(R,T,p,\gamma,\|X_0\|_{\mathbb{H}^{m+1}})$ such that*

$$\mathbb{E}\left(\sup_{t_{l-1}\leqslant t\leqslant t_l}\left\|\widetilde{e}_R^{l-1}(t)\right\|_{\mathbb{H}^m}^{2p}\right) \leqslant C_1(R,T,p,\gamma,\|X_0\|_{\mathbb{H}^{m+1}})\Delta t^p \qquad \forall\ l=1,\cdots,N.$$

Proof. Using the Duhamel formulation, we write $X_R(t)$ in terms of $\widetilde{X}_R^{l-1}$

$$X_R(t)-\widetilde{X}_R^{l-1} = \left(U\left(t,t_{l-1}\right)-\mathrm{Id}\right)\widetilde{X}_R^{l-1} + i\int_{t_{l-1}}^{t}\Theta_R^2\left(X_R(s)\right)U\left(t,s\right)F\left(X_R(s)\right)ds.$$

We deduce, using Lemma 4.2.1 and 4.3.1, that for any $p \geqslant 1$

$$\mathbb{E}\left(\sup_{t_{l-1}\leqslant t\leqslant t_l}\left\|\left(U\left(t,t_{l-1}\right)-\mathrm{Id}\right)\widetilde{X}_R^{l-1}\right\|_{\mathbb{H}^1}^{2p}\right) \leqslant C(R,T,p,\gamma,\|X_0\|_{\mathbb{H}^{m+1}})\Delta t^p.$$

We easily conclude, by definition of Θ_R,

$$\mathbb{E}\left(\sup_{t_{l-1}\leqslant t\leqslant t_l}\left\|\int_{t_{l-1}}^{t}\Theta_R^2\left(X_R(s)\right)U\left(t,s\right)F\left(X_R(s)\right)ds\right\|_{\mathbb{H}^1}^{2p}\right)\leqslant C(R)\Delta t^{2p}.$$

Finally, we get

$$\mathbb{E}\left(\sup_{t_{l-1}\leqslant t\leqslant t_l}\left\|\widetilde{e}_R^{l-1}(t)\right\|_{\mathbb{H}^1}^{2p}\right)\leqslant C_1(R,T,p,\gamma,\|X_0\|_{\mathbb{H}^{m+1}})\Delta t^{p}.$$

□

In order to obtain an estimate on the global error in $L^{2p}\left(\Omega\right)$, we decompose the term $X_R(s)-X_R^l$, appearing in $A_1^{l-1,l}$ and $A_3^{l-1,l}$, in two terms : $\widetilde{e}_R^l(s)$ and e_R^l. The first term gives the contribution to the final order and the second term may be handled by a fixed point procedure. Let us denote $\Theta_{X_R}^l=\Theta\left(\frac{\|X_R^l\|_{\mathbb{H}^1}}{R}\right)$ for any $l=0,\cdots n$. Writing

$$\begin{aligned}\Theta_R^2\left(X_R(s)\right)-\Theta_{X_R}^{l-1,l}=&\,\Theta_R\left(X_R(s)\right)\left(\Theta_R\left(X_R(s)\right)-\Theta_{X_R}^{l-1}\right)\\&-\Theta_{X_R}^{l-1}\left(\Theta_R\left(X_R(s)\right)-\Theta_{X_R}^{l}\right),\end{aligned}$$

using the isometry property of the random propagator U and the boundedness of Θ and Θ', we obtain the following bound

$$\begin{aligned}&\left\|\sum_{l=1}^{n}\int_{t_{l-1}}^{t_l}\Theta_R\left(X_R(s)\right)\left(\Theta_R\left(X_R(s)\right)-\Theta_{X_R}^{l-1}\right)U(t_n,s)F\left(X_R(s)\right)ds\right\|_{\mathbb{H}^1}^{2p}\\&\quad\leqslant\left(\sum_{l=1}^{n}\sup_{t_{l-1}\leqslant s\leqslant t_l}\left\|\Theta_R\left(X_R(s)\right)F\left(X_R(s)\right)\right\|_{\mathbb{H}^1}\int_{t_{l-1}}^{t_l}\left|\left(\Theta_R\left(X_R(s)\right)-\Theta_{X_R}^{l-1}\right)\right|ds\right)^{2p}\\&\quad\leqslant\left(\sum_{l=1}^{n}CR^3\int_{t_{l-1}}^{t_l}\frac{\|\Theta'\|_{L^\infty}}{R}\left(\left\|\widetilde{e}_R^{l-1}(s)\right\|_{\mathbb{H}^1}+\left\|e_R^{l-1}\right\|_{\mathbb{H}^1}\right)ds\right)^{2p}.\end{aligned}$$

By the same arguments, together with Lemma 4.3.1, we obtain

$$\begin{aligned}&\left\|\sum_{l=1}^{n}\int_{t_{l-1}}^{t_l}\Theta_{X_R}^{l-1}\left(\Theta_R\left(X_R(s)\right)-\Theta_{X_R}^{l}\right)U(t_n,s)F\left(X_R(s)\right)ds\right\|_{\mathbb{H}^1}^{2p}\\&\quad\leqslant\left(\sum_{l=1}^{n}CR^3\int_{t_{l-1}}^{t_l}\frac{\|\Theta'\|_{L^\infty}}{R}\left(\left\|\widetilde{X}_R^l-X_R(s)\right\|_{\mathbb{H}^1}+\left\|e_R^l\right\|_{\mathbb{H}^1}\right)ds\right)^{2p}.\end{aligned}$$

Combining the two above estimates we conclude, using Lemma 4.3.2, that

$$\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}A_1^{l-1,l}\right\|_{\mathbb{H}^1}^{2p}\right)\leqslant C(R,T,p)\left[\left(1+\gamma^p\right)\Delta t^p+\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\|e_R^n\|_{\mathbb{H}^1}^{2p}\right)\right].$$

Now, we may write

$$\begin{aligned} A_2^{l-1,l} &= \int_{t_{l-1}}^{t_l} U(t_n,s)\left(\mathrm{Id}-U(s,t_{l-1})\right)\Theta_{X_R}^{l-1,l}F(X_R(s))\,ds \\ &\quad+\int_{t_{l-1}}^{t_l}\left(U(t_n,t_{l-1})-\mathcal{U}_{\Delta t}^{n,l}\right)\Theta_{X_R}^{l-1,l}F(X_R(s))\,ds \\ &= A_{2,1}^{l-1,l}+A_{2,2}^{l-1,l}. \end{aligned}$$

We can simply estimate the first term in the above equality using Lemma 4.3.1 and 4.2.1. Indeed, using again the isometry property of the random propagator $U(t_n,s)$ and Hölder inequality, together with Lemma 4.3.1, we get

$$\begin{aligned} \left\|\sum_{l=1}^n A_{2,1}^{l-1,l}\right\|_{\mathbb{H}^1}^{2p} &\leqslant \left(\sum_{l=1}^n\int_{t_{l-1}}^{t_l}\|\mathrm{Id}-U(s,t_{l-1})\|_{\mathcal{L}(\mathbb{H}^2,\mathbb{H}^1)}\left\|\Theta_{X_R}^{l-1,l}F(X_R(s))\right\|_{\mathbb{H}^2}ds\right)^{2p} \\ &\leqslant n^{2p-1}\Delta t^{2p}C(R,T,p,\|X_0\|_{\mathbb{H}^2})\sum_{l=1}^n\sup_{t_{l-1}\leqslant s\leqslant t_l}\|\mathrm{Id}-U(s,t_{l-1})\|_{\mathcal{L}(\mathbb{H}^2,\mathbb{H}^1)}^{2p}. \end{aligned}$$

Hence, by Lemma 4.2.1

$$\begin{aligned} &\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^n A_{2,1}^{l-1,l}\right\|_{\mathbb{H}^1}^{2p}\right) \\ &\leqslant \mathbb{E}\left(N^{2p-1}\Delta t^{2p}C(R,T,p)\sum_{l=1}^N\sup_{t_{l-1}\leqslant s\leqslant t_l}\|\mathrm{Id}-U(s,t_{l-1})\|_{\mathcal{L}(\mathbb{H}^2,\mathbb{H}^1)}^{2p}\right) \\ &\leqslant C_1(R,T,p,\gamma,\|X_0\|_{\mathbb{H}^2})\Delta t^p. \end{aligned}$$

On the contrary, the second term $A_{2,2}^{l-1,l}$ cannot be bounded directly because we do not have an explicit representation (in Fourier space) of the random propagator $U(t,s)$, $t,s\in\mathbb{R}_+, t\geqslant s$, solution of the linear equation (4.2.1). Usually terms similar to $A_{2,2}^{l-1,l}$ are bounded thanks to the explicit expression of the semigroup generated by the Laplace operator using either the Fourier transform in the whole space or its eigenvalues on a smooth bounded domain [20, 95]. Writing

$$\mathcal{U}_{\Delta t}^{n,l}=U_{\Delta t,n}\cdots U_{\Delta t,l-1}\left(\mathrm{Id}-\frac{1}{2}H_{\Delta t,l-1}\right)^{-1},$$

we may split $A_{2,2}^{l-1,l}$ as follows

$$\begin{aligned} A_{2,2}^{l-1,l} &= \int_{t_{l-1}}^{t_l}\left(U(t_n,t_{l-1})-\mathcal{U}_{\Delta t}^{n,l}\right)\Theta_{X_R}^{l-1,l}\left(F(X_R(s))-F\left(\widetilde{X}_R^{l-1}\right)\right)ds \\ &\quad+\int_{t_{l-1}}^{t_l}\Theta_{X_R}^{l-1,l}\left(U(t_n,t_{l-1})-U_{\Delta t,n}\cdots U_{\Delta t,l-1}\right)F\left(\widetilde{X}_R^{l-1}\right)ds \\ &\quad+\int_{t_{l-1}}^{t_l}\Theta_{X_R}^{l-1,l}U_{\Delta t,n}\cdots U_{\Delta t,l-1}\left(\mathrm{Id}-\left(\mathrm{Id}-\frac{1}{2}H_{\Delta t,l-1}\right)^{-1}\right)F\left(\widetilde{X}_R^{l-1}\right)ds \\ &= A_{2,2,1}^{l-1,l}+A_{2,2,2}^{l-1,l}+A_{2,2,3}^{l-1,l}. \end{aligned}$$

The first term $A_{2,2,1}^{l-1,l}$ is easily bounded thanks to the local Lipschitz property of the nonlinear function F, the isometry property of both $U(t_n, t_{l-1})$ and $\mathcal{U}_{\Delta t}^{n,l}$, the boundedness of Θ and Lemma 4.3.1. This leads to

$$\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}A_{2,2,1}^{l-1,l}\right\|_{\mathbb{H}^1}^{2p}\right)\leqslant C_1(R,T,p,\gamma,\|X_0\|_{\mathbb{H}^1})\left[\Delta t^p+\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\|e_R^n\|_{\mathbb{H}^1}^{2p}\right)\right]. \tag{4.3.8}$$

Let us now consider the second term $A_{2,2,2}^{l-1,l}$ that can be bounded using the linear estimate (4.2.4) obtained in Proposition 4.2.2 together with Lemma 4.3.1. In this way,

$$\begin{aligned}&\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}A_{2,2,2}^{l-1,l}\right\|_{\mathbb{H}^1}^{2p}\right)\\&\quad\leqslant T^{2p}\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\max_{l\in[\![1,n]\!]}\|U(t_n,t_{l-1})-U_{\Delta t,n}\cdots U_{\Delta t,l-1}\|_{\mathcal{L}(\mathbb{H}^6,\mathbb{H}^1)}^{2p}\left\|F\left(\widetilde{X}_R^{l-1}\right)\right\|_{\mathbb{H}^6}^{2p}\right)\\&\quad\leqslant C_2(R,T,p,\gamma,\|X_0\|_{\mathbb{H}^6})\Delta t^p.\end{aligned} \tag{4.3.9}$$

An estimate on the last term $A_{2,2,3}^{l-1,l}$ is obtained thanks to the next result, whose proof is identical to Lemma 4.2.4.

Lemma 4.3.3. *For any n and $p\in\mathbb{N}$, there exists a positive random constant $C_n(\omega)<+\infty$ a.s. belonging to $L^{2p}(\Omega)$ such that for any $f\in\mathbb{H}^1$*

$$\left\|\left[Id-\left(Id-\frac{1}{2}H_{\Delta t,l-1}\right)^{-1}\right]f\right\|_{\mathbb{L}^2}\leqslant C_n(\omega)\sqrt{\Delta t}\,\|f\|_{\mathbb{H}^1}\qquad a.s.$$

From this Lemma, we easily obtain a bound on the last term $A_{2,2,3}$

$$\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}A_{2,2,3}^{l-1,l}\right\|_{\mathbb{H}^1}^{2p}\right)\leqslant C_3(R,T,p,\gamma,\|X_0\|_{\mathbb{H}^2},\omega)\Delta t^p. \tag{4.3.10}$$

Combining the above estimates (4.3.8), (4.3.9) and (4.3.10), we obtain an estimate on $A_{2,2}^{l-1,l}$

$$\begin{aligned}&\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\left\|\sum_{l=1}^{n}A_{2,2}^{l-1,l}\right\|_{\mathbb{H}^1}^{2p}\right)\\&\quad\leqslant C(R,T,p,\gamma,\|X_0\|_{\mathbb{H}^6})\left[\Delta t^p+\mathbb{E}\left(\max_{n\in[\![1,N]\!]}(C_n)^{2p}\right)\Delta t^p+\mathbb{E}\left(\max_{n\in[\![1,N]\!]}\|e_R^n\|_{\mathbb{H}^1}^{2p}\right)\right].\end{aligned}$$

Finally, we bound the last term $A_3^{l-1,l}$ dividing it as follows

$$\begin{aligned} A_3^{l-1,l} &= \int_{t_{l-1}}^{t_l} \mathcal{U}_{\Delta t}^{n,l} \Theta_{X_R}^{l-1,l} \left(F\left(X_R(s)\right) - F\left(\widetilde{X}_R^{l-1}, \widetilde{X}_R^{l-1}\right)\right) ds \\ &\quad + \int_{t_{l-1}}^{t_l} \mathcal{U}_{\Delta t}^{n,l} \Theta_{X_R}^{l-1,l} \left(F\left(\widetilde{X}_R^{l-1}, \widetilde{X}_R^{l-1}\right) - F\left(\widetilde{X}_R^{l-1}, \widetilde{X}_R^{l}\right)\right) ds \\ &\quad + \int_{t_{l-1}}^{t_l} \mathcal{U}_{\Delta t}^{n,l} \Theta_{X_R}^{l-1,l} \left(F\left(\widetilde{X}_R^{l-1}, \widetilde{X}_R^{l}\right) - F\left(X_R^{l-1}, X_R^{l}\right)\right) ds \\ &= A_{3,1}^{l-1,l} + A_{3,2}^{l-1,l} + A_{3,3}^{l-1,l}. \end{aligned}$$

Note that the last term $A_{3,3}^{l-1,l}$ is easily bounded using the same arguments as those used in the fixed point procedure. The first term $A_{3,1}^{l-1,l}$ is bounded using that $F\left(\widetilde{X}_R^{l-1}, \widetilde{X}_R^{l-1}\right) = F\left(\widetilde{X}_R^{l-1}\right)$. Hence, by Lemma 4.3.1 and Hölder inequality, we obtain

$$\begin{aligned} \left\| \sum_{l=1}^{n} A_{3,1}^{l-1,l} \right\|_{\mathbb{H}^1}^{2p} &\leqslant \left(\sum_{l=1}^{n} \int_{t_{l-1}}^{t_l} \left\| F\left(X_R(s)\right) - F\left(\widetilde{X}_R^{l-1}\right) \right\|_{\mathbb{H}^1} ds \right)^{2p} \\ &\leqslant n^{2p-1} \Delta t^{2p} C(R,T,p) \sum_{l=1}^{n} \sup_{t_{l-1} \leqslant s \leqslant t_l} \left\| X_R(s) - \widetilde{X}_R^{l-1} \right\|_{\mathbb{H}^1}^{2p}. \end{aligned}$$

Then, using the estimate given by Lemma 4.3.2, we deduce that

$$\begin{aligned} &\mathbb{E}\left(\max_{n \in [\![1,N]\!]} \left\| \sum_{l=1}^{n} A_{3,1}^{l-1,l} \right\|_{\mathbb{H}^1}^{2p} \right) \\ &\leqslant \mathbb{E}\left(N^{2p-1} \Delta t^{2p} C(R,T,p) \sum_{l=1}^{N} \sup_{t_{l-1} \leqslant s \leqslant t_l} \left\| X_R(s) - \widetilde{X}_R^{l-1} \right\|_{\mathbb{H}^1}^{2p} \right) \\ &\leqslant C_1(R,T,p,\gamma, \|X_0\|_{\mathbb{H}^2}) \Delta t^p. \end{aligned}$$

An estimate on the second term $A_{3,2}^{l-1,l}$ may be obtained noticing first that

$$\begin{aligned} &F\left(\widetilde{X}_R^{l-1}, \widetilde{X}_R^{l-1}\right) - F\left(\widetilde{X}_R^{l-1}, \widetilde{X}_R^{l}\right) \\ &\quad = \frac{2}{9}\left(\left|\widetilde{X}_R^{l-1}\right|^2 - \left|\widetilde{X}_R^{l}\right|^2 \right) \widetilde{X}_R^{l-1} + \frac{4}{9} \left|\widetilde{X}_R^{l-1}\right|^2 \left(\widetilde{X}_R^{l-1} - \widetilde{X}_R^{l} \right) + \frac{2}{9}\left(\left|\widetilde{X}_R^{l-1}\right|^2 - \left|\widetilde{X}_R^{l}\right|^2 \right) \widetilde{X}_R^{l}. \end{aligned}$$

We deduce, using the algebra property of $\mathbb{H}^1(\mathbb{R})$, Lemma 4.3.1 and Lemma 4.3.2, that

$$\mathbb{E}\left(\max_{n \in [\![1,N]\!]} \left\| \sum_{l=1}^{n} A_{3,2}^{l-1,l} \right\|_{\mathbb{H}^1}^{2p} \right) \leqslant C(R,T,p,\gamma, \|X_0\|_{\mathbb{H}^2}) \Delta t^p.$$

Collecting all the estimates leads to

$$\mathbb{E}\left(\max_{n \in [\![1,N]\!]} \|e_R^n\|_{\mathbb{H}^1}^{2p} \right) \leqslant C(R,T,p,\gamma, \|X_0\|_{\mathbb{H}^6}) \left[\Delta t^p + \mathbb{E}\left(\max_{n \in [\![1,N]\!]} \|e_R^n\|_{\mathbb{H}^1}^{2p} \right) \right].$$

Thus, for $T = T_1$ chosen sufficiently small so that $C(T_1, R, p, \gamma, \|X_0\|_{\mathbb{H}^6}) < 1$, we obtain

$$\mathbb{E}\left(\max_{n\in[\![1,N_{T_1}]\!]} \|e_R^n\|_{\mathbb{H}^1}^{2p}\right) \leqslant \frac{C(T_1, R, p, \gamma, \|X_0\|_{\mathbb{H}^6})}{1 - C(T_1, R, p, \gamma, \|X_0\|_{\mathbb{H}^6})}\Delta t^p.$$

Iterating this process on the time intervals $[T_1, 2T_1]$ and up to the final time T, we conclude that the scheme is of order $1/2$. □

4.3.2 Probability and almost sure order for the scheme (4.1.3)

Since the scheme is implicit, we had to use a cut off function of the approximation at times $n\Delta t$ and $(n+1)\Delta t$. In order to define a discrete solution to Equation (4.1.3), let us define the random variable

$$\tau_{\Delta t}^R = \inf\left\{n\Delta t, \left\|X_R^{n-1}\right\|_{\mathbb{H}^1} \geqslant R \quad \text{or} \quad \|X_R^n\|_{\mathbb{H}^1} \geqslant R\right\},$$

which is a $\mathcal{F}_{n\Delta t}$ stopping time. It is then clear that $(X_R^n)_{n=0,\cdots,n_0-1}$ satisfy the scheme (4.1.3) provided that $n_0\Delta t < \tau_{\Delta t}^R$. However, we do not know if a solution X_N^{n+1} to (4.1.3) exists and is unique.

Remark 4.3.1. *We cannot proceed as in the continuous case defining the blow-up time as the limit of $\tau_{\Delta t}^R$ when R goes to infinity because the time step Δt depends on the cut-off radius R as it is seen when constructing X_R^{n+1} from X_R^n.*

We follow the approach of [20] to construct a solution X_N^{n+1} to (4.1.3). The next Lemma gives a sufficient condition on the time step to extend the solution to $n+1$.

Lemma 4.3.4. *There exists a constant C_2 such that for any $\Delta t > 0$ and R_0 satisfying $\Delta t \leqslant C_2 R_0^{-2}$ and $n\Delta t \leqslant \tau_{\Delta t}^{R_0}$, there exists a unique adapted solution Z^{n+1} of*

$$Z^{n+1} = U_{\Delta t,n}X^n + i\Delta t T_{\Delta t,n}^{-1} F\left(X^n, Z^{n+1}\right) \tag{4.3.11}$$

such that $\|Z^{n+1}\|_{\mathbb{H}^1} \leqslant 4R_0$, provided $\|X^n\|_{\mathbb{H}^1} \leqslant R_0$.

Proof. We denote by $\mathcal{T}Z^{n+1}$ the right hand side of Equation (4.3.11). Since $\|X^n\|_{\mathbb{H}^1} \leqslant R_0$ by hypothesis, we easily obtain that

$$\left\|\mathcal{T}Z^{n+1}\right\|_{\mathbb{H}^1} \leqslant R_0 + C\frac{4\Delta t}{9}\left(R_0^3 + \left\|Z^{n+1}\right\|_{\mathbb{H}^1}^2 R_0 + \left\|Z^{n+1}\right\|_{\mathbb{H}^1} R_0^2 + \left\|Z^{n+1}\right\|_{\mathbb{H}^1}^3\right).$$

Taking Z^{n+1} in the closed ball of radius $4R_0$ and center 0 in $\mathbb{H}^1$, we then deduce that $\mathcal{T}$ maps the closed ball of radius $4R_0$ of $\mathbb{H}^1$ into itself, provided that $\Delta t \leqslant \frac{27}{85\times 4\times C}R_0^{-2}$. The existence of a unique adapted solution follows using the Banach fixed point theorem. □

Following the approach of [20], we now define a new process Y_R^{n+1}, solution of the truncated scheme (4.3.4) with $X_R^n = X^n$, and we define the random variable

$$R_{n+1} = \min\left\{R \in \mathbb{N}, \left\|Y_R^{n+1}\right\|_{\mathbb{H}^1} \leqslant R\right\}.$$

Fix any deterministic function $X_{\Delta t,\infty}$ such that $\left\|X_{\Delta t,\infty}\right\|_{\mathbb{H}^1} = 4R_0$. Thus, for $\Delta t \leqslant C_2 R_0^{-2}$, we can define a solution of Equation (4.1.3) as follows

$$X^{n+1} = \begin{cases} Z^{n+1} & \text{if} \quad \|X^n\|_{\mathbb{H}^1} \leqslant R_0 \\ Y^{n+1}_{R_{n+1}} & \text{if} \quad \|X^n\|_{\mathbb{H}^1} > R_0 \text{ and } R_{n+1} < +\infty \text{ and } X^n \neq X_{\Delta t,\infty} \\ X_{\Delta t,\infty} & \text{otherwise.} \end{cases} \tag{4.3.12}$$

Finally, let $\tau^*_{\Delta t}$ be the discrete stopping time such that $\tau^*_{\Delta t} = n_0 \Delta t$ and n_0 be the first integer such that $X^n = X_{\Delta t,\infty}$. In this way, we define a solution to (4.1.3) up to time $\tau^*_{\Delta t}$. Let us now prove a convergence in probability rather than in $\mathbb{L}^p(\Omega)$ following the definition given in [95].

Theorem 4.3.1. *Assume that $X_0 \in \mathbb{H}^6$, then for any stopping time $\tau < \tau^*$ almost surely we have*

$$\lim_{C \to +\infty} \mathbb{P}\left(\max_{n=0,\dots,N_\tau} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \geqslant C \Delta t^{1/2} \right) = 0,$$

uniformly in Δt. Then we say, according to [95], that the scheme has order $1/2$ in probability. Moreover, for any $p > 1$ and $\delta < \frac{1}{2} - \frac{1}{2p}$, there exists a random variable K_δ such that

$$\max_{n=0,\dots,N_\tau} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \leqslant K_\delta(T, R, p, \omega) \Delta t^\delta.$$

Note that from the above almost sure convergence, we get, for any stopping time $\tau < \tau^*$ a.s,

$$\lim_{\Delta t \to 0} \mathbb{P}\left(\tau^*_{\Delta t} < \tau\right) = 0.$$

Moreover, using the Fatou Lemma and the lower semicontinuity of the characteristic function $\mathbb{1}_{\tau^*_{\Delta t} < \tau}$, we may write

$$\liminf_{\Delta t \to 0} \mathbb{E}\left(\mathbb{1}_{\tau^*_{\Delta t} < \tau}\right) \geqslant \mathbb{E}\left(\liminf_{\Delta t \to 0} \mathbb{1}_{\tau^*_{\Delta t} < \tau}\right) \geqslant \mathbb{P}\left(\liminf_{\Delta t \to 0} \tau^*_{\Delta t} < \tau\right)$$

and we obtain that

$$\mathbb{P}\left(\liminf_{\Delta t \to 0} \tau^*_{\Delta t} \geqslant \tau^*\right) = 1.$$

Proof of Theorem 4.3.1. The proof in [20] can be adapted straightforwardly. We recall here for convenience the main steps of the proof.

Convergence in probability. For any stopping time $\tau < \tau^*$ *a.s* and any $\epsilon \in (0,1)$, there exists R_0 such that

$$\mathbb{P}\left(\sup_{t \in [0,\tau]} \|X(t)\|_{\mathbb{H}^1} \geqslant R_0 - 1 \right) \leqslant \epsilon/2.$$

Such a R_0 exists since $\tau < \tau^*$ *a.s.* Since for all $\epsilon \in (0,1)$ the following inclusion holds true

$$\begin{aligned}
&\left\{ \max_{n\in[\![0,N_\tau]\!]} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \geqslant \epsilon \right\} \\
&\subset \quad \left\{ \max_{n\in[\![0,N_\tau]\!]} \left\| \widetilde{X}^n \right\|_{\mathbb{H}^1} \geqslant R_0 - 1 \right\} \\
&\cup \quad \left[\left\{ \max_{n\in[\![0,N_\tau]\!]} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \geqslant \epsilon \right\} \cap \left\{ \max_{n\in[\![0,N_\tau]\!]} \left\| \widetilde{X}^n \right\|_{\mathbb{H}^1} < R_0 - 1 \right\} \right]
\end{aligned} \tag{4.3.13}$$

we obtain

$$\begin{aligned}
&\mathbb{P}\left(\max_{n\in[\![0,N_\tau]\!]} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \geqslant \epsilon \right) \\
&\quad \leqslant \frac{\epsilon}{2} + \mathbb{P}\left(\max_{n\in[\![0,N_\tau]\!]} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \geqslant \epsilon, \max_{n\in[\![0,N_\tau]\!]} \left\| \widetilde{X}^n \right\|_{\mathbb{H}^1} < R_0 - 1 \right).
\end{aligned}$$

As in [20, 95], we define the random variable

$$n_\epsilon = \min\left\{ n, \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \geqslant \epsilon \right\}.$$

If $\max_{n\in[\![0,N_\tau]\!]} \left\| \widetilde{X}^n \right\|_{\mathbb{H}^1} \leqslant R_0 - 1$, then for all $n \in [\![0, n_\epsilon - 1]\!]$, we get by the triangular inequality that

$$\max_{n\in[\![0,n_\epsilon-1]\!]} \|X^n\|_{\mathbb{H}^1} < \epsilon + R_0 - 1 < R_0.$$

Assuming now that $\Delta t \leqslant C_2 R_0^{-2}$, we obtain by Lemma 4.3.4 that $\|X^{n_\epsilon}\|_{\mathbb{H}^1} \leqslant 4R_0$. By construction, it follows that for all $n \in [\![0, n_\epsilon]\!]$, $X^n = X^n_{4R_0}$. We deduce that if $n_\epsilon \leqslant N_\tau$, then $\left\| X^{n_\epsilon}_{4R_0} - \widetilde{X}^{n_\epsilon}_{4R_0} \right\|_{\mathbb{H}^1} \geqslant \epsilon$. Therefore,

$$\max_{n\in[\![0,N_T]\!]} \left\| X^n_{4R_0} - \widetilde{X}^n_{4R_0} \right\|_{\mathbb{H}^1} \geqslant \epsilon.$$

Moreover, since the event

$$\left\{ \max_{n\in[\![0,N_\tau]\!]} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \geqslant \epsilon \right\} \cap \{ n_\epsilon > N_\tau \} = \emptyset,$$

we deduce that,

$$\left\{ \max_{n\in[\![0,N_\tau]\!]} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \geqslant \epsilon \right\} = \left\{ \max_{n\in[\![0,N_\tau]\!]} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \geqslant \epsilon \right\} \cap \{ n_\epsilon \leqslant N_\tau \}.$$

Hence, by Markov inequality and Proposition 4.3.2

$$\begin{aligned}
&\mathbb{P}\left(\max_{n\in[\![0,N_\tau]\!]}\left\|X^n-\widetilde{X}^n\right\|_{\mathbb{H}^1}\geqslant\epsilon, n_\epsilon\leqslant N_\tau, \max_{n\in[\![0,N_\tau]\!]}\left\|\widetilde{X}^n\right\|_{\mathbb{H}^1}<R_0-1\right)\\
&\leqslant\ \mathbb{P}\left(\max_{n\in[\![0,N_T]\!]}\left\|X^n_{4R_0}-\widetilde{X}^n_{4R_0}\right\|_{\mathbb{H}^1}\geqslant\epsilon\right)\\
&\leqslant\ \frac{1}{\epsilon^{2p}}\mathbb{E}\left(\max_{n\in[\![0,N_T]\!]}\left\|X^n_{4R_0}-\widetilde{X}^n_{4R_0}\right\|^{2p}_{\mathbb{H}^1}\right)\\
&\leqslant\ \frac{1}{\epsilon^{2p}}C\left(4R_0,T,p,\gamma,\|X_0\|_{\mathbb{H}^6}\right)\Delta t^p\\
&\leqslant\ \frac{\epsilon}{2},
\end{aligned}$$

for $\Delta t\leqslant\left(\frac{\epsilon^{2p+1}}{2C\left(4R_0,T,p,\gamma,\|X_0\|_{\mathbb{H}^6}\right)}\right)^{\frac{1}{p}}$. Hence,

$$\mathbb{P}\left(\max_{n\in[\![0,N_\tau]\!]}\left\|X^n-\widetilde{X}^n\right\|_{\mathbb{H}^1}\geqslant\epsilon\right)\leqslant\epsilon$$

and the convergence in probability is proved.

Order in probability. Let us now choose a radius $R_1\geqslant R_0-1$ such that for any $\Delta t>0$

$$\mathbb{P}\left(\max_{n\in[\![0,N_\tau]\!]}\|X^n\|_{\mathbb{H}^1}\geqslant R_1\right)\leqslant\frac{\epsilon}{2}.$$

Such a R_1 exists by the previous convergence and because $\{\Delta t\}$ is a discrete sequence. Let us define $\Delta t_1=C_2R_1^{-2}$. Using again the set inclusion (4.3.13) and arguing as above (replacing R_0 by R_1), we easily deduce that for any $\Delta t\leqslant\Delta t_1$

$$\begin{aligned}
&\mathbb{P}\left(\max_{n\in[\![0,N_\tau]\!]}\left\|X^n-\widetilde{X}^n\right\|_{\mathbb{H}^1}\geqslant C\Delta t^{1/2}\right)\\
&\leqslant\frac{\epsilon}{2}+\mathbb{P}\left(\max_{n\in[\![0,N_T]\!]}\left\|X^n_{4R_1}-\widetilde{X}^n_{4R_1}\right\|_{\mathbb{H}^1}\geqslant C\Delta t^{1/2}\right)\\
&\leqslant\frac{\epsilon}{2}+\frac{C\left(4R_1,T,p,\gamma,\|X_0\|_{\mathbb{H}^6}\right)}{C^{2p}}.
\end{aligned}$$

For C large enough such that $\frac{C\left(4R_1,T,p,\gamma,\|X_0\|_{\mathbb{H}^6}\right)}{C^{2p}}\leqslant\frac{\epsilon}{2}$, we get

$$\mathbb{P}\left(\max_{n\in[\![0,N_\tau]\!]}\left\|X^n-\widetilde{X}^n\right\|_{\mathbb{H}^1}\geqslant C\Delta t^{1/2}\right)\leqslant\epsilon,$$

uniformly in $\Delta t<\Delta t_1$. Let us now consider the case $\Delta t>\Delta t_1$. By assumption

$$\begin{aligned}
&\mathbb{P}\left(\max_{n\in[\![0,N_\tau]\!]}\left\|X^n-\widetilde{X}^n\right\|_{\mathbb{H}^1}\geqslant C\Delta t^{1/2}\right)\\
&\leqslant\ \mathbb{P}\left(\max_{n\in[\![0,N_\tau]\!]}\left\|X^n-\widetilde{X}^n\right\|_{\mathbb{H}^1}\geqslant C\Delta t_1^{1/2}\right)\\
&\leqslant\ \mathbb{P}\left(\max_{n\in[\![0,N_\tau]\!]}\|X^n\|_{\mathbb{H}^1}\geqslant C\frac{\Delta t_1^{1/2}}{2}\right)+\mathbb{P}\left(\max_{n\in[\![0,N_\tau]\!]}\left\|\widetilde{X}^n\right\|_{\mathbb{H}^1}\geqslant C\frac{\Delta t_1^{1/2}}{2}\right)\\
&\leqslant\ \epsilon
\end{aligned}$$

for $C \geqslant 2R_1/\Delta t_1^{1/2}$.

Almost sure convergence. From the $L^p(\Omega)$ order obtained in Proposition (4.3.2) and by the Borel Cantelli Lemma, we obtain that for any $p > 1$ and $\delta < \frac{1}{2} - \frac{1}{2p}$, there exists a random variable $K_\delta(T, R, p, \omega)$ such that

$$\max_{n=0,\cdots,N_\tau} \left\| X_R^n - \widetilde{X}_R^n \right\|_{\mathbb{H}^1} \leqslant K_\delta(T, R, p, \gamma, \omega) \Delta t^\delta. \tag{4.3.14}$$

For all stopping time τ such that $\tau < \tau^*$ almost surely, there exists a R_0 such that

$$R_0(\omega) \geqslant \sup_{t\in[0,\tau]} \|X(t)\|_{\mathbb{H}^1} \qquad \text{a.s.} \tag{4.3.15}$$

To prove the almost sure convergence of X^n to $\widetilde{X}^n$ in the $L^\infty(0,T;\mathbb{H}^1)$ norm, we use the same arguments as in step 1 and we proceed by contradiction. Let $\epsilon \in (0,1)$ and $\Delta t \leqslant C_2 R_0^{-2}(\omega)$ and assume that

$$\max_{n=0,\cdots,N_\tau} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \geqslant \epsilon.$$

Moreover, let us set

$$n_\epsilon = \min\left\{ n, \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \geqslant \epsilon \right\}.$$

Then $\|X^n\|_{\mathbb{H}^1} \leqslant R_0$ for $n = 0, \cdots, n_\epsilon - 1$. It follows by Lemma 4.3.4 that $\|X^{n_\epsilon}\|_{\mathbb{H}^1} \leqslant 4R_0$ and we deduce that $\|X^n\|_{\mathbb{H}^1} = \left\|X_{4R_0}^n\right\|_{\mathbb{H}^1}$ for $n = 0, \cdots, n_\epsilon$ from which we get that

$$\max_{n=0,\cdots,N_T} \left\| X_{4R_0}^n - \widetilde{X}_{4R_0}^n \right\|_{\mathbb{H}^1} \geqslant \epsilon.$$

By (4.3.14) this is impossible for Δt small enough. We deduce the almost sure convergence.

Almost sure order. From the above almost sure convergence, we know that almost surely for all $\epsilon > 0$, there exists a Δt_1 such that for all $\Delta t \leqslant \Delta t_1$

$$\max_{n=0,\cdots,N_\tau} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \leqslant \epsilon.$$

Thus, there exists $R_1(\omega) > R_0(\omega)$ such that

$$\max_{n=0,\cdots,N_\tau} \|X^n\|_{\mathbb{H}^1} \leqslant R_1(\omega), \tag{4.3.16}$$

and we deduce that if $\Delta t \leqslant C_2 R_1^{-2}$

$$\max_{n=0,\cdots,N_\tau} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} = \max_{n=0,\cdots,N_\tau} \left\| X_{R_1}^n - \widetilde{X}_{R_1}^n \right\|_{\mathbb{H}^1} \leqslant K_\delta(T, R_1, p, \omega) \Delta t^\delta.$$

If $C_2 R_0^{-2} \geqslant \Delta t > C_2 R_1^{-2}$, we have, thanks to (4.3.15) and (4.3.16), that

$$\max_{n=0,\cdots,N_\tau} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \leqslant 2R_1 \leqslant \frac{2R_1}{\Delta t_1^\delta} \Delta t^\delta.$$

□

Chapter 5

Numerical simulations of the stochastic Manakov equation and of the Polarization Mode Dispersion

5.1 Introduction

In this chapter, we are concerned with numerical simulations of the Stochastic Manakov equation, which appears as the asymptotic behaviour of the Manakov PMD equation (3.1.9). Using the physical notations, this equation is written

$$\begin{aligned} idX(z) + \left(\frac{d_0}{2}\frac{\partial^2 X(z)}{\partial t^2} + F\left(X(z)\right)\right) dz \\ + i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k \frac{\partial X(z)}{\partial t} \circ dW_k(z) = 0, \quad z \geqslant 0, t \in \mathbb{R}, \end{aligned} \tag{5.1.1}$$

where z is the distance along the fiber and x is the retarded time. Recall that γ is a small positive parameter given by the physics of the problem, d_0 is the group velocity dispersion (GVD) and $W = (W_1, W_2, W_3)$ is a 3d real valued Brownian motion. The matrices $\sigma_1, \sigma_2, \sigma_3$ are the Pauli matrices and the nonlinear term is given by $F\left(X(t)\right) = \frac{8}{9}\left|X\right|^2 X(t)$. This equation is of physical interest since the main dispersive effects can be identified. Indeed, the second term corresponds to the chromatic dispersion, the third term to the Kerr effect averaged over the Poincaré sphere, and the last term describes the linear PMD effects.

If the impact of the Polarization Mode Dispersion (PMD) on linear pulse with zero Group Velocity Dispersion (GVD) is quite well understood [39, 42] and easy to modelize; its interactions with nonlinear effects lead to complex behaviour [39, 72, 78]. Our aim in this work is twofold : we first intend to propose a careful numerical analysis for Equation (5.1.1) in order to capture the correct effects of the noise. In the absence of noise, the dynamics reduces to the well-known Manakov equation. This equation is a generalization of the scalar nonlinear Schrödinger equation in the sense that it is integrable and possesses soliton solutions. Thus, the first part of our work is to study the impact of this noise term on solitons and soliton wave-trains

propagation. This analysis shows different results than the usual ones for the NLS equation. The second objective of this chapter is to propose an efficient way to estimate the statistics of the Polarization Mode Dispersion. These estimates are obtained by Monte Carlo simulations. To increase the speed of the computations, we propose an appropriate variance reduction method with control variates. As we shall see, this method is well adapted to estimate expectation of random variables that are solutions of a stochastic partial differential equation. The great advantage of this method is that it does not require any insight on the physics and is relatively easy to implement. However, the choice of the optimal set of parameters is complex and a theoretical analysis should be performed to estimate them.

Our numerical analysis is divided as follows : In section 1, we introduce three numerical schemes which we use to simulate the solution of Equation (5.1.1). The construction of these schemes is based on the theoretical order analysis of Chapter 4. In particular, it follows the discussion about the pertinence of using an implicit discretization of the noise rather than an explicit one.

In section 2, we proceed with the validation of the above methods in deterministic situations. Therefore, we consider a nonlinear medium with constant birefringence. In this case, a good approximation for the slowly varying envelope is given by the Coupled Non-Linear Schrödinger equation (CNLS)

$$i\frac{\partial \mathbf{U}}{\partial z} + b\sigma_3\mathbf{U} + ib'\sigma_3\frac{\partial \mathbf{U}}{\partial t} + \frac{d_0}{2}\frac{\partial^2 \mathbf{U}}{\partial t^2} + \frac{5}{6}\left|\mathbf{U}\right|^2\mathbf{U} + \frac{1}{6}\left(\mathbf{U}^*\sigma_3\mathbf{U}\right)\sigma_3\mathbf{U} + \frac{1}{3}\begin{pmatrix} U_x^* U_y^2 \\ \\ U_y^* U_x^2 \end{pmatrix} = 0. \tag{5.1.2}$$

When the effect of the birefringence may be neglected, this equation appears to be a Hamiltonian perturbation of the Manakov equation [113, 114]. Although our aim is not to study these equations, the objective of this analysis is to emphasize some important aspects on soliton propagation (both propagation and collisions) launched at different polarization states and subject to uniform birefringence. As explained in Chapter 2, chromatic dispersion results from different velocities of the spectral components of the signal. Another important effect, occurring in optical fibers, is due to the high intensity of the incoming pulse and its containment in a small area (corresponding to the core area).

Depending on the value of the GVD parameter d_0, one of these two effects may dominate. If $d_0 < 0$, the dispersion regime is said to be normal; the high frequencies ("blue") propagate more slowly than the low frequencies ("red") and the signal becomes wider. Even for small values of the GVD, the two effects play along leading to larger and faster pulse spreading. On the contrary, if $d_0 > 0$, the dispersion regime is said to be anomalous, the opposite applies and the two effects tend to compensate. The exact balance between chromatic dispersion and Kerr effect produces a very special type of solutions : the solitons.

Remark 5.1.1. *In optical fibers, spatial solitons may also exist (in transverse variables) corresponding to the exact balance between diffraction and autofocalisation.*

Obviously in our case, we are only concerned with temporal solitons since the contribution of the transverse variables has been neglected.

This type of solutions arises in a large class of weakly nonlinear dispersive PDEs. Typical examples are the Korteweg-de Vries equation, known to describe wave motion on the surface of shallow water, and the Schrödinger equation. Solitons have the ability to maintain their shapes while travelling and are resistant to perturbations. Hasegawa and Tappert were the first, in 1973, to propose the use of NLS solitons as digital formats to encode bits in communication optical fibers [56, 93]. The first experimentations were done in 1980 by Mollenauer [84, 93] and coincide with the manufacturing of monomode fibers with negligible loss. Since the Kerr coefficient is positive, soliton may only be obtained in the anomalous dispersion regime that also requires the use of incident light with wavelengths larger than 1.3μm.

Finally, in Section 3, we are concerned with numerical experiments in optical fibers with randomly varying birefringence assuming that pulse propagation is well described by the stochastic Manakov equation (5.1.1). We first display numerical almost sure error curves and according to the analysis of Chapter 4, the almost sure order of the Crank-Nicolson scheme is 1/2. The almost sure order for the Relaxation and the Fourier split-step schemes also seems to be of order 1/2. Next, we give a short review on the statistical analysis of the PMD in linear medium with random birefringence and zero GVD. This case is simpler to handle than the nonlinear one since there is a direct correspondence between the Fourier transform and the propagation equations of the PMD vector and the DGD [42]. Indeed at a fixed frequency, the dynamical evolution of the PMD vector reduces to a simple SDE and its solution can be efficiently computed. Contrarily to the linear case, simple equations for PMD and DGD evolutions can not be obtained considering nonlinear dynamics. Their evolutions will be computed through numerical approximations of the solution of (5.1.1). We also propose a variance reduction method (based on control variates) allowing faster convergence of Monte Carlo methods (see [66, 86] and [45]). In the linear case with zero GVD, this algorithm is compared with usual Monte Carlo simulations together with theoretical results [42]. Finally, we study soliton propagation and collision in presence of random birefringence.

5.2 Fully discrete schemes for the stochastic Manakov equation

In this section, we introduce the fully discrete schemes that are used to perform our numerical simulations. They are constructed from classical schemes for the nonlinear Schrödinger equation. A large number of studies have been devoted to numerical simulations of the nonlinear Schrödinger equations (among them see [4, 5, 31, 64, 99, 103, 111]). These numerical schemes showed to be conservative and therefore, they do not introduce numerical dissipation. This property will be useful in order to study soliton propagation since these schemes succeed to preserve the balance between dispersive effects and nonlinear ones. Here, we consider that $d_0 = 1$.

We consider finite-difference approximation to simulate the complex valued solution $X = (X_1, X_2)$ of the Manakov system

$$i\frac{\partial X}{\partial z} + \frac{1}{2}\frac{\partial^2 X}{\partial t^2} + \frac{8}{9}|X|^2 X = 0, \quad z \geqslant 0, t \in \mathbb{R}, \tag{5.2.1}$$

where $|X|^2 = |X_1|^2 + |X_2|^2$. In practice, it is often approximated by one of the following scheme : the Crank-Nicolson, the Relaxation and the Fourier Split-step schemes. We define a final time $T > 0$ and a final distance $Z > 0$. The space step is $\Delta z = \frac{Z}{N} > 0$ and the retarded-time step is given by $\Delta t = \frac{2T}{M+1} > 0$. The grid is assumed to be homogeneous

$$(z_n, t_j) = (n\Delta z, j\Delta t), \quad n \in \{0, \ldots, N\}, \ j \in \{0, \ldots, M+1\}.$$

The computational domain $[-T, T]$ is taken sufficiently large to avoid numerical reflections. Depending on the scheme, we may consider either homogeneous Dirichlet boundary conditions or periodic boundary conditions. We denote $r = \Delta z/(\Delta t)^2$ and the solution $X = (X_1, X_2)$ of Equation (5.2.1), evaluated at (z_n, t_j), is approximated by $X_j^n = \left(X_{1,j}^n, X_{2,j}^n\right)$.

The Crank-Nicolson scheme. This scheme has been introduced in [31, 99]. For all $n \in [\![0, N-1]\!]$ and all $j \in [\![1, M]\!]$, it reads as follows

$$\begin{aligned} i\left(X_j^{n+1} - X_j^n\right) + \frac{r}{4}\left(X_{j-1}^{n+1} - 2X_j^{n+1} + X_{j+1}^{n+1}\right) + \frac{r}{4}\left(X_{j-1}^n - 2X_j^n + X_{j+1}^n\right) \\ + \Delta z F\left(X_j^n, X_j^{n+1}\right) X_j^{n+1/2} = 0, \end{aligned} \tag{5.2.2}$$

where $X_j^{n+1/2} = \left(X_j^{n+1} + X_j^n\right)/2$ and the nonlinear term F is given by the following approximation

$$F\left(X_j^n, X_j^{n+1}\right) = \frac{4}{9}\left(\left|X_{1,j}^n\right|^2 + \left|X_{1,j}^{n+1}\right|^2 + \left|X_{2,j}^n\right|^2 + \left|X_{2,j}^{n+1}\right|^2\right).$$

The scheme is implicit and, at each space step Δz, the numerical resolution of (5.2.2) is performed by a fixed point procedure. It is well-known that this scheme is of order 2 in the z variable (see [15, 33]). It also preserves the discrete mass and Hamiltonian, given respectively, for all $0 \leqslant n \leqslant N$, by

$$\|X^n\|_{\mathbb{L}^2}^2 = \Delta t \sum_{j=0}^{M+1}\left(\left|X_{1,j}^n\right|^2 + \left|X_{2,j}^n\right|^2\right) \tag{5.2.3}$$

$$\begin{aligned} H\left(X^n\right) = \frac{\Delta t}{4}\sum_{j=0}^{M}\left|\frac{X_{1,j+1}^n - X_{1,j}^n}{\Delta t}\right|^2 + \left|\frac{X_{2,j+1}^n - X_{2,j}^n}{\Delta t}\right|^2 \\ - \frac{2\Delta t}{9}\sum_{j=0}^{M}\left(\left|X_{1,j}^n\right|^2 + \left|X_{2,j}^n\right|^2\right)^2. \end{aligned} \tag{5.2.4}$$

However, this scheme is very costly in CPU time since it requires the computation of a sequence of Picard iterates to converge.

The relaxation scheme. Introduced by C. Besse [4], this method is based on the introduction of a new variable to linearize the equation. This scheme is very efficient since the computation of the new variable is explicit. The fully discrete Relaxation scheme is given by

$$\begin{cases} \dfrac{\Phi_j^{n+1/2}+\Phi_j^{n-1/2}}{2} = \left|X_j^n\right|^2 \\ i\left(X_j^{n+1}-X_j^n\right)+\dfrac{r}{4}\left(X_{j-1}^{n+1}-2X_j^{n+1}+X_{j+1}^{n+1}+X_{j-1}^n-2X_j^n+X_{j+1}^n\right) \\ \qquad +\Delta z\dfrac{8}{9}\Phi_j^{n+1/2}X_j^{n+1/2}=0, \end{cases} \tag{5.2.5}$$

where $\Phi_j^{-1}=\left|X_j^0\right|^2$. This scheme preserves the discrete mass and a modified discrete Hamiltonian

$$H\left(X^n\right)=\frac{\Delta t}{4}\sum_{j=0}^{M}\left|\frac{X_{1,j+1}^n-X_{1,j}^n}{\Delta t}\right|^2+\left|\frac{X_{2,j+1}^n-X_{2,j}^n}{\Delta t}\right|^2-\frac{2\Delta t}{9}\sum_{j=0}^{M}\Phi_j^{n+1/2}\Phi_j^{n-1/2}.$$

In [4], numerical results show evidences that the scheme is of order 2 in the z variable. But, there is no theoretical results yet proving this fact. Moreover, from the definition of the scheme, the property $\Phi^{n+1/2}\geqslant 0$ should hold for all n. However, this fact is not proved either.

Fourier split-step methods. These schemes are based on the decomposition of the flow of Equation (5.2.1) into two (or more) parts, each part being individually simpler to solve. Various decomposition of the flow of Equation (5.2.1) may be considered, the most famous being the Lie formula and the Strang formula. In [5], the authors proved that the Lie formula is of order 1, for initial data in $\mathbb{H}^2$. If the initial data belongs to $\mathbb{H}^4$, then the Strang formula is of order 2. The Lie split-step scheme is given below

$$\begin{cases} Y^{n+1}=S\left(\Delta z\right)X^n \\ X_j^{n+1}=\exp\left(i\dfrac{8}{9}\left|Y_j^{n+1}\right|^2\Delta z\right)Y_j^{n+1}, \end{cases} \tag{5.2.6}$$

where $S(\Delta t)$ defines the group associated with the free equation

$$i\partial_z X+\frac{1}{2}\partial_t^2 X=0. \tag{5.2.7}$$

This scheme is very simple to implement when it is used for the Manakov system (5.2.1). Indeed, the semigroup $S(\Delta z)$ has an explicit representation in Fourier space and can be efficiently solved using the FFT algorithm, which requires periodic boundary conditions. Moreover, the nonlinear part can be solved exactly since the modulus of the solution is preserved. However, in the case of the stochastic Manakov equation (5.1.1), there is no such explicit representation for the random propagator and the linear equation has to be approximated by a linear system. This

scheme preserves the discrete mass but fails to preserve exactly the energy and resonance effects may occur on long distance [24]. Finally, note that those schemes are unconditionally stable in the $\mathbb{L}^2$ norm since the discrete $\mathbb{L}^2$ norm is preserved.

We now adapt the three schemes to our stochastic setting. In order to propose an efficient scheme for Equation (5.1.1), we propose an implicit discretization of the Stratonovich integral by the midpoint rule. Concerning the space discretization, we choose a centered discretization since the random changes of the speed group velocity have not constant sign. Therefore, there is no reason a priori to prefer the upwind or the downwind discretization. For simplicity, we also set the group velocity dispersion parameter d_0 to 1. As in Chapter 4, we set

$$\chi_k^n = \frac{W_k\left((n+1)\Delta z\right) - W_k\left(n\Delta z\right)}{\sqrt{\Delta z}}, \qquad 0 \leqslant n \leqslant N-1,\ k = 1,2,3,$$

where $W = (W_1, W_2, W_3)$ is a 3d real valued Brownian motion.

The stochastic Crank-Nicolson and Relaxation schemes. The discretization of the linear part of Equation (5.1.1) is identical for the two schemes. The stochastic Crank-Nicolson scheme is given by

$$\begin{aligned} i\left(X_j^{n+1} - X_j^n\right) &+ \frac{r}{4}\left(X_{j-1}^{n+1} - 2X_j^{n+1} + X_{j+1}^{n+1}\right) + \frac{r}{4}\left(X_{j-1}^n - 2X_j^n + X_{j+1}^n\right) \\ &+ i\frac{\sqrt{\gamma r}}{4}\sum_{k=1}^{3}\sigma_k\left(X_{j+1}^{n+1} - X_{j-1}^{n+1} + X_{j+1}^n - X_{j-1}^n\right)\chi_k^n \\ &+ \Delta z F\left(X_j^n, X_j^{n+1}\right) X_j^{n+1/2} = 0, \end{aligned} \tag{5.2.8}$$

where $F\left(X_j^n, X_j^{n+1}\right)$ is given by

$$F\left(X_j^n, X_j^{n+1}\right) = \frac{4}{9}\left(\left|X_{1,j}^n\right|^2 + \left|X_{1,j}^{n+1}\right|^2 + \left|X_{2,j}^n\right|^2 + \left|X_{2,j}^{n+1}\right|^2\right).$$

In matrix formulation, the fully discrete relaxation scheme for Equation (5.1.1) reads

$$\begin{cases} \Phi^{n+1/2} = 2\left|X^n\right|^2 - \Phi^{n-1/2} \\ i\left(X^{n+1} - X^n\right) + \left(A_r + D_{\Delta z,\Delta t}^n + F_{\Delta z}^n\right)\left(X^{n+1} + X^n\right) = 0. \end{cases} \tag{5.2.9}$$

The matrix A_r is the usual $M \times M$ discretization matrix of the Laplace operator with homogeneous Dirichlet boundary conditions

$$A_r = \begin{pmatrix} -r/2 & r/4 & & & 0 \\ r/4 & -r/2 & r/4 & & \\ & \ddots & \ddots & \ddots & \\ & & r/4 & -r/2 & r/4 \\ 0 & & & r/4 & -r/2 \end{pmatrix}.$$

We introduce the coefficients $z_1^n = i\frac{\sqrt{\gamma r}}{4}(\chi_1^n + i\chi_2^n)$, $z_2^n = i\frac{\sqrt{\gamma r}}{4}(\chi_1^n - i\chi_2^n)$ and $z_3^n = i\frac{\sqrt{\gamma r}}{4}\chi_3^n$. The matrices $D^n_{\Delta z,\Delta t}$ and $F^n_{\Delta z}$, corresponding respectively to the discretization of the Stratonovich integrals and the linearized nonlinear term, are given by

$$F^n_{\Delta z} = \begin{pmatrix} \frac{4\Delta z}{9}\Phi_1^{n+1/2} & & & & 0 \\ & \frac{4\Delta z}{9}\Phi_1^{n+1/2} & & & \\ & & \ddots & & \\ & & & \frac{4\Delta z}{9}\Phi_M^{n+1/2} & \\ 0 & & & & \frac{4\Delta z}{9}\Phi_M^{n+1/2} \end{pmatrix}.$$

and

$$D^n_{\Delta z,\Delta t} = \begin{pmatrix} 0 & 0 & z_3^n & z_2^n & & \cdots & & 0 \\ 0 & 0 & z_1^n & -z_3^n & & \cdots & & \\ -z_3^n & -z_2^n & 0 & 0 & z_3^n & z_2^n & & \\ -z_1^n & z_3^n & 0 & 0 & z_1^n & -z_3^n & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ & & & \ddots & -z_3^n & -z_2^n & 0 & 0 \\ 0 & & & & -z_1^n & z_3^n & 0 & 0 \end{pmatrix}.$$

The stochastic Fourier split-step scheme. This scheme is based on the decomposition of the flow into two parts : one associated to the linear part of Equation (5.1.1) and the other to the nonlinear part. The scheme is given by

$$\begin{cases} i\left(\widehat{Y}_k^{n+1} - \widehat{X}_k^n\right) - \left(\frac{\Delta z h_k^2}{4} + \frac{\sqrt{\gamma\Delta z}h_k}{2}\sum_{l=1}^{3}\sigma_l\chi_l^n\right)\left(\widehat{Y}_k^{n+1} + \widehat{X}_k^n\right) = 0 \\ X_j^{n+1} = \exp\left(i\frac{8}{9}\left|Y_j^{n+1}\right|^2\Delta z\right)Y_j^{n+1}, \end{cases} \tag{5.2.10}$$

where $\widehat{X}_k^n$ is the discrete Fourier transform of X_j^n and is computed by the FFT and FFTSHIFT algorithms to center the zero frequency, i.e. $\widehat{X}^n = \text{fftshift}\,(\text{fft}\,(\text{ifftshift}\,(X^n)))$. Moreover, the vector $h = \frac{\pi}{T}\left(-\frac{M}{2}, \cdots, 0, \cdots, \frac{M}{2} - 1\right)^t$ contains the M Fourier modes. Equivalently, in matrices notations

$$\begin{cases} i\left(\widehat{Y}^{n+1} - \widehat{X}^n\right) + \left(\widehat{A}_{\Delta z,h} - \widehat{D}_{\Delta z,h}\right)\left(\widehat{Y}^{n+1} + \widehat{X}^n\right) = 0 \\ X^{n+1} = \exp\left(i\frac{8}{9}\left|Y^{n+1}\right|^2\Delta z\right)Y^{n+1}, \end{cases}$$

where $\widehat{A}_{\Delta z,h}$ contains the Fourier modes associated to the Laplace operator

$$\widehat{A}_{\Delta z,h} = \begin{pmatrix} -\Delta z h_1^2/4 & 0 & & & 0 \\ 0 & -\Delta z h_1^2/4 & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & 0 & -\Delta z h_M^2/4 & 0 \\ 0 & & & 0 & -\Delta z h_M^2/4 \end{pmatrix}$$

and $\widehat{D}^n_{\Delta z,h}$ contains the Fourier modes associated to the gradient terms

$$\widehat{D}^n_{\Delta z,h} = \begin{pmatrix} \widehat{z}_3^n h_1 & \widehat{z}_2^n h_1 & 0 & 0 & & & & 0 \\ \widehat{z}_1^n h_1 & -\widehat{z}_3^n h_1 & 0 & 0 & & & & \\ 0 & \ddots & \ddots & \ddots & \ddots & & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & 0 & 0 & \widehat{z}_3^n h_{M-1} & \widehat{z}_2^n h_{M-1} & 0 & 0 \\ & & 0 & 0 & \widehat{z}_1^n h_{M-1} & -\widehat{z}_3^n h_{M-1} & 0 & 0 \\ & & & \ddots & 0 & 0 & \widehat{z}_3^n h_M & \widehat{z}_2^n h_M \\ 0 & & & & 0 & 0 & \widehat{z}_1^n h_M & -\widehat{z}_3^n h_M \end{pmatrix},$$

where the coefficients are now given by

$$\widehat{z}_1^n = \frac{\sqrt{\gamma \Delta z}}{2}\left(\chi_1^n + i\chi_2^n\right), \quad \widehat{z}_2^n = \frac{\sqrt{\gamma \Delta z}}{2}\left(\chi_1^n - i\chi_2^n\right), \quad \widehat{z}_3^n = \frac{\sqrt{\gamma \Delta z}}{2}\chi_3^n.$$

In this case, the matrices to invert are block diagonal. Consequently, this scheme is less time consuming than the Relaxation scheme and the Crank-Nicolson scheme.

Remark 5.2.1. *Spectral methods are very often used in practice for the nonlinear Schrödinger equation because the group associated to the free equation* (5.2.7) *has an explicit and very simple form. However, the random propagator, solution of the linear Equation associated to Equation* (5.1.1)*, does not have an explicit formulation in Fourier space. The reason, already given in Chapter 3, is that the dynamics is driven by a multidimensional Brownian Motion and the Lie algebra associated to the linear equation does not satisfy a nilpotent hypothesis [112]. However, one may be tempted to divide the flow into five steps : one for the Laplace operator, one for each Stratonovich integrals and one for the nonlinear part. But this scheme does not converge numerically to the correct solution. The reason is that terms with similar nature have to be treated simultaneously. Therefore, the three Stratonovich integrals have to be treated together. This explains why we have to solve a linear system even when using a spectral scheme.*

5.3 Numerical simulations in homogeneous medium

In this section, we proceed to the numerical validation of the schemes described in Section 5.2 in two situations. In a first time, we study the numerical behaviour of the three schemes (5.2.2), (5.2.5) and (5.2.6) for the Manakov Equation (5.2.1). We test numerically the propagation of a single soliton and of a 2-soliton. We also compute the numerical errors for the three schemes, using the explicit expression of the Manakov solitons. In a second time, we study numerically the impact of constant birefringence on soliton propagation. We compare the numerical results to the theory.

In fibers with constant birefringence, light propagation is well described by the Coupled Non-Linear Schrödinger equation (5.1.2), which has been derived in Chapter 2 from the Maxwell equations. The behaviour of the nonlinear term $\left(U_x^* U_y^2, U_y^* U_x^2\right)^t$ depends on the strength of the birefringence b. Indeed, by a gauge transform $X = \exp\left(-i\sigma_3 bz\right)\mathbf{U}$, we obtain the new system

$$i\frac{\partial X}{\partial z} + ib'\sigma_3\frac{\partial X}{\partial t} + \frac{d_0}{2}\frac{\partial^2 X}{\partial t^2} + \frac{5}{6}\left|X\right|^2 X + \frac{1}{6}\left(X^*\sigma_3 X\right)\sigma_3 X + \frac{1}{3}\begin{pmatrix} X_1^* X_2^2 e^{-4ibz} \\ X_2^* X_1^2 e^{4ibz}\end{pmatrix} = 0.$$

In modern long distance communication fibers, the birefringence is assumed to be large. Consequently, the last nonlinear term is rapidly oscillating and will be neglected in the sequel. However, in weakly birefringent fibers these terms do not average out to zero and have to be kept in the equation. Therefore, considering large birefringence, the evolution of the electric field is well approximated by

$$i\frac{\partial X}{\partial z} + ib'\sigma_3\frac{\partial X}{\partial t} + \frac{d_0}{2}\frac{\partial^2 X}{\partial t^2} + \frac{5}{6}\left|X\right|^2 X + \frac{1}{6}\left(X^*\sigma_3 X\right)\sigma_3 X = 0. \tag{5.3.1}$$

5.3.1 Numerical schemes and error analysis

Propagation of solitons. The Manakov Equation (5.2.1) is a simplification of the CNLS Equation (5.3.1). Indeed, when the group velocity terms may be neglected ($b' = 0$), Equation (5.3.1) can be viewed as a nonlinear perturbation of the Manakov equation (5.2.1), since

$$\frac{5}{6}\left|X\right|^2 X + \frac{1}{6}\left(X^*\sigma_3 X\right)\sigma_3 X = \frac{8}{9}\left|X\right|^2 X + \frac{1}{9}\begin{pmatrix}\left(\left|X_1\right|^2 + 4\left|X_2\right|^2\right)X_1 \\ \left(4\left|X_1\right|^2 + \left|X_2\right|^2\right)X_2\end{pmatrix}. \tag{5.3.2}$$

In the Manakov Equation (5.2.1), the self-phase modulation and the cross-phase modulation (energy exchange between the two propagating modes) have the same strength and compensate exactly the chromatic dispersion, whereas it is not true for the CNLS equation (5.3.1). The Manakov equation has soliton solutions [54, 113, 114], that are of the form

$$X(z,t) = \sqrt{\frac{9}{8}}\begin{pmatrix}\cos\Theta/2\exp(i\phi_1) \\ \sin\Theta/2\exp(i\phi_2)\end{pmatrix}\eta\mathrm{sech}\eta(t-\tau(z))e^{-ik(t-\tau(z))+i\alpha(z)}. \tag{5.3.3}$$

Here, the polarization angle Θ, the phases ϕ_1, ϕ_2, the amplitude η and the group velocity $-k$ are arbitrary constants and the position τ and α are solutions of

$$\tau(z) = \tau_0 - kz \qquad \text{and} \qquad \alpha(z) = \alpha_0 + \frac{1}{2}\left(\eta^2 + k^2\right) z.$$

These type of solutions are thus of physical interest and will be used to check the validity of our numerical simulations. To test the efficiency of the schemes (5.2.2), (5.2.5) and (5.2.6), we define the relative errors in the $\mathbb{L}^2$ and $\mathbb{L}^\infty$ norms between the exact solution $\widetilde{X}^n$, evaluated at position z_n, and the approximated solution X^n, by either one of the previous scheme,

$$\mathrm{err}_2^n = \frac{\left\|\widetilde{X}^n - X^n\right\|_{\mathbb{L}^2}}{\|X_0\|_{\mathbb{L}^2}}, \qquad \mathrm{err}_\infty^n = \frac{\left\|\widetilde{X}^n - X^n\right\|_{\mathbb{L}^\infty}}{\left\|\widetilde{X}^n\right\|_{\mathbb{L}^\infty}}. \tag{5.3.4}$$

Moreover, it is also well-known that the Manakov equation possesses two invariants, which correspond to the mass and the Halmiltonian. A discrete version of these two invariants is given in Equation (5.2.3) and (5.2.4). To measure the ability of those schemes to preserve these two quantities, we introduce the errors

$$\mathrm{err}_{\mathbb{L}^2}^N = \max_{n\in[\![1,N]\!]} \left| \frac{\|X_N^n\|_{\mathbb{L}^2}^2 - \|X_0\|_{\mathbb{L}^2}^2}{\|X_0\|_{\mathbb{L}^2}^2} \right| \tag{5.3.5}$$

$$\mathrm{err}_H^N = \max_{n\in[\![1,N]\!]} \left| \frac{H\left(X^n\right) - H\left(X_0\right)}{H\left(X_0\right)} \right|. \tag{5.3.6}$$

For a first observation, we choose a soliton (5.3.3) as an input pulse, linearly polarized to 27.5°, i.e. $\Theta = \pi/4$ and $\phi_1 = \phi_2 = 0$. Moreover, we assume that the soliton is propagating with speed group 0.1, i.e. $k = -0.1$, and we set $\alpha_0 = \pi$. Table 5.1a compares the relative errors (5.3.4), (5.3.5) and (5.3.6) together with the CPU time for the three schemes. We observe that the Crank-Nicolson and the Relaxation schemes preserve exactly the discrete mass and the discrete Hamiltonian, while the Hamiltonian is not exactly preserved and oscillates for the Fourier split-step scheme. If the amplitude of the oscillations are relatively small (at the order of 10^{-4}) on this distance, it may become an issue on very long distance and for the study of soliton interactions. Numerically, we observe that the velocity and the amplitude of the soliton are well preserved, since the errors (5.3.4) are small for the three schemes. The computation time is bigger for the Crank-Nicolson scheme, the reason being that it is non-linearly implicit and has to be solved using a picard sequence. Table 5.1b displays numerical experiments of soliton propagation using the Relaxation scheme. Again, the errors (5.3.4), (5.3.5) and (5.3.6) are small, even if the grid is coarse. The Crank-Nicolson scheme and the Fourier split-step scheme give similar results. For the Hamiltonian conservation, the errors are respectively $1.7972e^{-11}$ and $3.3694e^{-7}$ in the first case, and respectively $2.4555e^{-11}$ and $1.3756e^{-7}$ in the second case. However, in some cases we observe that the numerical information does not travel as fast as the soliton. This situation happens when the value of the amplitude η or/and the group velocity are important. The numerical tests, displayed in Fig.

5.1a and 5.1b, were carried out choosing an amplitude $\eta \leqslant 1$ and a group velocity $|k| \leqslant 1$. In these cases, there is, a priori, no condition required on Δt and Δz for the numerical scheme to converge. However, if the amplitude or the group velocity are taken large, both Δt and Δz have to be carefuly chosen, depending on η and k, for the numerical information to follow the high group velocity of the wave packet. This fact is illustrated in Table 5.2.

Collision of two solitons. In this paragraph, we test soliton interference using multi-soliton initial data (5.3.7). This denomination comes from the fact that, when z goes to infinity, the solution becomes a superposition of several solitons, which are separated from each other. The initial $2-$solitons condition is simply chosen as a superposition of two Manakov solitons that is

$$X(0,t) = \mathbf{c}_1\eta_1 \mathrm{sech}\eta_1(t-\tau_{10})e^{-ik_1(t-\tau_{10})+i\alpha_{10}} + \mathbf{c}_2\eta_2 \mathrm{sech}\eta_2(t-\tau_{20})e^{ik_2(t-\tau_{20})+i\alpha_{20}}, \tag{5.3.7}$$

where the polarization vectors are given by

$$\mathbf{c}_1 = \sqrt{\frac{9}{8}}\begin{pmatrix}\cos\Theta_1/2\exp(i\phi_{11})\\ \sin\Theta_1/2\exp(i\phi_{12})\end{pmatrix} \qquad \text{and} \qquad \mathbf{c}_2 = \sqrt{\frac{9}{8}}\begin{pmatrix}\cos\Theta_2/2\exp(i\phi_{21})\\ \sin\Theta_2/2\exp(i\phi_{22})\end{pmatrix}.$$

In the case of the scalar nonlinear Schrödinger equation, solitons also have strong elasticity properties, i.e. if two solitons interact, they will emerge unchanged from the collision, except for a phase shift. However, the collision of Manakov solitons is far more complex than in the scalar case. The slow evolution of the soliton parameters (speed group k, polarization Θ, amplitude η, etc.) are derived analytically and numerically in [96, 113, 114, 115]. From analytical expressions, the polarization vectors $\mathbf{c}_k$, $k = 1, 2$ may change after the collision, due to a redistribution of energy between the two components. In the general case, the solitons emerge from the collision with different polarization states. However, the power of each soliton is preserved and, in that sense, we may say that the collision is elastic. In the particular case when the polarization vectors are parallel or orthogonal, the 2-soliton behaves similarly as in the scalar situation. Moreover, the velocity of the solitons may changed after the collision (see [96, 113, 114, 115]).

Now, we study the collision of two Manakov solitons. Accordingly, we choose (5.3.7) as an initial input for the Manakov equation (5.2.1) and for the Manakov Equation subject to the nonlinear perturbations (5.3.2). To test the interaction between two solitons, we take sufficiently distant supports to ensure that the interaction has not begun at initial state. Figure 5.1 displays the evolution of a 2-soliton with arbitrary polarization vectors and phases. Table 5.3a shows the soliton parameters used for the simulations. The simulations were done for the three schemes and the results are summarized in Table 5.3b that displays the relative errors

$$\mathrm{err}^N_{l,\mathbb{L}^2} = \max_{n\in[\![1,N]\!]}\left|\frac{\|X_l^n\|^2_{\mathbb{L}^2} - \|X_l(0)\|^2_{\mathbb{L}^2}}{\|X_l(0)\|^2_{\mathbb{L}^2}}\right| \qquad \text{for } l = 1, 2$$

measuring the ability of the schemes to preserve the mass of each component. When the two solitons emerge from the collision, we observe that their shapes have changed. According to the previous discussion, this is due to the change of polarization state induced by the collision. Moreover, we observe that the speed of the solitons has increased. Indeed, in this case, the group velocity of the 2-soliton is $k = 0.15$ and it propagates over a distance $Z = 120$, starting at two different locations $\tau_{10} = 10$ and $\tau_{20} = -8$ (see Table 5.3a). Therefore, if the speed of the 2-soliton were constant, the final locations should have been $\tau_1 = -8$ and $\tau_2 = 10$, whereas the maximum amplitudes are located at $\tau_1 = -11.57$ and $\tau_1 = 11.47$. From Table 5.3b, we observe that, for the three schemes, the mass of each component is preserved after the collision, but the Relaxation scheme gives more accurate results. As expected, the discrete Hamiltonian is not preserved by the Fourier split-step method, since the two dispersive effects are treated separately. On the contrary, both the Crank-Nicolson and the Relaxation schemes still preserve the discrete Hamiltonian. However, the results are less accurate than those obtained for the propagation of a single soliton. Indeed, when the interaction begins, the schemes fail to preserve the Hamiltonian exactly. The next numerical simulations are done using the Relaxation scheme. Figure 5.2 displays the evolution of a 2-soliton with orthogonal polarization vectors, i.e. $|\Theta_1 - \Theta_2| = \pi$. The two pulses emerge unchanged from the collision and we recover the usual behaviour of the NLS solitons. Finally, in the latter case (Fig. 5.3), we consider the same initial data and we study its propagation when subject to the nonlinear perturbations given by (5.3.2). In [113, 114], numerical simulations show that, depending on the speed of the solution, the solitons may either go through and continue or be reflected and go back. Indeed, as explained in [114], the speed of the soliton decreases during the collision and increases when emerging from the collision. If the initial speed is relatively small, then the sign of the velocity will change during the collision, resulting in reflection. Figure 5.3 shows that the higher soliton gets higher, while the smaller soliton gets smaller. This is entirely due to the nonlinear perturbations, since in Fig. 5.2, the two solitons pass through each other.

Numerical order of convergence. In section 5.2, we mentioned that the Crank-Nicolson and the Relaxation schemes are of order 2, in the z variable, whereas the Lie Fourier split-step is of order 1. In this section, we illustrate this fact numerically. Again, we choose a Manakov soliton (5.3.3) as an input pulse. The initial input is linearly polarized to 45°, i.e. $\Theta = \pi/2$ and $\phi_1 = \phi_2$, with amplitude η equal to 1. We compare the numerical approximations, by either one of the three schemes (5.2.2), (5.2.5) and (5.2.6), with the exact expression of the Manakov soliton (5.3.3). The discretization and soliton parameters used for the simulations are stated in Table 5.4a. In Figure 5.4, the convergence curves are plotted with respect to the space step Δz. To obtain such curves, we fix the time step Δt. Since we are only concerned with the space order of convergence, the time step Δt is taken sufficiently small, so that the time error does not interfere in the total error of the numerical approximation. The numerical approximation is computed on five different grids. The coarsest grid is given by the space step $\Delta z = 0.1$. From each grid, we define a finer grid, which is twice as small. Therefore, the smallest grid is given by $\Delta z = 0.00625$. The convergence curves, drawn in Figure 5.4, correspond to the logarithm of the

errors (5.3.4) at each space step. The slope of these curves is compared with the slope of the solid line, which gives the rate of convergence of the method. For the Crank-Nicolson and the relaxation schemes, the solid line represents the curve $\log\left((\Delta z/0.01)^2\right)$, while for the Fourier split-step method, the solid line represents the curve $\log(\Delta z/0.01)$. Table 5.4b displays the numerical errors (5.3.4) computed on the finest mesh.

	Fourier split-step	Crank-Nicolson	Relaxation
Parameters	$\eta = 1$, $\tau_0 = -5$, $T = 20$, $N = 4000$, $Z = 100$, $M = 10000$		
err_∞^N	$2.6443e^{-2}$	$8.7848e^{-4}$	$9.1168e^{-4}$
err_2^N	$2.8407e^{-2}$	$8.8481e^{-4}$	$9.2834e^{-4}$
$\text{err}_{\mathbb{L}^2}^N$	$9.4088e^{-13}$	$1.0545e^{-9}$	$2.9211e^{-14}$
err_H^N	$6.1812e^{-5}$	$3.2003e^{-9}$	$5.7923e^{-11}$
CPU time	39.40s	160.7s	41.32s

(a)

	Numerical experiments using the Relaxation scheme	
Parameters	$k = -1$, $\eta = 1/2$, $\tau_0 = -10$ $T = 40$, $Z = 25$ $N = 8000$, $M = 2500$	$k = -1$, $\eta = 1/6$, $\tau_0 = -25$ $T = 160$, $Z = 50$ $N = 3200$, $M = 500$
err_∞^N	$4.0568e^{-4}$	0.0298
err_2^N	$4.6158e^{-4}$	0.0305
$\text{err}_{\mathbb{L}^2}^N$	$1.4211e^{-14}$	$1.3175e^{-14}$
err_H^N	$1.0219e^{-11}$	$1.1964e^{-14}$

(b)

Table 5.1: Table 5.1a compares the schemes (5.2.2), (5.2.5) and (5.2.6). It displays the numerical approximation errors in the $\mathbb{L}^2$ and $\mathbb{L}^\infty$ norms together with the relative errors for the conservation of the two invariants. Table 5.1b displays numerical results using the Relaxation scheme.

	High amplitude	
Parameters	$k = -1$, $\eta = 10$, $\tau_0 = -2.5$ $T = 6$, $Z = 5$	
	$N = 1200$, $M = 500$	$N = 12000$, $M = 5000$
err_∞^N	1.0147	0.0412
err_2^N	1.4450	0.0502
$\text{err}_{\mathbb{L}^2}^N$	$1.2158e^{-14}$	$3.8764e^{-13}$
err_H^N	$2.4626e^{-5}$	$1.5741e^{-10}$

(a)

	High group velocity	
Parameters	$k = -10$, $\eta = 1$, $\tau_0 = -5$ $T = 20$, $Z = 1$	
	$N = 4000$, $M = 100$	$N = 40000$, $M = 1000$
err_∞^N	1.0260	0.0116
err_2^N	1.0365	0.0117
$\text{err}_{\mathbb{L}^2}^N$	$6.1186e^{-15}$	$3.5922e^{-14}$
err_H^N	$1.9192e^{-13}$	$1.8764e^{-12}$

(b)

Table 5.2: Various situations where the numerical information does not travel as fast as the soliton, when the space step Δz and the time step Δt is not small enough. In the first table 5.2a, the value of the amplitude η is chosen large, the other parameters being fixed. In the left column, the relative errors indicate that the numerical approximation is not accurate. In the right column, the results are more satisfactory. In the second table 5.2b, the value of the group velocity is taken large. The numerical simulations are performed using the Relaxation scheme.

	Fig. 5.1	Fig. 5.2 & Fig. 5.3
Soliton parameters	$\phi_{11} = \phi_{12} = -\pi$ $\Theta_1 = \pi/2 + 0.5, \quad \Theta_2 = \pi/2 - 0.3$ $\phi_{21} = \phi_{22} = \pi/6$ $\eta_1 = 1/3, \quad \eta_2 = 1$	$\phi_{11} = \phi_{12} = -\pi$ $\Theta_1 = 0.7, \quad \Theta_2 = \pi + 0.7$ $\phi_{21} = \phi_{22} = \pi$ $\eta_1 = 1/1.2, \quad \eta_2 = 1$
	$\alpha_{10} = \alpha_{20} = 0, \quad k = 0.15, \quad \tau_{10} = 10, \quad \tau_{20} = -8$	
Discretization parameters	$T = 50, \quad M = 6000, \quad Z = 120, \quad N = 6000$	

(a)

	Crank-Nicolson	Relaxation	Fourier split-step
$\text{err}^N_{1,\mathbb{L}^2}$	$1.8208e^{-8}$	$3.7345e^{-15}$	$5.5924e^{-9}$
$\text{err}^N_{2,\mathbb{L}^2}$	$1.9174e^{-8}$	$3.669e^{-14}$	$1.2868e^{-8}$
$\text{err}^N_{\mathbb{L}^2}$	$1.8642e^{-8}$	$2.5087e^{-14}$	$2.8729e^{-8}$
err^N_H	$8.1604e^{-8}$	$4.4385e^{-9}$	0.0107

(b)

Table 5.3: Table 5.3a displays discretization and 2-soliton parameters used for Fig. 5.1, Fig. 5.2 and Fig. 5.3. Table 5.3b displays the relative errors $\text{err}^N_{1,\mathbb{L}^2}$ and $\text{err}^N_{2,\mathbb{L}^2}$ related to the mass conservation of each component of the soliton after the collision. These errors are computed only for the numerical simulations displayed in Fig. 5.1.

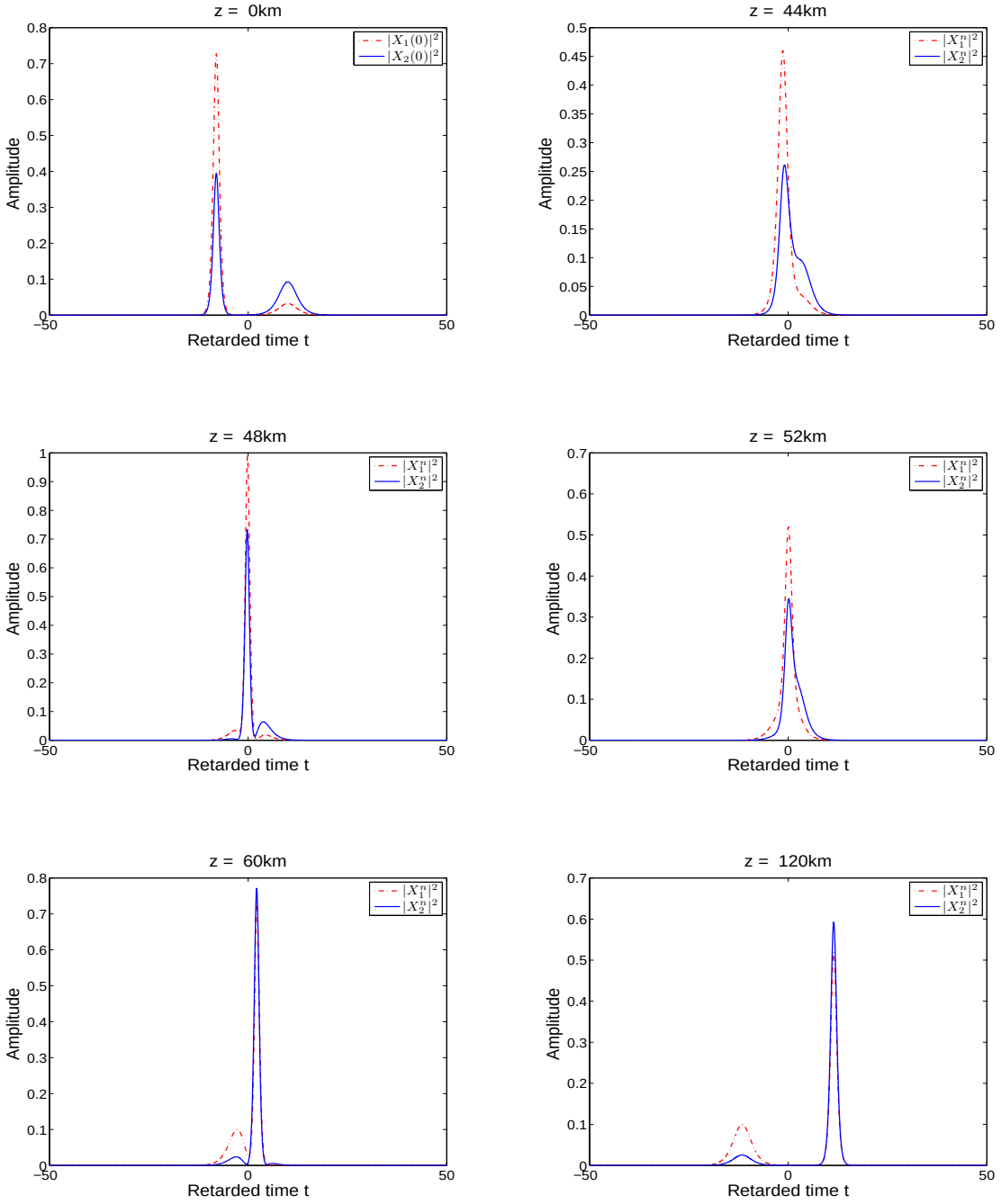


Figure 5.1: Elastic interaction between two Manakov solitons, solution of the Manakov equation (5.2.1), whose components have the same speed. At the end of the fiber, the polarization vectors of the 2−soliton have changed. However, the mass of each component remains conserved, as shown in Table 5.3b.

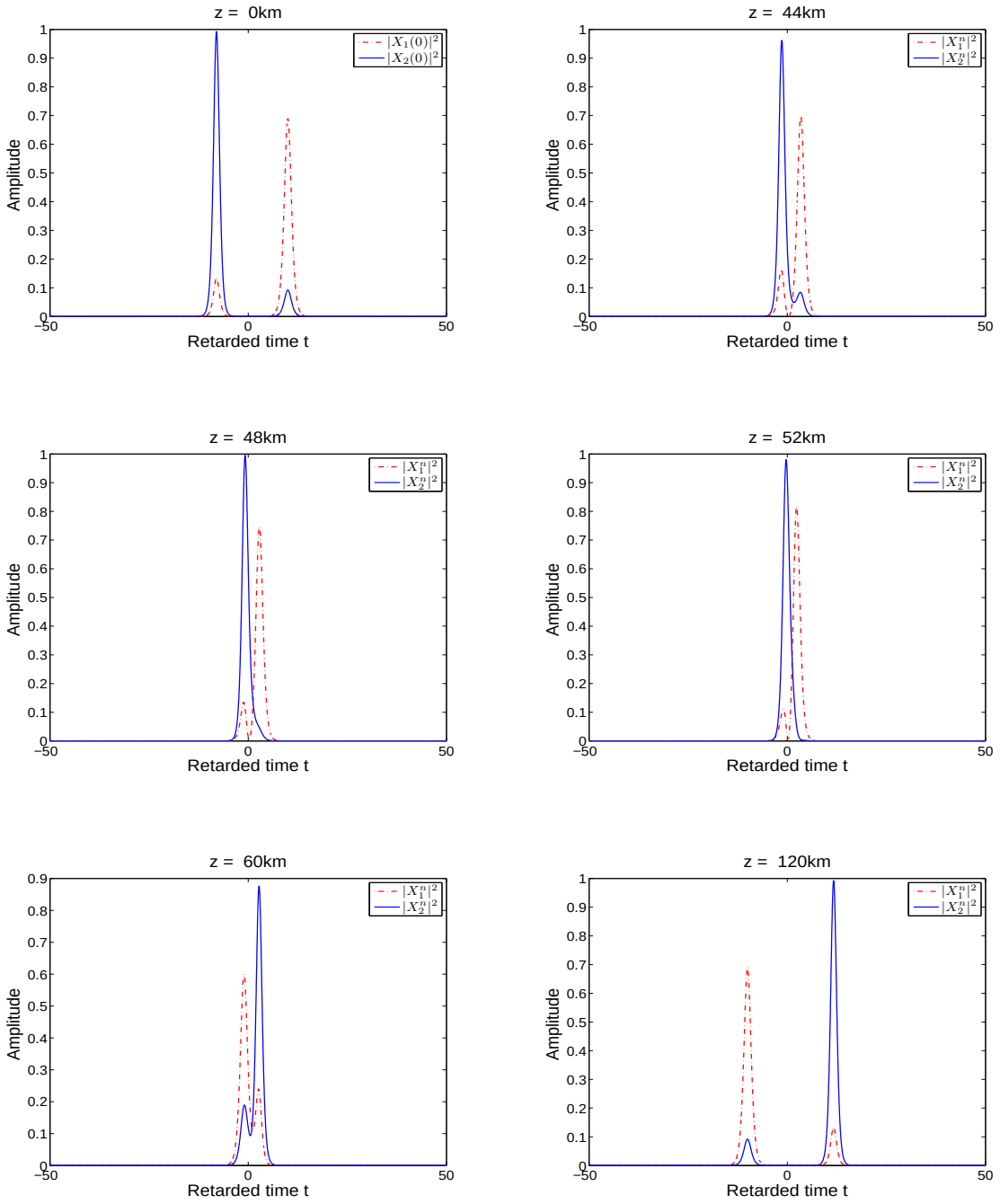


Figure 5.2: Elastic interaction between two Manakov solitons, whose polarization vectors are orthogonal. The two solitons remain unchanged except that they shift their positions. In this case, the behaviour is similar to the one observed for NLS solitons.

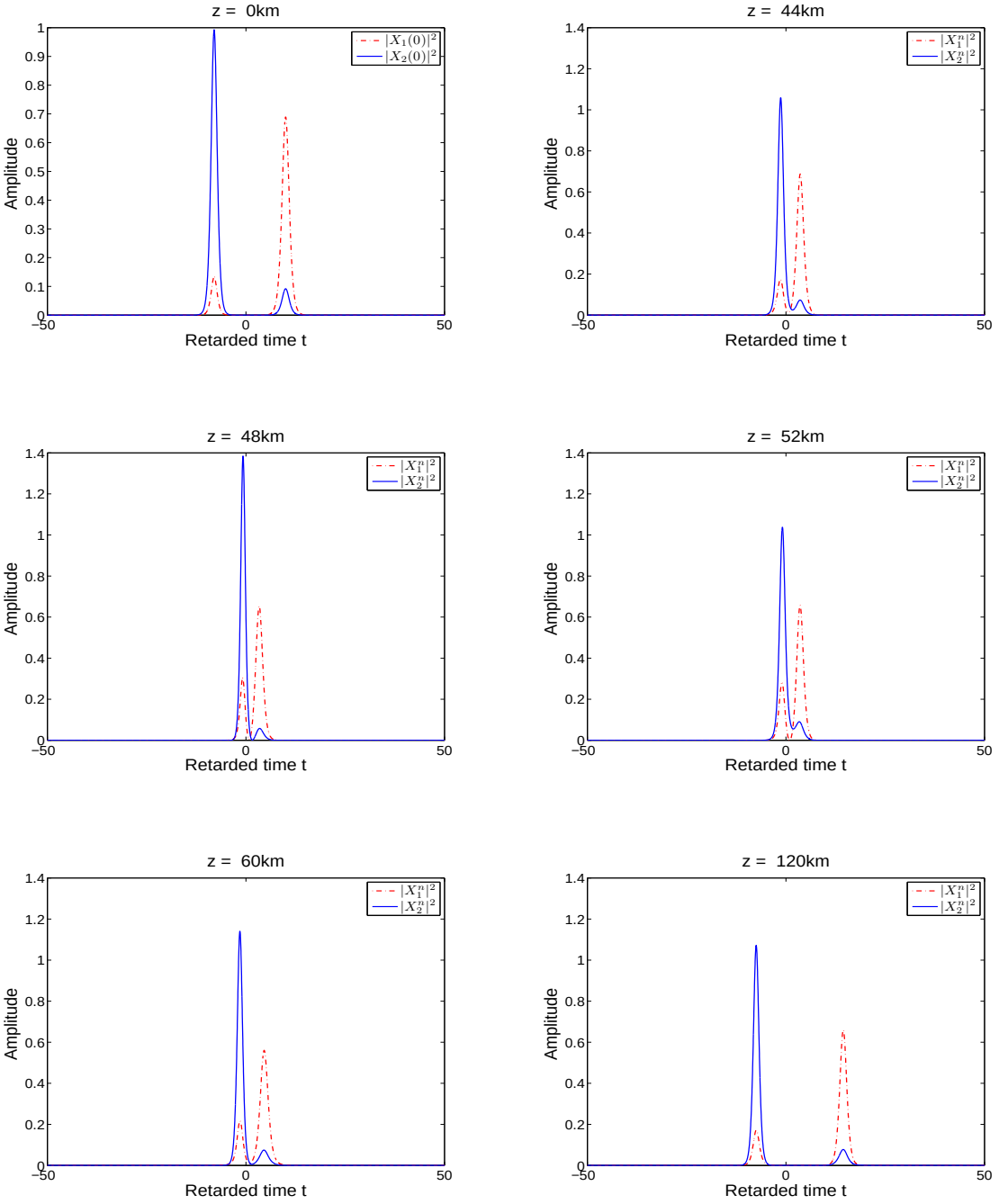


Figure 5.3: Inelastic interaction between two solitons of Manakov equation subject to nonlinear perturbations (5.3.2). Since the velocity of the pulses is relatively low, the two solitons repeal each other.

Soliton parameters	Discretization parameters
$\Theta = -\pi/2, \quad \phi_1 = \phi_2 = 0$	$T = 20, \quad M = 10000$
$\alpha_0 = \pi, \quad k = -0.5, \quad \eta = 1, \quad \tau = -2.5$	$Z = 4, \quad N_{\text{coarse}} = 20, \quad N_{\text{fine}} = 640$

(a)

	Crank-Nicolson	Relaxation	First order Fourier split-step
err_2^N	$6.1553e^{-6}$	$2.1230e^{-5}$	$2.666e^{-3}$
err_∞^N	$5.1282e^{-6}$	$1.9301e^{-5}$	$1.8587e^{-3}$
$\text{err}_{\mathbb{L}^2}^N$	$8.3353e^{-12}$	$3.1777e^{-14}$	$4.433e^{-13}$
err_H^N	$1.8394e^{-10}$	$3.5796e^{-12}$	$8.693e^{-5}$
CPU time	44.669s	30.8169s	17.4563s

(b)

Table 5.4: Table 5.4a displays soliton and discretization parameters used for the convergence curves in Fig. 5.4. Table 5.4b displays the numerical errors for $\Delta z = 0.00625$, which correspond to Fig. 5.4.

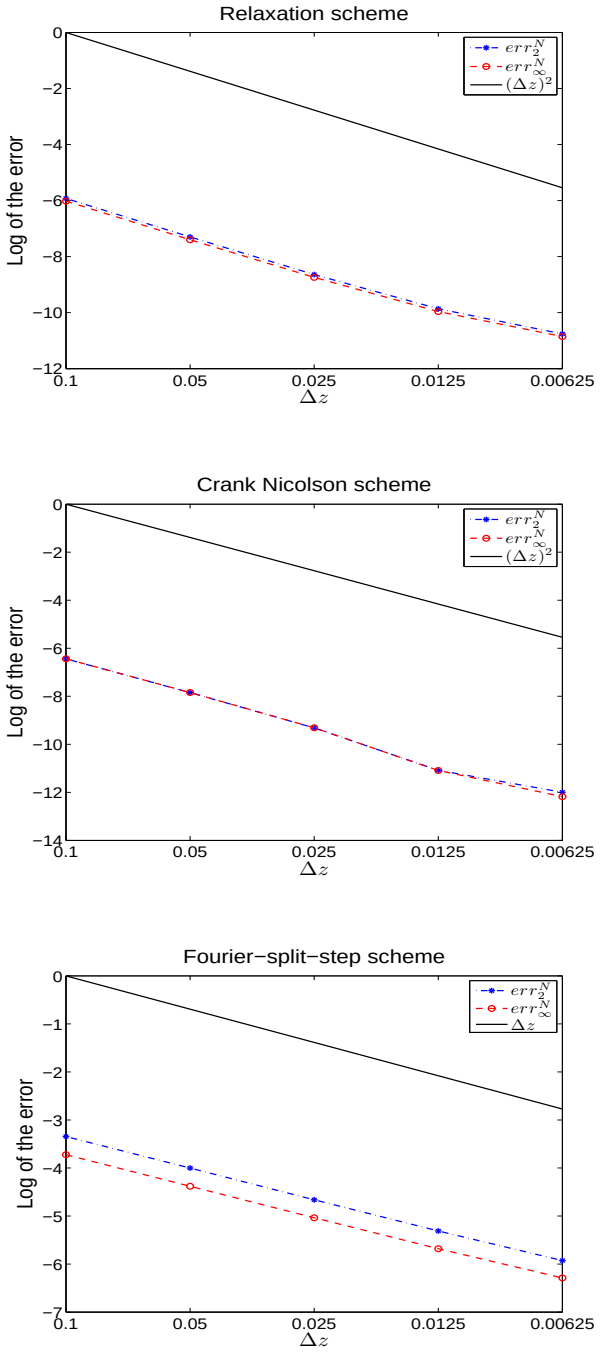


Figure 5.4: Plot of the log of the relative errors err_2^N and err_∞^N for various numerical schemes.

5.3.2 Wave propagation in optical fibers with constant birefringence

In this section, we proceed to the second step of the numerical validation process. Accordingly, we study the impact of uniform birefringence on soliton propagation and we compute the Differential Group Delay parameter (DGD). The numerical results are compared with the theory that may be found in [2, 14, 37, 39, 42].

As explained in Chapter 2, birefringence is a physical phenomenon appearing in transparent medium (such as silica and optical fibers) and resulting from internal and external stress or distortions. As a consequence, the light splits into two orthogonally polarized pulses propagating with different characteristics along two orthogonal axes, labelled as the slow and fast birefringence axes or eigenmodes. In the particular case of linear birefringence, the two polarization states correspond to linear polarizations. The description of these effects is often characterized by the PMD vector $\boldsymbol{\tau}$ and the square DGD parameter $\tau = |\boldsymbol{\tau}|^2$, which is the difference of group velocity between the two components. The square DGD parameter may be expressed either in time domain or in frequency domain. However, it is often convenient to express it in the frequency domain, thanks to the Fourier transform. We recall that the normalized Stokes vector of the Fourier transform of the electric field $\widehat{X}$ is given by

$$\widehat{s}(z,\omega) = \frac{\widehat{X}^* \overrightarrow{\sigma} \widehat{X}(z,\omega)}{\widehat{s}_0(z,\omega)}, \qquad z \geqslant 0, \tag{5.3.8}$$

where $\overrightarrow{\sigma} = (\sigma_0, \sigma_3, \sigma_1, \sigma_2)$ and the spectral intensity $\widehat{s}_0(z,\omega) = \left|\widehat{X}(z)\right|^2$. If the birefringence is assumed to be uniform (i.e. (b,θ) is constant) along the fiber, then the CNLS equation (5.3.1) is a good approximation of the evolution of the slowly varying envelope. In the absence of GVD and nonlinear effects, Equation (5.3.1) reduces to a transport equation

$$i\frac{\partial X}{\partial z} + ib'\sigma_3 \frac{\partial X}{\partial t} = 0. \tag{5.3.9}$$

When arbitrarily polarized wave is injected into such fibers, its components along the two eigenstates propagate without distortions. The two components travel with the same velocity but in opposite directions. In this case, the spectral intensity is constant along the fiber and the stokes vector is given by

$$\widehat{s}(z,\omega) = \frac{\widehat{X}^* \overrightarrow{\sigma} \widehat{X}(z,\omega)}{\widehat{s}_0(\omega)} \tag{5.3.10}$$

and the formal description of the birefringence in Stokes space is expressed by the set of equations [37, 39, 53]

$$\begin{aligned} \partial_z \widehat{s} &= \beta \times \widehat{s} \\ \partial_\omega \widehat{s} &= \boldsymbol{\tau} \times \widehat{s} \\ \partial_z \boldsymbol{\tau} &= \beta' + \beta \times \boldsymbol{\tau}, \end{aligned} \tag{5.3.11}$$

where β is the birefringence vector (and may be random), β' is the frequency-derivative of β and $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3)$ is the PMD vector. The $\times$ denotes the cross product which can be expressed as the product of a skew-symmetric matrix and a vector. Note that the last equation is obtained combining the first two equations. By a Taylor expansion around the carrier frequency, the dynamics of the PMD vector may be rewritten as [39, 46]

$$\partial_z \boldsymbol{\tau} = \beta' + \Delta\omega \beta' \times \boldsymbol{\tau}. \tag{5.3.12}$$

The birefringence vector, in Equation (5.3.9), is given by $\beta'(\omega_0) = (2b', 0, 0)^t$. The evolution of the stokes vector over the Poincaré sphere, at a fixed frequency, is a circle centered around the birefringence vector, whose orientation depends on the initial state of polarization (see Chapter 2 and [39]). If the birefringence is step-wise constant in z, the Stokes vector draws arches on the Poincaré sphere. Therefore the evolution of the PMD vector and the DGD parameter may be computed through the dynamical equation (5.3.12). It can also be computed through the viriel (or pulse width). Indeed, by the Fourier Plancherel formula

$$V(X(z)) = \int_{\mathbb{R}} t^2 \left|X(z,t)\right|^2 dt = \int_{\mathbb{R}} \widehat{R}(z,\omega)\widehat{s}_0(z,\omega) d\omega, \tag{5.3.13}$$

where $\widehat{R}(z,\omega) = \left|\widehat{X}'(z,\omega)\right|^2 / \widehat{s}_0(z,\omega)$ and the prime denotes the frequency derivative. From equation (5.3.9), we easily obtain that

$$\partial_z \widehat{R}(z,\omega) = -\frac{2b'}{\widehat{s}_0(\omega)} \mathcal{R}\mathrm{e}\left\{i\overline{\widehat{X}'}\sigma_3\widehat{X}\right\} = 2\left(b'\right)^2 z.$$

Moreover, from Equation (5.3.12), we obtain that $\partial_z \tau(z) = 8\left(b'\right)^2 z$. Thus the process $\widehat{R}(z,\omega)$ is the square DGD parameter, up to a multiplicative constant, and is given by

$$\tau(z) = \left|\boldsymbol{\tau}(z)\right|^2 = 4\left(\widehat{R}(z,\omega) - \widehat{R}(0,\omega)\right), \qquad z \geqslant 0.$$

For nonlinear pulses, the birefringence vector is nonlinear and not known, and the above description (5.3.12) is not valid. Therefore, the evolution of the DGD parameter is approximated as follows : we compute a numerical approximation of Equation (5.3.1) by one of the three schemes introduced in Section 5.2. From this approximation, we compute the evolution of the stochastic process $\widehat{R}(z,\omega)$.

We now illustrate numerically the physical facts mentioned above. The set of parameters used for the simulations are given in Table 5.5. The simulations are done using the Relaxation scheme, that is easily adapted to the CNLS equation (5.3.1). Figure 5.5 illustrates the evolution of the Stokes vector over the Poincaré sphere at $z = 2$km (first row) and $z = 10$km (second row). We consider input pulses with two different initial polarization states. On the left column, the pulse is initially circularly polarized i.e. the stokes vector is given by $(0, 0, 1)$. Therefore, it starts at the north pole. On the right hand side, the initial pulse is linearly polarized to 67.5°. The stokes vector is now given by $(-0.7071, -0.7071, 0)$ and starts from the equator. The blue line corresponds to the evolution of the polarization states of a linear

pulse, solution of Equation (5.3.9). Let us note that the length of the curve on the Poincaré sphere depends on the frequency. We clearly see that it evolves periodically from circular polarization to elliptic to linear, etc. This periodic evolution was also illustrated in Fig. 2.10. The red lines represent the evolution of the normalized stokes vector (5.3.8) related to the solution of Equation (5.3.1). The evolution of the stokes vector is now almost periodic and moves faster over the Poincaré sphere. The interaction between the birefringence and the phase modulation induced by Kerr effects has a small impact on the evolution of the polarization states.

Fig. 5.6 exhibits the evolution of the pulse intensity along the fiber. We still consider two different initial polarization states in both linear and nonlinear medium. In both cases and if $b' \neq 0$, the two components split apart and travel with the same velocity but in opposite directions. In the linear case with zero GVD, the shape of the pulses remains unchanged. The last row, in Figure 2.10, displays the evolution of the DGD along the fiber. As expected, the growth of the DGD is linear with the distance [39] and corresponds exactly to the spreading between the two pulses. Indeed, the velocity of the pulses is given by the frequency derivative of the birefringence strength i.e. $b' = 1.19$ps/km. Moreover, the two pulses are initially located at zero. Therefore, after a propagation distance of $Z = 10$km, one of the two components is located at $\tau = 11.9$ps, while the other is located at $\tau = -11.9$ps. Hence, at the end of the fiber, the DGD parameter, defined as the spreading between the two pulses, is equal to $\sqrt{\tau(Z)} = 2b'Z = 23.8$ps/km.

<table>
<tr><td></td><td>Circularly polarized</td><td>Linearly polarized to 67.5°</td></tr>
<tr><td rowspan="2">Soliton parameters</td><td colspan="2">$b' = 1.19, \quad \phi_1 = 0$
$\alpha_0 = 0, \quad k = 0, \quad \eta = 1/2, \quad \tau = 0.$</td></tr>
<tr><td>$\Theta = \pi/2, \quad \phi_2 = \pi/2$</td><td>$\Theta = -3\pi/4, \quad \phi_2 = 0$</td></tr>
<tr><td>Discretization parameters</td><td colspan="2">$T = 60, \quad M = 12000, \quad Z = 10, \quad N = 1000$</td></tr>
</table>

Table 5.5: Soliton and discretization parameters used for Fig. 5.5 and 5.6 .

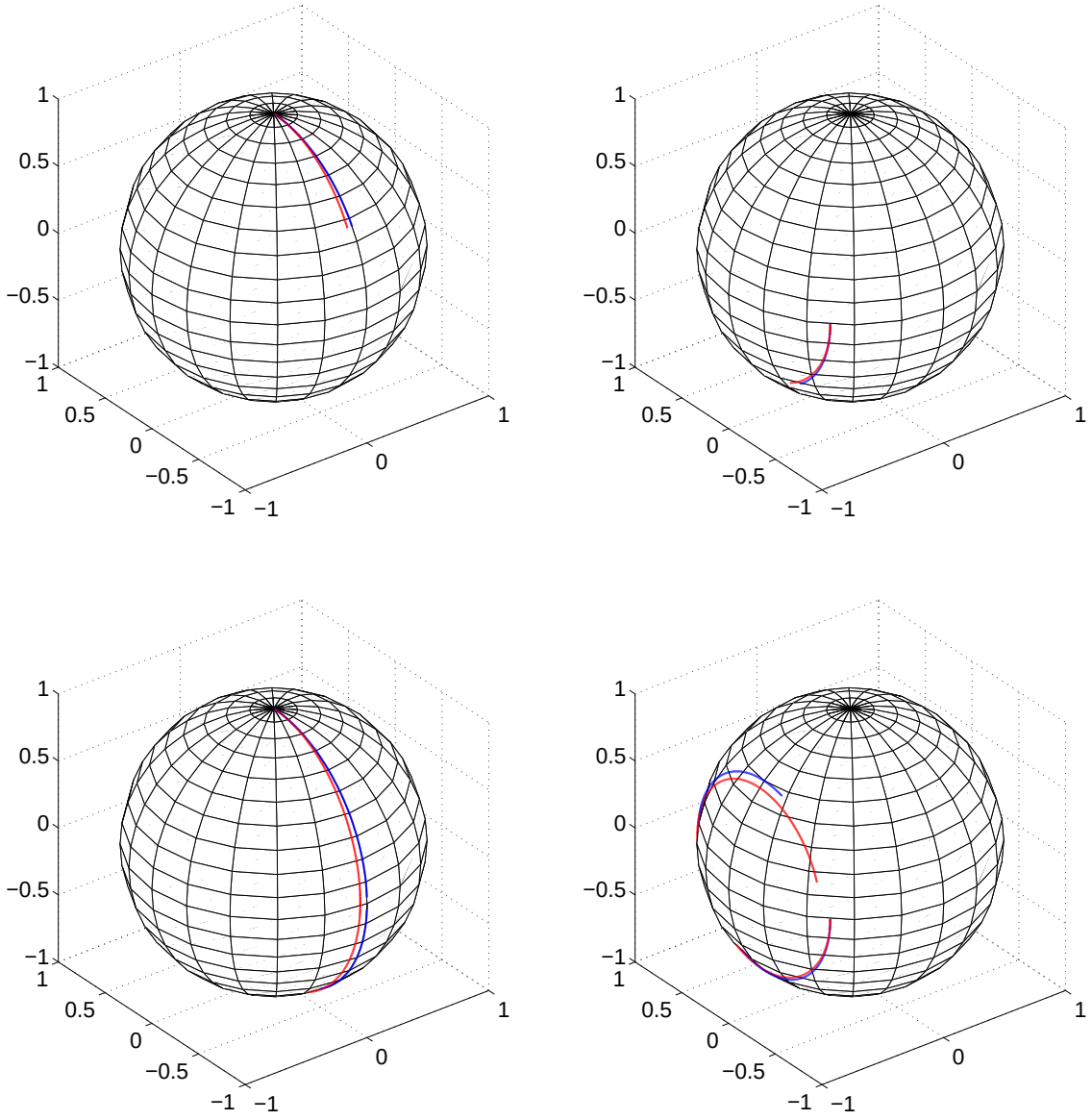


Figure 5.5: Evolution of the Stokes vector, at $z = 2$km (first row) and $z = 10$km (second row), over the Poincaré sphere. On the left, the pulse is initially circularly polarized, while on the right hand side, its is linearly polarized to 67.5°. The red curve corresponds to the normalized stokes evolution (5.3.8) in nonlinear medium and the blue curve is related to the stokes evolution in linear medium assuming zero GVD.

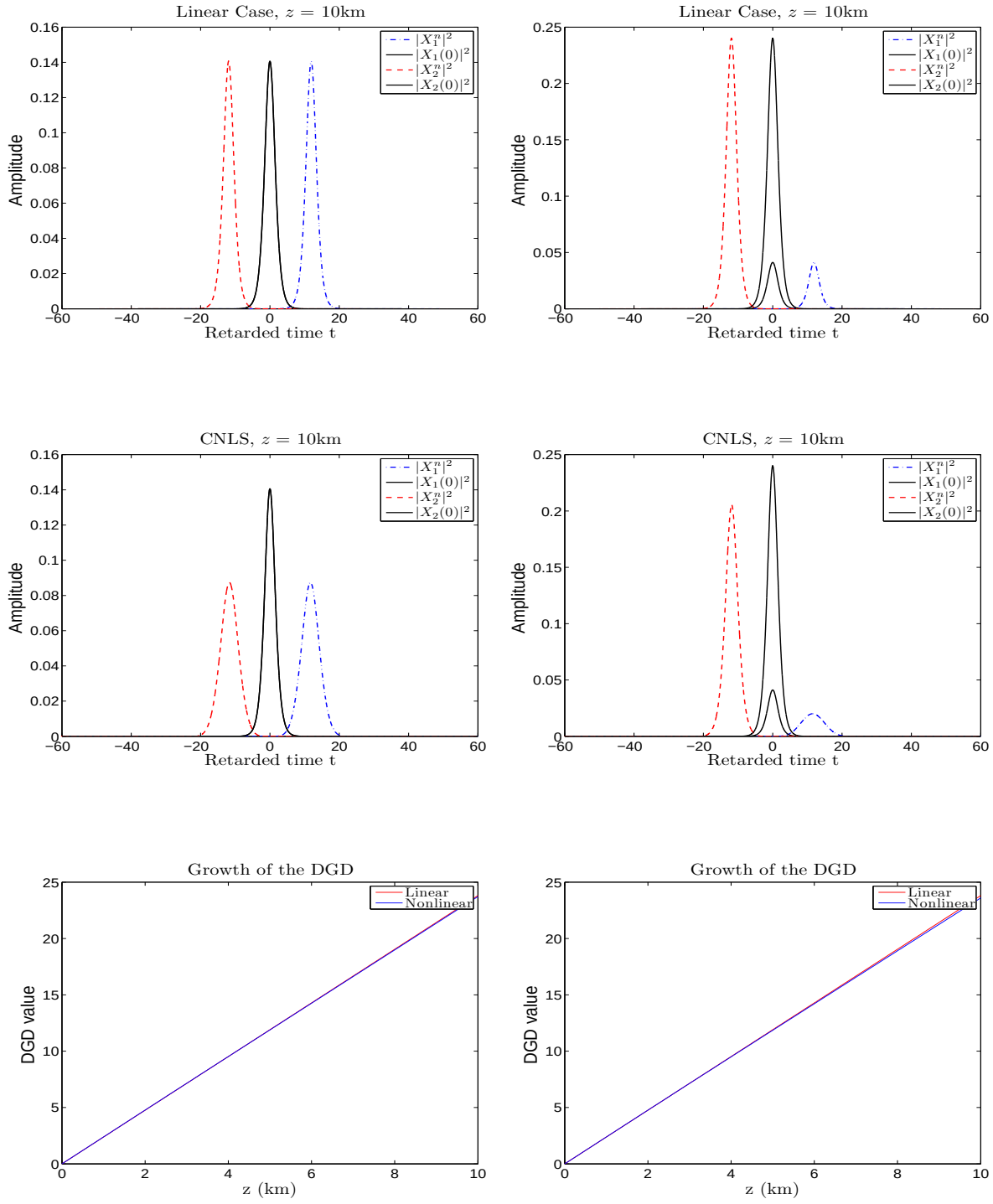


Figure 5.6: Evolution of the Manakov soliton subject to uniform birefringence in the linear case with zero GVD and the nonlinear case. The last row shows the evolution of the DGD along the fiber. As expected the growth of the DGD is linear with the distance and corresponds exactly to the spread between the two pulses.

5.4 Numerical simulations of light propagation in fibers with randomly varying birefringence

Polarization Mode Dispersion (PMD) does not simply reduce to the Differential Group Delay (DGD). Another significant manifestation of the PMD is the Polarization Mode coupling, that is the energy exchange between the two components. This dispersion effect is induced by the random variations of the birefringence and leads to pulse distortions. Thus, the model with constant birefringence (5.3.1) is too simple to completely describe the PMD effects. Physicists have proposed different models for the stochastic evolution of the birefringence [37, 60, 110] and a lot of research has been performed in order to determine theoretically and numerically the statistics of the PMD [7, 8, 30, 37, 39, 42, 53, 110].

5.4.1 Numerical almost sure error analysis

In this section, we numerically compute the almost sure order of convergence of the three schemes (5.2.8), (5.2.9) and (5.2.10), that have been introduced in Section 5.2. In Chapter 4, we obtained a theoretical a.s. order for a semi-discrete Crank-Nicolson scheme. This result states that the almost sure order is 1/2 i.e. for any stopping time $\tau < \tau^*$ and $\delta < \frac{1}{2}$, there exists a random variable K_δ such that

$$\max_{n=0,\dots,N_\tau} \left\| X^n - \widetilde{X}^n \right\|_{\mathbb{H}^1} \leqslant K_\delta(T,R,\omega)\Delta z^\delta,$$

where X^n is the approximation of Equation (5.1.1), by the scheme (4.1.3), and $\widetilde{X}^n$ is the exact solution evaluated at position z_n. The set of parameters used for the simulations are given in the following Table 5.6. Contrarily to the Manakov equation

	Almost-sure order
Soliton parameters	$\phi_1 = 0, \quad \phi_2 = 0, \quad \Theta = -\pi/2, \quad k = 0$ $\eta = 1/2, \quad \alpha_0 = \pi, \quad \tau = 0, \quad \gamma = 0.1$
Discretization parameters	$T = 30, \quad M = 20000, \quad Z = 4$ $N_{\text{coarse}} = 40, \quad N_{\text{fine}} = 2520$

Table 5.6: Set of parameters used to obtain the almost sure order.

(5.2.1), there is no explicit solutions for this equation. Therefore, we first compute an approximated solution X^n of Equation (5.1.1) on a fine mesh $\Delta z = Z/N_{\text{fine}}$, that we compare to approximations of the same equation on coarser grids. As in section 5.3, a coarser grid, in the z variable, is twice as big as the previous one. The Brownian path is keeping fixed for each approximation as well as the time step Δt. Figure 5.7 displays two convergence curves corresponding to the logarithm of the relative errors (5.3.4). The slopes of these curves are compared to a curve with slope

1/2. From Fig. 5.7, we see that the almost sure order of the Crank Nicolson scheme is 1/2 in the z variable. Hopefully, this result agrees with the theoretical analysis of Chapter 4. The two other schemes also seem to be of almost sure order 1/2.

	Crank-Nicolson	Relaxation	First order Fourier split-step
err_2^N	$3.514e^{-3}$	$2.667e^{-3}$	$4.494e^{-3}$
err_∞^N	$3.633e^{-3}$	$2.512e^{-3}$	$3.896e^{-3}$
$\mathrm{err}_{\mathbb{L}^2}^N$	$6.299e^{-11}$	$1.821e^{-14}$	$3.8813e^{-13}$
CPU time	251.87s	172.09s	100.79s

Table 5.7: Numerical values of relative errors for $\Delta z = 0.00625$.

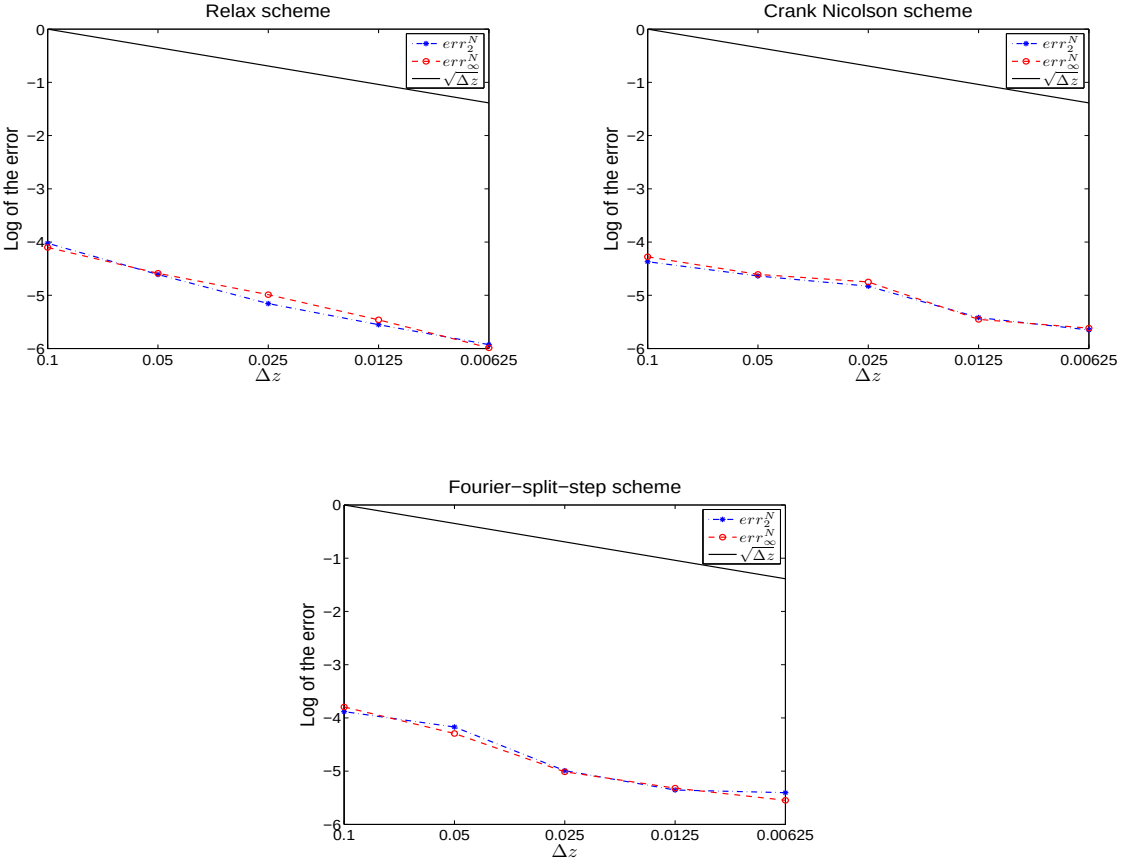


Figure 5.7: Plot of the log of the relative errors err_2^N and err_∞^N for the three schemes (5.2.8), (5.2.9) and (5.2.10).

5.4.2 Modelling of PMD due to random changes of birefringence for linear pulses in absence of GVD

5.4.2.1 Dynamical evolution of the PMD

In this subsection, we derive the evolution of the PMD vector (see [39, 42]) from the linear stochastic Manakov equation with zero GVD. This equation is obtained removing the nonlinear and second order terms from equation (5.1.1)

$$idX + i\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k\frac{\partial X}{\partial t}\circ dW_k(z) = 0. \tag{5.4.1}$$

This is a simplification of the nonlinear stochastic Manakov Equation (5.1.1) and is well adapted for studying PMD impact on pulse spreading and distortion. This simplified model allows to study dispersive effects resulting only from PMD.

In the previous section, we stated the dynamical equations (5.3.11) for the evolution of the stokes vector and the PMD vector. Physicists often assume that the birefringence vector β is a white noise [37, 39]. However, in this case the description (5.3.11) becomes formal and a meaning has thus to be given to those equations. In [37], the authors choose to interpret them as stochastic differential equation in the Stratonovich sense, so that the stokes vector $\widehat{s}$, given in Equation (5.3.10), stays at the surface of the sphere $\mathbb{S}^2 \subset \mathbb{R}^3$. Actually, they implicitly assume that the linear Stochastic Manakov equation (5.4.1) applies. Taking the Fourier transform (with mathematics convention) of equation (5.4.1) leads to the following dynamics in the frequency domain

$$d\widehat{X}(z,\omega) = -i\omega\sqrt{\gamma}\sum_{k=1}^{3}\sigma_k\widehat{X}(z,\omega)\circ dW_k(z). \tag{5.4.2}$$

Note that, at a fixed frequency, the modulus of $\widehat{X}$ is preserved over time i.e. $\widehat{s}_0(z) = \widehat{s}_0(0)$ since the three dimensional Brownian Motion is real valued. Applying the Stratonovich calculus, and by definition of the Stokes vector (5.3.10), we obtain an equation for $\widehat{s}_1(z,\omega)$

$$\begin{aligned} d\widehat{s}_1(z,\omega) &= 2\omega\sqrt{\gamma}\mathcal{R}\mathrm{e}\left(-i\overline{\widehat{X}}_1\widehat{X}_2 + i\overline{\widehat{X}}_2\widehat{X}_1\right)\circ dW_1(z) \\ &\quad -2\omega\sqrt{\gamma}\mathcal{R}\mathrm{e}\left(\overline{\widehat{X}}_1\widehat{X}_2 + \overline{\widehat{X}}_2\widehat{X}_1\right)\circ dW_2(z) \\ &= 2\omega\sqrt{\gamma}\left(\widehat{s}_3(z,\omega)\circ dW_1(z) - \widehat{s}_2(z,\omega)\circ dW_2(z)\right). \end{aligned}$$

Similar evolutions are obtained for the next two components

$$\begin{aligned} d\widehat{s}_2(z,\omega) &= 2\omega\sqrt{\gamma}\left(\widehat{s}_1(z,\omega)\circ dW_2(z) - \widehat{s}_3(z,\omega)\circ dW_3(z)\right) \\ d\widehat{s}_3(z,\omega) &= 2\omega\sqrt{\gamma}\left(\widehat{s}_2(z,\omega)\circ dW_3(z) - \widehat{s}_1(z,\omega)\circ dW_1(z)\right). \end{aligned}$$

Gathering the above results leads to the concise form

$$d\widehat{s}(z,\omega) = 2\omega\sqrt{\gamma}\ \dot{\mathbf{W}}\times\widehat{s}(z,\omega), \tag{5.4.3}$$

where $\dot{\mathbf{W}}= (dW_3, dW_1, dW_2)$ and the operations $\dot{\mathbf{W}}\ \times$ have to be understood in the Stratonovich sense. This equation says that the states of polarization evolve randomly (at fixed frequency) over the Poincaré sphere. Similarly as in Subsection 5.3.2, a dynamics of the PMD vector $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3)$ may be obtained from the viriel (5.3.13) and the square DGD parameter

$$\tau(z) = 4\left(\widehat{R}(z,\omega) - \widehat{R}(0,\omega)\right), \tag{5.4.4}$$

where $\widehat{R}(z,\omega) = \left|\widehat{X}'(z,\omega)\right|^2 / \widehat{s}_0(z,\omega)$. From equation (5.4.1), we easily obtain that the PMD vector is solution of the following SDE [42]

$$d\widehat{R}(z,\omega) = \sqrt{\gamma}\left(\tau_2 \circ dW_1(z) + \tau_3 \circ dW_2(z) + \tau_1 \circ dW_3(z)\right).$$

From Equation (5.4.1) and using the expression of τ_1, τ_2 et τ_3, we obtain an evolution for the PMD vector [42]

$$d\boldsymbol{\tau}(z,\omega) = 2\sqrt{\gamma}\omega\ \dot{\mathbf{W}} \times \boldsymbol{\tau}(z,\omega) + 2\sqrt{\gamma}\ \dot{\mathbf{W}}\ . \tag{5.4.5}$$

In Itô formulation it reads

$$d\boldsymbol{\tau}(z,\omega) = 2\sqrt{\gamma}\omega \begin{pmatrix} 0 & -dW_2(z) & dW_1(z) \\ dW_2(z) & 0 & -dW_3(z) \\ -dW_1(z) & dW_3(z) & 0 \end{pmatrix} \boldsymbol{\tau} + 2\sqrt{\gamma}\ \dot{\mathbf{W}} - 4\gamma\omega^2\boldsymbol{\tau}dz.$$

Applying the Itô formula, we write a dynamics for the square DGD parameter τ

$$\begin{aligned} d\tau(z,\omega) &= 4\sqrt{\gamma}\omega\left(\tau_1\tau_3 \circ dW_1(z) - \tau_1\tau_2 \circ dW_2(z)\right) + 4\sqrt{\gamma}\tau_1 dW_3(z) + 4\gamma dz \\ &\quad +4\sqrt{\gamma}\omega\left(\tau_2\tau_1 \circ dW_2(z) - \tau_2\tau_3 \circ dW_3(z)\right) + 4\sqrt{\gamma}\tau_2 dW_1(z) + 4\gamma dz \\ &\quad +4\sqrt{\gamma}\omega\left(\tau_3\tau_2 \circ dW_3(z) - \tau_3\tau_1 \circ dW_1(z)\right) + 4\sqrt{\gamma}\tau_3 dW_2(z) + 4\gamma dz \\ &= 4d\widehat{R}(z,\omega). \end{aligned}$$

The above calculations show again a direct correspondence between the Fourier transform of X and the PMD vector. It leads to an effective equation that can be easily implemented since, at a fixed frequency, it reduces to a simple SDE. However, in the nonlinear case, such simple dynamics for the PMD vector cannot be obtained and, as in section 5.3.2, the evolution of the DGD will be computed through the stochastic process $\widehat{R}$.

Figure 5.8 displays the evolution of the stokes vector, associated to the linear dynamics (5.4.1), at respectively $z = 0.2,\ 0.8,\ 1.2$ and 2km. The stokes vector now evolves randomly over the Poincaré sphere. Therefore, these changes of polarization are difficult to predict. This motivates the use of Monte Carlo simulations in the next section. Figure 5.9 displays the error curve between the numerical approximation of Equation (5.4.3) and the computation of the Stokes vector (5.3.10), where the solution X of Equation (5.4.1) is computed by the Relaxation scheme. It also displays the error curve between the computation of τ, using an approximation of

(5.4.5), and the computation of the stochastic process $\widehat{R}$, given by (5.4.4). We use an Euler scheme, with an implicit discretization of the Stratonovich integrals, for Equation (5.4.3) and (5.4.5). This numerical exemple shows that the approximation of the stokes vector and of the stochastic process $\widehat{R}$ by the Relaxation scheme is accurate.

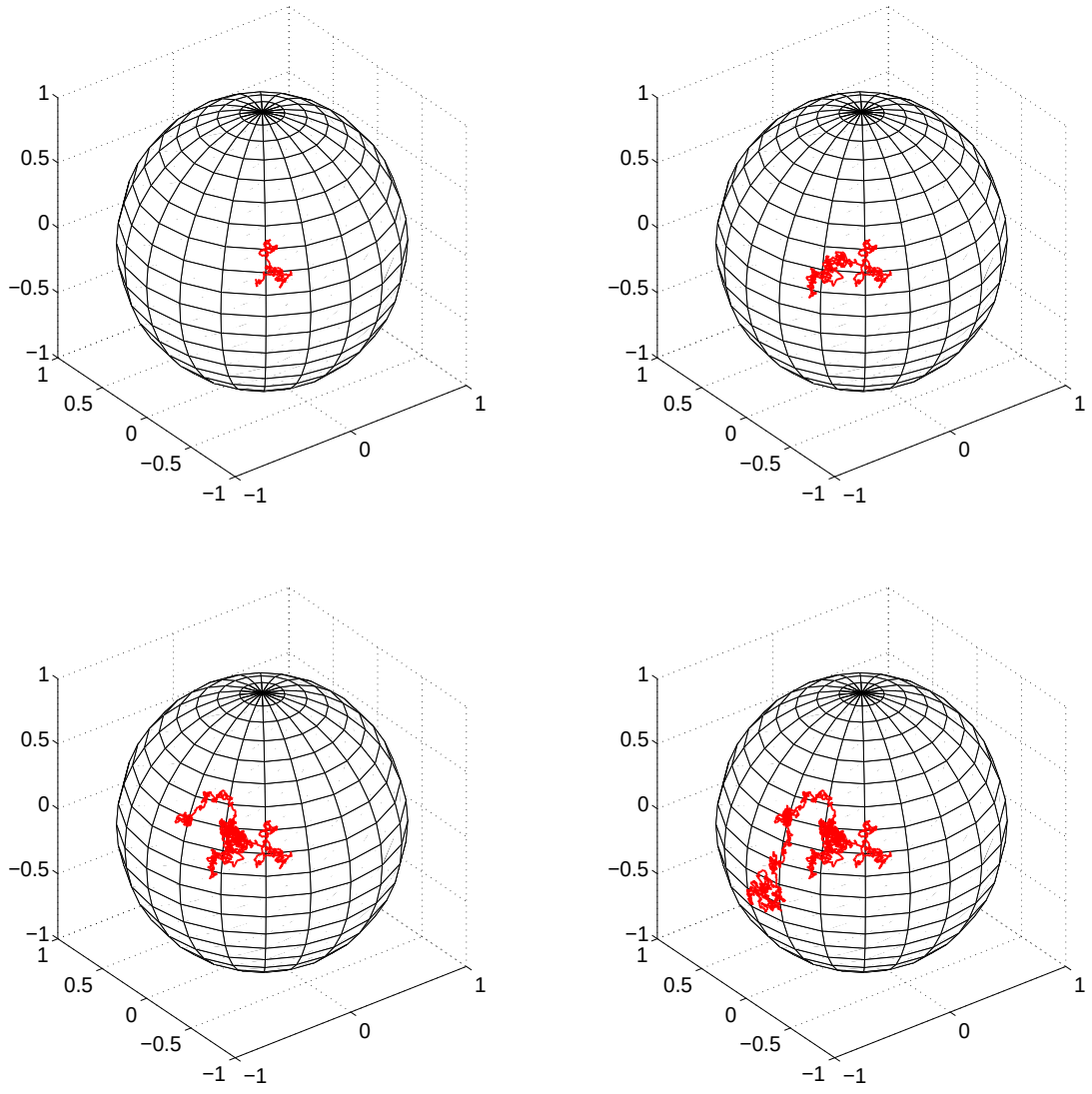


Figure 5.8: Stochastic evolution of the Stokes vector over the Poincaré sphere for linear waves, solution of (5.4.1), at different locations $z = 0.2, 0.8, 1.2, 2$km. We first compute the solution of (5.4.1) using the Relaxation scheme (5.2.9). Then, we compute the corresponding Stokes vector (5.3.10).

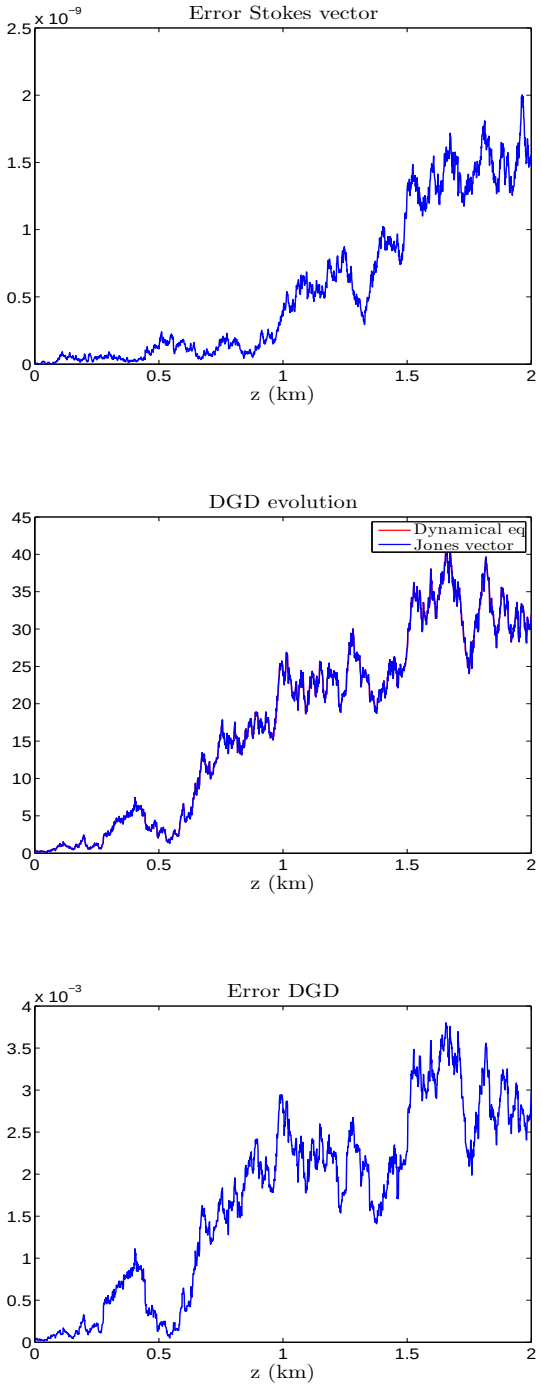


Figure 5.9: Plot of the error between the approximation of the dynamical equation (5.4.3) and the computation of the Stokes vector (5.3.10), where the solution of (5.4.1) is approximated by the Relaxation scheme. Plot of the evolution of the square DGD parameter $|\boldsymbol{\tau}|^2$ using Equation (5.4.5) and the stochastic process $\widehat{R}$, given in (5.4.4). The last plot gives the error curve between these two different approximations.

5.4.2.2 The Statistical Romberg method

Monte-Carlo methods for stochastic processes. The objective, in this paragraph, is to estimate the expectation of some functional f of the electric field envelope X, solution of the stochastic Manakov Equation (5.1.1). The quantities that are considered in this section are given in Table 5.8. Since the solution X of the stochastic Manakov equation (5.1.1) is not known explicitly, we consider a numerical approximation X_j^n of $X(z,t)$ by either one of the three schemes (5.2.8), (5.2.9) and (5.2.10), introduced in section 5.2. The estimation of expectations is performed by Monte Carlo simulations. The name "Monte Carlo simulations " gathers a large number of methods used to estimate expectations and, more generally, integrals. The basic principle of these methods is based on the Strong Law of Large Number and the Central Limit Theorem which gives an asymptotic rate of convergence of order $1/\sqrt{K}$, K being the size of the i.i.d sample. Even if various refinement of this result exists, assuming more integrability on the sequence of random variables, such as the Berry-Esseen and Bikelis theorems [47], the rate of convergence is still of order $1/\sqrt{K}$. From the Central Limit theorem, we easily obtain for K large enough,

$$\mathbb{P}\left(-\frac{1.96\sigma}{\sqrt{K}} \leqslant \mathbb{E}\left(X_1\right) - S_K \leqslant \frac{1.96\sigma}{\sqrt{K}}\right) \simeq 0.95$$

where $(X_k)_{k\geqslant 1}$ is an i.i.d sequence of square integrable random variables, $S_K = \frac{1}{K}\sum_{k=1}^K X_k$ and the variance σ^2 is estimated by the unbiased empirical estimate

$$\sigma_K^2 = \frac{1}{K-1}\sum_{k=1}^K \left(X_k - S_K\right)^2 = \frac{K}{K-1}\left(\frac{1}{K}\sum_{k=1}^K X_k^2 - S_K^2\right).$$

Thus, for a margin of error ε, the size of the random samples needed to obtain a 95% confidence interval is given by $K \geqslant (1.96\sigma/\varepsilon)^2$. The main drawback of the Monte Carlo simulations is this slow rate of convergence, since one needs to increase the size of the simulation by 4 in order to divide the error by 2. The fact that the size sample K depends on the variance σ^2 suggests that the efficiency of the Monte Carlo methods may be increased by reducing the variance of the variable to estimate.

The total error of the method may be decomposed as the sum of a discretization error, arising from the approximation of the exact solution by a discrete scheme, and a statistical error due to Monte Carlo approximation. Indeed,

$$\begin{aligned}
&\mathbb{E}\left(f\left(\widetilde{X}_j^n\right)\right) - \frac{1}{K}\sum_{k=1}^K f\left(X_{j,k}^n\right) \\
&\quad = \mathbb{E}\left(f\left(\widetilde{X}_j^n\right)\right) - \mathbb{E}\left(f\left(X_j^n\right)\right) + \mathbb{E}\left(f\left(X_j^n\right)\right) - \frac{1}{K}\sum_{k=1}^K f\left(X_{j,k}^n\right),
\end{aligned}$$

where $\left(X_{j,k}^n\right)_{1\leqslant k\leqslant K}$ is an i.i.d sample of size K of the random variable X_j^n and $\widetilde{X}_j^n$ is the exact solution evaluated at (z_n, t_j). Then, the total error of Monte Carlo approximations depends on various parameters such as the discretization steps $\Delta z = Z/N$,

$\Delta t = 2T/(M+1)$ and the size K of the i.i.d sample. Regarding the discretization error, the strong order of convergence is clearly not the one that is adapted to study convergence of Monte Carlo simulations, and the weak order of convergence should be analyzed. Usually, this rate of convergence is increased to 1, while the strong order is usually 1/2. For SPDEs, weak order of convergence has just begun to be investigated [23, 27, 59]. In the finite dimensional case, under some regularity hypotheses on the coefficients of the diffusion and of the function f, the authors proved in [106] that the weak error of the explicit Euler scheme has a Taylor expansion in $1/N$. This suggests that for a total error of order 1, the minimal size to run efficient Monte Carlo simulations is N^2. Thus, the complexity of such an algorithm is N^3. In the infinite dimensional case, the time discretization should be taken into account in the total error.

There exists a large range of literature on Monte Carlo simulations and variance reduction methods, which are applied in various situations (see for instance [8, 30, 39, 45, 47, 66, 85, 86]). Among them, the control variates method is a very simple approach that may be applied successfully in many contexts. The basic idea is the following : suppose that one wants to compute the expectation of a random variable U. Instead of computing this quantity by the basic Monte Carlo method, we introduce a new random variable V and denote

$$Q = U - (V - \mathbb{E}(V)).$$

Then, $\mathbb{E}(Q) = \mathbb{E}(U)$ and its variance is given by

$$\mathbb{V}(Q) = \mathbb{V}(U) + \mathbb{V}(V) - 2\mathrm{cov}(U, V).$$

The control variates method consists of finding a random variable V, not almost surely equal to U and positively correlated with U, such that

$$\mathbb{V}(Q) \ll \mathbb{V}(U).$$

In practice, the expectation $\mathbb{E}(V)$ is not known and should be computed at low cost. Moreover, the expectation of the random variable $U - V$ should be computed at the same cost as U. The variable V is called a control variates for U, and $V - \mathbb{E}(V)$ is used as a variance control. Note that the worst choice is to choose V independent of U.

The Statistical Romberg method. The Statistical Romberg method [66] is a variance reduction method based on control variates. In [66], it is applied to diffusion processes. In this paragraph, we introduce a generalization of this method in the context of SPDEs. Let us define a final time $T > 0$ and a final distance $Z > 0$. We denote by X_j^n the numerical approximation of $X(z_n, t_j)$, solution of the stochastic Manakov equation (5.1.1). This approximation is computed on a fine grid i.e. the space step is given by $\Delta z = Z/N_{\text{fine}} > 0$ and the retarded-time step is given by $\Delta t = 2T/(M_{\text{fine}} + 1) > 0$. Moreover, we denote by X_l^m the numerical solution of the same equation but computed on a coarser grid. This grid is such that $N_{\text{coarse}} \ll N_{\text{fine}}$ and $M_{\text{coarse}} \ll M_{\text{fine}}$. Accordingly, the space step and the time step

for the coarse grid are given by $\Delta z = Z/N_{\text{coarse}}$ and $\Delta t = 2T/\left(M_{\text{coarse}}+1\right)$. Finally, we denote by $I_{l,j}^{m,n}$ the linear interpolation operator from the coarse grid to the fine grid. A generalization of the Statistical Romberg method for Equation (5.1.1) would consist of computing the following control variates estimator

$$Q = f\left(X_j^n\right) - I_{l,j}^{m,n} f\left(X_l^m\right) + \mathbb{E}\left(I_{l,j}^{m,n} f\left(X_l^m\right)\right).$$

Thus, an estimate of $\mathbb{E}\left(f\left(X_j^n\right)\right)$ is given by

$$\frac{1}{N_f}\sum_{k=1}^{N_f} f\left(X_{j,k}^n\right) - I_{l,j}^{m,n} f\left(X_{l,k}^m\right) + \frac{1}{N_c}\sum_{k=1}^{N_c} I_{l,j}^{m,n} f\left(\widehat{X}_{l,k}^m\right),$$

where N_c and N_f are the number of sample paths used to perform the Monte Carlo simulations respectively on the coarse grid and the fine grid. The Brownian paths used to compute $\widehat{X}_{l,k}^m$ are different from those used for the computation of $\left(X_{j,k}^n, X_{l,k}^m\right)$. The difficult point in this method is to choose the optimal set of parameters $N_{\text{fine}}, N_{\text{coarse}}, M_{\text{fine}}, M_{\text{coarse}}, N_c, N_f$ in order to minimize the complexity of the algorithm. As in the finite dimensional case [66], the parameter N_{coarse} should not be too large, so that the estimation of $\mathbb{E}\left(f\left(X_l^m\right)\right)$ is not too costly, and not too small in order to have a significant variance reduction. The great advantage of this method is that it does not require knowledge of the solution to the stochastic equation. Moreover, this method is easy to adapt to a large class of problems, when concerned with random variables that are solutions to a stochastic equation (in finite or infinite dimensions). In this work, we want to estimate the average of the following quantities

pulse intensity	$\lvert X(z,t)\rvert^2$
spectral intensity	$\left\lvert \widehat{X}(z,\omega)\right\rvert^2$
barycenter (mass center)	$\langle t\rangle = T_c = \int t\,\lvert X(z,t)\rvert^2\,dt$
viriel (pulse width)	$\langle t^2\rangle = V = \int t^2\,\lvert X(z,t)\rvert^2\,dt$
root mean square	$\sigma_{rms} = \sqrt{\langle t^2\rangle - \langle t\rangle^2}$
degree of polarization	$D_p = \dfrac{\sqrt{S_1^2+S_2^2+S_3^2}}{S_0}$
$\widehat{R}(z,\omega)$	$\dfrac{\lvert\mathcal{F}\left(tX(t,z)\right)\rvert^2}{\left\lvert \widehat{X}(0,\omega)\right\rvert^2} = \dfrac{\left\lvert \widehat{X}'(z,\omega)\right\rvert^2}{\left\lvert \widehat{X}(0,\omega)\right\rvert^2}$
square Differential Group Delay (DGD)	$\tau = \lvert\boldsymbol{\tau}\rvert^2 = 4\left(\widehat{R}(z,\omega) - \widehat{R}(0,\omega)\right)$

Table 5.8: Quantities of interest for the numerical simulations.

Statistical properties of the Polarization Mode Dispersion. To test the efficiency of our method, we focus on the linear equation (5.4.1). In this case, explicit formulas are known on the statistics of the DGD parameter, the mean center and the viriel [42]. These exact formulas will be of help to study the effectiveness of the Monte Carlo simulations. They are given by

$$\begin{cases} \mathbb{E}(\sqrt{\tau(z)}) = \sqrt{\dfrac{32}{\pi}\gamma z} \\ \mathbb{E}(\tau(z)) = 12\gamma z \\ p(\tau) = \dfrac{\sqrt{\tau}}{\sqrt{2\pi}(4\gamma z)^{3/2}} \exp\left(-\dfrac{\tau}{8\gamma z}\right) \mathbb{1}_{\tau \geqslant 0} \\ \mathbb{E}\left(\langle t^2(z)\rangle\right) = V(X_0) + 3\gamma z. \end{cases} \tag{5.4.6}$$

The variances are estimated through [42]

$$\begin{cases} \dfrac{d\mathbb{E}\left(\langle t\rangle^2\right)}{dz} = \gamma \mathbb{E}\left(D_p^2\right) \\ \dfrac{d\mathbb{V}\left(\langle t^2\rangle\right)}{dz} = 12\gamma \mathbb{E}\left(\langle t\rangle\right)(z). \end{cases} \tag{5.4.7}$$

The mean-value of the DGD is called the PMD parameter and grows as the square root of the distance. Recall that the strength γ of the Brownian motion is related to b' and to the correlation length l_c by the relation $\gamma = (b')^2 L_c/6$. Moreover, $b = 2\pi\Delta n/\lambda_0$ and $\lambda_0 = 1.55$ μm is the wavelength of the carrier wave. Usually in fibers, the difference of the refractive index is of order $\Delta n/n \sim 10^{-5} - 10^{-7}$. Let us fix the birefringence strength to $b = 8$ m^{-1}, so that $L_B = 2\pi/b = 0.7854$ m. We now fix $b' = \frac{\partial b}{\partial \omega} = 0.2502$ ps.km^{-1}. Taking $L_c = 100$m as is usual in optical fibers, we get $\gamma = 1.0433e^{-3}$ ps^2.km^{-1}. The computations should be carried over long distances, typically $l = 1000$ km. To speed up the computations, we introduce a small parameter $1 \gg \epsilon > 0$ and the rescaled process $X_\epsilon(z) = X(\epsilon z, \epsilon t)$, whose evolution is given by

$$idX_\epsilon + i\sqrt{\frac{\gamma}{\epsilon}} \sum_{k=1}^{3} \sigma_k \frac{\partial X_\epsilon}{\partial t} \circ d\widetilde{W}_k(z) = 0,$$

where $\widetilde{W}_k(z) = \frac{1}{\sqrt{\epsilon}} W_k(\epsilon z)$ for $k = 1, 2, 3$. Therefore, similar results are obtained considering that $\gamma = 1.0433$ ps^2.km^{-1} over a fiber of 1km length. Note that in this case, the corresponding mean DGD is given by $\mathbb{E}(\sqrt{\tau(Z)}) = 3.2599$ps. We denote by μ the estimator of the expectation given either by the classical Monte Carlo method or by the Statistical Romberg method and τ^n the numerical approximation of $\tau(z_n)$. We introduce the relative errors for the average of the DGD and the average of the square DGD

$$\mathrm{err}_\tau = \max_{n \in [\![0,N]\!]} \left| \frac{\mu\left(\sqrt{\tau^n}\right) - \mathbb{E}\left(\sqrt{\tau}\left(z_n\right)\right)}{\mathbb{E}\left(\sqrt{\tau}\left(z_n\right)\right)} \right| \tag{5.4.8}$$

$$\mathrm{err}_{\tau^2} = \max_{n \in [\![0,N]\!]} \left| \frac{\mu\left(\tau^n\right) - \mathbb{E}\left(\tau\left(z_n\right)\right)}{\mathbb{E}\left(\tau\left(z_n\right)\right)} \right|. \tag{5.4.9}$$

We also introduce two relative errors corresponding to the statistical errors for the time displacement T_c and the pulse width

$$\mathrm{err}_{\langle t \rangle} = \max_{n \in [\![0,N]\!]} \left| \frac{\mu\left(T_c\left(X^n\right)\right) - \mathbb{E}\left(\langle t(z_n) \rangle\right)}{\mathbb{E}\left(\langle t(z_n) \rangle\right)} \right| \tag{5.4.10}$$

$$\mathrm{err}_{\langle t^2 \rangle} = \max_{n \in [\![0,N]\!]} \left| \frac{\mu\left(V\left(X^n\right)\right) - \mathbb{E}\left(\langle t^2(z_n) \rangle\right)}{\mathbb{E}\left(\langle t^2(z_n) \rangle\right)} \right|. \tag{5.4.11}$$

The Monte Carlo simulations have been performed using the parallel computing toolbox in Matlab (parfor loop). Here, we used 4 Matlab workers. Table 5.10a compares the computational times of the Monte Carlo method and the Statistical Romberg method and gives confidence intervals at 95% for important quantities, given in Table 5.8. Table 5.10b displays the relative errors (5.4.8), (5.4.9), (5.4.10) and (5.4.11). We observe that the errors are of the same order for the two methods. The results in these tables show that, with this set of parameters and for the same level of accuracy, the Statistical Romberg is a lot faster than the classical Monte Carlo method. Indeed, the CPU time has been divided by 6. Moreover, the estimations are more accurate since the confidence intervals are smaller than the ones obtained by the Monte Carlo method. Figure 5.11 and 5.12 display the numerical approximations together with the approximated variance of each random variable by the classical Monte Carlo method. Figure 5.13 and 5.14 display the approximation of the same quantities by the Statistical Romberg method. We observe that for all estimates, the variance has been significantly reduced. For the viriel, the variance has been divided by 34. The oscillations appearing in the variance are due to the linear interpolation from the coarse grid to the fine grid.

	Monte Carlo	Statistical Romberg
Soliton parameters	$\Theta = -2\pi/3, \quad \phi_1 = 0 \quad \phi_2 = -\pi/6$ $\alpha_0 = 0, \quad k = 0$ $\eta = 1/2, \quad \tau = 0$	
Discretization parameters	$T = 60, \quad M = 12000$ $Z = 1, \quad N = 200$ $\mathrm{iter} = 3.10^4, \quad \gamma = 1.0433$	$T = 60, \quad M_{\mathrm{fine}} = 12000$ $Z = 1, \quad N_{\mathrm{fine}} = 200$ $N_{\mathrm{coarse}} = 20, \quad M_{\mathrm{coarse}} = 400$ $N_f = 6.10^3, \quad N_c = 3.10^4$ $\gamma = 1.0433$

Table 5.9: Soliton and discretization parameters used to perform the Monte Carlo simulations.

	Monte Carlo	Statistical Romberg
CPU time	57740 seconds	9216 seconds
$\mathbb{E}(\tau(Z))$	[3.2421, 3.2732]	[3.2591, 3.2685]
$\mathbb{E}(\tau^2(Z))$	[12.383, 12.6142]	[12.4841, 12.5494]
$\mathbb{E}(T_c(Z))$	$[-6.4019e^{-03}, 1.4266e^{-02}]$	$[3.7944e^{-03}, 6.3369e^{-04}]$
$\mathbb{E}(V(X(Z)))$	[6.3922, 6.4457]	[6.4136, 6.4209]
$\mathbb{E}(\sigma_{rms}(Z))$	[5.563, 5.6071]	[5.5839, 5.5915]

(a)

	Monte Carlo	Statistical Romberg
err_{τ}	$6.1552e^{-03}$	$4.9401e^{-03}$
err_{τ^2}	$1.0342e^{-02}$	$9.7447e^{-03}$
$\text{err}_{\langle t \rangle}$	$5.3564e^{-03}$	$5.9144e^{-03}$
$\text{err}_{\langle t^2 \rangle}$	$2.9948e^{-03}$	$2.6603e^{-03}$

(b)

Figure 5.10: Comparison between the Monte Carlo method and the Statistical Romberg method. Table 5.10a compares the time complexity of the two methods and gives confidence intervals, at the level of 95%, for the expectation of some quantities given in Table 5.8. Table 5.10b compares the relative errors between the exact value and their approximations by one of the two methods.

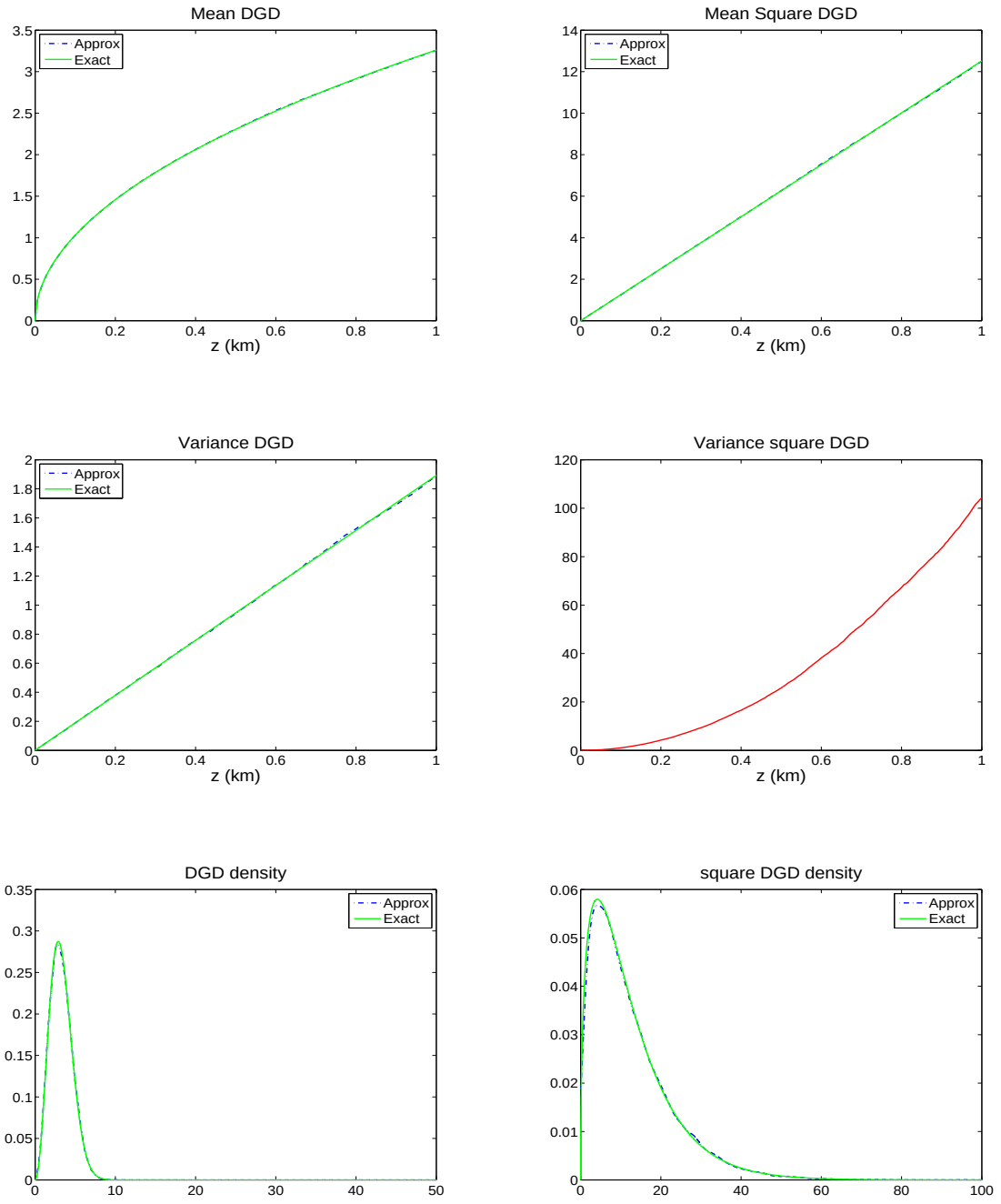


Figure 5.11: Statistics of the DGD and square DGD computed by the Monte-Carlo method. The numerical results are compared with the theoretical formulas (5.4.6) and (5.4.7).

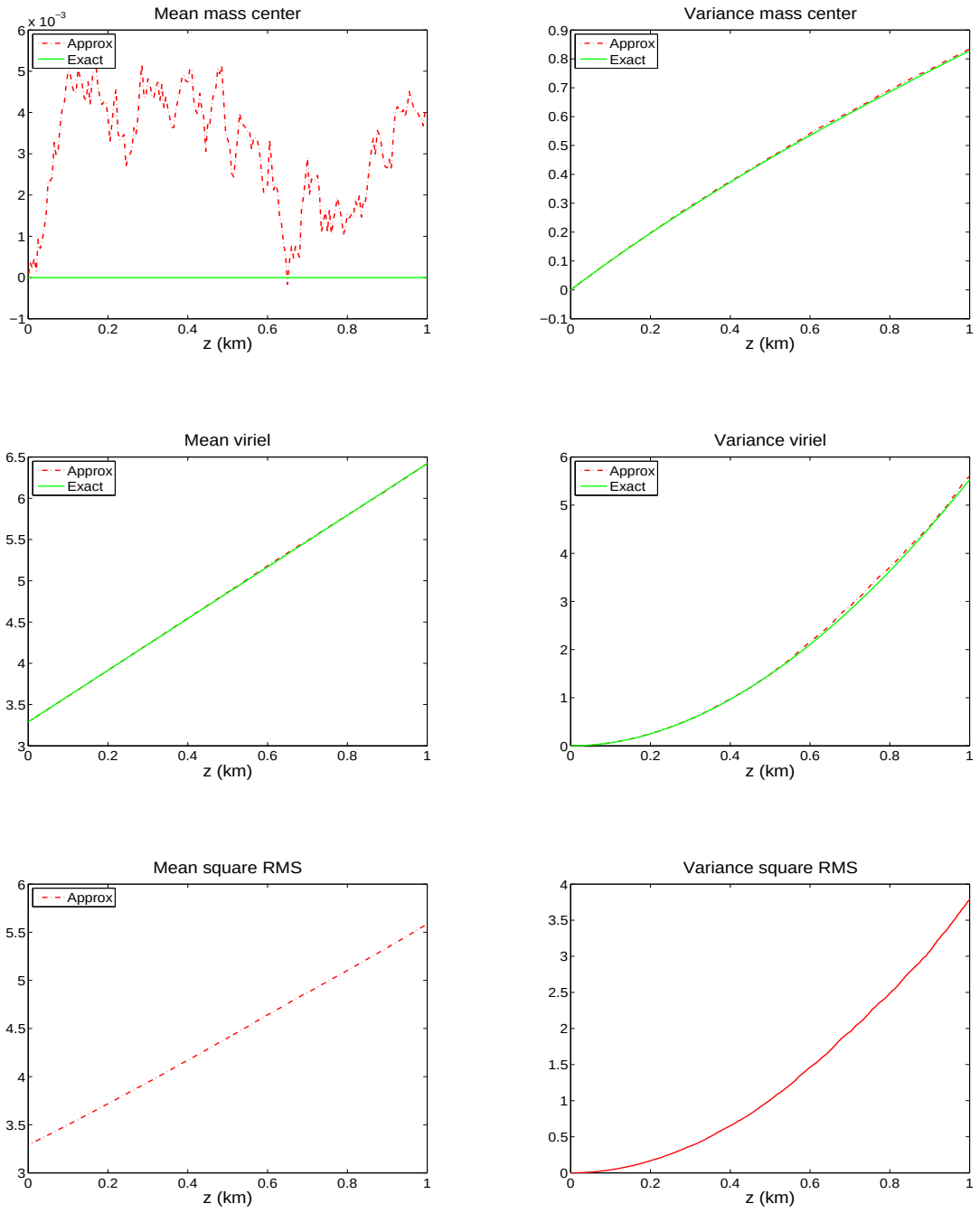


Figure 5.12: Statistics of the mean center, the viriel and the root mean square computed by the Monte-Carlo method. The numerical results are compared with the theoretical formulas (5.4.6) and (5.4.7).

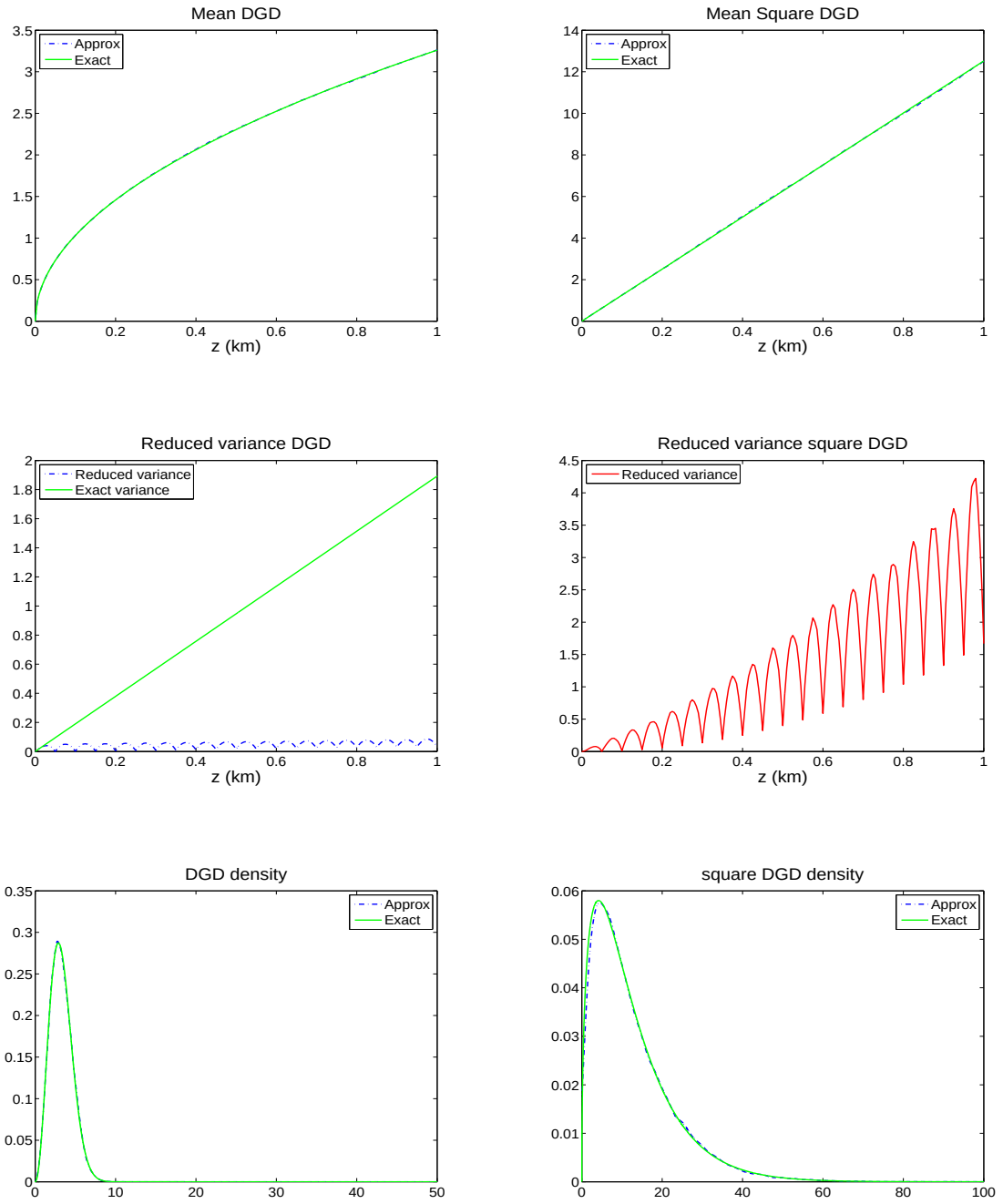


Figure 5.13: Statistics of the DGD and square DGD computed by the Statistical Romberg method. The numerical results are compared with the theoretical formulas (5.4.6).

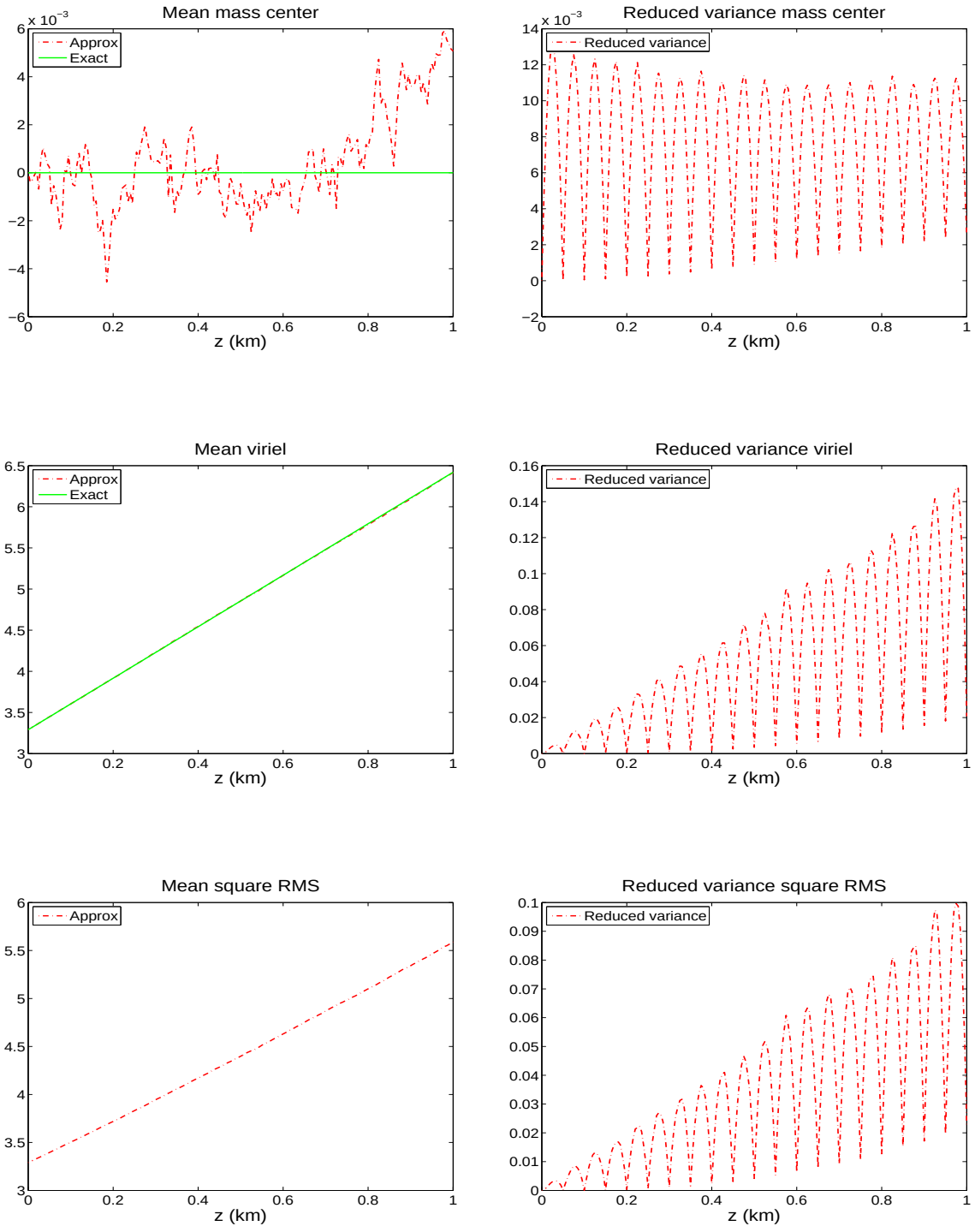


Figure 5.14: Statistics of the mean center, the viriel and the root mean square computed by the Statistical Romberg method. The numerical results are compared with the theoretical formulas (5.4.6).

5.4.3 Soliton propagation and collisions : interactions between nonlinear effects and PMD

In this subsection, we perform a qualitative study on Manakov soliton propagation when subject to random birefringence. This analysis is based on the one done in Section 5.3 and we use the numerical schemes introduced in Section 5.2. On one hand, we investigate the soliton's ability to resist to the perturbations induced by the random changes of the birefringence. Moreover, we want to understand the interactions between the PMD and the Kerr effect and we compare the evolution of nonlinear pulses with the dynamics of linear pulses. On the other hand, we study the collision of two Manakov solitons. Here, we want to understand the effect of the random changes of the polarization state in the collision process. Finally, we study the evolution of the intensity of pulse average and we perform a numerical analysis of the PMD statistics in nonlinear optical fibers.

Soliton propagation. We first explore the evolution of the Manakov soliton in presence of random birefringence. The numerical tests are performed considering two level of noises. Figure 5.15 displays the intensity of linear and nonlinear profiles at three different locations in the fiber i.e. $z = 6, 18$ and 36km. The linear pulse is solution of Equation (5.4.1), while the nonlinear one is solution of Equation (5.1.1). At the beginning of the fiber ($z = 6$km), the chromatic dispersion and the nonlinear effects have not come into play and the shape of the linear and nonlinear pulses are similar. If the soliton is not strongly altered by the noise, the polarization has already begun to change. The energy exchange is relatively low and we do not observe distortions of the two pulses. However, over longer distances, small perturbations arise at the basis of the soliton whose amplitude is small compared to the amplitude of the soliton. These perturbations do not exist for linear pulses whose support stays localized. Those radiations result from the energy exchange implied by the stochastic variations and the nonlinearity. This phenomenon is called Polarization Mode Coupling. The numerical simulations show that the devastating effect of the PMD is due to its accumulation along the fiber and not to its local value. Fig. 5.16 displays the evolution of the energy, the Dop, the mass center and the Viriel (see Table 5.8). It also displays the evolution of the stokes vector for linear and nonlinear pulses. We observe that the normalized stokes vector (5.3.8), associated to nonlinear pulses (red curve), moves quickly over the Poincaré sphere in comparison to the one for linear pulses (blue curve). This rapid rotation on the Poincaré sphere is induced by the interaction of the random birefringence and the nonlinear terms. Indeed, in absence of Kerr effects, the stokes vector stays localised over the Poincaré sphere. Moreover, in absence of noise, the polarization of the electric field stays constant and the normalized stokes vector reduces to a point over the sphere. Therefore, the phase difference induced by the birefringence and the phase modulation, induced by the Kerr effects, play along and lead to rapid rotations of the stokes vector over the Poincaré sphere. If we consider a stronger level of noise (Fig. 5.17), we observe the linear pulse is severely distorted while the shape of the nonlinear pulse is better preserved. The nonlinearity prevents the pulse from high distortions but the amplitude of the radiating waves are bigger.

Collision of two-solitons. In this paragraph, we study the interaction between two Manakov solitons and we consider the special case where the two polarization vectors are orthogonal. From the analysis of section 5.3, we know that, in the deterministic situation, the two solitons emerge unchanged from the collision and the mass of each component is preserved. Figure 5.18 displays the collision of two solitons in presence of noise and we observe that the perturbations induced by the random birefringence destroy the properties of the soliton. When the two solitons emerge from the collision, the mass of each component is no longer preserved, even if the $\mathbb{L}^2$ norm is still conserved. Indeed, $\text{err}^N_{\mathbb{L}^2} = 2.3685e^{-14}$ and the relative errors for the conservation of the mass of each component are $\text{err}^N_{1,\mathbb{L}^2} = 0.3362$ and $\text{err}^N_{2,\mathbb{L}^2} = 0.2925$. This is due to the creation of radiating waves leading to an energy exchange between the two components. Moreover, the shape of the solitons is not preserved due to the variation of the polarization vectors induced by the birefringence.

Pulse propagation in average and PMD estimations in nonlinear medium. Now, we perform Monte Carlo simulations using the Statistical Romberg method. The solution of Equation (5.1.1) is approximated by the Relaxation scheme. Figure 5.19 displays the evolution of the intensity of the average of X along the fiber. We observe that the small radiating waves, appearing at the basis of the solitons for one sample paths, disappear. Moreover, the shape of the soliton is well preserved and in average the polarization varies quickly. We compute the average of the following quantities : the DGD, the square DGD, the mean center and the viriel. Fig. 5.20 and 5.21 display their evolutions that we compare to the evolution of the same quantities for linear pulses. As expected, the linear model (5.4.1) is valid only for the first 5km and the growth of these quantities is no more linear. Actually, on this length, the nonlinear interactions have not yet begun. Moreover, we observe in Figure 5.20, that the tail of the probability density functions of the DGD and the square DGD is heavier than in the linear case. Thus, the configurations leading to large DGD are more likely to happen than in the linear case. An interesting question, that may be addressed, is to estimate the probability of these configurations, as it is done in the linear case in absence of GVD and for nonlinear schrödinger equations with additive noise (see [7, 8, 30, 85]).

	Fig. 5.15 & Fig. 5.16	Fig. 5.17
Soliton parameters	$\phi_1 = 0, \quad \phi_2 = \pi/3, \quad \Theta = \pi/2, \quad k = 0$ $\eta = 1, \quad \alpha_0 = \pi, \quad \tau = 0$	
Strength of the noise	$\gamma = 0.0261$	$\gamma = 0.1044$
Discretization parameters	$T = 80, \quad M = 32000, \quad Z = 40, \quad N = 8000$	

Table 5.10: Set of parameters used to simulate the propagation of solitons in optical fibers with random birefringence.

	Collision of two solitons (Fig. 5.18)
Soliton parameters	$\phi_{11} = \phi_{12} = -\pi, \quad \phi_{21} = \phi_{22} = \pi$ $\Theta_1 = 0.7, \quad \Theta_2 = \pi + 0.7, \quad \eta_1 = 1.2, \quad \eta_2 = 1$ $\alpha_{10} = \alpha_{20} = 0, \quad k = 0.35, \quad \tau_{10} = 10, \quad \tau_{20} = -8$
Discretization parameters	$T = 100, \quad M = 20000, \quad Z = 50, \quad N = 5000$

Table 5.11: Set of parameters used to simulate the collision of two solitons in optical fibers with random birefringence .

	Statistical Romberg (Fig. 5.20 & Fig. 5.21)
Soliton parameters	$\Theta = \pi/2.5, \quad \phi_1 = 0, \quad \phi_2 = \pi/3$ $\alpha_0 = \pi, \quad k = 0, \quad \eta = 1/2, \quad \tau = 0$
Discretization parameters	$T = 60, \quad M_{\text{fine}} = 12000, \quad Z = 40$ $N_{\text{fine}} = 2500, N_{\text{coarse}} = 250, \quad M_{\text{coarse}} = 400$ $N_f = 2.10^3, \quad N_c = 2.10^4, \quad \gamma = 2.6082e^{-2}$

Table 5.12: Soliton and discretization parameters used to perform Monte Carlo simulations by the Statistical Romberg method.

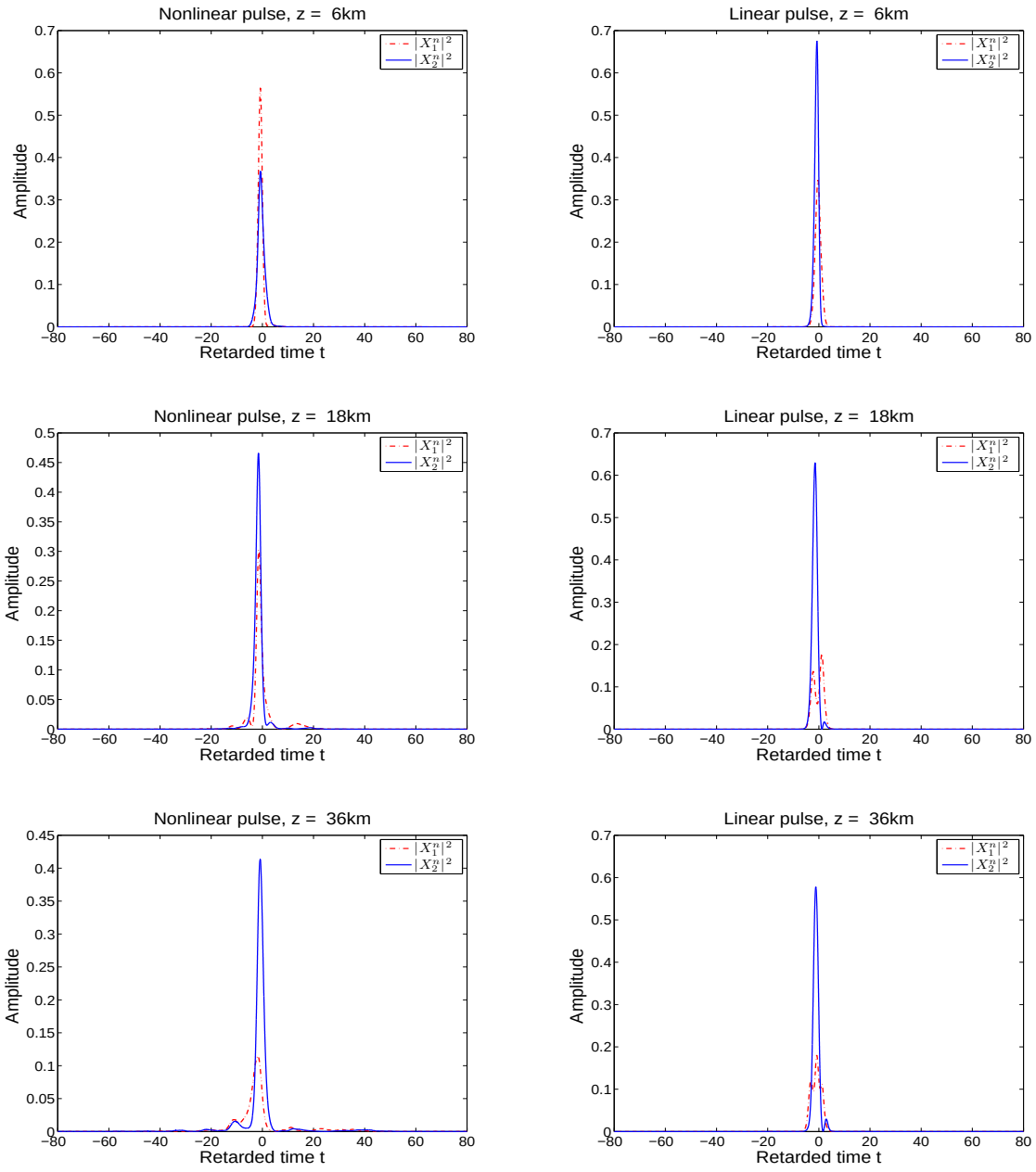


Figure 5.15: Evolution along the fiber of both nonlinear and linear pulses respectively solutions of Equation (5.1.1) and (5.4.1). The input profile is a Manakov soliton.

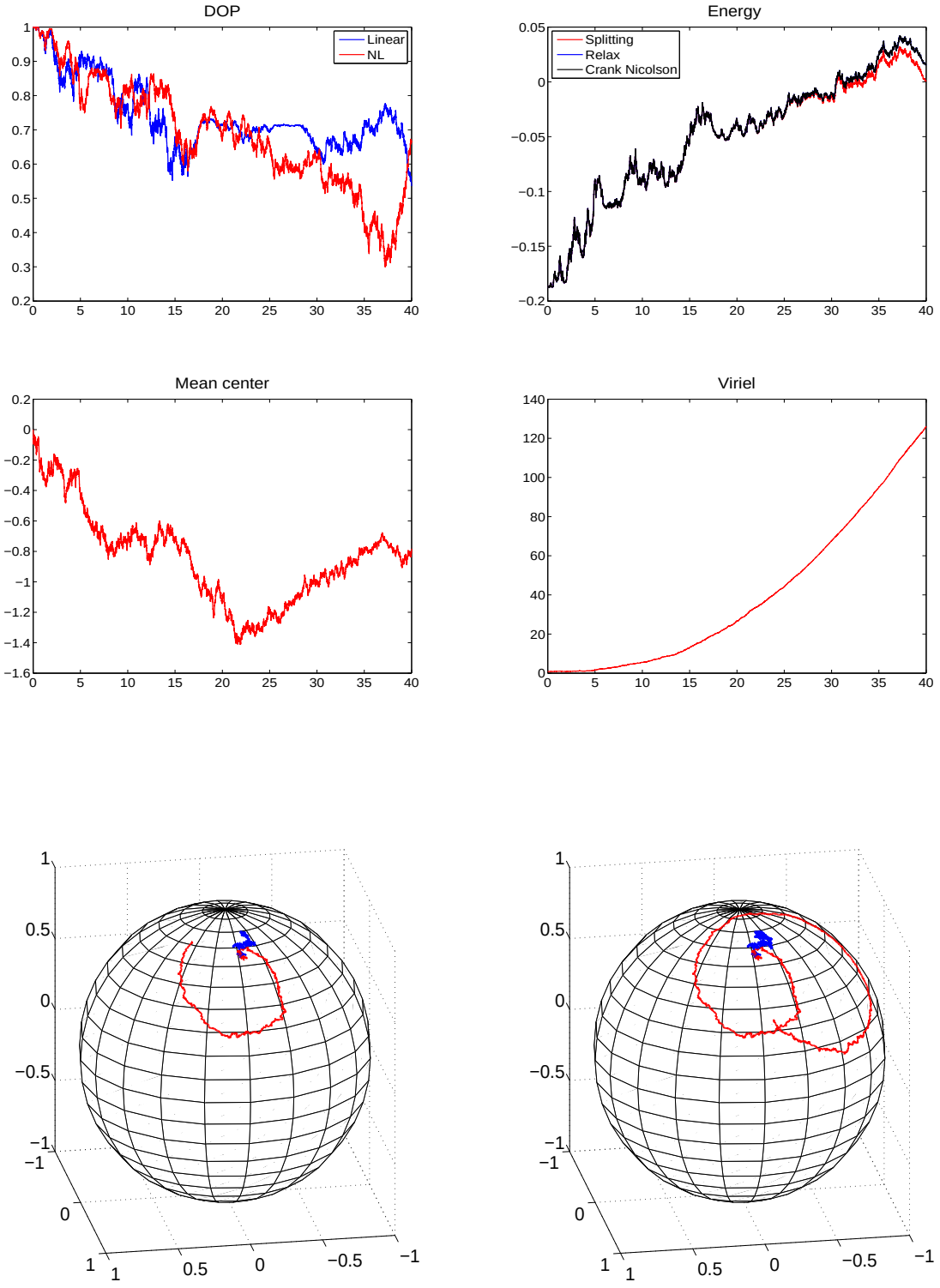


Figure 5.16: Evolution of the DoP, the energy, the mass center and the viriel for one sample paths. The last row displays the evolution of the stokes vector for nonlinear pulses (red curve) and the stokes vector for linear pulses (blue curve).

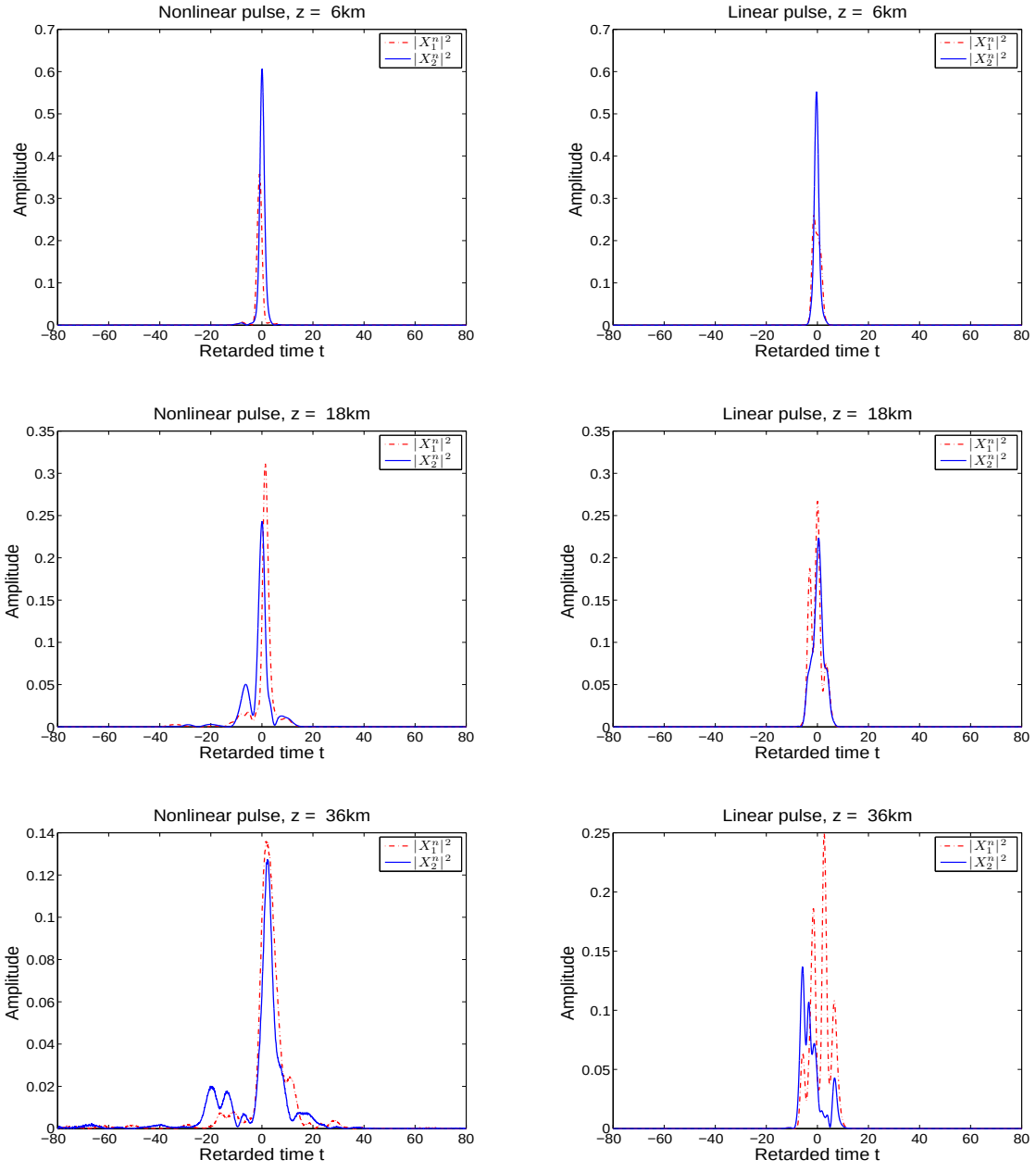


Figure 5.17: Evolution along the fiber of both nonlinear and linear pulses respectively solutions of Equation (5.1.1) and (5.4.1). The input profile is a Manakov soliton and the effective PMD parameter is $\gamma = 0.1044$.

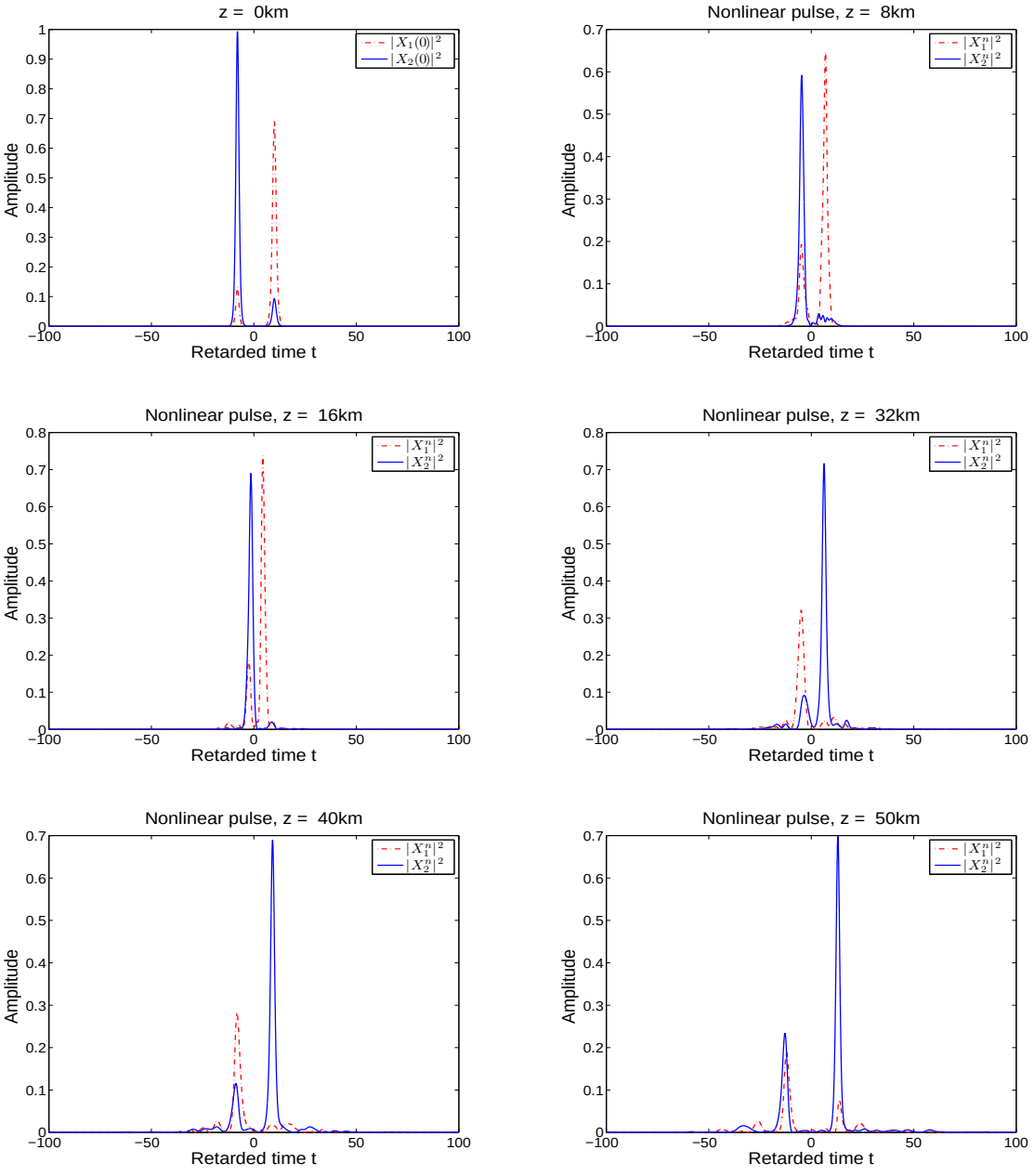


Figure 5.18: Evolution along the fiber of a 2-soliton. The polarization vectors are chosen orthogonal so that the changes of polarization are only induced by the random birefringence.

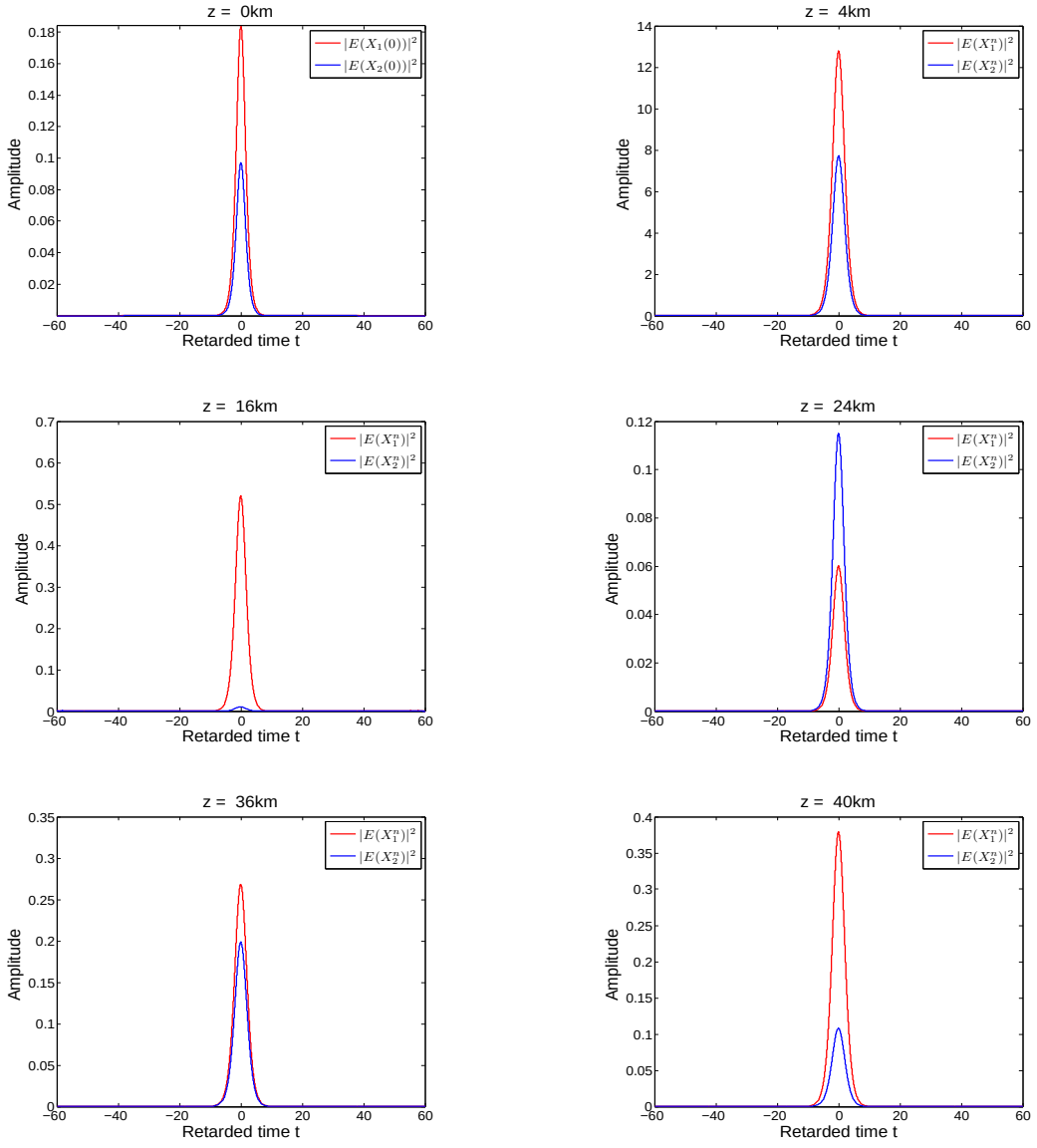


Figure 5.19: Evolution of the intensity of pulse average for Manakov soliton.

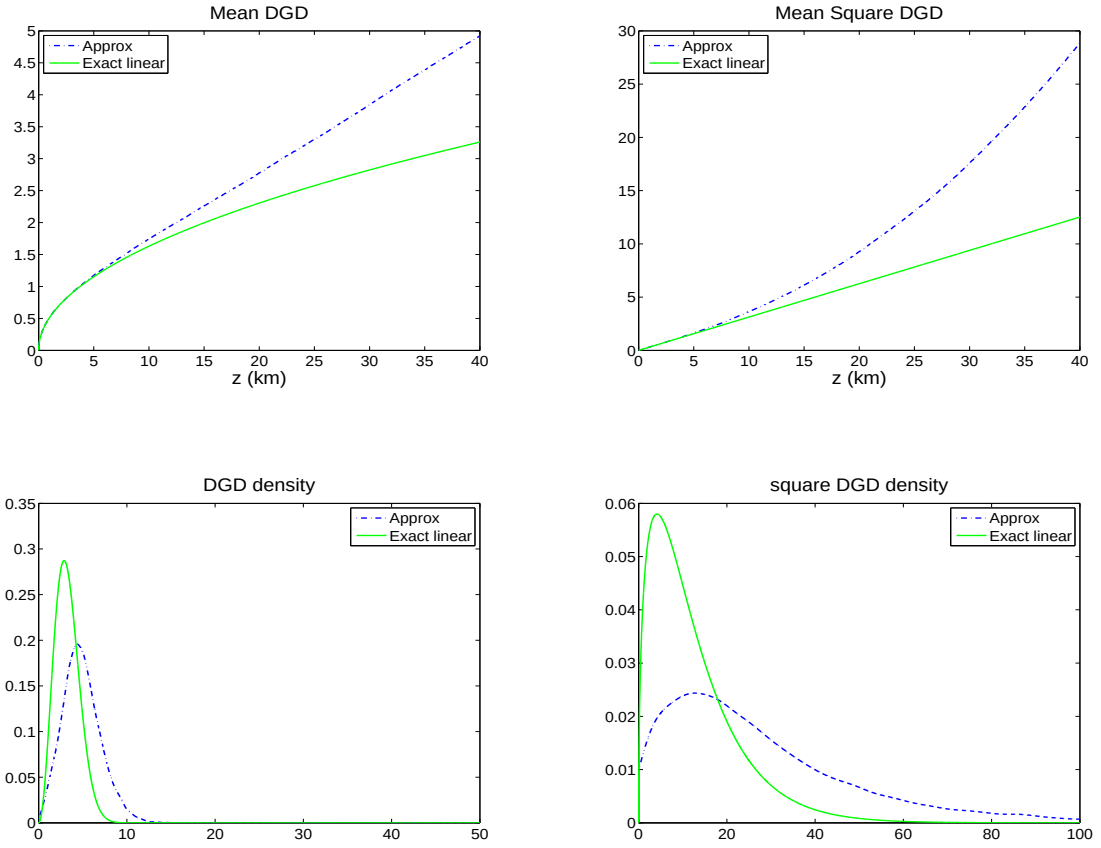


Figure 5.20: Average curves of the DGD and the square DGD for nonlinear pulses. The last row displays the empirical PDF of the DGD and the square DGD. The green curves correspond to the exact formulas in the linear case in absence of GVD (5.4.6). These quantities are computed using the Relaxation scheme and the Statistical Romberg method on a fiber of 40km length.

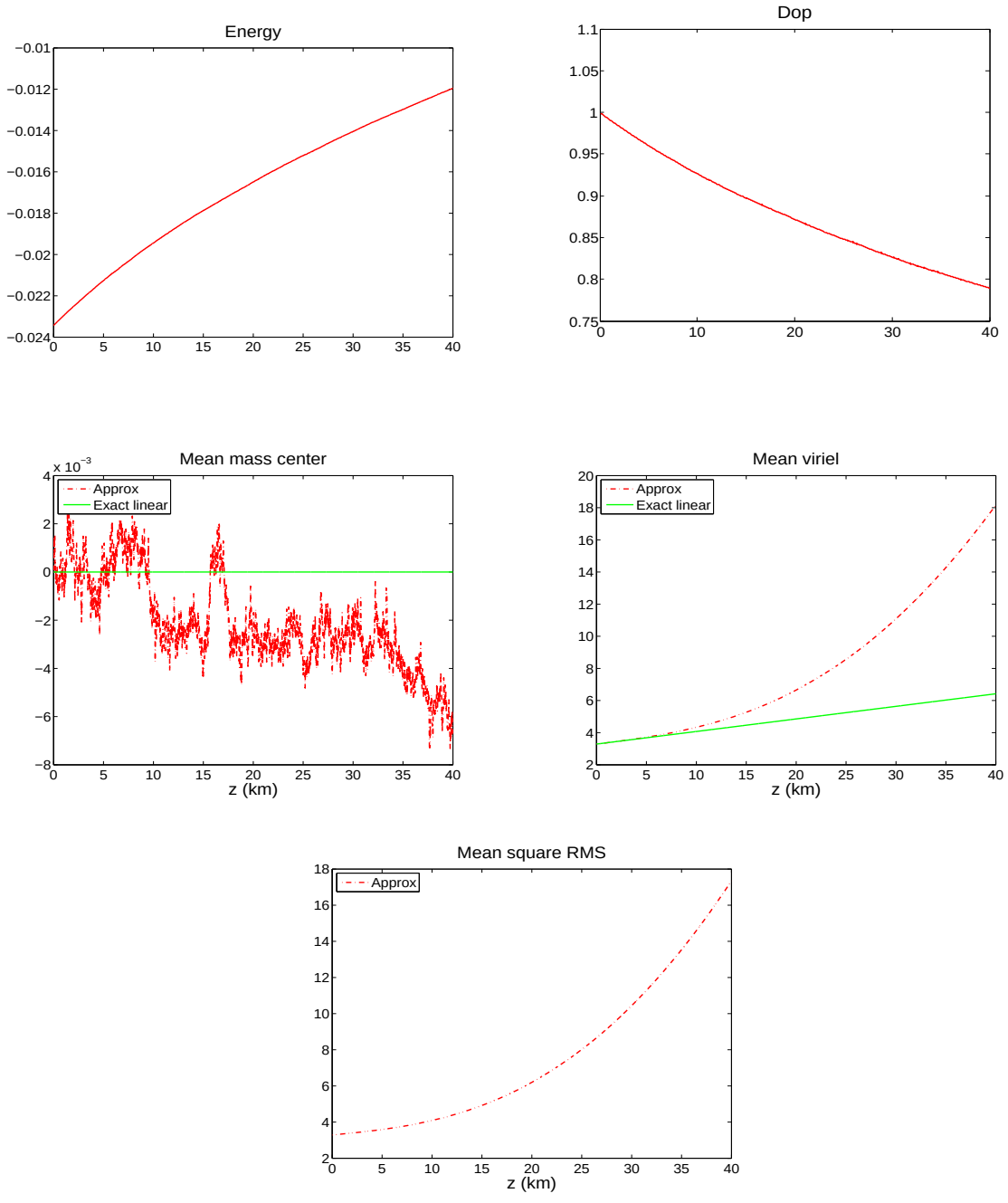


Figure 5.21: Average curves of the Energy, the Degree of Polarization, the mass center, the viriel and the root mean square in nonlinear optical fibers (see Table 5.8). These quantities are computed using the Relaxation scheme and the Statistical Romberg method on a fiber of 40km length.

Bibliography

[1] G. P. Agrawal. *Applications of Nonlinear Fiber Optics.* Academic Press, 2001.

[2] G. P. Agrawal. *Nonlinear Fiber Optics, third edition.* Academic Press, 2001.

[3] R. Azencott. Grandes déviations et applications. In *Eighth Saint Flour Probability Summer School—1978 (Saint Flour, 1978)*, volume 774 of *Lecture Notes in Math.*, pages 1–176. Springer, Berlin, 1980.

[4] C. Besse. A relaxation scheme for the nonlinear Schrödinger equation. *SIAM J. Numer. Anal.*, 42(3):934–952, 2004.

[5] C. Besse, B. Bidégaray, and S. Descombes. Order estimates in time of splitting methods for the nonlinear Schrödinger equation. *SIAM J. Numer. Anal.*, 40(1):26–40 (electronic), 2002.

[6] P. Billingsley. *Convergence of Probability measures.* Wiley, 1968.

[7] G. Biondini, W.L. Kath, and C. R. Menyuk. Importance sampling for polarization-mode dispersion. *Photonics Technology Letters, IEEE*, 14(3):310–312, 2002.

[8] G. Biondini, W.L. Kath, and C. R. Menyuk. Importance sampling for polarization-mode dispersion: techniques and applications. *Lightwave Technology, Journal of*, 22(4):1201 – 1215, 2004.

[9] G. Blankenship and G. C. Papanicolaou. Stability and control of stochastic systems with wide-band noise disturbances. I. *SIAM J. Appl. Math.*, 34(3):437–476, 1978.

[10] Z. Brzeźniak. Stochastic partial differential equations in M-type 2Banach spaces. *Potential Anal.*, 4(1):1–45, 1995.

[11] J. Bures. *Optique guidée, Fibres optiques et composants passifs tout-fibre.* Presses Internationales Polytechnique, 2009.

[12] K. Burrage, P. M. Burrage, and T. Tian. Numerical methods for strong solutions of stochastic differential equations: an overview. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 460(2041):373–402, 2004. Stochastic analysis with applications to mathematical finance.

[13] T. Cazenave. *Semilinear Schrödinger equations*, volume 10 of *Courant Lecture Notes in Mathematics*. New York University Courant Institute of Mathematical Sciences, New York, 2003.

[14] B. Clouet. *Etude de la dispersion modale de polarisation dans les systèmes régénérés optiquement*. Thèse de Doctorat, Université de Rennes I, 2007.

[15] T. Colin and P. Fabrie. Semidiscretization in time for nonlinear Schrödinger-waves equations. *Discrete Contin. Dynam. Systems*, 4(4):671–690, 1998.

[16] A. M. Davie and J. G. Gaines. Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations. *Math. Comp.*, 70(233):121–134, 2001.

[17] A. de Bouard and A. Debussche. A stochastic nonlinear Schrödinger equation with multiplicative noise. *Comm. Math. Phys.*, 205(1):161–181, 1999.

[18] A. de Bouard and A. Debussche. The stochastic nonlinear Schrödinger equation in H^1. *Stochastic Anal. Appl.*, 21(1):97–126, 2003.

[19] A. De Bouard and A. Debussche. A semi-discrete scheme for the stochastic nonlinear Schrödinger equation. *Numer. Math.*, 96(4):733–770, 2004.

[20] A. de Bouard and A. Debussche. Weak and strong order of convergence of a semidiscrete scheme for the stochastic nonlinear Schrödinger equation. *Appl. Math. Optim.*, 54(3):369–399, 2006.

[21] A. de Bouard and A. Debussche. The nonlinear Schrödinger equation with white noise dispersion. *J. Funct. Anal.*, 259(5):1300–1321, 2010.

[22] A. de Bouard and M. Gazeau. A diffusion approximation theorem for a nonlinear PDE with application to random birefringent optical fibers. *Ann. Appl. Probab.*, 22(6):2460–2504, 2012.

[23] A. Debussche. Weak approximation of stochastic partial differential equations: the nonlinear case. *Math. Comp.*, 80(273):89–117, 2011.

[24] A. Debussche and E. Faou. Modified energy for split-step methods applied to the linear Schrodinger equation. *SIAM J. Numer. Anal.*, 47(5):3705–3719, 2009.

[25] A. Debussche and J. Printems. Numerical simulation of the stochastic Korteweg-de Vries equation. *Phys. D*, 134(2):200–226, 1999.

[26] A. Debussche and J. Printems. Convergence of a semi-discrete scheme for the stochastic Korteweg-de Vries equation. *Discrete Contin. Dyn. Syst. Ser. B*, 6(4):761–781 (electronic), 2006.

[27] A. Debussche and J. Printems. Weak order for the discretization of the stochastic heat equation. *Math. Comp.*, 78(266):845–863, 2009.

[28] A. Debussche and Y. Tsutsumi. 1D quintic nonlinear Schrödinger equation with white noise dispersion. *J. Math. Pures Appl. (9)*, 96(4):363–376, 2011.

[29] A. Debussche and J. Vovelle. Diffusion limit for a stochastic kinetic problem. *Preprint*, 2011.

[30] P. Del Moral and J. Garnier. Genealogical particle analysis of rare events. *Ann. Appl. Probab.*, 15(4):2496–2534, 2005.

[31] M. Delfour, M. Fortin, and G. Payre. Finite-difference solutions of a nonlinear Schrödinger equation. *J. Comput. Phys.*, 44(2):277–288, 1981.

[32] H. Doss. Liens entre équations différentielles stochastiques et ordinaires. *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 13(2):99–125, 1977.

[33] A. Durán and J. M. Sanz-Serna. The numerical integration of relative equilibrium solutions. The nonlinear Schrödinger equation. *IMA J. Numer. Anal.*, 20(2):235–261, 2000.

[34] S. N. Ethier and T. G. Kurtz. *Markov processes, characterization and convergence*. Wiley, New York, 1986.

[35] O. Faure. *Simulation du Mouvement Brownien et des Diffusions.* Thèse de Doctorat, Ecole des Ponts et Chaussées, 1992.

[36] F. Flandoli and D. Gątarek. Martingale and stationary solutions for stochastic Navier-Stokes equations. *Probab. Theory Related Fields*, 102(3):367–391, 1995.

[37] G. J. Foschini and C. D. Poole. Statistical theory of polarization dispersion in single mode fibers. *Journal of Lightwave Technology*, 9:1439–1456, 1991.

[38] J. P. Fouque, J. Garnier, G. Papanicolaou, and K. Sølna. *Wave propagation and time reversal in randomly layered media*, volume 56 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2007.

[39] A. Galtarossa and C. R. Menyuk. *Polarization mode dispersion*. Optical and fiber communications reports (Berlin. Print). Springer, 2005.

[40] J. Garnier. Multi-scaled diffusion-approximation. Applications to wave propagation in random media. *ESAIM Probab. Statist.*, 1:183–206 (electronic), 1995/97.

[41] J. Garnier. Scattering, spreading, and localization of an acoustic pulse by a random medium. In *Three courses on partial differential equations*, volume 4 of *IRMA Lect. Math. Theor. Phys.*, pages 71–123. de Gruyter, Berlin, 2003.

[42] J. Garnier, J. Fatome, and G. Le Meur. Statistical analysis of pulse propagation driven by polarization-mode dispersion. *J. Opt. Soc. Am. B*, 19(9):1968–1977, Sep 2002.

[43] J. Garnier and R. Marty. Effective pulse dynamics in optical fibers with polarization mode dispersion. *Wave Motion*, 43(7):544–560, 2006.

[44] J. Ginibre. *Introduction aux équations de Schrödinger non linéaires*. Cours de DEA, 1994-1995.

[45] P. Glasserman. *Monte Carlo Methods in Financial Engineering (Stochastic Modelling and Applied Probability) (v. 53)*. Springer, 1 edition, 2003.

[46] J. P. Gordon and H. Kogelnik. Pmd fundamentals : Polarization mode dispersion in optical fibers. *PNAS 97*, 22(9):4541–4550, april 2000.

[47] C. Graham and D. Talay. *Simulation stochastique et méthodes de Monte carlo*. cours de l'Ecole Polytechnique, Départements de mathématiques, 2009.

[48] I. Gyöngy. Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. I. *Potential Anal.*, 9(1):1–25, 1998.

[49] I. Gyöngy and A. Millet. On discretization schemes for stochastic evolution equations. *Potential Anal.*, 23(2):99–134, 2005.

[50] I. Gyöngy and A. Millet. Rate of convergence of implicit approximations for stochastic evolution equations. In *Stochastic differential equations: theory and applications*, volume 2 of *Interdiscip. Math. Sci.*, pages 281–310. World Sci. Publ., Hackensack, NJ, 2007.

[51] I. Gyöngy and A. Millet. Rate of convergence of space time approximations for stochastic evolution equations. *Potential Anal.*, 30(1):29–64, 2009.

[52] I. Gyöngy and D. Nualart. Implicit scheme for quasi-linear parabolic partial differential equations perturbed by space-time white noise. *Stochastic Process. Appl.*, 58(1):57–72, 1995.

[53] J. N. Damask H. K. V. Lotsch. *Polarization Optics in Telecommunications*. Springer-Verlag, 2005.

[54] A. Hasegawa. Effect of polarization mode dispersion in optical soliton transmission in fibers. *Physica D: Nonlinear Phenomena*, 188(3-4):241–246, 2004.

[55] A. Hasegawa and M. Matsumoto. *Optical solitons in Fibers, third edition*. Springer, 2002.

[56] A. Hasegawa and F. Tappert. Transmission of stationary nonlinear optical pulses in dispersive dielectric fibers. ii. normal dispersion. *Applied Physics Letters*, 23:171–172, 1973.

[57] R. Z. Has′minskiĭ. A limit theorem for solutions of differential equations with a random right hand part. *Teor. Verojatnost. i Primenen*, 11:444–462, 1966.

[58] E. Hausenblas. Approximation for semilinear stochastic evolution equations. *Potential Anal.*, 18(2):141–186, 2003.

[59] E. Hausenblas. Weak approximation of the stochastic wave equation. *J. Comput. Appl. Math.*, 235(1):33–58, 2010.

[60] W. Huang. *Fiber birefringence modeling for polarization mode dispersion*. PhD thesis, University of Waterloo, 2007.

[61] S. Huard. *Polarisation de la lumière*. Dunod, 1994.

[62] L. I. Ignat. Fully discrete schemes for the Schrödinger equation. Dispersive properties. *Math. Models Methods Appl. Sci.*, 17(4):567–591, 2007.

[63] L. I. Ignat and E. Zuazua. Numerical dispersive schemes for the nonlinear Schrödinger equation. *SIAM J. Numer. Anal.*, 47(2):1366–1390, 2009.

[64] T. Jahnke and C. Lubich. Error bounds for exponential operator splittings. *BIT*, 40(4):735–744, 2000.

[65] T. Kato. On nonlinear Schrodinger equations. *Ann. Inst. H. Poincaré Phys. Théor.*, 46(1):113–129, 1987.

[66] A. Kebaier. Statistical Romberg extrapolation: a new variance reduction method and applications to option pricing. *Ann. Appl. Probab.*, 15(4):2681–2705, 2005.

[67] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.

[68] Y. Kodama. Optical solitons in a monomode fiber. *Journal of Statistical Physics*, 39:597–614, 1985. 10.1007/BF01008354.

[69] H. Kunita. *Stochastic flows and stochastic differential equations*. Cambridge University press, Cambridge, 1990.

[70] H. J. Kushner. *Approximation and weak convergence methods for random processes*. MIT press, Cambridge, 1984.

[71] D. Marcuse. Coupled-mode theory for anisotropic optical waveguides. *AT T Technical Journal*, 54:985–995, 1975.

[72] D. Marcuse, P. K. A. Wai, and C. R. Menyuk. Application of the Manakov-pmd equation to studies of signal propagation in optical fibers with randomly varying birefringence. *Journal of Lightwave Technology*, 15(9):1735–1746, 1997.

[73] R. Marty. *Problèmes d'évolution en milieux aléatoires : Théorèmes limites, Schémas numériques et applications en optique*. Thèse de Doctorat, Université Paul Sabatier, Toulouse III, 2005.

[74] R. Marty. On a splitting scheme for the nonlinear Schrödinger equation in a random medium. *Commun. Math. Sci.*, 4(4):679–705, 2006.

[75] C. R. Menyuk. Nonlinear pulse propagation in birefringence optical fibers. *IEEE J. Quantum Electronics*, 23:174–176, 1987.

[76] C. R. Menyuk. Pulse propagation in an elliptically birefringent kerr medium. *IEEE journal of quantum electronics*, 25(12):2674–2682, 1989.

[77] C. R. Menyuk. Application of multiple-length-scale methods to the study of optical fiber transmission. *Journal of Engineering Mathematics*, 36:113–136, 1999.

[78] C. R. Menyuk and B. S. Marks. Interaction of polarization mode dispersion and nonlinearity in optical fiber transmission systems. *Journal of Lightwave Technology*, 24:2806–2826, July 2006.

[79] A. Millet and P-L. Morien. On implicit and explicit discretization schemes for parabolic SPDEs in any dimension. *Stochastic Process. Appl.*, 115(7):1073–1106, 2005.

[80] G. N. Milstein. A method with second order accuracy for the integration of stochastic differential equations. *Teor. Verojatnost. i Primenen.*, 23(2):414–419, 1978.

[81] G. N. Milstein. Weak approximation of solutions of systems of stochastic differential equations. *Teor. Veroyatnost. i Primenen.*, 30(4):706–721, 1985.

[82] G. N. Milstein and M. V. Tretyakov. *Stochastic numerics for mathematical physics*. Scientific Computation. Springer-Verlag, Berlin, 2004.

[83] G.N. Milstein, Y.P. Repin, and M.V Tretyakov. Mean-square symplectic methods for hamiltonian systems with multiplicative noise.

[84] L. F. Mollenauer, R. H. Stolen, and J. P. Gordon. Experimental observation of picosecond pulse narrowing and solitons in optical fibers. *Phys. Rev. Lett.*, 45:1095–1098, 1980.

[85] R. O. Moore, G. Biondini, and W. L. Kath. A method to compute statistics of large, noise-induced perturbations of nonlinear schrodinger solitons. *SIAM Rev.*, 50(3):523–549, 2008.

[86] G. Pagès. Multi-step Richardson-Romberg extrapolation: remarks on variance control and complexity. *Monte Carlo Methods Appl.*, 13(1):37–70, 2007.

[87] G. C. Papanicolaou. Waves in one-dimensional random media. In *École d'Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, volume 1362 of *Lecture Notes in Math.*, pages 205–275. Springer, Berlin, 1988.

[88] G. C. Papanicolaou, D. Stroock, and S. R. S. Varadhan. Martingale approach to some limit theorems. pages ii+120 pp. Duke Univ. Math. Ser., Vol. III, 1977.

[89] E. Pardoux and A.L. Piatnitski. Homogenization of a nonlinear random parabolic partial differential equation. *Stochastic Process. Appl.*, 104(1):1–27, 2003.

[90] E. Pardoux and A. Yu. Veretennikov. On the Poisson equation and diffusion approximation. I. *Ann. Probab.*, 29(3):1061–1085, 2001.

[91] E. Pardoux and A. Yu. Veretennikov. On Poisson equation and diffusion approximation. II. *Ann. Probab.*, 31(3):1166–1192, 2003.

[92] E. Pardoux and A. Yu. Veretennikov. On the Poisson equation and diffusion approximation. III. *Ann. Probab.*, 33(3):1111–1133, 2005.

[93] M. Peyrard and T. Dauxois. *Physique des solitons*. EDP sciences, 2004.

[94] G. Da Prato and J. Zabczyk. *Stochastic Equations in Infinite Dimensions*, volume 44. Cambridge University Press, In Encyclopedia of Mathematics and Its Applications, 1992.

[95] J. Printems. On the discretization in time of parabolic stochastic partial differential equations. *M2AN Math. Model. Numer. Anal.*, 35(6):1055–1078, 2001.

[96] R. Radhakrishnan, M. Lakshmanan, and J. Hietarinta. Inelastic collision and switching of coupled bright solitons in optical fibers. *Phys. Rev. E*, 56:2213–2216, 1997.

[97] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, volume 293. Springer, A series of comprehensive Studies in Mathematics., 1999.

[98] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 1.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Foundations, Reprint of the second (1994) edition.

[99] J.M. Sanz-Serna. Methods for the numerical solution of the nonlinear Schroedinger equation. *Math. Comp.*, 43(167):21–27, 1984.

[100] D. W. Stroock and S. R. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1979.

[101] C. Sulem and P-L. Sulem. *The nonlinear Schrödinger equation, self-focusing and wave collapse*, volume 139 of applied Mathematical Sciences. Springer Verlag,New York, 1999.

[102] H. J. Sussman. On the gap between deterministic and stochastic differential equations. *Ann. Probability*, 6(1):19–41, 1978.

[103] T. R. Taha and M. J. Ablowitz. Analytical and numerical aspects of certain nonlinear evolution equations. II. Numerical, nonlinear Schrödinger equation. *J. Comput. Phys.*, 55(2):203–230, 1984.

[104] D. Talay. Résolution trajectorielle et analyse numérique des équations différentielles stochastiques. *Stochastics*, 9(4):275–306, 1983.

[105] D. Talay. Discrétisation d'une équation différentielle stochastique et calcul approché d'espérances de fonctionnelles de la solution. *RAIRO Modél. Math. Anal. Numér.*, 20(1):141–179, 1986.

[106] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991), 1990.

[107] P. K. A. Wai, W. L. Kath, and C. R. Menyuk. Nonlinear polarization mode dispersion in optical fibers with randomly varying birefringence. *Journal of Optical Society of America.*, 14(11):2697–2979, 1997.

[108] P. K. A. Wai and C. R. Menyuk. Polarization decorrelation in optical fibers with randomly varying birefringence. *Optics letters*, 19(19):1517–1519, 1994.

[109] P. K. A. Wai and C. R. Menyuk. Polarization evolution and dispersion in fibers with spatially varying birefringence. *J. Opt. Soc.*, 11(7):1288–1296, 1994.

[110] P. K. A. Wai and C. R. Menyuk. Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence. *Journal of Lightwave Technology*, 14(2):148–157, 1996.

[111] J. A. C. Weideman and B. M. Herbst. Split-step methods for the solution of the nonlinear Schrödinger equation. *SIAM J. Numer. Anal.*, 23(3):485–507, 1986.

[112] Y. Yamato. Stochastic differential equations and nilpotent Lie algebras. *Z. Wahrsch. Verw. Gebiete*, 47(2):213–229, 1979.

[113] J. Yang. Multisoliton perturbation theory for the Manakov equations and its applications to nonlinear optics. *Physical Review E*, 59(2):2393–2405, 1999.

[114] J. Yang. Suppression of Manakov soliton interference in optical fibers. *Phys. Rev. E*, 65:036606, 2002.

[115] J. Yang. *Nonlinear Waves in Integrable and Non-integrable Systems*. Society for Industrial and Applied Mathematics, 2010.

Analyse de modèles mathématiques pour la propagation de la lumière dans les fibres optiques en présence de biréfringence aléatoire

Résumé : L'étude de la propagation de la lumière dans les fibres optiques monomodes requiert la prise en compte de plusieurs phénomènes compliqués tels que la dispersion modale de polarisation et l'effet Kerr. Il s'est avéré que l'évolution de l'enveloppe lentement variable du champ électrique est bien décrite par un système couplé d'équations de Schrödinger non linéaires à coefficients aléatoires : l'équation de Manakov PMD. Cette équation fait intervenir différentes échelles dont le ratio est donné par un petit paramètre. La première partie de ce travail consiste à étudier le comportement asymptotique de la solution de l'équation de Manakov PMD lorsque ce petit paramètre tends vers zéro. En généralisant la théorie de l'Approximation-Diffusion au cadre de la dimension infinie, on a montré que la dynamique asymptotique est donnée par une équation aux dérivées partielles stochastiques dirigée par un mouvement brownien de dimension trois. Dans une seconde partie, nous proposons un schéma de différences finies de type Crank Nicolson pour cette équation pour lequel nous obtenons un ordre de convergence en probabilité d'ordre 1/2. La discrétisation du bruit doit être implicite afin d'obtenir un schéma conservatif et stable. Enfin la dernière partie est relative à la simulation numérique de la dispersion modale de polarisation et à ses effets sur la propagation et la collision de solitons de Manakov. Dans ce cadre, on propose une méthode de réduction de variance valable pour les équations aux dérivées partielles stochastiques.

Mots clés : Système couplé d'équations de Schrödinger non linéaires (Manakov PMD), Equations aux dérivées partielles stochastiques, Théorèmes limites, Schémas numériques, Ordre de convergence, Réduction de variance

Mathematical analysis of light propagation in optical fibers with randomly varying birefringence

Abstract : The study of light propagation in monomode optical fibers requires to take care of various complex phenomena such as the polarization mode dispersion (PMD) and the Kerr effect. It has been proved that the slowly varying envelope of the electric field is well described by a coupled non linear schrödinger equation with random coefficients called the Manakov PMD equation. The particularity of this equation is the presence of various length scales whose ratio is given by a small parameter. The first part of this thesis is concerned with the theoretical study of the asymptotic dynamics of the solution of the Manakov PMD equation as this parameter goes to zero. Generalizing the theory of the Diffusion Approximation in the infinite dimensional setting, we were able to prove that the asymptotic dynamics is given by a stochastic partial differential equation driven by three Brownian motions. In a second part, we propose a Crank Nicolson scheme for this equation and we prove that the order of convergence is 1/2. The discretization of the noise term is taken implicit so that the scheme is conservative and stable. Finally the last part is concerned with numerical simulations of the PMD and propagation and collision of Manakov solitons. The above scheme is implemented and we propose a variance reduction method valid in the context of stochastic partial differential equations.

Keywords : Coupled system of non linear Schrödinger equations (Manakov PMD), Stochastic partial differential equations, Limit theorems, Numerical schemes, Order of convergence, Variance reduction methods