The Interaction Picture method for solving the generalized nonlinear Schrödinger equation in optics

Stéphane Balac, Arnaud Fernandez, Fabrice Mahé, Florian Méhats, Rozenn Texier-Picard

THE INTERACTION PICTURE METHOD FOR SOLVING THE GENERALIZED NONLINEAR SCHRÖDINGER EQUATION IN OPTICS

STÉPHANE BALAC[†¶], ARNAUD FERNANDEZ[†], FABRICE MAHÉ[‡], FLORIAN MÉHATS[‡], AND ROZENN TEXIER-PICARD[§]

Abstract. The "interaction picture" (IP) method is a very promising alternative to Split-Step methods for solving certain type of partial differential equations such as the nonlinear Schrödinger equation used in the simulation of wave propagation in optical fibers. The method exhibits interesting convergence properties and is likely to provide more accurate numerical results than cost comparable Split-Step methods such as the Symmetric Split-Step method. In this work we investigate in detail the numerical properties of the IP method and carry out a precise comparison between the IP method and the Symmetric Split-Step method.

Key words. Interaction Picture method, Symmetric Split-Step method, Runge-Kutta method, Nonlinear optics, nonlinear Schrödinger equation.

AMS subject classifications. 65M12, 65M15, 65L06, 65T50, 78A60

1. Introduction. In this paper we study a mathematical model for the propagation of optical pulses in a single-mode fiber. We make the following usual assumptions, see for example [1] for a justification:

- the optical wave is assumed to be quasi-monochromatic, i.e. the spectral width of the pulse spectrum is small compared to the mean pulsation ω_0;
- the optical wave is supposed to maintain its polarization along the fiber length so that a scalar model (rather than a full vectorial one) is valid;
- the electric field $\mathbf{E}$ is linearly polarized along a direction $\mathbf{e_x}$ transverse to the direction of propagation $\mathbf{e_z}$ defined by the fiber axis and can be represented as a function of time τ and position $\mathbf{r} = (x, y, z)$ as

$$\mathbf{E}(\mathbf{r}, \tau) = A(z, t)\, F(x, y)\, \mathrm{e}^{-\mathrm{i}(\omega_0 \tau - kz)}\ \mathbf{e_x}$$

  where $A(z, t)$ represents the slowly varying pulse envelope, $F(x, y)$ is the modal distribution and k is the wavenumber. The pulse envelope $A(z, t)$ is expressed in a frame of reference, called the *retarded frame*, moving with the pulse at the "group velocity" $v_g = c/n_g$. The relation between the "local time" t in the retarded frame and the absolute time τ is: $t = \tau - z/v_g$.

Under these assumptions, the evolution of the slowly varying pulse envelope A is governed by the Generalized Nonlinear Schrödinger Equation (GNLSE) [1]

$$\frac{\partial}{\partial z} A(z,t) = -\frac{\alpha}{2} A(z,t) + \left(\sum_{n=2}^{N} \mathrm{i}^{n+1} \frac{\beta_n}{n!} \frac{\partial^n}{\partial t^n} A(z,t) \right) \tag{1.1}$$
$$+ \mathrm{i}\gamma \left(1 + \frac{\mathrm{i}}{\omega_0} \frac{\partial}{\partial t} \right) \left[A(z,t) \Big((1 - f_R) \left| A(z,t) \right|^2 + f_R \int_{-\infty}^{+\infty} h_R(s) |A(z, t-s)|^2 \, \mathrm{d}s \Big) \right].$$

The physical effects taken into account in (1.1) are the following. First, some linear effects are expressed through the linear attenuation/gain coefficient α and the linear dispersion coefficients β_n, $2 \leq n \leq N$ (it is assumed that $\beta_N \neq 0$), where e.g. β_2 expressed in units

[†]FOTON, Université de Rennes I, CNRS, UEB, Enssat, 6 rue de Kerampont, 22305 Lannion, France
[‡]IRMAR, Université de Rennes I, CNRS, UEB, Campus de Beaulieu, 35042 Rennes, France
[§]IRMAR, ENS Rennes, Université de Rennes I, CNRS, UEB, av. R. Schuman, 35170 Bruz, France
[¶]This work has been undertaken under the framework of the Green-Laser project and was partially supported by Conseil Régional de Bretagne, France.

$\mathrm{ps}^2\mathrm{km}^{-1}$ accounts for chromatic effects. In standard silica fibers we have $\alpha \sim 4\,10^{-2}\,\mathrm{km}^{-1}$ and $\beta_2 \sim 50\,\mathrm{ps}^2\mathrm{km}^{-1}$ for wavelengths in the visible region. In the anomalous dispersion regime we have $\beta_2 < 0$ (typically $\beta_2 \sim -20\,\mathrm{ps}^2\mathrm{km}^{-1}$ for wavelengths near $1.5\,\mu\mathrm{m}$) and the fiber can support optical soliton. Moreover, equation (1.1) includes the following nonlinear effects. The first-order partial derivative with respect to time takes into account the dispersion of the nonlinearity through the simplified optical shock parameter $\tau_{\mathrm{shock}} = 1/\omega_0$. Instantaneous Kerr effect manifests itself through the term $(1 - f_R)\,|A|^2$. The delayed Raman contribution in the time domain is taken into account through the convolution product between the instantaneous power $|A|^2$ and the Raman time response function h_R. For silica fibers, an expression for h_R is proposed in [1]. The constant f_R represents the fractional contribution of the delayed Raman response to nonlinear polarization and takes a value around 0.2. The nonlinear parameter γ typically takes values in the range 1 to $10\,\mathrm{W}^{-1}\mathrm{km}^{-1}$. Our interest for the GNLSE originates from a study of pulsed laser systems of MOPFA type (a master oscillator coupled with fiber amplifier usually a cladding-pumped high-power amplifier based on an ytterbium-doped fiber), see [22] for details. Equation (1.1) is to be solved for all z in a given interval $[0, L]$ where L denotes the length of the fiber and for all "local time" $t \in \mathbb{R}$. It is considered together with the boundary condition $A(0, t) = a_0(t)$ for all $t \in \mathbb{R}$, where a_0 is a "smooth" complex valued function, typically in the Schwartz space $\mathcal{S}(\mathbb{R})$ [1].

Recently a "fourth-order Runge-Kutta method in the interaction picture method" (RK4-IP method) has been proposed in [27] as an alternative to Split-Step methods for solving the GNLSE (1.1). The method has been numerically experimented on benchmark problems in optics in [27, 25]. This experimental investigation indicates that the RK4-IP method exhibits interesting convergence properties and provides more accurate numerical results than comparable Split-Step methods such as the Symmetric Split-Step method based on Strang formula. The scope of the present work is to investigate the mathematical and numerical features of the RK4-IP method for solving the GNLSE. The numerical tests in Section 4.2 are done with the general model given by equation (1.1). But the mathematical study of equation (1.1) itself and the set up of the corresponding functional framework for the study of the RK4-IP method is arduous due to the complicated expression of the nonlinear part of the equation. For this reason, to proceed with the mathematical justification of the RK4-IP method and its numerical analysis, we will consider the following simplified version of equation (1.1), corresponding to $\omega_0 = +\infty$ and $f_R = 0$:

$$\frac{\partial}{\partial z}A(z) = -\frac{\alpha}{2}A(z) + \left(\sum_{n=2}^{N} \mathrm{i}^{n+1}\frac{\beta_n}{n!}\,\frac{\partial^n}{\partial t^n}A(z)\right) + \mathrm{i}\gamma A(z)|A(z)|^2, \tag{1.2}$$

where $A(z)$ denotes the function $t \in \mathbb{R} \mapsto A(z, t) \in \mathbb{C}$. In (1.2), the nonlinear part is actually the same as in the standard nonlinear Schrödinger (NLS) equation in optics [1]

$$\frac{\partial}{\partial z}A(z) = -\frac{\alpha}{2}A(z) - \frac{\mathrm{i}}{2}\beta_2\frac{\partial^2}{\partial t^2}A(z) + \mathrm{i}\gamma\, A(z)\,|A(z)|^2\,. \tag{1.3}$$

In this work we investigate the numerical properties of the RK4-IP method for the Generalized Nonlinear Schrödinger Equation and we make a precise comparison between the RK4-IP method and the symmetric Split-Step Fourier method (SSF method) when used with the classical fourth-order Runge-Kutta (RK4) formula. At present time, the SSF method is the most frequently used method for simulating wave propagation in optical fibers, see e.g. [1, 32, 35]. We show that the RK4-IP and SSF-RK4 methods are equivalent regarding the numerical cost, due to similar computational inner structures, but the RK4-IP method exhibits a convergence rate of order 4 where with respect to the spatial discretization step-size

whereas the SSF method is limited by the second order accuracy of Strang splitting formula and exhibits a convergence rate of order 2.

The outline of the document is the following. In Section 2 we define the useful mathematical tools and we generalize to equation (1.2) the existence result known for (1.3). Section 3 deals with the IP method and the various aspects of its numerical implementation. We also analyze the numerical error of the RK4-IP method when applied to the simplified version (1.2) of the GNLSE. Our main result, which states the convergence rate of order 4 for the RK4-IP method, is given in Theorem 3.5. Finally, in Section 4, we present in a similar way the SSF method and we compare the two methods both from a theoretical point of view (Proposition 4.1) and on numerical simulation examples.

2. Mathematical toolbox. We denote by $\mathbb{L}^p(\mathbb{R},\mathbb{C})$, $p \in [1,+\infty[$ the space of complex-valued functions over $\mathbb{R}$ whose p-th powers are integrable and by $\mathbb{H}^m(\mathbb{R},\mathbb{C})$ for $m \in \mathbb{N}^*$ the Sobolev set of functions in $\mathbb{L}^2(\mathbb{R},\mathbb{C})$ with derivatives up to order m in $\mathbb{L}^2(\mathbb{R},\mathbb{C})$. For convenience, we will also use the notation $\mathbb{H}^0(\mathbb{R},\mathbb{C})$ for $\mathbb{L}^2(\mathbb{R},\mathbb{C})$ and $\mathbb{L}^\infty(\mathbb{R},\mathbb{C})$ for the space of essentially bounded functions. The Sobolev spaces $\mathbb{H}^m(\mathbb{R},\mathbb{C})$, $m \in \mathbb{N}$, are equipped with the usual norms denoted by $\| \ \|_m$. For $k,n \in \mathbb{N}$ and $I \subset \mathbb{R}$, we denote by $\mathcal{C}^k(I;\mathbb{H}^n(\mathbb{R},\mathbb{C}))$ the space of functions $u : z \in I \mapsto u(z) \in \mathbb{H}^n(\mathbb{R},\mathbb{C})$ with continuous derivatives up to order k (or just continuous when $k=0$). For any $m \in \mathbb{N}$ and any interval $I \subset \mathbb{R}$, we define

$$E_{m,N}(I) = \bigcap_{k=0}^{\lfloor m/N \rfloor} \mathcal{C}^k(I, \mathbb{H}^{m-Nk}(\mathbb{R},\mathbb{C})), \tag{2.1}$$

where $\lfloor s \rfloor$ denotes the integer part of $s \in \mathbb{R}_+$.

A comprehensive mathematical framework for the NLS equation (1.3) exists in the literature [12, 13]. Namely, it is known that for $a_0 \in \mathbb{H}^2(\mathbb{R},\mathbb{C})$ there exists a unique A belonging to $\mathcal{C}^0(\mathbb{R};\mathbb{H}^2(\mathbb{R},\mathbb{C})) \bigcap \mathcal{C}^1(\mathbb{R};\mathbb{L}^2(\mathbb{R},\mathbb{C}))$ solution of equation (1.3) satisfying $A(0) = a_0$. This result can be extended to the GNLSE with $\omega_0 = +\infty$ and $f_R = 0$, i.e. to (1.2), as follows.

THEOREM 2.1. *For all $a_0 \in \mathbb{H}^m(\mathbb{R},\mathbb{C})$, with $m \in \mathbb{N}^*$, there exists a unique maximal solution $A \in E_{m,N}([0,Z[)$, with $Z \in]0,+\infty]$, to the problem* (1.2). *This solution satisfies*

$$\|A(z)\|_0 = \mathrm{e}^{-\frac{\alpha}{2}z} \|a_0\|_0 \qquad \text{for all } z \in [0,Z[\tag{2.2}$$

and it is maximal in the sense that

$$\text{if} \quad Z < +\infty \quad \text{then} \quad \limsup_{z \to Z} \|A(z)\|_{\mathbb{L}^\infty(\mathbb{R},\mathbb{C})} = +\infty. \tag{2.3}$$

Moreover, if N is even and $m \geq N/2$ then the solution is global, i.e. $Z = +\infty$.

The proof of this result can be found in Appendix A.

In order to simplify the presentation of the interaction picture method, we now reformulate our problems (1.1) and (1.2) in a more abstract and unified way. To this aim, we need a few notations and technical results. We denote by $\mathcal{D}$ the unbounded linear operator on $\mathbb{L}^2(\mathbb{R},\mathbb{C})$ with domain $\mathbb{H}^N(\mathbb{R},\mathbb{C})$, $N \in \mathbb{N}^*$, defined as

$$\mathcal{D} : U \in \mathbb{H}^N(\mathbb{R},\mathbb{C}) \longmapsto \sum_{n=2}^{N} \mathrm{i}^{n+1} \frac{\beta_n}{n!} \frac{\partial^n}{\partial t^n} U \in \mathbb{L}^2(\mathbb{R},\mathbb{C}).$$

For $N = 2$, it is well known [12, 13] that this operator generates a continuous group of bounded operators on $\mathbb{L}^2(\mathbb{R},\mathbb{C})$, denoted by $\exp(z\mathcal{D})$ with $z \in \mathbb{R}$. For $N > 2$, the same property holds and we have the following lemma.

LEMMA 2.2. *Let $\varphi \in \mathbb{H}^m(\mathbb{R},\mathbb{C})$, where $m \in \mathbb{N}$. Then the problem*

$$\forall z \in \mathbb{R}, \quad \frac{\partial}{\partial z} U(z) = \mathcal{D}U(z), \qquad U(0) = \varphi \tag{2.4}$$

has a unique solution $U(z) = \exp(z\mathcal{D})\varphi$ with $U : z \in \mathbb{R} \mapsto U(z) \in E_{m,N}(\mathbb{R})$ and it satisfies for all $z \in \mathbb{R}$ the relations $\|U(z)\|_j = \|\varphi\|_j$ for all $j \in \{0,\dots,m\}$.

Let us now denote the two nonlinear operators appearing respectively in the GNLSE (1.1) and in its simplified version (1.2) by

$$\mathcal{N} : u \longmapsto -\frac{\alpha}{2}u + \mathrm{i}\gamma\left(1 + \frac{\mathrm{i}}{\omega_0}\frac{\partial}{\partial t}\right)\left[(1-f_R)u|u|^2 + f_R u\int_{-\infty}^{+\infty} h_R(s)|u(\cdot - s)|^2 \,\mathrm{d}s\right]$$

and $\mathcal{N}_0 : u \mapsto -\frac{\alpha}{2}u + \mathrm{i}\gamma\, u|u|^2$. Both $\mathcal{N}$ and $\mathcal{N}_0$ are considered as unbounded nonlinear operators on $\mathbb{L}^2(\mathbb{R},\mathbb{C})$. Note that $\mathcal{N}_0$ is nothing but $\mathcal{N}$ when $\omega_0 = +\infty$ and $f_R = 0$. Problems (1.1) and (1.2) then can be reformulated respectively as

$$\forall z \in \mathbb{R}, \quad \frac{\partial}{\partial z} A(z) = \mathcal{D}A(z) + \mathcal{N}(A)(z), \qquad A(0) = a_0, \tag{2.5}$$

and

$$\forall z \in \mathbb{R}, \quad \frac{\partial}{\partial z} A(z) = \mathcal{D}A(z) + \mathcal{N}_0(A(z)), \qquad A(0) = a_0. \tag{2.6}$$

It can be useful to note that another splitting is possible for $\mathcal{D}$ and $\mathcal{N}$: the term $-\frac{1}{2}\alpha A(z)$ can be added to the linear operator instead of the nonlinear one. Although it would seem more natural to add this term to the linear part, this would lead to a group of operators $\exp(z\mathcal{D})$ which does no longer preserve the H^j norms which would make the proof of Theorem 2.1 less straightforward. For the simplicity of this proof, we have chosen to add this term to the nonlinear part. In the numerical approach, both choices are equivalent regarding the computational cost. In particular, they do not impact the number of Fourier transforms to be computed in (3.14).

As we said above, the numerical experiments presented below are done on (1.1) (or equivalently, (2.5)), but the mathematical results concern the simplified problem (1.2) (or equivalently, (2.6)). Indeed, due to the time derivative, the nonlinearity $\mathcal{N}$ is not continuous on any Sobolev space (unless $\omega_0 = +\infty$) and the solution of the Cauchy problem for (1.1) would rely on the smoothing properties of the linear group connected to the higher-order dispersion, which goes beyond the scope of this paper on a numerical method. In contrast, the simplified nonlinearity $\mathcal{N}_0$ is locally Lipschitz continuous on every Sobolev space $\mathbb{H}^m$ of exponent $m \geq 1$. We recall indeed that we have the inclusion $\mathbb{H}^1(\mathbb{R},\mathbb{C}) \subset \mathbb{L}^\infty(\mathbb{R},\mathbb{C})$ and we summarize in the following lemma (stated without proof, see [11]) some classical useful properties of $\mathcal{N}_0$.

LEMMA 2.3. *The nonlinear operator $\mathcal{N}_0$ satisfies the following local Lipschitz condition. For all $M > 0$ and $m \in \mathbb{N}^*$, there exists $\Lambda_{m,M} > 0$ such that for all $u, v \in \mathbb{H}^m(\mathbb{R},\mathbb{C})$ such that $\|u\|_m \leq M$ and $\|v\|_m \leq M$, we have*

$$\|\mathcal{N}_0(u) - \mathcal{N}_0(v)\|_m \leq \Lambda_{m,M}\,\|u - v\|_m\,.$$

Moreover for all $m \in \mathbb{N}^*$ we have $\mathcal{N}_0 \in \mathcal{C}^\infty(\mathbb{H}^m(\mathbb{R},\mathbb{C}),\mathbb{H}^m(\mathbb{R},\mathbb{C}))$. Finally, there exists $\Lambda_{0,M} > 0$ such that, for all $u, v \in \mathbb{H}^1(\mathbb{R},\mathbb{C})$ such that $\|u\|_1 \leq M$ and $\|v\|_1 \leq M$, we have

$$\|\mathcal{N}_0(u) - \mathcal{N}_0(v)\|_0 \leq \Lambda_{0,M}\,\|u - v\|_0\,.$$

3. Solving the GNLSE by the Interaction Picture method.

3.1. Presentation of the numerical approach. The main idea of the Interaction Picture (IP) method is a change of unknown to transform the NLSE or GNLSE for the unknown A into a new equation where only remains an explicit reference to the partial derivative with respect to the space variable z and where the time variable t appears as a parameter. This new equation can be solved numerically using the usual methods for ordinary differential equations (ODE) such as the standard fourth order Runge-Kutta (RK4) method. Then, by using the inverse transform we obtain the approximate values of the unknown A at the grid points of a subdivision of the fiber length interval $[0, L]$. This numerical approach is referred to as the RK4-IP method.

The RK4-IP method has been developed by the Bose-Einstein condensate theory group of R. Ballagh from the Jack Dodd Centre at the University of Otago (New Zealand) in the 90's for solving the Gross-Pitaevskii equation which is ubiquitous in Bose condensation. It was described in the Ph.D. thesis of B. M. Caradoc-Davies [10] and M. J. Davis [17]. In this latter work an embedded Runge-Kutta scheme based on a Cash-Krap formula was additionally used in conjunction with the RK4-IP method for adaptive step-size control purposes but the efficiency of the method was judged disappointing. Recently an efficient embedded RK method based on Dormand and Prince RK4(3)-T formula [18] and specifically designed for the IP method has been proposed in [5].

The name "Interaction Picture" and the change of unknown at the heart of the method originate from quantum mechanics [34, 23] where it is usual to chose an appropriate "picture" in which the physical properties of the studied system can be easily revealed and the calculation made simpler. The interaction picture, considered as an intermediate between the Schrödinger picture and the Heisenberg picture, is useful in quantum optics for solving problems with time-dependent Hamiltonians in the form $H(t) = H_0 + V(t)$ where H_0 is a Hamiltonian independent of the time and its eigenvalues are easy to compute whereas V is a time-dependent potential which can be complicated.

The RK4-IP method can be interpreted as an exponential Runge-Kutta method according to the general form presented in the review article of Hochbruck and Ostermann [26] for semilinear parabolic or highly oscillatory problems. Exponential integrators are a family of classical tools for semilinear equations; they include in particular *Lawson methods* [29, 20], *integrating factor methods*, *exponential time-differencing methods* (see [16, 28] and the references therein), and *collocation methods*. These methods have raised a revived interest in the last decade, and have been widely applied to the Schrödinger equation, [19, 9, 8, 14, 15]. Our approach relies on the same change of unknown as the integrating factor method, but in our case the change is local on each subinterval instead of being global. This gives rise to a reduction of the number of Fourier transforms used in the numerical scheme and enables a direct comparison with usual Split-Step methods.

3.2. The idea of the Interaction Picture. In the following, a_0 is chosen once and for all and we suppose the existence of a unique solution of (2.5) on some interval $[0, L]$ with $L > 0$. This result is proved by theorem 2.1 when $\mathcal{N}$ is replaced by $\mathcal{N}_0$ and when $a_0 \in \mathbb{H}^m(\mathbb{R}, \mathbb{C})$ with $m \in \mathbb{N}^*$: in this case, we have a unique maximal solution on some interval $[0, Z)$, $Z > 0$. We suppose that the same result holds for the complete equation if a_0 has a sufficient regularity, and we choose $0 < L < Z$.

The integration interval $[0, L]$ is then divided into K subintervals. The spatial grid points are denoted z_k, $k \in \{0, \dots, K\}$ where $0 = z_0 < z_1 < \cdots < z_{K-1} < z_K = L$. For convenience we assume a constant grid spacing $h = L/K$ but this assumption is not a limitation of the method and an adaptive step-size version of the RK4-IP method is propounded in [5]. For $k \in \{0, \dots, K-1\}$ we set $z_{k+\frac{1}{2}} = z_k + \frac{h}{2}$.

We introduce the following auxiliary problems, for $0 \leq k \leq K-1$:

$$\frac{\partial}{\partial z} A_k(z) = \mathcal{D} A_k(z) + \mathcal{N}(A_k)(z) \quad \forall z \in [z_k, z_{k+1}], \qquad A_k(z_k) = a_k, \tag{3.1}$$

where a_k is a given function in $\mathbb{H}^m(\mathbb{R}, \mathbb{C})$, $m \in \mathbb{N}^*$. Solving problem (2.5) for $a_0 \in \mathbb{H}^m(\mathbb{R}, \mathbb{C})$ is equivalent to solving the sequence of connected problems (3.1) for $k \in \{0, \ldots, K-1\}$, with $a_k(t)$ defined for all $t \in \mathbb{R}$ and $k \geq 1$ by $a_k(t) = A_{k-1}(z_k, t)$, and

$$\forall z \in [z_k, z_{k+1}], \quad A(z) = A_k(z).$$

To solve problem (3.1) we introduce the new unknown mapping A_k^{ip} defined for $(z,t) \in [z_k, z_{k+1}] \times \mathbb{R}$ by

$$A_k^{\mathrm{ip}}(z,t) = \exp(-(z - z_{k+\frac{1}{2}})\mathcal{D}) A_k(z,t). \tag{3.2}$$

From Lemma 2.2, it can be deduced that $A_k \in E_{m,N}([z_k, z_{k+1}])$ is equivalent to $A_k^{\mathrm{ip}} \in E_{m,N}([z_k, z_{k+1}])$, where the space $E_{m,N}$ is defined by (2.1). From the chain rule, we deduce that A_k^{ip} satisfies the folowing problem

$$\forall z \in [z_k, z_{k+1}], \quad \frac{\partial}{\partial z} A_k^{\mathrm{ip}}(z) = \mathcal{G}_k(z, A_k^{\mathrm{ip}}(z)), \qquad A_k^{\mathrm{ip}}(z_k) = \exp(-(z_k - z_{k+\frac{1}{2}})\mathcal{D}) a_k, \tag{3.3}$$

where $\mathcal{G}_k$ is defined by

$$\mathcal{G}_k(z, v) = \exp(-(z - z_{k+\frac{1}{2}})\mathcal{D}) \left[\mathcal{N}\left(\exp((z - z_{k+\frac{1}{2}})\mathcal{D}) v \right) \right]. \tag{3.4}$$

Conversely, if A_k^{ip} is a solution to (3.3) then $A_k = \exp((z - z_{k+\frac{1}{2}})\mathcal{D}) A_k^{\mathrm{ip}}$ is a solution to (3.1). At this level of generality, this statement is formal. However, in the case $\omega_0 = +\infty$, $f_R = 0$, the nonlinearity $\mathcal{N}$ is reduced to $\mathcal{N}_0$ and we can state the following precise equivalence result.

Lemma 3.1. *Let $\omega_0 = +\infty$, $f_R = 0$. Then the problems* (3.1) *and* (3.3) *are equivalent, in the sense that, for any $a_k \in \mathbb{H}^m(\mathbb{R}, \mathbb{C})$, $m \in \mathbb{N}^*$, each one of them has a unique solution belonging to $E_{m,N}([z_k, z_{k+1}])$ and the solutions are related to each other through relation* (3.2).

To prove this equivalence, we first deduce from Lemmas 2.2 and 2.3 the following result on the mapping $\mathcal{G}_k$ defined in (3.4).

Lemma 3.2. *Let $\omega_0 = +\infty$, $f_R = 0$. Then, for all $m \geq 1$, the function $z \mapsto \mathcal{G}_k(z, v(z))$ belongs to $E_{m,N}([z_k, z_{k+1}])$ whenever $v \in E_{m,N}([z_k, z_{k+1}])$. Moreover, for all $M > 0$, let $B_{m,M} = \{w \in \mathbb{H}^m(\mathbb{R}, \mathbb{C}) : \|w\|_m \leq M\}$. For all $m \in \mathbb{N}^*$ and for all $u, v \in B_{m,M}$ we have the estimate*

$$\forall z \in \mathbb{R}, \qquad \|\mathcal{G}_k(z,u) - \mathcal{G}_k(z,v)\|_m \leq \Lambda_{m,M} \|u - v\|_m .$$

Finally, for all $u, v \in B_{1,M}$, we have

$$\|\mathcal{G}_k(z,u) - \mathcal{G}_k(z,v)\|_0 \leq \Lambda_{0,M} \|u - v\|_0 .$$

In this Lemma, the constants $\Lambda_{m,M}$ and $\Lambda_{0,M}$ are the same constants as in Lemma 2.3.

Proof. [of Lemma 3.1] Let $a_0 \in \mathbb{H}^m(\mathbb{R}, \mathbb{C})$ with $m \in \mathbb{N}^*$ and let $0 < L < Z$, where Z is defined in Theorem 2.1 (recall that, if N is even, then we have $Z = +\infty$ and then any $L > 0$ is allowed). Considering the unique solution $A(z)$ of (1.2) given by Theorem 2.1, we have $A \in E_{m,N}([0, L])$, and in particular we have the bound

$$M := \max_{z \in [0,L]} \|A(z)\|_m < +\infty. \tag{3.5}$$

Recalling that $a_k = A(z_k)$, we have then $\|a_k\|_m \leq M$ for all $0 \leq k \leq K = L/h$. From Theorem 2.1 and Lemma 3.2 one can deduce that each one of the problems (3.1) and (3.3) has a unique solution belonging to $E_{m,N}([z_k, z_{k+1}])$ and the solutions are related to each other through relation (3.2). Lemma 3.1 is proved. □

Remark 1. Although it is possible to choose any point in the interval $[z_k, z_{k+1}]$ instead of the middle point $z_{k+\frac{1}{2}}$ in (3.2), this particular choice is very relevant to save computations as detailed in section 3.3.

The major interest in doing the change of unknown (3.2) is that on the contrary to problem (3.1), the new problem (3.3) for the unknown A_k^{ip} does not anymore involve explicitly partial derivative with respect to the time variable t. Derivation with respect to time now occurs through the operators $\exp(\pm(z - z_{k+\frac{1}{2}})\mathcal{D})$. Problem (3.3) can be solved numerically using a standard quadrature scheme for ODE such as the classical 4th-order Runge-Kutta (RK) method [7].

We can summarize the IP method for solving problem (2.5) in the following way. The fiber length $[0, L]$ is divided into subintervals $[z_k, z_{k+1}]$, $k \in \{0, \ldots, K-1\}$, and over each subinterval $[z_k, z_{k+1}]$ the following three nested problems are solved:

$$\forall z \in [z_k, z_{k+\frac{1}{2}}], \quad \frac{\partial}{\partial z} A_k^+(z) = \mathcal{D} A_k^+(z), \qquad A_k^+(z_k) = A_{k-1}(z_k), \tag{3.6}$$

where $A_{k-1}(z_k)$, $k \geq 1$, represents the solution to (3.1) at grid point z_k computed at step $k-1$, and $A_{-1}(z_0) = a_0$;

$$\forall z \in [z_k, z_{k+1}], \quad \frac{\partial}{\partial z} A_k^{\mathrm{ip}}(z) = \mathcal{G}_k(z, A_k^{\mathrm{ip}}(z)) \qquad A_k^{\mathrm{ip}}(z_k, t) = A_k^+(z_{k+\frac{1}{2}}), \tag{3.7}$$

where $\mathcal{G}_k$ is defined by (3.4) and $A_k^+(z_{k+\frac{1}{2}}) = \exp(\frac{h}{2}\mathcal{D}) A_{k-1}(z_k)$ is the solution to (3.6) at grid point $z_{k+\frac{1}{2}}$;

$$\forall z \in [z_{k+\frac{1}{2}}, z_{k+1}], \quad \frac{\partial}{\partial z} A_k^-(z) = \mathcal{D} A_k^-(z), \qquad A_k^-(z_{k+\frac{1}{2}}) = A_k^{\mathrm{ip}}(z_{k+1}), \tag{3.8}$$

where $A_k^{\mathrm{ip}}(z_{k+1})$ represents the solution to (3.7) at grid point z_{k+1}. Finally, the solution of (2.5) at grid point z_{k+1} is given by $A_k(z_{k+1}) = A_k^-(z_{k+1})$.

3.3. The fourth-order Runge-Kutta scheme in the Interaction Picture method. For $k \in \{0, \ldots, K-1\}$ we denote by u_k^{ip} (resp. u_k) the approximation of the solution $A_k^{\mathrm{ip}}(z_k)$ (resp. $A_k(z_k)$) to problem (3.7) (resp. (3.1)) at grid point z_k. One step of the classical 4th-order RK formula is used to approximate the solution to problem (3.7) at grid point z_{k+1}:

$$A_k^{\mathrm{ip}}(z_{k+1}) \approx u_{k+1}^{\mathrm{ip}} = u_k^{\mathrm{ip}} + h\, \Phi(z_k, u_k^{\mathrm{ip}}; h) \tag{3.9}$$

where the mapping $\Phi(z_k, u_k^{\mathrm{ip}}; h) = \frac{1}{6}(\alpha_1 + 2\alpha_2 + 2\alpha_3 + \alpha_4)$ is the *increment function* of the RK4 method, with

$$\begin{aligned}
\alpha_1 &= \mathcal{G}_k(z_k, u_k^{\mathrm{ip}}) = \exp(\tfrac{h}{2}\mathcal{D})\, \mathcal{N}(\exp(-\tfrac{h}{2}\mathcal{D}) u_k^{\mathrm{ip}}) \\
\alpha_2 &= \mathcal{G}_k(z_{k+\frac{1}{2}}, u_k^{\mathrm{ip}} + \tfrac{h}{2}\alpha_1) = \mathcal{N}(u_k^{\mathrm{ip}} + \tfrac{h}{2}\alpha_1) \\
\alpha_3 &= \mathcal{G}_k(z_{k+\frac{1}{2}}, u_k^{\mathrm{ip}} + \tfrac{h}{2}\alpha_2) = \mathcal{N}(u_k^{\mathrm{ip}} + \tfrac{h}{2}\alpha_2) \\
\alpha_4 &= \mathcal{G}_k(z_{k+1}, u_k^{\mathrm{ip}} + h\alpha_3) = \exp(-\tfrac{h}{2}\mathcal{D})\, \mathcal{N}(\exp(\tfrac{h}{2}\mathcal{D})[u_k^{\mathrm{ip}} + h\alpha_3]).
\end{aligned} \tag{3.10}$$

Using the change of unknown (3.2) we deduce the following approximation of the solution to problem (3.1) at grid point z_{k+1}

$$A_k(z_{k+1}) \approx u_{k+1} = \exp(\tfrac{h}{2}\mathcal{D})\left(u_k^{\mathrm{ip}} + \tfrac{h}{6}\left(\alpha_1 + 2\alpha_2 + 2\alpha_3 + \alpha_4\right)\right). \tag{3.11}$$

Actually we are only interested in computing an approximate solution to problem (2.5) given by (3.11) and the use of the new unknown A_k^{ip} is a go-between in the computational approach. We can therefore recast the above approximation scheme in order to reduce the cost of the method as follows:

$$\begin{aligned}
u_k^{\mathrm{ip}} &= \exp(\tfrac{h}{2}\mathcal{D})u_k \\
\alpha_1 &= \exp(\tfrac{h}{2}\mathcal{D})\,\mathcal{N}(u_k), \quad \alpha_2 = \mathcal{N}(u_k^{\mathrm{ip}} + \tfrac{h}{2}\alpha_1), \quad \alpha_3 = \mathcal{N}(u_k^{\mathrm{ip}} + \tfrac{h}{2}\alpha_2), \\
\alpha_4' &= \mathcal{N}(\exp(\tfrac{h}{2}\mathcal{D})[u_k^{\mathrm{ip}} + h\alpha_3]) \\
u_{k+1} &= \exp(\tfrac{h}{2}\mathcal{D})\left(u_k^{\mathrm{ip}} + \tfrac{h}{6}\left(\alpha_1 + 2\alpha_2 + 2\alpha_3\right)\right) + \tfrac{h}{6}\alpha_4'.
\end{aligned} \tag{3.12}$$

Compared to the computational scheme (3.10)–(3.11), the new formulation saves one evaluation of $\exp(-\frac{h}{2}\mathcal{D})$. Of course the key-point in the computational procedure (3.12) lies in the way the linear operator $\exp(\frac{h}{2}\mathcal{D})$ and nonlinear operator $\mathcal{N}$ are computed.

According to Lemma 2.2, the expression $\exp(\frac{h}{2}\mathcal{D})f$ for $f \in \mathbb{H}^m(\mathbb{R},\mathbb{C})$ can be obtained in the usual way by solving problem (2.4) by the Fourier Transform (FT) approach. Namely, we have

$$\exp(\tfrac{h}{2}\mathcal{D})f = \mathcal{F}^{-1}\big[\widehat{f}\ \mathrm{e}^{\widehat{d}_\nu \frac{h}{2}}\big], \quad \text{where } \widehat{d}_\nu = \mathrm{i}\sum_{n=2}^{N}\frac{\beta_n}{n!}(2\pi\nu)^n, \tag{3.13}$$

$\widehat{f}$ is the FT of f and $\mathcal{F}^{-1}$ denotes the inverse FT operator. It accounts for 2 FT evaluations by the FFT algorithm. Similarly, each nonlinear term in (3.12) involving the nonlinear operator $\mathcal{N}$ can be computed by use of the FT method. Taking advantage of the properties of the FT with respect to derivation and convolution, it requires 4 FT evaluations by the FFT algorithm.

The computational procedure (3.12) can be recast as follows to reduce to 17 the number of FT to achieve at each computational step $k \in \{0,\dots,K-1\}$:

$$\begin{aligned}
\widehat{u}_k^{\mathrm{ip}} &= \mathrm{e}^{\widehat{d}_\nu \frac{h}{2}} \times \widehat{u}_k \\
\widehat{\alpha}_1 &= \mathrm{e}^{\widehat{d}_\nu \frac{h}{2}} \times \widehat{\mathcal{N}}(u_k), \quad \widehat{\alpha}_2 = \widehat{\mathcal{N}}\big(\mathcal{F}^{-1}(\widehat{u}_k^{\mathrm{ip}} + \tfrac{h}{2}\widehat{\alpha}_1)\big), \quad \widehat{\alpha}_3 = \widehat{\mathcal{N}}\big(\mathcal{F}^{-1}(\widehat{u}_k^{\mathrm{ip}} + \tfrac{h}{2}\widehat{\alpha}_2)\big) \\
\widehat{\alpha}_4' &= \widehat{\mathcal{N}}\big(\mathcal{F}^{-1}(\mathrm{e}^{\widehat{d}_\nu \frac{h}{2}} \times [\widehat{u}_k^{\mathrm{ip}} + h\widehat{\alpha}_3])\big) \\
\widehat{u}_{k+1} &= \mathrm{e}^{\widehat{d}_\nu \frac{h}{2}} \times \big(\widehat{u}_k^{\mathrm{ip}} + \tfrac{h}{6}\left(\widehat{\alpha}_1 + 2\widehat{\alpha}_2 + 2\widehat{\alpha}_3\right)\big) + \tfrac{h}{6}\widehat{\alpha}_4' \\
u_{k+1} &= \mathcal{F}^{-1}\big(\widehat{u}_{k+1}\big)
\end{aligned} \tag{3.14}$$

The algorithm of the RK4-IP method is given in Appendix B.

An important point to be mentioned about the RK4-IP method concerns the values of the *elementary quadrature nodes* $c_1 = 0$, $c_2 = \frac{1}{2}$, $c_3 = \frac{1}{2}$ and $c_4 = 1$ in the classical 4th-order RK formula for the efficiency of the RK4-IP method. Indeed, in conjunction with the choice of $z_{k+\frac{1}{2}} = z_k + \frac{h}{2}$ in the change of unknown (3.2), these particular values for the c_i coefficients enable the cancellation of 4 exponential operator terms in (3.10) compared to other possible sets of values, and therefore save up computations. As well, any other value

z' in the set $]z_k, z_{k+1}[$ could have been chosen rather than the particular value $z_{k+\frac{1}{2}}$ in the change of unknown (3.2); however the benefit of the cancellation of 4 exponential operator terms in (3.10) would have been lost.

A more immediate way of exploiting the Interaction Picture ideas would be to use a change of unknown similar to the one given by (3.2) but for the original problem (2.5), as in the usual *Integrating Factor* method. We mean by this to use a unique new unknown on the whole interval $[0, L]$ rather than using various change of unknowns on each of the subintervals $[z_k, z_{k+1}]$. The only difference between these 2 approaches, corresponding respectively to a local and a global change of unknown, lies in the way the subdivision of interval $[0, L]$ is introduced. In the first approach it is used upstream of the RK4 scheme to set an equivalent sequence of linked problems whereas in the second one it is inherent to the RK4 discretisation. The advantage of using the local change of unknown lies in the numerical evaluation of the $\exp(-(z - z')\mathcal{D})$ operator. We have seen that the $\exp(\frac{h}{2}\mathcal{D})$ operator can be efficiently computed by use of FT and that some cancellations happen reducing the number of terms to be evaluated. It would not be the case with the second approach where we would have to compute the operator $\exp(-(z_{k+1} - z')\mathcal{D})$ for all $k \in \{0, \ldots, K-1\}$.

3.4. Error analysis of the Interaction Picture method. In this section, we proceed to the mathematical analysis of the IP method. Hence, we only consider here the simplified version (1.2) of the GNLSE (1.1) for $a_0 \in \mathbb{H}^m(\mathbb{R}, \mathbb{C})$ with $m \geq 4N$.

As we have stated in Lemma 3.1, problems (3.1) and (3.3) for $a_k \in \mathbb{H}^m(\mathbb{R}, \mathbb{C})$, have a unique solution belonging to $E_{m,N}([z_k, z_{k+1}])$ and the solutions are related to each other through relation (3.2). One can also deduce from Lemma 3.2 that A_k^{ip} is slightly more regular than A_k in the sense that $A_k \in E_{m,N}([z_k, z_{k+1}])$, whereas we have

$$A_k^{\mathrm{ip}} \in \bigcap_{j=0}^{\lfloor m/N \rfloor} \mathcal{C}^{j+1}([0, L], \mathbb{H}^{m-Nj}([z_k, z_{k+1}], \mathbb{C})), \tag{3.15}$$

with uniformly bounded norms (i.e. independent of h).

The transformation of the initial problem (2.6) into the three nested problems (3.6)–(3.7)–(3.8) does not imply approximation. As mentioned earlier, problems (3.6) and (3.8) are solved by means of Fourier Transforms (FT) and the numerical accuracy of the computations is the one of the Fast Fourier Transform (FFT) algorithm for evaluating continuous FT of functions [2, 21]. Problem (3.7) is the only one solved using an approximation scheme and therefore (up to the very small spectral error of the FFT) the error in the IP method is essentially the approximation error when solving this ODE problem by the 4th-order Runge-Kutta method, with t as a parameter. Therefore, the approximation error in the RK4-IP method at grid point z_{k+1} (neglecting the FFT computational error) is given by

$$e_{k+1} = A(z_{k+1}) - u_{k+1} = \exp(\tfrac{h}{2}\mathcal{D})(A_k^{\mathrm{ip}}(z_{k+1}) - u_{k+1}^{\mathrm{ip}}) \tag{3.16}$$

where u_{k+1}^{ip} denotes the approximate solution to problem (3.7) computed at grid point z_{k+1} by one step of the RK4 method following the approximation scheme (3.9)–(3.10). Thus, the difference $A_k^{\mathrm{ip}}(z_{k+1}) - u_{k+1}^{\mathrm{ip}}$ for one step of the RK4-IP method coincides with the local error ℓ_k of the RK4 method defined by

$$A_k^{\mathrm{ip}}(z_{k+1}) = A_k^{\mathrm{ip}}(z_k) + h\,\Phi(z_k, A_k^{\mathrm{ip}}(z_k); h) + \ell_k. \tag{3.17}$$

From (3.15) and from the assumption $m \geq 4N$, we infer that the z-derivatives of $\mathcal{G}_k(z, A_k^{\mathrm{ip}}(z))$ up to the order 5 are uniformly bounded in $\mathbb{H}^{m-4N}(\mathbb{R}, \mathbb{C})$. Hence, the standard estimates for

the RK4 method [7, 24] give for the local error the following estimate

$$\forall k \in \{0, \dots, K\}, \qquad \|\ell_k\|_{m-4N} \leq Ch^5 \tag{3.18}$$

where $C > 0$ is independent of h. Note also that the z-derivatives of $\mathcal{G}_k(z, A_k^{\mathrm{ip}}(z))$ up to the order 4 are uniformly bounded in $\mathbb{H}^{m-3N}(\mathbb{R}, \mathbb{C}) \subset \mathbb{H}^1(\mathbb{R}, \mathbb{C})$, so we have additionally the estimate

$$\forall k \in \{0, \dots, K\}, \qquad \|\ell_k\|_1 \leq Ch^4. \tag{3.19}$$

This estimate will be useful in the case $m = 4N$ for the $\mathbb{L}^2$ error bound, since in this case Lemma 3.2 only provides a $\mathbb{H}^1$-conditional $\mathbb{L}^2$-stability property: as in [30], lower-order convergence (here, third-order) is used to provide a global $\mathbb{H}^1$ bound of the numerical solution and obtain fourth-order error in the $\mathbb{L}^2$ norm.

In order to estimate the global error for the RK4-IP method, we first need a stability estimate for the increment function Φ. The proof of the following lemma is straightforward from the local Lipschitz condition satisfied by $\mathcal{G}_k$.

LEMMA 3.3. *Let $\omega_0 = +\infty$, $f_R = 0$, $M > 0$ and $m \geq 4N$. Denote $m^* = \max(1, m-4N)$. Then there exists $\Lambda > 0$ such that the increment function Φ defined in* (3.10) *satisfies the following stability estimates: $\forall A, B \in B_{m^*,2M}$ and $\forall h < 1/(2\Lambda)$,*

$$\|\Phi(z, A; h) - \Phi(z, B; h)\|_\sigma \leq \Lambda\big(1 + \tfrac{1}{2}h\Lambda + \tfrac{1}{6}h^2\Lambda^2 + \tfrac{1}{24}h^3\Lambda^3\big)\, \|A - B\|_\sigma.$$

for $\sigma = 1$ and for $\sigma = m - 4N$.

In this Lemma, one can take $\Lambda = \max(\Lambda_{1,4M}, \Lambda_{m-4N,4M})$, where $\Lambda_{1,4M}$ and $\Lambda_{m-4N,4M}$ are the Lipschitz constants given by Lemma 2.3.

We also need the following lemma (its proof is straightforward by mathematical induction).

LEMMA 3.4. *Let $(\theta_k)_{k\in\mathbb{N}}$ and $(\varepsilon_k)_{k\in\mathbb{N}}$ be two non-negative sequences of real numbers and let h and λ be two non-negative real numbers such that $\forall k \in \mathbb{N}$, $\theta_{k+1} \leq (1+h\lambda)\theta_k + \varepsilon_k$. Then, for all $k \in \mathbb{N}^*$,*

$$\theta_k \leq \mathrm{e}^{kh\lambda}\, \theta_0 + \sum_{j=0}^{k-1} \mathrm{e}^{(k-1-j)h\lambda} \varepsilon_j.$$

We are now in position to prove our main result concerning the error analysis in the RK4-IP method.

THEOREM 3.5. *Let $a_0 \in \mathbb{H}^m(\mathbb{R}, \mathbb{C})$ with $m \geq 4N$, let $A \in E_{m,N}([0, Z[)$ be the maximal solution to* (1.2) *given by Theorem 2.1 and let $L \in]0, Z[$. Consider a constant step-size subdivision of interval $[0, L]$ into K subintervals by the points $z_0, \dots, z_K$ arranged in increasing order. Let us denote by $(u_k)_{k=0\dots,K}$ the sequence defined in* (3.12)*, with the initialization $u_0 = a_0$. Then there exists $h_0 > 0$ and $C > 0$ such that, if $0 < h < h_0$, we have*

$$\max_{k=0,\dots,K} \|A(z_k) - u_k\|_{m-4N} \leq CLh^4. \tag{3.20}$$

Proof. In this proof, we systematically assume that the numerical solution u_k defined by the computational procedure (3.12) satisfies the bound $\|u_k\|_{m^*} \leq 2M$, for all $0 \leq k \leq K$, where $m^* = \max(1, m - 4N)$. In fact, this bound can be a posteriori checked, for h small enough, thanks to (3.5), (3.20) and the third-order estimate (3.24) in the $\mathbb{H}^1$ norm (indeed,

in the case $m = 4N$, such estimate is required since (3.20) only bounds the $\mathbb{L}^2$ norm). Hence, with the notation of Lemma 3.2, we always have

$$\forall k \in \{0, \dots, K\}, \qquad A_k(z_k) \in B_{m,M} \subset B_{m^*,2M} \quad \text{and} \quad u_k \in B_{m^*,2M} \tag{3.21}$$

so the stability estimates of Lemma 3.3 holds for the functions $\exp(\frac{h}{2}\mathcal{D})A_k(z_k)$ and $\exp(\frac{h}{2}\mathcal{D})u_k$. When the RK4-IP method is applied for solving problem (2.6) the global error at grid point z_{k+1} is given by (3.16) and from (3.9) and (3.17) we have

$$\begin{aligned} A_k^{\mathrm{ip}}(z_{k+1}) &= \exp(\tfrac{h}{2}\mathcal{D})A_k(z_k) + h\,\Phi(z_k, \exp(\tfrac{h}{2}\mathcal{D})A_k(z_k); h) + \ell_k \\ u_{k+1}^{\mathrm{ip}} &= \exp(\tfrac{h}{2}\mathcal{D})u_k + h\,\Phi(z_k, \exp(\tfrac{h}{2}\mathcal{D})u_k; h) \end{aligned}$$

so that

$$\begin{aligned} e_{k+1} = &\exp(h\mathcal{D})[A_k(z_k) - u_k] + h\exp(\tfrac{h}{2}\mathcal{D})\big[\Phi(z_k, \exp(\tfrac{h}{2}\mathcal{D})A_k(z_k); h) \\ &- \Phi(z_k, \exp(\tfrac{h}{2}\mathcal{D})u_k; h)\big] + \exp(\tfrac{h}{2}\mathcal{D})\ell_k. \end{aligned} \tag{3.22}$$

Since Φ satisfies the Lipschitz condition of Lemma 3.3, since we have the bounds (3.21) and since $\exp(\frac{h}{2}\mathcal{D})$ is an isometry on $\mathbb{H}^\sigma(\mathbb{R}, \mathbb{C})$, we successively have, for $\sigma = 1$ and $\sigma = m - 4N$,

$$\begin{aligned} \|\Phi(z_k, \exp(\tfrac{h}{2}\mathcal{D})A_k(z_k); h) - \Phi(z_k, \exp(\tfrac{h}{2}\mathcal{D})u_k; h)\|_\sigma &\le \widetilde{\Lambda}\,\|\exp(\tfrac{h}{2}\mathcal{D})A_k(z_k) - \exp(\tfrac{h}{2}\mathcal{D})u_k\|_\sigma \\ &\le \widetilde{\Lambda}\,\|A_k(z_k) - u_k\|_\sigma \end{aligned} \tag{3.23}$$

where $\widetilde{\Lambda} = \Lambda(1 + \frac{1}{2}h\Lambda + \frac{1}{6}h^2\Lambda^2 + \frac{1}{24}h^3\Lambda^3)$ and Λ is defined in Lemma 3.3. Taking the norm of (3.22), using the triangle inequality and (3.23), yields $\|e_{k+1}\|_\sigma \le (1 + h\widetilde{\Lambda})\,\|e_k\|_\sigma + \|\ell_k\|_\sigma$. From Lemma 3.4, we deduce that for all $k \in \{0, \dots, K-1\}$

$$\|e_k\|_\sigma \le \mathrm{e}^{kh\widetilde{\Lambda}}\,\|e_0\|_\sigma + \sum_{j=0}^{k-1} \mathrm{e}^{(k-1-j)h\widetilde{\Lambda}}\|\ell_j\|_\sigma \le \mathrm{e}^{L\widetilde{\Lambda}}\left(\|e_0\|_\sigma + \sum_{j=0}^{k-1}\|\ell_j\|_\sigma\right).$$

Finally, from the error bounds (3.18) and (3.19) for the local error we conclude that

$$\max_{k\in\{0,\dots,K-1\}} \|e_k\|_{m-4N} \le \mathrm{e}^{L\widetilde{\Lambda}}\left(\|e_0\|_{m-4N} + CLh^4\right)$$

and

$$\max_{k\in\{0,\dots,K-1\}} \|e_k\|_1 \le \mathrm{e}^{L\widetilde{\Lambda}}\left(\|e_0\|_1 + CLh^3\right).$$

When we assume that the initial data is not perturbed, we obtain the $\mathbb{H}^{m-4N}$ estimate (3.20) and the $\mathbb{H}^1$ estimate

$$\max_{k\in\{0,\dots,K-1\}} \|A(z_k) - u_k\|_1 \le \mathrm{e}^{L\widetilde{\Lambda}} CLh^3. \tag{3.24}$$

The proof of Theorem 3.5 is complete. □

Remark 2. The IP method itself is exact since it is based on a change of unknown. The only source of errors lies in the way the nonlinear ODE problem (3.7) is solved and the proof of the theorem can be adapted to any other method than the Runge Kutta method used here. If the nonlinear problem is solved by a numerical scheme of order n, the IP method used in conjunction with this numerical scheme will be of order n. But the benefit of the cancellation of 4 exponential operator terms will be lost as explained in section 3.3.

4. Comparison of IP and Symmetric Split-Step methods.

4.1. Theoretical comparison. As shown in Section 3, the RK4-IP method is based on the change of unknown given by relation (3.2) leading to a set of three nested PDE problems (3.6)–(3.7)–(3.8) to be solved over each subinterval introduced with the discretisation of the fiber length. This computational structure is very similar to what is obtained when solving problem (2.5) by the Split-Step method based on the Strang splitting formula also termed the Symmetric Split-Step method. Namely, using the notations introduced in Section 3.2, the Symmetric Split-Step method consists in solving over each subinterval $[z_k, z_{k+1}]$ for $k \in \{0, \dots, K-1\}$, the following three nested problems with time variable t as a parameter:

$$\forall z \in [z_k, z_{k+\frac{1}{2}}], \quad \frac{\partial}{\partial z} A_k^+(z) = \mathcal{D} A_k^+(z), \qquad A_k^+(z_k) = A_{k-1}^s(z_k), \tag{4.1}$$

where $A_{k-1}^s(z_k)$ represents the approximate solution at grid point z_k computed at step $k-1$ for $k \geq 1$, and $A_{-1}^s(z_0) = a_0$;

$$\forall z \in [z_k, z_{k+1}], \quad \frac{\partial}{\partial z} B_k(z) = \mathcal{N}(B_k)(z), \qquad B_k(z_k) = A_k^+(z_{k+\frac{1}{2}}), \tag{4.2}$$

where $A_k^+(z_{k+\frac{1}{2}}, t)$ represents the solution to problem (4.1) at half grid point $z_{k+\frac{1}{2}}$;

$$\forall z \in [z_{k+\frac{1}{2}}, z_{k+1}], \quad \frac{\partial}{\partial z} A_k^-(z) = \mathcal{D} A_k^-(z) \qquad A_k^-(z_{k+\frac{1}{2}}) = B_k(z_{k+1}), \tag{4.3}$$

where $B_k(z_{k+1}, t)$ represents the solution to problem (4.2) at node z_{k+1}. An approximate solution to problem (2.5) at grid node z_{k+1} is then given by $A_k^s(z_{k+1}) := A_k^-(z_{k+1})$. The two linear PDE problems (4.1) and (4.3) are solved according to the computational procedure outlined in Section 3.3. When considering the NLSE, an explicit solution to the nonlinear problem (4.2) is known and in this particular case the Symmetric Split-Step method is much cheaper than the IP method. In the more general case of the GNLSE, there is not anymore an explicit solution to the nonlinear ODE problem (4.2) but its solution can be approximated using standard numerical schemes for ODE. Since the Symmetric Split-Step method is second order accurate, a natural choice would be to use a second order accurate numerical scheme for ODE such as the a second order Runge-Kutta method. However, in practical situations where the nonlinear physical effects taken into account in the GNLSE are high, the use of a second order accurate numerical scheme for solving problem (4.2) could put the global accuracy of the Symmetric Split-Step method at a disadvantage [4]. Therefore, in the literature in optics (see e.g. [1, 32]) it is common to combine the Symmetric Split-Step method with the classical fourth order Runge-Kutta method, referred as the SSF-RK4 method. It should be noted that the use of a fourth-order accurate Split-Step scheme with the classical fourth order Runge-Kutta method would increase the accuracy at the cost of increasing by a factor 3 the computational time of the method compared to the SSF method [31]. Moreover, numerical results in [31] show that the use of the Symmetric Split-Step scheme is a good compromise between accuracy and computational cost. Beyond these practical aspects, in the context of this paper the comparison of the RK4-IP method with SSF-RK4 method is dictated by the close form of the algorithm of these 2 methods.

Making use of the notations introduced in Section 3.2 the SSF-RK4 method at step k for

$k = 0, \dots, K-1$ reads

$$\begin{aligned}
\widehat{u}_{k+\frac{1}{2}} &= \mathrm{e}^{\widehat{d}_\nu \frac{h}{2}} \times \widehat{u}_k, \\
\widehat{\alpha}_1 &= \widehat{\mathcal{N}}(\mathcal{F}^{-1}(\widehat{u}_{k+\frac{1}{2}})), & \widehat{\alpha}_2 &= \widehat{\mathcal{N}}\big(\mathcal{F}^{-1}(\widehat{u}_{k+\frac{1}{2}} + \tfrac{h}{2}\widehat{\alpha}_1)\big), \\
\widehat{\alpha}_3 &= \widehat{\mathcal{N}}\big(\mathcal{F}^{-1}(\widehat{u}_{k+\frac{1}{2}} + \tfrac{h}{2}\widehat{\alpha}_2)\big), & \widehat{\alpha}_4 &= \widehat{\mathcal{N}}\big(\mathcal{F}^{-1}(\widehat{u}_{k+\frac{1}{2}} + h\widehat{\alpha}_3)\big), \\
\widehat{u}_{k+1} &= \mathrm{e}^{\widehat{d}_\nu \frac{h}{2}} \times \big(\widehat{u}_{k+\frac{1}{2}} + \tfrac{h}{6}\left(\widehat{\alpha}_1 + 2\widehat{\alpha}_2 + 2\widehat{\alpha}_3 + \widehat{\alpha}_4\right)\big), \\
u_{k+1} &= \mathcal{F}^{-1}\big(\widehat{u}_{k+1}\big).
\end{aligned}$$

Since the evaluation of the nonlinear operator $\widehat{\mathcal{N}}$ requires 4 FFT, the above computational scheme involves 17 FFT per step. Its cost is therefore the same as the cost of the computational scheme (3.14) used for solving the GNLSE (1.1) by the RK4-IP method.

From our presentation of both the RK4-IP and SSF methods, a formal comparison of the two methods is straightforward. Over one subinterval $[z_k, z_{k+1}]$ the three nested problems (3.6) – (3.7) – (3.8) are solved when the RK4-IP method is used whereas with the SSF-RK4 method the three nested problems (4.1) – (4.2) – (4.3) are solved. Since problems (3.6) and (4.1) are the same as well as problems (3.8) and (4.3), the difference between the two computational methods lies in problems (3.7) and (4.2). Both are solved here using the same 4th-order RK method. The splitting (3.6) – (3.7) – (3.8) is exact since it originates from the change of unknown (3.2) whereas the splitting (4.1) – (4.2) – (4.3) deduced from the Strang formula is second-order accurate. To be comprehensive, the relationship between the solution to problems (3.7) and (4.2) is given in the following proposition.

PROPOSITION 4.1. *Let $\omega_0 = +\infty$, $f_R = 0$. For all $k \in \{0, \dots, K-1\}$ let B_k denote the solution to problem* (4.2) *and A_k^{ip} denote the solution to problem* (3.7) *with the same initial data $A_k^+(z_{k+\frac{1}{2}})$ assumed to belong to $\mathbb{H}^m(\mathbb{R}, \mathbb{C})$ with $m \geq 2N$. Then, for all $k \in \{0, \dots, K-1\}$ we have the following estimate in $\mathbb{H}^{m-2N}(\mathbb{R}, \mathbb{C})$:*

$$B_k(z_{k+1}) = A_k^{\mathrm{ip}}(z_{k+1}) + \mathcal{O}(h^3). \tag{4.4}$$

Proof. For simplicity, denote $\tilde{A}_0 = A_k^+(z_{k+\frac{1}{2}})$. We remark that (4.2) and (3.7) take respectively the forms

$$\frac{\partial}{\partial z} B(z) = \mathcal{N}_0(B(z)) \qquad \text{and} \qquad \frac{\partial}{\partial z} A_k^{\mathrm{ip}}(z) = \mathcal{G}_k(z, A_k^{\mathrm{ip}}(z)) \tag{4.5}$$

with the same initial data $B_k(z_k) = \tilde{A}_0$ and $A_k^{\mathrm{ip}}(z_k) = \tilde{A}_0$. Writing the second-order Runge-Kutta scheme for both problems (4.5) yields the following standard estimate [24] in the $\mathbb{H}^{m-2N}$ norm (note indeed that this estimate involves two z-derivatives of $\mathcal{G}_k$, thus a loss of $2N$ derivatives in t),

$$B_k(z_{k+1}) = \tilde{A}_0 + h\,\mathcal{N}_0\left(\tilde{A}_0 + \frac{h}{2}\mathcal{N}_0(\tilde{A}_0)\right) + \mathcal{O}(h^3), \tag{4.6}$$

$$\begin{aligned}
A_k^{\mathrm{ip}}(z_{k+1}) &= \tilde{A}_0 + h\,\mathcal{G}_k\left(z_{k+\frac{1}{2}}, \tilde{A}_0 + \frac{h}{2}\mathcal{G}_k(z_k, \tilde{A}_0)\right) + \mathcal{O}(h^3) \\
&= \tilde{A}_0 + h\,\mathcal{N}_0\left(\tilde{A}_0 + \frac{h}{2}\mathcal{G}_k(z_k, \tilde{A}_0)\right) + \mathcal{O}(h^3)
\end{aligned} \tag{4.7}$$

where we used the key property $\mathcal{G}_k(z_{k+\frac{1}{2}}, v) = \mathcal{N}_0(v)$. Hence, since we have $\mathcal{G}_k(z_k, \tilde{A}_0) = \mathcal{N}_0(\tilde{A}_0) + \mathcal{O}(h)$, we deduce (4.4) from (4.6) and (4.7). □

As problems (3.6) and (4.1) are identical as well as problems (3.8) and (4.3), we deduce from Proposition 4.1 the following corollary.

COROLLARY 4.2. *Consider a subdivision of interval* $[0, L]$ *into* K *subintervals by the points* $z_0, \dots, z_K$ *arranged in increasing order. For all* $k \in \{0, \dots, K-1\}$ *denote by* A_k^{ip} *the solution to problem* (3.6)–(3.7)–(3.8) *over the subinterval* $[z_k, z_{k+1}]$ *and denote by* B_k *the approximation of the solution to problem* (2.6) *over the subinterval* $[z_k, z_{k+1}]$ *computed by solving the three nested problems* (4.1)–(4.2)–(4.3). *Then*

$$\sup_{k\in\{0,\dots,K-1\}} \|A_k^{\mathrm{ip}}(z_{k+1}) - B_k(z_{k+1})\|_{m-2N} = \mathcal{O}(h^2).$$

Remark 3. The convergence of the Split-Step methods applied to various forms of the Schrödinger equation is widely studied in the literature, see e.g. [6, 30, 33] where the authors prove that the convergence order of the SSF method is 2. As the change of unknown in the IP method does not imply approximation before discretization, the second order convergence of the SSF method can alternatively be deduced from Corollary 4.2.

4.2. Numerical comparison. As mentioned in the introduction, the IP method is aimed at solving the GNLSE. It is only due to difficulties related to the properties of the nonlinear operator $\mathcal{N}$ that we have restricted the mathematical analysis of the IP method achieved in section 3 to the simpler case when $\mathcal{N} = \mathcal{N}_0$ (i.e. to the NLSE). Thus, the numerical comparison of the IP method to the SSF method will be achieved on the GNLSE.

4.2.1. Numerical comparison of convergence rates. In order to illustrate the result of Theorem 3.5 we have solved the GNLSE (1.1) by the RK4-IP and SSF-RK4 methods with constant step-size values divided by 2 from one execution to the other. We have obtained the experimental convergence curves depicted in Figure 4.1 where the evolution of the relative quadratic error $\|A_{\mathrm{ref}}(L) - A_{K-1}(L)\|_0 / \|A_{\mathrm{ref}}(L)\|_0$ versus the step-size h is plotted (A_{ref} denotes a reference solution computed with a very small step-size).

The physical values used for the computation are the following: $\omega_0 = 1770\,\mathrm{Thz}$, $\gamma = 4.3\,\mathrm{W}^{-1}\mathrm{km}^{-1}$, $\beta_2 = 19.83\,\mathrm{ps}^2\mathrm{km}^{-1}$, $\beta_3 = 0.031\,\mathrm{ps}^3\mathrm{km}^{-1}$ and $\beta_n = 0$ for $n \geq 4$, $\alpha = 0.046\,\mathrm{km}^{-1}$, $L = 100\,\mathrm{m}$, $f_R = 0.245$. The source term is $a_0 : t \mapsto \sqrt{P_0}\,\mathrm{e}^{-\frac{1}{2}(t/T_0)^2}$ where $T_0 = 2.8365\,\mathrm{ps}$ is the pulse half-width and $P_0 = 1\,\mathrm{W}$ is the pulse peak power.

This experimental result is in a good agreement with the theoretical convergence behavior predicted by Theorem 3.5 for the RK4-IP method (convergence order 4) and by Remark 3 for the SSF-RK4 method (convergence order 2).

We would like to mention that in other simulations where the nonlinear parameter γ and the power of the source term were larger, we have obtained an experimental convergence order for the SSF-RK4 method close to 4. This can be easily explained as follows: when the "weight" of the nonlinear ODE in Strang splitting (4.1)–(4.2)–(4.3) is much larger than the "weight" of the linear PDE problems (4.1) and (4.3), the error due to the splitting formula (asymptotically in $\mathcal{O}(h^2)$) can be much lower than the error of the RK4 formula (asymptotically in $\mathcal{O}(h^4)$) which dominates. Therefore, over the range of values for the step-size h used to draw the convergence curve, the observed convergence rate is close to 4. Of course, if computations with much smaller step-size values were achieved, a convergence curve with slope 2 would finally be obtained (unfortunately in practice the error reaches the numerical precision quicker than this). We refer to [4] for details on this phenomenon including numerical simulations. This behavior of the numerical error explains why in optics the Symmetric Split-Step method is widely used in conjunction with a RK4 scheme for solving the nonlinear problem (4.2) rather than with a RK2 scheme. It improves the accuracy of the results when nonlinear phenomena predominate during the propagation of the optical wave in a fiber. As a consequence, when

non linear effects predominate in the GNLSE, the RK4-IP and SSF-RK4 methods give very similar results with respect to accuracy and computation time.

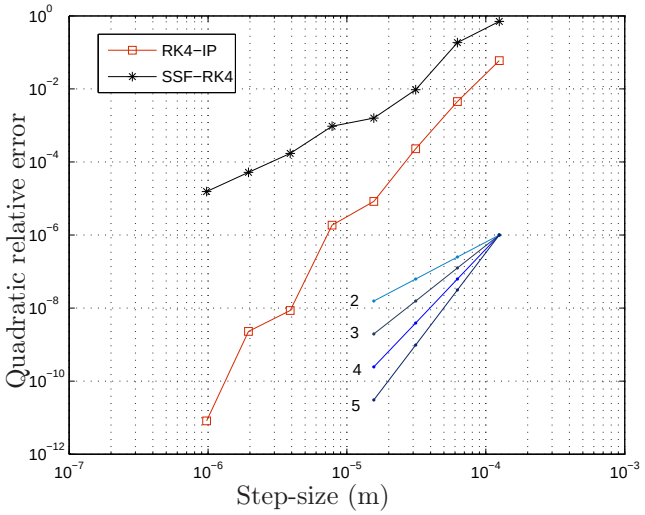

FIG. 4.1. *Experimental convergence curves for the RK4-IP and SSF-RK4 methods.*

4.2.2. Numerical comparison for simulation of wave propagation in optical fibers. In order to numerically compare the RK4-IP method to the SSF-RK4 method, we have solved the GNLSE on a test example chosen to match with a typical case of high speed data propagation through a $L = 20\,\mathrm{km}$ single mode fibre in optical telecommunication with a data's carrier frequency located in the C band of the infrared spectrum ($f_0 = 193\,\mathrm{Thz}$). The following set of fibre's parameters were used for the simulation: $\alpha = 0.046\,\mathrm{km}^{-1}$, $\gamma = 4.3\,\mathrm{W}^{-1}\mathrm{km}^{-1}$, $f_R = 0.245$, $\beta_2 = -19.83\,\mathrm{ps}^2\mathrm{km}^{-1}$, $\beta_3 = 0.031\,\mathrm{ps}^3\mathrm{km}^{-1}$ and $\beta_n = 0$ for $n \geq 4$. The source term $a_0 = A(z = 0)$ was represented as a Gaussian pulse: $a_0 : t \mapsto \sqrt{P_0}\,\mathrm{e}^{-\frac{1}{2}(t/T_0)^2}$ where T_0 is the pulse half-width at 1/e intensity point and P_0 is the pulse peak power. Simulations were carried out for a pulse-width $T_0 = 6.8\,\mathrm{ps}$ and for a peak power $P_0 = 5\,\mathrm{mW}$.

Both algorithms were tested on a Intel Core i5-4200M with 8Go RAM. The CPU time and relative quadratic error $\|A(L) - A_{K-1}(L)\|_0/\|A(L)\|_0$ are compared in Table 4.1 for a step-size of 100m and 10m. For a step-size of 100m the RK4-IP method gives an error of $1.4957\,10^{-9}$ in 1.42s but to obtain an equivalent precision with the SSF-RK4 method it is necessary to take a step-size of 2.5m leading to a CPU time of 70.17s (which is fifty times more). For a more complete experimental comparison between the RK4-IP method and the SSF-RK4 method for solving the GNLSE in optics we refer to [27]. We have presented above results obtained with a constant step-size; using an embedded Runge-Kutta scheme, local error estimates can be computed and used for adaptive step-size purposes in the IP method, see [5, 3].

5. Conclusion. We have presented an alternative method to the Split-Step approach for solving the GNLSE in optics. The Interaction Picture (IP) method has a form very similar to the one of the Symmetric Split-Step method. However it is based on a change of unknown rather than on a splitting formula and therefore does not contain any approximation at this stage. Actually, the error in the IP method is in the use of an approximation scheme for solving the nonlinear ODE problem resulting from the change of unknown. We have carried

TABLE 4.1
Comparison of the RK4-IP and SSF-RK4 methods for solving the GNLSE on a test problem.

Method	Step-size (m)	CPU time (s)	Relative quadratic error
RK4-IP	100	1.42	$1.4957\,10^{-9}$
SSF-RK4	100	1.48	$2.5582\,10^{-6}$
RK4-IP	10	13.85	$4.6192\,10^{-13}$
SSF-RK4	10	14.49	$2.555\,10^{-8}$
SSF-RK4	2.5	70.17	$1.5968\,10^{-9}$

out a theoretical and experimental study of the IP method and we have compared it to the Symmetric Split-Step method. It is worth mentioning that the IP method can be used for a larger number of PDEs than only the GNLSE; actually it is suitable to solve all PDEs where Split-Step methods are generally used.

Appendix A. Proof of Theorem 2.1. We analyze the well-posedness of the Cauchy problem (1.2) (formulated under the form (2.6)). For simplicity, we restrict the solution to $z \geq 0$. Since the equation is time reversible, this result can be easily extended to $z \leq 0$. We proceed into 3 steps.

Step 1. (Local well-posedness of the Cauchy problem.) Let $a_0 \in \mathbb{H}^m(\mathbb{R},\mathbb{C})$, with $m \geq 1$. To prove the local existence of a unique solution A to (2.6), we first transform it into an ODE in infinite dimension, with new unknown $V = \mathrm{e}^{-z\mathcal{D}}A$, solution of

$$\forall z \in \mathbb{R}^+, \quad \frac{\partial}{\partial z}V(z) = \mathcal{F}(z,V(z)), \qquad V(0) = a_0,$$

where $\mathcal{F}(z,V) = \exp(-z\mathcal{D})\mathcal{N}_0(\exp(z\mathcal{D})V)$. From Lemma 2.2 and Lemma 2.3, it can be directly deduced that $\mathcal{F}$ is locally Lipschitz continuous on $\mathbb{H}^m(\mathbb{R},\mathbb{C})$. More precisely, for all $M > 0$, there exists a constant $C_{m,M}$ such that, if $\|u\|_m \leq M$ and $\|v\|_m \leq M$, then

$$\forall z \in \mathbb{R}^+, \qquad \|\mathcal{F}(z,u) - \mathcal{F}(z,v)\|_m \leq C_{m,M}\|u-v\|_m.$$

The mapping $\mathcal{F}$ being also continuous on $\mathbb{R} \times \mathbb{H}^m(\mathbb{R},\mathbb{C})$, the Cauchy-Lipschitz theorem in Banach spaces gives the local existence of a unique maximal solution $V \in \mathcal{C}^1([0,Z[,\mathbb{H}^m(\mathbb{R},\mathbb{C}))$ such that if $Z < +\infty$ then $\limsup_{z\to Z}\|V(z)\|_m = +\infty$. Coming back to A, one has $A \in \mathcal{C}^0([0,Z[,\mathbb{H}^m(\mathbb{R},\mathbb{C}))$ and, by differentiating (2.6), one gets $A \in E_{m,N}([0,Z[)$. Moreover, since $\|A(z)\|_m = \|V(z)\|_m$, one has clearly

$$\text{if} \quad Z < +\infty \quad \text{then} \quad \limsup_{z\to Z}\|A(z)\|_m = +\infty. \tag{A.1}$$

In fact, the condition (A.1) can be replaced by (2.3). To prove this, we recall the following classical *tame estimate*, see [11]: for all $M > 0$, there exists $C_{m,M} > 0$ such that, for all $u \in \mathbb{H}^m(\mathbb{R},\mathbb{C})$ with $\|u\|_{\mathbb{L}^\infty} \leq M$, one has

$$\|\mathcal{N}_0(u)\|_m \leq C_{m,M}\|u\|_m. \tag{A.2}$$

Let us prove (2.3) by contradiction. We assume that $Z < +\infty$ with $M = \|A\|_{\mathbb{L}^\infty((0,Z)\times\mathbb{R})} < +\infty$. From (A.2) and from the Duhamel formula

$$A(z) = \exp(z\mathcal{D})a_0 + \int_0^z \exp((z-\zeta)\mathcal{D})\mathcal{N}_0(A(\zeta))\,\mathrm{d}\zeta, \tag{A.3}$$

one gets

$$\|A(z)\|_m \le \|a_0\|_m + \int_0^z \|\mathcal{N}_0(A(\zeta))\|_m \,\mathrm{d}\zeta \le \|a_0\|_m + C_{m,M}\int_0^z \|A(\zeta)\|_m \,\mathrm{d}\zeta.$$

Hence, by the Gronwall lemma, $\forall z \in [0, Z[$, $\|A(z)\|_m \le e^{ZC_{m,M}}\|a_0\|_m$, which implies $\limsup_{z\to Z}\|A(z)\|_m < +\infty$, and then, by (A.1), $Z = +\infty$, which is a contradiction. The proof of (2.3) is complete.

Step 2. ($\mathbb{L}^2$ estimate.) Introduce the new unknown $U(z) = \mathrm{e}^{\frac{\alpha}{2}z}A(z)$, satisfying

$$\forall z \in \mathbb{R}^+, \quad \frac{\partial}{\partial z}U(z) = \mathcal{D}U(z) + \widetilde{\mathcal{N}}_0(U)(z), \qquad U(0) = a_0, \tag{A.4}$$

where $\widetilde{\mathcal{N}}_0(U)(z) = \mathrm{i}\gamma\mathrm{e}^{-\alpha z}U(z)|U(z)|^2$. Multiplying (A.4) by $\mathrm{i}\,\overline{U}$ and integrating with respect to t, we obtain for the imaginary part

$$\Im\left(\mathrm{i}\int_{\mathbb{R}}\overline{U}(z)\,\partial_z U(z)\,\mathrm{d}t\right) = \frac{1}{2}\frac{d}{dz}\|U(z)\|_0^2 = 0$$

because $\Im(\mathrm{i}\partial_z U(z)\,\overline{U}(z)) = \Re(\partial_z U(z)\,\overline{U}(z)) = \frac{1}{2}\partial_z(U(z)\,\overline{U}(z))$. It follows that

$$\forall z \in [0, Z[, \qquad \|U(z)\|_0 = \|a_0\|_0 \tag{A.5}$$

and therefore $\|A(z)\|_0 = \mathrm{e}^{-\frac{\alpha}{2}z}\|a_0\|_0$, which is (2.2).

Step 3. (A priori bound in $\mathbb{H}^N(\mathbb{R},\mathbb{C})$ and global existence for $N = 2P$.) From now on, we assume that $N = 2P$, with $P \in \mathbb{N}^*$, and that $m \ge P$. To prove that $Z = +\infty$, by (2.3) it suffices to obtain an a priori estimate on the $\mathbb{H}^P(\mathbb{R},\mathbb{C})$ norm of $U(z)$, which will imply an $\mathbb{L}^\infty$ estimate by Sobolev embeddings.

To prove that $\|U(z)\|_P$ is bounded, we derive a second conservation law for (A.4). Multiply (A.4) by $\mathrm{i}\partial_z\overline{U}(z)$ and integrate with respect to t. The real part reads, $\forall z \in [0, Z[$,

$$\Re\left(-\sum_{n=2}^{2P}\mathrm{i}^n\frac{\beta_n}{n!}\int_{\mathbb{R}}\partial_t^n U(z)\,\partial_z\overline{U}(z)\,\mathrm{d}t - \gamma\mathrm{e}^{-\alpha z}\int_{\mathbb{R}}|U(z)|^2\,U(z)\,\partial_z\overline{U}(z)\,\mathrm{d}t\right) = 0. \tag{A.6}$$

We set $I_n = -\int_{\mathbb{R}}\Re\left(\mathrm{i}^n\frac{\beta_n}{n!}\partial_t^n U(z)\,\partial_z\overline{U}(z)\right)\mathrm{d}t$. Using integrations by parts, respectively for $n = 2j$ and $n = 2j+1$, we obtain

$$I_{2j} = -\frac{\beta_{2j}}{2(2j)!}\frac{d}{dz}\int_{\mathbb{R}}|\partial_t^j U(z)|^2\,\mathrm{d}t, \qquad I_{2j+1} = -\frac{\mathrm{i}\beta_{2j+1}}{2(2j+1)!}\,\frac{d}{dz}\int_{\mathbb{R}}\partial_t^j\overline{U}(z)\,\partial_t^{j+1}U(z)\,\mathrm{d}t.$$

The last part of equation (A.6) is rewritten as

$$\int_{\mathbb{R}}|U(z)|^2\,U(z)\,\partial_z\overline{U}(z)\,\mathrm{e}^{-\alpha z}\,\mathrm{d}t = \frac{1}{4}\,\frac{d}{dz}\int_{\mathbb{R}}|U(z)|^4\,\mathrm{e}^{-\alpha z}\,\mathrm{d}t + \frac{\alpha}{4}\int_{\mathbb{R}}|U(z)|^4\,\mathrm{e}^{-\alpha z}\,\mathrm{d}t.$$

Equation (A.6) then reads

$$-\frac{\partial}{\partial z}\mathcal{B}(z) - \frac{\gamma\alpha}{2}\int_{\mathbb{R}}|U(z)|^4\,\mathrm{e}^{-\alpha z}\,\mathrm{d}t = 0,$$

where

$$\mathcal{B}(z) = \sum_{j=1}^{P-1} \left(\frac{\beta_{2j}}{(2j)!} \int_{\mathbb{R}} |\partial_t^j U(z)|^2 \, \mathrm{d}t + \mathrm{i} \frac{\beta_{2j+1}}{(2j+1)!} \int_{\mathbb{R}} \partial_t^{j+1} U(z) \, \partial_t^j \overline{U}(z) \, \mathrm{d}t \right) + \frac{\beta_{2P}}{(2P)!} \int_{\mathbb{R}} |\partial_t^P U(z)|^2 \, \mathrm{d}t + \frac{\gamma}{2} \int_{\mathbb{R}} |U(z)|^4 \, \mathrm{e}^{-\alpha z} \, \mathrm{d}t. \tag{A.7}$$

It follows that

$$\mathcal{B}(z) = \mathcal{B}(0) - \frac{\gamma\alpha}{2} \int_0^z \int_{\mathbb{R}} |U(\zeta)|^4 \, \mathrm{e}^{-\alpha\zeta} \, \mathrm{d}t \, \mathrm{d}\zeta. \tag{A.8}$$

In particular, if the attenuation/gain coefficient α vanishes, then $\mathcal{B}$ is independent of z. Moreover, this identity implies that the mapping $z \in \mathbb{R}^+ \mapsto \mathcal{B}(z) \in \mathbb{R}$ is decreasing in the case $\alpha \geq 0$ and $\gamma \geq 0$.

Let us now derive from (A.8) an a priori bound of U in $\mathbb{H}^P(\mathbb{R}, \mathbb{C})$, showing that $\mathcal{B}(z)$ is greater than a quantity depending on $\|\partial_t^P U(z)\|_0^2$. Using Gagliardo-Nirenberg inequality and Young inequality, we get for all $\varepsilon > 0$ and for all $(p_j, p_j') \in \mathbb{R}^+ \times \mathbb{R}^+$ such that $\frac{1}{p_j} + \frac{1}{p_j'} = 1$,

$$\left| \int_{\mathbb{R}} \partial_t^j \overline{U}(z) \, \partial_t^{j+1} U(z) \, \mathrm{d}t \right| \leq \frac{1}{p_j} \varepsilon^{p_j} \, \|\partial_t^P U(z)\|_0^{\frac{2j+1}{P} p_j} + \frac{1}{p_j' \varepsilon^{p_j'}} \, \|U(z)\|_0^{(2 - \frac{2j+1}{P}) p_j'}$$

which becomes, for $p_j = \frac{2P}{2j+1}$ and $p_j' = \frac{2P}{2P-2j-1}$,

$$\left| \int_{\mathbb{R}} \partial_t^j \overline{U}(z) \, \partial_t^{j+1} U(z) \, \mathrm{d}t \right| \leq \frac{1}{p_j} \varepsilon^{p_j} \|\partial_t^P U(z)\|_0^2 + \frac{1}{p_j' \varepsilon^{p_j'}} \|U(z)\|_0^2.$$

In the same way, with $q_j = \frac{P}{j}$ and $q_j' = \frac{P}{P-j}$ we obtain for all $\varepsilon > 0$

$$\int_{\mathbb{R}} \left| \partial_t^j U(z) \right|^2 \, \mathrm{d}t \leq \frac{1}{q_j} \varepsilon^{q_j} \, \|\partial_t^P U(z)\|_0^2 + \frac{1}{q_j' \varepsilon^{q_j'}} \, \|U(z)\|_0^2.$$

Remark that, for all $j \leq P-1$, we have $1 < q_j, q_j', p_j, p_j' < +\infty$. Without loss of generality, we assume that $\beta_{2P} > 0$ (otherwise, change z to $-z$) and we choose ε small enough such that

$$\sum_{j=1}^{P-1} \left(\frac{|\beta_{2j}|}{(2j)!} \frac{1}{q_j} \varepsilon^{q_j} + \frac{|\beta_{2j+1}|}{(2j+1)!} \frac{1}{p_j} \varepsilon^{p_j} \right) \leq \frac{1}{2} \frac{\beta_{2P}}{(2P)!}.$$

Then, from (A.7) and the previous inequalities, we deduce that

$$\mathcal{B}(z) \geq \frac{1}{2} \frac{\beta_{2P}}{(2P)!} \|\partial_t^P U(z)\|_0^2 - C_1 \|U(z)\|_0^2 + \frac{\gamma}{2} \int_{\mathbb{R}} |U(z)|^4 \, \mathrm{e}^{-\alpha z} \, \mathrm{d}t$$

where $C_1 = \sum_{j=1}^{P-1} \left(\frac{|\beta_{2j}|}{(2j)!} \frac{1}{q_j' \varepsilon^{q_j'}} + \frac{|\beta_{2j+1}|}{(2j+1)!} \frac{1}{p_j' \varepsilon^{p_j'}} \right)$. Setting $C_2 = \frac{2(2P)!}{\beta_{2P}}$, it follows from (A.8) that

$$\begin{aligned} \|\partial_t^P U(z)\|_0^2 \leq & C_2 \mathcal{B}(0) + C_1 C_2 \|U(z)\|_0^2 \\ & - C_2 \frac{\gamma}{2} \left(\int_{\mathbb{R}} |U(z)|^4 \, \mathrm{e}^{-\alpha z} \, \mathrm{d}t + \alpha \int_0^z \int_{\mathbb{R}} |U(\zeta)|^4 \, \mathrm{e}^{-\alpha\zeta} \, \mathrm{d}t \, \mathrm{d}\zeta \right). \end{aligned} \tag{A.9}$$

Next, the growth of $\|\partial_t^P U(z)\|_0$ can be controlled for all $z \in [0, Z[$ thanks to the Gagliardo-Nirenberg inequality

$$\int_{\mathbb{R}} |U(z)|^4 \mathrm{e}^{-\alpha z}\ \mathrm{d}t \leq \int_{\mathbb{R}} |U(z)|^4\ \mathrm{d}t \leq C\|U(z)\|_0^{4-\frac{1}{P}} \|\partial_t^P U(z)\|_0^{\frac{1}{P}}.$$

From (A.9), we get

$$\begin{aligned}\|\partial_t^P U(z)\|_0^2 \leq & C_2\mathcal{B}(0) + C_1C_2\|U(z)\|^2 \\ & + C_3\left(\|U(z)\|_0^{4-\frac{1}{P}}\|\partial_t^P U(z)\|_0^{\frac{2}{P}} + \int_0^z \|U(\zeta)\|_0^{4-\frac{1}{P}}\|\partial_t^P U(\zeta)\|_0^{\frac{2}{P}}\ \mathrm{d}\zeta\right) \\ \leq & C_4 + C_5\left(\|\partial_t^P U(z)\|_0^{\frac{2}{P}} + \int_0^z \|\partial_t^P U(\zeta)\|_0^{\frac{2}{P}}\ \mathrm{d}\zeta\right)\end{aligned}$$

where C_3, C_4 and C_5 are positive constants. The following version of Gronwall's lemma implies that $\|\partial_t^P U(z)\|_0$ is bounded on every finite interval $[0, L[\subset [0, Z[$ and then, that $\|U(z)\|_P$ is bounded on the same interval. This is enough to conclude that $Z = +\infty$ (if $Z < +\infty$, take $L = Z$). The proof of Theorem 2.1 is complete.

LEMMA A.1. *Let $a, b > 0$, m be a positive integer and y be a positive function with regularity $\mathcal{C}^1$ satisfying $y(t) \leq a + by(t)^{\frac{1}{m}} + b\int_0^t y(s)^{\frac{1}{m}}\ \mathrm{d}s$. Then $\forall t \in [0, T]$,*

$$y(t) \leq \left((a + b\,y(0))^{1-\frac{1}{m}} + \frac{(m-1)}{m}\,b\,t + \frac{b(m-1)}{m^2}\,\ln\left(\frac{y(t)}{y(0)}\right)\right)^{m/(m-1)}.$$

Appendix B. The RK4-IP algorithm for solving the GNLSE.

Input: Array u contains the input signal amplitude sampled over the time window.
Array $[\nu_j]_{j=1,\ldots,J}$ contains the frequency sampling points.
Array $[z_k]_{k=0,\ldots,K}$ contains the spatial grid points.
Array $\widehat{h}_R$ containing the sampling of the FT of the Raman response function.
{Initialisation}
for $j = 1, \ldots, J$ **do**
  $\widehat{d}[j] \leftarrow \mathrm{i}\sum_{n=2}^{N} \frac{\beta_n}{n!}(2\pi\nu_j)^n$
  $tfexpd[j] \leftarrow \exp(\frac{h}{2}\widehat{d}[j])$
end for
$\widehat{u}_1 \leftarrow \mathrm{FFT}(u, forward)$
{Loop over the propagation subinterval}
for $k = 1, \ldots, K$ **do**
  for $j = 1, \ldots, J$ **do**
    $\widehat{u}_{ip}[j] \leftarrow tfexpd[j] \times \widehat{u}_1[j]$
  end for
  $\widehat{\alpha}_1 \leftarrow \mathrm{COMPUTE\_TFN}(u, \widehat{u}_1)$
  for $j = 1, \ldots, J$ **do**
    $\widehat{\alpha}_1[j] \leftarrow tfexpd[j] \times \widehat{\alpha}_1[j]$
    $\widehat{u}_2[j] \leftarrow \widehat{u}_{ip}[j] + \frac{h}{2}\widehat{\alpha}_1[j]$
  end for
  $u_2 \leftarrow \mathrm{FFT}(\widehat{u}_2, backward)$
  $\widehat{\alpha}_2 \leftarrow \mathrm{COMPUTE\_TFN}(u_2, \widehat{u}_2)$
  for $j = 1, \ldots, J$ **do**
    $\widehat{u}_3[j] \leftarrow \widehat{u}_{ip}[j] + \frac{h}{2}\widehat{\alpha}_2[j]$
  end for
  $u_3 \leftarrow \mathrm{FFT}(\widehat{u}_3, backward)$

$\widehat{\alpha}_3 \leftarrow$ COMPUTE_TFN$(u_3, \widehat{u}_3)$
for $j = 1, \ldots, J$ **do**
$\widehat{u}_4[j] \leftarrow tfexpd[j] \times (\widehat{u}_{ip}[j] + h\widehat{\alpha}_3[j])$
end for
$u_4 \leftarrow$ FFT$(\widehat{u}_4, backward)$
$\widehat{\alpha}_4 \leftarrow$ COMPUTE_TFN$(u_4, \widehat{u}_4)$
for $j = 1, \ldots, J$ **do**
$\widehat{u}_1[j] \leftarrow tfexpd[j] \times (\widehat{u}_{ip}[j] + \frac{h}{6}\widehat{\alpha}_1[j] + \frac{h}{3}\widehat{\alpha}_2[j] + \frac{h}{3}\widehat{\alpha}_3[j]) + \frac{h}{6}\widehat{\alpha}_4[j]$
end for
$u \leftarrow$ FFT$(\widehat{u}_1, backward)$ {Array u contains the values $[A_k(z_{k+1}, t_j)]_{j=1,\ldots,J}$ the sampling of the signal amplitude at step z_k}
end for

FFT$(u, forward)$ stands for a call to the Fast Fourier Transform (FFT) algorithm to compute the Discrete Fourier Transform (DFT) of array u, FFT$(u, backward)$ stands for a call to FFT algorithm to compute the inverse DFT of array u, and COMPUTE_TFN refers to the following function.

FUNCTION $\widehat{g}$ = COMPUTE_TFN(f, $\widehat{f}$)
{Compute the Fourier Transform of $g : t \mapsto \mathcal{N}(f)(z,t)$ for a given z}
Input: Array f contains the time sampling of function f for the given z.
Array $\widehat{f}$ contains the sampled FT of f.
Array $[\nu_j]_{j=1,\ldots,J}$ contains the frequency sampling points.
Array $\widehat{h}_R$ containing the sampling of the FT of the Raman response function.
Output: Array $\widehat{g}$ contains the sampled FT of g.
for $j = 1, \ldots, J$ **do**
$op_1[j] \leftarrow |f[j]|^2$
end for
$\widehat{op}_1 \leftarrow$ FFT$(op_1, forward)$
for $j = 1, \ldots, J$ **do**
$\widehat{op}_2[j] \leftarrow \widehat{op}_1[j] \times \widehat{h}_R[j]$
end for
$op_2 \leftarrow$ FFT$(\widehat{op}_2, backward)$ {Array op_2 contains the convolution product $h_R \star |f|^2$}
for $j = 1, \ldots, N$ **do**
$op_3[j] \leftarrow f[j] \times \big((1 - f_R)op_1[j] + f_R\, op_2[j]\big)$
end for
$\widehat{op}_3 \leftarrow$ FFT$(\widehat{op}_3, forward)$
for $j = 1, \ldots, J$ **do**
$\widehat{op}_4[j] \leftarrow \mathrm{i}\gamma(1 - \frac{2\mathrm{i}\pi\nu_j}{\omega_0})\widehat{op}_3[j]$
end for
for $j = 1, \ldots, J$ **do**
$\widehat{g}[j] \leftarrow -\frac{\alpha}{2}\widehat{f}[j] + \widehat{op}_4[j]$
end for

REFERENCES

[1] G. Agrawal. *Nonlinear fiber optics*. Academic Press, 4th edition, 2006.
[2] L. Auslander and F.A. Grunbaum. The Fourier transform and the discrete Fourier transform. *Inverse Probl*, 5(2):149–164, 1989.
[3] S. Balac. High order Embedded Runge-Kutta scheme for step-size control in the Interaction Picture method. *J. KSIAM*, 17(4):238–266, 2013.
[4] S. Balac and A. Fernandez. Mathematical analysis of adaptive step-size techniques when solving the nonlinear Schrödinger equation for simulating light-wave propagation in optical fibers. *Opt. Commun.*, 329:1–9, 2014.

[5] S. Balac and F. Mahé. Embedded Runge-Kutta scheme for step-size control in the Interaction Picture method. *Comput. Phys. Commun.*, 184:1211–1219, 2013.
[6] 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, 2002.
[7] J.C. Butcher. *Numerical methods for ordinary differential equations*. John Wiley and Sons, 2008.
[8] B. Cano and A. González-Pachón. Plane waves numerical stability of some explicit exponential methods for cubic Schrödinger equation. See `http://hermite.mac.cie.uva.es/bego/cgp3.pdf`, 2013.
[9] B. Cano and A. González-Pachón. Exponential time integration of solitary waves of cubic Schrödinger equation. *Applied Numerical Mathematics*, 91(0):26–45, 2015.
[10] B.M. Caradoc-Davies. *Vortex dynamics in Bose-Einstein condensate*. PhD thesis, University of Otago (NZ), 2000.
[11] R. Carles. *Semi-classical Analysis for Nonlinear Schrödinger Equations*. World Scientific, 2008.
[12] T. Cazenave. *Semilinear Schrödinger Equations*. Courant Lecture Notes in Mathematics, AMS, New York, 2003.
[13] T. Cazenave and A. Haraux. *Introduction aux problèmes d'évolution semi-linéaires*. Ellipses, Paris, 1990.
[14] E. Celledoni, D. Cohen, and B. Owren. Symmetric exponential integrators with an application to the cubic Schrödinger equation. *Found. Comput. Math.*, 8:303–317, 2008.
[15] D. Cohen and L. Gauckler. One-stage exponential integrators for nonlinear Schrödinger equations over long times. *BIT*, 52:877–903, 2012.
[16] S.M. Cox and P. C. Matthews. Exponential time-differencing for stiff systems. *J. Comput. Phys.*, 176:430–455, 2002.
[17] M.J. Davis. *Dynamics in Bose-Einstein condensate*. PhD thesis, University of Oxford (UK), 2001.
[18] J.R. Dormand and P.J. Prince. A family of embedded Runge-Kutta formulae. *J. Comput. Appl. Math.*, 6:19–26, 1980.
[19] G. Dujardin. Exponential Runge-Kutta methods for the Schrödinger equation. *Appl. Numer. Math.*, 59:1839–1857, 2009.
[20] B. L. Ehle and J. D. Lawson. Generalized Runge-Kutta processes for stiff initial-value problems. *J. Inst. Math. Appl.*, 16:11–21, 1975.
[21] C.L. Epstein. How well does the finite Fourier transform approximate the Fourier transform? *Comm. Pure Appl. Math.*, 58(10):1421–1435, 2005.
[22] A. Fernandez, S. Balac, A. Mugnier, F. Mahé, R. Texier-Picard, T. Chartier, and D. Pureur. Numerical simulation of incoherent optical wave propagation in nonlinear fibers. *Eur. Phys. J. - Appl. Phys.*, 64(2):24506/1–11, 2013.
[23] M. Guenin. On the interaction picture. *Commun. Math. Phys.*, 3:120–132, 1966.
[24] E. Hairer, S. P. Norsett, and G. Wanner. *Solving ordinary differential equations I: nonstiff problems*. Springer-Verlag New York, Inc., New York, NY, USA, 1993.
[25] A. Heidt. Efficient adaptive step size method for the simulation of supercontinuum generation in optical fibers. *J. Lightwave Technol.*, 27(18):3984–3991, 2009.
[26] M. Hochbruck and A. Ostermann. Exponential integrators. *Acta Numer.*, 19:209–286, 2010.
[27] J. Hult. A fourth-order Runge–Kutta in the Interaction Picture method for simulating supercontinuum generation in optical fibers. *J. Lightwave Technol.*, 25(12):3770–3775, 2007.
[28] A.-K. Kassam and L. N. Trefethen. Fourth-order time-stepping for stiff PDEs. *SIAM J. Sci. Comput.*, 26(4):1214–1233, 2005.
[29] J. D. Lawson. Generalized Runge-Kutta processes for stable systems with large Lipschitz constants. *SIAM J. Numer. Anal.*, 4:372–380, 1967.
[30] C. Lubich. On splitting methods for Schrödinger-Poisson and cubic nonlinear Schrödinger equations. *Math. Comp.*, 77:2141–2153, 2008.
[31] G.M. Muslu and H.A. Erbay. A split-step Fourier method for the complex modified Korteweg de Vries equation. *Comput. Math. Appl.*, 45(1-3):503 – 514, 2003.
[32] O.V. Sinkin, R. Holzlöhner, J. Zweck, and C.R. Menyuk. Optimization of the Split-Step Fourier method in modeling optical-fiber communications systems. *J. Lightwave Technol.*, 21(1):61, 2003.
[33] M. Thalhammer. Convergence analysis of high-order time-splitting pseudospectral methods for nonlinear Schrödinger equations. *SIAM J. Numer. Anal.*, 50(6):3231–3258, 2012.
[34] J.S. Townsend. *A modern approach to quantum mechanics*. International series in pure and applied physics. University Science Books, 2000.
[35] J.A.C. Weideman and B.M. Herbst. Split-step methods for the solution of the nonlinear Schrodinger equation. *SIAM J. Numer. Anal.*, 23(3):485–507, 1986.