Stochastic Regularity of a Quadratic Observable of High Frequency Waves

G. Malenová · M. Motamed · O. Runborg

Abstract We consider high frequency waves satisfying the scalar wave equation with highly oscillatory initial data. The wave speed as well as the phase and amplitude of the initial data are assumed to be uncertain, described by a finite number of random variables with known probability distributions. We define quantities of interest (QoIs), or observables, as local averages of the squared modulus of the wave solution. We aim to quantify the regularity of these QoIs in terms of the input random parameters, and the wave length, i.e. to estimate the size of their derivatives. The regularity is important for uncertainty quantification methods based on interpolation in the stochastic space. In particular, the size of the derivatives should be bounded independently of the wave length. In this paper we are able to show that when these QoIs are approximated by Gaussian beam superpositions, they indeed have this property, despite the highly oscillatory character of the waves.

Keywords Uncertainty quantification · High frequency wave propagation · Stochastic regularity · Gaussian beam superposition

1 Introduction

Many physical phenomena can be described by propagation of high frequency waves. Such problems arise for example in optics, acoustics, geophysics or oceanography. By high frequency it is understood that the wavelength is very short compared to the distance traveled by the wave. Uncertainty may often be introduced through one or multiple sources: either the model parameters are incomplete or it

G. Malenová, O. Runborg
Department of Mathematics and Swedish e-Science Research Center (SeRC), KTH,
100 44, Stockholm, Sweden
E-mail: {malenova,olofr}@kth.se

M. Motamed
Department of Mathematics and Statistics, The University of New Mexico, Albuquerque,
NM 87131, USA
E-mail: motamed@math.unm.edu

is due to the intrinsic variability of the problem. An example of such uncertain high frequency wave propagation is an earthquake where seismic waves with uncertain epicenter travel through the layers of the Earth with uncertain soil characteristics.

As a simplified model of the wave propagation, we consider the scalar wave equation

$$u_{tt}^{\varepsilon}(t, \mathbf{x}, \mathbf{y}) - c(\mathbf{x}, \mathbf{y})^2 \, \Delta u^{\varepsilon}(t, \mathbf{x}, \mathbf{y}) = 0, \qquad \text{in } [0, T] \times \mathbb{R}^n \times \Gamma, \qquad (1a)$$

$$u^{\varepsilon}(0, \mathbf{x}, \mathbf{y}) = B_0(\mathbf{x}, \mathbf{y}) \, e^{i \, \varphi_0(\mathbf{x}, \mathbf{y})/\varepsilon}, \qquad \text{in } \mathbb{R}^n \times \Gamma, \qquad (1b)$$

$$u_t^{\varepsilon}(0, \mathbf{x}, \mathbf{y}) = \frac{1}{\varepsilon} B_1(\mathbf{x}, \mathbf{y}) \, e^{i \, \varphi_0(\mathbf{x}, \mathbf{y})/\varepsilon}, \qquad \text{in } \mathbb{R}^n \times \Gamma, \qquad (1c)$$

with highly oscillatory initial data represented by the wavelength ε, which is assumed to be much smaller than both the typical scale of variations in the wave speed and the wave propagation distance. Such initial data generate high frequency waves propagating in low frequency media. By $t \in [0, T]$ we denote the time and by $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ the spatial variable. The uncertainty in the model is described by the random variables $\mathbf{y} = (y_1, \ldots, y_N) \in \Gamma \subset \mathbb{R}^N$ where N is a positive integer. Their joint probability density is assumed to be bounded. Two sources of uncertainty are considered: the local speed, $c = c(\mathbf{x}, \mathbf{y})$, and the initial data, $B_0 = B_0(\mathbf{x}, \mathbf{y})$, $B_1 = B_1(\mathbf{x}, \mathbf{y})$, $\varphi_0 = \varphi_0(\mathbf{x}, \mathbf{y})$. The solution is therefore also a function of the random parameter, $u^{\varepsilon} = u^{\varepsilon}(t, \mathbf{x}, \mathbf{y})$. The initial data and the local speed are all assumed to be smooth functions.

In this work, we are interested in *observables*, or *quantities of interest* (QoI) defined as functionals of the solution u^{ε} to (1) for each $\mathbf{y} \in \Gamma$. These observables often carry important physical information. They are functions of the random variable $\mathbf{y}$ and we want to determine their statistical moments, for example the expected value and the variance. This means that we first need to solve (1), then evaluate the functional and finally integrate in $\mathbf{y}$ over Γ, which typically is in a high dimensional space.

Hence, when computing the statistical moments, we are facing two problems: high frequencies and high dimensional integrals. Firstly, modeling of a high frequency problem by direct methods is not feasible. The accuracy of direct methods is determined by the number of points per wavelength. For a given accuracy, the wave resolution cost grows rapidly when shrinking the wavelength. As a consequence, direct numerical methods become very expensive at high frequencies. Secondly, the dimension of the random space is typically large which results in a fast growth of the number of mesh points in standard numerical integration schemes. This phenomenon is referred to as the *curse of dimensionality*.

In [18], we proposed a method for computing the statistics of quantities of interest that deals efficiently with both difficulties. It consists of two techniques: the *Gaussian beam superposition method* for treating the high frequencies and the *sparse stochastic collocation* for treating the uncertainty. The Gaussian beam method is an asymptotic technique largely based on the geometrical optics approach. It has been extensively studied in [4,23,29,16]. The sparse grid collocation method has been used in forward propagation of uncertainty in many PDE models, see [16,25]. The two methods separately are in use, but their combination offers a new powerful tool.

In order for the sparse stochastic collocation to converge fast, high *stochastic regularity* of the QoI is crucial. That is, the derivatives of the QoI with respect to $\mathbf{y}$

should exist and be small. In the high frequency setting, the QoI will depend on the wavelength ε and it is important that the derivatives are bounded independently of ε. In particular, the QoI should not oscillate with period ε. The objective of this paper is to show such bounds for a certain observable.

Natural choices of QoI include the solution itself in a fixed point $(T, \mathbf{x}_0)$,

$$u^\varepsilon(T, \mathbf{x}_0, \mathbf{y}),$$

or more generally the weighted mean value of the solution,

$$\int_{\mathbb{R}^n} u^\varepsilon(T, \mathbf{x}, \mathbf{y})\psi(\mathbf{x})d\mathbf{x}, \qquad \psi \in C_c^\infty(\mathbb{R}^n).$$

In our case, however, neither of these QoIs is interesting due to the high frequency nature of u^ε. In the first case, the point-wise QoI is itself highly oscillatory in $\mathbf{y}$ and its derivatives cannot be bounded by constants that are independent of ε. In the second case, the mean value tends to zero as $\varepsilon \to 0$ due to the stationary phase lemma [9], and little new information is given by the QoI.

Quadratic quantities present another choice of observables. In this paper we will consider:

$$\mathcal{Q}^\varepsilon(\mathbf{y}) = \int_{\mathbb{R}^n} |u^\varepsilon(T, \mathbf{x}, \mathbf{y})|^2 \, \psi(\mathbf{x}) \, d\mathbf{x}, \tag{2}$$

where ψ is again a given real-valued compactly supported smooth function referred to as the *QoI test function*. QoI (2) has an important physical motivation. Since the square of the wave amplitude $|u^\varepsilon|^2$ represents energy density, the weighted integration of the wave amplitude represents the weighted average strength of the wave inside the support of ψ. For example, if the solution u^ε describes the acoustic pressure, $\mathcal{Q}^\varepsilon$ represents the local acoustic potential energy.

It was conjectured in [18] that the QoI (2) computed by the Gaussian beam method, $\mathcal{Q}_{\mathrm{GB}}^\varepsilon$, has high stochastic regularity, more precisely, that

$$\sup_{\mathbf{y}\in\Gamma} \left|\partial_{\mathbf{y}}^{\boldsymbol{\sigma}} \mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y})\right| \leq C_{\boldsymbol{\sigma}}, \quad \forall\boldsymbol{\sigma} \in \mathbb{N}_0^N. \tag{3}$$

The result (3) may seem counterintuitive as the solution itself oscillates with ε. However, formal theoretical arguments and numerical examples presented in [18] showed that $\mathcal{Q}_{\mathrm{GB}}^\varepsilon$ is indeed non-oscillatory. As a consequence the convergence of the sparse stochastic collocation method was exponential, contrasting the rather slow convergence for Monte Carlo simulations.

The main result of this work is Theorem 1 in Section 5, where (3) is proved. This shows the suitability of the sparse grid collocation method for computation of the statistics of the QoI. The proof is based on techniques presented in [16]. We stress that (3) is shown for the approximate solution. Whether it holds also for the exact solution with matching initial data is an open question. A simple argument based on the accuracy of the Gaussian beam solution gives that at least $|\mathcal{Q}_{\mathrm{GB}}^\varepsilon - \mathcal{Q}^\varepsilon| = O(\varepsilon^{k/2-1})$, for kth order beams; see Remark 1. Hence, for sufficiently high order beams $\mathcal{Q}_{\mathrm{GB}}^\varepsilon$ is close to $\mathcal{Q}^\varepsilon$. Based on numerical experiments we conjecture that it is in fact close for all beam orders.

The paper is organized as follows: In Section 2, we introduce precise assumptions. In Section 3, we present the sparse collocation technique and a numerical example. Section 4 contains the Gaussian beam approximation description and properties. In Section 5 we prove the main result on stochastic regularity of the QoI.

2 Preliminaries and assumptions

In this section, we introduce notation used in the paper and formulate precise assumptions for the high frequency wave propagation problem (1) with stochastic parameters. We write $|\mathbf{x}|$ for the Euclidean norm of a vector $\mathbf{x} \in \mathbb{R}^n$. For a multi-index $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$, we however use the standard convention that $|\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_n$. We frequently use the simple estimate,

$$|\mathbf{x}^{\boldsymbol{\alpha}}| \leq |\mathbf{x}|^{|\boldsymbol{\alpha}|}, \qquad \mathbf{x} \in \mathbb{R}^n, \quad \boldsymbol{\alpha} \in \mathbb{N}_0^n.$$

For a function $f : \mathbb{R}^n \mapsto \mathbb{R}$ we let $\nabla f(\mathbf{x})$ denote its gradient, $\Delta f(\mathbf{x})$ represents the Laplacian of f with respect to the spatial variable $\mathbf{x}$, and $D^2 f(\mathbf{x})$ its Hessian. Partial derivatives of order $\boldsymbol{\alpha}$ are written as $\partial_{\mathbf{x}}^{\boldsymbol{\alpha}} f(\mathbf{x})$, where

$$\partial_{\mathbf{x}}^{\boldsymbol{\alpha}} f := \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}, \qquad \mathbf{x} = (x_1, \ldots, x_n)^{\top}.$$

For a vector-valued function $\mathbf{f} : \mathbb{R}^n \mapsto \mathbb{R}^n$ we denote the Jacobian matrix by $D\mathbf{f}(\mathbf{x})$. Finally, we let a^* denote the complex conjugate of a complex number a.

We make the following precise assumptions:

(A1) Strictly positive, smooth and bounded speed of propagation,

$$c \in C^{\infty}(\mathbb{R}^n \times \Gamma), \quad 0 < c_{\min} \leq c(\mathbf{x}, \mathbf{y}) \leq c_{\max} < \infty, \quad \forall \mathbf{x} \in \mathbb{R}^n, \quad \forall \mathbf{y} \in \Gamma.$$

(A2) Smooth and (uniformly) compactly supported initial amplitudes,

$$B_{\ell} \in C^{\infty}(\mathbb{R}^n \times \Gamma), \qquad \operatorname{supp} B_{\ell}(\,\cdot\,, \mathbf{y}) \subset K_0, \qquad \ell = 0, 1, \quad \forall \mathbf{y} \in \Gamma,$$

where $K_0 \subset \mathbb{R}^n$ is a compact set.

(A3) Smooth initial phase with non-zero gradient,

$$\varphi_0 \in C^{\infty}(\mathbb{R}^n \times \Gamma), \qquad |\nabla \varphi_0(\mathbf{x}, \mathbf{y})| > 0, \qquad \forall \mathbf{x} \in \mathbb{R}^n, \quad \forall \mathbf{y} \in \Gamma.$$

(A4) Compact support in random space: the set Γ is the closure of a bounded open set.

(A5) High frequency,

$$0 < \varepsilon \leq 1.$$

(A6) Smooth and compactly supported QoI test function,

$$\psi \in C_c^{\infty}(\mathbb{R}^n).$$

3 Forward propagation of uncertainty

There are in general two types of numerical methods for propagating uncertainty in PDE models subject to stochastic uncertainty: 1) intrusive methods; and 2) non-intrusive methods. Intrusive methods, such as stochastic Galerkin [33, 2], are based on a weak formulation of the problem and are combined with direct simulations such as the finite element method. Consequently, their application to high frequency waves is not feasible, since direct solvers require a large number of elements to resolve the wave's high oscillations. On the contrary, non-intrusive methods, such as Monte Carlo [6] and stochastic collocation [32, 1, 20], are sample-based approaches and rely on a set of deterministic models corresponding to a set of realizations. Consequently, asymptotic techniques for approximating deterministic high frequency waves [5, 27, 22], such as Gaussian beam superposition (presented in Section 4), can be employed within the framework of a non-intrusive method.

In this section, we discuss and motivate the applicability and efficiency of two popular non-intrusive methods, Monte Carlo and stochastic collocation, for approximating the statistics of the QoI in (2).

3.1 Non-intrusive numerical methods

Two popular non-intrusive methods, combined with Gaussian beam superposition, have been proposed in the high frequency wave community: a Monte Carlo method [31] and a sparse stochastic collocation method [18]. In both methods, the solution to problem (1) is approximated in space and time by the Gaussian beam superposition (see Section 4). This gives the Gaussian beam QoI $\mathcal{Q}_{\mathrm{GB}}^{\varepsilon}(\mathbf{y})$ in (13), which is evaluated at some set of points $\{\mathbf{y}^{(m)}\}_{m=1}^{\eta}$. The two methods differ in treating the Gaussian beam QoI in the stochastic space, including the selection of the set of points $\{\mathbf{y}^{(m)}\}_{m=1}^{\eta}$. Here, we briefly review and compare the convergence rates of the two methods.

Monte Carlo. The Monte Carlo method approximates the statistical moments of a stochastic quantity by sample averages. For instance, let $\{\mathbf{y}^{(m)}\}_{m=1}^{\eta} \in \Gamma$ be a set of η independent and identically distributed samples of the random vector $\mathbf{y}$. Then, the Monte Carlo estimator [31] computes the first moment of the quantity $\mathcal{Q}^{\varepsilon}$ by sample averages of the Gaussian beam QoI:

$$\mathbb{E}[\mathcal{Q}^{\varepsilon}] \approx \mathbb{E}[\mathcal{Q}_{\mathrm{GB}}^{\varepsilon}] \approx \frac{1}{\eta} \sum_{m=1}^{\eta} \mathcal{Q}_{\mathrm{GB}}^{\varepsilon}(\mathbf{y}^{(m)}),$$

where $\mathcal{Q}_{\mathrm{GB}}^{\varepsilon}(\mathbf{y}^{(m)})$ is the Gaussian beam QoI (13) evaluated at $\mathbf{y}^{(m)}$.

Sparse stochastic collocation. Stochastic collocation is another non-intrusive technique, which exploits the stochastic regularity of the QoI, if available, to obtain a much faster, possibly spectral convergence. In stochastic collocation, $\mathcal{Q}_{\mathrm{GB}}^{\varepsilon}(\mathbf{y})$ in (13), is collocated in the points $\{\mathbf{y}^{(m)}\}_{m=1}^{\eta} \in \Gamma$, giving the η Gaussian beam quantities $\{\mathcal{Q}_{\mathrm{GB}}^{\varepsilon}(\mathbf{y}^{(m)})\}_{m=1}^{\eta}$. Then, a global polynomial approximation, $\mathcal{Q}_{\mathrm{GB},\eta}^{\varepsilon}$, is built upon those evaluations,

$$\mathcal{Q}_{\mathrm{GB},\eta}^{\varepsilon}(\mathbf{y}) = \sum_{m=1}^{\eta} \mathcal{Q}_{\mathrm{GB}}^{\varepsilon}(\mathbf{y}^{(m)}) \, L_m(\mathbf{y}), \qquad (4)$$

for suitable multivariate polynomials, $\{L_m(\mathbf{y})\}_{m=1}^{\eta}$, such as Lagrange interpolating polynomials. It is important to note that the statistical moments of the quantity $\mathcal{Q}^\varepsilon$ can directly be computed without forming the polynomial approximation $\mathcal{Q}_{\mathrm{GB},\eta}^\varepsilon(\mathbf{y})$ in (4). For instance, the first moment of the quantity $\mathcal{Q}^\varepsilon$ is computed by

$$\mathbb{E}[\mathcal{Q}^\varepsilon] \approx \mathbb{E}[\mathcal{Q}_{\mathrm{GB}}^\varepsilon] \approx \sum_{m=1}^{\eta} \theta_m \, \mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y}^{(m)}), \quad \theta_m = \int_\Gamma L_m(\mathbf{y}) \, \rho(\mathbf{y}) \, d\mathbf{y}. \tag{5}$$

The choice of the set of collocation points, $\{\mathbf{y}^{(m)}\}_{m=1}^{\eta}$, i.e., the type of computational grid in the N-dimensional stochastic space, is crucial in the computational complexity of stochastic collocation. A full tensor grid, based on the Cartesian product of mono-dimensional grids, can only be used when the dimension of the stochastic space, N, is small, since the computational cost grows exponentially fast with N (the curse of dimensionality). In this case, the grid consists of $\eta = (p+1)^N$ collocation points, where p is the polynomial degree in each direction. Alternatively, *sparse grids*, such as total degree and hyperbolic cross grids can reduce the curse of dimensionality. In both full tensor and sparse grids, typical choices of grid points to be used include Gauss abscissas [30], Clenshaw-Curtis abscissas [30], and Leja abscissas [24]. The statistical moments of the QoI can then be computed using Gauss or Clenshaw-Curtis quadrature formulas applied to (4). See [18] for more details.

3.2 Convergence of numerical methods

By the Central Limit Theorem [6], if the variance of $\mathcal{Q}_{\mathrm{GB}}^\varepsilon$ is bounded, then the convergence rate of Monte Carlo estimator is $\mathcal{O}(\eta^{-1/2})$. Therefore, while being very flexible and easy to implement, Monte Carlo sampling features a very slow convergence rate. Multilevel Monte Carlo [7,19], multi-index Monte Carlo [8], and quasi Monte Carlo methods [11] have been proposed to accelerate the convergence of Monte Carlo sampling. It is to be noted that these recent techniques have not been applied to high frequency wave propagation problems.

The rate of convergence of stochastic collocation for approximating a stochastic function $v(\mathbf{y})$ depends on the regularity of the function in stochastic space, i.e., the regularity of the mapping $v : \Gamma \to \mathbb{R}$. Fast convergence with respect to the number of collocation points is attained in the presence of high stochastic regularity. For instance, for a stochastic function $v(\mathbf{y})$ with $s \geq 1$ bounded mixed derivatives, $\partial_{y_1}^s \partial_{y_2}^s \ldots \partial_{y_N}^s v \in L^\infty(\Gamma)$, the error in stochastic collocation on sparse Smolyak grids based on Gauss-Legendre abcissas satisfies (see Theorem 5 in [20])

$$||v - v_\eta||_{L_\rho^2(\Gamma)} \leq C \left(1 + \log_2 \frac{\eta}{N}\right)^{2N} \eta^{-s \frac{\log_e 2}{\xi + \log_e N}}, \quad \xi = 1 + (\log_e 2)(1 + \log_2 1.5) \approx 2.1,$$

where $\| \cdot \|_{L_\rho^2(\Gamma)}$ is the weighted L^2-norm with bounded weight function ρ and

$$C = C_{s,N} \, ||\rho||_\infty^{1/2} \max_{n=1,\ldots,N} \max_{0 \leq k_1,\ldots,k_n \leq s} ||\partial_{y_1}^{k_1} \ldots \partial_{y_n}^{k_n} v||_{L^2(\Gamma)}.$$

Here, the constant C depends on s, N, ρ, and the function v, but not on η. Therefore, the larger s, or equivalently the higher the regularity of v, the faster the

convergence rate in η. In particular, when the stochastic function has infinitely many bounded mixed derivatives, the convergence rate is close to spectral. See also [20, 21, 3].

In the case of high frequency waves with a highly oscillatory wave solution, the constant C depends in general also on the wavelength ε. To maintain fast convergence, it is therefore crucial that C remains bounded as $\varepsilon \to 0$. Since C depends on the size of the $\mathbf{y}$-derivatives of v on Γ, we need uniform bounds for the $\mathbf{y}$-derivatives of $\mathcal{Q}_{\mathrm{GB}}^{\varepsilon}(\mathbf{y})$ independent of ε to obtain fast convergence for all ε. We emphasize that since the Gaussian beam quantity $\mathcal{Q}_{\mathrm{GB}}^{\varepsilon}(\mathbf{y})$ is used in the stochastic collocation approximation (4)-(5), it is the stochastic regularity of the Gaussian beam quantity $\mathcal{Q}_{\mathrm{GB}}^{\varepsilon}(\mathbf{y})$ that determines the convergence rate, not the regularity of the analytical quantity $\mathcal{Q}^{\varepsilon}(\mathbf{y})$.

3.3 Example

Let us present a numerical example to illustrate the discussion in Section 3.2. We introduce two 2D pulses, nearly disjoint in the beginning and moving towards each other. At a finite time, they start to overlap which leads to oscillations of order ε in the magnitude of the solution. We consider three uniformly distributed independent random variables $\mathbf{y} = (y_1, y_2, y_3)$ characterizing a random speed, $c = c(\mathbf{y}) = y_1$, and a random initial position of the second pulse, $\mathbf{s}_2 = \mathbf{s}_2(\mathbf{y}) = (y_2, y_3)$. We use the smooth test function

$$\psi(\mathbf{x}) = \begin{cases} e^{-\frac{|\mathbf{x}|^2}{1/4 - |\mathbf{x}|^2}}, & |\mathbf{x}| < 1/2. \\ 0, & \text{otherwise.} \end{cases}$$

The initial data in (1) read

$$\varphi_0(\mathbf{x}) = |x_1|, \qquad B_0(\mathbf{x}, \mathbf{y}) = e^{-5|\mathbf{x} - \mathbf{s}_1|^2} + e^{-5|\mathbf{x} - \mathbf{s}_2(\mathbf{y})|^2},$$

where $\mathbf{s}_1 = (-1, 0)$. We consider the random variables to span

$$y_1 \sim \mathcal{U}(0.8, 1), \qquad y_2 \sim \mathcal{U}(1, 1.5), \qquad y_3 \sim \mathcal{U}(0, 0.5).$$

The amplitude B_1 is chosen such that the pulses move towards each other, without producing any backward propagating waves. The first order Gaussian beam method from Section 4 has been utilized to compute the solution u^{ε}. Figure 1 shows the absolute value of u^{ε} with $\mathbf{y} = (0.8, 1, 0.5)$ and wavelength $\varepsilon = 1/40$ at various times. At time $T = 1$, the pulses reach the central circle indicating the support of the QoI test function with oscillations emerging in the overlap area.

In Figure 2, we plot the quantity of interest (13) and its first and second derivatives along the line $\mathbf{y}(r) = (0.8 + 0.2r, 1 + 0.5r, 0.5r)$, $r \in [0, 1]$ at time $T = 1$. We note that the solution u^{ε} oscillates with ε, however, this is not the case for the QoI and its derivatives. This suggests bounds of the type (3).

Figure 3 shows the relative error of the approximate expected value of the Gaussian beam QoI

$$e(\eta) := \frac{\left| \mathbb{E}[\mathcal{Q}_{\mathrm{GB}, \eta_{\mathrm{ref}}}^{\varepsilon}] - \mathbb{E}[\mathcal{Q}_{\mathrm{GB}, \eta}^{\varepsilon}] \right|}{\left| \mathbb{E}[\mathcal{Q}_{\mathrm{GB}, \eta_{\mathrm{ref}}}^{\varepsilon}] \right|},$$
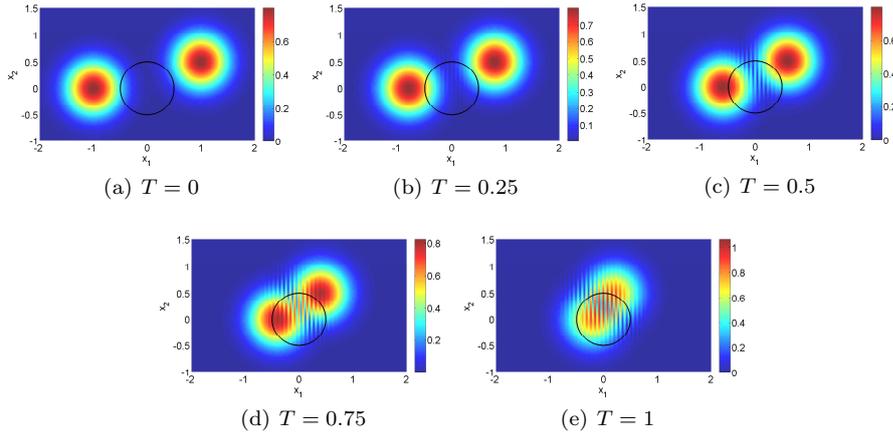
Fig. 1 Time evolution of $|u^\varepsilon(T, \mathbf{x}, \mathbf{y})|$ with wavelength $\varepsilon = 1/40$ and $\mathbf{y} = (0.8, 1, 0.5)$ in the example in Section 3.3. The central circle indicates the support of the QoI test function.
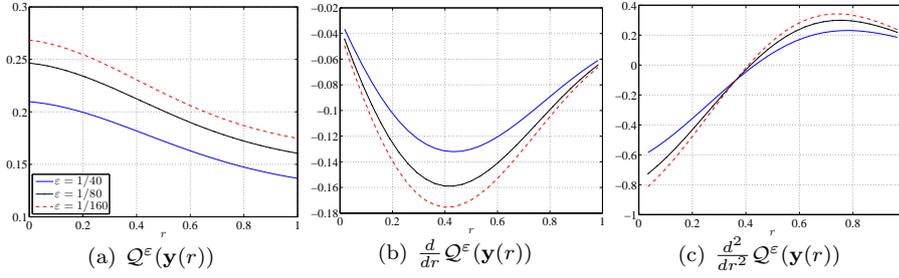
Fig. 2 QoI $\mathcal{Q}^\varepsilon_{\mathrm{GB}}$ and its first and second derivatives along the line $\mathbf{y}(r) = (0.8 + 0.2r, 1 + 0.5r, 0.5r)$ where $r \in [0, 1]$, for wavelengths $\varepsilon = [1/40, 1/80, 1/160]$.

versus the number of collocation points η for various wavelengths. The Gaussian beam quantity is computed by (4) on the sparse Smolyak grid at time $T = 1$. To obtain a reference solution, we use $\eta_{\mathrm{ref}} = 30465$ collocation points. The figure also compares convergence of the Monte Carlo and the stochastic collocation methods. A simple linear regression through the Monte Carlo data points shows that the rate of convergence of Monte Carlo is approximately 0.46. We observe a fast spectral convergence of the stochastic collocation error due to the high stochastic regularity of the Gaussian beam QoI. Moreover, the decay rate does not deteriorate as ε decreases.

4 Gaussian beams

In this section, we briefly describe the Gaussian beam approximation. We restrict the description to the points that are relevant for the analysis in Section 5. For a more detailed account with a general derivation for hyperbolic equations, dispersive wave equations and Helmholtz equation, we refer to [26, 28, 16, 13, 12, 14].
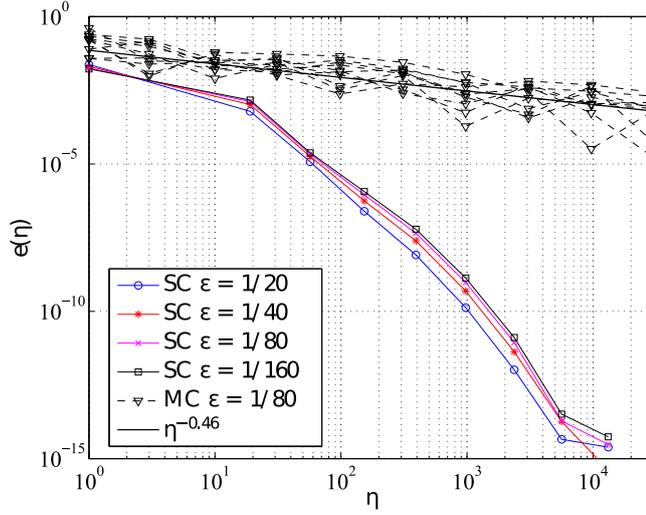
Fig. 3 Relative error $e(\eta)$ at time $T = 1$ versus the number of collocation points η (or the number of samples in the case of Monte Carlo sampling), for various wavelengths.

Individual Gaussian beams are asymptotic solutions to the wave equation that concentrate around a *central ray* in space-time. We denote the k-th order Gaussian beam and the central ray starting at $\mathbf{z} \in K_0$ by $v_k(t, \mathbf{y}; \mathbf{z})$ and $\mathbf{q}(t, \mathbf{y}; \mathbf{z})$ respectively. The beam has the following form,

$$v_k(t, \mathbf{x}, \mathbf{y}; \mathbf{z}) = A_k(t, \mathbf{x} - \mathbf{q}(t, \mathbf{y}; \mathbf{z}), \mathbf{y}; \mathbf{z})e^{i\Phi_k(t, \mathbf{x} - \mathbf{q}(t, \mathbf{y}; \mathbf{z}), \mathbf{y}; \mathbf{z})/\varepsilon}, \qquad (6)$$

where

$$\Phi_k(t, \mathbf{x}, \mathbf{y}; \mathbf{z}) = \phi_0(t, \mathbf{y}; \mathbf{z}) + \mathbf{x} \cdot \mathbf{p}(t, \mathbf{y}; \mathbf{z}) + \frac{1}{2}\mathbf{x} \cdot M(t, \mathbf{y}; \mathbf{z})\mathbf{x} + \sum_{|\boldsymbol{\beta}|=3}^{k+1}\frac{1}{\boldsymbol{\beta}!}\phi_{\boldsymbol{\beta}}(t, \mathbf{y}; \mathbf{z})\mathbf{x}^{\boldsymbol{\beta}},$$

is the k-th order phase function and

$$A_k(t, \mathbf{x}, \mathbf{y}; \mathbf{z}) = \sum_{j=0}^{\lceil\frac{k}{2}\rceil-1}\varepsilon^j\bar{a}_{j,k}(t, \mathbf{x}, \mathbf{y}; \mathbf{z}), \qquad (7)$$

is the k-th order amplitude function, with

$$\bar{a}_{j,k}(t, \mathbf{x}, \mathbf{y}; \mathbf{z}) = \sum_{|\boldsymbol{\beta}|=0}^{k-2j-1}\frac{1}{\boldsymbol{\beta}!}a_{j,\boldsymbol{\beta}}(t, \mathbf{y}; \mathbf{z})\mathbf{x}^{\boldsymbol{\beta}} . \qquad (8)$$

Single beams are summed together to form the k-th order Gaussian beam superposition solution $u_k(t, \mathbf{x}, \mathbf{y})$,

$$u_k(t, \mathbf{x}, \mathbf{y}) = \left(\frac{1}{2\pi\varepsilon}\right)^{\frac{n}{2}}\int_{K_0}v_k(t, \mathbf{x}, \mathbf{y}; \mathbf{z})\varrho_\eta(\mathbf{x} - \mathbf{q}(t, \mathbf{y}; \mathbf{z}))d\mathbf{z}, \qquad (9)$$

where the integration in $\mathbf{z}$ is over the support of the initial data $K_0 \subset \mathbb{R}^n$. Since the wave equation is linear, the superposition of beams is still an asymptotic solution.

The function $\varrho_\eta \in C^\infty(\mathbb{R}^n)$ is a real-valued *cutoff* function with radius $0 < \eta \leq \infty$ satisfying,

$$\varrho_\eta(\mathbf{x}) \geq 0 \quad \text{and} \quad \varrho_\eta(\mathbf{x}) = \begin{cases} \begin{array}{l} 1 \text{ for } |\mathbf{x}| \leq \eta, \\ 0 \text{ for } |\mathbf{x}| \geq 2\eta, \end{array} & \text{for } 0 < \eta < \infty, \\ 1, & \text{for } \eta = \infty. \end{cases} \tag{10}$$

As shown below in (P4), if $\eta > 0$ is sufficiently small, it is ensured that Im $\Phi_k > 0$ on the support of ϱ_η and the Gaussian beam superposition is well-behaved. For first order beams, $k = 1$, the cutoff function is not needed and we can take $\eta = \infty$.

Since the wave equation (1) is a second order equation, two modes and two Gaussian beam superpositions are needed, one for forward and one for backward propagating waves. We denote the corresponding coefficients by $^+$ and $^-$, respectively, and write the solution

$$\tilde{u}_k(t, \mathbf{x}, \mathbf{y}) = \left(\frac{1}{2\pi\varepsilon}\right)^{\frac{n}{2}} \int_{K_0} \Big[v_k^+(t, \mathbf{x}, \mathbf{y}; \mathbf{z}) \varrho_\eta(\mathbf{x} - \mathbf{q}^+(t, \mathbf{y}; \mathbf{z}))$$
$$+ v_k^-(t, \mathbf{x}, \mathbf{y}; \mathbf{z}) \varrho_\eta(\mathbf{x} - \mathbf{q}^-(t, \mathbf{y}; \mathbf{z})) \Big] d\mathbf{z},$$

where $v_k^\pm$ are built from the central rays $\mathbf{q}^\pm(t, \mathbf{y}; \mathbf{z})$ and coefficients $\phi_0^\pm$, $\mathbf{p}^\pm$, $M^\pm$, $\phi_{\boldsymbol{\beta}}^\pm$, $a_{j,\boldsymbol{\beta}}^\pm$.

The central rays $\mathbf{q}^\pm(t, \mathbf{y}; \mathbf{z})$ and all the coefficients of the two modes $\phi_0^\pm$, $\mathbf{p}^\pm$, $M^\pm$, $\phi_{\boldsymbol{\beta}}^\pm$ and $a_{j,\boldsymbol{\beta}}^\pm$ satisfy ODEs in t. The dependence on $\mathbf{z}$ is only via the initial data. The leading order ODEs are

$$\partial_t \mathbf{q}^\pm = \pm c(\mathbf{q}^\pm) \frac{\mathbf{p}^\pm}{|\mathbf{p}^\pm|}, \tag{11a}$$

$$\partial_t \mathbf{p}^\pm = \mp \nabla c(\mathbf{q}^\pm) |\mathbf{p}^\pm|, \tag{11b}$$

$$\partial_t \phi_0^\pm = 0, \tag{11c}$$

$$\partial_t M^\pm = \mp (E + B^T M^\pm + M^\pm B + M^\pm C M^\pm), \tag{11d}$$

$$\partial_t a_0^\pm = \pm \frac{1}{2|\mathbf{p}^\pm|} \left(\mathbf{p}^\pm \cdot \nabla c(\mathbf{q}^\pm) + \frac{c(\mathbf{q}^\pm) \, \mathbf{p}^\pm \cdot M \mathbf{p}^\pm}{|\mathbf{p}^\pm|^2} - c(\mathbf{q}^\pm) \operatorname{Tr}(M^\pm) \right) a_0^\pm, \tag{11e}$$

with

$$E = |\mathbf{p}^\pm| D^2 c(\mathbf{q}^\pm), \quad B = \frac{\mathbf{p}^\pm \otimes \nabla c(\mathbf{q}^\pm)}{|\mathbf{p}^\pm|}, \quad C = \frac{c(\mathbf{q}^\pm)}{|\mathbf{p}^\pm|} \operatorname{Id}_{n \times n} - \frac{c(\mathbf{q}^\pm)}{|\mathbf{p}^\pm|^3} \mathbf{p}^\pm \otimes \mathbf{p}^\pm.$$

There are similar linear ODEs for higher order coefficients $\phi_{\boldsymbol{\beta}}^\pm$ and $a_{j,\boldsymbol{\beta}}^\pm$; see e.g. [26,28].

Each Gaussian beam $v_k^\pm(t, \mathbf{x}, \mathbf{y}; \mathbf{z})$ requires initial values for the central ray and all of the amplitude and phase Taylor coefficients. The appropriate choice of these initial values will make $\tilde{u}_k(0, \mathbf{x}, \mathbf{y})$ asymptotically converge to the initial conditions

in (1). As shown in [16], initial data for the central ray and phase coefficients should be chosen as follows.

$$\mathbf{q}^{\pm}(0, \mathbf{y}; \mathbf{z}) = \mathbf{z}, \tag{12a}$$

$$\mathbf{p}^{\pm}(0, \mathbf{y}; \mathbf{z}) = \nabla \varphi_0(\mathbf{z}, \mathbf{y}), \tag{12b}$$

$$\phi_0^{\pm}(0, \mathbf{y}; \mathbf{z}) = \varphi_0(\mathbf{z}, \mathbf{y}), \tag{12c}$$

$$M^{\pm}(0, \mathbf{y}; \mathbf{z}) = D^2 \varphi_0(\mathbf{z}, \mathbf{y}) + i \, \mathrm{Id}_{n \times n}, \tag{12d}$$

$$\phi_{\boldsymbol{\beta}}^{\pm}(0, \mathbf{y}; \mathbf{z}) = \partial_{\mathbf{x}}^{\boldsymbol{\beta}} \varphi_0(\mathbf{z}, \mathbf{y}), \qquad |\boldsymbol{\beta}| = 3, \dots, k+1 \; . \tag{12e}$$

The initial data for the amplitude coefficients is more complicated and we refer the reader to [16] for more detailed description.

In the remainder of the paper we will restrict ourselves to only one of the two modes and drop the $\pm$ subscript. This means that we only consider high frequency wave solutions that propagate in one direction from the initial source. Note that the interaction between the two modes could potentially introduce more oscillations in $\mathcal{Q}_{\mathrm{GB}}^{\varepsilon}$. An analysis of this will be reported elsewhere.

4.1 Gaussian beam quantity of interest

Instead of the full quantity of interest defined in (2) we here consider the corresponding quantity of interest computed based on the Gaussian beam approximation with one mode,

$$\mathcal{Q}_{\mathrm{GB}}^{\varepsilon}(\mathbf{y}) = \int_{\mathbb{R}^n} |u_k(T, \mathbf{x}, \mathbf{y})|^2 \, \psi(\mathbf{x}) \, d\mathbf{x}, \tag{13}$$

where u_{k} is given by (9) and v_k is either v_k^+ or v_k^-, with corresponding $\mathbf{q}^+$ or $\mathbf{q}^-$ respectively. Since u_k is a good approximation of u^{ε} the QoI $\mathcal{Q}_{GB}^{\varepsilon}$ in (13) is a good approximation to $\mathcal{Q}^{\varepsilon}$; see Remark 1.

We get

$$\mathcal{Q}_{\mathrm{GB}}^{\varepsilon}(\mathbf{y}) = \left(\frac{1}{2\pi\varepsilon}\right)^n \int_{\mathbb{R}^n} \int_{K_0 \times K_0} v_k(T, \mathbf{x}, \mathbf{y}; \mathbf{z}) \varrho_\eta(\mathbf{x} - \mathbf{q}(T, \mathbf{y}; \mathbf{z}))$$
$$\times v_k^*(T, \mathbf{x}, \mathbf{y}; \mathbf{z}') \varrho_\eta(\mathbf{x} - \mathbf{q}(T, \mathbf{y}; \mathbf{z}')) \, \psi(\mathbf{x}) d\mathbf{z} \, d\mathbf{z}' d\mathbf{x}$$
$$=: \left(\frac{1}{2\pi\varepsilon}\right)^n \int_{K_0 \times K_0} I^{\varepsilon}(T, \mathbf{y}; \mathbf{z}, \mathbf{z}') \, d\mathbf{z} \, d\mathbf{z}'. \tag{14}$$

Since $(ae^{i\phi})^* = a^* e^{-i\phi^*}$, we can write using (6),

$$I^{\varepsilon}(T, \mathbf{y}; \mathbf{z}, \mathbf{z}') =$$

$$\int_{\mathbb{R}^n} \sum_{j,m=0}^{\lceil k/2 \rceil - 1} \sum_{|\boldsymbol{\alpha}|=0}^{k-2j-1} \sum_{|\boldsymbol{\beta}|=0}^{k-2m-1} \varepsilon^{j+m} g_{j,m,\boldsymbol{\alpha},\boldsymbol{\beta}}(T, \mathbf{x}, \mathbf{y}; \mathbf{z}, \mathbf{z}') \, (\mathbf{x} - \mathbf{q}(T, \mathbf{y}; \mathbf{z}))^{\boldsymbol{\alpha}}$$

$$\times (\mathbf{x} - \mathbf{q}(T, \mathbf{y}; \mathbf{z}'))^{\boldsymbol{\beta}} \, e^{i[\Phi_k(T, \mathbf{x} - \mathbf{q}(T, \mathbf{y}; \mathbf{z}), \mathbf{y}; \mathbf{z}) - \Phi_k^*(T, \mathbf{x} - \mathbf{q}(T, \mathbf{y}; \mathbf{z}'), \mathbf{y}; \mathbf{z}')]/\varepsilon} \, d\mathbf{x},$$

where $g_{j,m,\boldsymbol{\alpha},\boldsymbol{\beta}}$ is given by

$$g_{j,m,\boldsymbol{\alpha},\boldsymbol{\beta}} = \frac{1}{\boldsymbol{\alpha}! \, \boldsymbol{\beta}!} a_{j,\boldsymbol{\alpha}}(T, \mathbf{y}; \mathbf{z}) a_{m,\boldsymbol{\beta}}^*(T, \mathbf{y}; \mathbf{z}') \varrho_\eta(\mathbf{x} - \mathbf{q}(T, \mathbf{y}; \mathbf{z})) \varrho_\eta(\mathbf{x} - \mathbf{q}(T, \mathbf{y}; \mathbf{z}')) \psi(\mathbf{x}).$$

To simplify the expressions, we introduce the symmetrizing variables

$$\bar{\mathbf{q}} = \bar{\mathbf{q}}(T, \mathbf{y}; \mathbf{z}, \mathbf{z}') = \frac{\mathbf{q}(T, \mathbf{y}; \mathbf{z}) + \mathbf{q}(T, \mathbf{y}; \mathbf{z}')}{2},$$

$$\Delta\mathbf{q} = \Delta\mathbf{q}(T, \mathbf{y}; \mathbf{z}, \mathbf{z}') = \frac{\mathbf{q}(T, \mathbf{y}; \mathbf{z}) - \mathbf{q}(T, \mathbf{y}; \mathbf{z}')}{2}, \tag{15}$$

and write

$$I^\varepsilon(T, \mathbf{y}; \mathbf{z}, \mathbf{z}') = \tag{16}$$

$$\int_{\mathbb{R}^n} \sum_{j,m=0}^{\lceil k/2 \rceil - 1} \sum_{|\boldsymbol{\alpha}|=0}^{k-2j-1} \sum_{|\boldsymbol{\beta}|=0}^{k-2m-1} \varepsilon^{j+m} \bar{g}_{j,m,\boldsymbol{\alpha},\boldsymbol{\beta}} \left(\mathbf{x} - \Delta\mathbf{q}\right)^{\boldsymbol{\alpha}} (\mathbf{x} + \Delta\mathbf{q})^{\boldsymbol{\beta}} e^{i\Psi/\varepsilon} \, d\mathbf{x},$$

where

$$\Psi(T, \mathbf{x}, \mathbf{y}; \mathbf{z}, \mathbf{z}') = \Phi_k(T, \mathbf{x} - \Delta\mathbf{q}, \mathbf{y}; \mathbf{z}) - \Phi_k^*(T, \mathbf{x} + \Delta\mathbf{q}, \mathbf{y}; \mathbf{z}') \tag{17}$$

is the symmetrized phase and

$$\bar{g}_{j,m,\boldsymbol{\alpha},\boldsymbol{\beta}}(T, \mathbf{x}, \mathbf{y}; \mathbf{z}, \mathbf{z}') =$$

$$\frac{1}{\boldsymbol{\alpha}!\boldsymbol{\beta}!} \, a_{j,\boldsymbol{\alpha}}(T, \mathbf{y}; \mathbf{z}) \, a_{m,\boldsymbol{\beta}}^*(T, \mathbf{y}; \mathbf{z}') \, \varrho_\eta(\mathbf{x} - \Delta\mathbf{q}) \, \varrho_\eta(\mathbf{x} + \Delta\mathbf{q}) \, \psi(\mathbf{x} + \bar{\mathbf{q}}). \tag{18}$$

Remark 1 In [16] it was proved that under the given smoothness conditions and for a small enough cutoff radius η,

$$\sup_{0 \le t \le T} ||u^\varepsilon(t, \cdot) - u_k(t, \cdot)||_{H^1} \le C(T)\varepsilon^{k/2-1},$$

if initial data is sufficiently well approximated by the beams. This immediately gives the same estimate for $\mathcal{Q}_{\mathrm{GB}}^\varepsilon$ when $\psi \ge 0$. Indeed, for a fixed $\mathbf{y}$,

$$|\mathcal{Q}^\varepsilon(\mathbf{y}) - \mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y})| = \int \left| |u^\varepsilon(T, \mathbf{x}, \mathbf{y})|^2 - |u_k(T, \mathbf{x}, \mathbf{y})|^2 \right| \psi(\mathbf{x}) d\mathbf{x}$$

$$\le \int |u^\varepsilon(T, \mathbf{x}, \mathbf{y}) - u_k(T, \mathbf{x}, \mathbf{y})| \cdot (|u^\varepsilon(T, \mathbf{x}, \mathbf{y})| + |u_k(T, \mathbf{x}, \mathbf{y})|) \, \psi(\mathbf{x}) d\mathbf{x}$$

$$\le ||u^\varepsilon(T, \cdot, \mathbf{y}) - u_k(T, \cdot, \mathbf{y})||_{L^2} \left(\sqrt{\mathcal{Q}^\varepsilon(\mathbf{y})} + \sqrt{\mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y})} \right)$$

$$\le C(T, \mathbf{y}) \, \varepsilon^{k/2-1} \left(2\sqrt{1 + \mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y})} + |\mathcal{Q}^\varepsilon(\mathbf{y}) - \mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y})| \right).$$

For small enough ε and $k \ge 3$ we can move the second term over to the left hand side and get the estimate,

$$|\mathcal{Q}^\varepsilon(\mathbf{y}) - \mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y})| \le \tilde{C}(T, \mathbf{y})\varepsilon^{k/2-1}\sqrt{1 + \mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y})}.$$

Below we show that $\mathcal{Q}_{\mathrm{GB}}^\varepsilon$ is bounded independently of ε, and the estimate therefore implies that $\mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y})$ is close to $\mathcal{Q}^\varepsilon(\mathbf{y})$ for small ε, at least if we use beams of third order or higher. We do not believe this estimate is sharp, however. In particular, in computations we see a small difference already with first order beams.

4.2 Gaussian beam phase properties

In this section we present some consequences of assumptions (A1)–(A4) for the Gaussian beam coefficient functions, the whole phase Φ_k and the symmetrized phase Ψ. The properties are of great importance in the subsequent estimates in Section 5. They also show the significance of the cutoff function introduced in Section 4, and indicate how the width η should be chosen. Following (P4) below we make the definition:

Definition 1 The cutoff width η used for the Gaussian beam approximation is called *admissible* for Φ_k if it is small enough in the sense of property (P4) in Proposition 1.

Throughout this paper, we will consider the cutoff function ϱ_η in the Gaussian beam approximation (9) to be chosen such that η is admissible in the sense of Definition 1.

Because of the cutoff-function, the terms $g_{j,m,\boldsymbol{\alpha},\boldsymbol{\beta}}$ in (18) are zero unless $|\mathbf{x} - \Delta\mathbf{q}|$ and $|\mathbf{x} + \Delta\mathbf{q}|$ are both small enough. We introduce the subsets $\Omega_\mu \subset \mathbb{R}^n$, parameterized by μ, to describe this

$$\Omega_\mu = \Omega_\mu(t, \mathbf{y}; \mathbf{z}, \mathbf{z}') = \{\mathbf{x} \in \mathbb{R}^n \ : \ |\mathbf{x} + \Delta\mathbf{q}| \leq 2\mu \text{ and } |\mathbf{x} - \Delta\mathbf{q}| \leq 2\mu\}.$$

We note that since $|\mathbf{x}|^2 \leq |\mathbf{x}|^2 + |\Delta\mathbf{q}|^2 = |\mathbf{x} + \Delta\mathbf{q}|^2/2 + |\mathbf{x} - \Delta\mathbf{q}|^2/2$, we have the inclusion

$$\Omega_\mu \subset B_{2\mu}, \qquad B_\mu := \{\mathbf{x} \in \mathbb{R}^n \ : \ |\mathbf{x}| \leq \mu\}. \tag{19}$$

The components of the Gaussian beam approximation then have the following properties.

Proposition 1 *Under assumptions (A1)–(A4),*

(P1) ϕ_0, $\mathbf{p}$, $\mathbf{q}$, M, $\phi_{\boldsymbol{\beta}}$, $a_{j,\boldsymbol{\beta}} \in C^\infty(\mathbb{R} \times \Gamma \times \mathbb{R}^n)$ and

$$\operatorname{supp} a_{j,\boldsymbol{\beta}}(t, \mathbf{y}; \cdot) \subset K_0, \qquad \forall t \in \mathbb{R}, \ \mathbf{y} \in \Gamma. \tag{20}$$

(P2) $\Phi_k \in C^\infty(\mathbb{R} \times \mathbb{R}^n \times \Gamma \times \mathbb{R}^n)$,
(P3) For each $T > 0$, there is a constant $C > 0$ such that

$$|\mathbf{p}(t, \mathbf{y}; \mathbf{z}) - \mathbf{p}(t, \mathbf{y}; \mathbf{z}')| + |\mathbf{q}(t, \mathbf{y}; \mathbf{z}) - \mathbf{q}(t, \mathbf{y}; \mathbf{z}')| \geq C|\mathbf{z} - \mathbf{z}'| \ ,$$

for all $t \in [0, T]$, $\mathbf{z}, \mathbf{z}' \in K_0$ and $\mathbf{y} \in \Gamma$.
(P4) For each $T > 0$ and beam order k, there is a Gaussian beam cutoff width $\eta > 0$ and constant $\delta > 0$, such that for all $\mathbf{x} \in B_{2\eta}$,

$$\operatorname{Im} \Phi_k(t, \mathbf{x}, \mathbf{y}; \mathbf{z}) \geq \frac{1}{2}\delta|\mathbf{x}|^2 \ , \qquad \forall t \in [0, T], \ \mathbf{y} \in \Gamma, \ \mathbf{z} \in K_0, \tag{21}$$

Moreover, for the symmetrized phase, when $\mathbf{x} \in \Omega_\eta$,

$$\operatorname{Im} \Psi(t, \mathbf{x}, \mathbf{y}; \mathbf{z}, \mathbf{z}') \geq \frac{1}{2}\delta\left[|\mathbf{x} + \Delta\mathbf{q}|^2 + |\mathbf{x} - \Delta\mathbf{q}|^2\right] = \delta|\mathbf{x}|^2 + \delta|\Delta\mathbf{q}|^2, \tag{22}$$

for all $t \in [0, T]$, $\mathbf{y} \in \Gamma$ and $\mathbf{z}, \mathbf{z}' \in K_0$. When $k = 1$ the width η can be any positive value and (21, 22) hold for all $\mathbf{x} \in \mathbb{R}^n$.

(P5) *The following statement holds for all $t \in \mathbb{R}$, $\mathbf{y} \in \Gamma$ and $\mathbf{z} \in \mathbb{R}^n$,*

$$\nabla_{\mathbf{z}} \phi_0(\mathbf{y}; \mathbf{z}) = D_{\mathbf{z}} \mathbf{q}(t, \mathbf{y}; \mathbf{z})^T \mathbf{p}(t, \mathbf{y}; \mathbf{z}).$$

(P6) *Let $\mathbf{z}, \mathbf{z}' \in K_0$ be such that $|\mathbf{q}(t, \mathbf{y}; \mathbf{z}) - \mathbf{q}(t, \mathbf{y}; \mathbf{z}')| \leq \theta |\mathbf{z} - \mathbf{z}'|$ for fixed $t \in [0, T]$ and $\mathbf{y} \in \Gamma$. Then, for small enough θ and μ, there exists a constant $C(\theta, \mu) > 0$ such that*

$$\inf_{\mathbf{x} \in \Omega_\mu} |\nabla_{\mathbf{x}} \Psi(t, \mathbf{x}, \mathbf{y}; \mathbf{z}, \mathbf{z}')| \geq C(\theta, \mu) |\mathbf{z} - \mathbf{z}'|.$$

The constant is independent of $\mathbf{z}$, $\mathbf{z}'$, t and $\mathbf{y}$.

The proof of these properties essentially follows the same steps as in the corresponding lemmas in [15, 16]. The details are provided in [17].

5 Stochastic regularity

I this section we prove the main result of the paper, namely that the derivative $\partial_{\mathbf{y}}^{\boldsymbol{\sigma}} \mathcal{Q}_{\mathrm{GB}}^{\varepsilon}$ is uniformly bounded in $\mathbf{y}$ and ε. More precisely, we show the following theorem.

Theorem 1 *Under assumptions (A1)-(A6), for the Gaussian beam quantity of interest (13) where the cutoff parameter η is admissible in the sense of Definition 1, there are constants $C_{\boldsymbol{\sigma}}$, independent of ε, such that*

$$\sup_{\mathbf{y} \in \Gamma} \left| \frac{\partial^{\boldsymbol{\sigma}} \mathcal{Q}_{\mathrm{GB}}^{\varepsilon}(\mathbf{y})}{\partial \mathbf{y}^{\boldsymbol{\sigma}}} \right| \leq C_{\boldsymbol{\sigma}}, \quad \forall \boldsymbol{\sigma} \in \mathbb{N}_0^N.$$

In Section 5.1, we prove various prerequisite lemmas needed for the proof of Theorem 1 and introduce notation used in the rest of the paper. Section 5.2 is devoted to rewriting the QoI and its derivatives. In Section 5.3, several important estimates are presented. In Section 5.4, the actual proof of Theorem 1 is carried out.

5.1 Preliminaries

Let $\mathcal{U}_\eta$ denote the space

$$\mathcal{U}_\eta = \{ f \in C^\infty(\mathbb{R} \times \mathbb{R}^n \times \Gamma \times \mathbb{R}^n \times \mathbb{R}^n) : \operatorname{supp} f(t, \cdot, \mathbf{y}; \mathbf{z}, \mathbf{z}') \subset \Omega_\eta(t, \mathbf{y}; \mathbf{z}, \mathbf{z}') \}.$$

We also use the shorthand

$$\mathcal{V} = C^\infty(\mathbb{R} \times \Gamma \times \mathbb{R}^n \times \mathbb{R}^n),$$

and note that if $u, v \in \mathcal{U}_\eta$ and $w \in \mathcal{V}$, then both $uv \in \mathcal{U}_\eta$, and $uw \in \mathcal{U}_\eta$. Consequently, for $\bar{g}_{j,m,\boldsymbol{\alpha},\boldsymbol{\beta}}$ in (18), it holds that $\bar{g}_{j,m,\boldsymbol{\alpha},\boldsymbol{\beta}} \in \mathcal{U}_\eta$ due to (A6), (10) and (P1).

We first note that the Taylor expansion remainder of a smooth function belongs to $\mathcal{V}$.

Lemma 1 *Let f be a scalar function such that $f \in C^\infty(\mathbb{R} \times \Gamma \times \mathbb{R}^n)$, then its Taylor expansion to k-th order reads*

$$f(t, \mathbf{y}; \mathbf{z}') = \sum_{|\boldsymbol{\alpha}| \leq k} \frac{\partial_{\mathbf{z}}^{\boldsymbol{\alpha}} f(t, \mathbf{y}; \mathbf{z})}{\boldsymbol{\alpha}!} (\mathbf{z} - \mathbf{z}')^{\boldsymbol{\alpha}} + \sum_{|\boldsymbol{\beta}|=k+1} R_{\boldsymbol{\beta}}(t, \mathbf{y}; \mathbf{z}, \mathbf{z}')(\mathbf{z} - \mathbf{z}')^{\boldsymbol{\beta}}, \quad (23)$$

where $R_{\boldsymbol{\beta}} \in \mathcal{V}$. Similarly, for a vector valued function $\mathbf{g} \in C^\infty(\mathbb{R} \times \Gamma \times \mathbb{R}^n)$, there exists a matrix valued function $G \in \mathcal{V}$ so that

$$\mathbf{g}(t, \mathbf{y}; \mathbf{z}') = \mathbf{g}(t, \mathbf{y}; \mathbf{z}) + G(t, \mathbf{y}; \mathbf{z}, \mathbf{z}')(\mathbf{z} - \mathbf{z}'). \qquad (24)$$

Proof This follows from Taylor's theorem by considering the integral form of the remainder terms.

We will often use the following lemma.

Lemma 2 *For a matrix $A \in \mathcal{V}$ and $\mathbf{u} \in \mathbb{R}^n$, there exists function $c_{\boldsymbol{\beta}} \in \mathcal{V}$ such that*

$$\left(A(t, \mathbf{y}; \mathbf{z}, \mathbf{z}')\,\mathbf{u}\right)^{\boldsymbol{\alpha}} = \sum_{|\boldsymbol{\beta}|=|\boldsymbol{\alpha}|} c_{\boldsymbol{\beta}, \boldsymbol{\alpha}}(t, \mathbf{y}; \mathbf{z}, \mathbf{z}')\mathbf{u}^{\boldsymbol{\beta}}. \qquad (25)$$

Proof We recall that for any vectors $\mathbf{a}, \mathbf{a}_j$, $j = 1, \ldots n$, scalar u and multiindices $\boldsymbol{\alpha}, \boldsymbol{\gamma}_j$,

$$(\mathbf{a}_1 + \ldots + \mathbf{a}_n)^{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\gamma}_1 + \ldots + \boldsymbol{\gamma}_n = \boldsymbol{\alpha}} \frac{\boldsymbol{\alpha}!}{\boldsymbol{\gamma}_1! \ldots \boldsymbol{\gamma}_n!} \mathbf{a}_1^{\boldsymbol{\gamma}_1} \ldots \mathbf{a}_n^{\boldsymbol{\gamma}_n}, \qquad (u\mathbf{a})^{\boldsymbol{\alpha}} = u^{|\boldsymbol{\alpha}|}\mathbf{a}^{\boldsymbol{\alpha}}.$$

Letting $\mathbf{a}_j$ be the column vectors of A and u_j the elements of $\mathbf{u}$, we then have

$$(A\mathbf{u})^{\boldsymbol{\alpha}} = (u_1\mathbf{a}_1 + u_2\mathbf{a}_2 + \ldots + u_n\mathbf{a}_n)^{\boldsymbol{\alpha}}$$

$$= \sum_{\boldsymbol{\gamma}_1 + \ldots + \boldsymbol{\gamma}_n = \boldsymbol{\alpha}} \frac{\boldsymbol{\alpha}!}{\boldsymbol{\gamma}_1! \ldots \boldsymbol{\gamma}_n!} (u_1\mathbf{a}_1)^{\boldsymbol{\gamma}_1} \ldots (u_n\mathbf{a}_n)^{\boldsymbol{\gamma}_n}$$

$$= \sum_{\boldsymbol{\gamma}_1 + \ldots + \boldsymbol{\gamma}_n = \boldsymbol{\alpha}} \frac{\boldsymbol{\alpha}!}{\boldsymbol{\gamma}_1! \ldots \boldsymbol{\gamma}_n!} \mathbf{a}_1^{\boldsymbol{\gamma}_1} \ldots \mathbf{a}_n^{\boldsymbol{\gamma}_n} u_1^{|\boldsymbol{\gamma}_1|} \ldots u_n^{|\boldsymbol{\gamma}_n|}.$$

Hence, $(A\mathbf{u})^{\boldsymbol{\alpha}}$ will be of the form

$$(A\mathbf{u})^{\boldsymbol{\alpha}} = \sum_{|\boldsymbol{\beta}|=|\boldsymbol{\alpha}|} c_{\boldsymbol{\beta}, \boldsymbol{\alpha}}\mathbf{u}^{\boldsymbol{\beta}},$$

where $c_{\boldsymbol{\beta}, \boldsymbol{\alpha}} \in \mathcal{V}$.

5.2 Derivatives of the quantity of interest

By (14), differentiating $\mathcal{Q}_{\mathrm{GB}}^\varepsilon$ with respect to $\mathbf{y}$ gives

$$\partial_{\mathbf{y}}^{\boldsymbol{\sigma}} \mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y}) = \left(\frac{1}{2\pi\varepsilon}\right)^n \int_{K_0 \times K_0} \partial_{\mathbf{y}}^{\boldsymbol{\sigma}} I^\varepsilon(T, \mathbf{y}; \mathbf{z}, \mathbf{z}')\, d\mathbf{z}d\mathbf{z}'. \qquad (26)$$

In this section we first rewrite the integral I^ε (Proposition 2) employing Lemma 3, and phase Ψ (Lemma 4). The form of $\partial_{\mathbf{y}}^{\boldsymbol{\sigma}} I^\varepsilon$ then follows from Proposition 3.

Lemma 3 *There exists $f_{\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\alpha},\boldsymbol{\beta}} \in \mathcal{V}$ such that*

$$(\mathbf{x} - \Delta\mathbf{q})^{\boldsymbol{\alpha}} (\mathbf{x} + \Delta\mathbf{q})^{\boldsymbol{\beta}} = \sum_{|\boldsymbol{\mu}+\boldsymbol{\nu}|=|\boldsymbol{\alpha}+\boldsymbol{\beta}|} f_{\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\alpha},\boldsymbol{\beta}}(t,\mathbf{y};\mathbf{z},\mathbf{z}')(\mathbf{z} - \mathbf{z}')^{\boldsymbol{\mu}}\mathbf{x}^{\boldsymbol{\nu}}. \qquad (27)$$

Proof Let us recall the definition of $\Delta\mathbf{q}$ in (15). From (P1), $\mathbf{q} \in C^{\infty}(\mathbb{R}^n \times \Gamma \times \mathbb{R}^n)$ and hence according to (24), we conclude that

$$\Delta\mathbf{q}(t,\mathbf{y};\mathbf{z},\mathbf{z}') = Q(t,\mathbf{y};\mathbf{z},\mathbf{z}')(\mathbf{z} - \mathbf{z}'),$$

with $Q \in \mathcal{V}$. From (25), we then have

$$\Delta\mathbf{q}^{\boldsymbol{\alpha}} = (Q(t,\mathbf{y};\mathbf{z},\mathbf{z}')(\mathbf{z} - \mathbf{z}'))^{\boldsymbol{\alpha}} = \sum_{|\boldsymbol{\sigma}|=|\boldsymbol{\alpha}|} f_{\boldsymbol{\sigma},\boldsymbol{\alpha}}(t,\mathbf{y};\mathbf{z},\mathbf{z}')(\mathbf{z} - \mathbf{z}')^{\boldsymbol{\sigma}}, \qquad (28)$$

for some $f_{\boldsymbol{\sigma},\boldsymbol{\alpha}} \in \mathcal{V}$. By the generalized binomial theorem and (28), $(\mathbf{x} - \Delta\mathbf{q})^{\boldsymbol{\alpha}}$ reads

$$(\mathbf{x} - \Delta\mathbf{q})^{\boldsymbol{\alpha}} = \sum_{\boldsymbol{\beta}+\boldsymbol{\gamma}=\boldsymbol{\alpha}} \binom{\boldsymbol{\alpha}}{\boldsymbol{\gamma}} (-1)^{\boldsymbol{\beta}} \Delta\mathbf{q}^{\boldsymbol{\beta}}\mathbf{x}^{\boldsymbol{\gamma}} = \sum_{|\boldsymbol{\sigma}+\boldsymbol{\gamma}|=|\boldsymbol{\alpha}|} h_{\boldsymbol{\sigma},\boldsymbol{\gamma},\boldsymbol{\alpha}}(t,\mathbf{y};\mathbf{z},\mathbf{z}') (\mathbf{z} - \mathbf{z}')^{\boldsymbol{\sigma}}\mathbf{x}^{\boldsymbol{\gamma}},$$

where $h_{\boldsymbol{\sigma},\boldsymbol{\gamma},\boldsymbol{\alpha}} \in \mathcal{V}$. The terms $(\mathbf{x} + \Delta\mathbf{q})^{\boldsymbol{\beta}}$ can be expanded in a similar manner:

$$(\mathbf{x} + \Delta\mathbf{q})^{\boldsymbol{\beta}} = \sum_{|\boldsymbol{\delta}+\boldsymbol{\omega}|=|\boldsymbol{\beta}|} \widetilde{h}_{\boldsymbol{\delta},\boldsymbol{\omega},\boldsymbol{\beta}}(t,\mathbf{y};\mathbf{z},\mathbf{z}') (\mathbf{z} - \mathbf{z}')^{\boldsymbol{\delta}}\mathbf{x}^{\boldsymbol{\omega}},$$

for some $\widetilde{h}_{\boldsymbol{\delta},\boldsymbol{\omega},\boldsymbol{\alpha}} \in \mathcal{V}$. Then

$$
\begin{aligned}
(\mathbf{x} - \Delta\mathbf{q})^{\boldsymbol{\alpha}} (\mathbf{x} + \Delta\mathbf{q})^{\boldsymbol{\beta}} &= \sum_{|\boldsymbol{\sigma}+\boldsymbol{\gamma}|=|\boldsymbol{\alpha}|} \sum_{|\boldsymbol{\delta}+\boldsymbol{\omega}|=|\boldsymbol{\beta}|} h_{\boldsymbol{\sigma},\boldsymbol{\gamma},\boldsymbol{\alpha}}\widetilde{h}_{\boldsymbol{\delta},\boldsymbol{\omega},\boldsymbol{\beta}} (\mathbf{z} - \mathbf{z}')^{\boldsymbol{\sigma}+\boldsymbol{\delta}}\mathbf{x}^{\boldsymbol{\gamma}+\boldsymbol{\omega}} \\
&= \sum_{|\boldsymbol{\mu}+\boldsymbol{\nu}|=|\boldsymbol{\alpha}+\boldsymbol{\beta}|} f_{\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\alpha},\boldsymbol{\beta}}(\mathbf{z} - \mathbf{z}')^{\boldsymbol{\mu}}\mathbf{x}^{\boldsymbol{\nu}},
\end{aligned}
$$

where $f_{\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\alpha},\boldsymbol{\beta}} \in \mathcal{V}$. This concludes the proof.

Utilizing the results of Lemma 3 we are ready to reformulate I^{ε} in (16) in terms of $\mathbf{x}$ and $\mathbf{z} - \mathbf{z}'$.

Proposition 2 *The quantity I^{ε} in (16) can be rewritten as*

$$I^{\varepsilon}(T,\mathbf{y};\mathbf{z},\mathbf{z}') = \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (29)$$
$$\sum_{j=0}^{2\lceil k/2\rceil-2} \sum_{|\boldsymbol{\mu}+\boldsymbol{\nu}|=0}^{2(k-j-1)} \varepsilon^{j}(\mathbf{z} - \mathbf{z}')^{\boldsymbol{\mu}} \int_{\mathbb{R}^n} \mathbf{x}^{\boldsymbol{\nu}}\, g_{j,\boldsymbol{\mu},\boldsymbol{\nu}}(T,\mathbf{x},\mathbf{y};\mathbf{z},\mathbf{z}')\, e^{i\Psi(T,\mathbf{x},\mathbf{y};\mathbf{z},\mathbf{z}')/\varepsilon}\, d\mathbf{x},$$

with $g_{j,\boldsymbol{\mu},\boldsymbol{\nu}} \in \mathcal{U}_{\eta}$.

Proof We start by noting that the expression for I^{ε} in (16) can be simplified to

$$I^{\varepsilon}(T,\mathbf{y};\mathbf{z},\mathbf{z}') = \sum_{j=0}^{2\lceil k/2\rceil-2} \sum_{|\boldsymbol{\alpha}+\boldsymbol{\beta}|=0}^{2(k-j-1)} \varepsilon^{j} \int_{\mathbb{R}^n} \bar{g}_{j,\boldsymbol{\alpha},\boldsymbol{\beta}} (\mathbf{x} - \Delta\mathbf{q})^{\boldsymbol{\alpha}}(\mathbf{x} + \Delta\mathbf{q})^{\boldsymbol{\beta}}e^{i\Psi/\varepsilon}\, d\mathbf{x},$$

where $\bar{g}_{j,\boldsymbol{\alpha},\boldsymbol{\beta}} \in \mathcal{U}_\eta$ are either sums of $\bar{g}_{j,m,\boldsymbol{\alpha},\boldsymbol{\beta}} \in \mathcal{U}_\eta$ in (18) or identically zero. Recalling Lemma 3, this can be reformulated as

$$
\begin{aligned}
I^\varepsilon(T, \mathbf{y}; \mathbf{z}, \mathbf{z}') & \\
&= \sum_{j=0}^{2\lceil k/2 \rceil - 2} \sum_{|\boldsymbol{\alpha}+\boldsymbol{\beta}|=0}^{2(k-j-1)} \varepsilon^j \int_{\mathbb{R}^n} \sum_{|\boldsymbol{\mu}+\boldsymbol{\nu}|=|\boldsymbol{\alpha}+\boldsymbol{\beta}|} \bar{g}_{j,\boldsymbol{\alpha},\boldsymbol{\beta}} f_{\boldsymbol{\mu},\boldsymbol{\nu},\boldsymbol{\alpha},\boldsymbol{\beta}} (\mathbf{z}-\mathbf{z}')^{\boldsymbol{\mu}} \mathbf{x}^{\boldsymbol{\nu}}\, e^{i\Psi/\varepsilon}\, d\mathbf{x} \\
&=: \sum_{j=0}^{2\lceil k/2 \rceil - 2} \sum_{|\boldsymbol{\mu}+\boldsymbol{\nu}|=0}^{2(k-j-1)} \varepsilon^j (\mathbf{z}-\mathbf{z}')^{\boldsymbol{\mu}} \int_{\mathbb{R}^n} g_{j,\boldsymbol{\mu},\boldsymbol{\nu}} \mathbf{x}^{\boldsymbol{\nu}}\, e^{i\Psi/\varepsilon}\, d\mathbf{x},
\end{aligned}
$$

where $g_{j,\boldsymbol{\mu},\boldsymbol{\nu}} \in \mathcal{U}_\eta$. This concludes the proof.

Similarly as for I^ε, let us first express the Gaussian beam phase Ψ in terms of $\mathbf{x}$ and $\mathbf{z} - \mathbf{z}'$.

Lemma 4 *For the k-th order Gaussian beam phase Ψ, there exist $g_{\boldsymbol{\alpha},\boldsymbol{\beta}} \in \mathcal{V}$ such that*

$$
\Psi(T, \mathbf{x}, \mathbf{y}; \mathbf{z}, \mathbf{z}') = \sum_{|\boldsymbol{\alpha}+\boldsymbol{\beta}|=2}^{k+1} g_{\boldsymbol{\alpha},\boldsymbol{\beta}}(T, \mathbf{y}; \mathbf{z}, \mathbf{z}')\, (\mathbf{z}-\mathbf{z}')^{\boldsymbol{\alpha}}\, \mathbf{x}^{\boldsymbol{\beta}}. \tag{30}
$$

Proof Recall the definition of Ψ in (17). For k-th order beams, we write (dropping T)

$$
\begin{aligned}
\Psi(\mathbf{x}, \mathbf{y}; \mathbf{z}, \mathbf{z}') = {}& F(\mathbf{y}; \mathbf{z}, \mathbf{z}') + \mathbf{x}^T[\mathbf{p}(\mathbf{y}; \mathbf{z}) - \mathbf{p}(\mathbf{y}; \mathbf{z}')] + \\
& + \frac{1}{2}(\mathbf{x} - \Delta\mathbf{q})^T M(\mathbf{y}; \mathbf{z})(\mathbf{x} - \Delta\mathbf{q}) - \frac{1}{2}(\mathbf{x} + \Delta\mathbf{q})^T M^*(\mathbf{y}; \mathbf{z}')(\mathbf{x} + \Delta\mathbf{q}) + \\
& + \sum_{|\boldsymbol{\gamma}|=3}^{k+1} \frac{1}{\boldsymbol{\gamma}!} \phi_{\boldsymbol{\gamma}}(\mathbf{z})(\mathbf{x} - \Delta\mathbf{q})^{\boldsymbol{\gamma}} - \sum_{|\boldsymbol{\gamma}|=3}^{k+1} \frac{1}{\boldsymbol{\gamma}!} \phi_{\boldsymbol{\gamma}}^*(\mathbf{z}')(\mathbf{x} + \Delta\mathbf{q})^{\boldsymbol{\gamma}}, \tag{31}
\end{aligned}
$$

where

$$
F(\mathbf{y}; \mathbf{z}, \mathbf{z}') = \phi_0(\mathbf{y}; \mathbf{z}) - \phi_0(\mathbf{y}; \mathbf{z}') - \frac{1}{2}(\mathbf{q}(\mathbf{y}; \mathbf{z}) - \mathbf{q}(\mathbf{y}; \mathbf{z}'))(\mathbf{p}(\mathbf{y}; \mathbf{z}) + \mathbf{p}(\mathbf{y}; \mathbf{z}')).
$$

We show that all the terms of Ψ in (31) are of the form (30). First, we note that for all $\mathbf{y}$

$$
F(\mathbf{y}; \mathbf{z}, \mathbf{z}) = 0, \qquad \nabla_{\mathbf{z}'} F(\mathbf{y}; \mathbf{z}, \mathbf{z}) = -\nabla_{\mathbf{z}} \phi_0(\mathbf{y}; \mathbf{z}) + D_{\mathbf{z}}\mathbf{q}(T, \mathbf{y}; \mathbf{z})^T \mathbf{p}(\mathbf{y}; \mathbf{z}) = 0,
$$

where the second equality follows from (P5). By (P1), $F \in \mathcal{V}$, hence due to (23) there exists a function $g_{\boldsymbol{\alpha}} \in \mathcal{V}$ such that,

$$
F(\mathbf{y}; \mathbf{z}, \mathbf{z}') = \sum_{|\boldsymbol{\alpha}|=2} g_{\boldsymbol{\alpha}}(\mathbf{y}; \mathbf{z}, \mathbf{z}')\, (\mathbf{z}-\mathbf{z}')^{\boldsymbol{\alpha}}.
$$

The second term in (31) includes $\mathbf{x}, \mathbf{p}$ that are both in $C^\infty(\mathbb{R} \times \Gamma \times \mathbb{R}^n)$ by (P1) and hence (24) can be applied:

$$
\mathbf{x}^T(\mathbf{p}(\mathbf{y}; \mathbf{z}) - \mathbf{p}(\mathbf{y}; \mathbf{z}')) = \mathbf{x}^T G(\mathbf{y}; \mathbf{z}, \mathbf{z}')(\mathbf{z} - \mathbf{z}') = \sum_{|\boldsymbol{\alpha}|=|\boldsymbol{\beta}|=1} h_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\mathbf{y}; \mathbf{z}, \mathbf{z}')\, (\mathbf{z}-\mathbf{z}')^{\boldsymbol{\alpha}}\, \mathbf{x}^{\boldsymbol{\beta}},
$$

with $h_{\boldsymbol{\alpha},\boldsymbol{\beta}} \in \mathcal{V}$. Similarly,

$$
\begin{aligned}
(\mathbf{x} - \Delta\mathbf{q})^T M(\mathbf{y};\mathbf{z})(\mathbf{x} - \Delta\mathbf{q}) &= \sum_{|\boldsymbol{\alpha}|=2} m_{\boldsymbol{\alpha}}(\mathbf{y};\mathbf{z},\mathbf{z}')(\mathbf{x} - \Delta\mathbf{q})^{\boldsymbol{\alpha}} \\
&= \sum_{|\boldsymbol{\alpha}+\boldsymbol{\beta}|=2} \ell_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\mathbf{y};\mathbf{z},\mathbf{z}')\,(\mathbf{z}-\mathbf{z}')^{\boldsymbol{\alpha}}\,\mathbf{x}^{\boldsymbol{\beta}},
\end{aligned}
$$

where $m_{\boldsymbol{\alpha}}, \ell_{\boldsymbol{\alpha},\boldsymbol{\beta}} \in \mathcal{V}$. Again, we used (P1) and (24). The term $(\mathbf{x}+\Delta\mathbf{q})^T M^*(\mathbf{y};\mathbf{z})(\mathbf{x}+\Delta\mathbf{q})$ can be formulated in a similar fashion. The last two terms in (31) include the quantity $(\mathbf{x} \pm \Delta\mathbf{q})^{\boldsymbol{\gamma}}$, which can be expanded by (27), namely

$$
\sum_{|\boldsymbol{\gamma}|=3}^{k+1} \frac{1}{\boldsymbol{\gamma}!} \phi_{\boldsymbol{\gamma}}(\mathbf{z})(\mathbf{x} - \Delta\mathbf{q})^{\boldsymbol{\gamma}} = \sum_{|\boldsymbol{\alpha}+\boldsymbol{\beta}|=3}^{k+1} p_{\boldsymbol{\alpha},\boldsymbol{\beta}}(\mathbf{y};\mathbf{z},\mathbf{z}')\,(\mathbf{z}-\mathbf{z}')^{\boldsymbol{\alpha}}\,\mathbf{x}^{\boldsymbol{\beta}},
$$

with $p_{\boldsymbol{\alpha},\boldsymbol{\beta}} \in \mathcal{V}$. Analogously, $\sum_{|\boldsymbol{\gamma}|=3}^{k+1} \frac{1}{\boldsymbol{\gamma}!} \phi_{\boldsymbol{\gamma}}^*(\mathbf{z}')(\mathbf{x} + \Delta\mathbf{q})^{\boldsymbol{\gamma}}$ is of the same form.

To summarize, all the terms of (31) are indeed of the form (30). Their sums are finite hence the union is also finite, which concludes the claim of the lemma.

The first step in proving Theorem 1 is to express $\partial_{\mathbf{y}}^{\boldsymbol{\sigma}} I^\varepsilon(T,\mathbf{y};\mathbf{z},\mathbf{z}')$ using the lemmas above.

Proposition 3 *There exist functions $b_{j,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}} \in \mathcal{U}_\eta$ such that the derivatives of I^ε in (16) with respect to $\mathbf{y}$ read:*

$$
\partial_{\mathbf{y}}^{\boldsymbol{\sigma}} I^\varepsilon(T,\mathbf{y};\mathbf{z},\mathbf{z}') = \tag{32}
$$
$$
\sum_{j=-|\boldsymbol{\sigma}|}^{2\lceil k/2\rceil-2} \sum_{|\boldsymbol{\alpha}+\boldsymbol{\beta}|=\max(-2j,0)}^{(|\boldsymbol{\sigma}|+2)(k-1)-2j} \varepsilon^j (\mathbf{z}-\mathbf{z}')^{\boldsymbol{\alpha}} \int_{\mathbb{R}^n} \mathbf{x}^{\boldsymbol{\beta}} b_{j,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}}(T,\mathbf{x},\mathbf{y};\mathbf{z},\mathbf{z}') e^{i\Psi/\varepsilon} d\mathbf{x}.
$$

Proof The proof will be carried out by mathematical induction. Let us define an index set $\mathcal{J}^{\boldsymbol{\sigma}}$ which corresponds to the double sum in (32), namely

$$
\mathcal{J}^{\boldsymbol{\sigma}} = \{(j,\boldsymbol{\alpha},\boldsymbol{\beta}) : -|\boldsymbol{\sigma}| \le j \le 2(\lceil k/2\rceil - 1),\quad 0 \le |\boldsymbol{\alpha}+\boldsymbol{\beta}| + 2j \le (|\boldsymbol{\sigma}|+2)(k-1)\},
$$

for all $\boldsymbol{\sigma} \in \mathbb{N}_0^n$. First, we show that the statement (32) is valid for $\boldsymbol{\sigma} \equiv 0$. Next, while assuming that the claim holds for $\boldsymbol{\sigma} \in \mathbb{N}_0^n$ we prove that it is valid for any $\bar{\boldsymbol{\sigma}} = \boldsymbol{\sigma} + \mathbf{e}_\ell$.

1. $\boldsymbol{\sigma} = 0$: The sums in (29) are the same as the ones in (32) in this case, with $b_{j,\boldsymbol{\alpha},\boldsymbol{\beta},0} = g_{j,\boldsymbol{\alpha},\boldsymbol{\beta}}$.
2. $\bar{\boldsymbol{\sigma}} = \boldsymbol{\sigma} + \mathbf{e}_\ell$: We assume (32) to hold for $\boldsymbol{\sigma} \in \mathbb{N}_0^n$. Then

$$
\begin{aligned}
\partial_{\mathbf{y}}^{\bar{\boldsymbol{\sigma}}} I^\varepsilon(T,\mathbf{y};\mathbf{z},\mathbf{z}') &= \partial_{y_\ell}\left(\partial_{\mathbf{y}}^{\boldsymbol{\sigma}} I^\varepsilon(T,\mathbf{y};\mathbf{z},\mathbf{z}')\right) \\
&= \sum_{(j,\boldsymbol{\alpha},\boldsymbol{\beta})\in\mathcal{J}^{\boldsymbol{\sigma}}} \varepsilon^j (\mathbf{z}-\mathbf{z}')^{\boldsymbol{\alpha}} \int_{\mathbb{R}^n} \mathbf{x}^{\boldsymbol{\beta}}\, \partial_{y_\ell} b_{j,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}}(T,\mathbf{x},\mathbf{y};\mathbf{z},\mathbf{z}')\, e^{i\Psi/\varepsilon}\, d\mathbf{x} + \tag{33}
\end{aligned}
$$
$$
+ \sum_{(j,\boldsymbol{\alpha},\boldsymbol{\beta})\in\mathcal{J}^{\boldsymbol{\sigma}}} \varepsilon^j (\mathbf{z}-\mathbf{z}')^{\boldsymbol{\alpha}} \int_{\mathbb{R}^n} \mathbf{x}^{\boldsymbol{\beta}}\, b_{j,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}}(T,\mathbf{x},\mathbf{y};\mathbf{z},\mathbf{z}')\, \partial_{y_\ell} e^{i\Psi/\varepsilon}\, d\mathbf{x}. \tag{34}
$$

In (33), $\partial_{y_\ell} b_{j,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}} \in \mathcal{U}_\eta$ and $\mathcal{J}^{\boldsymbol{\sigma}} \subset \mathcal{J}^{\bar{\boldsymbol{\sigma}}}$. The derivative $\partial_{y_\ell} e^{i\Psi/\varepsilon}$ in (34) can be expressed from Lemma 4 as

$$\partial_{y_\ell} e^{i\Psi/\varepsilon} = \frac{i}{\varepsilon} e^{i\Psi/\varepsilon} \partial_{y_\ell} \Psi = \sum_{|\boldsymbol{\gamma}+\boldsymbol{\delta}|=2}^{k+1} \varepsilon^{-1} \, (\mathbf{z}-\mathbf{z}')^{\boldsymbol{\gamma}} \, \mathbf{x}^{\boldsymbol{\delta}} \, h_{\boldsymbol{\gamma},\boldsymbol{\delta},\mathbf{e}_\ell}(T,\mathbf{y};\mathbf{z},\mathbf{z}') \, e^{i\Psi/\varepsilon},$$

with $h_{\boldsymbol{\gamma},\boldsymbol{\delta},\mathbf{e}_\ell} = i\, \partial_{y_\ell} g_{\boldsymbol{\gamma},\boldsymbol{\delta}} \in \mathcal{V}$, with $g_{\boldsymbol{\gamma},\boldsymbol{\delta}}$ from (30). The terms in (34) are hence of the form

$$\varepsilon^{\bar{j}} (\mathbf{z}-\mathbf{z}')^{\bar{\boldsymbol{\alpha}}} \int_{\mathbb{R}^n} \mathbf{x}^{\bar{\boldsymbol{\beta}}} \, g(T,\mathbf{x},\mathbf{y};\mathbf{z},\mathbf{z}') \, e^{i\Psi/\varepsilon} d\mathbf{x}$$

with $g := b_{j,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}} \, h_{\boldsymbol{\gamma},\boldsymbol{\delta},\mathbf{e}_\ell} \in \mathcal{U}_\eta$ and $\bar{j} = j-1$, $\bar{\boldsymbol{\alpha}} = \boldsymbol{\alpha}+\boldsymbol{\gamma}$, $\bar{\boldsymbol{\beta}} = \boldsymbol{\beta}+\boldsymbol{\delta}$. Note that $2 \le |\boldsymbol{\gamma}+\boldsymbol{\delta}| \le k+1$ and since $(j,\boldsymbol{\alpha},\boldsymbol{\beta}) \in \mathcal{J}^{\boldsymbol{\sigma}}$,

$$0 \le |\boldsymbol{\alpha}+\boldsymbol{\beta}| + 2j + |\boldsymbol{\gamma}+\boldsymbol{\delta}| - 2 = |\bar{\boldsymbol{\alpha}}+\bar{\boldsymbol{\beta}}| + 2\bar{j} \le (|\boldsymbol{\sigma}|+3)(k-1),$$

where $-|\boldsymbol{\sigma}| - 1 \le \bar{j} \le 2\lceil k/2 \rceil - 2$. Therefore $(\bar{j}, \bar{\boldsymbol{\alpha}}, \bar{\boldsymbol{\beta}}) \in \mathcal{J}^{\bar{\boldsymbol{\sigma}}}$ which concludes the induction and proves Proposition 3.

5.3 Estimates for derivatives of the quantity of interest

We build the proof of Theorem 1 on estimates presented in this section for the derivatives of the quantity of interest (13). We first note that the n-dimensional Gaussian integral can be computed as

$$\int_{\mathbb{R}^n} |\mathbf{x}|^{|\boldsymbol{\alpha}|} e^{-a|\mathbf{x}|^2/\varepsilon} \, d\mathbf{x} = C\varepsilon^{n/2+|\boldsymbol{\alpha}|/2}, \tag{35}$$

where C does not depend on ε. We will use this result in several places below.

Lemma 5 *Let $g \in \mathcal{U}_\eta$, and η be admissible in the sense of Definition 1. Then, for any $t \in [0,T]$, $\boldsymbol{\beta} \in \mathbb{N}_0^n$, there is a constant C and $\delta > 0$, independent of t, $\mathbf{z}$, $\mathbf{z}'$, $\mathbf{y}$ and ε such that*

$$\left| \int_{\mathbb{R}^n} \mathbf{x}^{\boldsymbol{\beta}} \, g(t,\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{z}') \, e^{i\Psi(t,\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{z}')/\varepsilon} \, d\mathbf{x} \right| \le C\, \varepsilon^{\frac{n+|\boldsymbol{\beta}|}{2}} e^{-\delta|\Delta\mathbf{q}(t,\mathbf{y};\mathbf{z},\mathbf{z}')|^2/\varepsilon}. \tag{36}$$

Proof Integral (36) reads

$$I := \left| \int_{\mathbb{R}^n} \mathbf{x}^{\boldsymbol{\beta}} \, g \, e^{i\Psi/\varepsilon} \, d\mathbf{x} \right| = \left| \int_{\Omega_{2\eta}} \mathbf{x}^{\boldsymbol{\beta}} \, g \, e^{i\Psi/\varepsilon} \, d\mathbf{x} \right|.$$

By (P4) and since g is uniformly bounded in $\mathbf{y}$ by (A4), we have

$$I \le \int_{\Omega_{2\eta}} \mathbf{x}^{\boldsymbol{\beta}} \, |g| \, e^{-\mathrm{Im}\, \Psi/\varepsilon} \, d\mathbf{x} \le C\, e^{-\frac{\delta}{\varepsilon}|\Delta\mathbf{q}|^2} \int_{\Omega_{2\eta}} |\mathbf{x}|^{|\boldsymbol{\beta}|} \, e^{-\frac{\delta}{\varepsilon}|\mathbf{x}|^2} \, d\mathbf{x}$$
$$\le C\, \varepsilon^{\frac{n+|\boldsymbol{\beta}|}{2}} \, e^{-\frac{\delta}{\varepsilon}|\Delta\mathbf{q}|^2},$$

where we recalled (35) in the last inequality. We note that the constant C may depend on n and $\boldsymbol{\beta}$, but is independent of $\mathbf{z}$, $\mathbf{z}'$, $\mathbf{y}$ and ε.

Lemma 6 *Let $g \in \mathcal{U}_\eta$, and η be admissible in the sense of Definition 1. Suppose further that for fixed $t \in [0, T]$ and $\mathbf{y} \in \Gamma$ it holds that $|\mathbf{q}(t, \mathbf{y}; \mathbf{z}) - \mathbf{q}(t, \mathbf{y}; \mathbf{z}')| \leq \theta|\mathbf{z} - \mathbf{z}'|$ for some $\theta > 0$ and $\mathbf{z}, \mathbf{z}' \in K_0$. Then, for any $\boldsymbol{\beta} \in \mathbb{N}_0^n$, $K \in \mathbb{N}_0$, and if θ is small enough, there is a constant C, independent of t, $\mathbf{z}$, $\mathbf{z}'$, $\mathbf{y}$ and ε such that*

$$\left| \int_{\mathbb{R}^n} \mathbf{x}^{\boldsymbol{\beta}} \, g(t, \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{z}') \, e^{i \Psi(t, \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{z}')/\varepsilon} \, d\mathbf{x} \right| \leq C \, \varepsilon^{\frac{n+|\boldsymbol{\beta}|}{2}} \frac{1}{1 + \left| \frac{\mathbf{z} - \mathbf{z}'}{\sqrt{\varepsilon}} \right|^K}. \qquad (37)$$

Proof Take μ and θ such that they are small enough in the sense of (P6), and also $\mu < \eta$. We simplify our notation by omitting the functions' arguments and introduce two smooth test functions $\varphi_1, \varphi_2 \in C_c^\infty(\mathbb{R}^n)$ with the following properties:

$$\varphi_1(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in \Omega_{\mu/2}, \\ 0 & \mathbf{x} \in \Omega_\eta \setminus \Omega_\mu, \end{cases} \qquad \varphi_2(\mathbf{x}) = 1 - \varphi_1(\mathbf{x}),$$

and let φ_1 be monotonic on $\Omega_\mu \setminus \Omega_{\mu/2}$. By forming a smooth partition of unity, we split the function g into two parts:

$$g = (\varphi_1 + \varphi_2) \, g = \varphi_1 \, g + \varphi_2 \, g =: g_1 + g_2,$$

where $\operatorname{supp}(g_1) = \Omega_\mu \subset \Omega_\eta$, which implies that $g_1 \in \mathcal{U}_\eta$, and $\operatorname{supp}(g_2) = \Omega_\eta \setminus \Omega_{\mu/2}$; see Figure 4.
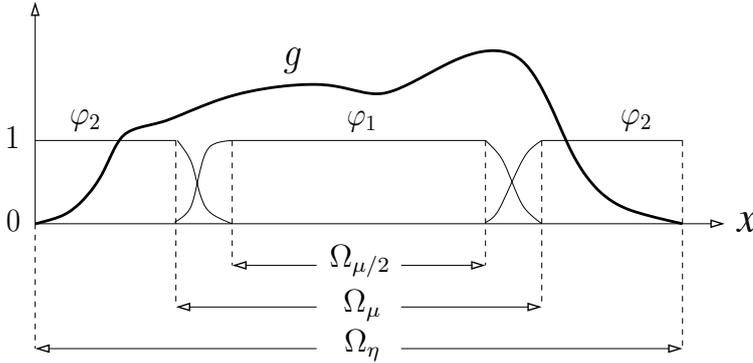
Fig. 4 A one-dimensional schematic representation of partitioning the function g.

Accordingly, we can split the integral on the left hand side of the inequality in (37) into two parts and using the triangular inequality write:

$$I := \left| \int_{\mathbb{R}^n} \mathbf{x}^{\boldsymbol{\beta}} \, g \, e^{i \Psi/\varepsilon} \, d\mathbf{x} \right| \leq \left| \int_{\Omega_\mu} \mathbf{x}^{\boldsymbol{\beta}} \, g_1 \, e^{i \Psi/\varepsilon} \, d\mathbf{x} \right| + \left| \int_{\Omega_\eta \setminus \Omega_{\mu/2}} \mathbf{x}^{\boldsymbol{\beta}} \, g_2 \, e^{i \Psi/\varepsilon} \, d\mathbf{x} \right| =: I_1 + I_2.$$

We first estimate I_1. We recall Lemma 5 and write

$$I_1 \leq C \, \varepsilon^{\frac{n+|\boldsymbol{\beta}|}{2}} e^{-\frac{\delta}{\varepsilon} |\Delta \mathbf{q}|^2} \leq C \, \varepsilon^{\frac{n+|\boldsymbol{\beta}|}{2}}.$$

We note that the constant C may depend on $(n, \boldsymbol{\beta})$, but is independent of $\mathbf{y}$ and ε. This proves the needed estimate for I_1 for the case $K = 0$ as well as the case

$K > 0$ when $\mathbf{z} = \mathbf{z}'$. Therefore, in the remainder of the proof we will consider the case $K \neq 0$ and $\mathbf{z} \neq \mathbf{z}'$.

From (P6), since μ and θ are small enough, there exists a constant $C > 0$ such that $\inf_{\mathbf{x} \in \Omega_\mu} |\nabla \Psi| \geq C |\mathbf{z} - \mathbf{z}'| > 0$. Since the gradient of Ψ does not vanish, we employ the non-stationary phase lemma (see [10]). Since $g_1 \in \mathcal{U}_\mu$ we get

$$
\begin{aligned}
I_1 &\leq C_K \, \varepsilon^K \sum_{|\mathbf{m}| \leq K} \int_{\Omega_\mu} \frac{|\partial_{\mathbf{x}}^{\mathbf{m}} (\mathbf{x}^{\boldsymbol{\beta}} g_1)|}{|\nabla \Psi|^{2K - |\mathbf{m}|}} \, e^{-\operatorname{Im} \Psi / \varepsilon} \, d\mathbf{x} \\
&\leq C_K \, \varepsilon^K \sum_{|\mathbf{m}| \leq K} |\mathbf{z} - \mathbf{z}'|^{|\mathbf{m}| - 2K} \underbrace{\int_{\Omega_\mu} |\partial_{\mathbf{x}}^{\mathbf{m}} (\mathbf{x}^{\boldsymbol{\beta}} g_1)| \, e^{-\operatorname{Im} \Psi / \varepsilon} \, d\mathbf{x}}_{(*)},
\end{aligned}
$$

where we used (P6). Employing the general Leibniz rule, $(*)$ can be estimated as

$$
\begin{aligned}
(*) &\leq C \sum_{\substack{\mathbf{m}_1 + \mathbf{m}_2 = \mathbf{m} \\ \mathbf{m}_1 \leq \boldsymbol{\beta}}} \int_{\Omega_\mu} |\partial_{\mathbf{x}}^{\mathbf{m}_1} \mathbf{x}^{\boldsymbol{\beta}}| \, |\partial_{\mathbf{x}}^{\mathbf{m}_2} g_1| \, e^{-\operatorname{Im} \Psi / \varepsilon} \, d\mathbf{x} \\
&\leq C \sum_{\substack{\mathbf{m}_1 + \mathbf{m}_2 = \mathbf{m} \\ \mathbf{m}_1 \leq \boldsymbol{\beta}}} \int_{\Omega_\mu} |\mathbf{x}|^{|\boldsymbol{\beta}| - |\mathbf{m}_1|} \, e^{-\operatorname{Im} \Psi / \varepsilon} \, d\mathbf{x},
\end{aligned}
$$

where the last inequality follows because g_1 has bounded derivatives with respect to $\mathbf{x}$, independent of $\mathbf{y}$ and ε, which is a consequence of (A4) and the smoothness of φ_1 and g on the compact sets Ω_μ and Ω_η, respectively. Using (P4) and (35), we have

$$
\begin{aligned}
\int_{\Omega_\mu} |\mathbf{x}|^{|\boldsymbol{\beta}| - |\mathbf{m}_1|} \, e^{-\operatorname{Im} \Psi / \varepsilon} \, d\mathbf{x} &\leq e^{-\frac{\delta}{\varepsilon} |\Delta \mathbf{q}|^2} \int_{\Omega_\mu} |\mathbf{x}|^{|\boldsymbol{\beta}| - |\mathbf{m}_1|} \, e^{-\frac{\delta}{\varepsilon} |\mathbf{x}|^2} \, d\mathbf{x} \\
&\leq C \, \varepsilon^{\frac{n + |\boldsymbol{\beta}| - |\mathbf{m}_1|}{2}} \leq C \, \varepsilon^{\frac{n + |\boldsymbol{\beta}| - |\mathbf{m}|}{2}},
\end{aligned}
$$

where we used that $\mathbf{m}_1 \leq \mathbf{m}$ and (A5) in the last step. Consequently, we obtain

$$
I_1 \leq C_K \, \varepsilon^{\frac{n + |\boldsymbol{\beta}|}{2}} \sum_{|\mathbf{m}| \leq K} \left(\frac{|\mathbf{z} - \mathbf{z}'|}{\sqrt{\varepsilon}} \right)^{|\mathbf{m}| - 2K}.
$$

Now since when $K = 0$: $I_1 \leq C \, \varepsilon^{\frac{n + |\boldsymbol{\beta}|}{2}}$, in the general case of $K \geq 0$, we can write

$$
I_1 \leq C_K \, \varepsilon^{\frac{n + |\boldsymbol{\beta}|}{2}} \min\left(1, \sum_{|\mathbf{m}| \leq K} \left(\frac{|\mathbf{z} - \mathbf{z}'|}{\sqrt{\varepsilon}} \right)^{|\mathbf{m}| - 2K} \right).
$$

Since for positive numbers $a, b, c > 0$, the following inequalities hold:

$$
\min(a, b + c) \leq \min(a, b) + \min(a, c) \quad \text{and} \quad \min(1, 1/a) \leq 2/(1 + a),
$$

then

$$
I_1 \leq C_K \, \varepsilon^{\frac{n + |\boldsymbol{\beta}|}{2}} \sum_{|\mathbf{m}| \leq K} \frac{2}{1 + \left(\frac{|\mathbf{z} - \mathbf{z}'|}{\sqrt{\varepsilon}} \right)^{2K - |\mathbf{m}|}} \leq C_K \, \varepsilon^{\frac{n + |\boldsymbol{\beta}|}{2}} \frac{1}{1 + \left(\frac{|\mathbf{z} - \mathbf{z}'|}{\sqrt{\varepsilon}} \right)^K}.
$$

It remains to show that I_2 can be estimated in a similar way. By (P4) and since g_2 is uniformly bounded in $\mathbf{y}$ and by (A4), we have

$$
\begin{aligned}
I_2 &\leq C\, e^{-\frac{\delta}{\varepsilon}\,|\Delta\mathbf{q}|^2} \int_{\Omega_\eta\setminus\Omega_{\mu/2}} |\mathbf{x}|^{|\boldsymbol{\beta}|}\, e^{-\frac{\delta}{\varepsilon}\,|\mathbf{x}|^2}\, d\mathbf{x} \\
&= C\, e^{-\frac{\delta}{2\varepsilon}\,|\Delta\mathbf{q}|^2} \int_{\Omega_\eta\setminus\Omega_{\mu/2}} |\mathbf{x}|^{|\boldsymbol{\beta}|}\, e^{-\frac{\delta}{2\varepsilon}\,(|\mathbf{x}|^2+|\Delta\mathbf{q}|^2)} e^{-\frac{\delta}{2\varepsilon}\,|\mathbf{x}|^2}\, d\mathbf{x} \\
&\leq C \int_{\Omega_\eta\setminus\Omega_{\mu/2}} |\mathbf{x}|^{|\boldsymbol{\beta}|} e^{-\frac{\delta}{4\varepsilon}\,(|\mathbf{x}+\Delta\mathbf{q}|^2+|\mathbf{x}-\Delta\mathbf{q}|^2)} e^{-\frac{\delta}{2\varepsilon}\,|\mathbf{x}|^2}\, d\mathbf{x},
\end{aligned}
$$

where the last inequality follows from $|\mathbf{x}|^2 + |\Delta\mathbf{q}|^2 = |\mathbf{x} + \Delta\mathbf{q}|^2/2 + |\mathbf{x} - \Delta\mathbf{q}|^2/2$ and $e^{-\frac{\delta}{2\varepsilon}\,|\Delta\mathbf{q}|^2} \leq 1$. Since on $\Omega_\eta \setminus \Omega_{\mu/2}$, at least one of the following inequalities holds:

$$
|\mathbf{x} + \Delta\mathbf{q}| > \mu \quad \text{or} \quad |\mathbf{x} - \Delta\mathbf{q}| > \mu \quad \text{or} \quad |\mathbf{x} \pm \Delta\mathbf{q}| > \mu,
$$

we have

$$
I_2 \leq C e^{-\frac{\delta}{4\varepsilon}\,\mu^2} \int_{\Omega_\eta\setminus\Omega_{\mu/2}} |\mathbf{x}|^{|\boldsymbol{\beta}|} e^{-\frac{\delta}{2\varepsilon}\,|\mathbf{x}|^2}\, d\mathbf{x}.
$$

From (35), this yields

$$
I_2 \leq C e^{-\frac{\delta}{4\varepsilon}\,\mu^2}\, \varepsilon^{\frac{|\boldsymbol{\beta}|+n}{2}} \leq C_s\, \varepsilon^s, \quad \forall s > 0.
$$

This means that I_2 is of order $\mathcal{O}(\varepsilon^\infty)$. Let us find s such that I_2 satisfies the estimate (37). We have

$$
I_2 \leq C_s\, \varepsilon^s \leq C_s\, \varepsilon^s \frac{1 + \mathrm{diam}(K_0)^K}{\varepsilon^{K/2} + \mathrm{diam}(K_0)^K} = C_s\, \varepsilon^{s-K/2} \frac{1 + \mathrm{diam}(K_0)^K}{1 + \left(\frac{\mathrm{diam}(K_0)}{\sqrt{\varepsilon}}\right)^K},
$$

where $\mathrm{diam}(K_0) < \infty$ denotes the diameter of K_0, which is finite since K_0 is compact (see (A2)). As $|\mathbf{z} - \mathbf{z}'| \leq \mathrm{diam}(K_0)$, we get

$$
I_2 \leq \tilde{C}_s\, \frac{\varepsilon^{s-K/2}}{1 + \left(\frac{|\mathbf{z}-\mathbf{z}'|}{\sqrt{\varepsilon}}\right)^K},
$$

so we have the same estimate as (37) if we take $s = (K + n + |\boldsymbol{\beta}|)/2$. The proof of Lemma 6 is complete.

5.4 Proof of Theorem 1

Let us recall the form of $\partial_{\mathbf{y}}^{\boldsymbol{\sigma}} \mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y})$ in (26). We express $\partial_{\mathbf{y}}^{\boldsymbol{\sigma}} I^\varepsilon(\mathbf{y}; \mathbf{z}, \mathbf{z}')$ from Proposition 3 and estimate $|\partial_{\mathbf{y}}^{\boldsymbol{\sigma}} \mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y})|$, dropping the T-argument, as

$$
\begin{aligned}
\left| \partial_{\mathbf{y}}^{\boldsymbol{\sigma}} \mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y}) \right| &\leq \left(\frac{1}{2\pi\varepsilon}\right)^n \sup_{\mathbf{z}'\in K_0} \int_{K_0} \left| \partial_{\mathbf{y}}^{\boldsymbol{\sigma}} I^\varepsilon(\mathbf{y}; \mathbf{z}, \mathbf{z}') \right| d\mathbf{z} \\
&\leq \left(\frac{1}{2\pi\varepsilon}\right)^n \sup_{\mathbf{z}'\in K_0} \sum_{j=-|\boldsymbol{\sigma}|}^{2(\lceil k/2\rceil-1)} \sum_{|\boldsymbol{\alpha}+\boldsymbol{\beta}|=\max(-2j,0)}^{(|\boldsymbol{\sigma}|+2)(k-1)-2j} \\
&\qquad \varepsilon^j \underbrace{\int_{K_0} |\mathbf{z}-\mathbf{z}'|^{|\boldsymbol{\alpha}|} \left| \int_{\mathbb{R}^n} \mathbf{x}^{\boldsymbol{\beta}}\, b_{j,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}}(\mathbf{x}, \mathbf{y}; \mathbf{z}, \mathbf{z}')\, e^{i\Psi/\varepsilon}\, d\mathbf{x} \right| d\mathbf{z}}_{(*)}. \quad (38)
\end{aligned}
$$

To begin with $(*)$, let us first divide the set $K_0 \times K_0$ into two sets: the non-caustic region where $|\Delta \mathbf{q}| \geq \theta |\mathbf{z} - \mathbf{z}'|$ and the caustic region where $|\Delta \mathbf{q}| < \theta |\mathbf{z} - \mathbf{z}'|$, for small enough θ in the sense of (P6).

1. Non-caustic region $|\Delta \mathbf{q}| \geq \theta |\mathbf{z} - \mathbf{z}'|$.
   The function $b_{j,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}}$ in (38) belongs to $\mathcal{U}_\eta$ and we can apply the results of Lemma 5. Then $(*)$ reads:

$$\int_{K_0} |\mathbf{z} - \mathbf{z}'|^{|\boldsymbol{\alpha}|} \left| \int_{\mathbb{R}^n} \mathbf{x}^{\boldsymbol{\beta}} b_{j,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}}(\mathbf{x},\mathbf{y};\mathbf{z},\mathbf{z}') e^{i\Psi/\varepsilon} d\mathbf{x} \right| d\mathbf{z}$$

$$\leq C\varepsilon^{\frac{n+|\boldsymbol{\beta}|}{2}} \int_{K_0} |\mathbf{z} - \mathbf{z}'|^{|\boldsymbol{\alpha}|} e^{-\delta|\Delta\mathbf{q}|^2/\varepsilon} d\mathbf{z}$$

$$\leq C\varepsilon^{\frac{n+|\boldsymbol{\beta}|}{2}} \int_{K_0} |\mathbf{z} - \mathbf{z}'|^{|\boldsymbol{\alpha}|} e^{-\delta\theta^2|\mathbf{z}-\mathbf{z}'|^2/\varepsilon} d\mathbf{z} \leq C\varepsilon^{\frac{n+|\boldsymbol{\beta}|}{2}} \varepsilon^{\frac{n+|\boldsymbol{\alpha}|}{2}},$$

where we used the fact that $|\Delta \mathbf{q}| \geq \theta |\mathbf{z} - \mathbf{z}'|$ and recalled the Gaussian integral formula (35) in the last step. Consequently, (38) reduces to

$$\left| \partial_{\mathbf{y}}^{\boldsymbol{\sigma}} \mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y}) \right| \leq C_{\boldsymbol{\sigma}} \sum_{j=-|\boldsymbol{\sigma}|}^{2(\lceil k/2 \rceil - 1)} \sum_{|\boldsymbol{\alpha}+\boldsymbol{\beta}|=\max(-2j,0)}^{(|\boldsymbol{\sigma}|+2)(k-1)-2j} \varepsilon^{j+\frac{|\boldsymbol{\beta}|+|\boldsymbol{\alpha}|}{2}}. \tag{39}$$

2. Caustic region $|\Delta \mathbf{q}| < \theta |\mathbf{z} - \mathbf{z}'|$.
   Relation (38) fulfills the prerequisites of Lemma 6, hence $(*)$ reads, for all $K \geq 0$,

$$\int_{K_0} |\mathbf{z} - \mathbf{z}'|^{|\boldsymbol{\alpha}|} \left| \int_{\mathbb{R}^n} \mathbf{x}^{\boldsymbol{\beta}} b_{j,\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\sigma}}(\mathbf{x},\mathbf{y};\mathbf{z},\mathbf{z}') e^{i\Psi/\varepsilon} d\mathbf{x} \right| d\mathbf{z}$$

$$\leq C\varepsilon^{\frac{n+|\boldsymbol{\beta}|}{2}} \int_{K_0} \frac{|\mathbf{z} - \mathbf{z}'|^{|\boldsymbol{\alpha}|}}{1 + (|\mathbf{z}-\mathbf{z}'|/\sqrt{\varepsilon})^K} d\mathbf{z} \leq C\varepsilon^{\frac{n+|\boldsymbol{\beta}|}{2}} \varepsilon^{\frac{n+|\boldsymbol{\alpha}|}{2}} \int_{\mathbb{R}^n} \frac{|t|^{|\boldsymbol{\alpha}|}}{1 + |t|^K} dt.$$

The integral is bounded when $K > n + |\boldsymbol{\alpha}|$, hence we arrive at the same conclusion as in (39).

In both the caustic and non-caustic regions, the same inequality has been obtained. The exponent of ε is always positive as $0 \leq |\boldsymbol{\alpha} + \boldsymbol{\beta}| + 2j$, hence

$$\left| \partial_{\mathbf{y}}^{\boldsymbol{\sigma}} \mathcal{Q}_{\mathrm{GB}}^\varepsilon(\mathbf{y}) \right| \leq C_{\boldsymbol{\sigma}},$$

which concludes the proof of Theorem 1.

References

1. Babuska, I., Nobile, F., Tempone, R.: A stochastic collocation method for elliptic partial differential equations with random input data. SIAM review **52**(2), 317–355 (2010)
2. Babuska, I., Tempone, R., Zouraris, G.E.: Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation. Computer Methods in Applied Mechanics and Engineering **194**, 1251–1294 (2005)
3. Beck, J., Nobile, F., Tamellini, L., Tempone, R.: Convergence of quasi-optimal stochastic Galerkin methods for a class of PDEs with random coefficients. Computers and Mathematics with Applications **67**, 732–751 (2014)

4. Cervený, V., Popov, M.M., Pšenčík, I.: Computation of wave fields in inhomogeneous media — Gaussian beam approach. Geophys. J. R. Astr. Soc. **70**, 109–128 (1982)
5. Engquist, B., Runborg, O.: Computational high frequency wave propagation. Acta Numerica **12**, 181–266 (2003)
6. Fishman, G.S.: Monte Carlo: Concepts, Algorithms, and Applications. Springer- Verlag, New York (1996)
7. Giles, M.B.: Multilevel Monte Carlo path simulation. Oper. Res. **56**, 607–617 (2008)
8. Haji-Ali, A.L., Nobile, F., Tempone, R.: Multi-index Monte Carlo: when sparsity meets sampling. To appear in Numerische Mathematik (2016)
9. Hörmander, L.: The Analysis of Linear Partial Differential Operators I: Distribution Theory and Fourier Analysis. Springer-Verlag (1983)
10. Hörmander, L.: The Analysis of Linear Partial Differential Operators III: Pseudo-Differential Operators. Springer-Verlag (1994)
11. Hou, T.Y., Wu, X.: Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications. J. Comput. Phys. **230**, 3668–3694 (2011)
12. Jin, S., Markowich, P., Sparber, C.: Mathematical and computational models for semiclassical Schrödinger equations. Acta Numerica **20**, 121–209 (2011)
13. Liu, H., Ralston, J.: Recovery of high frequency wave fields from phase space–based measurements. Multiscale Model. Sim. **8**(2), 622–644 (2010)
14. Liu, H., Ralston, J., Runborg, O., Tanushev, N.M.: Gaussian beam method for the Helmholtz equation. SIAM J. Appl. Math. **74**(3), 771–793 (2014)
15. Liu, H., Runborg, O., Tanushev, N.: Sobolev and max norm error estimates for Gaussian beam superpositions. Commun. Math. Sci. **14**(7), 2037–2072 (2016)
16. Liu, H., Runborg, O., Tanushev, N.M.: Error estimates for Gaussian beam superpositions. Math. Comp. **82**, 919–952 (2013)
17. Malenová, G.: Uncertainty quantification for high frequency waves. Licentiate thesis, KTH Royal Institute of Technology (2016)
18. Malenova, G., Motamed, M., Runborg, O., Tempone, R.: A sparse stochastic collocation technique for high-frequency wave propagation with uncertainty. SIAM J. Uncertainty Quantification **4**, 1084–1110 (2016)
19. Mishra, S., Schwab, C.: Sparse tensor multi-level Monte Carlo finite volume methods hyperbolic conservation laws with random initial data. Mathematics of Computation **81**, 1979–2018 (2012)
20. Motamed, M., Nobile, F., Tempone, R.: A stochastic collocation method for the second order wave equation with a discontinuous random speed. Numer. Math. **123**, 493–536 (2013)
21. Motamed, M., Nobile, F., Tempone, R.: Analysis and computation of the elastic wave equation with random coefficients. Computers and Mathematics with Applications **70**, 2454–2473 (2015)
22. Motamed, M., Runborg, O.: Asymptotic approximations of high frequency wave propagation problems. In: B. Engquist, A. Fokas, E. Hairer, A. Iserles (eds.) Highly Oscillatory Problems, *London Mathematical Society Lecture Note Series*, vol. 366, pp. 72–97. Cambridge University Press (2009)
23. Motamed, M., Runborg, O.: Taylor expansion and discretization errors in Gaussian beam superposition. Wave Motion **47**, 421–439 (2010)
24. Nobile, F., Tamellini, L., Tempone, R.: Comparison of Clenshaw-Curtis and Leja quasi-optimal sparse grids for the approximation of random PDEs. Technical report mathicse-41.2014, MATHICSE, EPFL, Lausanne, Switzerland (2014)
25. Nobile, F., Tempone, R., Webster, C.G.: An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data. SIAM J. Numer. Anal. **46**, 2411–2442 (2008)
26. Ralston, J.: Gaussian beams and the propagation of singularities. Studies in partial differential equations **23**, 206–248 (1982)
27. Runborg, O.: Mathematical models and numerical methods for high frequency waves. Commun. Comput. Phys. **2**, 827–880 (2007)
28. Tanushev, N.M.: Superpositions and higher order Gaussian beams. Commun. Math. Sci. **6(2)**, 449–475 (2008)
29. Tanushev, N.M., Qian, J., Ralston, J.V.: Mountain waves and Gaussian beams. Multiscale Model. Simul. **6(2)**, 688–709 (2007)
30. Trefethen, L.N.: Is Gauss quadrature better than Clenshaw-Curtis? SIAM Rev. **50**, 67–87 (2008)

31. Tsuji, P., Xiu, D., Ying, L.: Fast method for high-frequency acoustic scattering from random scatterers. International Journal for Uncertainty Quantification **1**, 99–117 (2011)
32. Xiu, D., Hesthaven, J.S.: High-order collocation methods for differential equations with random inputs. SIAM J. Sci. Comput. **27**, 1118–1139 (2005)
33. Xiu, D., Karniadakis, G.E.: Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos. Computer Methods in Applied Mechanics and Engineering **191**, 4927–4948 (2002)