Numer. Math. (2009) 113:1–71
DOI 10.1007/s00211-009-0226-6

Numerische Mathematik

Wave atoms and time upscaling of wave equations

Laurent Demanet · Lexing Ying

Received: 17 March 2007 / Revised: 22 February 2009 / Published online: 12 May 2009

Abstract We present a new geometric strategy for the numerical solution of hyperbolic wave equations in smoothly varying, two-dimensional time-independent periodic media. The method consists in representing the time-dependent Green's function in wave atoms, a tight frame of multiscale, directional wave packets obeying a precise parabolic balance between oscillations and support size, namely wavelength $\sim$(diameter).^{2} Wave atoms offer a uniquely structured representation of the Green's function in the sense that

- the resulting matrix is *universally sparse* over the class of C^∞ coefficients, even for "large" times;
- the matrix has a natural *low-rank block-structure* after separation of the spatial indices.

The parabolic scaling is essential for these properties to hold. As a result, it becomes realistic to accurately build the full matrix exponential in the wave atom frame, using repeated squaring up to some time typically of the form $\Delta t \sim \sqrt{\Delta x}$, which is bigger than the standard CFL timestep. Once the "expensive" precomputation of the Green's function has been carried out, it can be used to perform unusually large, upscaled, "cheap" time steps. The algorithm is relatively simple in that it does not require an underlying geometric optics solver. We prove accuracy and complexity results based on a priori estimates of sparsity and separation ranks. On a N-by-N grid, the "expensive" precomputation takes somewhere between $O(N^3 \log N)$ and $O(N^4 \log N)$ steps

L. Demanet (✉)
Department of Mathematics, Stanford University, 450 Serra Mall, Stanford, CA 94305, USA
e-mail: demanet@gmail.com

L. Ying
Department of Mathematics, University of Texas at Austin,
1 University Station/C1200, Austin, TX 78712, USA

depending on the separability of the acoustic medium. The complexity of upscaled timestepping, however, beats the $O(N^3 \log N)$ bottleneck of pseudospectral methods on an N-by-N grid, for a wide range of physically relevant situations. In particular, we show that a naive version of the wave atom algorithm provably runs in $O(N^{2+\delta})$ operations for arbitrarily small δ—but for the final algorithm we had to slightly increase the exponent in order to reduce the large constant. As a result, we get estimates between $O(N^{2.5} \log N)$ and $O(N^3 \log N)$ for upscaled timestepping. We also show several numerical examples. In practice, the current wave atom solver becomes competitive over a pseudospectral method in regimes where the wave equation should be solved hundreds of times with different initial conditions, as in reflection seismology. In academic examples of accurate propagation of bandlimited wavefronts, if the precomputation step is factored out, then the wave atom solver is indeed faster than a pseudospectral method by a factor of about 3–5 at $N = 512$, and a factor 10–20 at $N = 1{,}024$, for the same accuracy. Very similar gains are obtained in comparison versus a finite difference method.

Mathematics Subject Classification (2000) Primary 65M99 · Secondary 42C99

1 Introduction

In this paper we address the design of fast algorithms for some wave propagation problems. The two-dimensional (2D) wave equation in periodic media with sound speed $c(x)$ can be written as

$$p_{tt} - c^2(x)\Delta p = 0, \quad p(x,0) = p_0, \quad p_t(x,0) = p_1.$$

We take the domain to be $[0,1]^2$ with periodic boundary conditions. Here and throughout, $c(x)$ is of class C^∞, and positive bounded away from zero. Very often, we write the wave equation as a linear system of two first-order equations through the additional unknown $q = p_t$. A shorthand notation for this system of equations is

$$u_t = Au, \quad u(t=0) = u_0,$$

where $u = (p, q)$ and A is a 2-by-2 matrix of operators. Naturally, the solution is $u(t) = e^{tA}u_0$.

The wave equation in heterogeneous media is a sensible model for acoustic waves traveling through the Earth, for example.

1.1 Time upscaling

Typical numerical methods for the above wave equation, say in the periodic square $[0,1]^2$ with initial conditions on a N-by-N grid, consist in evolving the solution using small time-steps Δt constrained by the CFL condition,

$$\Delta t < \frac{1}{c_{\max} N},$$

where $c_{\max} = \max_x c(x)$. When the wave equation is written as a system $u_t = Au$, the Euler explicit time discretization heuristically converges because

$$u(t) = e^{TA}u_0 \simeq (I + \Delta t\, A)^{\frac{T}{\Delta t}} u_0. \tag{1}$$

If the wave equation is to be solved until time 1, say, the method needs to perform $O(N)$ time steps.

One typical numerical method is the (very simple) split-step pseudospectral method, where $c^2(x)\Delta$ is discretized by applying Δ in the frequency domain, and $c^2(x)$ in the spatial domain. Since this method requires one FFT and one IFFT per time step, the overall complexity of solving the wave equation until time 1 is $O(N^3 \log N)$. In what follows we will use this pseudospectral method as a benchmark for our comparisons.

The first message of this paper is that the CFL condition can be bypassed, and complexity lowered below $O(N^3)$, if we can find a representation of the small time Green's function (also called propagator) $e^{\tau A}$ that is significantly different and better than a low-order polynomial in τA.

We call *time upscaling* the possibility of constructing, in a fast and accurate manner, such a representation where τ is much greater than the CFL timestep Δt—yet smaller than the time T up to which the wave equation needs to be solved.[1] The solution at $t = T$ can be obtained by performing $\frac{T}{\tau}$ "upscaled" time steps, consisting of repeated applications of the Green's function:

$$e^{TA}u_0 = \left(e^{\tau A}\right)^{\frac{T}{\tau}} u_0.$$

The question now becomes that of finding a good representation, in which computing $e^{\tau A}$ is an accessible task. The authors have shown in a series of papers [7,8,15,16] (see also [35]), that there exist isometric change of variables, or tight frames of $L^2(\mathbb{R}^2)$, in which the wave equation Green's function e^{tA} is *optimally sparse* even though these change of variables don't depend on $c(x)$. Sparsity can for instance be expressed as the ℓ_p to ℓ_p boundedness of the matrix of the Green's function, for *every* $p > 0$. Two such representations are wave packet families called *curvelets* and *wave atoms*: they are related to but quite different from wavelets. There are good reasons to believe that, in view of the current state of knowledge in applied harmonic analysis, no other transform architecture than curvelets or wave atoms will provide optimal sparsity results.

Note in passing that perfect time upscaling would be obtained in the basis of eigenfunctions of A. In that case $e^{tA} = \sum_j e^{t\lambda_j} P_j$ is a diagonal operation in each eigenspace with projector P_j. There is, at present, no known fast numerical procedure to compute the eigendecomposition of A in compressed form, let alone expand u_0 in

[1] It is important to understand the sense in which a numerical method could qualify as "beating the CFL condition". After all, one could discretize u_0 by finite differences, group small time steps two by two in (1), use $(I + \Delta t A)^2$ as propagator and declare that the new time step is $2\Delta t$. This operation of course does not qualify as time upscaling, because the matrix representation of the propagator fills up to compensate the larger time step, so that no overall simplification occurs. Progress is achieved only if a representation can be found in which the Green's function *stays* simple, even for times greater than Δt.

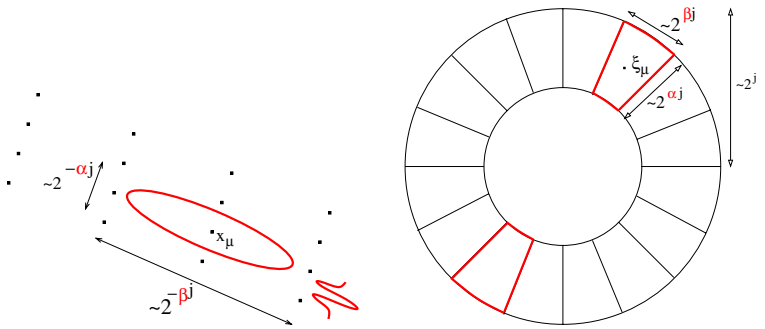

Fig. 1 Essential support of a wave packet with parameters (α, β), in space (*left*), and in frequency (*right*). The parameter α indexes the multiscale nature of the transform, from 0 (*uniform*) to 1 (*dyadic*). The parameter β measures the wave packet's directional selectivity, from $\beta = 0$ (*best selectivity*) to $\beta = 1$ (*poor selectivity*). Wave atoms are the special case $\alpha = 1/2$, $\beta = 1/2$

eigenfunctions. Although curvelets or wave atoms are not eigenfunctions of A, they offer a *fixed* frame of $L^2([0, 1]^2)$ with reliable expansion algorithms and good sparsity properties for e^{tA}.

1.2 Wave atoms, curvelets, and other phase-space tilings

An introduction to the curvelet transform is in [11], and to the wave atom transform in [16]. In this paper, we would like to introduce both transforms as a special case of "phase-space tiling".

We use two parameters to index a lot of known wave packet architectures: α to index whether the decomposition is "multiscale" ($\alpha = 1$) or not ($\alpha = 0$); and β to indicate whether basis elements should be isotropic ($\beta = \alpha$) or, on the contrary, elongated and anisotropic ($\beta < \alpha$).

We will require that each wave packet obey specific aspect ratios governed by α and β. Call $\varphi_\mu(x)$ a wave packet indexed by μ. The subscript μ is an aggregate of the scale, rotation, and translation parameters. At scale $j \geq 0$,

- the essential support of $\varphi_\mu(x)$ should be of size $\sim 2^{-\alpha j}$ versus $2^{-\beta j}$, with oscillations of wavelength $\sim 2^{-j}$ tranverse to the ridge; and
- the essential support of $\hat{\varphi}_\mu(\xi)$ should be of size $\sim 2^{\alpha j}$ versus $2^{\beta j}$, at a distance $\sim 2^j$ from the origin.

Figure 1 summarizes these "microlocalization" properties. A collection of wave packets is then built from these basic atoms by "tiling" phase-space (positions and frequency) while preserving the aspect ratios introduced above (Fig. 2).

We hope that a description in terms of α and β will clarify the connections between various transforms of modern harmonic analysis. Curvelets [8–10,20,33] correspond to $\alpha = 1, \beta = 1/2$, wavelets [13,27,30,33] are $\alpha = \beta = 1$, ridgelets [6,21] are $\alpha = 1$, $\beta = 0$, the Gabor transform is $\alpha = \beta = 0$, and the new "wave atoms" are the point $\alpha = \beta = 1/2$. A continuous Gaussian transform with geometry $\alpha = \beta = 1/2$ was introduced by Cordóba and Fefferman [14]. The situation is summarized in Fig. 3.

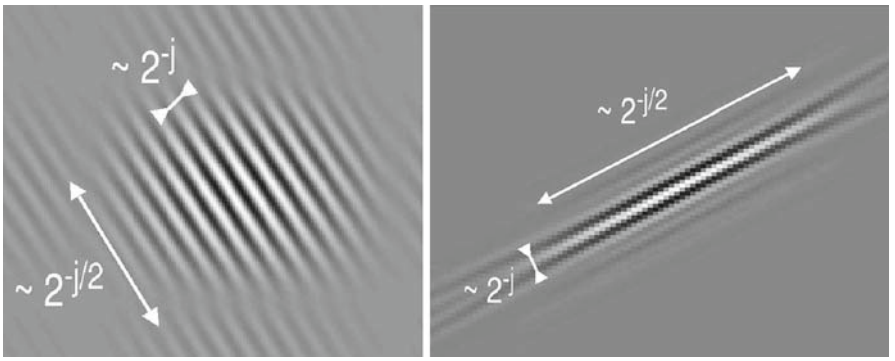

Fig. 2 *Left*: a digital wave atom, from the Waveatom software [16]. *Right*: a digital curvelet, from the Curvelab software [9]

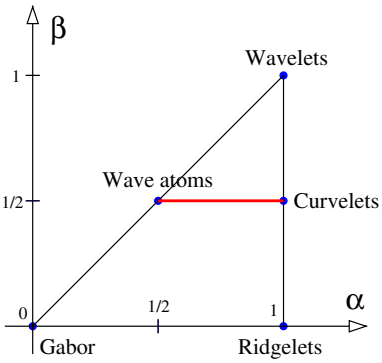

Fig. 3 Identification of various transforms as (α, β) families of wave packets. The horizontal segment at $\beta = 1/2$ indicates the only wave packet families that yield sparse decompositions of Fourier Integral Operators and wave equation Green's functions

A precise definition of wave atoms will be given later, but let us observe for now that they have an isotropic aspect ratio $\sim 2^{-j/2} \times 2^{-j/2}$ in space, with oscillations of wavelength $\sim 2^{-j}$ in the codirection ξ_μ. The subscript μ contains $(j, \mathbf{m}, \mathbf{n})$, where j is scale, $\mathbf{m}$ indexes the different wave vectors at scale j, and $\mathbf{n}$ indexes the position vector. There exist a fast transform, with a fast and accurate inverse. Wave atoms form a tight frame, in the sense that, for all $f \in L^2(R^2)$, we have the Plancherel identity

$$\|f\|_2^2 = \sum_\mu |\langle f, \varphi_\mu \rangle|^2$$

(those two quantities would only be within a constant factor of each other in the case of a general frame). As a result, any function $f \in L^2(R^2)$ expands as

$$f = \sum_\mu c_\mu \varphi_\mu, \quad c_\mu = \langle f, \varphi_\mu \rangle,$$

and any operator $A : L^2(\mathbb{R}^2) \to L^2(\mathbb{R}^2)$ expands as

$$Af = \sum_{\mu,\mu'} c_\mu A_{\mu,\mu'} \varphi_{\mu'}, \quad A_{\mu,\mu'} = \langle \varphi_{\mu'}, A\varphi_\mu \rangle.$$

The transforms corresponding to the line segment $\beta = 1/2$, and $1/2 \le \alpha \le 1$, including wave atoms and curvelets, are the only transforms that offer an optimally sparse representation of the wave equation Green's function, among those in the triangle. In short, other scalings than the parabolic scaling do not work.

More can be said. The wave atom matrix of the Green's function e^{tA} has the following structure: for every column index μ' (indexing position and frequency of a wave atom), the only row indices giving rise to significant matrix elements are of the form $\mu \simeq \mu'_\nu(t)$, where we have used the shorthand label $\mu'_\nu(t)$ for the image of the point $(x_{\mu'}, \xi_{\mu'})$ in phase-space (the "center" of the wave atom $\varphi_{\mu'}$) under the Hamiltonian dynamics in phase-space, or sound rays, up to time t. For the wave equation, there are two values of ν called $\pm$ (because there are two Hamiltonian flows), hence $\mu \simeq \mu'_\nu(t)$ defines two "shifted diagonals", around which the large matrix elements cluster, and away from which they decay almost exponentially.

An important matrix approximation result follows if we discard the small entries away from the shifted band-diagonals.

Corollary 1 *Let $E_B(t)$ be the operator resulting from truncation of the wave atom matrix of $E(t) = e^{tA}$ to shifted band-diagonals of size B. Then*

$$\|E(t) - E_B(t)\|_{L^2 \to L^2} \le C_{t,M} B^{-M}, \qquad \text{for all } M > 0. \tag{2}$$

The constant $C_{t,M}$ grows at most exponentially in t. The same result holds in the wave atom representation, in all $\ell_p \to \ell_p$ norms, for $1 \le p \le \infty$.

For more details, see Sect. 2. For the justification of Corollary 1, see the original paper [8].

At this point, both curvelets and wave atoms offer the same sparsity properties, but it turns out that wave atoms are more adequate than curvelets for the numerical simulations of wave equations because of their low separation rank. Why this is an asset will be made clear in due course. We will focus exclusively on wave atoms in the sequel. Let us now explain how the Green's function should be constructed numerically.

1.3 Repeated squaring

In the spirit of [19] and [32], we form the matrix exponential $e^{\tau A}$ by *repeated squaring* from a small time approximation. Let $t_n = 2^n \Delta t$ for some small Δt, and assume $\tau = t_{n^*}$ for some n^*. Then the basic relation underlying our algorithm is the time-doubling group property

$$e^{t_{n+1}A} = \left(e^{t_n A}\right)^2.$$

As mentioned earlier, this equation should be understood in a tight frame of wave atoms; once the matrix of an operator is available in the transformed domain, we can use the standard sequence (1) transform (2) apply the matrix, (3) inverse transform, to compute the action of the operator.

Sparsity needs to be imposed by an adequate *truncation* step after each time doubling. As we mentioned earlier, the large matrix elements occur near two shifted diagonals defined from the flows $\mu'_{\pm}(t)$. Let us call B the desired band size, such that the significant matrix elements live within

$$\omega(\mu, \mu'_{\pm}(t)) \leq C_B, \tag{3}$$

where C_B is a constant depending on B. Elements outside of those shifted band diagonals should not be accessed or computed at all. The proper notion of distance ω in wave-atom phase space will be defined later.

Prediction of the location of the shifted diagonals for $O(1)$ times is not a priori obvious. We believe the Phase Flow Method (PFM) is perfectly suited for this task [37]. This new method is an important improvement over raytracing which allows to compute a multitude of rays at once. PFM is an interpolation-based repeated squaring strategy to compute the whole *phase flow*, that is, the diffeomorphism of phase-space generated by the Hamiltonian ODE system.

We are glossing over lots of details, see Sect. 3.1 for those, but we can already state that this simple repeated squaring procedure has near-optimal asymptotic complexity, in the sense that it requires

$$C_{\epsilon,\delta} N^{2+\delta} \tag{4}$$

operations to build the propagator on an N-by-N grid, for a resulting ℓ^2 accuracy ϵ on fairly oscillatory initial data, and for arbitrarily small $\delta > 0$. The constant $C_{\epsilon\delta}$ depends on ϵ like $C_M \epsilon^{-1/M}$, signaling what is usually called spectral accuracy in numerical analysis. Of course, $C_{\epsilon\delta}$ also blows up as $\delta \to 0$. Notice that reading the initial data already takes N^2 operations, so the above complexity estimate is asymptotically near-optimal in N.

Heuristically, the complexity result (4) follows directly from the compression result (2): multiplication of sparse matrices with size N^2 and band size B has complexity $O(B^2 N^2)$. Justifying (4) means showing that $B = O(N^{\delta/2})$ for small δ suffices to control the error from successive thresholdings and repeated squarings of matrices.

The actual proof of (4), albeit in a slightly idealized setting, can be found in Appendix B. A by-product of our analysis is the important observation that the overall accuracy ϵ of the solver is uniform in the choice of τ between $O(1/N)$ (CFL level) and $O(1)$ (complete time upscaling). Of course complexity is not and we detail this dependence in the sequel.

1.4 Curse of dimensionality

To our knowledge, the complexity claim (4) for the repeated squaring is the first to break the asymptotic $O(N^3 \log N)$ bottleneck of standard methods in two dimensions,

and by a wide margin. In the spirit of spectral methods, universally good accuracy over oscillatory initial conditions is a result of discretizing differential operators in the Fourier domain. These encouraging result shows that wave packet analysis brings fundamentally new insights into the numerical analysis of wave equations.

Yet, the repeated squaring algorithm as introduced above does not perform as expected, regardless of whether curvelets or wave atoms are used. A typical band size B to obtain ℓ^2 accuracy $\simeq 10^{-2}$ in (2) would be $B \simeq 500$. As a result, storing the compressed Green's function on grids larger than 128-by-128 requires more memory than what most 2006 desktop computers can offer (2–4 Gb).

There is no contradiction: asymptotic estimates like (2) and (4) are valid, but with large constants. In two space dimensions these large constants makes sense if we observe that B is the total number of elements inside a ball in four-dimensional phase-space, as in Eq. (3), hence the relation $B \simeq C_B^4$. If $C_B \simeq 5$ elements define a decent neighborhood in phase-space, then $B \simeq 625$. This curse of dimensionality is probably also responsible, although perhaps to a lesser extent, for the relatively large constants found in wavelet-based numerical methods for PDE [12]. We believe it would be timely if some of the "computational harmonic analysis" literature would start addressing this high-dimensional phenomenon.

An encouraging thought, perhaps, is that the curse of phase-space dimensionality for the wave equation can be overcome with an adequate *separation of variables* strategy in the wave atom frame, which we now explain.

1.5 The separated wave atom representation

Consider a tight frame of wave atoms $\varphi_\mu(x)$, with $\mu = (j, \mathbf{m}, \mathbf{n})$. The actual wave atom representation of $E(t)$ is the (infinite) matrix

$$E(t; \mu, \nu; \mu', \nu') = \langle E(t)\varphi_{\mu'}\mathbf{e}_{\nu'}, \varphi_\mu \mathbf{e}_\nu \rangle.$$

where $\mathbf{e}_\nu$ are the canonical basis vectors in $\mathbb{R}^2$ and $\nu = \pm$. In the above matrix, consider the submatrix left after fixing ν, ν' and the wave vectors $(j, \mathbf{m})$ and $(j', \mathbf{m}')$. The remaining indices are those of the position vectors $\mathbf{n} = (n_1, n_2)$ and $\mathbf{n}' = (n_1', n_2')$. The *separated wave atom representation* is obtained by seeking a low-rank approximation corresponding to separation of the spatial indices along x_1 versus x_2,

$$E(t; j, \mathbf{m}, \mathbf{n}, \nu; j', \mathbf{m}', \mathbf{n}', \nu') = \sum_{k=1}^{r} \sigma_k u^k_{n_1, n_1'} v^k_{n_2, n_2'} + O(\epsilon),$$

where u^k and v^k have been normalized to unit ℓ_2 norm. Of course u^k and v^k depend on $j, \mathbf{m}, \nu; j', \mathbf{m}', \nu'$. The most efficient such decomposition, in the sense that the ℓ^2 norm of the residual is minimized for fixed r, is the singular value decomposition (SVD) of the block $(j, \mathbf{m}, \nu; j', \mathbf{m}', \nu')$ after *reorganization* of the matrix elements to make the row and column indices $(n_1, n_1'; n_2, n_2')$ instead of $(n_1, n_2; n_1', n_2')$.

Conversion from the standard to the separated wave atom representation, as an SVD factorization of the reorganized submatrix, is however never done in practice.

Instead, we modify the repeated squaring strategy so that all computations are done on separated components without ever forming the standard submatrix. We explain later how both initialization and matrix multiplication can be realized in this context, using small QR and SVD decompositions.

The separated wave atom scheme performs much better than the standard repeated squaring, both in terms of memory and time savings—hence feasibility on larger grids. It even competes with the standard pseudospectral method in regimes where a given wave equation should be solved several times with different initial conditions. Forming the Green's function should be seen as a precomputation that can be amortized over the several runs. For instance in Sect. 4, we take for $c(x)$ a smooth wave guide and observe that about 500 runs is enough to amortize the precomputation. For the examples we tried, upscaled timestepping alone—not counting the precomputation—runs about 3 times faster than a pseudospectral method, see Sect. 4.

Complexity of the separated wave atom scheme is very well understood. We give precise estimates of ϵ-separation ranks [r in Eq. (13)] as a function of the upscaled time step τ, the scale j and accuracy level ϵ. The resulting number of operations for repeated squaring (RS) and upscaled timestepping (UTS) are reported in Sect. 3.3. Although not optimal anymore, estimates for UTS still beat the $O(N^3 \log N)$ bottleneck in a variety of physically interesting situations.

The methods of proof of rank estimates rely on understanding the information content of oscillatory functions in high dimensions—or their Fourier dual, functions with singularities—and could be of independent interest in numerical analysis.

Let us conclude this section by mentioning that the composition rules we developed for wave atom submatrices are remindful of the calculus of H-matrices [22]. In fact, high-dimensional numerical analysis using separated representations is a promising emerging idea, see the paper by Beylkin and Mohlenkamp [4], and citations thereof. Partitioned low rank representations for the wave equation were already used by Beylkin and Sandberg [5], with nice results, but in a setup in which it is not obvious that advantageous rank estimates would hold. As a result we believe that the time upscaling capabilities of the approach in [5] are in fact quite limited. On the other hand, the sophistication of the approach in Beylkin and Sandberg is in the choice of basis functions (PSWF) for representing bandlimited functions on non-periodic domains, a refinement that our method does not address.

1.6 Applications and opportunities

We aimed at a low-complexity solver for extremely large 2D wave problems in heterogeneous media, which need to be solved at least several hundreds of time with different data. The initial data is allowed to be fairly oscillatory. We anticipate that some formulations of the linearized inverse problem in reflection seismology (how to find $c(x)$ from wavefield measurements), like reverse-time migration, could be cast in a form in which such a solver may become useful.[2]

[2] We would like to thank William Symes and Maarten de Hoop for interesting discussions on reflection seismology.

Beyond the prospect of an algorithmic speedup also lie opportunities of linking the geometries of the transform, the Green's function and the data. *The wave atom algorithm does not just compute the Green's function, it offers a representation where its phase-space structure is explicit.*

For instance, sparsity of the solution wavefield in wave atoms directly translates into complexity gains for the upscaled timestepping. If the initial condition can be accurately represented using a fraction $\rho < 1$ of all wave atoms, then applying the Green's function in wave atoms only requires considering a fraction ρ of all rows. In particular, we can show that the "bandlimited wavefronts" considered in Sect. 4 remain so in time and satisfy $\rho = O(\frac{1}{\sqrt{N}})$.

Hence we see that wave atoms, or curvelets, provide a unique opportunity for having a representation giving enhanced sparsity of wave groups, and *simultaneously* of the solution space. In this respect, curvelets are ideal for representing wavefront phenomena [11], or objects which display curve-punctuated smoothness—smoothness except for discontinuity along a general curve with bounded curvature [10,11], and wave atoms are ideal at representing classes of warped oscillatory patterns [16]. We believe that these joint sparsity properties will eventually be of great practical significance for applications in fields which are great consumers of these mathematical models, e.g., seismic imaging [9,16,17,23–25].

Finally, we would like to mention that [35] presents theory that deals with acoustic media of limited differentiability, e.g., of class $C^{1,1}$. It is important for applications e.g. to exploration seismology that the medium contain interesting features which may not be well modeled by the class C^∞. Like in [35], we anticipate that all estimates of sparsity and separability deteriorate when the medium becomes non-smooth. The numerical solver will not crash if higher frequencies are introduced in the medium, but its performance clearly deteriorates in the same fashion as the estimates would.

2 Wave atoms and wave equations

The mathematical requirements we put on a family of basis function to be called "wave atoms" are quite stringent and will be made precise below. They have to do with uniform space-frequency localization, and put the general architecture introduced earlier on solid ground. To the best of our knowledge, none of the transforms in the repertoire of modern computational harmonic analysis satisfies these localization conditions.

2.1 Definition of wave atoms

We write wave atoms as $\varphi_\mu(x)$, with subscript $\mu = (j, \mathbf{m}, \mathbf{n}) = (j, m_1, m_2, n_1, n_2)$. All five quantities j, m_1, m_2, n_1, n_2 are integer-valued and index a point (x_μ, ξ_μ) in phase-space, as

$$x_\mu = 2^{-j}\mathbf{n}, \quad \xi_\mu = \pi 2^j \mathbf{m}, \quad C_1 2^j \le \max_{i=1,2} |m_i| \le C_2 2^j. \tag{5}$$

where C_1, C_2 are two positive constants left unspecified for convenience, but whose values will be implied by the specifics of the implementation. Heuristically, the *position vector* x_μ is the center of $\varphi_\mu(x)$ and the *wave vector* ξ_μ determines the centers of both bumps of $\hat{\varphi}_\mu(\xi)$ as $\pm\xi_\mu$. Note that the range of $\mathbf{m}$ needs to be further reduced to $m_2 > 0$, (or $m_2 = 0$ and $m_1 > 0$,) to account for the central symmetry of the Fourier transform of real-valued functions about the origin in ξ. Some further restriction on $\mathbf{n}$ (cutoff in space) and j (cutoff in scale), are of course necessary in practice, but not for the description of a frame of L^2.

Wave atoms then need to obey a localization condition around the phase-space point (x_μ, ξ_μ).

Definition 1 (*Wave Atoms*) Let x_μ and ξ_μ be as in Eq. (5) for some C_1, C_2. The elements of a frame of wave packets $\{\varphi_\mu\}$ are called *wave atoms* when

$$|\hat{\varphi}_\mu(\xi)| \le C_M \cdot 2^{-j}(1 + 2^{-j}|\xi - \xi_\mu|)^{-M} + C_M \cdot 2^{-j}(1 + 2^{-j}|\xi + \xi_\mu|)^{-M}, \quad \text{for all } M > 0, \tag{6}$$

and

$$|\varphi_\mu(x)| \le C_M \cdot 2^{j}(1 + 2^{j}|x - x_\mu|)^{-M}, \quad \text{for all } M > 0. \tag{7}$$

It is of course possible to restrict the decay order or even moderately alter the definition of x_μ and ξ_μ—and still call the basis functions "wave atoms"—but this is a refinement we will not address here.

The parabolic scaling is encoded in the localization conditions as follows: at scale 2^{-2j}, or frequency $|\xi_\mu| \sim 2^{2j}$, the essential frequency support is of size $\sim 2^j$ (for each bump) and the essential spatial support is of size $\sim 2^{-j}$. Note that the subscript j indexes the different "dyadic coronae", whereas the additional subscript $\mathbf{m}$ labels the different wave numbers ξ_μ within each dyadic corona.

2.2 Properties

In the same spirit as curvelets [34,35], there is a natural notion of pseudodistance in phase-space associated to wave atoms.

Definition 2 Let $\mu = (j, \mathbf{m}, \mathbf{n})$ and $\mu' = (j', \mathbf{m}', \mathbf{n}')$ be two wave atom subscripts. The *wave atom pseudodistance* $\omega(\mu, \mu')$ is defined as

$$\omega(\mu, \mu') = 1 + 2^{\min(j,j')}|x_\mu - x_{\mu'}| + 2^{-\max(j,j')}|\xi_\mu - \xi_{\mu'}|. \tag{8}$$

The motivation for this definition is the interpretation of ω a *lattice distance* in phase-space. Consider the graph with wave packet indices μ as nodes, and connection between two nodes if and only if the corresponding wave packets are neighbors, that is if

- either $j = j'$, $\mathbf{m} = \mathbf{m}'$ and $|\mathbf{n} - \mathbf{n}'| = 1$;
- or μ and μ' correspond to adjacent frequency tiles and $|x_\mu - x_{\mu'}| = \min_{\mu''=(j',\mathbf{m}',\mathbf{n}'')} |x_\mu - x_{\mu''}|$.

Then $\omega(\mu, \mu')$ is proportional to the minimum number of edges needed to connect μ and μ'.

Of course ω is not a distance in the strict sense, but it is symmetric, satisfies the quasi-triangle inequality $\omega(\mu, \mu'') \le C(\omega(\mu, \mu') + \omega(\mu', \mu''))$, and is invariant under Hamiltonian flows, $\omega(\mu(t), \mu'(t)) \asymp \omega(\mu(0), \mu'(0))$.

The main purpose of ω is that it allows us to formulate a key almost-orthogonality estimate.

Lemma 1 *Let φ_μ and $\tilde{\varphi}_\mu$ be two collections of wave atoms, in the sense of Definition 1. Then for every $M > 0$ there exists a constant $C_M > 0$ such that*

$$|\langle \varphi_\mu, \tilde{\varphi}_{\mu'} \rangle| \le C_M \cdot \omega(\mu, \mu')^{-M}.$$

Proof We will make use of the following generic bump convolution convolution inequality (see [31, p. 56]), valid when $a \ge a'$ and, say, $M \ge 2$,

$$\int_{-\infty}^{\infty} (1 + a|x|)^{-M} (1 + a'|x - x_0|)^{-M} \, dx \le \frac{C}{a} (1 + a'|x_0|)^{-M}. \tag{9}$$

In two dimensions, we get $\frac{C}{a^2}$ in place of $\frac{C}{a}$, and $M \ge 3$. Assume without loss of generality that $j \le j'$. Combining (9) with the frequency localization estimate (6), we obtain

$$\int |\hat{\varphi}_\mu(\xi) \hat{\tilde{\varphi}}_{\mu'}(\xi)| \, d\xi \le C_M \cdot 2^{j-j'} (1 + 2^{-j'} |\xi_\mu - \xi_{\mu'}|)^{-M}. \tag{10}$$

Similarly, combining (9) with the spatial localization estimate (7), we obtain

$$\int |\varphi_\mu(x) \tilde{\varphi}_{\mu'}(x)| \, dx \le C_M \cdot 2^{j-j'} (1 + 2^{j} |x_\mu - x_{\mu'}|)^{-M}. \tag{11}$$

The conclusion follows by taking the geometric mean of (10) and (11), and noticing that

$$\begin{aligned} &(1 + 2^{-j'} |\xi_\mu - \xi_{\mu'}|)^{-M} (1 + 2^{j} |x_\mu - x_{\mu'}|)^{-M} \\ &\le (1 + 2^{-j'} |\xi_\mu - \xi_{\mu'}| + 2^{j} |x_\mu - x_{\mu'}|)^{-M}. \end{aligned}$$

□

We are now ready to state the main sparsity result. Almost orthogonality is one ingredient in its proof, hence the resemblance of statements. We will not give the proof of this result, but is a straightforward modification of the argument for curvelets. The interested reader is invited to refer to the original paper [8] and check that all statements completely carry over to the case of wave atoms.

Theorem 1 *Let $E(t) = e^{tA}$, with $c(x)$ of class C^∞ and bounded away from zero. Let $E(t; \mu, \nu; \mu', \nu')$ be the representation of $E(t)$ in a wave atom frame φ_μ. Denote by $\mu'_\nu(t)$ be image of μ', flowed along the νth Hamiltonian flow. Then, for all times $t > 0$, for all $M > 0$,*

$$|E(t; \mu, \nu; \mu', \nu')| \leq C_{t,M} \cdot \sum_{\nu''=1}^{2} \omega(\mu, \mu'_{\nu''}(t))^{-M}.$$

The constant $C_{t,M}$ grows in time at most like $C_1 e^{C_2 t}$ for some $C_1, C_2 > 0$ themselves depending on M.

The notion of a subscript being "flowed" along the trajectory of a dynamical system is explained in [8].

We can finally make the heuristics of the introduction precise: Theorem 1 is a sparsity result, because the decay of the matrix elements is super-algebraic (near-exponential) away from the shifted diagonal $\mu \simeq \mu'_{\nu''}(t)$, in the sense of the pseudo-distance ω. This is the Theorem valid for the segment at $\beta = 1/2$ in Fig. 3. Note that Corollary 1 in the Introduction is a consequence of Theorem 1.

Another corollary of Theorem 1 is that the wave atom matrix of $E(t)$ is bounded on ℓ_p for all $p > 0$; see [8] for a justification in the case of curvelets.

2.3 Specifics on the implementation

How wave atoms are realized as a digital transform is explained in [16]. There are two variants: (1) a tight frame with redundancy 2, in which each wave atom has 2 bumps in frequency; and (2) an orthobasis, in which each wave atom has 4 bumps in frequency. Which variant to choose has no bearing on the theory. In both cases, the wave atom transform on a N-by-N grid has complexity $O(N^2 \log N)$ with a small constant. The complexity of the inverse transform is the same, and the accuracy of the sequence direct-then-inverse transform is empirically 14 digits in double precision, for grid sizes N on the order of 1,000 or less. In practice, we chose the orthobasis variant for all our numerical experiments.

3 Main algorithm

The description of the main algorithm will be split into two parts. The basic repeated squaring algorithm only exploits sparsity of the wave atom matrix of the propagator $E(t)$ and is detailed in Sect. 3.1. The refinement based on separation of spatial indices comes as a modification of the basic repeated squaring algorithm, and is explained in Sect. 3.2.

3.1 Basic repeated squaring

Let us denote $u(t)$ for the couple $(p(t), \frac{\partial p}{\partial t})$, and write the wave equation as the first-order system $\frac{\partial u}{\partial t} = Au$ with initial condition $u(0) = u_0$. The generator is

$$A = \begin{pmatrix} 0 & I \\ c^2(x)\Delta & 0 \end{pmatrix}. \tag{12}$$

We define the propagator $E(t)$ from $u(t) = E(t)u_0 = e^{tA}u_0$.

Since the solution $u(t)$ has two components, we need to introduce $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$. The wave atom matrix elements are

$$E(t; \mu\nu; \mu'\nu') = \langle E(t)\varphi_{\mu'}\mathbf{e}_{\nu'}, \varphi_\mu \mathbf{e}_\nu \rangle.$$

We write $\tilde{E}(t; \mu\nu; \mu'\nu')$ for its numerical approximation. As mentioned in the introduction we aim at building this matrix at dyadic times $t_n = 2^n \Delta t$, using a repeated-squaring strategy, based on the group property

$$E(2t; \mu\nu; \mu'\nu') = \sum_{\mu'',\nu''} E(t; \mu\nu; \mu''\nu'')E(t; \mu''\nu''; \mu'\nu'),$$

which in turn comes from $E(2t) = (E(t))^2$ and the tight-frame property. The squaring is efficient because the numerical approximation of the wave atom matrices is kept sparse at all dyadic times, by putting to zero the small entries below a prescribed threshold.

Algorithm 1 (Wave-Atom Repeated Squaring) *Choose a small time step Δt and a small tolerance ϵ. Denote by* Trunc *the operation of putting to zero all matrix elements below ϵ in absolute value.*

- **Initialization***: Obtain $\tilde{A}(\mu\nu; \mu'\nu')$ an approximation to the wave atom matrix of the generator A, then*

$$\tilde{E}(\Delta t, \mu\nu; \mu'\nu') = \delta_{\mu\nu;\mu'\nu'} + \Delta t \; Trunc(\tilde{A}(\mu\nu; \mu'\nu')).$$

- **Iteration***: Forecast the biggest entries' location, then compute them as*

$$\tilde{E}(2^{n+1}\Delta t; \mu\nu; \mu'\nu') = Trunc \sum_{\mu'',\nu''} E(2^n \Delta t; \mu\nu; \mu''\nu'')E(2^n \Delta t; \mu''\nu''; \mu'\nu').$$

- **Terminate** *at time $\tau = 2^{n^*}\Delta t$.*

To compute the solution $u(\tau)$ at time τ, start with the coefficients

$$c_{\mu\nu}(0) = \langle u_0, \varphi_\mu \mathbf{e}_\nu \rangle,$$

perform the matrix-vector multiplication,

$$\tilde{c}_{\mu\nu}(\tau) = \sum_{\mu'\nu'} \tilde{E}(\tau; \mu\nu; \mu'\nu')c_{\mu'\nu'}(0),$$

and inverse transform, $\tilde{u}(\tau) = \sum_{\mu\nu} \tilde{c}_{\mu\nu}(\tau)\varphi_\mu \mathbf{e}_\nu$. For times larger than τ, one should perform several applications of $E(\tau)$ to the initial condition.

As alluded to in the introduction section, prediction of the location of the large matrix elements is efficiently done using the PFM, see [37]. The truncation should be done to keep at most B elements per row (hence also per column), with B a moderately large constant.

For the initialization, how to compute an approximation to the generator $A_{\mu\nu;\mu'\nu'}$ in an efficient manner is best understood in the context of the separated wave atom representation, so we postpone this discussion until the next section.

Let us finally remark that, in view of Theorem 1, the *Trunc* operation consists in keeping track of *two* shifted diagonals, because there are two Hamiltonian flows. If instead of the standard wave atoms $\varphi_\mu^{(1)}$ or $\varphi_\mu^{(2)}$ we use the orthobasis variation of Sect. 2.3, $\varphi_\mu^+ = \varphi_\mu^{(1)} + \varphi_\mu^{(2)}$, we could expect to have to trace *four* bumps. For small times, namely $t \le \frac{1}{\sqrt{N}}$ or a multiple thereof, we will see later in Sect. 3.3 that tracing is useful but not necessary, so the gain due to lower redundancy may offset the frequency entangling. For larger times, "clean" wave atoms with two bumps in frequency may be more appropriate.

3.2 The separated wave atom representation

In the wave atom representation of $E(t)$, consider the submatrix left after fixing ν, ν' and the wave vectors $(j, \mathbf{m})$ and $(j', \mathbf{m}')$. The remaining indices are those of the position vectors $\mathbf{n} = (n_1, n_2)$ and $\mathbf{n}' = (n_1', n_2')$. The *separated wave atom representation* is obtained by seeking a low-rank approximation corresponding to separation of the spatial indices along x_1 versus x_2,

$$E(t; j, \mathbf{m}, \mathbf{n}, \nu; j', \mathbf{m}', \mathbf{n}', \nu') = \sum_{k=1}^{r} \sigma_k u^k_{n_1,n_1'} v^k_{n_2,n_2'} + O(\epsilon), \tag{13}$$

where u^k and v^k have been normalized to unit ℓ_2 norm. Of course u^k and v^k depend on $j, \mathbf{m}, \nu; j', \mathbf{m}', \nu'$. The most efficient such decomposition, in the sense that the ℓ^2 norm of the residual is minimized for fixed r, is the singular value decomposition of the block $(j, \mathbf{m}, \nu; j', \mathbf{m}', \nu')$ after *reorganization* of the matrix elements to make the row and column indices $(n_1, n_1'; n_2, n_2')$ instead of $(n_1, n_2; n_1', n_2')$.

Conversion from the standard to the separated wave atom representation, as an SVD factorization of the reorganized submatrix, is however never done in practice. Instead, we modify the repeated squaring strategy so that all computations are done on separated components without ever forming the standard submatrix. Let us explain how both initialization and matrix multiplication can be realized in this context.

3.2.1 Initialization

The wave atom representation of the generator reads $\langle \varphi_{\mu\nu}, A\varphi_{\mu'\nu'} \rangle$, where $\nu = 1$ refers to the $p(t)$ component, whereas $\nu = 2$ refers to the $\frac{\partial p}{\partial t}$ component. The only non-trivial or non-precomputable contribution is

$$\langle \varphi_{\mu,2}, A\varphi_{\mu',1}\rangle = \int_{[0,1]^2} \varphi_\mu(x) c^2(x) \Delta\varphi_{\mu'}(x)\, dx. \tag{14}$$

Our initialization strategy is based on separation of the integrand in x_1 versus x_2. Since $c^2(x)$ is a C^∞ periodic function, its ϵ-separation rank is a small constant $C_\epsilon = O(\epsilon^{-1/M})$ for all $M > 0$ (see Lemma 2) so we can write

$$c^2(x) = \sum_{k=1}^{C_\epsilon} \gamma_k^{(1)}(x_1)\gamma_k^{(2)}(x_2).$$

The Laplacian operator is also nicely separated,

$$\Delta = \frac{\partial^2}{\partial x_1^2} \otimes I + I \otimes \frac{\partial^2}{\partial x_2^2}. \tag{15}$$

As for wave atoms themselves, let us assume that we are using the separable "orthobasis" variation, as in Sect. 2.3. For the full wave atoms there would be two separated terms to write down instead.

We can then split the matrix element (14) into a finite number of separated components,

$$\begin{aligned}\langle \varphi_\mu, c^2(x)\Delta\varphi_{\mu'}\rangle = \sum_{k=1}^{C_\epsilon} &\left\langle \psi^j_{m_1,n_1}, \gamma_k^{(1)} \frac{\partial^2}{\partial x_1^2} \psi^{j'}_{m_1',n_1'} \right\rangle \left\langle \psi^j_{m_2,n_2}, \gamma_k^{(2)} \psi^{j'}_{m_2',n_2'} \right\rangle \\ &+ \left\langle \psi^j_{m_1,n_1}, \gamma_k^{(1)} \psi^{j'}_{m_1',n_1'} \right\rangle \left\langle \psi^j_{m_2,n_2}, \gamma_k^{(2)} \frac{\partial^2}{\partial x_2^2} \psi^{j'}_{m_2',n_2'} \right\rangle,\end{aligned}$$

where all the inner products in the right-hand side are one-dimensional. Observe that the above formula is exactly in "separated wave atom" form, as in Eq. (13).

The initialization algorithm computes all the factors in the above decomposition as follows. Assume the segment [0, 1] has been discretized into N equispaced points. For each (j', m_1', n_1'),

1. Form $\psi^{j'}_{m_1',n_1'}(x_1)$ for x_1 on the grid by applying the inverse 1D wavelet packet transform to the sequence of coefficients

$$c_{j,m_1,n_1} = \begin{cases} 1 & \text{if } j = j', m_1 = m_1', n_1 = n_1', \\ 0 & \text{otherwise.} \end{cases}$$

2. Apply $\gamma_k^{(1)} \frac{\partial^2}{\partial x_1^2}$ to $\psi^{j'}_{m_1',n_1'}(x)$. For accuracy purposes, all derivatives are discretized in the Fourier domain.

3. Apply a direct wavelet packet transform to the result and obtain at once the inner products with all the $\psi^j_{m_1,n_1}(x_1)$.

Repeat over all indices (j', m_1', n_1'). Repeat the algorithm, *mutatis mutandis*, for the inner products involving no derivatives and the inner products involving x_2 instead of x_1. Do not sum over k or multiply the factors as we are interested in the separated form only.

In practice, we may use a different time-integration scheme than Euler explicit for the first time step. We have found the leap-frog scheme to be quite efficient.

3.2.2 Matrix multiplication

We seek a fast algorithm for

$$\begin{aligned} &E(2t;\, j, \mathbf{m}, \mathbf{n}, \nu;\, j'', \mathbf{m}'', \mathbf{n}'', \nu'') \\ &\quad = \sum_{j', \mathbf{m}', \mathbf{n}', \nu'} E(t;\, j, \mathbf{m}, \mathbf{n}, \nu;\, j', \mathbf{m}', \mathbf{n}', \nu') E(t;\, j', \mathbf{m}', \mathbf{n}', \nu';\, j'', \mathbf{m}'', \mathbf{n}'', \nu''), \end{aligned} \tag{16}$$

where each factor is given by (13). We fix the row index $(j, \mathbf{m}, \nu)$ as well as the column index $(j'', \mathbf{m}'', \nu'')$, and for simplicity omit to write them in what follows.

We start by directly computing each sum over n_1' and n_2'. There is one such sum for each value of the intermediate index $(j', \mathbf{m}', \nu')$. Let us introduce

$$U^{\mathbf{k}}_{n_1,n_1''}(j', \mathbf{m}', \nu') = \sum_{n_1'} u^k_{n_1,n_1'} u^{k'}_{n_1',n_1''} \tag{17}$$

and, similarly,

$$V^{\mathbf{k}}_{n_2,n_2''}(j', \mathbf{m}', \nu') = \sum_{n_2'} v^k_{n_2,n_2'} v^{k'}_{n_2',n_2''}. \tag{18}$$

We grouped (k, k') into one single index $\mathbf{k}$. If we also let $\sigma_{\mathbf{k}} = \sigma_k \sigma_{k'}$, and form the resulting diagonal matrix $\Sigma_{\mathbf{k}}$, then the matrix element (16) can be written as

$$\begin{aligned} &\sum_{j', \mathbf{m}', \nu'} U(j', \mathbf{m}', \nu')\, \Sigma\, V^t(j', \mathbf{m}', \nu') \\ &\quad \equiv \sum_{\mathbf{k}} \sum_{j', \mathbf{m}', \nu'} U^{\mathbf{k}}_{n_1,n_1''}(j', \mathbf{m}', \nu') \sigma_{\mathbf{k}} V^{\mathbf{k}}_{n_2,n_2''}(j', \mathbf{m}', \nu'), \end{aligned} \tag{19}$$

Call K the maximum number of different values of the couple $\mathbf{k}$; if $k \le r$, then $K \le r^2$. Call M the maximum number of different values of the triple $(j', \mathbf{m}', \nu')$. To obtain the desired separated wave atom form we need not only compute those sums, but also

factor the result into its singular value decomposition,

$$\tilde{U}\tilde{\Sigma}\tilde{V}^t \equiv \sum_{\tilde{k}} \tilde{U}^{\tilde{k}}_{n_1,n_1''}\tilde{\sigma}_{\tilde{k}}\tilde{V}^{\tilde{k}}_{n_2,n_2''}, \tag{20}$$

where $\tilde{U}$ and $\tilde{V}$ are isometric matrices.

So we are faced with the problem of computing the SVD of a sum of matrices which are almost in SVD form—because the columns of $U(j',\mathbf{m}',\nu')$ and $V(j',\mathbf{m}',\nu')$ are not in general orthogonal. We will proceed in two steps:

- We start by turning each U and V into isometric matrices. For each $(j',\mathbf{m}',\nu')$, perform a QR decomposition to obtain an isometric matrix $Q_U(j',\mathbf{m}',\nu')$ and an upper triangular matrix $R_U(j',\mathbf{m}',\nu')$ such that

$$U(j',\mathbf{m}',\nu') = Q_U(j',\mathbf{m}',\nu')R_U(j',\mathbf{m}',\nu').$$

  Similarly, factor

$$V(j',\mathbf{m}',\nu') = Q_V(j',\mathbf{m}',\nu')R_V(j',\mathbf{m}',\nu').$$

  Gather the small factors in the middle and perform an SVD:

$$R_U(j',\mathbf{m}',\nu')\,\Sigma\,R_V^t(j',\mathbf{m}',\nu') = U^\sharp(j',\mathbf{m}',\nu')\,\Sigma^\sharp(j',\mathbf{m}',\nu')\,[V^\sharp]^t(j',\mathbf{m}',\nu').$$

  Put to zero the small singular values in $\Sigma^\sharp(j',\mathbf{m}',\nu')$ below some threshold τ, so as to keep at most $O(r)$ of them. Then compute

$$\begin{aligned} U(j',\mathbf{m}',\nu') &:= Q_U(j',\mathbf{m}',\nu')\,U^\sharp(j',\mathbf{m}',\nu'),\\ V(j',\mathbf{m}',\nu') &:= Q_V(j',\mathbf{m}',\nu')\,V^\sharp(j',\mathbf{m}',\nu'),\end{aligned}$$

  and for simplicity call $\Sigma(j',\mathbf{m}',\nu') := \Sigma^\sharp(j',\mathbf{m}',\nu')$. We have just orthogonalized (19) at the (benign) expense of making each Σ matrix depend on $(j',\mathbf{m}',\nu')$.
- We can now simplify the sum over $(j',\mathbf{m}',\nu')$. Let us group terms two by two and notice that a sum of two SVDs can be rewritten in matrix form as

$$U_1\,\Sigma_1\,V_1^t + U_2\,\Sigma_2\,V_2^t = \begin{pmatrix} U_1 & U_2\end{pmatrix}\begin{pmatrix}\Sigma_1 & 0\\ 0 & \Sigma_2\end{pmatrix}\begin{pmatrix}V_1^t\\ V_2^t\end{pmatrix}. \tag{21}$$

  The same strategy as above, involving two QR and one SVD decomposition, can be invoked to compute the SVD of the right-hand side in the above equation. The procedure can be applied to each couple of terms and repeated at the next level. This way, the whole sum (19) can be reduced in a binary fashion into its SVD form, leaving us with (20).

Standard linear algebra routines have been used for QR and SVD. It does not appear that iterative algorithms for sparse SVD offer any improvement, in the present context, over the standard algorithms.

3.2.3 Upscaled timestepping

Once the separated wave atom representation of the wave propagator at time τ is available, we can apply it to the initial condition as follows.

- Apply the wave atom transform to each component of the initial condition u_0.
- For each $j, \mathbf{m}, \nu$ and $j', \mathbf{m}', \nu'$, unfold the separated form to obtain the classical wave atom representation. Not all matrix elements in the classical form need to be computed, however. For a given (n_1, n_2), only a certain subset of positions subscripts (n_1', n_2') will be relevant—only those for which the wave has been given enough time to travel from x_μ to $x_{\mu'}$.
  - For times $\tau \lesssim 2^{-j}$, a wave atom cannot travel essentially farther than its own diameter, hence the restriction $|\mathbf{n}' - \mathbf{n}| \leq C$ for some constant C to be determined empirically.
  - For times $\tau \gtrsim 2^{-j}$, this rule becomes $|\mathbf{n}' - \mathbf{n}| \leq C \cdot t 2^j$.

  Once the restricted submatrices have been formed, we can compute the large matrix-vector product of the propagator with the coefficients of the initial condition.
- Apply an inverse wave atom transform to get $u(t)$ from its coefficients.

The same procedure, without the initial and final transforms, can be iterated to perform timestepping with the "upscaled" time step τ, generally much larger than the CFL timestep. As we will see from the complexity analysis in the next section, a reasonable choice of time step is $\tau \simeq \frac{1}{\sqrt{N}}$, whereas the CFL timestep is at most $\Delta t < \frac{1}{c_{\max} N}$ when $c(x) \leq c_{\max}$. We call "time upscaling" the possibility of using such a large time step, offered by an explicit precomputation of the propagator.

3.3 Complexity analysis

In this section we derive the total complexity for the repeated squaring scheme in separated form (RS), as well as the subsequent upscaled timestepping (UTS). We will first formulate the total computational cost as a function of N—the initial data is on an N-by-N grid—as well as the various values of the ϵ-ranks r of submatrices corresponding to different wave vectors $(j', \mathbf{m}', \nu')$. In later sections we will carefully analyze how those ranks themselves depend on N and on the geometry of the speed of sound, $c(x)$.

We would like to make clear that the complexity estimates we are about to derive refer to the total number of operations when we fix some small threshold ϵ below which the singular values of submatrices are discarded. In practice, we observe that ϵ is very well correlated to the overall L^2 accuracy of the method (see Sect. 4), although we do not prove the connection on a rigorous level in this paper. This observation of course assumes that the first time step in the initialization is itself made sufficiently accurate by taking Δt sufficiently small. Deriving accuracy estimates would imply dealing with sampling issues, namely that a function cannot be compactly supported both in space and in frequency. To obtain an L^2 accuracy $\overline{\epsilon}$ at time $T = 1$ we suspect that the threshold ϵ needs to depend on N and $\overline{\epsilon}$ like $O(\overline{\epsilon} N^{-1})$ as $N \to \infty$, in order to

compensate for the cumulative error introduced by repeated squaring. In this scenario, the power of N in each complexity estimates would need to be incremented by an arbitrarily small number δ—at the expense of a constant depending on the choice of δ.

As always, we assume that $c(x)$ is C^∞. We measure complexity in terms of elementary floating point operations (flops). Let us remark once and for all that the two token indices ν and ν' do not play a role in the complexity analysis since we are interested in asymptotic results—up to constants.

3.3.1 Initialization

We start by observing that the initialization step of the repeated squaring can be done in $O(N^2 \log N)$ steps. Indeed, applying an inverse wavelet packet transform to find $\psi^j_{m_1,n_1}(x_1)$ takes $O(N \log N)$ operations. Performing multiplication by a function or differentiation takes at most the time of a FFT, $O(N \log N)$. Finally, applying a direct transform costs $O(N \log N)$ again. Since there are $O(N)$ values of the indices (j, m_1, n_1), and a constant ϵ-separation rank C_ϵ, the overall complexity is $O(C_\epsilon N^2 \log N)$.

As we will see, initialization happens to make for a negligible fraction of the total computing time.

3.3.2 Matrix multiplication

Let us now consider the iteration step in the repeated squaring procedure, as described in Sect. 3.2.2. Fix a fine time step Δt and an upscaled time step $\tau = 2^{n^*}\Delta t$. We need to consider each scale j separately and recall that, by sparsity, we can always assume that j' is comparable to j. Ranks of submatrices are simply denoted by r, but let us keep in mind that they depend on the time step τ and the frequency indices $j, \mathbf{m}, j', \mathbf{m}'$, as well as on the desired accuracy level ϵ.

1. For fixed wave vectors, each subscript n_1 or n_2, or their counterpart with primes, takes on $O(2^j)$ values. Since both k and k' take on r values, a matrix such as $U^{\mathbf{k}}$ or $V^{\mathbf{k}}$ is of size $O(2^{2j}r^2)$. Each element of $U^{\mathbf{k}}$ or $V^{\mathbf{k}}$ takes $O(2^j)$ operations to compute, so the total complexity for forming two such matrices is $O(2^{3j}r^2)$.
2. One QR decomposition of $U^{\mathbf{k}}$ or $V^{\mathbf{k}}$ takes on $O(2^{2j}r^4)$ operations. Each middle factor $R_U \Sigma R_V^t$ is of size at most $O(r^2)$-by-$O(r^2)$, so the SVD of their product costs $O(r^6)$.
3. After performing the center SVD, only $O(r)$ singular values are kept (above the threshold ϵ), so we can trim both R_U and R_V to $O(r)$ columns. Computing each $U = Q_U R_U$ and $V = Q_V R_V$ takes $O(2^{2j}r^3)$ operations.
4. In the binary reduction using QR and SVDs, each matrix U has size $O(2^{2j})$-by-r, and let us grossly over-estimate r by its maximum $\max r$ over $(j', \mathbf{m}')$. Set M the total number of relevant indices $(j', \mathbf{m}')$. The total complexity for the binary reduction is $O(M(2^{2j}(\max r)^2 + (\max r)^3))$.

Steps 1, 2 and 3 above are to be repeated for each value of $(j', \mathbf{m}')$. In addition, there is an outer loop on the output wave vector $(j, \mathbf{m})$. The total complexity for one

time doubling in the separated wave-atom repeated squaring algorithm is therefore

$$\text{Compl}(RS, \text{one step}) \le C \cdot \sum_{j,\mathbf{m}} \sum_{j',\mathbf{m}'} \left(2^{3j} r^2 + 2^{2j} r^4 + r^6 + 2^{2j} (\max r)^2 + (\max r)^3 \right).$$

For each j there are $O(2^{2j})$ different values of $\mathbf{m}$. We can use the inequality

$$\sum_{j',\mathbf{m}'} r^p \le (\max r)^{p-1} \left(\sum_{j',\mathbf{m}'} r \right),$$

as well as the obvious $\max r \le \sum_{j',\mathbf{m}'} r$, simplify and obtain

$$\begin{aligned} &\text{Compl}(RS, \text{one step}) \\ &\quad \le C \cdot \sum_{j} \left[(2^{5j} \max r + 2^{4j} (\max r)^3 + 2^{2j} (\max r)^5) \left(\sum_{j',\mathbf{m}'} r \right) \right]. \end{aligned} \tag{22}$$

The number of time doublings is small and depends remarkably little on the choices we make for Δt and τ. As long as both quantities are taken to depend inverse polynomially on N, the number of grid steps per dimension, the number of time doublings is $O(\log N)$. The total complexity for the repeated squaring is therefore $\log N$ times the right-hand side in Eq. (22).

We give very precise estimates for $\max_{j',\mathbf{m}'} r$ in Sect. 5 and $\sum_{j',\mathbf{m}'} r$ in Sect. 6, as a function of τ, ϵ and j (uniformly over $\mathbf{m}$). In Sect. 7, we improve the bounds of Sects. 5 and 6 to take into account important special cases. In all cases, ranks and sums of ranks depend weakly on ϵ, namely they are all of order $O(\epsilon^{-1/M})$ for all $M > 0$ (with a constant depending on M.) This slow growth rate is the signature of spectral accuracy.

The simple choice $\tau = \frac{1}{\sqrt{N}}$, for example, is advantageous. We show in Theorem 2 that the worst-case estimate is $\sum r \simeq \max r \le C_\epsilon \cdot 2^j$. In that case, and with the choice $\tau = \frac{1}{\sqrt{N}}$, the total complexity becomes

$$\text{Compl}(RS, \text{worst}) \le C_\epsilon N^4 \log N.$$

For wave guides (when $c(x)$ depend only on x_1 or x_2, but not both,) the estimates become $\sum r \simeq \max r \le C_\epsilon \cdot 2^{j/2}$, and with $\tau = \frac{1}{\sqrt{N}}$ we have

$$\text{Compl}(RS, \text{wave guide}) \le C_\epsilon N^3 \log N.$$

Other choices for τ give rise to a variety of different complexity estimates. The dependence of complexity as a function of τ is studied in great detail in the sequel, is summarized in Fig. 12.

3.3.3 Upscaled timestepping

As we saw earlier, the complexity of the wave atom transform is $O(N^2 \log N)$.

The separated wave atom matrix needs to be unfolded into its classical form at every upscaled time step. For fixed wave vectors, the submatrix $E^{n_2,n_2'}_{n_1,n_1'}$ is of size $O(2^{2j})$-by-$O(2^{2j})$ and comes in separated form with rank r. Because of the restriction on nearby positions, all but $B_j = \max(2^j, \tau 2^{2j})$ rows and columns are kept, around $n_1' = n_1$ and $n_2' = n_2$. These rows and columns are easy to identify in the separated components as well, see Eq. (13), resulting in matrices u^k and v^k of size B_j-by-r. Explicitly forming $E^{n_2,n_2'}_{n_1,n_1'}$ from its separated components is a matrix-matrix product which takes $O(B_j^2 r)$ operations. The re-indexing of the relevant $O(B_j^2)$ elements into the classical form $E^{n_1',n_2'}_{n_1,n_2}$ takes $O(B_j^2)$ operations. Note that the latter matrix is band-diagonal with band $O(2^{-2j} B_j^2)$.

One matrix-vector product involving $E^{n_1',n_2'}_{n_1,n_2}$ then takes B_j^2 operations. Assuming that the solution $u(t)$ has a full set of wave atom coefficients (no particular sparsity pattern), then unfolding must be done for each wave vectors $(j', \mathbf{m}')$ (indexing columns) and $(j, \mathbf{m})$ (indexing rows), resulting in a total complexity

$$\text{Compl}(UTS, \text{one step}) \leq C \cdot \sum_{j,\mathbf{m}} B_j^2 \left(\sum_{j',\mathbf{m}'} r \right),$$

which can be rewritten more explicitly as

$$\text{Compl}(UTS, \text{one step}) \leq C \cdot \sum_{j} \left[(2^{4j} + \tau^2 2^{6j}) \left(\sum_{j',\mathbf{m}'} r \right) \right]. \tag{23}$$

Since T/τ upscaled time steps are necessary to reach time T (a multiple of τ), then the total complexity is the right-hand side of (23) multiplied by τ^{-1}.

For example, when τ is chosen as $\frac{1}{\sqrt{N}}$, then inspection of Theorems 2 and 5 reveals that the complexity estimate becomes

$$\text{Compl}(UTS, \text{worst case}) \leq C \cdot N^3$$

in the worst case, and

$$\text{Compl}(UTS, \text{wave guide}) \leq C \cdot N^{2.75}$$

in the case of wave guides. Estimates for different τ are summarized in Fig. 12. By comparison, recall that a pseudospectral method would be $O(N^3 \log N)$.

Complexity and computational times can yet be improved when the wave atom expansion of the solution is uniformly sparse in time. Assume that $u(t)$, $0 \leq t \leq T$,

can be approximated to accuracy ϵ in ℓ_2 using a fraction $\rho < 1$ of all wave numbers ξ_μ—not necessarily the same ones for different times. Then only the submatrices corresponding to those wave numbers must be computed at all, resulting in a direct net improvement of ρ of the complexity estimate for timestepping. For example, when $u(t)$ is a single bandlimited wavefront, then we can expect $\rho \simeq N^{-1/2}$. The corresponding total complexity becomes $O(N^{2.5})$ in general and $O(N^{2.25})$ for wave guides.

Complexity gains due to sparsity of the solution are harder to obtain for the repeated squaring, because it would demand identifying in advance which wave vectors are going to contribute in the yet unknown solution at dyadic times. These wave vectors are part of a "fat" manifold in phase-space. Such information could be obtained from a geometrical optics solver such as the phase-flow method [37], but we do not consider such a refinement in the present study.

Even though our complexity estimates may appear somewhat pessimistic, in particular for the core repeated squaring, it is worth keeping in mind that the result of the computation is not just one solution to the wave equation—*it is the whole Green's function in compressed form.* In particular, physical information of propagation of high frequencies can be read directly from the wave atom matrix representation.

4 Numerical implementation and examples

In this section, we apply the algorithm of Sect. 3.2 to several sample media. Theoretical studies of some of these representative media will be presented in Sect. 7. We used the orthonormal basis variation in all numerical experiments in this section.

We study four representative velocity fields defined over the unit square $[0, 1)^2$:

- Wave guide (Fig. 4a). The index of refraction is defined by

$$c^{-1}(x_1, x_2) = 1 + \exp\left(-64 \times \left(x_1 - \frac{1}{2}\right)^2\right).$$

- Bumps (Fig. 4b). The wave speed is a simple trigonometric polynomial,

$$c(x_1, x_2) = \frac{(3 + \sin(4\pi x_1)) \cdot (3 + \sin(4\pi x_2))}{16}.$$

- Converging lens (Fig. 4c). The index of refraction is given by

$$c^{-1}(x_1, x_2) = 1 + \exp(-64 \times \left(\left(x_1 - \frac{1}{2}\right)^2 + \left(x_2 - \frac{1}{2}\right)^2\right).$$

- Linear mirror (Fig. 4d),

$$c(x_1, x_2) = 0.75 + \rho(x_1, x_2) \cdot \left(x_1 - \frac{1}{2}\right)$$

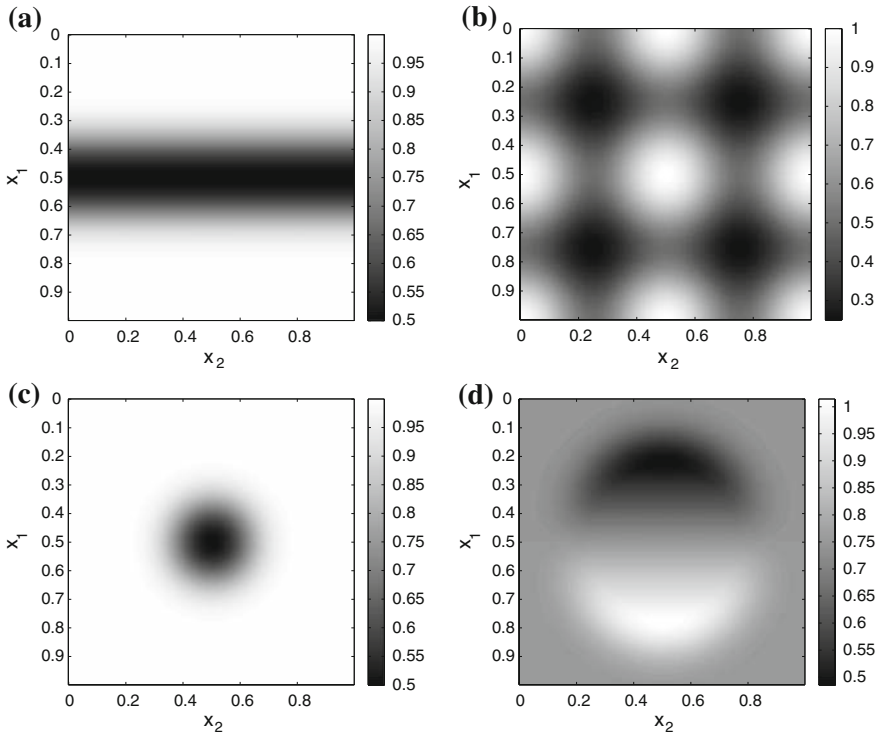

Fig. 4 Four representative acoustic media. **a** wave guide, **b** bumps, **c** Gaussian converging lens, and **d** linear mirror

where ρ is a radial window function which smoothly extracts the center part of the unit square $[0, 1)^2$.

In each of these four cases, we apply the algorithm presented in Sect. 3.2 to generate the propagator $E(\tau)$ at time $\tau = 1/16$. The initial time step Δt used is set to 2^{-10}. The thresholding constant ϵ for the size of discarded singular values is chosen to be 10^{-4}, and the grid size N is 128. As we pointed out already, the matrix $E(\tau)$ is organized as a collection of submatrices, which are indexed by row index $(j, \mathbf{m}, \nu)$ and column index $(j', \mathbf{m}', \nu')$. For each of the four media, the corresponding plot in Fig. 5 describes the time dependence of the number of their singular values greater than ϵ, i.e., their ϵ-rank. The solid curve is the maximum ϵ-rank over all submatrices, while the broken curve is the maximum of the sums of the ϵ-ranks over all column indices $(j', \mathbf{m}', \nu')$ (for a fixed row index $(j, \mathbf{m}, \nu)$). We compute these values at the dyadic time steps appeared in the construction $E(\tau)$, namely $t_n = 2^n \cdot \Delta t$, and linearly interpolate the value at other times.

We use two typical initial conditions to study our upscaled timestepping algorithm. The "line" initial condition (Fig. 6a) is a smoothed indicator of $\{(x_1, x_2) : x_2 = \frac{1}{2}\}$ while the "pulse" (Fig. 6b) is a smoothed delta function at the center of the domain. Both initial conditions, which have significant energy in the high frequency modes, are adequate for testing the numerical dispersion.

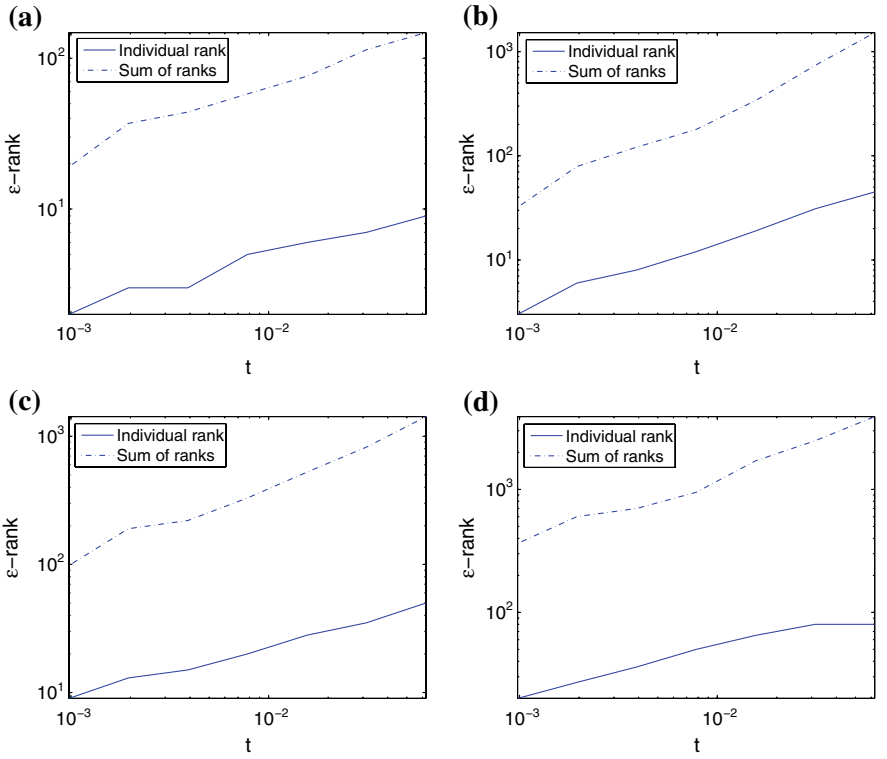

Fig. 5 ϵ-rank of the submatrices. **a** wave guide, **b** bumps, **c** Gaussian converging lens, and **d** linear mirror

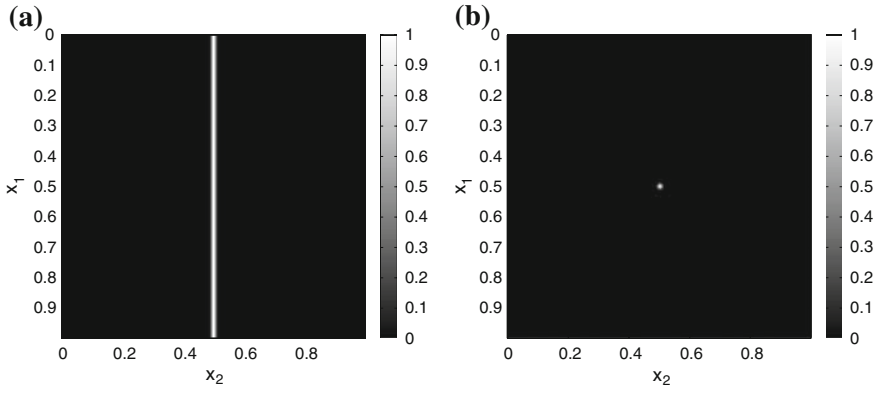

Fig. 6 Initial condition used in the upscaled time-stepping algorithm

For each acoustic medium, we apply the upscaled time-stepping algorithm on these two initial conditions. We are particularly interested in conservation of the energy and accuracy of the solution. Since we start from the equation

$$p_{tt} - c^2(x)\Delta p = 0,$$

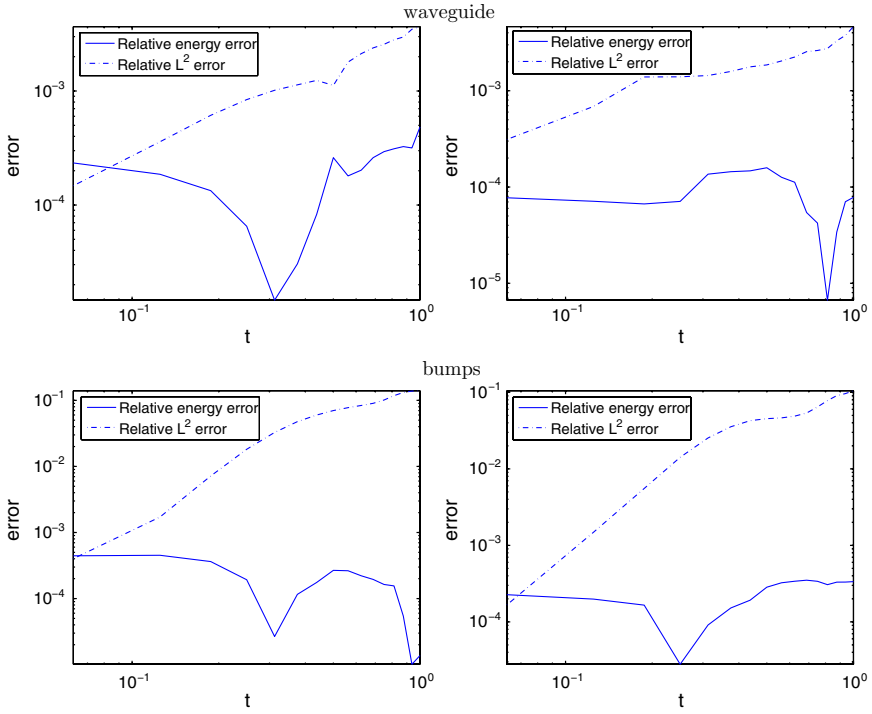

Fig. 7 Relative error of the energy integral and the wave field. *Left*: "line" initial condition. *Right*: "pulse" initial condition

the correct conserved energy is

$$\int \frac{|p_t|^2}{c^2(x)} + |\nabla p|^2 dx.$$

Figure 7 summarizes the time dependency of the relative errors of the energy integral (the solid curve) and the wave profile (the broken curve).

It is well known that standard finite difference methods for hyperbolic equations often suffer from the problem of excessive numerical dispersion [29]. This is particularly obvious when one uses a typical central-difference leapfrog scheme. In the following two experiments, we compare the numerical dispersion phenomenon in our upscaled time-stepping algorithm and the standard leapfrog algorithm. The time step and the grid size are chosen to be the same for both algorithms (Fig. 8).

The first experiment involves the waveguide acoustic media and the "pulse" initial condition. The three images in Fig. 9 show the solution at $t = 1/2$ and $t = 1$ computed using our method and the solution at $t = 1$ computed using the Leapfrog algorithm respectively. Notice that the ripples, which are the direct consequence of the numerical dispersion, are clearly observable in the leapfrog solution. The second experiment (Fig. 10), which uses the bump acoustic media and the "line" initial condition, demonstrates the same phenomenon.

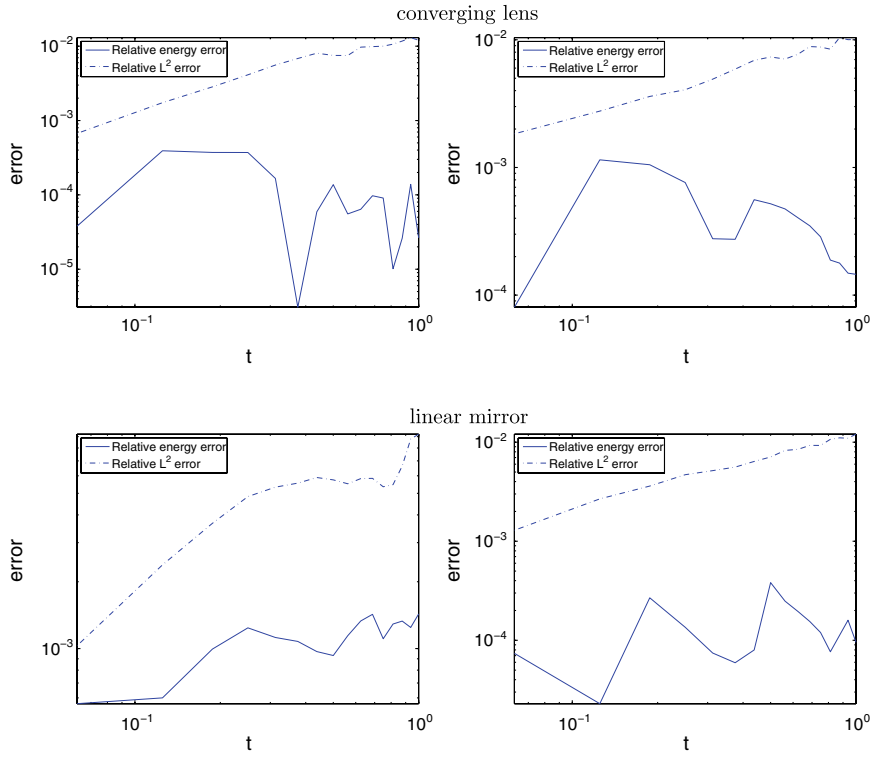

Fig. 8 Relative error of the energy integral and the wave field. *Left*: "line" initial condition. *Right*: "pulse" initial condition

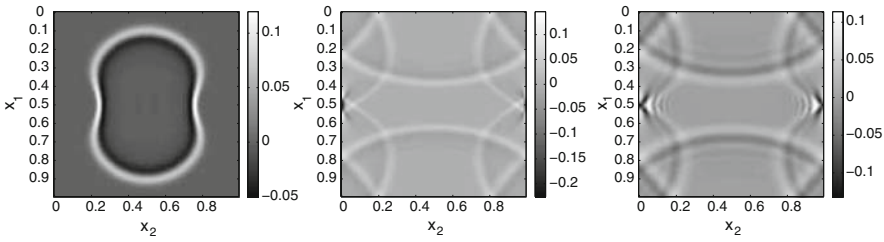

Fig. 9 Numerical dispersion. Waveguide media and "pulse" initial condition. *Left*: $t = 1/2$, wave atom method. *Middle*: $t = 1$, wave atom method. *Right*: $t = 1$, finite difference method

In the last experiment, we study the complexity of the upscaled time-stepping algorithm. As stated in Sect. 3.3, for certain types of acoustic media (e.g., wave guides), the upscaled timestepping algorithm has lower complexity compared to the standard spectral or pseudospectral methods, especially when the spatial discretization is refined. In fact, we are able to observe this fact even when N is relatively small. Figure 11 presents the time spent on applying a single upscaled time-step for various discretization size. For both the waveguide and bump media, the curve of the upscaled time-stepping algorithm grows much more slowly, and it becomes more efficient than the standard

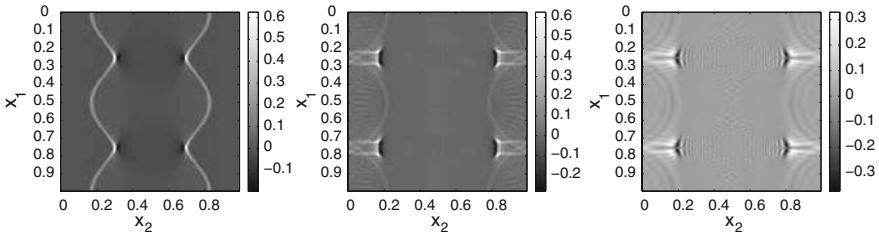

Fig. 10 Numerical dispersion. Bump media and "line" initial condition. *Left*: $t = 1/2$, wave atom method. *Middle*: $t = 1$, wave atom method. *Right*: $t = 1$, finite difference method

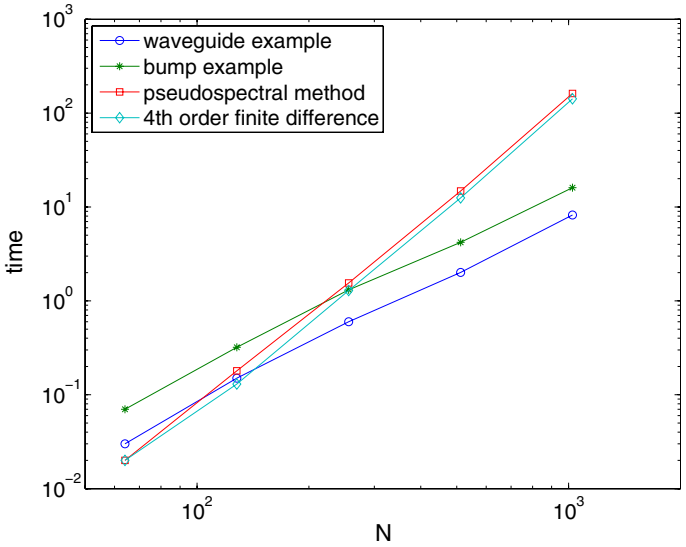

Fig. 11 Computational time of a single upscaled time step. In all cases, the small time step is $\Delta t = 1/1{,}024$

spectral method when N is larger than 256. In the examples that we have tried at $N = 1{,}024$, the efficiency gain is a factor 10–20 depending on the geometry of the smooth medium.These observations are in complete conformity with the complexity estimates in Sect. 3.3.

5 Rank estimates

The ϵ-rank r of a (possibly infinite) matrix A_{ij} is the smallest number r such that A_{ij} can be approximated up to accuracy ϵ by a matrix of rank r in ℓ^2,

$$\left\| A_{ij} - \sum_{k=1}^{r} u_i^k v_j^k \right\|_2 \leq \epsilon.$$

The ϵ-separation rank, or just ϵ-rank r of a function $f(x_1, x_2)$ is the smallest number of separated components $u^k(x_1)v^k(x_2)$ necessary to approximate $f(x_1, x_2)$ up to accuracy ϵ in L^2, i.e.,

$$\left\| f(x_1, x_2) - \sum_{k=1}^{r} u^k(x_1)v^k(x_2) \right\|_2 \leq \epsilon.$$

The main theoretical result of this section is a sharp bound on the ϵ-rank of reordered submatrices of the propagator in the wave atom frame. As detailed earlier each submatrix of interest has row index (n_1, n_2) versus column index (n_1', n_2'), but the separation isolates (n_1, n_1') versus (n_2, n_2'). Hence the necessity of *reordering* the entries, to prepare the submatrix for standard low-rank approximation. Notice that the size of the remainder, no more than ϵ, is however measured in ℓ^2 in the *original* form (n_1, n_2) versus (n_1', n_2'), in complete conformity with the goal of bounding the overall ℓ^2 norm of the error on the propagator.

Theorem 2 *Assume the velocity profile $c(x)$ is C^∞. Consider the submatrix $E_{j\mathbf{m}\nu;\, j'\mathbf{m}'\nu'}(t)$ obtained by fixing $j, \mathbf{m}, \nu$ and $j', \mathbf{m}', \nu'$ in the wave atom representation of the propagator $E(t)$. It is of size $O(2^{2j})$-by-$O(2^{2j})$, where $|\xi_\mu| = O(2^{2j})$ and at finest scale $2^{2j} \simeq N$. After reordering $(n_1, n_2; n_1', n_2') \to (n_1, n_1'; n_2, n_2')$, $E_{j\mathbf{m}\nu;\, j'\mathbf{m}'\nu'}(t)$ has ϵ-rank r bounded as follows.*

- *for $t \lesssim 2^{-j}$, $r \leq C_\epsilon \cdot (1 + t2^{2j})$,*
- *for $2^{-j} \lesssim t \lesssim 2^{-j/2}$, $r \leq C_\epsilon \cdot 2^j$,*
- *for $2^{-j/2} \lesssim t \leq T$, $r \leq C_\epsilon \cdot t^2 2^{2j}$,*

with $C_\epsilon \leq C_M \epsilon^{-1/M}$, for all $M > 0$, and C_ϵ also depends on T.

The various values taken on by the bound on r are summarized in Fig. 12. Notice that for large times the rank r is always obviously bounded by $C2^{2j}$, the size of the submatrix.

Before proving this result, we need to recall that the propagator $E(t)$ can be approximated by an oscillatory integral, called the Lax parametrix, [28] of the form

$$E(t)u(x) = \sum_{\ell=\pm} \int e^{i\Phi_\ell(x,\xi,t)} a_\ell(x, \xi, t)\hat{u}(\xi)\, d\xi + R_1(t)u(x).$$

This formula is only valid for small times $0 \leq t < T$ before caustics [18,26]. For smooth C^∞ media, each phase function is C^∞, positive-homogeneous of degree 1 in ξ, and solves a Hamilton–Jacobi equation,

$$\frac{\partial \Phi_\pm(x, \xi, t)}{\partial t} = \pm c(x)|\nabla_x \Phi_\pm(x, \xi, t)|, \quad \Phi_\pm(x, \xi, 0) = x \cdot \xi. \tag{24}$$

We assume that the hyperbolic system is written in the proper form so that each matrix-valued amplitude a_ℓ is a symbol of order 0 and type $(1, 0)$, i.e., componentwise,

$$|\partial_\xi^\alpha \partial_x^\beta a_\ell(x, \xi, t)| \leq C_{\alpha,\beta,t}(1 + |\xi|)^{-|\alpha|}, \quad \text{for all } \alpha.$$

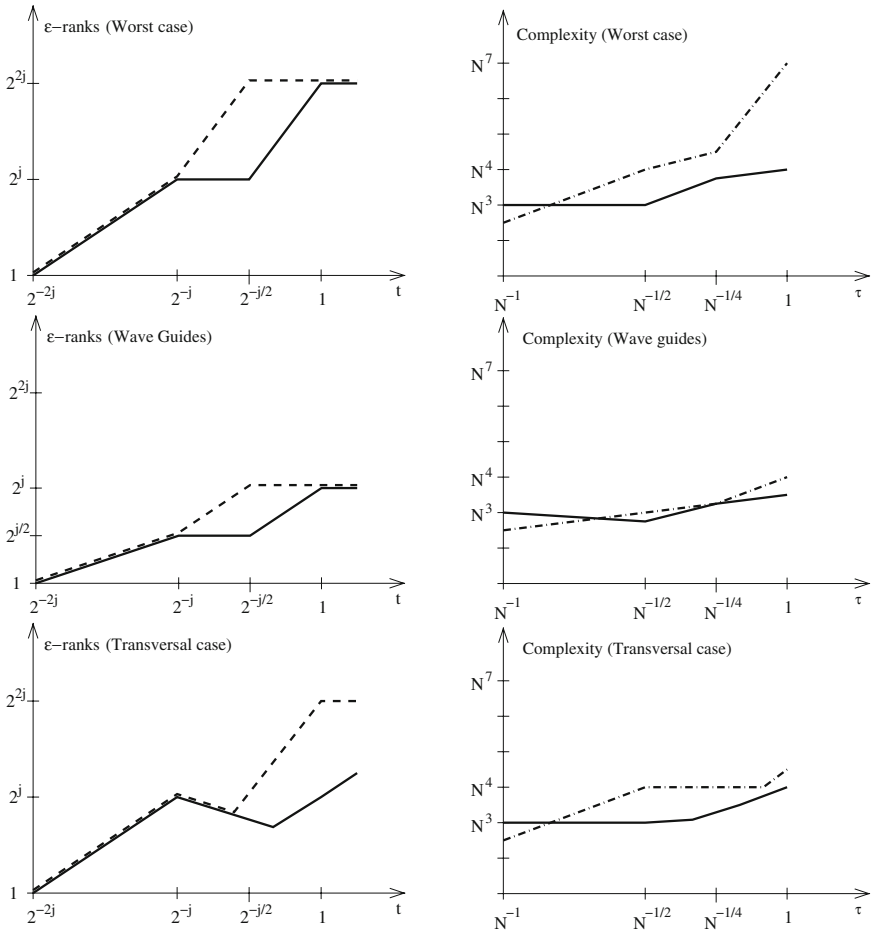

Fig. 12 *Left*: schematic illustration of the bounds on the ϵ-ranks of submatrices. *Solid line*: bound on individual ranks, as in Theorem 2. *Dashed line*: bound on sums of ranks over wave vectors, as in Corollary 2. *Right*: schematic illustration of the overall complexity. *Solid line*: upscaled time-stepping. *Dash-dotted line*: repeated squaring. All values of time, ranks and complexity should be understood modulo multiplicative constants and even $\log N$ factors. The portion of the t-axis $2^{-j} \lesssim \tau \lesssim 2^{-j/2}$ corresponds to the region of interest for numerical computations

This condition is denoted $a_\ell \in S^0$. The remainder $R_1(t)$ is smoothing in the sense that it turns tempered distribution into C^∞ functions. This is the same setting as in [8], to which we refer for details and justifications.

In what follows, we consider $x \in \mathbb{R}^2$, but $u(t, x)$ with support in a subset Ω inside the open unit square $]0, 1[^2$. Without loss of generality, we can make $a_\ell(x, \xi, t)$ compactly supported in x through multiplication by an adequate cutoff equal to one on Ω and tapering smoothly to zero outside $[0, 1]^2$.

We will need the following two simple lemmas in the sequel.

Lemma 2 *Let $f \in C^\infty(]0, 1[^2)$. Then the ϵ-separation rank of f obeys $r_\epsilon(f) \leq C_M \epsilon^{-1/M}$ for all $M > 0$, and some constant $C_M > 0$. Furthermore, for each positive*

integer $s > 0$ there exists a constant C_s such that the same rank-$r_\epsilon(f)$ decomposition has error $C_s\epsilon$ when measured in the Sobolev space $W^{s,p}$, for all $1 \le p \le \infty$.

Proof It suffices to notice that the Fourier series coefficients of f decay like

$$|\hat{f}[k]| \le C_M |k|^{-M}.$$

Each Fourier mode is separable. Truncating the Fourier series to $|k| \le K$, by means of $O(K^2)$ terms, results in a squared L^2 error

$$\sum_{|k|>K} |\hat{f}[k]|^2 \le C_M^{(0)} K^{-M}.$$

for all $M > 0$, and some other constant $C_M^{(0)} > 0$. Hence K can be chosen less than $C_M \epsilon^{-1/M}$ for some adequate choice of constant C_M.

The expression of the square of the error in $W^{s,2}$ is (up to constants)

$$\sum_{|k|>K} |k|^{2s} |\hat{f}[k]|^2 \le C_M^{(s)} K^{-M},$$

for all $M > 0$ but where the constant $C_M^{(s)}$ is likely larger than $C_M^{(0)}$. Since $K \le C_M \epsilon^{-1/M}$, we certainly have a $W^{s,2}$ error bounded by $C_s\epsilon$ for some constant C_s depending on s.

When $p \ne 2$, we conclude using the continuous Sobolev inclusion $W^{2+s,2} \subset W^{s,p}$, valid for all $1 \le p \le \infty$ in two dimensions. □

Lemma 3 *Let $\Phi(x, \xi, t)$ solve either Hamilton–Jacobi equation (24). Then, for $0 \le t \le T$ while the equation is well posed, the Hessian obeys*

$$\nabla_x \nabla_x \Phi(x, \xi, t) = t\psi(x, \xi, t),$$

where each component of ψ is C^∞ away from $\xi = 0$, and positive-homogeneous of degree 1 in ξ.

Proof Write $\nabla_x \nabla_x \Phi(x, \xi, t)$ as the integral $\pm \int_0^t \nabla_x \nabla_x c(x) |\nabla_x \Phi(x, \xi, s)|\, ds$, where the integrand has the same smoothness and homogeneity properties as Φ itself. □

Proof of Theorem 2 Let $\epsilon > 0$. We seek a bound on the number r of separated terms in

$$\langle \varphi_\mu \mathbf{e}_\nu, E(t) \varphi_{\mu'} \mathbf{e}_{\nu'} \rangle = \sum_{k=1}^{r} u^k_{n_1, n_1'}(t) v^k_{n_2, n_2'}(t) + R_{\mathbf{n},\mathbf{n}'}(t),$$

where $R_{\mathbf{n},\mathbf{n}'}(t)$, as a matrix with row subscript $\mathbf{n}$ and column subscript $\mathbf{n}'$, has ℓ^2 norm less than ϵ. Note that all the quantities r, u, v and R depend on the parameters j, $\mathbf{m}$, ν and j', $\mathbf{m}'$, ν', but we drop this dependence for simplicity of notations.

By construction, wave atoms have rank 2 in the frequency domain, namely

$$\hat{\varphi}_\mu(\xi) = \left[\hat{\psi}^j_{m_1,+}(\xi_1)\hat{\psi}^j_{m_2,+}(\xi_2) + \hat{\psi}^j_{m_1,-}(\xi_1)\hat{\psi}^j_{m_2,-}(\xi_2) \right] e^{-i2^{-j}n_1\xi_1} e^{-i2^{-j}n_2\xi_2},$$

Without loss of generality, at the expense of at most quadrupling the constants in front of each estimate, we will drop the second term in the above parenthesis. This results in considering each wave atom as having only one separated bump in the frequency plane, i.e, having rank 1. We keep the notation $\hat{\varphi}_\mu(\xi)$ for these "amputated" atoms.

Call S_μ the support of $\hat{\varphi}_\mu(\xi)$; it can be inscribed in a ball centered at ξ_μ, and of radius equal to $2^{j+1}\sqrt{2}\pi$. We will denote by $\chi_\mu(\xi)$ a smooth and separable indicator function, equal to one on S_μ, and zero on the complement of the larger set $S_\mu + \{\xi : |\xi| \le 2^j\}$.

Denote by $E(\xi, \eta, t)$ the frequency kernel of $E(t)$, namely

$$\widehat{E(t)u}(\xi) = \int E(\xi, \eta, t)\hat{u}(\eta)\, d\eta.$$

By Parseval, the matrix elements are

$$\begin{aligned}\langle \varphi_\mu, E(t)\varphi_{\mu'} \rangle = \int\int \hat{\psi}^j_{m_1,+}(\xi_1)\hat{\psi}^j_{m_2,+}(\xi_2) e^{-i2^{-j}n_1\xi_1} e^{-i2^{-j}n_2\xi_2} E(\xi, \eta, t) \\ \times \overline{\hat{\psi}^{j'}_{m'_1,+}(\eta_1)\hat{\psi}^{j'}_{m'_2,+}(\eta_2)} e^{i2^{-j'}n'_1\eta_1} e^{-i2^{-j'}n'_2\eta_2}\, d\xi_1 d\xi_2 d\eta_1 d\eta_2, \quad (25)\end{aligned}$$

where the kernel is

$$E(\xi, \eta, t) = K(\xi, \eta, t) + R_1(\xi, \eta, t), \quad (26)$$

where

$$K(\xi, \eta, t) = \sum_{\ell=\pm} \int e^{i(\Phi_\ell(x,\eta,t) - x\cdot\xi)} a_\ell(x, \eta, t)\, dx. \quad (27)$$

In what follows we will drop the sum and the subscript ℓ (at the expense of doubling the separation rank,) because $\ell = -$ is totally analogous to $\ell = +$. We now seek results of separation of $K(\xi, \eta, t)$ in both ξ and η, on the frequency support of each wave atom, i.e.,

$$\chi_\mu(\xi)\chi_{\mu'}(\eta)K(\xi, \eta, t) = \sum_{k=1}^{r} K^{(1)}_k(\xi_1, \eta_1, t)K^{(2)}_k(\xi_2, \eta_2, t) + R_2(\xi, \eta, t). \quad (28)$$

The following lemma shows that the size of the remainder R_2 directly translates into a remainder of comparable size for the submatrix of interest.

Lemma 4 *Let $T(\xi, \eta)$ be any kernel defining by extension a bounded operator T on $L^2(\mathbb{R}^2)$. As usual, we denote $\mu = (j, \mathbf{m}, \mathbf{n})$ and $\boldsymbol{\varphi}_{\mu\nu} = \varphi_\mu \mathbf{e}_\nu$. For any $(j, \mathbf{m}, \nu)$ and $(j', \mathbf{m}', \nu')$,*

$$\| \langle \boldsymbol{\varphi}_{\mu\nu}, T\boldsymbol{\varphi}_{\mu'\nu'} \rangle \|_{\ell^2_{\mathbf{n}'} \to \ell^2_{\mathbf{n}}} \leq \|T\|_{L^2_{\eta\nu'} \to L^2_{\xi\nu}} \tag{29}$$

Proof By the tight frame property,

$$\|T\|_{L^2_{\eta\nu'} \to L^2_{\xi\nu}} = \| \langle \boldsymbol{\varphi}_{\mu\nu}, T\boldsymbol{\varphi}_{\mu'\nu'} \rangle \|_{\ell^2_{\mu'\nu'} \to \ell^2_{\mu\nu}}.$$

Choose any set of indices $(j, \mathbf{m}, \nu)$ and $(j', \mathbf{m}', \nu')$. The conclusion follows by restricting the wave atoms matrix of T to rows indexed by $(j, \mathbf{m}, \nu)$ and columns indexed by $(j', \mathbf{m}', \nu')$. □

Notice that the remainder R_1 does not pose any difficulty. Since it corresponds to a smoothing operator on a bounded domain, we have the bound

$$|R_1(\xi, \eta, t)| \leq C_M (1 + |\xi| + |\eta|)^{-M}, \quad \text{for all } M > 0,$$

so, in the spirit of Lemmas 2 and 4, a constant number $C_\epsilon \sim \epsilon^{-1/M}$ of (separable) Fourier modes suffices to approximate the submatrix coming from R_1 to accuracy $\epsilon/2$ in $\ell^2_{\mathbf{n}'} \to \ell^2_{\mathbf{n}}$.

It is also important to notice that for each fixed wave vector ξ_μ, only a few wave vectors $\xi_{\mu'}$ give rise to nonnegligible matrix elements. This is due to sparsity, and quantified in Sect. 6. For the time being we only need to observe that, for those nonnegligible entries, the wave vectors are comparable; very conservatively, $|j - j'| \leq$ const. Also, the particular value of $\mathbf{m}$ will be seen not to play any significant role. As a consequence, ϵ-separation ranks essentially only depend on one of the two numbers j and j', say j. In the sequel, we will look for a bound on r which depends solely on j, understanding that it holds uniformly over all j', $\mathbf{m}'$ and $\mathbf{m}$.

We are now ready to split the proof into three parts, corresponding respectively to (1) coarse scales, (2) fine scales in the regions of nonstationary phase, and (3) fine scales near the locus of stationary phase.

5.1 Coarse scales

The case of coarse scales, i.e., say $j = j' = 0$, needs to be considered separately because the phase $\Phi(x, \xi, t)$ has in general a kink at the origin in ξ, that is, a discontinuity in the gradient.

Let $g(x, \eta, t) = e^{i\Phi(x,\eta,t)} a(x, \eta, t) \chi_\mu(\eta)$, so that

$$K(\xi, \eta, t) = \int e^{-ix\cdot\xi} g(x, \eta, t)\, dx = \hat{g}(\xi, \eta, t),$$

where the Fourier transform is taken over the first variable only.

Take $\{\psi_\lambda\}$ a 2D separable wavelet orthonormal basis, with super-algebraic decay in both space and frequency, and expand $g(x,\cdot,t)$:

$$g(x,\eta,t) = \left(\sum_{\lambda\in\Lambda_1} + \sum_{\lambda\in\Lambda_2}\right) c_\lambda(x,t)\psi_\lambda(\eta).$$

Determine the subset of subscripts Λ_2 such that

$$\sup_x \sup_{0\le t\le T} \sum_{\lambda\in\Lambda_2} |c_\lambda(x,t)|^2 \le \frac{\epsilon^2}{16\pi^2}.$$

Since $\nabla_\eta g(x,\cdot,t)$ is discontinuous at the origin, but otherwise C^∞ and compactly supported in a $O(1)$ region, it is a classical result from wavelet analysis that

$$|\Lambda_1| \le C_M \epsilon^{-1/M}, \quad \text{for all } M > 0.$$

This constant number of important subscripts in Λ_1 correspond to large scales as well as locations near the singularity.

Therefore,

$$K(\xi,\eta,t)\chi_\mu(\eta) = \sum_{\lambda\in\Lambda_1} \hat{c}_\lambda(\xi,t)\psi_\lambda(\eta) + R_2(\xi,\eta,t).$$

Each coefficient $c_\lambda(x,t)$ inherits the C^∞ smoothness of g, and is essentially supported near the unit cube. By Lemma 2, the ϵ-separation rank of $\hat{c}_\lambda(\xi,t)$ is therefore $O(\epsilon^{-1/M})$ for all $M > 0$. In addition, each $\psi_\lambda(\eta)$ is separable, and the sum runs over at most $O(\epsilon^{-1/M})$ terms. So the overall separation rank for the sum is $O(\epsilon^{-1/M})$ as well.

Let us now check that $R_2(\xi,\eta,t)$ generates an error which is the correct fraction of ϵ in L^2. The squared Hilbert–Schmidt norm of R_2, which bounds the squared L^2 norm, is

$$\begin{aligned}
\int\int |R_2(\xi,\eta,t)|^2 \, d\xi d\eta &= \int \sum_{\lambda\in\Lambda_2} |\hat{c}_\lambda(\xi,t)|^2 \, d\xi \quad \text{by Plancherel-wavelets,} \\
&= \sum_{\lambda\in\Lambda_2} \int |c_\lambda(x,t)|^2 \, dx \quad \text{by Plancherel-Fourier,} \\
&\le \sup_{x\in[0,1]^2} \sum_{\lambda\in\Lambda_2} |c_\lambda(x,t)|^2 \le \frac{\epsilon^2}{16\pi^2}.
\end{aligned}$$

We can now apply Lemma 4 to obtain a remainder of size $\epsilon/2$ for each wave atom submatrix of R_2, with indices $\mathbf{n}, \mathbf{n}'$. Together with R_1's submatrices, also of size $\epsilon/2$, the overall remainder is of size at most ϵ. This finishes the proof for the coarse scales, with the result that $r = O(\epsilon^{-1/M})$.

5.2 Fine scales, stationary phase

Consider μ and μ' such that $\chi_\mu(\xi) = \chi_{\mu'}(\xi) = 0$ in a neighborhood of the origin $\xi = 0$. We expect that the main contribution to the integral in Eq. (27) comes from the points of stationary phase, i.e. $\xi = \nabla_x \Phi(x, \eta, t)$. For each $\delta > 0$, consider the sets

$$X_\xi^\eta(\delta) = \{x \in [0,1]^2;\, |\xi - \nabla_x \Phi(x, \eta, t)| \le 2^j \delta\},$$

and their union

$$X_\mu^{\mu'}(\delta) = \{x \in [0,1]^2; \text{there exist } \xi \in \text{supp}\chi_\mu \text{ and } \eta \in \text{supp}\chi_{\mu'},\ |\xi - \nabla_x \Phi(x, \eta, t)| \le 2^j \delta\}. \tag{30}$$

Our aim is to find a smooth indicator $p(x)$ equal to one for $x \in X_\mu^{\mu'}(\delta)$, and for which the restricted kernel

$$K_{\text{nonstat}}(\xi, \eta, t) = \int (1 - p(x)) e^{i\Phi(x,\eta,t) - ix\cdot\xi} a(x, \eta, t)\, dx$$

is negligible in the L^2 sense, $\|K_{\text{nonstat}}\|_2 \le \epsilon/4$. We will see in the next section that such an estimate holds provided δ is chosen large enough; let us accept for the moment that it can be taken of the form $\delta = O(\epsilon^{-1/M})$.

In this section we show how to build $p(x)$ as a sum of functions $q_k(x)$, which define kernels

$$K_k(\xi, \eta, t) = \int q_k(x) e^{i\Phi(x,\eta,t) - ix\cdot\xi} a(x, \eta, t)\, dx, \tag{31}$$

such that each $K_k(\xi, \eta, t)\chi_\mu(\xi)\chi_{\mu'}(\eta)$ has ϵ-separation rank of order $O(\epsilon^{-1/M})$ in ξ and η, for all $M > 0$. An estimate on the overall rank is then expected, for then

$$K = K_{\text{nonstat}} + \sum_{k=1}^{N_B} K_k \tag{32}$$

will be well separated by $O(N_B \epsilon^{-1/M})$ terms for all $M > 0$. In the rest of this section we intend to estimate N_B as a function of t as well as justify smallness of the non-separated remainder.

The first observation is that the union in the definition of $X_\mu^{\mu'}(\delta)$ is not essential. More precisely, let σ be the Lyapunov exponent of the bicharacteristic Hamiltonian system,

$$\sigma = \sup_{t \ge 0} \frac{1}{t} \log \left(\sup_{x \in [0,1]^2} \sup_{\xi \in \mathbb{R}^2} |\nabla_x \nabla_\xi \Phi(x, \xi, t)| \right).$$

For any $\xi_0, \xi \in \mathrm{supp}\chi_\mu$ and $\eta_0, \eta \in \mathrm{supp}\chi_{\mu'}$, we have the estimates

$$|\xi_0 - \xi| \le C2^j \quad \text{and} \quad |\eta_0 - \eta| \le C'2^{j'} \le C2^j.$$

A Taylor expansion of Φ around η_0 then reveals

$$X_\mu^{\mu'}(\delta) \subset X_{\xi_0}^{\eta_0}(Ce^{\sigma t}\delta). \tag{33}$$

This observation is important because it shows that the condition $|\xi - \nabla_x \Phi(x, \eta, t)| \le 2^j\delta$ is the strongest definition of the neighborhood of the locus of stationary phase which still makes it independent of ξ and η.

The next step is to linearize the phase $\Phi(x, \eta, t)$ in η near some point $\eta_0 \in \mathrm{supp}\chi_{\mu'}$. The whole point of partitioning the frequency plane into indicators of radius $O(2^j)$, when $|\eta| \sim 2^{2j}$, is precisely to make the remainder non-oscillatory. More precisely, for $\eta \in \mathrm{supp}\chi_{\mu'}$, homogeneity of degree one in η implies the estimate

$$\partial_\eta^\alpha \left[\Phi(x, \eta, t) - \eta \cdot \nabla_\eta \Phi(x, \eta_0, t)\right] = O(|\eta|^{-|\alpha|/2}).$$

For a proof, see [8, Appendix B (p. 55)]; or [36, Chapter IX, pp. 406–407]. This nonlinear remainder can be absorbed in the amplitude, which we still denote $a(x, \eta, t)$ for simplicity,

$$a(x, \eta, t)\chi_{\mu'}(\eta) := e^{i(\Phi(x,\eta,t) - \eta\cdot\nabla_\eta\Phi(x,\eta_0,t))}a(x, \eta, t)\chi_{\mu'}(\eta),$$

without essentially changing its properties: the new amplitude $a\chi_{\mu'}$ is still of order zero and type $(1/2, 0)$, i.e.,

$$|\partial_\eta^\alpha \partial_x^\beta a(x, \eta, t)\chi_{\mu'}(\eta)| \le C_{\alpha,\beta}(1 + |\eta|)^{-|\alpha|/2}. \tag{34}$$

The central argument now consists in performing Taylor expansions of the (linearized) phase in x within adequately small balls $B_{x_k}(\rho_k)$. Call $f(x) = \eta \cdot \nabla_\eta \Phi(x, \eta_0, t)$. Then

$$f(x) = f(x_k) + (x - x_k)\nabla f(x_k) + \frac{1}{2}(x - x_k)^t \nabla\nabla f(y)(x - x_k), \tag{35}$$

where $x, y \in B_{x_k}(\rho_k)$, and $\nabla\nabla f$ denotes the Hessian. The first genuinely non-separable contribution comes from the off-diagonal quadratic term x_1x_2. We can still have control over this term if we make it nonoscillatory, i.e., if we take ρ_k small enough that

$$\rho_k^2|\nabla\nabla f(x)| \le C \quad \text{for } x \in B_{x_k}(\rho_k). \tag{36}$$

The point is that the constant C is independent of j. The quadratic term can then be absorbed in the amplitude without essentially changing the latter, as was done previously for the linearization in η.

We are then led to the geometric problem of covering the set $X_\mu^{\mu'}(\delta)$ with the smallest possible number of balls $B_{x_k}(\rho_k)$ in which the quadratic term is non-oscillatory. Let us first lighten notations by writing $g(x)$ for either $\frac{\partial}{\partial x_1}\nabla_\eta\Phi(x,\eta_0,t)$ or $\frac{\partial}{\partial x_2}\nabla_\eta\Phi(x,\eta_0,t)$. Uniform boundedness of the quadratic term, as above, can be expressed as

$$|\nabla g(x)| \le C\cdot 2^{-2j}\rho_k^{-2}. \tag{37}$$

As we saw in Eq. (33), the condition $x \in X_\mu^{\mu'}(\delta)$ can be reduced to $x \in X_\xi^{\eta_0}(C\cdot\delta)$, which in turn reads

$$|g(x)| \le C\cdot 2^{-j}. \tag{38}$$

Notice that $g(x)$, like the sound speed $c(x)$, is C^∞ for times $t < T$ before breakdown of the Hamilton–Jacobi equation on plane wave initial conditions. We then claim that, for any such smooth $g(x)$, the set where (38) holds can be covered by $N_B = O(2^j)$ balls in which (37) holds. The construction of such a covering necessarily depends on $g(x)$ itself, so we apologize to the reader for the following argument being a bit technical.

We switch to a continuous description of the problem by introducing a local *ball radius density* $\rho(x,j)$ which will help determine $\rho_k = \rho(x_k,j)$ at a collection of points x_k still to be determined. We set

$$\rho(x,j) = \frac{1}{\sqrt{2^j + 2^{2j}|\nabla g(x)|}}. \tag{39}$$

Two basic properties motivate this formula, namely that

- $|\nabla g(x_k)| \le 2^{-2j}\rho(x_k,j)^{-2}$, as required, and
- $C\cdot 2^{-j} \le \rho(x,j) \le 2^{-j/2}$, for all $x\in[0,1]^2$.

It is important for what follows to check that formula (39) is consistent as a definition of local radius, in the sense that

$$\sup_{x\in B_{x_k}(\rho_k)} \rho(x,j) \le C_{c.o.}\cdot\rho(x_k,j). \tag{40}$$

This result is an easy consequence of Landau's inequality and is justified in the appendix. We call it the *constant overlap property*.

The collection of ball centers x_k is now determined as follows. Start from a Cartesian lattice $y_k = (k_1,k_2)b2^{-j}$ with k_1,k_2 integers and some small $b>0$ to be determined. Assign a ball of center $\tilde\rho_k = \rho_k/5 = \rho(x_k,j)/5$ to y_k. The constant b is taken so that the union of all the balls $B_{y_k}(\tilde\rho_k)$ covers $[0,1]^2$. In general the balls significantly overlap and the covering needs pruning, for instance by means of the following elementary covering lemma.

Lemma 5 *Let $\mathcal{G} = \{B_{y_k}(\tilde{\rho}_k)\}$ be a family of closed balls with uniformly bounded radius. Then there is a subfamily $\mathcal{F} \subset \mathcal{G}$ of pairwise disjoint balls such that*

$$\bigcup_{B_{y_k}(\tilde{\rho}_k)\in\mathcal{G}} B_{y_k}(\tilde{\rho}_k) \subset \bigcup_{B_{y_k}(\tilde{\rho}_k)\in\mathcal{F}} B_{y_k}(5\tilde{\rho}_k).$$

Proof See [38, p. 7]. □

The collection x_k then emerges as the centers of the remaining balls and the radii are chosen as $\rho_k = 5\tilde{\rho}_k$.

Notice that, by construction, each point in the unit square is covered by at most a constant number of balls $B_{x_k}(\rho_k)$ (independent of j or ρ_k). This is because the constant overlap property (40) can be iterated to yield

$$\sup_{x\in B_{x_k}(2\rho_k)} \rho(x, j) \asymp \rho(x_k, j).$$

(The notation $A \asymp B$ means $A \le C \cdot B$ and $B \le C \cdot A$ for some positive C which may depend on some parameters, depending on context.) The balls overlapping with $B_{x_k}(\rho_k)$ therefore have radius comparable to ρ_k, so there can only be a constant number of them.

We are now ready to estimate the number N_B of balls which cover $X \equiv X_\mu^{\mu'}(\delta)$. To every lattice point y_k, assign a weight

$$w_k = 2^{-2j} \sum_{x_\ell\in X : y_k\in B_{x_\ell}(\rho_\ell)} \frac{1}{\rho_\ell^2}. \tag{41}$$

Since there are $O(2^{2j}\rho_\ell^2)$ grid points y_k inside the ball $B_{x_\ell}(\rho_\ell)$, it is straightforward to check that

$$N_B \le C \cdot \sum_k w_k.$$

On the other hand, the constant overlap property [Eq. (40)] entitles us to see $\sum_k w_k$ as a Riemann sum and bound

$$\sum_k w_k \le C \cdot \int_X \frac{1}{\rho^2(x, j)}\, dx.$$

Using the definition (39), we get

$$N_B \le C \cdot \left(2^j + 2^{2j} \int_X |\nabla g(x)|\, dx\right).$$

We claim that $\int_X |\nabla g(x)|\,dx \le C \cdot 2^{-j}$. This fact follows from the following lemma, which is a simple reformulation of the co-area formula for BV functions. For our application, we let $\tilde{\epsilon} = 2^{-j}$.

Lemma 6 *Let $g \in C^2([0,1]^2)$. For all $\tilde{\epsilon} > 0$, let $X_{\tilde{\epsilon}} = \{x \in [0,1]^2, |g(x)| \le \tilde{\epsilon}\}$. Then*

$$\int_{X_{\tilde{\epsilon}}} |\nabla g(x)|\,dx \le C \cdot \tilde{\epsilon},$$

where $C = 2\sup_{t\in\mathbb{R}} H^1(\partial X_t)$ and H^1 is the Hausdorff measure, or length.

Proof See Appendix A. □

Note that $\sup_{t\in\mathbb{R}} H^1(\partial X_t)$ may in principle be infinite for generic C^2 functions, but in our case this quantity remains bounded by the smoothness properties of solutions of Hamilton–Jacobi equations with C^∞ coefficients, for small times.

We have shown that $N_B \le C \cdot 2^j$. Let us now translate this result into a separation rank for the kernel $K(\xi,\eta,t)$, by means of the smooth partition of unity $q_k(x)$ already alluded to earlier in this section. Specifically, take a C^∞ function $\chi(x)$ such that $\chi(x) > 0$ for $|x| < 1$ and $\chi(x) = 0$ for $|x| \ge 1$. Consider the collection x_k of all ball centers, including those outside the set $X_\mu^{\mu'}(\delta)$. Then for each x_k define

$$\tilde{q}_k(x) = \chi\left(\frac{x - x_k}{\rho_k}\right).$$

By Lemma 2, each $\tilde{q}_k(x)$ has ϵ-separation rank of order $O(\epsilon^{-1/M})$ for all $M > 0$. The partition of unity is then, in the usual manner, defined as

$$q_k(x) = \frac{\tilde{q}_k(x)}{\sum_k \tilde{q}_k(x)}.$$

The constant overlap property, valid in a neighborhood of $X_\mu^{\mu'}(\delta)$, ensures that the smoothness constants of $q_k(x)$ are comparable to those of $\tilde{q}_k(x)$, as long as x_k is in or near $X_\mu^{\mu'}(\delta)$. As a matter of illustration, Lemma 2 would apply to those $q_k(x)$ near $X_\mu^{\mu'}(\delta)$ and yields an ϵ-separation rank of order $O(\epsilon^{-1/M})$ for all $M > 0$. (In truth, we will apply Lemma 2 later to a more complicated amplitude involving $q_k(x)$).

At this point, recall that we are trying to separate the restricted kernel (31) on supp $\chi_\mu \times$ supp $\chi_{\mu'}$, that we have linearized the phase in η and that we are linearizing it in x as in Eq. (35). The point of $q_k(x)$ is that the quadratic contribution can be absorbed in the amplitude without changing the symbol properties of the latter [Eq. (34)]. The new amplitude a_k is defined from

$$q_k(x)a_k(x,\eta,t)\chi_{\mu'}(\eta) = q_k(x)e^{\frac{i}{2}(x-x_k)^t\nabla_x\nabla_x\eta\cdot\nabla_\eta\Phi(y(x),\eta_0,t)(x-x_k)}a(x,\eta,t)\chi_{\mu'}(\eta).$$

The constant and linear contributions to the phase are

$$\eta \cdot \nabla_\eta \Phi(x_k, \eta_0, t) + (x - x_k) \cdot \nabla_x (\eta \cdot \nabla_\eta \Phi(x_k, \eta_0, t)) - x \cdot \xi.$$

The first term, $\eta \cdot \nabla_\eta \Phi(x_k, \eta_0, t)$, and the third term $-x_k \cdot \nabla_x (\eta \cdot \nabla_\eta \Phi(x_k, \eta_0, t))$ are both independent of x and separable in η, so we can ignore them. What remains is a modified kernel of the form

$$K_k(\xi, \eta, t) = \int q_k(x) e^{ix \cdot (A(t)\eta - \xi)} a_k(x, \eta, t)\, dx, \tag{42}$$

where $A(t) = \nabla_x \nabla_\eta \Phi(x_k, \eta_0, t)$. For sufficiently small times, that is $t = O(2^{-j/2})$, it turns out that K_k "looks enough like a pseudodifferential operator" and has constant ϵ-separation rank. When t gets larger than $2^{-j/2}$, this property quickly degrades, however. In order to justify these claims, consider the changes of variables

$$x' = \frac{x - x_k}{\rho_k}, \quad \xi' = \frac{\xi - \xi_\mu}{2^j}, \quad \eta' = \frac{\eta - \eta_{\mu'}}{2^j}. \tag{43}$$

Translations and dilations do not affect separation ranks. Their effect is to normalize the kernel so that the integral in x' is in a region of size at most $O(1)$ in x, and the range for ξ' and η' is a ball centered at the origin, with $O(1)$ radius. The new amplitude

$$b_k(x', \eta', t) = q_k(x(x')) a_k(x(x'), \eta(\eta'), t) \chi_{\mu'}(\eta(\eta'))$$

is a C^∞ function whose smoothness constants do not depend on j or j' anymore, because in the new variables, the symbol conditions (34) read

$$|\partial_{\eta'}^\alpha \partial_{x'}^\beta b_k(x', \eta', t)| \le C_{\alpha,\beta} 2^{j|\alpha|} (1 + |\eta_{\mu'} + 2^j \eta'|)^{-|\alpha|/2} \le C_{\alpha,\beta}.$$

(We have used $|\eta_{\mu'}| \asymp 2^{2j}$.) As for the phase, we have $A(t) = I + tP(t)$ by Lemma 3, with $P(t) = O(1)$ componentwise. Therefore,

$$x \cdot (A(t)\eta - \xi) = \rho_k 2^j x' \cdot (\eta' - \xi') + t\rho_k 2^j x' \cdot (P(t)\eta' - \xi') + \text{OK}. \tag{44}$$

The term "OK" refers to quantities that depend either on x', or on (η', ξ') – but not on all three at the same time, hence absorbable in the amplitude.

Let us now distinguish three subcases, depending on how t asymptotically compares to $2^{-j/2}$. Recall that $C2^{-j} \le \rho_k \le C2^{-j/2}$.

5.2.1 Typical times, $2^{-j} \lesssim t \lesssim 2^{-j/2}$

If $t \lesssim 2^{-j/2}$, then $t\rho_k 2^j \le C$ and hence the second term in (44) is non-oscillatory and can be absorbed in the amplitude b in a now standard manner. What remains is

$$\int e^{i\rho_k 2^j x' \cdot (\eta' - \xi')} b_k(x', \eta', t) dx' = (2\pi)^2 \hat{b}_k(\rho_k 2^j (\eta' - \xi'), \eta')$$

and can be seen to have ϵ-separation rank $O(\epsilon^{-1/M})$, by applying Lemma 2 to the properly supported C^∞ function $\hat{b}_k$ (the diagonal scaling by $\rho_k 2^j$ is harmless.) The overall separation rank is proportional to the number of balls used to cover the set $X_\mu^{\mu'}(\delta)$, hence of order $O(2^j)$, as claimed in Theorem 2.

Note that Lemma 2 should actually be invoked with an adequate fraction of ϵ, to make sure that

$$\|R_2\|_{L^2 \to L^2} \leq \frac{\epsilon}{4}. \tag{45}$$

In the appendix we settle an inconspicuous complication arising in the justification of (45), having to do with the fact that the separation remainder is actually a sum over $O(2^j)$ contributions, as in Eq. (32).

An application of Lemma 4 now shows that each wave atom submatrix formed from R_2, with indices $\mathbf{n}, \mathbf{n}'$, has $\ell^2_{\mathbf{n}'} \to \ell^2_{\mathbf{n}}$ norm at most $\epsilon/4$.

5.2.2 Large times, $t \gtrsim 2^{-j/2}$

If asymptotically $t \geq 2^{-j/2}$ then $t\rho_k 2^j$ grows in j and a different definition of $q_k(x)$ is necessary. More precisely, we repeat the covering argument of $X_\mu^{\mu'}(\delta)$ with a smaller local ball radius density, given by

$$\rho(x, j) = \frac{1}{\sqrt{t^2 2^{2j} + 2^{2j}|\nabla g(x)|}}.$$

All ball radii now obey $\rho_k \leq \frac{1}{t2^j}$, hence the phase becomes $\rho_k 2^j x' \cdot (\eta' - \xi')$+non-oscillatory, as required. By repeating the previous counting argument, their total number is $O(t^2 2^{2j})$. The rest of the argument is otherwise identical.

The conclusion is the same as before: the overall separation rank is proportional to the number of balls used in the main partitioning argument, here $O(t^2 2^{2j})$. The justification that R_2 gives rise to submatrices of norm $\epsilon/4$ is the same as before.

5.2.3 Small times, $t \lesssim 2^{-j}$

For small times, the same argument would apply, but a major simplification of the problem's geometry allows us to prove a stronger result. By Lemma 3,

$$\nabla_x \Phi(x, \eta, t) = \eta + O(t|\eta|) = \eta + O(t2^{2j}).$$

For $t \leq C \cdot 2^{-j}$ there exists a value of δ for which the set $X_\eta^\eta(\delta)$ defined by the condition $|\nabla_x \Phi(x, \eta, t) - \eta| \leq 2^j \delta$ covers $[0, 1]^2$. So will $X_\mu^{\mu'}(\delta)$, which is bigger than $X_\eta^\eta(\delta)$. The neighborhood of the locus of stationary phase is, therefore, the whole unit square.

We follow the same reasoning as before, and try to find a covering of $[0, 1]^2$ with balls of radius ρ_k in which the second-order term in the x-expansion of the phase is

non-oscillatory. For $t = O(2^{-j})$ it suffices to take $\rho_k = \rho_0$, identically equal to

$$\rho_0 = \frac{1}{\sqrt{t}2^j}.$$

Indeed, by Lemma 3,

$$\rho_0^2 |\nabla_x \nabla_x \Phi(t, \eta, t)| \leq C \cdot \rho_0^2 t|\eta| \leq C.$$

The collections of ball centers x_k can be taken as the Cartesian grid

$$x_k = (k_1, k_2)\frac{1}{2}t^{-1/2}2^{-j}, \quad k_1, k_2 \in \mathbb{Z}.$$

This corresponds to $O(1+t2^{2j})$ balls $B_{x_k}(\rho_0)$. The exact same reasoning as in the more general case applies, and yields an overall $\epsilon/4$-separation rank of order $O(1 + t2^{2j})$.

5.3 Fine scales, nonstationary phase

Let us now show that the nonstationary phase part yields a negligible contribution. Recall that we have defined, for each $\delta > 0$,

$$X_\mu^{\mu'}(\delta) = \{x \in [0, 1]^2; \text{ there exist } \xi \in \text{supp } \chi_\mu \text{ and } \eta \in \text{supp } \chi_{\mu'}, \ |\xi - \nabla_x \Phi(x, \eta, t)| \leq 2^j \delta\}. \tag{46}$$

The partition of unity $\{q_k(x)\}$ introduced in the previous section can be used as smooth indicators for the complement of $X_\mu^{\mu'}(\delta)$. Let S_{out} be the set $\{x_k : B_{x_k}(\rho_k) \cap X_\mu^{\mu'}(\delta) \neq \emptyset\}$, and

$$p(x) = \sum_{x_k \in S_{\text{out}}} q_k(x).$$

Of course, $p(x)$ depends on j, j' and δ but keeping track of this fact would make the notations unnecessarily heavy. It follows from the definition of $q_k(x)$ that we have the "maximal" smoothness condition

$$\sup_{x \in [0,1]^2} |\partial_x^\alpha p(x)| \leq C_\alpha \cdot 2^{j|\alpha|}.$$

We can now readily estimate

$$R_3(\xi, \eta, t) = \int e^{i(\Phi(x,\eta,t) - x\cdot\xi)}(1 - p(x))a(x, \eta, t)\, dx.$$

Indeed, we claim that an adequate choice of δ implies $\|R_3\|_2 \leq \epsilon/4$ in L^2. To this end, let us first check L^2 boundedness. The smoothness property of $p(x)$, along with the estimate $2^{2j} \sim |\xi|$, imply that the amplitude

$$\sigma(x, \eta, t) = (1 - p(x))a(x, \eta, t)$$

is a symbol of order zero and type $(1, 1/2)$, in the sense that

$$|\partial_\xi^\alpha \partial_x^\beta \sigma(x, \xi, t)| \leq C_{\alpha,\beta}(1 + |\xi|)^{-|\alpha|+|\beta|/2}.$$

As mentioned earlier, it is a beautiful application of the wave atom sparsity Theorem that Fourier integral operators of type $(1/2, 1/2)$, and in particular the kernel $e^{i\Phi}\sigma$ with σ as above, are bounded on L^2.

Let us now show that the L^2 bound can be made arbitrarily small, by an adequate choice of δ. Consider the differential operator

$$L = \frac{1}{|\xi - \nabla_x \Phi(x, \eta, t)|^2} (\Delta_x - i\Delta_x \Phi(x, \eta, t)I),$$

which is chosen so that $Le^{i(\Phi(x,\eta,t)-x\cdot\xi)} = e^{i(\Phi(x,\eta,t)-x\cdot\xi)}$. The operator L can be applied any number of times to the exponential factor, and then moved to $\sigma = (1-p)a$ by integration by parts. The effect on the amplitude σ is the following:

- Every $\frac{1}{|\xi-\nabla_x\Phi(x,\eta,t)|^2}$, on the support of $(1 - p(x))\chi_{\mu'}(\eta)$, brings in a factor $\frac{1}{\delta^2 2^{2j}}$, thanks to the definition of the set $X_\mu^{\mu'}(\delta)$.
- Every $L(1-p(x))$ yields a factor 2^{2j}, because of the smoothness property of $p(x)$.
- Every $\Delta_x\Phi$ yields a factor 2^{2j}, by homogeneity.
- After integration by parts, the new amplitude obeys the same smoothness assumptions as σ, hence is still a symbol of type $(1, 1/2)$.

Therefore, we conclude that

$$\delta^{2M} L^M \sigma(x, \eta, t)$$

is of type $(1, 1/2)$, with smoothness constants depending on M, but independent of δ. Invoking the general theory of FIOs, the L^2 bound on R_3 is therefore of the form

$$\|R_3\|_2 \leq C_M \delta^{-2M} \tag{47}$$

For fixed M, this bound can be made less than $\frac{\epsilon}{8\pi}$ by choosing δ as

$$\delta \geq C'_{M'} \epsilon^{-\frac{1}{M'}}, \tag{48}$$

with $M' = 2M$ and for some constant $C'_{M'}$ related to C_M. The combination of this result and Lemma 4 translates into a boundedness result for the corresponding submatrix in $\mathbf{n}$, $\mathbf{n}'$, namely that its ℓ^2 norm is bounded by $\epsilon/4$.

The proof is now complete, because the remainders R_1, R_2 and R_3 are of size at most $\epsilon/2$, $\epsilon/4$ and $\epsilon/4$ respectively, hence add up to ϵ. □

6 Scattering estimates

The objective of this section is to quantify the interactions, or energy transfer from an input wave vector $\xi_{\mu'}$ to other output wave vectors ξ_μ. As a result, we will obtain estimates on the *sum* of ranks of submatrices, either on $j, \mathbf{m}, \nu$ or $j', \mathbf{m}', \nu'$.

Theorem 3 *Let $E_{j\mathbf{m}\nu;\, j'\mathbf{m}'\nu'}(t)$ be the submatrix corresponding to $j, \mathbf{m}, \nu$ and $j', \mathbf{m}', \nu'$ in the separated wave atom representation of $E(t)$. For any $\epsilon > 0$, given $(j', \mathbf{m}')$, let $\Omega_{j',\mathbf{m}'}(t)$ be the smallest set of wave vectors $(j, \mathbf{m})$ such that setting $E(t)_{j\mathbf{m}\nu;\, j'\mathbf{m}'\nu'} = 0$ for $(j, \mathbf{m}) \notin \Omega_{j',\mathbf{m}'}(t)$ and all ν, ν' results in an error less than ϵ in matrix ℓ_2 norm. Then the cardinality of $\Omega_{j',\mathbf{m}'}(t)$ obeys the bound*

$$|\Omega_{j',\mathbf{m}'}(t)| \leq C_\epsilon \cdot (1 + t^2 2^{2j'}),$$

where $C_\epsilon \leq C_M \epsilon^{-1/M}$, for all $M > 0$.

Proof Fix $\epsilon > 0$ and $M > 0$. For this proof, we will exploit the compression properties of the wave propagator as in Theorem 1. The wave atom representation $\tilde{E}_{B,N}(t)$ of the propagator $E(t)$ is constructed as a matrix with two shifted band diagonals indexed by $\nu = \pm$, each of them corresponding to a ball in phase space centered about $h_{t,\nu}(\mu')$, and defined through the wave atom metric ω. More precisely, the "shifted band diagonals" are defined as the following set of wave atom subscripts:

$$SBD(\mu') = \bigcup_{\nu=\pm} \{\mu : \omega(\mu, h_{t,\nu}(\mu')) \leq r\},$$

with r chosen such that $|\{\mu : \omega(\mu, h_{t,\nu}(\mu')) \leq r\}| \asymp B$. Take B large enough so that the right-hand side of the error estimate (2) obeys $C_M B^{-M} \leq \epsilon$. Then of course $r \leq C_M \epsilon^{-1/M}$. Note that an error ϵ in L^2 for operators translates into an error ϵ in ℓ^2 for the wave atom matrix, by the tight frame property. In turn, restriction to a certain subset of rows and columns implies an error smaller than ϵ in ℓ^2 for the submatrix corresponding to $j, \mathbf{m}, \nu$ and $j', \mathbf{m}', \nu'$.

To estimate the size of $\Omega_{j',\mathbf{m}'}(t)$ it suffices to count the number of wave vectors $(j, \mathbf{m})$ which are part of at least one element $\mu = (j, \mathbf{m}, \mathbf{n})$ of the union of the shifted band diagonals over all $\mathbf{n}'$,

$$\bigcup_{\mathbf{n}':\mu'=(j',\mathbf{m}',\mathbf{n}')} SBD(\mu').$$

To this effect, recall that the local wave vector $\xi(t) = \nabla_x \Phi(t, x)$ is obtained from the solution of the Hamilton–Jacobi equation $\Phi_t = c(x)|\nabla_x \Phi|$ with initial condition $\Phi(0, x) = x \cdot \xi_{\mu'}$. The range of all such wave vectors defines a region in the frequency plane, which can be inscribed in a ball Q_0 centered at $\xi_{\mu'}$ and of radius majorized by

$C \cdot t|\xi_{\mu'}| \le C \cdot t 2^{2j'}$. The set of wave vectors ξ_μ defined through $SBD(\mu')$ is slightly larger however, because the radius r is nonzero, but can certainly be inscribed in a larger ball Q_r of radius bounded by $C \cdot t 2^{2j'} + C_\epsilon 2^{j'}$.

It remains to count the number of tiling indicators $\chi_\mu(\xi)$ whose supports intersect the ball Q_r. Near $\xi_{\mu'}$, the support of each indicator has radius $O(2^{j'})$, so it suffices to use a number of indicators bounded by

$$C \cdot \left(\frac{C \cdot t 2^{2j'} + C_\epsilon \cdot 2^{j'}}{2^{j'}} \right)^2 \le C_\epsilon \cdot (1 + t^2 2^{2j'}).$$

This is the desired bound on the cardinality of $\Omega_{j',\mathbf{m}'}(t)$. □

A simple counting argument now allows us to formulate the following result, companion to Theorem 2. The collection of bounds is summarized in Fig. 12.

Corollary 2 *Consider the submatrix $E_{j\mathbf{m}\nu;j'\mathbf{m}'\nu'}(t)$ obtained by fixing $(j,\mathbf{m})$ and $(j',\mathbf{m}')$ in the wave atom representation of the propagator $E(t)$ after reordering $(n_1,n_2;n_1',n_2') \to (n_1,n_1';n_2,n_2')$. Denote by $r_{j\mathbf{m}}^{j'\mathbf{m}'}$ the maximum over ν,ν' of the ϵ-rank of $E_{j\mathbf{m}\nu;j'\mathbf{m}'\nu'}(t)$. Then we have the bounds*

- *for $t \lesssim 2^{-j}$, $\sum_{j\mathbf{m}} r_{j\mathbf{m}}^{j'\mathbf{m}'} \le C_\epsilon \cdot (1 + t 2^{2j})$,*
- *for $2^{-j} \lesssim t \lesssim 2^{-j/2}$, $\sum_{j\mathbf{m}} r_{j\mathbf{m}}^{j'\mathbf{m}'} \le C_\epsilon \cdot t^2 2^{3j}$,*
- *for $2^{-j/2} \lesssim t \le T$, $\sum_{j\mathbf{m}} r_{j\mathbf{m}}^{j'\mathbf{m}'} \le C_\epsilon \cdot 2^{2j}$,*

with $C_\epsilon \le C_M \epsilon^{-1/M}$, for all $M > 0$, and C_ϵ also depends on T. The same bounds are valid for $\sum_{j'\mathbf{m}'} r_{j\mathbf{m}}^{j'\mathbf{m}'}$.

Proof For $t \le 2^{-j/2}$, or a constant multiple thereof, we can combine Theorem 2 with the scattering estimate (3) to obtain the first two bounds. For $t \ge 2^{-j/2}$, it suffices to notice that the rank of each submatrix $E_{j\mathbf{m}\nu;j'\mathbf{m}'\nu'}(t)$ must be smaller than the number of nonzero elements. After thresholding at level ϵ in ℓ^2, the number of nonzero elements in any of the matrices $E_{j\mathbf{m}\nu;j'\mathbf{m}'\nu'}(t)$, for fixed $\mathbf{m}'$, is bounded by $C_\epsilon \cdot 2^{2j}$, by sparsity (Theorem 1). The third bound follows.

The same bounds on $\sum_{j'\mathbf{m}'} r_{j\mathbf{m}}^{j'\mathbf{m}'}$ stem from the observation that the adjoint operator $E^*(t)$ is obtained from the backward-in-time wave equation, which admits the same sparsity and separation properties. Note that formulating bounds in terms of j or j' does not make any difference since $j \asymp j'$ by sparsity. □

7 Special cases

In this section we continue the study of three of the four representative sample media introduced in Sect. 4, as well as another medium called "misaligned wave guide," this time in the light of the rank estimates just obtained. In two cases (Wave Guide and Bumps) the rank and complexity estimates turn out to be quite pessimistic and we are

able to prove better bounds under certain conditions. In the two other less favorable cases (Misaligned Wave Guide and Linear Mirror), we give heuristic arguments that the rank bounds of Sects. 5 and 6 are in fact attained.

7.1 Wave guides

We refer to a wave guide as an acoustic medium whose speed of sound depends only on one coordinate, either x_1 or x_2. As always, it is also assumed to be C^∞.

The rank bounds can be significantly improved for wave guides. In short, we show that rank majorants for wave guides are in general the *square root* of the rank majorants in the worst case.

Theorem 4 *Assume the velocity profile depends only on x_2 and is C^∞. Consider the submatrix $E_{j\mathbf{m}\nu;\, j'\mathbf{m}'\nu'}(t)$ obtained by fixing $(j, \mathbf{m})$ and $(j', \mathbf{m}')$ in the wave atom representation of the propagator $E(t)$, after reordering $(n_1, n_2; n_1', n_2') \to (n_1, n_1'; n_2, n_2')$. Then the ϵ-rank of $E_{j\mathbf{m}\nu;\, j'\mathbf{m}'\nu'}(t)$ obeys*

- *for $t \lesssim 2^{-j}$, $r \le C_\epsilon \cdot (1 + \sqrt{t} 2^j)$,*
- *for $2^{-j} \lesssim t \lesssim 2^{-j/2}$, $r \le C_\epsilon \cdot 2^{j/2}$,*
- *for $2^{-j/2} \lesssim t \le T$, $r \le C_\epsilon \cdot t 2^j$,*

with $C_\epsilon \le C_M \epsilon^{-1/M}$, for all $M > 0$, and C_ϵ also depends on T.

Proof When the velocity profile $c(x)$ does not depend on x_1, it is easy to check that the local wave numbers $\nabla_x \Phi_\pm(x, \xi, t)$ do not depend on x_1 either (although $\Phi_\pm$ itself does). The steps of the proof are then the same as for Theorem 2, except that the definition of indicators $q_k(x)$ is a bit different. Instead of considering balls $B_{x_k}(\rho_k)$, we will consider horizontal *strips* $S_{x_{2,k}}(\rho_k)$ centered at height $x_2 = x_{2,k}$ and of width $2\rho_k$. Equations (35) through (38) then carry through unchanged, but a major simplification occurs in the counting argument for N_S, the number of strips necessary to make the restrictions of the phase non-oscillatory on each $q_k(x)$. The problem is now one-dimensional, g depends on x_2 only, so the local "strip width density" can be defined as

$$\rho(x_2, j) = \frac{1}{\sqrt{2^j + 2^{2j} |g'(x_2)|}}, \tag{49}$$

and the lattice y_k can be replaced by a simpler one-dimensional sequence $y_{2,k} = kb2^{-j}$. In contrast with Eq. (41), the weights w_k assigned to $y_{2,k}$ must now be defined as

$$w_k = 2^{-j} \sum_{x_{2,\ell} \in X : y_{2,k} \in S_{x_{2,\ell}}(\rho_\ell)} \frac{1}{\rho_\ell}.$$

There are $O(2^j \rho_\ell)$ points $y_{2,k}$ inside the interval $[x_{2,\ell} - \rho_\ell, x_{2,\ell} + \rho_\ell]$, so we have

$$N_S \le C \cdot \sum_k w_k.$$

The corresponding integral is

$$N_S \le C \cdot \int_X \frac{1}{\rho(x_2, j)}\, dx_2.$$

We should now use (49) in combination with the bound $\sqrt{2^j + 2^{2j}|g'(x_2)|} \le 2^{j/2}(1 + \frac{1}{2}2^j|g'(x_2)|)$ and Lemma 6—also valid in dimension one—to obtain the improved bound

$$N_S \le C \cdot 2^{j/2}.$$

The rest of the proof then proceeds in an analogous way.

- For typical times $2^{-j} \lesssim t \lesssim 2^{-j/2}$, the bound on r is the same as that for N_S, namely $O(2^{j/2})$.
- For small times $t \lesssim 2^{-j}$, the strip heights can be taken equispaced and equal to

$$x_{2,k} = kbt^{-1/2}2^{-j},$$

  yielding $N_S = O(1 + \sqrt{t}2^j)$ strips and a comparable rank r.
- For large times $t \gtrsim 2^{-j/2}$,

$$\rho(x_2, j) = \frac{1}{\sqrt{t^2 2^{2j} + 2^{2j}|g'(x_2)|}},$$

  so $r \simeq N_S = O(t2^j)$ by the previous argument. □

These rank bounds are summarized in Fig. 12. Let us remark at this point that the rank plateaus at a value $O(2^j)$ for $t \simeq 1$, although the size of the matrix is $\simeq 2^{2j}$-by-2^{2j}. This is obviously a consequence of the above theorem for times before caustics, but it turns out the same result is also valid *after* caustics start forming. The justification of this more general claim will follow from the analysis of the stronger bound on the sum of ranks over j' and $\mathbf{m}'$, which we now present.

Theorem 5 *Assume the velocity profile depends only on x_2 and is C^∞. Consider the submatrix $E_{j\mathbf{m}\nu;\, j'\mathbf{m}'\nu'}(t)$ obtained by fixing $(j, \mathbf{m})$ and $(j', \mathbf{m}')$ in the wave atom representation of the propagator $E(t)$, after reordering $(n_1, n_2; n_1', n_2') \to (n_1, n_1'; n_2, n_2')$. Denote by $r_{j\mathbf{m}}^{j'\mathbf{m}'}$ the maximum over ν, ν' of the ϵ-rank of $E_{j\mathbf{m}\nu;\, j'\mathbf{m}'\nu'}(t)$. Then we have the bounds*

- *for $t \lesssim 2^{-j}$, $\sum_{j\mathbf{m}} r_{j\mathbf{m}}^{j'\mathbf{m}'} \le C_\epsilon \cdot (1 + \sqrt{t}2^j)$,*
- *for $2^{-j} \lesssim t \lesssim 2^{-j/2}$, $\sum_{j\mathbf{m}} r_{j\mathbf{m}}^{j'\mathbf{m}'} \le C_\epsilon \cdot t2^{3j/2}$,*
- *for $t \gtrsim 2^{-j/2} \le T$, $\sum_{j\mathbf{m}} r_{j\mathbf{m}}^{j'\mathbf{m}'} \le C_\epsilon \cdot 2^j$,*

with $C_\epsilon \le C_M \epsilon^{-1/M}$, for all $M > 0$, and C_ϵ also depends on T. The same bounds are valid for $\sum_{j'\mathbf{m}'} r_{j\mathbf{m}}^{j'\mathbf{m}'}$.

Proof We need a stronger version of the scattering estimate in Theorem 3, in the special case of wave guides. The question is to determine the number of balls of radius $\simeq 2^j$ (each containing a wave atom bump in frequency) necessary to cover the locus of local wave vectors $\nabla_x \Phi_\pm(x,\xi,t)$, when a union is taken over all possible values of x. We know from the general case that this local wave vector cannot wander too far off ξ, namely $|\xi - \nabla_x \Phi_\pm(x,\xi,t)| \le C \cdot t 2^{2j}$, resulting in a covering by at most $O(1+t^2 2^{2j})$ balls.

In the case of wave guides, however, this locus is for each phase a *one-dimensional smooth curve* Γ_ξ, generated by the union of all wave vectors over the single coordinate x_2 (because the local wave vector is independent of x_1). In addition, Γ_ξ inherits the homogeneity of degree one of Φ, which makes it homothetic in $|\xi|$. As a result, the length of Γ_ξ is in fact comparable to the diameter of the locus in the general case, $O(t2^{2j})$, so it only takes $O(t2^j)$ balls of radius $\simeq 2^j$ to cover Γ_ξ. As a result, the cardinality of the set of participating wave vectors, in analogy with Theorem 3, is

$$|\Omega_{j',\mathbf{m}'}(t)| \le C_\epsilon \cdot (1 + t2^j).$$

The argument bounding sums of ranks over j' and $\mathbf{m}'$ then goes on to follow from the proof of Corollary 2, and we obtain

- for $t \lesssim 2^{-j}$, $\sum_{j\mathbf{m}} r_{j\mathbf{m}}^{j'\mathbf{m}'} \le C_\epsilon \cdot (1 + \sqrt{t} 2^j)$,
- for $2^{-j} \lesssim t \lesssim T$, before caustics, $\sum_{j\mathbf{m}} r_{j\mathbf{m}}^{j'\mathbf{m}'} \le C_\epsilon \cdot t 2^{3j/2}$.

Since we have so far relied on the existence of the phase functions $\Phi_\pm$ in our reasoning, we took the precaution of mentioning that the result is valid before the formation of caustics (on plane wave initial conditions). The same bounds also hold when the sum is taken over $(j',\mathbf{m}')$ instead of $(j,\mathbf{m})$, for the same reasons as previously.

We however claim that a stronger estimate holds: $\sum_{j',\mathbf{m}'} r_{j,\mathbf{m}}^{j',\mathbf{m}'} \le C_\epsilon \cdot 2^j$, regardless of t, even after caustics. This improves on the earlier bound when $t \gtrsim 2^{-j/2}$. In order to justify this claim, we need to understand the effect of the wave guide structure on the submatrices of interest, $E_{j\mathbf{m}\nu;j'\mathbf{m}'\nu'}(t)$ with row index $\mathbf{n} = (n_1, n_2)$ and column index $\mathbf{n}' = (n_1', n_2')$. For short, when the other parameters are encumbering, we also denote the submatrix by $E_{n_1,n_2}^{n_1';n_2'}$.

The subscript ν takes on two values ($\pm$) so we omit it in what follows. Recall the central sparsity result, Theorem 1, which states that for fixed $\mu = (j,\mathbf{m},\mathbf{n})$ the number of matrix elements above a threshold ϵ (in absolute value), spanned by the remaining indices $(j',\mathbf{m}',\mathbf{n}')$, is a constant $C_\epsilon = O(\epsilon^{-1/M})$ for all $M > 0$. Let us now make the exercise of only fixing $(j,\mathbf{m},n_1)$: the number of elements above ϵ spanned by the other indices $(j',\mathbf{m}',\mathbf{n}',n_2)$ is proportional to the number of n_2's, that is $C_\epsilon \cdot 2^j$. Fixing n_1 means considering only a subset of the rows, i.e. "mutilating" each submatrix $E_{n_1,n_2}^{n_1',n_2'}$. Surely, for fixed n_1 the sum of ranks of those mutilated matrices over $j',\mathbf{m}'$ cannot exceed the total number of elements, $C_\epsilon \cdot 2^j$. Re-ordering the submatrices as $E_{n_1,n_1'}^{n_2,n_2'}$ does not change that fact.

As we now consider different values of n_1 (still for fixed $j,\mathbf{m}$), we introduce *no new information*. Because of the invariance of the problem under translations in x_1,

we obtain the same wave atom matrix elements, albeit shifted circularly in n_1'. More precisely, the invariance property reads

$$E_{n_1,n_2}^{n_1',n_2'} = E_{n_1+p,n_2}^{n_1'+p,n_2'},$$

where p is any integer and addition is understood modulo the bound on the number of n_1. Consequently, the rank of $E_{n_1,n_2}^{n_1',n_2'}$ does depend on whether it is mutilated to a certain subset of n_1's or not. The same is true for the sum of ranks over $(j', \mathbf{m}')$, so the claim follows.

Again, the same bounds also hold when the sum is taken over $(j, \mathbf{m})$ instead of $(j', \mathbf{m}')$. The proof is complete. □

7.2 Bumps

The example "Bumps" belongs to a larger class of nondegenerate oscillating profiles, which can be formalized as follows.

Definition 3 (Transversality) A smooth velocity profile $c(x) > 0$ is said to be *transversal* when the following two conditions are satisfied:

1. The locus where the Hessian $\nabla\nabla c$ is singular is the union of a finite number of smooth curves.
2. For every point x for which there exists two unit vectors d, d' such that $(d \cdot \nabla)^2 c(x) = 0$ and $(d' \cdot \nabla)^3 c(x) = 0$, we have $d \cdot d' \neq 0$.

As can easily be checked, examples of transversal profiles include smooth and separable functions $c(x_1, x_2) = \gamma_1(x_1)\gamma_2(x_2) > 0$ with γ_k''' nonzero when γ_k'' vanishes, $k = 1$ or 2. In the "Bump" example, we have taken $\gamma_1(x) = \gamma_2(x) = \frac{3+\sin(4\pi x)}{16}$. We also expect a sum of wide bumps with random location and random positive amplitude to satisfy the transversality condition with high probability.

A notable example of non-transversal profile, on the other hand, would be the innocent-looking

$$c(x_1, x_2) = 2 + \sin(2\pi x_1)\sin(2\pi x_2),$$

for which condition 2 in Definition 3 is violated.

The rationale for introducing "transversal" profiles is the following (obvious) asymptotic relation for the phase Hessian,

$$\nabla_x \nabla_x \Phi_\pm(x, \xi, t) = \pm t \nabla\nabla c(x)\,|\xi| + O(t^2|\xi|).$$

For small times $t = o(1)$, the locus of singularity of $\nabla_x \nabla_x \Phi_\pm$ is a deformation of that of $\nabla\nabla c$. Such information allows to characterize the locus $X_\xi^\eta(\delta)$ of stationary phase in a much more precise way than was done in the proof of Theorem 2. As a result, the rank estimates can be strengthened as follows. The results are reported in Fig. 12.

Theorem 6 *Assume $c(x)$ is smooth and transversal, in the sense of Definition 3. Consider the submatrix $E_{j\mathbf{m}\nu;\,j'\mathbf{m}'\nu'}(t)$ obtained by fixing $(j,\mathbf{m})$ and $(j',\mathbf{m}')$ in the wave atom representation of the propagator $E(t)$, after reordering $(n_1,n_2;n_1',n_2') \to (n_1,n_1';n_2,n_2')$. Then the ϵ-rank of $E_{j\mathbf{m}\nu;\,j'\mathbf{m}'\nu'}(t)$ obeys*

- *for $t \lesssim 2^{-j}$, $r \le C_\epsilon \cdot (1+t2^{2j})$,*
- *for $2^{-j} \lesssim t \lesssim 2^{-j/3}$, $r \le C_\epsilon \cdot \frac{2^{j/2}}{\sqrt{t}}$,*
- *for $2^{-j/3} \lesssim t \le T = o(1)$, $r \le C_\epsilon \cdot t2^j$,*

with $C_\epsilon \le C_M \epsilon^{-1/M}$, for all $M > 0$.

Proof As alluded to earlier, the condition $T = o(1)$ ensures that the phases $\Phi_\pm$ satisfy the same transversality conditions as $c(x)$.

The proof of the rank bound for $t \lesssim 2^{-j}$ is the same as previously, so let us consider $t \gtrsim 2^{-j}$. As alluded to earlier, the condition $T = o(1)$ ensures that the phases $\Phi_\pm$ satify the same transversality conditions as $c(x)$. The purpose of the transversality condition is to allow a much more explicit description of the loci $X_\mu^{\mu'}(\delta)$ of stationary phase than in the proof of Theorem 2.

Consider one phase function, say $\Phi = \Phi_+$. Given a wave number η and a point $x^* \in [0,1]^2$, only three scenarios can occur.

1. Assume $\nabla_x \nabla_x \Phi(x^*,\eta,t) = 0$. By transversality, we necesarily have

$$|(d\cdot\nabla)^3 \Phi(x^*,\xi,\eta)| \ge C_{\text{trans}} t|\eta|, \tag{50}$$

uniformly over all unit vectors d. Let $\xi_0 = \nabla_x \Phi(x^*,\eta,t)$. We would like to find good bounds for the set

$$X_{\xi_0}^{\eta}(\delta) = \{x \in [0,1]^2 : |\nabla_x \Phi(x,\eta,t) - \xi_0| \le \delta 2^j\}.$$

Once this is done, we can identify the wave atom subscripts μ, μ' such that ξ_μ is closest to ξ_0, η_μ is closest to η_0, and assert that $X_\mu^{\mu'}(\delta)$ has about the same size, up to a constant, as $X_{\xi_0}^{\eta}(\delta)$. See the reasoning leading to Eq. (33).
Using a Taylor expansion around x^* and Lemma 3 we first obtain

$$\begin{aligned}\nabla_x \Phi(x,\eta,t) = \xi_0 + \frac{1}{2}\sum_{k_1,k_2}(x-x^*)_{k_1}(x-x^*)_{k_2}\frac{\partial^2}{\partial x_{k_1}\partial x_{k_2}}\nabla_x\Phi(x^*,\eta,t)\\ + O(|x-x^*|^3 t|\eta|).\end{aligned} \tag{51}$$

We can take the dot product of this relation with $d(x) = \frac{x-x^*}{|x-x^*|}$ to get

$$\begin{aligned}d(x)\cdot(\nabla_x\Phi(x,\eta,t)-\xi_0) = \frac{1}{2}|x-x^*|^2 (d(x)\cdot\nabla_x)^3\Phi(x^*,\eta,t)\\ + O(|x-x^*|^3 t|\eta|).\end{aligned}$$

The magnitude of a gradient is certainly greater than the absolute value of any directional derivative, so

$$|\nabla_x \Phi(x,\eta,t) - \xi_0| \geq \frac{1}{2} C_{\text{trans}} |x - x^*|^2 t|\eta| - O(|x - x^*|^3 t|\eta|).$$

When $|x - x^*| = o(1)$ as the scale j or equivalently $|\eta| \simeq 2^{2j}$ grows, then the $O(|x - x^*|^3 t|\eta|)$ remainder is asymptotically negligible and the behavior of Φ near x^* is governed by its third spatial derivatives. If we let $x \in X^{\eta}_{\xi_0}(\delta)$ then the condition defining the latter set implies

$$C \cdot t|\eta||x - x^*|^2 \leq 2^j \delta,$$

which in turn shows that $X^{\eta}_{\xi_0}(\delta)$ is included in a ball centered at x^*, with radius ρ_X proportional to $\frac{2^{-j/2}}{\sqrt{t}}$. As t asymptotically exceeds 2^{-j} we are indeed in the regime where $|x - x^*| = o(1)$, validating smallness of the Taylor remainder.
With this information on the extent of the set of near-stationary phase points, we are ready to repeat the ball counting argument of Sect. 5.2. The argument consists in exhibiting balls $B_{x_k}(\rho_k)$ over which the phase is non-oscillatory in the sense that for $x \in B_{x_k}(\rho_k)$, it holds that

$$\rho_k^2 |\nabla_x \nabla_x \Phi(x,\eta,t)| \leq C. \tag{52}$$

In the neighborhood of x^* the phase Hessian obeys, componentwise,

$$|\nabla_x \nabla_x \Phi(x,\eta,t)| \leq C \cdot t|\eta||x - x^*|,$$

which means that for $x \in X^{\eta}_{\xi_0}(\delta)$ we have

$$|\nabla_x \nabla_x \Phi(x,\eta,t)| \leq C \cdot 2^{3j/2}\sqrt{t}.$$

To satisfy the non-oscillation condition (52), it suffices to take the ball radii r_k uniformly equal to

$$\rho_k \simeq 2^{-3j/4} t^{-1/4}. \tag{53}$$

This choice corresponds to a covering of $X^{\mu'}_{\mu}(\delta)$ by N_B balls, where

$$N_B \leq C \cdot \left(\frac{\rho_X}{\rho_k}\right)^2 = C \cdot 2^{j/2} t^{-1/2}. \tag{54}$$

This bound on N_B will be interpreted later as a rank estimate, because the zero Hessian scenario turns out to be the worst case (largest bound on N_B.). To this end, we now intend to review and compare the other two scenarios.

2. Assume now that there exists a direction d along which

$$d \cdot \nabla_x \Phi(x^*, \eta, t) = 0 \quad \text{but} \quad d^\perp \cdot \nabla_x \Phi(x^*, \eta, t) \neq 0.$$

In the direction d, we can repeat the argument of scenario 1 to conclude that the spatial extent of $X_\mu^{\mu'}(\delta)$ is of order $\rho_X(d) = 2^{-j/2} t^{-1/2}$. In the direction $d^\perp$, the situation is simpler because the Taylor expansion of $\nabla_x \Phi$ is the usual

$$\nabla_x \Phi(x, \eta, t) = \xi_0 + \sum_k (x - x^*)_k \frac{\partial}{\partial x_k} \nabla_x \Phi(x^*, \eta, t) + O(|x - x^*|^2 t|\eta|).$$

Repeating the sequence of steps leading up to (53), we obtain instead

$$\rho_X(d^\perp) \simeq 2^{-j} t^{-1}.$$

It is straightforward to check that the phase is always non-oscillatory in the sense of (52) over balls of radius $\rho_k = 2^{-j} t^{-1}$. We conclude that $X_\mu^{\mu'}(\delta)$ can be covered by N_B balls $B_{x_k}(\rho_k)$, with

$$N_B \leq C \cdot \frac{\rho_X(d) \rho_X(d^\perp)}{\rho_k^2} = C \cdot 2^{j/2} \sqrt{t}. \tag{55}$$

For times $t = O(1)$ this bound is always smaller than (54), obtained in scenario 1.

3. Finally, assume that the phase Hessian is nonsingular. By the same argument as above, the set $X_\mu^{\mu'}(\delta)$ can be inscribed in a ball of radius $\rho_X \simeq 2^{-j} t^{-1}$, over which the phase is non-oscillatory, resulting in

$$N_B \leq C, \tag{56}$$

independently of j. This latter bound is always smaller than (54) for times $t \gtrsim 2^{-j}$.

The conclusion of the above analysis is that the worst-case scenario arises when the Hessian vanishes, for which $N_B \leq C \cdot 2^{j/2} t^{-1/2}$. Before translating this bound into a rank estimate, we must make sure that the off-diagonal linear term in the phase [see Eq. (44)] is itself non-oscillatory. Recall that the normalizing change of variables (43) for x' was chosen so that $x' = O(1)$ as long as $x \in B_{x_k}(\rho_k)$. In our case, we can choose it as

$$x' = \frac{x - x_k}{2^{-3j/4} t^{-1/4}},$$

resulting in

$$x \cdot (A(t)\eta - \xi) = (2^j t^{-1})^{1/4} x' \cdot (\eta' - \xi') + (2^j t^3)^{1/4} x' \cdot (P(t)\eta' - \xi'),$$

where $A(t) = I + tP(t)$ (compare with (44)). The term involving $P(t)$ is of order $O(1)$ as long as $t \lesssim 2^{-j/3}$, therefore allowing to view the bound on N_B as a rank estimate. That is the content of the second bullet in Theorem 6.

For times $t \gtrsim 2^{-j/3}$, we resort to the same reasoning as previously, namely modifying the change of variables as

$$x' = \frac{x - x_k}{(t2^j)^{-1}}.$$

This choice imposes a covering of $X_\mu^{\mu'}(\delta)$ by balls of radius $\rho_k = 2^{-j}t^{-1}$, resulting in

$$N_B \le C \cdot \left(\frac{2^{j/2}t^{-1/2}}{2^{-j}t^{-1}}\right)^2 = C \cdot t2^j.$$

The corresponding rank estimate follows (bullet 3 in Theorem 6.) This concludes the proof. □

The corresponding result for sums of ranks is the following.

Theorem 7 *Assume the velocity profile is tranversal and C^∞. Consider the submatrix $E_{j\mathbf{m}\nu;\, j'\mathbf{m}'\nu'}(t)$ obtained by fixing $(j, \mathbf{m})$ and $(j', \mathbf{m}')$ in the wave atom representation of the propagator $E(t)$, after reordering $(n_1, n_2; n_1', n_2') \to (n_1, n_1'; n_2, n_2')$. Denote by $r_{j\mathbf{m}}^{j'\mathbf{m}'}$ the maximum over ν, ν' of the ϵ-rank of $E_{j\mathbf{m}\nu;\, j'\mathbf{m}'\nu'}(t)$. Then we have the bounds*

- *for $t \lesssim 2^{-j}$, $\sum_{j\mathbf{m}} r_{j\mathbf{m}}^{j'\mathbf{m}'} \le C_\epsilon \cdot (1 + t2^{2j})$,*
- *for $2^{-j} \lesssim t \lesssim 2^{-3j/5}$, $\sum_{j\mathbf{m}} r_{j\mathbf{m}}^{j'\mathbf{m}'} \le C_\epsilon \cdot \frac{2^{j/2}}{\sqrt{t}}$,*
- *for $2^{-3j/5} \lesssim t \le T = o(1)$, $\sum_{j\mathbf{m}} r_{j\mathbf{m}}^{j'\mathbf{m}'} \le C_\epsilon \cdot t^2 2^{2j}$,*

with $C_\epsilon \le C_M \epsilon^{-1/M}$, for all $M > 0$. The same bounds are valid for $\sum_{j'\mathbf{m}'} r_{j\mathbf{m}}^{j'\mathbf{m}'}$.

Proof The justification is a combination of the bounds of Theorem 6 with a scattering estimate, counting the number of wave vectors $\xi_{\mu'}$ involved in each scenario on the phase Hessian (see the proof of Theorem 6.) Fix a wave vector ξ_μ. The count is as follows:

1. We claim that the locus where the Hessian $\nabla\nabla c$ is identically zero contains at most a finite number of points, in the case of transversal velocity profiles. Assume by contradiction that it is not the case. By compactness there exists a sequence of points x_i in $[0, 1]^2$ converging to some limit $x^* \in [0, 1]^2$, such that $x_i \ne x^*$ and $\nabla\nabla c(x_i) = 0$. Necessarily, by continuity, $\nabla\nabla c(x^*) = 0$. Denote $d_i = \frac{x_i - x^*}{|x_i - x^*|}$. Since the unit circle is compact, there exists a subsequence d_{i_j} converging to some $d \in S^1$. It is then a simple matter to check to check that $(d \cdot \nabla)\nabla\nabla c = 0$, contradicting the transversality condition in Definition 3.
   The same property transfers to the phase Hessian for times $t = o(1)$. Each point x where $\nabla_x \nabla_x \Phi_\pm(x, \xi_\mu, t)$ vanishes identically corresponds to one wave vector,

$\xi_0 = \nabla_x \Phi_\pm(x, \xi_\mu, t)$. As a consequence, there are at most a constant number of wave vectors $\xi_{\mu'}$ which belong in scenario 1, yielding a total combined rank

$$\sum_{(j,\mathbf{m})\in I} r_{j,\mathbf{m}}^{j',\mathbf{m}'} \le C \cdot \max\{2^{j/2} t^{-1/2}, 2^j t\}. \tag{57}$$

2. For scenario 2, we directly obtain from the transversality condition that the locus $\mathcal{L}$ where the phase Hessian is singular is a one-dimensional manifold. So is the locus Γ_{ξ_μ} of wave vectors $\xi = \nabla_x \Phi_\pm(x, \xi_\mu, t)$, where $x \in \mathcal{L}$. As in the proof of Theorem 5, the intersection of Γ_{ξ_μ} with the "scattering" ball $B_{\xi_\mu}(Ct2^{2j})$ can be covered by at most $O(t2^j)$ indicators $\chi_{\mu'}(\xi)$. As a result, the sum of ranks over $(j, \mathbf{m})$ for scenario 2 is

$$\sum_{(j,\mathbf{m})\in II} r_{j,\mathbf{m}}^{j',\mathbf{m}'} \le C \cdot 2^j t \cdot 2^{j/2} t^{1/2} = C \cdot 2^{3j/2} t^{3/2}. \tag{58}$$

3. Scenario 3 corresponds to all the wave vectors ξ_μ that are left out from scenarios 1 and 2. By Theorem 3, there are at most $O(t^2 2^{2j})$ of them. Each of those wave vectors corresponds to a submatrix with rank bounded by a $O(1)$ constant, so the total count is

$$\sum_{(j,\mathbf{m})\in III} r_{j,\mathbf{m}}^{j',\mathbf{m}'} \le C \cdot 2^{2j} t^2. \tag{59}$$

It now remains to add Eqs. (57)–(59). The two last bullets in Theorem 7 follow from the observation that (57) is asymptotically dominant when $t \lesssim 2^{-3j/5}$, but (59) dominates when $t \gtrsim 2^{-3j/5}$. □

7.3 Misaligned wave guide

A "misaligned wave guide" is an essentially one-dimensional profile $c(x)$ whose redundant coordinate is not aligned with x_1 or x_2. One such example is

$$c(x_1, x_2) = 2 - \cos(2\sqrt{2}\pi(x_1 - x_2)),$$

which depends only on $x_1 - x_2$. We take the precaution to name those profiles essentially one-dimensional, because they should also be smooth and periodic on the torus, a requirement incompatible with being a wave guide in other directions than vertical, horizontal, or diagonal at 45 degrees as above.

The performance of our solver on "misaligned wave guide" is rather poor so we chose not to report it in Sect. 4.

We intent to justify, albeit not in a rigorous manner, that misaligned wave guides probably saturate the rank bound $r \lesssim 2^j$ of Theorem 2, when $t \simeq 2^{-j}$. We hope that this example may help illustrate a central piece of the argument behind Theorem 2.

Locally near the diagonal $x_1 = x_2$, we have $c(x_1, x_2) \simeq 1 + 4\pi^2(x_1 - x_2)^2$. The phases $\Phi_\pm$ therefore obey the small-time (and small $|x_1 - x_2|$) asymptotic relations

$$\Phi_\pm(x, \xi, t) \simeq x \cdot \xi \pm t(1 + 4\pi^2(x_1 - x_2)^2)|\xi|.$$

Let us now explain why the most expensive contribution in the phase, in terms of the resulting ranks, is the off-diagonal term proportional to $t x_1 x_2 |\xi|$. We had already alluded to this fact in Sect. 5.2. We remind the reader that $|\xi| \simeq 2^{2j}$, so we will simply consider the phase $2^j x_1 x_2$.

In view of the proof of Theorem 2, we would like to bound the cardinality of a covering of the locus $X^\eta_\xi(\delta)$ of near-stationary phase by balls inside which the phase satisfies the stronger requirement of being non-oscillatory, see (36). For any given $\xi = \eta$ and large δ it is easy to see that the locus $X^\eta_\xi(\delta)$ actually covers the whole unit square. The phase Hessian is

$$\nabla_x \nabla_x \Phi(x_1, x_2) = 2^j \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

which implies a uniform ball radius $\rho_k \simeq 2^{-j/2}$. It takes $O(2^j)$ balls of radius ρ_k to cover the whole unit square, resulting in the announced bound $r \simeq 2^j$ for the rank.

7.4 Linear mirror

A "linear mirror" is a profile $c(x)$ which is locally of the form $C + x \cdot \lambda$ for some vector λ. Of course $x \cdot \lambda$ is not compatible with smoothness and periodicity on the torus; see Sect. 4 for a good compromise.

Linear mirrors are representative of a class of profiles for which the rank bound of Theorem 2 is expected to be sharp. Again, we will not provide a rigorous proof but only give indications towards this claim.

In the region where $c(x) = C + x \cdot \lambda$, the phases can be solved for explicitly,

$$\Phi_\pm(x, \eta, t) = x \cdot \eta \pm Ct|\xi| \pm x \cdot \frac{\lambda}{|\lambda|}|\eta|(e^{t|\lambda|} - 1).$$

In analogy with Eq. (44), this expression can be linearized in η and rewritten as $x \cdot A(t)\eta + \text{OK}$. In our case, the matrix elements of $A(t)$ are, for small time t, given as

$$A_{ij}(t) = \delta_{ij} + t\frac{\lambda_i}{|\lambda|}\frac{\eta_j}{|\eta|} + O(t^2).$$

In the notations of Sect. 5.2, we identify $P_{ij} = \frac{\lambda_i}{|\lambda|}\frac{\eta_j}{|\eta|}$. This is a prototypical non-diagonal matrix. This example leads us to believe that the linear part of the phase genuinely affects the rank estimates, and that we are not in presence of a proof artifact.

8 Discussion

So far we have assumed periodic boundary conditions for the wave equation inside the unit square $[0, 1]^2$, but simple modifications will allow the wave atom algorithm to work in slightly more general settings.

First, we can consider standard boundary conditions like Dirichlet ($u = 0$ on the boundary) or Neumann ($\frac{\partial u}{\partial n} = 0$) in the same domain $[0, 1]^2$. The two cases can be handled in a straightforward manner by mirror extension of the computational domain to the periodized square $[0, 2]^2$ with velocity

$$\tilde{c}(x_1, x_2) = \begin{cases} c(x_1, x_2) & \text{if } \ 0 \le x_1, x_2 < 1, \\ c(2 - x_1, x_2) & \text{if } \ 1 \le x_1 < 2, \ 0 \le x_2 < 1, \\ c(x_1, 2 - x_2) & \text{if } \ 0 \le x_1 < 1, \ 1 \le x_2 < 2, \\ c(2 - x_1, 2 - x_2) & \text{if } \ 1 \le x_1, x_2 < 2. \end{cases}$$

The ideal situation is when this extension does not create gradient discontinuities. The wave equation can then be solved up to some time T for $\tilde{u}$ in the periodized extended square $[0, 2]^2$, and $\tilde{u}$ mirror folded back onto $[0, 1]^2$ using the rule

$$u(x_1, x_2) = \tilde{u}(x_1, x_2) - \tilde{u}(2 - x_1, x_2) - \tilde{u}(x_1, 2 - x_2) + \tilde{u}(2 - x_1, 2 - x_2)$$

if u is to satisfy Dirichlet boundary conditions, or

$$u(x_1, x_2) = \tilde{u}(x_1, x_2) + \tilde{u}(2 - x_1, x_2) + \tilde{u}(x_1, 2 - x_2) + \tilde{u}(2 - x_1, 2 - x_2)$$

if u is to satisfy Neumann boundary conditions. Some other choice of signs are possible and would lead, for example, to Dirichlet on two opposite sides and Neumann on the two other sides. For the wave atom algorithm to perform accurately on the extended domain, we need to ensure sure that $\tilde{c}(x_1, x_2)$ remains sufficiently smooth after mirror extension as above.

The increase in complexity resulting from the doubling of N, the number of grid points per dimension, may however be unacceptable in some applications. Readers interested in a more elegant treatment of boundary considerations, in the context of some other basis of bandlimited functions (prolate spheroidal wavefunctions) should refer to the recent work of Beylkin and Sandberg [5].

More generally, if the computational domain can be mapped onto the unit square by means of a smooth diffeomorphism, then it is only a matter of changing variables and re-using the same algorithm on the transformed equation. More complicated geometries or topologies would pose a significant challenge to wave-packet-type methods and their treatment would go far beyond the scope of this paper.

Finally, wave atoms seem to be a promising tool for implementing absorbing boundary conditions in the regime of high-frequency solutions. Assume for a moment that the wavefield $u(t, x)$ has frequency support obeying $|\xi| \ge \lambda$, and that the profile $c(x)$ is near constant near the edges of the unit square. Then the computational domain can be extended to include a surrounding buffer strip of width $O(\frac{1}{\sqrt{\lambda}})$ and constant sound speed, in which outgoing wave atoms can be safely removed from the solution

by putting the corresponding matrix elements to zero. This should work provided the upscaled time step τ is of order $\tau = O(\frac{1}{\sqrt{N}})$.

Acknowledgments We would like to thank Emmanuel Candès for insightful discussions, advice and generous support. We are indebted to the referees for their thoughtful comments, which generated a much improved manuscript.

Appendix A: Additional proofs

Proof of inequality (40) In what follows the notation *sup* refers to the supremum taken over all $x \in B_{x_k}(r_k)$, and over all components of vector or matrix arguments. Put $r_k = r(x_k, j)$. We need to show that

$$\sup |r(x, j) - r_k| \le \sup \left| \frac{\delta r}{\delta |\nabla g|} \right| \sup |\nabla g(x) - \nabla g(x_k)| \le C \cdot r_k.$$

On the one hand,

$$\frac{\delta r}{\delta |\nabla g|} = -\frac{1}{2} \frac{2^{2j}}{(2^j + 2^{2j} |\nabla g(x)|)^{3/2}} = -\frac{1}{2} 2^{2j} r_k^3. \tag{60}$$

On the other hand,

$$|\nabla g(x) - \nabla g(x_k)| \le r_k \sup |Hg(x)|, \tag{61}$$

where $Hg(x)$ is the Hessian of g. In order to estimate $|Hg(x)|$, recall Landau's inequality for the interval $[0, 1]$ which reads

$$\|f'\|_\infty \le \frac{2}{h} \|f\|_\infty + \frac{h}{4} \|f''\|_\infty,$$

for all $0 \le h \le 1$ (see for example [1]). This inequality needs to be extended to two dimensions and applied twice with $\partial^\alpha g$ in place of f, $\alpha = (1, 0)$ and $(0, 1)$ respectively (where g is understood to be adequately extended to zero outside of $B_{x_k}(r_k)$). Upon choosing $h = C \cdot \|\partial^\alpha g\|_\infty^{1/2} \le C \cdot 2^{-j} r_k^{-1} \le 1$, with the constant C determined by the condition $h \le 1$, it follows that (the sup is still over $x \in B_{x_k}(r_k)$),

$$\sup |Hg(x)| \le C \cdot \sup |\nabla g(x)|^{1/2} \le C \cdot 2^{-j} r_k^{-1}. \tag{62}$$

As always the constant C changes from line to line. From Eqs. (60)–(62), we check that

$$\sup |r(x, j) - r_k| \le C \cdot 2^j r_k^3$$

This is dominated by $C \cdot r_k$ because $r_k \le 2^{-j/2}$, and we are done.

Proof of Lemma 6 The coarea for BV functions in the unit square $\Omega \subset \mathbb{R}^2$ is

$$\int |\nabla g(x)|\, dx = \int_{-\infty}^{\infty} H^1(g^{-1}(t) \cap \Omega)\, dt. \tag{63}$$

Written as above, the formula is valid for Lipschitz functions, the quantity $|\nabla g(x)|$ must be interpreted in a suitable measure-theoretic sense and the proof is rather technical. For C^2 functions, the proof is more accessible and can be found in [38, pp. 76] and following.

In our case $g \in C^2([0,1]^2)$ so the level sets $g^{-1}(t) \cap \Omega$ of g have bounded Hausdorff-H^1 measure for almost every t, and

$$g^{-1}(t) \cap \Omega \equiv \partial N_t, \quad \text{a.e. } t,$$

where $N_t = \{x \in \Omega, g(x) \leq t\}$. We can let $X_t = N_t \backslash N_{-t}$ as in the wording, and apply the coarea formula to the function defined as

$$\tilde{g}(x) = \begin{cases} g(x) & \text{if } |g(x)| \leq t, \\ -t & \text{if } g(x) < -t, \\ t & \text{if } g(x) > t. \end{cases}$$

Since sets of zero measure do not contribute in the integral in t, we obtain

$$\int_{X_t} |\nabla g(x)|\, dx = \int_{-t}^{t} H^1(\partial N_t)\, dt \leq 2t \sup_u H^1(\partial N_u).$$

We leave it as an exercise to the interested reader to prove that there is another, perhaps more visual way to derive the above formula from the Reynolds transport theorem. □

Proof of inequality (45) Let $\tilde{\epsilon} > 0$ and $\tilde{r}_k(x', \xi', t)$ be the separation remainder of $b_k(x', \xi', t)$ for that $\tilde{\epsilon}$ in L^2. We invoke the strong version of Lemma 2 to obtain control on $\tilde{r}_k$ in $W^{s,\infty}$,

$$|\partial_{\xi'}^{\alpha} \partial_{x'}^{\beta} \tilde{r}_k(x', \xi', t)| \leq C_{\alpha\beta} \tilde{\epsilon}.$$

In the original variables x and ξ, let $r_k(x, \xi, t) = \tilde{r}_k(x', \xi', t)$ so the condition becomes

$$|\partial_{\xi}^{\alpha} \partial_{x}^{\beta} r_k(x, \xi, t)| \leq C_{\alpha\beta} \tilde{\epsilon} (1 + |\xi|)^{-|\alpha|/2 + |\beta|/4},$$

i.e., $\frac{r_k}{\tilde{\epsilon}}$ is a symbol of order zero and type $(1/2, 1/4)$. Owing to the decomposition $K_{\text{stat}} = \sum_k K_k$ using indicators $q_k(x)$, the total separation remainder is actually the sum $r = \sum_k r_k$. Although each sum contains $O(2^j)$ terms, by the constant overlap property (40) for each given x there is a constant number of terms (independent of j) contributing in $\sum_k K_k$. Likewise, the separated components of $q_k(x)$ are all supported

on balls centered at x_k with radius twice the diameter of supp(q_k), so for each given x there is a constant number of terms contributing in $\sum_k r_k$. Hence the symbol property transfers to r,

$$|\partial_\xi^\alpha \partial_x^\beta r(x,\xi,t)| \le C_{\alpha\beta}\tilde{\epsilon}(1+|\xi|)^{-|\alpha|/2+|\beta|/4}.$$

We conclude by standard pseudo-differential calculus that r is bounded in L^2 with a norm not exceeding $C\tilde{\epsilon}$ for some constant C, which by choosing $\tilde{\epsilon}$ small enough can be made less than $\epsilon/4$ as in Eq. (45). The point of the analysis is that the L^2 bound on r is not only small but independent of j.

Appendix B: Analysis of the repeated squaring algorithm

We make two idealizations:

- we assume exact arithmetic,
- we ignore accuracy issues due to truncation of the computational domain.

The second restriction can be heuristically removed if we observe that solutions of the wave equation can be bandlimited, hence sampled without errors, and at the same time compactly supported well inside $[0,1]^2$ up to an arbitrarily small error (in any L^p norm). Or, alternatively, we can observe that truncation of the spatial domain is simply the way to encode periodic boundary conditions. We apologize to the reader for deferring this minor technical point for the time being.

B.1 Discretization by sampling

We assume throughout this appendix that the reader is familiar with Shannon's sampling theory.

Discretization of a scalar function $f(x)$, $x \in \mathbb{R}^2$, to $O(N^2)$ degrees of freedom is defined as $f_{ij} = f(x_{ij})$ on the N-by-N Cartesian grid $x_{ij} = (i,j)/N$ covering $[0,1]^2$. As announced, we assume that all our functions decay sufficiently fast at infinity so that samples outside the grid contribute a negligible error. Discretization of a vector function is defined similarly, component-by-component.

One can form back the bandlimited interpolant $\overline{f}(x)$ for $x \in [0,1]^2$ as

$$\overline{f}(x) = \sum_{ij} h(x - x_{ij}) f_{ij},$$

where h is the Dirichlet kernel. Intuitively, $\overline{f}$ is a lowpass-filtered, periodized version of f.

As before, we call wave atoms φ_μ. We have a result of faithfulness of the wave atom transform:

Lemma 7 *The transform is implemented on the array indexed by the integers* $0 \le t_1, t_2 < n$. *Let* $\varphi^D_{j,\mathbf{m}} = \varphi_{j,\mathbf{m},0}$ *be the "mother wave atom" at scale* j *and wave vector*

indexed by $\mathbf{m}$*, and* $\varphi^{\sharp}_{j,\mathbf{m}}$ *denote its periodization over the unit square* $[0,1]^2$*,*

$$\varphi^{\sharp}_{j,\mathbf{m}}(x_1,x_2) = \sum_{m_1\in\mathbb{Z}}\sum_{m_2\in\mathbb{Z}} \varphi^{D}_{j,\mathbf{m}}(x_1+m_1, x_2+m_2).$$

Let $L_j = 2^{-j+1}$ *(twice the spatial sampling at scale* j*). In exact arithmetic, the discrete coefficients are given by*

$$c^D(j,\mathbf{m},\mathbf{n}) = \frac{1}{n^2}\sum_{t_1=0}^{n-1}\sum_{t_2=0}^{n-1} f[t_1,t_2]\,\overline{\varphi^{\sharp}_{j,\mathbf{m}}}\left(\frac{t_1}{n}-\frac{k_1}{L_j}, \frac{t_2}{n}-\frac{k_2}{L_j}\right). \tag{64}$$

This is a discrete circular convolution if and only if L_j *divides* n*.*

The proof is very similar to that of a corresponding result for curvelets (Proposition 6.1 in [9]), because both transforms use the same "wrapping" trick for the implementation. The interested reader can transpose line-for-line the proof of the curvelet result to obtain the proof for wave atoms.

Applying the fast wave atom transform to the data f_{ij}, Lemma 7 asserts that (in exact arithmetic) the discrete coefficients can be given the interpretation of continuous inner products,

$$c^D_\mu = \langle \varphi^{\sharp}_\mu, \overline{f}\rangle_{L^2([0,1]^2)}, \tag{65}$$

where $\sharp$ denotes periodization. Equivalently,

$$c^D_\mu = \langle \varphi_\mu, \overline{f}\rangle_{L^2(\mathbb{R}^2)},$$

where $\overline{f}$ is defined outside of $[0,1]^2$ by periodic extension. The set of subscripts corresponding to a grid of size N-by-N is

$$\mathcal{M}_N = \{\mu : x_\mu \in [0,1]^2 \quad \text{and} \quad \xi_\mu \in [-\pi N, \pi N]^2\}.$$

Under the idealization that truncation of the spatial domain affects accuracy in a negligible manner, we allow ourselves to identify $\varphi^{\sharp}_\mu$ with φ_μ, and use the tight frame property to obtain

$$\sum_{\mu\notin\mathcal{M}_N} |c_\mu|^2 + \sum_{\mu\in\mathcal{M}_N} |c_\mu - c^D_\mu|^2 = \int_{[0,1]^2} |f(x)-\overline{f}(x)|^2\,dx + \text{negligible}.$$

The first term can be put to zero if furthermore we assume that f is bandlimited inside the square $[-\frac{\pi}{2}N, \frac{\pi}{2}N]$, a cell *twice* smaller than what is allowed by the Shannon sampling theorem. This is because wave atoms are compactly supported in frequency, so if $\hat{f}$ is supported away from the boundary $\max(|\xi_1|,|\xi_2|) = \pi N$ then f has zero component along the φ_μ corresponding to subgrid scales. For this reason, the wave

atom transform should always be implemented on a grid *twice finer* than the one on which the initial data is given.

When $\|f - \overline{f}\| \leq \epsilon \|f\|$, and has the proper bandlimit so that its coefficients also obey $\|c - c^D\| \leq \epsilon \|f\|$, then we say f is ϵ-*discretizable* on the N-by-N Cartesian grid x_{ij} spanning $[0, 1]^2$.

Note in passing that the initialization step for the wave atom repeated squaring (Sect. 3.2.1) uses the wave atom transform, therefore inherits its accuracy.

B.2 Main result

Recall that we are solving $u_t = Au$, where u has m components and A obeys all the assumptions of Theorem 1. The following theorem establishes that if the solution is ϵ-discretizable then the repeated squaring will obtain it with comparable accuracy in near-optimal asymptotic complexity.

Define $\tilde{u}(T, x)$ for $x \in [0, 1]^2$ as the bandlimited interpolant of the discrete solution $u_{ij}(T)$ obtained from the repeated squaring algorithm.

Theorem 8 *Take for the discrete initial data the samples $u_0(x_{ij})$ of a continuous function u_0 on the Cartesian N-by-N grid covering $[0, 1]^2$. For every $\epsilon > 0$ and $T > 0$ such that u_0 is $\epsilon/2$-discretizable in $\mathcal{M}_N$, for all $0 \leq t \leq T$, there exists a choice of N_g, J, B, such that:*

- *we have accuracy*

$$\|u(T) - \tilde{u}(T)\|_{L^2([0,1]^2, \mathbb{R}^m)} \leq \epsilon \|u_0\|_{L^2(\mathbb{R}^2)},$$

  and
- *for every $\delta > 0$, there exists $C_{\epsilon\delta} > 0$ such that the repeated squaring algorithm runs in less than $C_{\epsilon\delta} N^{2+\delta}$ flops. Furthermore,*

$$C_{\epsilon\delta} \leq C_{\delta M} \epsilon^{-1/M},$$

  for every $M > 0$.

Remark We will not rigorously study the dependence of the constants $C_{\epsilon\delta}$ or $C_{\delta M}$ on δ. Let us only observe that $C_{\epsilon\delta} \to \infty$ as $\delta \to 0$, naturally. We believe that the rate of blowup may go like $e^{1/\delta}$, or even worse.

Remark N is not constrained to be large in the above result, other than by the fact that the initial data—and by way of consequence the solution itself—need to be "ϵ-discretizable" on the N-by-N grid.

Proof We claim that since u_0 is $\epsilon/2$-discretizable, we may focus on the approximation of the Green's function. Indeed, denote c the sequence of coefficients of u_0, c^D its numerical counterpart, E the wave atom matrix of the Green's function and $\tilde{E}$ its

numerical counterpart. Notice that $\|E\|_{\ell^2\to\ell^2} = 1$ by anti-symmetry of the generator A and the tight frame property. If for the initial condition

$$\|c - c^D\|_{\ell^2} \le \epsilon/2\|c\|_{\ell^2},$$

and if we can prove $\|E - \tilde{E}\|_{\ell^2\to\ell^2} \le \epsilon/4$, then

$$\begin{aligned}\|Ec - \tilde{E}c^D\| &\le \|E\|\,\|c - c^D\| + \|E - \tilde{E}\|\|c\| + \|E - \tilde{E}\|\,\|c - c^D\| \\ &\le \left(\frac{3\epsilon}{4} + \frac{\epsilon^2}{8}\right)\|c\|\end{aligned}$$

and we would be done (if $\epsilon \le 2$ then $\frac{3\epsilon}{4} + \frac{\epsilon^2}{8} \le \epsilon$ as required, otherwise change the fractions of ϵ accordingly.)

B.3 Geometric squeezing

We need to control *geometric squeezing*, the phenomenon that waves can acquire a smaller wavelength when they are slowed down. This aspect is important if we wish to sample the solution correctly, above the Nyquist sampling rate. As a result, *the grid on which computations are performed often needs to be finer than the grid on which the initial data is defined.* Hence our notation $N_g \ge N$ for the actual number of grid points per dimension.

Denote P_N the projector onto wave numbers in the square $[-\pi N, \pi N]^2$,

$$\widehat{P_N f}(\xi) = \begin{cases} f(\xi) & \text{if } \xi \in [-\pi N, \pi N]^2, \\ 0 & \text{otherwise.}\end{cases}$$

Define σ as the Lyapunov exponent of geometrical optics, i.e. the smallest number σ such that

$$\max_{\pm} \sup_{x_0} \max_{|\xi_0|=1} |\xi(t)| \le e^{\sigma t},$$

when $(x(t), \xi(t))$ solve either Hamiltonian system

$$\begin{cases} \dot{x}(t) = \pm c(x(t))\frac{\xi(t)}{|\xi(t)|}, & x(0) = x_0, \\ \dot{\xi}(t) = \mp\nabla c(x)|\xi(t)|, & \xi(0) = \xi_0. \end{cases} \tag{66}$$

indexed by $\pm$. As a result, if $\Phi(x, \xi, t)$ solves either Hamilton–Jacobi equation

$$\frac{\partial\Phi_\pm}{\partial t}(t, x, \xi) = \pm c(x)|\nabla_x\Phi_\pm(t, x, \xi)|, \quad \Phi_\pm(0, x, \xi) = x \cdot \xi,$$

then

$$\sup_{\pm}\sup_{x}\sup_{\xi}|\nabla_x\nabla_\xi\Phi(t,x,\xi)|\le e^{\sigma t}. \tag{67}$$

When relevant, we term σ the *geometric squeezing* parameter. Its value directly depends on the size of the derivatives of the medium's physical properties (coefficients in the generator A.)

To the best of our knowledge, the following important result is (surprisingly) absent from the numerical analysis literature.

Lemma 8 *Let $E(t)=e^{tA}$ under the assumptions of Theorem 1, as an operator from L^2 to L^2. For every $\alpha>1$,*

$$\|(I-P_{\alpha Ne^{\sigma t}})E(t)P_N\|_{L^2\to L^2}\le C_M((\alpha-1)N)^{-M}.$$

Proof of Lemma 8 Let us consider the kernel of $E(t)$ in frequency:

$$\hat u(t,\xi)=\int\hat E(t,\xi,\eta)\hat u_0(\xi)\,d\xi$$

We know from [8] that

$$\hat E(t,\xi,\eta)=\int e^{-ix\cdot\xi}e^{i\Phi(t,x,\eta)}a(t,x,\eta)\,dx+\text{ smoothing.} \tag{68}$$

Let us treat the oscillatory integral first. Define the 'no-action' differential operator

$$L=\frac{-\Delta_x+i\Delta_x\Phi(t,x,\eta)}{|\nabla_x\Phi(t,x,\eta)-\xi|^2},$$

chosen so that it leaves the exponential invariant in (68). We can create M copies of it, integrate by parts, and define $a^{(M)}$ from the following equation for the modified amplitude,

$$L^Ma(t,x,\xi)=a^{(M)}(t,x,\xi)\left(\frac{(1+|\eta|^2)^{1/2}}{|\nabla_x\Phi(t,x,\eta)-\xi|^2}\right)^M. \tag{69}$$

It is easy to see that $a^{(M)}$ is a symbol $S^0_{1,0}$ of order zero. The rationale for introducing L is that:

- the presence of the first projector $I-P_{\alpha Ne^{\sigma t}}$ implies

$$|\xi|\ge\alpha\pi Ne^{\sigma t},$$

whereas

- the second projector P_N implies $|\eta| \leq \pi N$. Invoking $\nabla_x \Phi(t, x, \eta) = \eta \cdot \nabla_x \nabla_\eta \Phi(t, x, \eta)$ and Eq. (67), we deduce the bound

$$|\nabla_x \Phi(t, x, \xi)| \leq \pi N e^{\sigma t}.$$

As a result, the two wave numbers ξ and $\nabla_x \Phi(t, x, \eta)$ are geometrically separated by a distance greater than $(\alpha - 1)\pi N e^{\sigma t}$. The factor inside the big parentheses in Eq. (69) is therefore bounded by $C_t(\alpha - 1)^{-2} N^{-1}$. We can produce as many (M) such factors as desired, at the expense of the smoothness constants of $a^{(M)}$.

The 'smoothing' term can be written in FIO form as well, with trivial phase $\Phi(t, x, \eta) = x \cdot \eta$. The same reasoning as above yields an even bigger separation between wave numbers:

$$|\xi - \eta| \geq \alpha \pi N e^{\sigma t} - \pi N.$$

The same conclusions are valid for the smoothing and oscillatory terms.

Up to constants depending on M, the kernel $\frac{E(t,\xi,\eta)}{(\alpha-1)^{-2M} N^{-M}}$ is bounded on L^2, uniformly in α and N. Replacing $(\alpha - 1)^2$ by $(\alpha - 1)$ poses no problem. This concludes the proof of Lemma 8. □

Presumably, in practice α in Lemma 8 can be taken small (very close to 1), particularly when N is large. As a result of this lemma, and of the remark made earlier on how the transform should be implemented on a grid twice finer than initially, we see that it suffices to inflate the grid to

$$N_g \geq (2 + \alpha) N e^{\sigma T},$$

(or the power of two immediately greater) to make sure the sampling rate is above Nyquist at all times.

Let us get back to the proof of Theorem 8.

B.4 Initialization

A simple Euler explicit time step suffices for the initilaization of the Green's function at time Δt.

We put $E_n = E(t_n)$ at $t_n = 2^n \Delta t$. All the computations are done in the wave atom domain, and so will the analysis.

In what follows, we will omit the indices $\mu, \nu; \mu', \nu'$ yet understand that we are working with the wave atom representations of our operators.

By $\mathrm{Trunc}_B E(t)$ we mean truncation to a band-diagonal, or shifted band diagonal form in the wave atoms representation, where B elements per row and column are kept. Whether the diagonals are shifted or not depends on time. The operator Q_J denotes the projector onto wave atom scales $0 \leq j \leq J$. We will use ℓ^p norms in the transformed domain; for operators $\|E\|_p$ means the $\ell^p \to \ell^p$ norm. Also, statements

referring to $\|E\|_{p,q}$ are valid both in ℓ^p and ℓ^q. In the sequel, $1 \le p, q \le \infty$ will suffice.

The scale cutoff J should be chosen so that the Nyquist sampling rate is respected on a N-by-N grid. In wave atoms $2^{2J} \simeq N$. In that case we say "J is compatible with N".

Lemma 9 *Define $\tilde{E}_0$ in wave atoms as*

$$\tilde{E}_0 = Trunc_B\,(Q_J + \Delta t\, Q_J A Q_J).$$

Choose $\Delta t \le N^{-1}$, and J compactible with N. Then

$$\|Q_J E_0 Q_J - \tilde{E}_0\|_2 \le C \cdot (\Delta t)^2 N^2 + C_M B^{-M},$$

for all $M > 0$. The same bound holds if the L^2 norm is replaced by a ℓ_1 or ℓ_∞ norm in wave atoms.

Proof of Lemma 9 The thresholding operation Trunc_B is best controlled in ℓ^1 or ℓ^∞ because those norms operate directly on the matrix elements.

We consider the intermediate approximation $\mathrm{Trunc}_B(Q_J E_0 Q_J)$. An immediate consequence of Corollary 1 is that

$$\|Q_J E_0 Q_J - \mathrm{Trunc}_B(Q_J E_0 Q_J)\|_{1,\infty} \le C_M B^{-M}.$$

For the other contribution, take a coefficient sequence $c \equiv c_{\mu,\nu}$ of the form $c = Q_J c$, for a scale J corresponding to N-by-N grid points. In wave atoms, $J \simeq \frac{1}{2}\log_2 N$ up to a constant. Choose a time step $\Delta t \lesssim N^{-1}$ or perhaps much smaller than that bound (the bound is the CFL condition).

$$\|(\mathrm{Trunc}_B(Q_J E_0 Q_J) - \tilde{E}_0)c\|_{1,\infty} \le \|(e^{\Delta t A} - I - \Delta t A)c\|_{1,\infty}$$

At this point we would like to convey the idea that A is differential of order one, hence produces a factor N when acting on $Q_J c_0$, and that $e^{\Delta t\lambda} - I - \Delta t\lambda = O((\Delta t)^2\lambda^2)$. We are in the presence of functions of operators, however, so we need to be more careful. Fortunately, we can avoid invoking spectral theory. It suffices to notice that exponentials obey the identity

$$e^{\Delta t A} - I - \Delta t A = \int_0^{\Delta t}\int_0^{s} A^2 e^{uA}\, du\, ds,$$

and use Minkowski's integral inequality twice. We are left with the bound

$$C \cdot (\Delta)^2 \|A^2 c\|_{1,\infty}.$$

The operator $A^2\Delta^{-1}$ is pseudodifferential of order zero, componentwise, so it comes as a very special case of Theorem 1 that it is bounded on ℓ^p for all $0 < p \le \infty$. The

action of the Laplacian Δ on c is obviously to turn each wave atom into a molecule with weight 2^{2j} whithout otherwise affecting sparsity. By the scale cutoff $J \simeq \log N$, we can pull out a factor $2^{2J} \simeq N^2$, and obtain the bound

$$C(\Delta)^2 N^2 \|c\|_{1,\infty}.$$

By grouping the bounds we have obtained so far we get

$$\|Q_J E_0 Q_J - \tilde{E}_0\|_{1,\infty} \leq C \cdot (\Delta t)^2 N^2 + C_M B^{-M}.$$

The same conclusion is valid in ℓ^2 by interpolation between ℓ^p spaces. □

Lemma 9 can probably be improved if a higher-order time integration method is used, like a Runge-Kutta method. The proof would however become more complicated than it would appear, mostly due to the fact that $AQ_JA \neq A^2$.

B.5 Iterations: thresholding and scale cutoff

For the main part of the proof, our strategy is to set up the difference $R_n = E_n - \tilde{E}_n$ between actual and computed propagator at dyadic time t_n, and to study it in an adequate norm in wave atom space. Since scale cutoffs introduce numerical errors that propagate from the smaller to the larger scales, we'll introduce a *weak* norm on the wave atom coefficients to forgive this effect, at the expense of increasing the grid size N_g.

We need to define a sequence of *decreasing* scales $J_0 \geq J_1 \geq \ldots \geq J_{n^*}$, which are meant to define concentric "balls" in the discrete frequency plane ξ_μ through

$$|\mu| \leq J_n \quad \Leftrightarrow \quad |\xi_\mu| \leq 2^{2J_n},$$

such that (1) the restriction $Q_{J_{n^*}}$ to $|\mu| \leq J_{n^*}$ in wave atom space still captures the solution very accurately at all times of interest, and such that (2) two successive scales are "almost orthogonal" in the sense that, for all ν'',

$$\omega(\mu, \mu'_{\nu''}(t_{n+1})) \geq C_S, \quad |\mu| \leq J_{n+1}, \ |\mu'| \geq J_n, \tag{70}$$

for some constant $C_S > 0$ to be determined. (Note that J plays a role similar to j in the discrete construction of wave atoms.)

The property (1) that J_{n^*} is large enough reads, in the notations of Lemma 8,

$$\|(E(t) - Q_{J_{n^*}} E(t)) P_N u_0\| \leq C_M((\alpha-1)N)^{-M}, \quad \text{for all } M > 0, \text{for all } t \leq T. \tag{71}$$

When the norm is $L^2 \to L^2$, or $\ell_2 \to \ell_2$ in wave atoms, Eq. (71) is precisely justified by Lemma 8, compact support of wave atoms in frequency space, and, say, the choice $2J_{n^*} \geq 2\alpha N e^{\sigma T}$. When the norm is $\ell_1 \to \ell_1$ or $\ell_\infty \to \ell_\infty$ in wave atoms, the justification is similarly done in the frequency plane, and left to the reader.

The property (2) relating to almost-orthogonality requires J_n to be sufficiently greater than J_{n+1}. Essentially, we need to make sure that a wave atom at scale $|\mu'| \leq 2^{2J_{n+1}}$, after evolution under $E(t_{n+1})$, is still almost orthogonal to all wave atoms with $|\mu| \geq 2^{2J_n}$. In view of Lemma 8, and the general sparsity theory, it suffices to impose

$$2^{2J_{n+1}} e^{\sigma t_{n+1}} \leq 2^{2J_n} - 2^{J_n} C_S. \tag{72}$$

Indeed, a substitution in (8) (further using the bound $||\xi_\mu| - |\xi_{\mu'}|| \leq |\xi_\mu - \xi_{\mu'}|$) shows that (70) holds. The sequence of scales is then defined iteratively from J_{n^*} down to J_0 using (72). It is immediate to deduce from (72) that—as soon as $C_S = O(N^{1/4})$, say, which we will check later—we have the uniform bound

$$2^{2J_n} \leq C\alpha e^{5\sigma T} N \log N, \quad 0 \leq n \leq n^*,$$

where C is a number. This dictates an inflation of the grid size to $N_g' = 2^{2J_0}$, which is only a $C \cdot \log N$ factor away from the old value, a harmless modification from the viewpoint of asymptotic complexity estimates.

The J of the previous sections and Lemmas should now be chosen as the slightly larger J_0. The consequence of a scale cutoff in terms of wave vectors ξ_μ should now be understood as $|\xi_\mu| \leq 2^{2J_n}$.

It is now time to state that the "Trunc" operation in the description of the algorithm should in fact be accompanied, at time t_n, by the scale cutoff Q_{J_n}. This is perhaps not necessary in practice (and in fact decreases the algorithmic complexity), but we found it useful for the analysis. So we let, by induction,

$$\tilde{E}_{n+1} = Q_{J_{n+1}} \mathrm{Trunc}_B(\tilde{E}_n^2) Q_{J_{n+1}}.$$

Put $R_n = E_n - \tilde{E}_n$. We will measure R_n in the wave atom $\ell_1 \to \ell_1$ norm, but only when we let it act on functions u with admissible frequency, meaning $u = Q_{J_n} u$. So we are interested in $\|R_n Q_{J_n}\|_1$. (This is the sense in which the norm is "weak", as announced earlier). By Lemmas 9 and 8, we have

$$\begin{aligned} \|R_0 Q_{J_0}\|_1 &\leq \|(E_0 - Q_{J_0} E_0) Q_{J_0}\|_1 + \|(Q_{J_0} E_0 - \tilde{E}_0) Q_{J_0}\|_1, \\ &\leq C_M((\alpha - 1)N)^{-M} + C(\Delta t)^2 N^2 + C_M B^{-M} \equiv \epsilon_0. \end{aligned} \tag{73}$$

The choice of parameters B, α and Δt that makes ϵ_0 sufficiently small will be explained later. We let $\epsilon_n = \|R_n Q_{J_n}\|_1$ and seek to bound it by induction. We start by noticing that not much can be said in the absence of the Q_{J_n} factor:

$$\|R_n\|_1 \leq \|R_n Q_{J_n}\|_1 + \|R_n(I - Q_{J_n})\|_1 \leq \epsilon_n + \|E_n\|_1 \leq \epsilon_n + C, \tag{74}$$

where C is a constant depending at most on n by Corollary 1 (hence a universal constant for $0 \leq n \leq n^*$).

For the induction analysis, we decompose

$$\begin{aligned} R_{n+1} &= E_{n+1} - Q_{J_n}\mathrm{Trunc}_B(\tilde{E}_n^2) \\ &= (E_{n+1} - Q_{J_n}\mathrm{Trunc}_B(E_{n+1})) + Q_{J_n}\mathrm{Trunc}_B(R_nE_n + E_nR_n + R_n^2) \\ &= (E_{n+1} - \mathrm{Trunc}_B(E_{n+1})) + \mathrm{Trunc}_B((I - Q_{J_n})E_{n+1}) \\ &\quad + Q_{J_n}\mathrm{Trunc}_B(R_nE_n + E_nR_n + R_n^2). \end{aligned} \tag{75}$$

This expression is postmultiplied by $Q_{J_{n+1}}$ in the expression of ϵ_{n+1}. Let us consider the five terms separately:

- The first term is bounded by $C_M B^{-M}$, by means of Corollary 1.
- For the other terms, we will systematically drop the Trunc operation since the $\ell_1 \to \ell_1$ norm operates directly on the matrix elements. In the second term, we are faced with $(I - Q_{J_n})E_{n+1}Q_{J_n}$: recall that the scales J_{n+1} and J_n have been chosen so that (70) holds, hence by Corollary 1 we have the bound $C_M C_S^{-M}$ for all $M > 0$.
- The third term in (75) is $R_nE_nQ_{J_{n+1}} = R_n(Q_{J_n} + I - Q_{J_n})E_nQ_{J_{n+1}}$. In the first contribution, relative to Q_{J_n}, we identify the bound ϵ_n times a constant C, since $\|E(t)\|_1$ is bounded uniformly in $0 \le t \le T$ by Corollary 1. The second contribution, relative to $I - Q_{J_n}$, is like above and yields another $C_M C_S^{-M}$. The remaining $\|R_n\|_1$ is handled by the coarse bound (74).
- The fourth term is handled in a similar manner: $R_nQ_{J_{n+1}}$ is a restriction of $R_nQ_{J_n}$ and therefore yields ϵ_n in the ℓ_1 norm. As before, we use $\|E(t)\|_1 \le C$.
- The fifth term contains $R_n^2Q_{J_{n+1}}$. The rightmost $R_nQ_{J_n}$ yields ϵ_n, and the leftmost R_n is handled by the coarse bound (74).

We summarize our majorations as follows: there exists a constant D (depending at most on n), such that

$$\epsilon_{n+1} \le D\epsilon_n + C_M B^{-M} + C_M C_S^{-M}.$$

We solve this difference inequation and its initial condition (73) as

$$\epsilon_n \le D^n \left[C_M B^{-M} + C_M C_S^{-M} + C_M((\alpha - 1)N)^{-M} + C(\Delta t)^2 N^2 \right].$$

Notice that the same bound holds for $\tilde{\epsilon}_n = \|R_nQ_{J_n}\|_\infty$ in the ℓ_∞ norm, by transposition because the adjoint $E^*(t)$ solves the backward hyperbolic equation, for which the exact same sparsity results hold. The same bound therefore holds for $\|R_nQ_{J_n}\|_2$, by interpolation in ℓ_p spaces.

The first factor in the expression of ϵ_n can be bounded by

$$D^n \le D^{n^*} \le D^{C\log N} = N^{C'},$$

for some constant C'. The values for B, α, C_S and Δt need to be chosen to not only balance the growth $N^{C'}$ in N, but also to make sure that $\epsilon_{n^*} \le \epsilon/4$ (see the discussion

after Theorem 8). These conditions are satisfied as soon as we take

$$B \geq CN^{\delta}, \quad C_S \geq CN^{\delta}, \quad \alpha \geq C, \tag{76}$$

for arbitrary small $\delta > 0$, but some adequately large constant C, and

$$\Delta t \leq C_1 N^{-C_2},$$

for, again, some adequately large C_1, C_2. We check that this value of Δt still implies a reasonable number of time steps: $n^* \asymp \log N$. We also check that $C_S \lesssim N^{\delta}$ is compatible with the bound $C_S \lesssim N^{1/4}$ required above. Furthermore, it is easy to see that all the constants in the above equations depend on $\epsilon^{-1/M}$, for all $M > 0$.

The conclusion of this section is that the overall error on the discrete Green's function can be made less than $\epsilon/4$, provided the parameters B and N'_g (via α) of the algorithm are chosen adequately large, through (76), and Δt sufficiently small.

B.6 Complexity

Now that the accuracy estimate has been established, the complexity analysis itself is quite simple. It is mostly a special case of the analysis in Sect. 3.3. Let us review the different components of the algorithm:

- The wave atom transform is an $N^2 \log N$ operation, and needs to be applied once to the initial condition. The inverse wave atom transform is also $N^2 \log N$, and needs to be applied once to the final solution.
- The initialization should be done using the separation trick of Sect. 3.3.1. The complexity of this step is $O(N^2 \log N)$.
- The repeated squaring is done on matrices of size $O(N'_g) \times O(N'_g)$, but sparse with band size $B = O(N^{\delta})$. Recall that $N'_g = O(N^2 \log N)$, and that δ can be made arbitrary small. So one matrix-matrix multiplication costs $O(N^{2+2\delta} \log N)$ operations. Since there are $O(\log N)$ time steps, we have a total complexity of $O(N^{2+2\delta} \log^2 N)$.
- Recall that we are not dealing with banded matrix, strictly speaking, but matrices with a fixed number of "shifted band-diagonals", with an especially adapted notion of distance between matrix indices. The question of accurately predicting the *location* of these band diagonals can probably be answered in several ways. The solution we chose is an explicit tracking of all the trajectories $(x_\mu(t), \xi_\mu(t))$ by the *phase-flow method*, developed by two of the authors in [37]. The phase-flow method was precisely designed with the task of tracking wave atoms (or curvelets) in mind. Optimal complexity of the phase-flow method ($O(N^2)$ in our context) has been fully discussed in [37].

After renaming δ, we can read the optimal complexity as $O(N^{2+\delta})$, with a constant that depends on δ, and also on ϵ like $\epsilon^{-1/M}$. The proof is complete. □

References

1. Babenko, Yu.V.: Pointwise inequalities of Landau-Kolmogorov-type for functions defined on a finite segment. Ukr. Math. J. **53**(2), 270–275 (2001)
2. Bacry, E., Mallat, S., Papanicolaou, G.: A wavelet based space-time adaptive numerical method for partial differential equations. Math. Model. Num. Anal. **26**(7), 793 (1992)
3. Beylkin, G., Coifman, R.R., Rokhlin, V.: Fast wavelet transforms and numerical algorithms. Comm. Pure Appl. Math. **44**, 141–183 (1991)
4. Beylkin, G., Mohlenkamp, M.J.: Algorithms for numerical analysis in high dimensions. SIAM J. Sci. Comput. **26**(6), 2133–2159 (2005)
5. Beylkin, G., Sandberg, K.: Wave propagation using bases for bandlimited functions. Wave Motion **41**(3), 263–291 (2005)
6. Candès, E.J.: Harmonic analysis of neural networks. Appl. Comput. Harmon. Anal. **6**, 197–218 (1999)
7. Candès, E.J., Demanet, L.: Curvelets and Fourier integral operators. C. R. Acad. Sci. Paris Ser. I **336**, 395–398 (2003)
8. Candès, E.J., Demanet, L.: The curvelet representation of wave propagators is optimally sparse. Comm. Pure Appl. Math. **58**(11), 1472–1528 (2005)
9. Candès, E.J., Demanet, L., Donoho, D.L., Ying, L.: Fast discrete curvelet transforms. SIAM Mult. Model. Sim. (2005, submitted)
10. Candès, E.J., Donoho, D.L.: Curvelets—a surprisingly effective nonadaptive representation for objects with edges. In: Rabut, C., Cohen, A., Schumaker, L.L. (eds.) Curves and Surfaces, pp. 105–120. Vanderbilt University Press, Nashville (2000)
11. Candès, E.J., Donoho, D.L.: New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. Comm. Pure Appl. Math. **57**, 219–266 (2004)
12. Cohen, A.: Numerical Analysis of Wavelet Methods. North-Holland/Elsevier, Amsterdam (2003)
13. Cohen, A., Dahmen, W., DeVore, R.: Adaptive wavelet methods for elliptic operator equations—Convergence rates. Math. Comp. **70**(233), 27–75 (2000)
14. Córdoba, A., Fefferman, C.: Wave packets and Fourier integral operators. Comm. PDE **3**(11), 979–1005 (1978)
15. Demanet, L.: Curvelets, wave atoms, and wave equations. Ph.D. Thesis, California Institute of Technology, May (2006)
16. Demanet, L., Ying, L.: Wave atoms and sparsity of oscillatory patterns (2006, submitted)
17. Douma, H., de Hoop, M.V.: Wave-character preserving prestack map migration using curvelets. Presentation at the Society of Exploration Geophysicists, Denver, CO (2004)
18. Duistermaat, J.: Fourier Integral Operators. Birkhauser, Boston (1996)
19. Engquist, B., Osher, S., Zhong, S.: Fast wavelet based algorithms for linear evolution equations. SIAM J. Sci. Comput. **15**(4), 755–775 (1994)
20. Fefferman, C.: A note on spherical summation multipliers. Israel J. Math. **15**, 44–52 (1973)
21. Flesia, A.G., Hel-Or, H., Averbuch, A., Candès, E.J., Coifman, R.R., Donoho, D.L.: Digital implementation of ridgelet packets. In: Stoeckler, J., Welland, G.V. (eds.) Beyond Wavelets. Academic Press, London (2003)
22. Hackbusch, W.: A sparse matrix arithmetic based on H-matrices. Part I: introduction to H-matrices. Computing **62**, 89–108 (1999)
23. Hennenfent, G., Herrmann, F.J.: Seismic denoising with unstructured curvelets. Comput. Sci. Eng. (2006, to appear)
24. Herrmann, F.J., Moghaddam, P.P., Stolk, C.C.: Sparsity- and continuity-promoting seismic image recovery with curvelet frames. (2006, submitted)
25. de Hoop, M.V., le Rousseau, J.H., Wu, R.: Generlization of the phase-screen approximation for the scattering of acoustic waves. Wave Motion **31**(1), 43–70 (2000)
26. Hörmander, L.: The Analysis of Linear Partial Differential Operators, vol. 4. Springer, Heidelberg (1985)
27. Jaffard, S.: Wavelet methods for fast resolution of elliptic problems. SIAM J. Numer. Anal. **29**(4), 965–986 (1992)
28. Lax, P.: Asymptotic solutions of oscillatory initial value problems. Duke Math J. **24**, 627–646 (1957)
29. LeVeque, R.J.: Finite Volume Methods for Hyperbolic Problems. Cambridge University Press, London (2002)
30. Meyer, Y.: Ondelettes et Opérateurs. Hermann, Paris (1990)

31. Meyer, Y., Coifman, R.R.: Wavelets, Calderón-Zygmund and Multilinear Operators. Cambridge University Press, Cambridge (1997)
32. Moler, C., Van Loan, C.: Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. SIAM Rev. **45**(1), 3–49 (2003)
33. Seeger, A., Sogge, C., Stein, E.: Regularity properties of Fourier integral operators. Ann. Math. **134**, 231–251 (1991)
34. Smith, H.: A Hardy space for Fourier integral operators. J. Geom. Anal. **8**, 629–653 (1998)
35. Smith, H.: A parametrix construction for wave equations with $C^{1,1}$ coefficients. Ann. Inst. Fourier (Grenoble) **48**, 797–835 (1998)
36. Stein, E.: Harmonic Analysis. Princeton University Press, Princeton (1993)
37. Ying, L., Candès, E.J.: The phase-flow method. J. Comput. Phys. (2006, to appear)
38. Ziemer, W.: Weakly Differentiable Functions. Springer, Heidelberg (1989)