Abstract

This paper introduces the synchrosqueezed wave packet transform as a method for analyzing 2D images. This transform is a combination of wave packet transforms of a certain geometric scaling, a reallocation technique for sharpening phase space representation, and clustering algorithms for mode decomposition. For a function that is a superposition of several wave-like components satisfying certain separation conditions, we prove that the synchrosqueezed wave packet transform identifies these components and estimates their instantaneous wavevectors. A discrete version of this transform is discussed in detail and numerical results are given to demonstrate properties of the proposed transform.

Keywords. Wave packet transform, synchrosqueezing, clustering, instantaneous wavevector, phase space representation, empirical mode decomposition.

AMS subject classifications.

1 Introduction

In many applications, a typical time signal can be viewed as a superposition of several simple components, each of which is localized in the time frequency (or phase space) representation and exhibits well-defined, often non-stationary, instantaneous frequency. An important task in analyzing these signals is to identify these simple components and estimate their instantaneous frequencies. Time frequency analysis provides a wide range of tools for this task. Most of these tools fall into two categories: linear and quadratic methods, each of which has its own strengths and weaknesses. The linear methods are typically efficient and easy to reconstruct, but provide poor resolution. The quadratic methods, on the other hand, provide better resolution but suffer from higher computational cost, more difficult reconstruction process, and non-physical interference between multiple components. Among the approaches proposed to remedy this problem, the reallocation (or reassignment) methods [1, 4, 5, 7] can be viewed as standard linear methods, followed by reassigning values of the time frequency representation based on their local oscillation. One such method is the synchrosqueezed wavelet transform, which was proposed in [7] and given rigorous justification for an important class of signals (superpositions of approximate sinusoidal waves with well-separated frequencies at each location) in [6].
An obvious question is whether the synchrosqueezing idea can be extended to 2D images. A naïve attempt would simply combine the 2D wavelet transform with the synchrosqueezing approach. The resulting synchrosqueezed 2D wavelet transform would be capable of separating components that have different frequencies at each location, just as the 1D transform does for 1D signals. However, in many situations this is not enough since a typical 2D image can have components whose wavevectors have the same magnitude but point in different directions, as shown in Figure 1(left). In fact, images from many applications related to high-frequency wave propagation have this feature. Since the 2D wavelet transform is isotropic and non-directional, this naïve synchrosqueezed 2D wavelet transform cannot distinguish these different modes that have the same frequency but different wavevectors. In order to distinguish these modes, we propose a synchrosqueezed wave packet transform that combines the synchrosqueezing idea with wave packets of an appropriate geometric scaling. The reason why it works is that these wave packets have finer and, more importantly, directional support in the Fourier domain, which allows them to distinguish components oscillating in different directions, as shown in Figure 1(right).

![Figure 1: Comparison of the resolutions of (continuous) wavelets (left) and (continuous) wave packets (right) in the Fourier domain. Consider the superposition of two plane waves $e^{2\pi ip \cdot x}$ and $e^{2\pi iq \cdot x}$ with the same frequency ($|p| = |q|$) but different wavevectors ($p \neq q$). The two dots in each plot show the support of the Fourier transforms of the superposition. The gray region in each plot stands for the support of a continuous wavelet (left) and the one of a continuous wave packet (right). Since the isotropic support of each wavelet either covers or misses both points $p$ and $q$, the wavelet transform is not able to distinguish these two plane waves (left). On the other hand, as long as $p$ and $q$ are well-separated, they are in the support of two different wave packets (right). Hence these two plane waves can be distinguished from each other by the wave packet transform.](image)

### 1.1 Synchrosqueezed wave packet transform

In what follows, we introduce the synchrosqueezed wave packet transform, alongside several simple motivating examples. Let $w(x)$ for $x \in \mathbb{R}^2$ be the mother wave packet, which is used to define all wave packets through scaling, modulation, and translation. Suppose that $w(x)$ is in the Schwartz class and the Fourier transform $\hat{w}(\xi)$ is a non-negative, radial, real-valued,
smooth function with support equal to the unit ball $B_1(0)$ in the Fourier domain. Based on $w(x)$, we can define a family of wave packets through scaling, modulation, and translation as follows, controlled by a geometric parameter $s$.

**Definition 1.1.** Given the mother wave packet $w(x)$ and the parameter $s \in (1/2, 1)$, the family of wave packets $\{w_{pb}(x), p, b \in \mathbb{R}^2\}$ are defined as

$$w_{pb}(x) = |p|^s w(|p|^s(x - b)) e^{2\pi i (x-b) \cdot p},$$

or equivalently in Fourier domain

$$\hat{w}_{pb}(\xi) = |p|^{-s} e^{-2\pi i b \cdot \xi} \hat{w}(|p|^{-s}(\xi - p)).$$

Based on this definition, we see the Fourier transform $\hat{w}_{pb}(\xi)$ is supported in $B_{|p|^s}(p)$, a ball centered at $p$ with radius $|p|^s$. On the other hand, $w_{pb}(x)$ is centered in space at $b$ with an essential support of width $O(|p|^{-s})$. $\{w_{pb}(x), p, b \in \mathbb{R}^2\}$ are all appropriately scaled to have the same $L^2$ norm with the mother wave packet $w(x)$. Notice that if $s$ were equal to 1, these functions would be qualitatively similar to the standard 2D wavelets. On the other hand, if $s$ were equal to 1/2, we would obtain the wave atoms defined in [8]. As we shall see, it is essential for our purposes that $s \in (1/2, 1)$. Equipped with this family of wave packets, we can define the wave packet transform as follows.

**Definition 1.2.** The wave packet transform of a function $f(x)$ is a function

$$W_f(p, b) = \langle w_{pb}, f \rangle = \int \overline{w_{pb}(x)} f(x) dx$$

for $p, b \in \mathbb{R}^2$.

If the Fourier transform of $\hat{f}(\xi)$ vanishes for $|\xi| < 1$, it is easy to check that the $L^2$ norms of $W_f(p, b)$ and $f(x)$ are equivalent, up to a uniform constant factor, i.e.,

$$\int |W_f(p, b)|^2 dpdb \simeq \int |f(x)|^2 dx.$$

As a simple example, let us consider the wave packet transform for a plane wave function

$$f(x) = \alpha e^{2\pi i N \beta x},$$

where $\alpha$ and $\beta$ are non-zero constants of order $O(1)$ and $N$ is a sufficiently large constant. The instantaneous wavevector is $N \beta$ and we have

$$W_f(p, b) = \int_{\mathbb{R}^2} \alpha e^{2\pi i N \beta x} |p|^s w(|p|^s(x - b)) e^{-2\pi i (x-b) \cdot p} dx$$

$$= |p|^{-s} \alpha \int_{\mathbb{R}^2} e^{2\pi i N \beta \cdot (b+|p|^{-s}y)} w(y) e^{-2\pi i p \cdot |p|^{-s}y} dy$$

$$= |p|^{-s} \alpha e^{2\pi i N \beta \cdot b} \hat{w}(|p|^{-s}(N \beta - p)).$$

Since $\hat{w}(\xi)$ is compactly supported in the unit ball, for each fixed $b$ the coefficients $W_f(p, b)$ are non-zero if $p$ satisfies

$$|p - N \beta| \leq |p|^s.$$
This implies that, for each \( b \), \( W_f(p, b) \) has a spreading of width \( O(\|N\beta\|^s) \) around the wavevector \( N\nu \) in the \( p \) variable. The essential observation of synchrosqueezing is that the oscillation of \( W_f(p, b) \) in the \( b \) variable in fact encodes the correct wavevector \( \nu \), independently of the amplitude \( \alpha \) or the position \( b \). More precisely, the derivative of \( W_f(p, b) \) with respect to \( b \) and \( W_f(p, b) \) satisfy the following equation:

\[
\frac{\nabla_b W_f(p, b)}{2\pi i W_f(p, b)} = \frac{2\pi i N\beta |p|^{-s} \alpha e^{2\pi i N\beta b \pi i \xi}}{2\pi i |p|^{-s} \alpha e^{2\pi i N\beta b \pi i \xi}} = N\beta
\]

for \( W_f(p, b) \neq 0 \).

Let us consider now a general function of form

\[
f(x) = \alpha(x)e^{2\pi i N\phi(x)}
\]

with smooth amplitude \( \alpha(x) \), smooth phase \( \phi(x) \), and sufficiently large \( N \). As we shall see, for each \( b \) the wave packet transform \( W_f(p, b) \) is essentially supported in the following set

\[
\{p : |p - N\nabla \phi(b)| \lesssim |p|^s\}.
\]

This motivates us to define the instantaneous wavevector for a general function \( f(x) \) as follows.

**Definition 1.3.** The instantaneous wavevector estimation of a function \( f(x) \) at \( (p, b) \) is

\[
v_f(p, b) = \frac{\nabla_b W_f(p, b)}{2\pi i W_f(p, b)}
\]

for \( p, b \in \mathbb{R}^2 \) such that \( W_f(p, b) \neq 0 \).

Given the wavevector estimation \( v_f(p, b) \), the synchrosqueezing step reallocates the information in the phase space and provides a sharpened phase space representation of \( f(x) \).

**Definition 1.4.** Given \( f(x) \), \( W_f(p, b) \), and \( v_f(p, b) \), the synchrosqueezed energy distribution \( T_f(v, b) \) is

\[
T_f(v, b) = \int |W_f(p, b)|^2 \delta(v_f(p, b) - v) dp
\]

for \( v, b \in \mathbb{R}^2 \).

As we shall see, for \( f(x) = \alpha(x)e^{2\pi i N\phi(x)} \) with sufficiently smooth amplitude \( \alpha(x) \) and sufficiently steep phase \( N\phi(x) \), we can show that for each \( b \) the estimation \( v_f(p, b) \) indeed approximates \( N\nabla \phi(b) \), independently of \( p \). As a direct consequence, for each \( b \), the support of \( T_f(v, b) \) in \( v \) variable concentrates near \( N\nabla \phi(b) \) (see Figure 2 for an example). For \( f(x) \) with Fourier transform vanishing for \( |\xi| < 1 \), the following norm equivalence holds

\[
\int T_f(v, b) dv db = \int |W_f(p, b)|^2 dp db \approx \|f\|^2_2
\]

as a consequence of the \( L^2 \) norm equivalence between \( W_f(p, b) \) and \( f(x) \).

Let us now informally discuss why the synchrosqueezed wave packet transform allows one to identify individual pieces of a superposition of multiple components. For simplicity, let

\[
f(x) = \alpha_1(x)e^{2\pi i N\phi_1(x)} + \alpha_2(x)e^{2\pi i N\phi_2(x)},
\]
with smooth amplitudes $\alpha_1(x)$ and $\alpha_2(x)$, smooth phases $N\phi_1(x)$ and $N\phi_2(x)$ for sufficiently large $N$ (see Figure 3(top-left)). Let us assume that at each position the instantaneous wavevectors $N\nabla\phi_1(x)$ and $N\nabla\phi_2(x)$ are sufficiently large and well-separated from each other (this will be made precise later). From the above discussion, we know that for each $p$ the wave packet transform $W_f(p,b)$ is essentially supported in two sets

$$\{(p,b) : |p - N\nabla\phi_1(b)| \lesssim |p|^s\}, \quad \{(p,b) : |p - N\nabla\phi_2(b)| \lesssim |p|^s\}.$$ 

Since both $|N\nabla\phi_1(b)|$ and $|N\nabla\phi_2(b)|$ are large, the first set is within distance $O(|N\nabla\phi_1(b)|^s)$ from $N\nabla\phi_1(b)$ and the second one is within distance $O(|N\nabla\phi_2(b)|^s)$ from $N\nabla\phi_2(b)$. Since $N\nabla\phi_1(x)$ and $N\nabla\phi_2(b)$ are sufficiently well-separated, these two sets are disjoint. Therefore,

$$v_f(p,b) \approx N\nabla\phi_1(b)$$

for $p$ in the first set and

$$v_f(p,b) \approx N\nabla\phi_2(b)$$

for $p$ in the second one (see Figure 3(top-right) for an example). After synchrosqueezing, the energy distribution $T_f(v,b)$ then essentially concentrates near $S_1 = \{(N\nabla\phi_1(b), b) : b \in \mathbb{R}^2\}$ and $S_2 = \{(N\nabla\phi_2(b), b) : b \in \mathbb{R}^2\}$, each of which is a two-dimensional surface in the four-dimensional phase space. Since $S_1$ and $S_2$ are well-separated, one expects the essential support of $T_f(v,b)$ to separate into two disjoint regions $U_1$ and $U_2$, where $U_1 \supset S_1$ contains the support of the synchrosqueezed coefficients from the first mode $\alpha_1(x)e^{2\pi i N\phi_1(x)}$, while $U_2 \supset S_2$ contains the support of the coefficients of the second one $\alpha_2(x)e^{2\pi i N\phi_2(x)}$. Computationally, these two sets $U_1$ and $U_2$ can be identified with standard clustering algorithms (see the second row of Figure 3).
Figure 3: Synchrosqueezed wave packet transform applied to the superposition of two deformed plane waves, \( f(x) = e^{2\pi i N \phi_1(x)} + e^{2\pi i N \phi_2(x)} \), which is explained in Example 2 in Section 4. Top-left: The essential support of the wave packet transform \( W_f(p, b) \) at a fixed \( b_1 \) value. Top-right: The essential support of the synchrosqueezed energy distribution \( T_f(v, b) \) at the same \( b_1 \) value. At a fixed \( b_1 \) value, \( T_f(v, b) \) is more concentrated near two curves. In general, in the full 4D phase space \( T_f(v, b) \) is concentrated near two separated 2D surfaces in the 4D phase space, which makes it easier to separate these components with clustering techniques. The second row shows the support of two sets \( U_1 \) and \( U_2 \) after the clustering algorithm is applied. Each set contains the synchrosqueezed energy distribution of one deformed plane wave.

Once \( U_1 \) and \( U_2 \) are identified, we can extract individual modes with

\[
\begin{align*}
    f_1(x) &= \int_{v_f(p, b) \in U_1} \tilde{w}_{pb}(x) W_f(p, b) dp db, \\
    f_2(x) &= \int_{v_f(p, b) \in U_2} \tilde{w}_{pb}(x) W_f(p, b) dp db,
\end{align*}
\]

where the set of functions \( \{ \tilde{w}_{pb}(x), p, b \in \mathbb{R}^2 \} \) is a dual frame of \( \{ w_{pb}(x), p, b \in \mathbb{R}^2 \} \).
1.2 Related work

Extracting individual components and estimating instantaneous wavevectors is an essential problem in adaptive data analysis and understanding. There has been a long history behind applying linear and quadratic methods from applied harmonic analysis to this problem and this line of research is summarized in [13] for example. The synchrosqueezing method was originally proposed for auditory signal processing in [7] and has been shown to provide good results for 1D signals even under substantial amount of noise. The recent work in [6] provides an important step in understanding of the synchrosqueezing approach. The method of this paper is motivated by extending the work of [7, 6] to higher dimensional cases.

A different but related approach for component extraction is the empirical mode decomposition (EMD) initiated and refined by Huang et al [11, 12]. Given a superposition of simple components, named intrinsic mode functions, this approach recursively extracts these components starting from the most oscillatory. Typically, the most oscillatory component is estimated by computing local minima and maxima and applying spline interpolation to these extrema to estimate the envelope function. Though this method has been widely used for climate data analysis, it is rather sensitive to noise and its mathematical analysis is still under development [9, 10, 16].

The rest of the paper is organized as follows. Section 2 contains the main theoretical result of this paper. We prove that when the instantaneous wavevectors of multiple components are well-separated at each position, the synchrosqueezed wave packet transform is able to estimate the instantaneous wavevectors. In Section 3 a discrete version of the synchrosqueezed wave packet transform is introduced in detail. Section 4 provides several numerical examples on instantaneous wavevector estimation and mode decomposition. Finally, we conclude with some discussions in Section 5.

2 Analysis of the transform

In this section we show that, for a superposition of multiple components with well-separated instantaneous wavevectors, the synchrosqueezed wave packet transform is able to estimate these instantaneous wavevectors. We start by providing precise definitions for these components and the superposition in 2D, following the model used in Huang et al [11] and Daubechies et al [6].

Definition 2.1. A function $f(x) = \alpha(x)e^{2\pi iN\phi(x)}$ is an intrinsic mode function of type $(M, N)$ if $\alpha(x)$ and $\phi(x)$ satisfy

$$\alpha(x) \in C^\infty, \quad |\nabla \alpha| \leq M, \quad 1/M \leq \alpha \leq M$$

$$\phi(x) \in C^\infty, \quad 1/M \leq |\nabla \phi| \leq M, \quad |\nabla^2 \phi| \leq M.$$

Definition 2.2. A function $f(x)$ is a well-separated superposition of type $(M, N, K)$ if

$$f(x) = \sum_{k=1}^{K} f_k(x)$$

where each $f_k(x) = \alpha_k(x)e^{2\pi iN\phi_k(x)}$ is an intrinsic mode function of type $(M, N)$ and they satisfy the separation condition

$$|N\nabla \phi_k(b) - N\nabla \phi_l(b)| \leq 2^{1+s}(|N\nabla \phi_k(b)|^s + |N\nabla \phi_l(b)|^s)$$
We claim that when \( N \in 1 \) for any \( 1 \leq k \leq K \). We denote by \( F(M, N, K) \) the set of all such functions.

Let us recall that \( W_f(p, b) \) is the wave packet transform with geometric scaling parameter \( s \in (1/2, 1) \) of a function \( f(x) \) and \( v_f(p, b) \) is the instantaneous wavevector estimation. The following theorem is our main theoretical result for the synchrosqueezed wave packet transform.

**Theorem 2.3.** For a function \( f(x) \) and \( \varepsilon > 0 \), we define

\[
R_{f, \varepsilon} = \{(p, b) : |W_f(p, b)| \geq |p|^{-s} \sqrt{\varepsilon}\}
\]

and

\[
Z_{f, k} = \{(p, b) : |p - N\nabla \phi_k(b)| \leq |p|^s\}
\]

for \( 1 \leq k \leq K \). For fixed \( M \) and \( K \), there exists a constant \( \varepsilon_0(M, K) > 0 \) such that for any \( \varepsilon \in (0, \varepsilon_0) \) there exists a constant \( N_0(M, K, \varepsilon) > 0 \) such that for any \( N > N_0(M, K, \varepsilon) \) and \( f(x) \in F(M, N, K) \) the following statements hold.

(i) \( \{Z_{f, k} : 1 \leq k \leq K\} \) are disjoint and \( R_{f, \varepsilon} \subset \bigcup_{1 \leq k \leq K} Z_{f, k} \);

(ii) For any \( (p, b) \in R_{f, \varepsilon} \cap Z_{f, k} \),

\[
|v_f(p, b) - N\nabla \phi_k(b)| \leq |N\nabla \phi_k(b)| \lesssim \sqrt{\varepsilon}.
\]

In what follows, when we write \( O(\cdot), \lesssim, \text{ or } \gtrsim \), the implicit constants may depend on \( M \) and \( K \). The proof of the theorem relies on several lemmas. The following one estimates \( W_f(p, b) \).

**Lemma 2.4.** Under the assumption of the theorem, we have

\[
W_f(p, b) = \begin{cases} |p|^{-s}O(\varepsilon), & |p| \notin \left[ \frac{N}{2M}, 2MN \right] \\ |p|^{-s} \left( \sum_{k=1}^{K} \alpha_k(b) e^{2\pi i N \phi_k(b)} \right) \hat{\omega} (|p|^{-s} (p - N\nabla \phi_k(b))) + O(\varepsilon) \end{cases}, \quad |p| \in \left[ \frac{N}{2M}, 2MN \right].
\]

\[
(5)
\]

**Proof.** Let us first estimate \( W_f(p, b) \) assuming that \( f(x) \) contains a single intrinsic mode function of type \( (M, N) \)

\[
f(x) = \alpha(x) e^{2\pi i N \phi(x)}.
\]

Using the definition of the wave packet transform, we have the following expression for \( W_f(p, b) \).

\[
W_f(p, b) = \int \alpha(x) e^{2\pi i N \phi(x)} |p|^s w(|p|^s (x - b)) e^{-2\pi i (x-b) \cdot p} dx
\]

\[
= \int \alpha(b + |p|^{-s} y) e^{2\pi i N \phi(b + |p|^{-s} y)} |p|^s w(y) e^{-2\pi i p \cdot |p|^{-s} y} dy
\]

\[
= |p|^{-s} \int \alpha(b + |p|^{-s} y) w(y) e^{2\pi i (N \phi(b + |p|^{-s} y) - |p|^{-s} y \cdot p)} dy.
\]

We claim that when \( N \) is sufficiently large

\[
W_f(p, b) = \begin{cases} |p|^{-s}O(\varepsilon), & |p| \notin \left[ \frac{N}{2M}, 2MN \right] \\ |p|^{-s} \left( \alpha(b) e^{2\pi i N \phi(b)} \right) \hat{\omega} (|p|^{-s} (p - N\nabla \phi(b))) + O(\varepsilon) \end{cases}, \quad |p| \in \left[ \frac{N}{2M}, 2MN \right].
\]

\[
(6)
\]
First, let us consider the case \(|p| \notin [\frac{N}{2M}, 2MN]\). Consider the integral of form
\[
\int h(y)e^{i g(y)} dy
\]
for real smooth functions \(h(y)\) and \(g(y)\), along with the differential operator
\[
L = \frac{1}{i} \frac{\langle \nabla g, \nabla \rangle}{|\nabla g|^2}.
\]
If \(|\nabla g|\) does not vanish, we have
\[
Le^{ig} = \frac{\langle \nabla g, i \nabla e^{ig} \rangle}{i|\nabla g|^2} = e^{ig}.
\]
Assuming that \(h(y)\) decays sufficiently fast at infinity, we perform integration by parts \(r\) times to get
\[
\int h e^{ig} dy = \int h(L^r e^{ig}) dy = \int ((L^*)^r h) e^{ig} dy,
\]
where \(L^*\) is the adjoint of \(L\). In the current setting, \(W_f(p, b) = |p|^{-s} \int h(y)e^{i g(y)} dy\) with
\[
h(y) = \alpha(b + |p|^{-s} y) w(y), \quad g(y) = 2\pi (N\phi(b + |p|^{-s} y) - |p|^{-s} y \cdot p),
\]
where \(h(y)\) clearly decays rapidly at infinity since \(w(y)\) is in the Schwartz class. In order to understand the impact of \(L\) and \(L^*\), we need to bound the norm of
\[
\nabla g(y) = 2\pi \left( N \nabla \phi(b + |p|^{-s} y) - p \right) |p|^{-s}
\]
from below when \(|p| \notin [\frac{N}{2M}, 2MN]\). If \(|p| < \frac{N}{2M}\), then
\[
|\nabla g| \gtrsim (|N\nabla \phi| - |p|)|p|^{-s} \gtrsim |N\nabla \phi||p|^{-s}/2 \gtrsim N^{1-s}.
\]
If \(|p| > 2MN\), then
\[
|\nabla g| \gtrsim (|p| - |N\nabla \phi|)|p|^{-s} \gtrsim |p| \cdot |p|^{-s}/2 \gtrsim (|p|)^{1-s} \gtrsim N^{1-s}.
\]
Hence \(|\nabla g| \gtrsim N^{1-s}\) if \(|p| \notin [\frac{N}{2M}, 2MN]\). Since \(|\nabla g| \neq 0\) and each \(L^*\) contributes a factor of order \(1/|\nabla g|\)
\[
\left| \int e^{ig(y)} ((L^*)^r h)(y) dy \right| \lesssim N^{-1-s} r.
\]
When \(N \gtrsim \varepsilon^{-1/(1-s)r}\), we obtain
\[
\left| \int e^{ig(y)} ((L^*)^r h)(y) dy \right| \lesssim \varepsilon.
\]
Using the fact \(W_f(p, b) = |p|^{-s} \int h(y)e^{i g(y)} dy\), we have \(|W_f(p, b)| \lesssim |p|^{-s} \varepsilon\).

Second, let us address the case \(|p| \in [\frac{N}{2M}, 2MN]\). We want to approximate \(W_f(p, b)\) with
\[
|p|^{-s} \alpha(b) e^{2\pi i N \phi(x)} \hat{w} \left( |p|^{-s} (p - N \nabla \phi(b)) \right).
\]
Since \( w(y) \) is in the Schwartz class, we can assume that \( |w(y)| \leq \frac{C_m}{|y|^{m}} \) for some sufficient large \( m \) with \( C_m \) for \( |y| \geq 1 \). Therefore, the integration over \( |y| \gtrsim \varepsilon^{-1/m} \) results a contribution at most of order \( O(\varepsilon) \). We can then estimate

\[
|W_f(p,b)| = |p|^{-s} \left( \int_{|y| \leq \varepsilon^{-1/m}} \alpha(b + |p|^{-s}y)w(y)e^{2\pi i(N\phi(b) + |p|^{-s}y - |p|^{-s}y \cdot p)}dy + O(\varepsilon) \right)
\]

A Taylor expansion of \( \alpha(x) \) and \( \phi(x) \) shows that

\[
\alpha(b + |p|^{-s}y) = \alpha(b) + \nabla\alpha(b^*) \cdot |p|^{-s}y
\]

and

\[
\phi(b + |p|^{-s}y) = \phi(b) + \nabla\phi(b) \cdot (|p|^{-s}y) + \frac{1}{2}(|p|^{-s}y)^t \nabla^2\phi(b^*)(|p|^{-s}y),
\]

where in each case \( b^* \) is a point between \( b \) and \( b + |p|^{-s}y \). We want to drop the last term from the above formulas without introducing a relative error larger than \( O(\varepsilon) \). We begin with the estimate

\[
\int_{|y| \leq \varepsilon^{-1/m}} |\nabla\alpha \cdot |p|^{-s}y w(y)|dy \lesssim \varepsilon,
\]

which holds if \( \varepsilon^{-2/m} |\nabla\alpha \cdot |p|^{-s}y| \lesssim \varepsilon \). This is true when \( |p|^{-s} \lesssim \varepsilon^{1+3/m} \). Since \( |p| \in [\frac{N}{2M}, 2MN] \), the above holds if

\[
N \gtrsim \varepsilon^{-(1+3/m)/s}.
\]

We also need

\[
\int_{|y| \leq \varepsilon^{-1/m}} |\alpha(b)w(y) e^{2\pi i(N\phi(b) + N\nabla\phi(b) \cdot |p|^{-s}y - |p|^{-s}y \cdot p)} | \cdot |e^{2\pi iN/2(|p|^{-s}y)^t \nabla^2\phi(|p|^{-s}y) - 1}|dy \lesssim \varepsilon.
\]

Since \( |e^{ix} - 1| \leq |x| \), the above inequality is equivalent to

\[
\int_{|y| \leq \varepsilon^{-1/m}} \alpha(b)w(y)e^{2\pi i(N\phi(b) + N\nabla\phi(b) \cdot |p|^{-s}y - |p|^{-s}y \cdot p)} |2\pi N/2(|p|^{-s}y)^t \nabla^2\phi(|p|^{-s}y)|dy \lesssim \varepsilon,
\]

which is true if \( \varepsilon^{-2/m} N(|p|^{-s}y)^t \nabla^2\phi(|p|^{-s}y) \lesssim \varepsilon \), which in turn holds if \( N|p|^{-2s} |y|^2 \lesssim \varepsilon^{1+2/m} \). Because \( |y| \lesssim \varepsilon^{-\frac{1}{M}} \) and \( |p| \in [\frac{N}{2M}, 2MN] \), then the above inequality is true when

\[
N \gtrsim \varepsilon^{-(1+4/m)(1-2s)}.
\]

In summary, when \( |p| \in [\frac{N}{2M}, 2MN] \) we have

\[
W_f(p,b) = |p|^{-s} \left( \int_{|y| \leq \varepsilon^{-1/m}} \alpha(b)w(y)e^{2\pi i(N\phi(b) + N\nabla\phi(b) \cdot |p|^{-s}y - |p|^{-s}y \cdot p)}dy + O(\varepsilon) \right)
\]

\[
= |p|^{-s} \left( \int_{|y| \leq \varepsilon^{-1/m}} \left( \alpha(b)e^{2\pi iN\phi(b)} \right) w(y)e^{2\pi i(N\nabla\phi(b) - p) \cdot |p|^{-s}y}dy + O(\varepsilon) \right)
\]

\[
= |p|^{-s} \left( \int_{\mathbb{R}^2} \left( \alpha(b)e^{2\pi iN\phi(b)} \right) w(y)e^{2\pi i(N\nabla\phi(b) - p) \cdot |p|^{-s}y}dy + O(\varepsilon) \right)
\]

\[
= |p|^{-s} \left( \alpha(b)e^{2\pi iN\phi(b)} \hat{w} \left( |p|^{-s}(p - N\nabla\phi(b)) \right) + O(\varepsilon) \right).
\]
where the third line uses the fact that the integration of \( w(y) \) outside the set \( \{ y : |y| \lesssim \varepsilon^{-1/m} \} \) is again of order \( O(\varepsilon) \).

Now let us return to the general cases, where \( f(x) \) is a superposition of \( K \) well-separated intrinsic mode components:

\[
f(x) = \sum_{k=1}^{K} f_k(x) = \sum_{k=1}^{K} \alpha_k(x) e^{2\pi i N \phi_k(x)}
\]

in the Definition 2.2. By linearity of the wave packet transform and (6), we get

\[
W_f(p, b) = \begin{cases}
|p|^{-s} O(\varepsilon), & |p| \notin \left[ \frac{N}{2M}, 2MN \right] \\
|p|^{-s} \left( \sum_{k=1}^{K} \alpha_k(b) e^{2\pi i N \phi_k(b)} \hat{w}(|p|^{-s}(p - N\nabla \phi_k(b))) + O(\varepsilon) \right), & |p| \in \left[ \frac{N}{2M}, 2MN \right].
\end{cases}
\]

(7)

The next lemma estimates \( \nabla_b W_f(p, b) \) when \( |p| \in \left[ \frac{N}{2M}, 2MN \right] \), i.e., the case \( W_f(p, b) \) is non-negligible.

**Lemma 2.5.** Under the assumption of the theorem, we have

\[
\nabla_b W_f(p, b) = 2\pi i N |p|^{-s} \left( \sum_{k=1}^{K} \nabla \phi_k(b) \alpha_k(b) e^{2\pi i N \phi_k(b)} \hat{w}(|p|^{-s}(p - N\nabla \phi_k(b))) + O(\varepsilon) \right)
\]

(8)

when \( |p| \in \left[ \frac{N}{2M}, 2MN \right] \).

**Proof.** The proof is similar to the one of Lemma 2.4. Assume that \( f(x) \) contains a single intrinsic mode function, i.e.,

\[
f(x) = \alpha(x) e^{2\pi i N \phi(x)},
\]

we have

\[
\nabla_b W_f(p, b) = \int_{\mathbb{R}^2} \alpha(x) e^{2\pi i N \phi(x)} |p|^s (\nabla w(|p|^s(x - b))(-|p|^s) + 2\pi ipw(|p|^s(x - b))) e^{-2\pi i(x-b) \cdot p} dx
\]

\[
= \int_{\mathbb{R}^2} \alpha(b + |p|^{-s} y) e^{2\pi i N \phi(b + |p|^{-s} y)} |p|^{-s} (\nabla w(y)(-|p|^s) + 2\pi ipw(y)) e^{-2\pi i|p|^{-s} y \cdot p} dy
\]

\[
= \int_{\mathbb{R}^2} \alpha(b + |p|^{-s} y) e^{2\pi i N \phi(b + |p|^{-s} y)} |p|^{-s} \nabla w(y)(-|p|^s) e^{-2\pi i|p|^{-s} y \cdot p} dy
\]

\[
+ \int_{\mathbb{R}^2} \alpha(b + |p|^{-s} y) e^{2\pi i N \phi(b + |p|^{-s} y)} |p|^{-s} 2\pi ipw(y) e^{-2\pi i|p|^{-s} y \cdot p} dy.
\]

Forming a Taylor expansion and following the same argument in the proof of Lemma 2.4 gives the following approximation for \( |p| \in \left[ \frac{N}{2M}, 2MN \right] \)

\[
\nabla_b W_f(p, b) = \left( -2\pi i |p|^{-s}(p - N\nabla \phi(b)) \alpha(b) e^{2\pi i N \phi(b)} \hat{w}(|p|^{-s}(p - N\nabla \phi(b))) + O(\varepsilon) \right)
\]

\[
+ 2\pi i |p|^{-s} \left( \alpha(b) e^{2\pi i N \phi(b)} \hat{w}(|p|^{-s}(p - N\nabla \phi(b))) + O(\varepsilon) \right)
\]

\[
= 2\pi i N |p|^{-s} \left( \nabla \phi(b) \alpha(b) e^{2\pi i N \phi(b)} \hat{w}(|p|^{-s}(p - N\nabla \phi(b))) + O(\varepsilon) \right).
\]
For \( f(x) = \sum_{k=1}^{K} f_k(x) = \sum_{k=1}^{K} \alpha_k(x) e^{2\pi i N \phi_k(x)} \), taking sum over \( K \) terms gives

\[
\nabla_b W_f(p, b) = 2\pi i N |p|^{-s} \left( \sum_{k=1}^{K} \left( \nabla \phi_k(b) \alpha_k(x) e^{2\pi i N \phi_k(b)} \tilde{w}(|p|^{-s}(p - N \nabla \phi_k(b))) \right) + O(\varepsilon) \right)
\]

for \( |p| \in \left[ \frac{N}{2M}, 2MN \right] \).

We are now ready to prove the theorem.

**Proof.** Let us first consider \((i)\). Suppose there exists \((p, b) \in Z_{f,k} \cap Z_{f,l} \) with \( k \neq l \). Then

\[
|p - N \nabla \phi_k(b)| \leq |p|^s, \quad |p - N \nabla \phi_l(b)| \leq |p|^s,
\]

which implies

\[
|p| \leq |p|^s + |N \nabla \phi_k(b)|, \quad |p| \leq |p|^s + |N \nabla \phi_l(b)|.
\]

Since \( s < 1 \),

\[
|p| \leq 2 |N \nabla \phi_k(b)|, \quad |p| \leq 2 |N \nabla \phi_l(b)|.
\]

Therefore,

\[
|N \nabla \phi_k(b) - N \nabla \phi_l(b)| \leq |p - N \nabla \phi_k(b)| + |p - N \nabla \phi_l(b)| \leq 2 |p|^s \leq 2 \cdot 2^s (|N \nabla \phi_k(b)|^s + |N \nabla \phi_l(b)|^s),
\]

which contradicts the separation assumption. Thus, all \( Z_{f,k} \) are disjoint. Let \((p, b)\) be a point in \( R_{f,\varepsilon} = \{ (p, b) : W_f(p, b) \geq |p|^{-s} \sqrt{\varepsilon} \} \). From the above lemma, we see that for \( \varepsilon \) sufficiently small, if \( |p| \in \left[ \frac{N}{2M}, 2MN \right] \), we have

\[
W_f(p, b) = |p|^{-s} \left( \sum_{k=1}^{K} \alpha_k(b) e^{2\pi i N \phi_k(b)} \tilde{w} \left( |p|^{-s}(p - N \nabla \phi_k(b)) \right) + O(\varepsilon) \right).
\]

Therefore, there exists \( k \) between 1 and \( K \) such that \( \tilde{w} \left( |p|^{-s}(p - N \nabla \phi_k(b)) \right) \) is non-zero. From the definition of \( \tilde{w}(\xi) \), we see that this implies \( (p, b) \in Z_{f,k} \). Hence \( R_{f,\varepsilon} \subset \bigcup_{k=1}^{K} Z_{f,k} \).

To show \((ii)\), let us recall that \( v_f(p, b) \) is defined as

\[
v_f(p, b) = \frac{\nabla_b W_f(p, b)}{2\pi i W_f(p, b)}
\]

for \( W_f(p, b) \neq 0 \). If \((p, b) \in R_{f,\varepsilon} \cap Z_{f,k} \), then

\[
W_f(p, b) = |p|^{-s} \left( \alpha_k(b) e^{2\pi i N \phi_k(b)} \tilde{w} \left( |p|^{-s}(p - N \nabla \phi_k(b)) \right) + O(\varepsilon) \right)
\]

and

\[
\nabla_b W_f(p, b) = 2\pi i N |p|^{-s} \left( \nabla \phi_k(b) \alpha_k(b) e^{2\pi i N \phi_k(b)} \tilde{w}(|p|^{-s}(p - N \nabla \phi_k(b))\right) + O(\varepsilon)
\]

as the other terms drop out since \( \{ Z_{f,k} \} \) are disjoint. Hence

\[
v_f(p, b) = \frac{N \nabla \phi_k(b) \alpha_k(b) e^{2\pi i N \phi_k(b)} \tilde{w} \left( |p|^{-s}(p - N \nabla \phi_k(b)) \right) + O(\varepsilon))}{\alpha_k(b) e^{2\pi i N \phi_k(b)} \tilde{w} \left( |p|^{-s}(p - N \nabla \phi_k(b)) \right) + O(\varepsilon))}.
\]
Let us denote the term \( \alpha_k(b) e^{2\pi i N \phi_k(b)} \hat{w} (|p|^{-s}(p - N \nabla \phi_k(b))) \) by \( g \). Then
\[
v_f(p, b) = \frac{N \nabla \phi_k(b) (g + O(\varepsilon))}{g + O(\varepsilon)}.
\]
Since \(|W_f(p, b)| \geq |p|^{-s} \sqrt{\varepsilon} \) for \((p, b) \in R_{f, \varepsilon}, |g| \gtrsim \sqrt{\varepsilon} \). Therefore
\[
\left| \frac{v_f(p, b) - N \nabla \phi_k(b)}{|N \nabla \phi_k(b)|} \right| \lesssim \left| \frac{O(\varepsilon)}{g + O(\varepsilon)} \right| \lesssim \sqrt{\varepsilon}.
\]

The assumption \( s \in (1/2, 1) \) is essential to the proof. The upper bound \( s < 1 \) allows the wave packets to detect oscillations in different directions. The lower bound \( s > 1/2 \) ensures that the support of the wave packets is sufficiently small in space so that the second order properties of the phase function (such as the curvature of the wave front) do not affect the synchrosqueezing estimate of the instantaneous wavevectors.

### 3 Implementation of the transform

In this section, we describe in detail the discrete synchrosqueezed wave packet transform. Let us first recall the continuous setting. For a given superposition \( f(x) \) of several well-separated components, the synchrosqueezed wave packet transform consists of the following steps:

(i) Apply the wave packet transform to obtain \( W_f(p, b) \) and the gradient \( \nabla_b W_f(p, b) \);

(ii) Compute the approximate instantaneous wavevector \( v_f(p, b) \) and perform synchrosqueezing to get \( T_f(v, b) \);

(iii) Use a clustering algorithm to identify the support of the new phase space representation \( T_f(v, b) \) of different intrinsic mode functions;

(iv) Reconstruct each intrinsic mode function using the dual frame.

In order to realize these steps in the discrete setting, we first introduce a discrete implementation of the wave packet transform in Section 3.1. The full discrete algorithm will then be discussed in Section 3.2.

#### 3.1 Discrete wave packet transform

For simplicity, we consider functions that are periodic over the unit square \([0, 1)^2\) in 2D. Let
\[
X = \{ (n_1/L, n_2/L) : 0 \leq n_1, n_2 < L, n_1, n_2 \in \mathbb{Z} \}
\]
be the \( L \times L \) spatial grid at which these functions are sampled. The corresponding \( L \times L \) Fourier grid is
\[
\Xi = \{ (\xi_1, \xi_2) : -L/2 \leq \xi_1, \xi_2 < L/2, \xi_1, \xi_2 \in \mathbb{Z} \}.
\]
For a function \( f(x) \in \ell^2(X) \), the discrete forward Fourier transform is defined by
\[
\hat{f}(\xi) = \frac{1}{L} \sum_{x \in X} e^{-2\pi i x \cdot \xi} f(x).
\]
For a function \( g(\xi) \in \ell^2(\Xi) \), the discrete inverse Fourier transform is

\[
\hat{g}(x) = \frac{1}{L} \sum_{\xi \in \Xi} e^{2\pi i x \cdot \xi} g(\xi).
\]

In both transforms, the factor \( 1/L \) ensures that these discrete transforms are isometries between \( \ell_2(X) \) and \( \ell_2(\Xi) \).

In order to design a discrete wave packet transform, we need to specify how to decimate the momentum space \( p \) and the position space \( b \). Let us first consider the momentum space \( p \). In the continuous setting, the Fourier transform \( \hat{w}_{pb}(\xi) \) of the wave packets for a fixed \( p \) value have the profile

\[
|p|^{-s} \hat{w}(|p|^{-s}(\xi - p)),
\]

modulo complex modulation. In the discrete setting, we sample the Fourier domain \([-L/2, L/2]^2\) with a set of points \( P \) and associate with each \( p \in P \) a window function \( g_p(\xi) \) that behaves qualitatively as \( |p|^{-s} \hat{w}(|p|^{-s}(\xi - p)) \). More precisely, \( g_p(\xi) \) is required to satisfy the following conditions:

- \( g_p(\xi) \) is non-negative and centered at \( p \) with a compact support of width \( L_p = O(|p|^s) \);
- \( g_p(|p|^s \tau + p) \) is a sufficiently smooth function of \( \tau \), thus making the discrete wave packets decay rapidly in the spatial domain;
- \( C_1 \leq \int |g_p(|p|^s \tau + p)|^2 d\tau \leq C_2 \) for positive constants \( C_1 \) and \( C_2 \), independent of \( p \);
- In addition, for any \( \xi \in [-L/2, L/2]^2 \), \( \sum_{p \in P} |g_p(\xi)|^2 = 1 \).

One possible way to specify the set \( P \) and the functions \( \{g_p(\xi), p \in P\} \) is to follow the constructions of the the wave atom frame in [8] or the Gaussian wave packets of [15]. In both constructions, the parabolic scaling \( s = 1/2 \) is used in order to represent the oscillatory patterns efficiently. However, in the current setting, the proposed wave packet transform requires \( s \in (1/2, 1) \) and hence one needs to increase the support of \( g_p(\xi) \) accordingly. We refer to [8, 15] for more complete discussions. The above conditions of \( g_p(\xi), p \in P \) also impose a constraint on the sampling density of the set \( P \). In the frequency plane, the set \( P \) becomes dense near the origin and sparser for large \( \xi \). A straightforward calculation shows that the total number of samples in \( P \) is of order \( O(L^{2 - 2s}) \).

The decimation of the position space \( b \) is much easier; we simply discretize it with an \( L_B \times L_B \) uniform grid as follows:

\[
B = \{(n_1/L_B, n_2/L_B) : 0 \leq n_1, n_2 < L_B, n_1, n_2 \in \mathbb{Z}\}.
\]

As we shall see, the only requirement is that \( L_B \geq \max_{p \in P} L_p \) so that the discrete wave atoms can form a frame.

For each fixed \( p \in P \) and \( b \in B \), the discrete wave packet, still denoted by \( w_{pb}(x) \) without causing much confusion, is defined through its Fourier transform as

\[
\hat{w}_{pb}(\xi) = \frac{1}{L_p} e^{-2\pi ib \cdot \xi} g_p(\xi)
\]

for \( \xi \in \Xi \). Since \( g_p(\xi) \) is centered at \( p \) and has a support of width \( L_p = O(|p|^s) \), this function fits into the scaling of the wave packet. Applying the discrete inverse Fourier transform provides its spatial description

\[
w_{pb}(x) = \frac{1}{L \cdot L_p} \sum_{\xi \in \Xi} e^{2\pi i (x - b) \cdot \xi} g_p(\xi).
\]
For a function $f(x)$ defined on $x \in X$, the discrete wave packet transform is a map from $\ell_2(X)$ to $\ell_2(P \times B)$, defined by

$$W_f(p,b) = \langle w_{pb}, f \rangle = \frac{1}{L_p} \sum_{\xi \in \Xi} e^{2\pi ib \cdot \xi} g_p(\xi) \hat{f}(\xi). \quad (10)$$

We can introduce an inner product on the space $\ell_2(P \times B)$ as follows: for any two functions $g(p, b)$ and $h(p, b)$,

$$\langle g, h \rangle = \sum_{p \in P, b \in B} \overline{g(p, b)} h(p, b) (L_p / L_B)^2.$$

The following result shows that $\{w_{pb}, (p, b) \in P \times B\}$ forms a tight frame when equipped with this inner product.

**Proposition 3.1.** For any function $f(x)$ for $x \in X$, we have

$$\sum_{p \in P, b \in B} |W_f(p,b)|^2 (L_p / L_B)^2 = \|f\|_2^2.$$

**Proof.** From the definition of the wave packet transform, we have

$$\sum_{p \in P, b \in B} |W_f(p,b)|^2 (L_p / L_B)^2 = \sum_{p \in P, b \in B} \left| \sum_{\xi \in \Xi} \frac{1}{L_p} e^{2\pi ib \cdot \xi} g_p(\xi) \hat{f}(\xi) \right|^2 (L_p / L_B)^2$$

$$= \sum_{p \in P, b \in B} \left| \sum_{\xi \in \Xi} \frac{1}{L_B} e^{2\pi ib \cdot \xi} g_p(\xi) \hat{f}(\xi) \right|^2$$

$$= \sum_{p \in P} \sum_{\xi \in \Xi} \left| g_p(\xi) \hat{f}(\xi) \right|^2$$

$$= \sum_{\xi \in \Xi} |\hat{f}(\xi)|^2.$$

$\square$

For a function $h(p, b)$ in $\ell_2(P \times B)$, the transpose of the wave packet transform is given by

$$W_t^h(x) := \sum_{p \in P, b \in B} h(p, b) w_{pb}(x) (L_p / L_B)^2. \quad (11)$$

The next result shows that this transpose operator allows us to reconstruct $f(x), x \in X$ from its wave packet transform $W_f(p, b), (p, b) \in P \times B$.

**Proposition 3.2.** For any function $f(x)$ with $x \in X$,

$$f(x) = \sum_{p \in P, b \in B} W_f(p,b) w_{pb}(x) (L_p / L_B)^2.$$
Proof. Let us consider the Fourier transform of the right hand side. It is equal to

\[
\sum_{p \in P, b \in B} \left( \sum_{\eta \in \Xi} \frac{1}{L_p} e^{2\pi i b \cdot \eta} g_p(\eta) \hat{f}(\eta) \right) \cdot \frac{1}{L_p} e^{-2\pi i b \cdot \xi} g_p(\xi) \left( L_p / L_B \right)^2
\]

\[
= \sum_{p \in P} \left( \sum_{\eta \in \Xi} \frac{1}{L_B^2} \left( \sum_{b \in B} e^{2\pi i b \cdot (\eta - \xi)} g_p(\eta) \hat{f}(\eta) \right) \right) g_p(\xi)
\]

\[
= \sum_{p \in P} (g_p(\xi))^2 \hat{f}(\xi) = \hat{f}(\xi),
\]

where the second step uses the fact that in the \( \eta \) sum only the term with \( \eta = \xi \) yields a non-zero contribution. \( \square \)

Let us now turn to the discrete approximation of \( \nabla_b W_f(p, b) \). From the continuous definition (1), we have

\[
\nabla_b W_f(p, b) = \nabla_b \langle \hat{w}_{pb}, \hat{f} \rangle = \langle -2\pi i \xi \hat{w}_{pb}(\xi), \hat{f}(\xi) \rangle.
\]

Therefore, we define the discrete gradient \( \nabla_b W_f(p, b) \) in a similar way

\[
\nabla_b W_f(p, b) = \sum_{\xi \in \Xi} \frac{1}{L_p} 2\pi i \xi e^{2\pi i b \cdot \xi} g_p(\xi) \hat{f}(\xi).
\]

(12)

The above definitions give rise to fast algorithms for computing the forward wave packet transform, its transpose, and the discrete gradient operator. All three algorithms heavily rely on the fast Fourier transform (FFT). For the forward transform, writing (10) as

\[
W_f(p, b) = \frac{L_B}{L_p} \cdot \left( \frac{1}{L_B} \sum_{\xi \in \Xi} e^{2\pi i b \cdot \xi} g_p(\xi) \hat{f}(\xi) \right)
\]

suggests the following algorithm

**Algorithm 3.3.** Forward transform from \( f(x) \) to \( W_f(p, b) \)

1: Compute \( \hat{f}(\xi) \) with \( \xi \in \Xi \) from \( f(x) \) with \( x \in X \) using an \( L \times L \) forward FFT.
2: for each \( p \in P \) do
3: Form \( g_p(\xi) \hat{f}(\xi) \) on the support of \( g_p(\xi) \)
4: Wrap the result modulo \( L_B \) onto the domain \([-L_B/2, L_B/2)^2\]
5: Apply an \( L_B \times L_B \) inverse FFT to the wrapped result
6: Multiple the result by \( L_B / L_p \) to get \( W_f(p, b) \) for all \( b \in B \)
7: end for

The transpose operator (11) can be written equivalently in the Fourier domain as

\[
\hat{W}_h^t(\xi) = \sum_{p \in P, b \in B} h(p, b) \frac{1}{L_p} e^{-2\pi i b \cdot \xi} g_p(\xi) \left( L_p / L_B \right)^2 = \sum_{p \in P} \left( \sum_{b \in B} \frac{1}{L_B} h(p, b) \frac{L_p}{L_B} e^{-2\pi i b \cdot \xi} \right) g_p(\xi).
\]

This suggests the following algorithm for the transpose operator

**Algorithm 3.4.** Transpose operator from \( h(p, b) \) to \( \hat{W}_h^t(x) \)
1: for each $p \in P$ do
2: Multiply $h(p, b)$ for each $b \in B$ by $L_p / L_B$
3: Apply an $L_B \times L_B$ forward FFT to the product
4: Unwrap the result modulo $L_B$ onto the support of $g_p(\xi)$
5: Multiply the unwrapped data with $g_p(\xi)$ and add the product to get $\hat{f}(\xi)$
6: end for
7: Compute $f(x)$ with $x \in X$ from $\hat{f}(\xi)$ with $\xi \in \Xi$ using an $L \times L$ inverse FFT.

To implement the discrete gradient operator, we rewrite (12) as

$$\nabla_b W_f(p, b) = \frac{L_B}{L_p} \left( \sum_{\xi \in \Xi} \frac{1}{L_B} e^{2\pi i b \cdot \xi} 2\pi i g_p(\xi) \hat{f}(\xi) \right).$$

This suggests the following algorithm

Algorithm 3.5. Discrete gradient operator from $f(x)$ to $\nabla_b W_f(p, b)$
1: Compute $\hat{f}(\xi)$ with $\xi \in \Xi$ from $f(x)$ with $x \in X$ using an $L \times L$ forward FFT.
2: for each $p \in P$ do
3: Form $2\pi i g_p(\xi) \hat{f}(\xi)$ on the support of $g_p(\xi)$
4: Wrap the result modulo $L_B$ onto the domain $[-L_B/2, L_B/2]^2$
5: Apply an $L_B \times L_B$ inverse FFT to each component of the wrapped result
6: Multiple the result by $L_B / L_p$ to get $\nabla_b W_f(p, b)$ for all $b \in B$
7: end for

As we mentioned earlier, the conditions on $\{g_p(\xi), p \in P\}$ imply that there are $O(L^{2(1-s)})$ samples in set $P$. A straightforward calculation shows that the computational cost of all three algorithms is $O(L^2 \log L + L^{2(1-s)}L_B^2 \log L_B)$ with $L_B \geq \max_{p \in P} L_p = O(L^s)$. If we choose $L_B$ to be of the same order as $L^s$, the complexity of these algorithms is $O(L^2 \log L)$, which is the cost of an FFT on an $L \times L$ Cartesian grid.

3.2 Description of the full algorithm

With the discrete transforms and their fast algorithms available, we now go through the steps of the synchrosqueezed wave packet transform.

For a given function $f(x)$ defined on $x \in X$, we apply Algorithm 3.3 to compute $W_f(p, b)$ and Algorithm 3.5 to compute $\nabla_b W_f(p, b)$. The approximate instantaneous wavevector $v_f(p, b)$ is then estimated by

$$v_f(p, b) = \frac{\nabla_b W_f(p, b)}{2\pi i W_f(p, b)}$$

for $p \in P, b \in B$ with $W_f(p, b) \neq 0$.

To specify the synchrosqueezed energy distribution $T_f(v, b)$, we first place in the Fourier domain a two dimensional Cartesian grid of stepsize $\Delta$:

$$V = \{(n_1 \Delta, n_2 \Delta) : n_1, n_2 \in \mathbb{Z}\}.$$

At each $v = (n_1 \Delta, n_2 \Delta) \in V$, we associate a cell $D_v$ centered at $v$

$$D_v = \left[ \left(n_1 - \frac{1}{2}\right) \Delta, (n_1 + \frac{1}{2}) \Delta \right] \times \left[ \left(n_2 - \frac{1}{2}\right) \Delta, (n_2 + \frac{1}{2}) \Delta \right].$$
Then the discrete synchrosqueezed energy distribution is defined as
\[ T_f(v, b) = \sum_{(p, b) : v_f(p, b) \in D_v} |W_f(p, b)|^2 \cdot \left( \frac{L_p}{L_B} \right)^2. \]

It is straightforward to check that
\[ \sum_{v \in V, b \in B} T_f(v, b) = \sum_{p \in P, b \in B} |W_f(p, b)|^2 \left( \frac{L_p}{L_B} \right)^2 = \| f \|^2_2. \]

Suppose that \( f(x) \) is a superposition of \( K \) well-separated intrinsic mode functions:
\[ f(x) = \sum_{k=1}^{K} f_k(x) = \sum_{k=1}^{K} \alpha_k(x) e^{2\pi i N \phi_k(x)}. \]

From the previous discussion, we know that, for each \( b \in B, v_f(p, b) \) points approximately to one of \( N \nabla \phi_k(b) \), depending on \( p \). Therefore, after synchrosqueezing, \( T_f(v, b) \) is essentially supported in the phase space near the \( K \) “discrete” surfaces \( \{ (N \phi_k(b), b), b \in B \} \). The next step is to decompose the essential support of \( T_f(v, b) \) into \( K \) clusters, one for each intrinsic mode function, through a spectral clustering method. We make use of the algorithm proposed in [14]. For a fixed set \( S \) of \( n \) points \( \{ s_1, \ldots, s_n \} \) in \( \mathbb{R}^l \) and an integer \( K \), this method partitions these points into \( K \) clusters as follows.

**Algorithm 3.6. General spectral clustering on set \( S = s_1, \ldots, s_n \)**

1. Construct the matrix \( A = (\alpha_{ij})_{ij} \in \mathbb{R}^{n \times n} \) with distance function \( \alpha_{ij} = \exp(-|s_i - s_j|^2/\sigma^2) \) if \( i \neq j \), and \( \alpha_{ii} = 0 \), \( \forall i \). Here \( \sigma \) is an input parameter.
2. Let \( D \) be a diagonal matrix such that \( D_{ii} = \sum_{j=1}^{n} \alpha_{ij} \) and define the Laplacian-type matrix \( L = D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \).
3. Choose the \( K \) largest orthogonal eigenvectors of \( L \), say \( v_1, \ldots, v_K \), and stack them horizontally to get the matrix \( V = [v_1, v_2, \ldots, v_K] \in \mathbb{R}^{n \times K} \). The entries of \( V \) are denoted by \( v_{ij} \).
4. Define the matrix \( M = (m_{ij}) \) with \( m_{ij} = v_{ij}/(\sum_j v_{ij}^2)^{1/2} \), which means normalizing the rows of \( V \).
5. Consider each row of \( M \) as a point in \( \mathbb{R}^K \) and then partition these \( n \) points into \( K \) clusters with the \( K \)-means algorithm.
6. If row \( i \) of \( M \) is assigned to cluster \( j \), then assign the original point \( s_i \) to cluster \( j \).

In the current setting, we choose a threshold parameter \( \eta > 0 \), define the set \( S \) to be
\[ \{ (v, b) : v \in V, b \in B, T_f(v, b) \geq \eta \}, \]
and apply the above algorithm to \( S \). The resulting clusters are defined to be \( U_1, \ldots, U_K \). In many cases, the number of components \( K \) is not known a priori and needs to be discovered from the function \( T_f(v, b) \). To do that, we use the set of non-negligible entries for each \( b \), i.e., \( \{ v \in V, T_f(v, b) \geq \eta \} \), to decide the local number of clusters and take the maximum number of all \( b \in B \) to be the cluster number \( K \).

In the final step, we recover each intrinsic mode function by computing.
\[ f_k(x) = \sum_{(p, b) : v_f(p, b) \in U_k} W_f(p, b) w_{ph}(x) \left( \frac{L_p}{L_B} \right)^2. \]

This step can be carried out efficiently by restricting \( W_f(p, b) \) to the set \( \{ (p, b) : v_f(p, b) \in U_k \} \) and applying Algorithm 3.4 to the restriction for each \( k \).
4 Numerical Results

This section presents several numerical examples to illustrate the proposed synchrosqueezed wave packet transforms. Throughout all examples, the threshold value $\varepsilon$ is $10^{-4}$ and the size $L$ of the Cartesian grid $X$ of the discrete algorithm is 512. In the implementation of the discrete wave packet transforms, the scaling parameter $s$ is equal to $2/3$.

4.1 Instantaneous wavevector extraction

We first test the accuracy of the estimated instantaneous wavevector $v_f(p,b)$. Let $f(x)$ be a deformed plane wave

$$f(x) = \alpha(x)e^{2\pi i N\phi(x)}.$$ 

We have seen that for each point $b$ the estimate $v_f(p,b)$ approximates the instantaneous wavevector at $b$ if $|W_f(p,b)| \geq |p|^{-s}\sqrt{\varepsilon}$. In the discrete setting, since $L_p = O(|p|^s)$, the corresponding threshold criteria becomes $|W_f(p,b)L_p| \geq \sqrt{\varepsilon}$. In order to further stabilize the estimate, we define the mean estimated instantaneous frequency at $b$ to be

$$v^m_f(b) = \frac{\sum_p |W_f(p,b)|^2 v_f(p,b)}{\sum_p |W_f(p,b)|^2}$$

where the sum in $p$ is taken over all $p$ that satisfy $|W_f(p,b)L_p| \geq \sqrt{\varepsilon}$. Using this estimate, we can define the (discrete) relative error $R(b)$ between $v^m_f(b)$ and the exact instantaneous frequency $N\nabla\phi(b)$ as

$$R(b) = \frac{|v^m_f(b) - N\nabla\phi(b)|}{|N\nabla\phi(b)|}.$$

Example 1. We perform the above test on a deformed plane wave $f(x)$ with $\alpha(x) = 1$, $\phi(x) = \phi(x_1, x_2) = x_1 + x_2 + \beta \sin(2\pi x_1) + \beta \sin(2\pi x_2)$ with $\beta = 0.1$, and $N = 135$. The relative error $R(b)$ shown in Figure 4 is of order $10^{-2}$, which agrees with Theorem 2.3 on that the relative approximation error is of order $O(\sqrt{\varepsilon})$.

![Figure 4: Example 1. Relative error $R(b)$ of instantaneous wavevector estimation.](image)
4.2 Intrinsic mode decomposition

**Example 2.** Here $f(x)$ is a sum of two deformed plane waves

$$f(x) = e^{2\pi i N \phi_1(x)} + e^{2\pi i N \phi_2(x)}$$

$$\phi_1(x) = \phi_1(x_1, x_2) = x_1 + x_2 + \beta \sin(2\pi x_1) + \beta \sin(2\pi x_2)$$

$$\phi_2(x) = \phi_2(x_1, x_2) = -x_1 + x_2 - \beta \sin(2\pi x_1) + \beta \sin(2\pi x_2)$$

with $N = 135$ and $\beta = 0.1$. The algorithm described in Section 3.2 is applied to $f(x)$ to extract these two components. Figure 5 summarizes the results of this test. The first row shows the superposition $f(x)$ (left) and the synchrosqueezed energy distribution $T_f(v, b)$ with $b_1$ fixed at 1 (right). For a fixed $b_1$ value $T_f(v, b)$ concentrates near two curves. More generally, in phase space $T_f(v, b)$ concentrates near two 2D surfaces. The second row shows the two sets $U_1$ and $U_2$ after the clustering steps. Finally, the third row plots the two reconstructed components.

The proposed synchrosqueezed wave packet transform is also rather robust to noise. To demonstrate this, let $f(x)$ be the superposition of two deformed plane waves and a noise term

$$f(x) = e^{2\pi i N \phi_1(x)} + e^{2\pi i N \phi_2(x)} + n(x),$$

where $n(x)$ is an isotropic complex Gaussian random noise with zero mean and variance $\sigma^2 = 0.5$. In order to reduce the influence of noise, we set up a threshold parameter $\delta \approx 3\sigma^2$ and keep only the values of $T_f(v, b)$ that are greater than $\delta$. Figure 6 summarizes the results of this test. The first row shows the noisy superposition $f(x)$ (left) and the synchrosqueezed energy distribution $T_f(v, b)$ (after thresholding with $\delta$) with $b_1$ fixed at 1 (right). We observe that the support of $T_f(v, b)$ in this case is almost indistinguishable from the noiseless case, thus demonstrating the robustness of the proposed method against the noise. The second row plots the two sets $U_1$ and $U_2$ after the clustering steps. Finally, the last row gives the two reconstructed components and we see that they are very similar to the ones obtained in the noiseless case.

**Example 3.** In our analysis, we have assumed so far that each component of the superposition covers the whole domain. In many real applications, one or more components of the superposition might be incomplete, i.e., covering only part of the domain. The analysis of this general case is more complicated due to the boundary of these incomplete components. On the other hand, one expects that the synchrosqueezed energy distribution $T_f(v, b)$ should still be supported near several 2D surfaces in the phase space, though some of them might be incomplete. In this example, we show that the synchrosqueezed wave packet transform still works quite well under this more general setting. Here we choose $f(x)$ to be the superposition of two components, one of which is incomplete:

$$f(x) = \chi(x) \cdot e^{2\pi i N \phi_1(x)} + e^{2\pi i N \phi_2(x)},$$

$$\phi_1(x) = \phi_1(x_1, x_2) = -(x_1 + \beta \sin(2\pi x_1)) + (x_2 + \beta \sin(2\pi x_2)),$$

$$\phi_2(x) = \phi_2(x_1, x_2) = (x_1 + \beta \sin(2\pi x_1)) - (x_2 + \beta \sin(2\pi x_2)),$$

where $N = 135$, $\beta = 0.1$, and $\chi(x)$ is an indicator function of an ellipse in $[0,1]^2$. Figure 7 summarizes the results of this example. The first row shows the superposition $f(x)$ without noise. The second row plots the two reconstructed components. We note that the boundary of the second incomplete component is accurately captured by the proposed method.
Figure 5: Example 2. Mode decomposition without noise. Top-left: A superposition of two deformed plane waves. Top-right: Synchrosqueezed energy distribution $T_f(v, b)$ at $b_1 = 1$. Second row: The support of $T_f(v, b)$ is clustered into two subsets. Third row: the two reconstructed components.

5 Discussion

The method proposed here is an initial step of mode decomposition for higher dimensional signals. Several possible directions for future research are listed below.
Figure 6: Example 2. Mode decomposition with noise. Top-left: A superposition of two deformed plane waves. Top-right: Synchrosqueezed energy distribution $T_f(v, b)$ at $b_1 = 1$. Second row: The support of $T_f(v, b)$ are clustered into two subsets. Third row: the two reconstructed components. The noise has mostly been removed in the reconstructed components.
Figure 7: Example 3. Mode decomposition with incomplete component. Top: A superposition of a complete deformed plane wave and an incomplete one. Second row: the two reconstructed components. The sharp cutoff boundary of the second incomplete component is clearly shown.

The synchrosqueezed wave packet transform has a geometric scaling parameter $s$, which is in $(1/2, 1)$. As we mentioned earlier, the case $s = 1/2$ is the wave atom construction proposed in [8]. Wave atoms provide better angular resolution as the support of wave atoms in the Fourier domain are more refined. However, as we pointed out, the synchrosqueezing step is no longer sufficient for instantaneous wavevector estimation, as the wave atoms are large enough in space to see the second order effects of the phase function. One natural question is whether it is possible to generalize or modify the synchrosqueezing idea so that it will work for the wave atom case.

Another closely related set of analyzing functions is the curvelet frame [3, 2]. Curvelets have shown to be the optimal tool for representing images that are smooth except for at isolated curvilinear discontinuities. The current approach can be extended to the curvelets as long as the parabolic scaling case $s = 1/2$ can be addressed. A potential advantage of a synchrosqueezed curvelet transform is that it should be able to optimally identify isolated wavefronts, as opposed to the extended wave fields used in the current paper.

So far we have assumed that the instantaneous wavevector of the different intrinsic
mode functions are well-separated at each point. In practice, this assumption might not hold, as many images might have multiple nearby wavevectors at isolated points or curves. It would be useful to find more robust clustering algorithms to address such situations.

The current approach can be easily extended to 3D or higher dimensions. This direction should be relevant for applications, such as seismic imaging.

Acknowledgments. H.Y. was partially supported by NSF grant CDI-1027952. L.Y. was partially supported by NSF grants CAREER DMS-0846501, DMS-1027952, and CDI-1027952. L.Y. thanks Jianfeng Lu and Hau-Tieng Wu for discussion and Jack Poulson for comments on the manuscript.

References


