Journal of Fourier Analysis and Applications (2024) 30:62
https://doi.org/10.1007/s00041-024-10098-8

Coorbit Theory of Warped Time-Frequency Systems in $\mathbb{R}^d$

Nicki Holighaus[1] · **Felix Voigtlaender**[2]

Received: 15 September 2022 / Revised: 4 April 2024 / Accepted: 24 May 2024 /
Published online: 25 October 2024

Abstract

Warped time-frequency systems have recently been introduced as a class of structured continuous frames for functions on the real line. Herein, we generalize this framework to the setting of functions of arbitrary dimensionality. After showing that the basic properties of warped time-frequency representations carry over to higher dimensions, we determine conditions on the warping function which guarantee that the associated Gramian is well-localized, so that associated families of coorbit spaces can be constructed. We then show that discrete Banach frame decompositions for these coorbit spaces can be obtained by sampling the continuous warped time-frequency systems. In particular, this implies that sparsity of a given function f in the discrete warped time-frequency dictionary is equivalent to membership of f in the coorbit space. We put special emphasis on the case of radial warping functions, for which the relevant assumptions simplify considerably.

Keywords Time-frequency representations · Frequency-warping · Anisotropic function systems · Integral transforms · Coorbit spaces · Discretization · Sampling · Banach frames · Atomic decompositions · Mixed-norm spaces

Mathematics Subject Classification Primary 42B35 · 42C15; Secondary 46F05 · 46F12 · 94A20

Dedicated to Karlheinz Gröchenig on the occasion of his 65th birthday.

Communicated by Stephan Dahlke.

✉ Nicki Holighaus
nicki.holighaus@oeaw.ac.at

Felix Voigtlaender
felix.voigtlaender@ku.de

[1] Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12–14, 1040 Vienna, Austria

[2] Lehrstuhl Reliable Machine Learning, Katholische Universität Eichstätt-Ingolstadt, Ostenstraße 26, 85072 Eichstätt, Germany

Birkhäuser

1 Introduction

Time-frequency representations[1] (TF representations) are versatile tools for the analysis and decomposition of general functions (or signals) with respect to simpler, structured building blocks. They provide rich and intuitive information about a function's time-varying spectral behavior in settings where both time-series and stationary Fourier transforms are insufficient.

Important fields relying on time-frequency representations include signal processing [4, 18, 71, 77] and image processing [20, 24, 71, 86], medical imaging [69, 98], the numerical treatment of PDEs [26, 56], and quantum mechanics [76]. In particular, short-time Fourier transforms [53] and wavelet transforms [32] are widely and successfully used in these fields.

Yet, the limitations of such rigid schemes, considering only translations and modulations (resp. simple scalar dilations) of a single prototype function, are often considered detrimental to their representation performance. Therefore, numerous more flexible time-frequency representations have been proposed and studied in the last decades. As the most prominent of such systems, we mention curvelets [21, 23], shearlets [31, 67], ridgelets [22], and α-modulation systems [29, 30, 41, 55, 73, 84].

In the present article, we consider a more flexible scheme for constructing time-frequency representations, namely the framework of *warped time-frequency systems* that was recently introduced for dimension $d = 1$ in [61, 62]. To motivate this construction, note that the systems mentioned above are all examples of so-called *generalized translation-invariant (GTI) systems* [58, 65, 81], i.e., each of these systems is of the form $(\mathbf{T}_x \psi_i)_{i \in I, x \in Z_i}$ for certain generators $\psi_i \in \mathbf{L}^2(\mathbb{R}^d)$ and subgroups $Z_i \subset \mathbb{R}^d$. Here, $\mathbf{T}_x \psi(y) = \psi(y - x)$ denotes the translation of ψ by x. Although it is not required that the Z_i are discrete, they are often taken to be lattices, i.e., $Z_i = T_i \mathbb{Z}^d$, with $T_i \in \mathrm{GL}(\mathbb{R}^d)$. The various systems differ in the way in which the generators ψ_i and the lattices Z_i are chosen. But in each case there is a *finite* set of prototypes, often a single prototype, such that each ψ_i is a certain dilated and/or modulated version of one of the prototypes. Here, the dilations might be anisotropic, as is the case for shearlets.

As two canonical examples, we note that for a Gabor system, we have $\psi_k(x) = e^{2\pi i \alpha \langle k, x \rangle} \cdot \psi(x)$ for $k \in I = \mathbb{Z}^d$, while for a (homogeneous) wavelet system, we have $\psi_j(x) = 2^{dj/2} \cdot \psi(2^j x)$ for $j \in I = \mathbb{Z}$. Thus, the two systems differ with respect to the frequency localization of the generators ψ_i: For a Gabor system, the (essential) frequency supports of the generators ψ_k form a *uniform* covering of the frequency space $\mathbb{R}^d$— in contrast to the case of wavelets, where the (essential) frequency supports form a *dyadic* covering.

Warped time-frequency systems are motivated by the crucial observation that the dyadic covering corresponds to a uniform covering *with respect to a logarithmic scaling of the frequency space*. This suggests the following general construction: Starting from a *warping function* Φ—i.e., a diffeomorphism $\Phi : D \subset \mathbb{R}^d \to \mathbb{R}^d$—and a

[1] The term *time-frequency representation* is used in a wide sense here, also covering time-scale representations like wavelets.

prototype function $\theta \in \mathbf{L}^2(\mathbb{R}^d)$, we consider the associated *warped time-frequency system* $\mathcal{G}(\theta, \Phi) = \big(g_{y,\omega}\big)_{y\in\mathbb{R}^d,\omega\in D}$ given by

$$g_{y,\omega} = \mathbf{T}_y \left[\mathcal{F}^{-1} g_\omega\right] \quad \text{with} \quad g_\omega = c_\omega \cdot \big(\mathbf{T}_{\Phi(\omega)}\theta\big) \circ \Phi \qquad \text{for} \qquad (y, \omega) \in \mathbb{R}^d \times D. \tag{1.1}$$

Here, the function $c_\omega \cdot \big(\mathbf{T}_{\Phi(\omega)}\theta\big) \circ \Phi : D \subset \mathbb{R}^d \to \mathbb{C}$ is extended trivially to a map defined on all of $\mathbb{R}^d$ before applying the (inverse) Fourier transform $\mathcal{F}^{-1}$ to it, and the constant $c_\omega > 0$ is chosen such that the resulting family $(g_{y,\omega})_{y\in\mathbb{R}^d,\omega\in D}$ forms a *tight* frame for the space $\mathbf{L}^{2,\mathcal{F}}(D) = \mathcal{F}^{-1}(\mathbf{L}^2(D))$ of all $\mathbf{L}^2$-functions with Fourier transform vanishing outside of D.

At first sight, this construction might seem intimidating, but it can be unraveled as follows: The warping function Φ provides a map from the frequency space D to the warped frequency space $\mathbb{R}^d$. Thus, θ serves as a prototype for the Fourier transform of the GTI generators $\mathcal{F}^{-1} g_\omega$, but in warped coordinates. In that sense, g_ω can be understood as a shifted version of θ, but the shift is performed in warped (frequency) coordinates. In order to build further intuition for this construction, it is helpful to consider the case in which θ is (essentially) concentrated at 0, so that $\mathbf{T}_{\Phi(\omega)}\theta$ is concentrated at $\Phi(\omega)$, whence g_ω is concentrated at ω. Put briefly, the warping function Φ determines the frequency scale and, with it, the frequency-bandwidth relationship of the resulting warped time-frequency system.

As a further illustration, let us explain how wavelet systems fit into the above construction. Define $D := (0, \infty)$ and $\Phi : D \to \mathbb{R}, x \mapsto \ln(x)$. Then

$$\big([\,\mathbf{T}_{\Phi(\omega)}\theta\,] \circ \Phi\big)(\xi) = \theta\big(\ln(\xi) - \ln(\omega)\big) = [\theta \circ \ln]\big(\xi/\omega\big),$$

and hence, with $\psi = \mathcal{F}^{-1}(\theta \circ \ln)$, it holds that $\mathcal{F}^{-1} g_\omega = c_\omega \cdot \omega \cdot \big[\mathcal{F}^{-1}(\theta \circ \ln)\big]\,(\omega\bullet) = c_\omega\, \omega \cdot \psi(\omega\bullet)$, so that $(g_{y,\omega})_{y\in\mathbb{R},\omega\in D} = \big(c_\omega\, \omega \cdot \mathbf{T}_y[\psi(\omega\bullet)]\big)_{y\in\mathbb{R},\omega\in D}$ is a continuous wavelet system, for an appropriate choice of c_ω. Finally, since translations in the frequency domain correspond to modulations in the time domain, continuous Gabor systems can be obtained by choosing $\Phi : \mathbb{R}^d \to \mathbb{R}^d$ to be the identity function.

1.1 Contribution

The overall goal of the present article is to start an in-depth study of the properties of warped time-frequency systems on $\mathbb{R}^d$. Some essential, basic characteristics of warped time-frequency systems on $\mathbb{R}^d$ are obtained analogously to the one-dimensional case treated in [61]. In particular, $\mathcal{G}(\theta, \Phi)$ forms, under mild assumptions on θ and Φ, a *continuous* tight frame for $\mathbf{L}^{2,\mathcal{F}}(D)$. However, our main objective, verifying the applicability of *general coorbit theory* (as developed in [46, 66, 78]) to the continuous frame $\mathcal{G}(\theta, \Phi)$ is *decidedly* more involved in the higher-dimensional case that we consider here than in the case $d = 1$. Therefore, the main results presented in this work are concerned with establishing a set of assumptions on θ and Φ, such that the rich discretization theory for the coorbit spaces associated with $\mathcal{G}(\theta, \Phi)$ is accessible.

To make the latter point more precise, let us briefly recall the main points of coorbit theory related to the present setting. The main tenet of coorbit theory is to quantify the regularity of a function f using a certain norm $\|f\|_{\mathrm{Co}(Y)} := \|V_{\theta,\Phi} f\|_Y$ of the *voice transform* $V_{\theta,\Phi} f(y,\omega) = \langle f, g_{y,\omega}\rangle$. The *coorbit space* associated with a Banach space $Y \subset \mathbf{L}^1_{\mathrm{loc}}(\mathbb{R}^d \times D)$ is then given by $\mathrm{Co}_{\theta,\Phi}(Y) = \{f : V_{\theta,\Phi} f \in Y\}$.

Of course, the general theory of coorbit spaces as developed in [46, 66, 78] does not consider the special frame $\mathcal{G}(\theta, \Phi)$, but a general continuous frame $\Psi = (\psi_\lambda)_{\lambda\in\Lambda}$. Coorbit theory then provides (quite technical) conditions concerning the frame Ψ which ensure that the associated coorbit spaces $\mathrm{Co}_\Psi(Y)$ are indeed well-defined Banach spaces. We will verify these conditions in the setting of the warped time-frequency systems $\mathcal{G}(\theta, \Phi)$. Precisely, we shall derive verifiable conditions concerning θ and Φ which ensure that coorbit theory is applicable.

Additionally, coorbit spaces come with a powerful *discretization theory*: Under suitable conditions on the frame $\Psi = (\psi_\lambda)_{\lambda\in\Lambda}$, taken from an appropriate test function space, and on the discrete set $\Lambda_d \subset \Lambda$, coorbit theory shows that the sampled frame $\Psi_d = (\psi_\lambda)_{\lambda\in\Lambda_d}$ forms a *Banach frame decomposition* for the coorbit space $\mathrm{Co}_\Psi(Y)$. The precise definition of this concept will be given later. Here, we just note that it implies the existence of sequence spaces $Y_d^\flat \subset \mathbb{C}^{\Lambda_d}$ and $Y_d^\sharp \subset \mathbb{C}^{\Lambda_d}$ such that

$$\|f\|_{\mathrm{Co}(Y)} \asymp \left\|\big(\langle f, \psi_\lambda\rangle\big)_{\lambda\in\Lambda_d}\right\|_{Y_d^\flat} \asymp \inf\left\{\left\|(c_\lambda)_{\lambda\in\Lambda_d}\right\|_{Y_d^\sharp} : f = \sum_{\lambda\in\Lambda_d} c_\lambda \cdot \psi_\lambda\right\}.$$

Hence, for a generalized notion of *sparsity*, membership of f in $\mathrm{Co}_\Psi(Y)$ is *simultaneously* equivalent to *analysis sparsity* and *synthesis sparsity* of f with respect to the discretized frame Ψ_d. Specifically, a sequence c is considered sparse if $c \in Y_d^\flat$ or $c \in Y_d^\sharp$. This is most closely related to classical sparsity if $Y_d^\flat$ and $Y_d^\sharp$ coincide with certain (weighted) ℓ^p spaces.

We indeed show under suitable conditions concerning θ and Φ that the discretization theory applies to $\mathcal{G}(\theta, \Phi)$. Therefore, the coorbit spaces $\mathrm{Co}_{\theta,\Phi}(Y)$ characterize sparsity with respect to the (suitably discretized) warped time-frequency system $\mathcal{G}(\theta, \Phi)$. As a byproduct, we also show that the space $\mathrm{Co}_{\theta,\Phi}(Y)$ is essentially independent of the choice of *sufficiently regular* θ.

1.2 Related Work: Warped Time-Frequency Systems

Warped time-frequency systems have already been considered before, though only for the one-dimensional case $d = 1$. In particular, in [61], the authors essentially obtain the results that we just outlined, i.e., that warped time-frequency systems form tight frames and that the assumptions of generalized coorbit theory can be satisfied, at least for coorbit spaces associated to the (weighted) Lebesgue spaces $Y = \mathbf{L}^p_\kappa(\mathbb{R} \times D)$. We generalize these results to higher dimensions $d > 1$ and to the weighted *mixed* Lebesgue spaces $\mathbf{L}^{p,q}_\kappa(\mathbb{R}^d \times D)$, equipped with the norm $\|F\|_{\mathbf{L}^{p,q}_\kappa} = \left\|\omega \mapsto \|(\kappa \cdot F)(\bullet, \omega)\|_{\mathbf{L}^p(\mathbb{R}^d)}\right\|_{\mathbf{L}^q(D)}$. Furthermore, we relax some of the assumptions imposed in

[61]. The generalization to higher dimensions is, as we will see, by no means trivial. The extension to the spaces $\mathbf{L}_{\kappa}^{p,q}(\mathbb{R}^d \times D)$ relies on our recent work [60].

Hilbert space frames obtained by sampling warped time-frequency systems were examined in [62], where different necessary or sufficient frame conditions similar to those for Gabor and wavelet frames were obtained. In the same paper, the authors also derive readily verifiable conditions under which the sampled warped time-frequency system satisfies the local integrability condition, thereby providing access to useful results from the theory of GTI systems.

1.3 Related Work: GTI Systems

Warped time-frequency representations are GTI systems [58, 65, 81], and they could be analyzed within this abstract framework. However, fully general GTI systems include a considerable number of—usually undesired— pathological cases [49, 93]; these can be excluded by imposing additional structure-enforcing conditions. The most general and well-known such condition is the *local integrability condition* (LIC) of Hernandez et al. [58], further investigated in [65, 93].

In practice, GTI systems are mostly generated from one (or few) prototype functions through the application of a family of operators—like modulations or dilations—that promote a given frequency-bandwidth relationship, such as the constant frequency/bandwidth ratio for classical wavelet systems. Naturally, such systems are well suited for representing functions with certain frequency-domain properties.

In our case, structure is imposed by the choice of the prototype and warping function that determine the frequency-bandwidth relation and the distribution of GTI generators in the frequency domain. In this sense, warped time-frequency systems provide a unified framework for studying structured time-frequency representations. We will see that warped time-frequency systems, despite their generality, satisfy many beneficial properties that are not simply trivial consequences of them being GTI systems.

As other related time-frequency systems, we mention dictionaries obtained by combining multiple TF dictionaries, either globally [5, 8, 99], or locally in *weaved* phase space covers [34, 37, 80]. Furthermore, nonstationary Gabor systems [10, 35, 36, 59] are closely related to GTI systems via the Fourier transform.

1.4 Related Work: Function Space Theory

The joint study of integral transforms and appropriate (generalized) function spaces is a classical topic in Fourier- and harmonic analysis. In particular, localization and smoothness properties of functions and their Fourier transforms have received much attention. Indeed, from the distribution theory of Laurent Schwartz [82, 83] to Paley-Wiener spaces [17], Sobolev spaces [3, 68, 88] and Besov spaces [16, 88, 91], a large number of classical function spaces can be meaningfully characterized through their Fourier transform properties. Other examples include the family of modulation spaces [40, 53]— defined through the short-time Fourier transform [50, 53]—as well as spaces of (poly-)analytic functions [1, 12] and the Bargmann [13, 14] and Bergman transforms [2].

A powerful general framework for studying function spaces associated with a certain transform is provided by coorbit theory, originally introduced by Feichtinger and Gröchenig [43, 44, 52]. As described above, the underlying idea for this theory is to measure the regularity of a function or distribution in terms of growth or decay properties of an abstract *voice transform*. In the original approach of Feichtinger and Gröchenig, the voice transform is defined through an integrable group representation acting on a suitable prototype function. Prime examples of different transforms and the associated coorbit spaces are the short-time Fourier transform [50, 53] and modulation spaces, associated with the (reduced) Heisenberg group, and the wavelet transform [32] and (homogeneous) Besov spaces [16, 91], associated with the $ax + b$ group.

Fornasier and Rauhut [46] realized that the group structure on which classical coorbit theory relies can be discarded completely. Instead, one can consider the voice transform associated with a general *continuous frame* [6, 7], the Gramian kernel of which is required to satisfy certain integrability and oscillation conditions. Since the introduction of this *general coorbit theory*, these results have been improved and expanded [11, 66, 78], as well as successfully applied, e.g., to Besov and Triebel-Lizorkin spaces [88, 90, 91] or α-modulation spaces [51]; see e.g. [78, 92] and [30, 84].

1.5 Structure of the Paper

We begin with a brief introduction to general coorbit theory in Sect. 2. We then formally introduce warped time-frequency systems in Sect. 3, in which we also discuss several concrete examples. Section 4 is concerned with conditions on the warping function Φ and the prototype θ which ensure that the continuous frame $\mathcal{G}(\theta, \Phi)$ satisfies the assumptions of (general) coorbit theory.

To show that the continuous frame $\mathcal{G}(\theta, \Phi)$ can be sampled to obtain discrete Banach frame decompositions of the associated coorbit spaces, we will need certain coverings of the phase space $\Lambda = \mathbb{R}^d \times D$ associated with the warping function Φ. These coverings are studied in Sect. 5. In Sect. 6, we prove the existence of discrete Banach frame decompositions for the coorbit spaces $\mathrm{Co}_{\theta,\Phi}(Y)$. Finally, in Sect. 8 we investigate warped time-frequency systems generated by *radial* warping functions on $\mathbb{R}^d$. In particular, we show that admissible symmetric warping functions on $\mathbb{R}$ give rise to admissible radial warping functions on $\mathbb{R}^d$.

1.6 Notation and Fundamental Definitions

We use the notation $\underline{n} := \{1, \dots, n\}$ for $n \in \mathbb{N}$. We write $\mathbb{R}^+ = (0, \infty)$ for the set of positive real numbers, and $S^1 := \{z \in \mathbb{C} \colon |z| = 1\}$. For the composition of functions f and g we use the notation $f \circ g$ defined by $f \circ g(x) = f(g(x))$. For a subset $M \subset X$ of a fixed base set X (which is usually understood from the context), we use the indicator function $\mathbb{1}_M$ of the set M, where $\mathbb{1}_M(x) = 1$ if $x \in M$ and $\mathbb{1}_M(x) = 0$ otherwise.

The (topological) dual space of a (complex) topological vector space X (i.e., the space of all continuous linear functions $\varphi : X \to \mathbb{C}$) is denoted by X', while the (topological) *anti*-dual of a Banach space X (i.e., the space of all *anti-linear* continuous functionals on X) is denoted by $X^{\urcorner}$. A superscript asterisk (*) is used to denote the adjoint of an operator between Hilbert spaces.

We use the convenient short-hand notations $\lesssim$ and $\asymp$, where $A \lesssim B$ means $A \leq C \cdot B$, for some constant $C > 0$ that depends on quantities that are either explicitly mentioned or clear from the context. $A \asymp B$ means $A \lesssim B$ and $B \lesssim A$.

1.6.1 Norms and Related Notation

We write $|x|$ for the Euclidean norm of a vector $x \in \mathbb{R}^d$, and we denote the operator norm of a linear operator $T : X \to Y$ by $\|T\|_{X \to Y}$, or by $\|T\|$, if X, Y are clear from the context. In the expression $\|A\|$, a matrix $A \in \mathbb{R}^{n \times d}$ is interpreted as a linear map $(\mathbb{R}^d, |\bullet|) \to (\mathbb{R}^n, |\bullet|)$. The open (Euclidean) ball around $x \in \mathbb{R}^d$ of radius $r > 0$ is denoted by $B_r(x)$.

1.6.2 Fourier-Analytic Notation

The Lebesgue measure of a (measurable) subset $M \subset \mathbb{R}^d$ is denoted by $\mu(M)$. The Fourier transform is given by $\widehat{f}(\xi) = \mathcal{F}f(\xi) = \int_{\mathbb{R}^d} f(x)\, e^{-2\pi i \langle x, \xi \rangle}\, dx$, for all $f \in \mathbf{L}^1(\mathbb{R}^d)$. It is well-known that $\mathcal{F}$ extends to a unitary automorphism of $\mathbf{L}^2(\mathbb{R}^d)$. The inverse Fourier transform is denoted by $\check{f} := \mathcal{F}^{-1} f$. We write $\mathbf{L}^{2,\mathcal{F}}(D) := \mathcal{F}^{-1}(\mathbf{L}^2(D))$ for the space of square-integrable functions whose Fourier transform vanishes (a.e.) outside of $D \subset \mathbb{R}^d$. In addition to the Fourier transform, the *modulation* and *translation operators* $\mathbf{M}_\omega f = f \cdot e^{2\pi i \langle \omega, \bullet \rangle}$ and $\mathbf{T}_y f = f(\bullet - y)$, will be used frequently.

1.6.3 Matrix Notation

For matrix-valued functions $A : U \to \mathbb{R}^{d \times d}$, the notation $A(x)\langle y \rangle := A(x) \cdot y$ denotes the multiplication of the matrix $A(x)$, $x \in U$, with the vector $y \in \mathbb{R}^d$ in the usual sense. Likewise, for a set $M \subset \mathbb{R}^d$, we write

$$A(x)\langle M \rangle := \big\{ A(x)\langle y \rangle \ : \ y \in M \big\}.$$

Moreover, we define $A^{-1}(\tau) := [A(\tau)]^{-1}$ and similarly $A^{\pm T}(\tau) := [A(\tau)]^{\pm T}$. Here, as in the remainder of the paper, the notation A^T denotes the transpose of a matrix A. We will denote the elements of the standard basis of $\mathbb{R}^d$ by $e_1, \ldots, e_d$.

1.6.4 Convention for Variables

Throughout this article, $x, y, z \in \mathbb{R}^d$ will be used to denote variables in time/position space, $\xi, \omega, \eta \in D$ in frequency space, $\lambda, \rho, \nu \in \mathbb{R}^d \times D$ in phase space, and finally $\sigma, \tau, \iota \in \mathbb{R}^d$ denote variables in warped frequency space. Unless otherwise stated, this also holds for subscript-indexed variants; precisely, subscript indices (i.e., x_i) are used to denote the i-th element of a vector $x \in \mathbb{R}^d$. In some cases, we also use subscripts to enumerate multiple vectors, e.g., $x_1, \dots, x_n \in \mathbb{R}^d$. In this case, we denote the components of x_i by $(x_i)_j$.

1.6.5 Solid Spaces, Integral Kernels, and Mixed Lebesgue Spaces

Unless noted otherwise, we will always consider $\Lambda = \mathbb{R}^d \times D$ (with an open set $D \subset \mathbb{R}^d$), equipped with the Lebesgue measure μ. A Banach space $Y \subset \mathbf{L}^1_{\mathrm{loc}}(\Lambda)$ will be called *solid* if it satisfies the following: whenever $F, G : \Lambda \to \mathbb{C}$ are measurable with $|F| \leq |G|$ almost everywhere and with $G \in Y$, then $F \in Y$ and $\|F\|_Y \leq \|G\|_Y$. Y is *rich*, if it contains all (locally) integrable, compactly supported functions. The analogous definitions apply for general locally compact measure spaces, and in particular to sequence spaces (where the index set is equipped with the discrete topology).

A *kernel* on Λ is a (measurable) function $K : \Lambda \times \Lambda \to \mathbb{C}$. Its application to a (measurable) function $F : \Lambda \to \mathbb{C}$ is denoted by

$$K(F)(\lambda) := \int_\Lambda K(\lambda, \rho) F(\rho)\, d\mu(\rho), \qquad \text{whenever the integral exists.} \tag{1.2}$$

We will identify two kernels if they agree almost everywhere. As usual, K^* denotes the *adjoint kernel* $K^*(\lambda, \rho) = \overline{K(\rho, \lambda)}$, and K^T denotes the *transposed kernel*, given by $K^T(\lambda, \rho) = K(\rho, \lambda)$.

Since $\Lambda = \mathbb{R}^d \times D$ has a product structure, it is natural to consider the weighted, *mixed* Lebesgue spaces $\mathbf{L}^{p,q}_\kappa(\Lambda)$, for $1 \leq p, q \leq \infty$, that consist of all (equivalence classes of almost everywhere equal) measurable functions $F : \Lambda \to \mathbb{C}$ for which

$$\|F\|_{\mathbf{L}^{p,q}_\kappa} := \big\| \lambda_2 \mapsto \|(\kappa \cdot F)(\bullet, \lambda_2)\|_{\mathbf{L}^p(\mathbb{R}^d)} \big\|_{\mathbf{L}^q(D)} < \infty. \tag{1.3}$$

Here, $\kappa : \Lambda \to \mathbb{R}^+$ is a (measurable) weight function.

2 Frames, Coverings and Coorbit Spaces

In this section, we prepare our investigation of warped time-frequency systems by recalling several notions and results related to the theory of continuous frames and general coorbit theory.

A collection $\Psi = (\psi_\lambda)_{\lambda \in \Lambda}$ of elements $\psi_\lambda \in \mathcal{H}$ of a separable Hilbert space $\mathcal{H}$ is called a *tight continuous frame (for $\mathcal{H}$)*, if there exists $A \in \mathbb{R}^+$ such that

$$A \cdot \|f\|^2_{\mathcal{H}} = \int_\Lambda |\langle f, \psi_\lambda \rangle_{\mathcal{H}}|^2 d\mu(\lambda) \quad \text{for all } f \in \mathcal{H}, \tag{2.1}$$

and if furthermore the map $\lambda \mapsto \psi_\lambda$ is weakly measurable, meaning that $\lambda \mapsto \langle f, \psi_\lambda \rangle$ is measurable for each $f \in \mathcal{H}$. For the warped time-frequency systems considered later, we will see that $\lambda \mapsto \psi_\lambda$ is in fact continuous (see Proposition 3.4). We say that Ψ is a *Parseval frame* if $A = 1$ in Eq. (2.1).

The *voice transform* with respect to a tight continuous frame $\Psi = (\psi_\lambda)_{\lambda \in \Lambda}$ is given by

$$V_\Psi \ : \ \mathcal{H} \to \mathbf{L}^2(\Lambda), \quad \text{defined by} \quad V_\Psi f(\lambda) := \langle f, \psi_\lambda \rangle_{\mathcal{H}} \text{ for all } \lambda \in \Lambda. \tag{2.2}$$

The adjoint of the voice transform is given by

$$V_\Psi^* \ : \ \mathbf{L}^2(\Lambda) \to \mathcal{H}, \quad V_\Psi^* G = \int_\Lambda G(\lambda)\, \psi_\lambda d\mu(\lambda), \tag{2.3}$$

with the integral understood in the weak sense (see [53, Page 43]). Finally, the *frame operator* of Ψ is given by $\mathbf{S}_\Psi := V_\Psi^* \circ V_\Psi : \mathcal{H} \to \mathcal{H}$, so that

$$\mathbf{S}_\Psi f = \int_\Lambda \langle f, \psi_\lambda \rangle_{\mathcal{H}} \psi_\lambda d\mu(\lambda).$$

It follows from (2.1) that $\mathbf{S}_\Psi f = A \cdot f$ for all $f \in \mathcal{H}$; see [25].

Essentially all of coorbit theory is based on certain regularity properties of the *reproducing kernel* K_Ψ associated to the continuous frame Ψ. It is given by

$$K_\Psi : \quad \Lambda \times \Lambda \to \mathbb{C}, \quad (\lambda, \rho) \mapsto A^{-1} \langle \psi_\rho, \psi_\lambda \rangle_{\mathcal{H}}. \tag{2.4}$$

Without loss of generality, we will henceforth assume that $A = 1$, i.e., Ψ is a Parseval frame. We remark that K_Ψ is measurable with respect to the product σ-algebra. Indeed, since $\mathcal{H}$ is separable, we can choose a countable orthonormal basis $(\eta_j)_{j \in J} \subset \mathcal{H}$, so that $K_\Psi(\lambda, \rho) = \sum_{j \in J} \langle \psi_\rho, \eta_j \rangle_{\mathcal{H}} \langle \eta_j, \psi_\lambda \rangle_{\mathcal{H}}$ is seen to be measurable as a convergent, countable series of measurable functions.

A *(discrete) frame* for $\mathcal{H}$ is a countable family $\Psi_d = (\psi_j)_{j \in J} \subset \mathcal{H}$ for which there exist $0 < A \leq B < \infty$ such that

$$A \cdot \|f\|_{\mathcal{H}}^2 \leq \sum_{j \in J} |\langle f, \psi_j \rangle_{\mathcal{H}}|^2 \leq B \cdot \|f\|_{\mathcal{H}}^2 \quad \text{for all } f \in \mathcal{H}. \tag{2.5}$$

This implies (cf. [25] for details) that every $f \in \mathcal{H}$ can be expanded with respect to Ψ_d; that is, for each $f \in \mathcal{H}$ there exists a sequence $(c_j)_{j \in J} \in \ell^2(J)$ such that

$$f = \sum_{j \in J} c_j\, \psi_j. \tag{2.6}$$

2.1 Banach Frame Decompositions

When the Hilbert space $\mathcal{H}$ is exchanged for a Banach space $(B, \| \bullet \|_B)$, and $\ell^2(J)$ is replaced by a suitable sequence space $B^\flat \subset \mathbb{C}^J$, then validity of the (modified) frame inequality $\big\| \big(\langle f, \psi_j \rangle_{B,B'} \big)_{j \in J} \big\|_{B^\flat} \asymp \|f\|_B$ does *not* necessarily imply a statement similar to (2.6) (among other things because in general $\psi_j \in B'$ and not $\psi_j \in B$). Therefore, the dual concepts of Banach frames and atomic decompositions have been introduced; see [43, 44, 52]. To reduce the number of required definitions, in this article we only consider the combined concept of a *Banach frame decomposition*, which unifies both concepts, under some mild assumptions on B that allow one to make sense of the intersection $B \cap B'$; see Appendix Appendix A for details.

Definition 2.1 Let $(B, \|\bullet\|_B)$ be a Banach space. A family $\Psi_d = (\psi_j)_{j\in J} \subset B \cap B'$ is called a *Banach frame decomposition* for B if there exist a dual family $E_d = (e_j)_{j\in J} \subset B \cap B'$ and solid, rich Banach sequence spaces $(B^\sharp, \|\bullet\|_{B^\sharp})$ and $(B^\flat, \|\bullet\|_{B^\flat})$ over J, i.e., $B^\sharp, B^\flat \subset \mathbb{C}^J$, with the following properties:

- The *coefficient operators*

$$\begin{aligned} \mathcal{C}_{\Psi_d}: \quad & B \to B^\flat, \quad f \mapsto \big(\langle f, \psi_j\rangle_{B,B'}\big)_{j\in J} \qquad \text{and} \\ \mathcal{C}_{E_d}: \quad & B \to B^\sharp, \quad f \mapsto \big(\langle f, e_j\rangle_{B,B'}\big)_{j\in J} \end{aligned}$$

  are well-defined and bounded.
- The *reconstruction operators*

$$\begin{aligned} \mathcal{R}_{\Psi_d}: \quad & B^\sharp \to B, \quad (c_j)_{j\in J} \mapsto \sum_{j\in J} c_j\, \psi_j \quad \text{and} \\ \mathcal{R}_{E_d}: \quad & B^\flat \to B, \quad (c_j)_{j\in J} \mapsto \sum_{j\in J} c_j\, e_j \end{aligned}$$

  are well-defined and bounded, with unconditional convergence of the defining series in a suitable topology.
- We have $\mathcal{R}_{\Psi_d} \circ \mathcal{C}_{E_d} = \mathrm{id}_B = \mathcal{R}_{E_d} \circ \mathcal{C}_{\Psi_d}$, or in other words

$$f = \sum_{j\in J} \langle f, e_j\rangle_{B,B'}\ \psi_j = \sum_{j\in J} \langle f, \psi_j\rangle_{B,B'}\ e_j \qquad \text{for all } f \in B.$$

Remark 2.2 In some recent works, atomic decompositions of Banach spaces are defined by a pair of systems $(\Psi_d, \widetilde{\Psi_d})$, with $\Psi_d \in B'$ providing the analysis, and $\widetilde{\Psi_d} \in B$ the synthesis operation, e.g., [25, Definition 24.3.1]. In that sense, Definition 2.1 is not dissimilar to stating that both (Ψ_d, E_d) and (E_d, Ψ_d) are atomic decompositions of B. Nonetheless, a Banach frame decomposition, which implies the existence of a class of test functions embedded into B and B', is distinct, since it places additional assumptions on the sequence spaces $B^\sharp, B^\flat$ on which the reconstruction operators are further required to be unconditionally convergent.

2.2 Coverings and Weight Functions

For applying the discretization results of (general) coorbit theory, we will have to construct special coverings of the phase space $\Lambda = \mathbb{R}^d \times D$. To allow for a more streamlined development later on, the present subsection discusses the required properties of these coverings. The most basic of these properties are admissibility and moderateness.

Definition 2.3 Let $\mathcal{O} \neq \varnothing$ be a set. A family $\mathcal{V} = (V_j)_{j \in J}$ of non-empty subsets of $\mathcal{O}$ is called an *admissible covering* of $\mathcal{O}$, if we have $\mathcal{O} = \bigcup_{j \in J} V_j$ and if

$$\mathcal{N}(\mathcal{V}) := \sup_{j \in J} |j^*| < \infty \qquad \text{where} \qquad j^* := \{i \in J : V_i \cap V_j \neq \varnothing\} \quad \text{for } j \in J. \tag{2.7}$$

If $\mathcal{O}$ is a topological space, we say that a family $\mathcal{V}$ as above is *topologically admissible* if it is admissible and if each $V_j \subset \mathcal{O}$ is open and relatively compact (i.e., $\overline{V_j} \subset \mathcal{O}$ is compact).

Remark We remark that every topologically admissible covering is locally finite: Given $x \in \mathcal{O}$, we have $x \in V_{j_0}$ for some $j_0 \in J$. Since V_{j_0} is open and since $V_j \cap V_{j_0} \neq \varnothing$ can only hold for $i \in j_0^*$ with $j_0^* \subset J$ finite, we see that $\mathcal{V}$ is indeed a locally finite covering.

In the special case where $\mathcal{O} = \Lambda$ has a product structure, we will also use the following class of coverings.

Definition 2.4 [60, Def. 2.12] Let $\Lambda = \Lambda_1 \times \Lambda_2$, where each Λ_j is equipped with a measure μ_j and $\mu = \mu_1 \otimes \mu_2$. We say that a family $\mathcal{U} = (U_j)_{j \in J}$ is a *product-admissible covering* of Λ, if it satisfies the following: J is countable, $\Lambda = \bigcup_{j \in J} U_j$, each U_j is non-empty and of the form $U_j = U_{1,j} \times U_{2,j}$ with $U_{\ell,j} \subset \Lambda_\ell$ open, and there is a constant $C > 0$ such that the *covering weight* $w_{\mathcal{U}}$ defined by

$$(w_{\mathcal{U}})_j := \min\left\{1,\ \mu_1(U_{1,j}),\ \mu_2(U_{2,j}),\ \mu(U_j)\right\} \qquad \text{for } j \in J \tag{2.8}$$

satisfies $(w_{\mathcal{U}})_j \leq C \cdot (w_{\mathcal{U}})_\ell$ for all $j, \ell \in J$ with $U_j \cap U_\ell \neq \varnothing$.

Given a product-admissible covering $\mathcal{U} = (U_j)_{j \in J}$ and a measurable function $u : \Lambda \to \mathbb{R}^+$, we say that u is $\mathcal{U}$*-moderate* if there is a constant $C' > 0$, such that $u(\lambda) \leq C' \cdot u(\rho)$ for all $j \in J$ and all $\lambda, \rho \in U_j$.

If $\mathcal{U} = (U_j)_{j \in J}$ is a product-admissible covering of Λ, then with $w_{\mathcal{U}}$ as defined in (2.8), it is easy to see that there exists a measurable function $w_{\mathcal{U}}^c : \Lambda \to \mathbb{R}^+$ such that

$$(w_{\mathcal{U}})_j \asymp w_{\mathcal{U}}^c(\lambda) \quad \text{for all } j \in J \text{ and } \lambda \in U_j. \tag{2.9}$$

Furthermore, any two such weights $w_{\mathcal{U}}^c$, $\widetilde{w_{\mathcal{U}}^c}$ satisfy $w_{\mathcal{U}}^c \asymp \widetilde{w_{\mathcal{U}}^c}$. We refer to [60, Theorem 2.13] for the details.

In addition to such coverings, the study of specific coorbit spaces and their properties relies on certain weighted spaces that are compatible with the given coverings in a suitable way. The following classes of weight functions are of particular importance.

Definition 2.5 1. Any measurable function $v : \mathcal{O} \to \mathbb{R}^+$ on a measurable space $\mathcal{O}$ will be called a *weight*, or a *weight function*.

2. A weight $m : \mathcal{O} \times \mathcal{O} \to \mathbb{R}^+$ is called *symmetric* if $m(\lambda, \rho) = m(\rho, \lambda)$ for all $\lambda, \rho \in \mathcal{O}$.
3. Given any weight $v : \mathcal{O} \to \mathbb{R}^+$, the *associated weight* $m_v : \mathcal{O} \times \mathcal{O} \to \mathbb{R}^+$ is defined by

$$m_v(\lambda, \rho) := \max\left\{\frac{v(\lambda)}{v(\rho)}, \frac{v(\rho)}{v(\lambda)}\right\}, \qquad \text{for all } \lambda, \rho \in \mathcal{O}. \tag{2.10}$$

4. A weight function v on $\mathbb{R}^d$ is called *submultiplicative*, if

$$v(\lambda + \rho) \leq v(\lambda) \cdot v(\rho), \qquad \text{for all } \lambda, \rho \in \mathbb{R}^d.$$

   Given such a submultiplicative weight v, another weight function $\widetilde{v} : \mathbb{R}^d \to \mathbb{R}^+$ is called v*-moderate* if

$$\widetilde{v}(\lambda + \rho) \leq v(\lambda) \cdot \widetilde{v}(\rho), \qquad \text{for all } \lambda, \rho \in \mathbb{R}^d. \tag{2.11}$$

5. We say that a weight v on $\mathbb{R}^d$ is *radially increasing* if $v(\lambda) \leq v(\rho)$ whenever $\lambda, \rho \in \mathbb{R}^d$ with $|\lambda| \leq |\rho|$. This in particular implies that $v(\lambda)$ only depends on $|\lambda|$.

Remark 2.6 If v_1, v_2 are v_0-moderate weights and $v_0(\lambda) = v_0(-\lambda)$ for all $\lambda \in \mathbb{R}^d$, then a simple derivation shows that $1/v_1$, $\max\{v_1, v_2\}$, and $\min\{v_1, v_2\}$ are v_0-moderate as well.

2.3 Kernel Spaces

The main prerequisite of general coorbit theory is that the reproducing kernel K_Ψ—and some additional kernels derived from it—must satisfy appropriate decay conditions. These are formulated in terms of certain *Banach spaces of integral kernels* that we review in this subsection.

Let (Λ, μ) be a σ-finite measure space. Recall from Sect. 1.6.5 that a kernel is any measurable map $K : \Lambda \times \Lambda \to \mathbb{C}$. Given such a kernel and a symmetric weight m on $\Lambda \times \Lambda$, we define $\|K\|_{\mathcal{A}_m(\Lambda)} := \|K\|_{\mathcal{A}_m}$, where

$$\begin{aligned} \|K\|_{\mathcal{A}_m} := \max\Bigg\{ &\operatorname*{ess\,sup}_{\rho \in \Lambda} \int_\Lambda \big|m(\rho, \lambda) \cdot K(\rho, \lambda)\big| \, d\mu(\lambda), \\ &\operatorname*{ess\,sup}_{\lambda \in \Lambda} \int_\Lambda \big|m(\rho, \lambda) \cdot K(\rho, \lambda)\big| \, d\mu(\rho) \Bigg\}, \end{aligned} \tag{2.12}$$

and we define $\mathcal{A}_m := \mathcal{A}_m(\Lambda) := \big\{K : \Lambda \times \Lambda \to \mathbb{C} : \ K \text{ measurable and } \|K\|_{\mathcal{A}_m} < \infty\big\}$. In the case where $m \equiv 1$, we use the notation $\mathcal{A}_1$.

For most applications, it is not enough to know that $K_\Psi \in \mathcal{A}_m$; rather, it is required that the integral operator associated to K_Ψ or $|K_\Psi|$ (defined in Equation (1.2)) acts boundedly on a given solid Banach space $Y \subset \mathbf{L}^1_{\mathrm{loc}}(\Lambda)$. Precisely, given a kernel $K : \Lambda \times \Lambda \to \mathbb{C}$, we set $\big\| \, |K| \, \big\|_{Y \to Y} := \infty$ if the integral operator associated to $|K|$ does *not* define a bounded linear map on Y; otherwise, we denote by $\big\| \, |K| \, \big\|_{Y \to Y}$ the operator norm of this integral operator. With this convention, we define

$$\mathcal{A}_{m,Y} := \left\{ K \in \mathcal{A}_m \ : \ \big\| \, |K| \, \big\|_{Y \to Y} < \infty \right\}, \quad \text{with norm}$$
$$\|K\|_{\mathcal{A}_{m,Y}} := \max \left\{ \|K\|_{\mathcal{A}_m}, \big\| \, |K| \, \big\|_{Y \to Y} \right\}.$$

Remark 2.7 (cf. [66, Lemma 2.45]) If K is measurable and if $|K|$ induces a bounded operator $Y \to Y$, then so does K itself, since Y is solid. A similar argument shows that $\mathcal{A}_{m,Y}$ is a solid space of kernels: Let K, L be measurable with $K \in \mathcal{A}_{m,Y}$ and $|L| \leq |K|$ almost everywhere (with respect to the product measure). Then, for μ-almost every $\lambda \in \Lambda$, $|L(\lambda, \bullet)| \leq |K(\lambda, \bullet)|$ μ-almost everywhere, implying

$$||L|(F)(\lambda)| \leq |L|(|F|)(\lambda) \leq |K|(|F|)(\lambda) \quad \mu\text{-almost everywhere.}$$

Noting that $\|F\|_Y = \||F|\|_Y$ due to solidity of Y, the first inequality implies that to determine the operator norm $\||L|\|_{Y \to Y}$, it suffices to consider nonnegative functions $F \in Y$. On the other hand, for such functions, the second inequality implies $\||L|(F)\|_Y \leq \||K|(F)\|_Y$, by solidity of Y. Hence, we have established $\||L|\|_{Y \to Y} \leq \||K|\|_{Y \to Y}$, and therefore $\||L|\|_{\mathcal{A}_{m,Y}} \leq \||K|\|_{\mathcal{A}_{m,Y}}$ follows with solidity of $\mathcal{A}_m$, which is clear from the definition.

Finally, we remark that our definition of $\mathcal{A}_{m,Y}$ is different from the definition in [66, Section 2.4] in that we take the norm $\big\| \, |K| \, \big\|_{Y \to Y}$ instead of $\|K\|_{Y \to Y}$. Nevertheless, if a kernel K satisfies $K \in \mathcal{A}_{m,Y}$ with our definition, it also satisfies $K \in \mathcal{A}_{m,Y}$ according to the definition in [66, Section 2.4], so that the slightly different definition will not cause problems.

For applications of coorbit theory, one has to verify $K_\Psi \in \mathcal{A}_{m,Y}$ for the space Y of interest and a certain weight m. In many cases, it turns out to be easier to verify $K_\Psi \in \mathcal{B}_{m_0}$, where $\mathcal{B}_{m_0}$ is a smaller space of kernels that satisfies $\mathcal{B}_{m_0} \hookrightarrow \mathcal{A}_{m,Y}$, possibly with $m_0 = m$. Precisely, since we are mostly interested in the product setting of kernels on $\Lambda = \Lambda_1 \times \Lambda_2$, we will use the following spaces $\mathcal{B}_m$ introduced in [60].

Definition 2.8 Let $(\Lambda, \mu) = (\Lambda_1 \times \Lambda_2, \mu_1 \otimes \mu_2)$, where (Λ_1, μ_1), (Λ_2, μ_2) are σ-finite measure spaces. Given a kernel $K : \Lambda \times \Lambda \to \mathbb{C}$, we define

$$K^{(\lambda_2, \rho_2)}(\lambda_1, \rho_1) := K(\lambda, \rho) \quad \text{for} \quad \lambda = (\lambda_1, \lambda_2), \rho = (\rho_1, \rho_2) \in \Lambda. \tag{2.13}$$

Using this notation, we define

$$\|K\|_{\mathcal{B}_1} := \|K\|_{\mathcal{B}_1(\Lambda)} := \Big\| (\lambda_2, \rho_2) \mapsto \big\| K^{(\lambda_2, \rho_2)} \big\|_{\mathcal{A}_1(\Lambda_1)} \Big\|_{\mathcal{A}_1(\Lambda_2)} \in [0, \infty],$$

and $\mathcal{B}_1 := \mathcal{B}_1(\Lambda) := \big\{K : \Lambda \times \Lambda \to \mathbb{C} : K \text{ measurable and } \|K\|_{\mathcal{B}_1} < \infty\big\}$. Finally, given a symmetric weight $m : \Lambda \times \Lambda \to \mathbb{R}^+$, we define $\mathcal{B}_m := \mathcal{B}_m(\Lambda) := \big\{K : \Lambda \times \Lambda \to \mathbb{C} : m \cdot K \in \mathcal{B}_1\big\}$, with norm $\|K\|_{\mathcal{B}_m} := \|m \cdot K\|_{\mathcal{B}_1}$.

As shown in [60, Propositions 2.5 and 2.6], $\mathcal{B}_m$ is a solid Banach space of integral kernels that satisfies $\|K^T\|_{\mathcal{B}_m} = \|K\|_{\mathcal{B}_m}$ and furthermore $\|K\|_{\mathcal{A}_m} \leq \|K\|_{\mathcal{B}_m}$ for every kernel K. If the weight m additionally satisfies $m(x, z) \leq C m(x, y) m(y, z)$, for all $x, y, z \in \Lambda$ and some $C > 0$, then it is easy to see that $\mathcal{A}_m, \mathcal{B}_m$ are algebrae with respect to the standard kernel product, defined by

$$K_1 \cdot K_2 = \int_\Lambda K_1(\bullet_1, \lambda) K_2(\lambda, \bullet_2)\, d\mu(\lambda).$$

Most importantly for us, the integral operators associated to kernels in $\mathcal{B}_{m_\kappa}$ act boundedly on the mixed-norm Lebesgue spaces $\mathbf{L}^{p,q}_\kappa(\Lambda)$; see the following proposition.

Proposition 2.9 (see [60, Proposition 2.7]) *Let Λ as in Definition 2.8, let κ be a weight on Λ, and let $m_\kappa : \Lambda \times \Lambda \to \mathbb{R}^+$ be as in Equation (2.10). Then, for each kernel $K \in \mathcal{B}_{m_\kappa}(\Lambda)$ and arbitrary $p, q \in [1, \infty]$, the associated integral operator $K(\bullet)$ defined in Equation (1.2) restricts to a bounded linear operator $K(\bullet) : \boldsymbol{L}^{p,q}_\kappa(\Lambda) \to \boldsymbol{L}^{p,q}_\kappa(\Lambda)$, with absolute convergence almost everywhere of the defining integral, and with*

$$\|K(F)\|_{\boldsymbol{L}^{p,q}_\kappa(\Lambda)} \leq \|K\|_{\mathcal{B}_{m_\kappa}} \cdot \|F\|_{\boldsymbol{L}^{p,q}_\kappa(\Lambda)} \qquad \forall\, F \in \boldsymbol{L}^{p,q}_\kappa(\Lambda). \tag{2.14}$$

In particular, this implies for $Y = \boldsymbol{L}^{p,q}_\kappa(\Lambda)$ and any (symmetric) weight m with $m \geq m_\kappa$ that $\|K\|_{\mathcal{A}_{m,Y}} \leq \|K\|_{\mathcal{B}_m}$.

2.4 General Coorbit Spaces

In this subsection, we give a brief crash-course to general coorbit theory. Our treatment is essentially based on [66], but incorporates additional simplifications (from [60]) that are on the one hand due to using the kernel space $\mathcal{B}_m$ instead of $\mathcal{A}_{m,Y}$, and on the other hand due to imposing slightly more restrictive assumptions than in [66]. For the warped time-frequency systems that we consider, these assumptions are automatically satisfied, justifying this restriction.

To formulate our assumptions for the applicability of coorbit theory, we need one final ingredient.

Definition 2.10 Let $\mathcal{V} = (V_j)_{j \in J}$ be an arbitrary open covering of $\Lambda = \mathbb{R}^d \times D$. The *maximal kernel* $\mathrm{M}_\mathcal{V} K$ associated to a given kernel $K : \Lambda \times \Lambda \to \mathbb{C}$, given by

$$\begin{aligned} &\mathrm{M}_\mathcal{V} K : \quad \Lambda \times \Lambda \to [0, \infty], \quad (\lambda, \rho) \mapsto \sup_{\nu \in \mathbf{V}_\lambda} |K(\nu, \rho)| \quad \text{where} \\ &\mathbf{V}_\lambda := \bigcup_{j \in J \text{ with } \lambda \in V_j} V_j. \end{aligned} \tag{2.15}$$

In what follows, we shall always work in the following setting:

Assumption 2.11 Let $D \subset \mathbb{R}^d$ be open, and let $\Lambda = \mathbb{R}^d \times D$, equipped with the Borel σ-algebra and the Lebesgue measure μ. We assume that

1. $\mathcal{U} = (U_j)_{j \in J}$ is a product-admissible covering of Λ;
2. $u : \Lambda \to \mathbb{R}^+$ is continuous and $\mathcal{U}$-moderate;
3. $m_0 : \Lambda \times \Lambda \to \mathbb{R}^+$ is continuous and symmetric and satisfies $m_0(\lambda, \rho) \leq C^{(0)} \cdot u(\lambda)\, u(\rho)$ for all $\lambda, \rho \in \Lambda$ and some $C^{(0)} > 0$;
4. $\Psi = (\psi_\lambda)_{\lambda \in \Lambda}$ is a continuous Parseval frame for $\mathbf{L}^{2,\mathcal{F}}(D)$, and the map $\Lambda \to \mathbf{L}^2(\mathbb{R}^d)$, $\lambda \mapsto \psi_\lambda$ is continuous;
5. $v : \Lambda \to [1, \infty)$ is continuous and satisfies $v(\lambda) \geq c \cdot \max \big\{ \|\psi_\lambda\|_{\mathbf{L}^2},\ u(\lambda)/w_{\mathcal{U}}^c(\lambda) \big\}$ for some $c > 0$ and all $\lambda \in \Lambda$, with $w_{\mathcal{U}}^c$ as in Eq. (2.9);
6. $Y \subset \mathbf{L}^1_{\text{loc}}(\Lambda)$ is a rich, solid Banach space such that $\|K(\bullet)\|_{Y \to Y} \leq \|K\|_{\mathcal{B}_{m_0}}$ for all $K \in \mathcal{B}_{m_0}$;
7. The kernel K_Ψ defined in Eq. (2.4) satisfies

$$K_\Psi \in \mathcal{A}_{m_v} \quad \text{and} \quad \mathrm{M}_{\mathcal{U}} K_\Psi \in \mathcal{B}_{m_0}. \tag{2.16}$$

with m_v as defined in Eq. (2.10).

By Proposition 2.9, Condition (6) is satisfied for $Y = \mathbf{L}_\kappa^{p,q}(\Lambda)$, as long as $\frac{\kappa(\lambda)}{\kappa(\rho)} \leq m_0(\lambda, \rho)$ for all $\lambda, \rho \in \Lambda$.

Remark 2.12 If the kernel K is continuous in the second component (as is the case for the reproducing kernel K_Ψ, under the conditions in Assumption 2.11), then $\mathrm{M}_{\mathcal{U}} K$ is lower semicontinuous and hence measurable. To see this, let $\alpha \in \mathbb{R}$ and $(\lambda_0, \rho_0) \in \Lambda \times \Lambda$ with $\mathrm{M}_{\mathcal{U}} K(\lambda_0, \rho_0) > \alpha$. Then there are $j \in J$ with $\lambda_0 \in U_j$ and some $\nu \in U_j$ such that $|K(\nu, \rho_0)| > \alpha$. By continuity of $K(\nu, \bullet)$, there is thus an open set $V \subset \Lambda$ with $\rho_0 \in V$ and such that $|K(\nu, \rho)| > \alpha$ for all $\rho \in V$. Overall, we see for $(\lambda, \rho) \in U_j \times V$ that $\mathrm{M}_{\mathcal{U}} K(\lambda, \rho) \geq |K(\nu, \rho)| > \alpha$. Since $\mathcal{U}$ is a product-admissible covering, U_j is open; thus, we have shown that $\mathrm{M}_{\mathcal{U}} K$ is indeed lower semicontinuous.

The next theorem shows that the conditions in Assumption 2.11 ensure that one can extend the voice transform to a suitably defined space of distributions.

Theorem 2.13 *Under Assumption 2.11, the following hold: The space*

$$\mathcal{H}_v^1 := \mathcal{H}_v^1(\Psi) := \big\{ f \in \boldsymbol{L}^{2,\mathcal{F}}(D) \ :\ V_\Psi f \in \boldsymbol{L}_v^1 \big\}, \text{ with the norm } \|f\|_{\mathcal{H}_v^1} := \|V_\Psi f\|_{\boldsymbol{L}_v^1}, \tag{2.17}$$

is a Banach space satisfying $\mathcal{H}_v^1 \hookrightarrow \boldsymbol{L}^{2,\mathcal{F}}(D)$, with dense image. Furthermore, there is some $C' > 0$ such that $\|\psi_\lambda\|_{\mathcal{H}_v^1} \leq C' \cdot v(\lambda) < \infty$ for all $\lambda \in \Lambda$. In fact, $\mathcal{H}_v^1$ is the minimal Banach space with that property.

Finally, for each $f \in (\mathcal{H}_v^1)^\urcorner$, the extended voice transform

$$V_\Psi f : \quad \Lambda \to \mathbb{C}, \quad \lambda \mapsto \langle f, \psi_\lambda \rangle_{(\mathcal{H}_v^1)^\urcorner, \mathcal{H}_v^1} = f(\psi_\lambda) \tag{2.18}$$

satisfies $V_\Psi f \in \boldsymbol{L}_{1/v}^\infty(\Lambda)$. In fact, the expression $\|V_\Psi f\|_{\boldsymbol{L}_{1/v}^\infty}$ defines an equivalent norm on $(\mathcal{H}_v^1)^\urcorner$.

Proof Define $C' := \|K_\Psi\|_{\mathcal{A}_{m_v}}$. Then, [66, Lemma 2.13] shows that $\|\psi_\lambda\|_{\mathcal{H}_v^1} \leq C' \cdot v(\lambda)$ holds for all $\lambda \in \Lambda \setminus N$, if $\lambda \mapsto \psi_\lambda$ is weakly measurable. If $\lambda \mapsto \psi_\lambda$ and v are continuous, their proof is easily seen to hold pointwise for all $\lambda \in \Lambda$ and hence $\Psi \subset \mathcal{H}_v^1$. Since Ψ is a continuous frame for $\mathbf{L}^{2,\mathcal{F}}(D)$, this in particular implies that $\mathcal{H}_v^1 \subset \mathbf{L}^{2,\mathcal{F}}(D)$ is dense. The completeness of $(\mathcal{H}_v^1, \|\bullet\|_{\mathcal{H}_v^1})$ and the continuity of the embedding $\mathcal{H}_v^1 \hookrightarrow \mathbf{L}^{2,\mathcal{F}}(D)$ follow from [60, Lemma 8.1]. The minimality property of $\mathcal{H}_v^1$ is shown in [46, Corollary 1].

For $\varphi \in (\mathcal{H}_v^1)^\urcorner$, [60, Lemma 8.1] shows that $V_\Psi\varphi$ is measurable with respect to the *Lebesgue σ-algebra*, and that $\varphi \mapsto \|V_\Psi\varphi\|_{\mathbf{L}^\infty_{1/v}}$ defines an equivalent norm on $(\mathcal{H}_v^1)^\urcorner$. Thus, we only show that $V_\Psi\varphi$ is in fact measurable with respect to the *Borel σ-algebra*. To see this, define $W := \{V_\Psi f : f \in \mathcal{H}_v^1\} \subset \mathbf{L}_v^1(\Lambda)$ and $\gamma : W \to \mathbb{C}, V_\Psi f \mapsto \overline{\varphi(f)}$, noting that this is a well-defined, bounded linear functional since $|\gamma(V_\Psi f)| = |\varphi(f)| \leq C \cdot \|f\|_{\mathcal{H}_v^1} = C \cdot \|V_\Psi f\|_{\mathbf{L}_v^1}$. By combining the Hahn-Banach theorem with the characterization of the dual of $\mathbf{L}_v^1(\Lambda)$, we thus see that there exists $G \in \mathbf{L}^\infty_{1/v}(\Lambda)$ satisfying

$$V_\Psi\varphi(\lambda) = \varphi(\psi_\lambda) = \overline{\gamma(V_\Psi\psi_\lambda)} = \overline{\int_\Lambda G(\rho) V_\Psi\psi_\lambda(\rho)\, d\rho}.$$

Now, since $(\lambda, \rho) \mapsto V_\Psi\psi_\lambda(\rho) = \langle\psi_\lambda, \psi_\rho\rangle = K_\Psi(\rho, \lambda)$ is measurable and since $G \in \mathbf{L}^\infty_{1/v}$ and $V_\Psi\psi_\lambda \in \mathbf{L}_v^1$ (as shown above), the measurability of $V_\Psi\varphi$ is an easy consequence of the Fubini-Tonelli theorem (see [27, Proposition 5.2.1]). □

Now that we have constructed the "reservoir" $(\mathcal{H}_v^1)^\urcorner$, we can use it to define the coorbit space associated to the frame Ψ and a solid Banach space Y.

Theorem 2.14 *Suppose that Assumption 2.11 is satisfied. Then the* coorbit of Y with respect to Ψ*,*

$$\operatorname{Co} Y := \operatorname{Co}(\Psi, Y) := \left\{ f \in (\mathcal{H}_v^1)^\urcorner \ : \ V_\Psi f \in Y \right\}, \tag{2.19}$$

is a Banach space with natural norm $\|f\|_{\operatorname{Co} Y} := \|V_\Psi f\|_Y$.

Additionally, for any $G \in Y$, the property $G = K_\Psi(G)$ is equivalent to $G = V_\Psi f$ for some $f \in \operatorname{Co} Y$. The map $V_\Psi : \operatorname{Co} Y \to Y$ is an isometry of Co Y *onto the closed subspace $K_\Psi(Y)$ of Y. Finally, the inclusion* Co $Y \hookrightarrow (\mathcal{H}_v^1)^\urcorner$ *is continuous.*

Proof This follows from [60, Proposition 8.6] together with [66, Sections 2.3 and 2.4]. □

Note that the definition of Co Y is independent of the weight v in the following sense: If $\widetilde{v}$ is another weight such that Assumption 2.11 holds, then (2.19) defines the same space, see [66, Lemma 2.26]. Furthermore, according to [66, Lemma 2.32], we have the following special cases:

$$\operatorname{Co} \mathbf{L}_v^1 = \mathcal{H}_v^1, \quad \operatorname{Co} \mathbf{L}^\infty_{1/v} = (\mathcal{H}_v^1)^\urcorner \quad \text{and} \quad \operatorname{Co} \mathbf{L}^2 = \mathbf{L}^2.$$

The coorbit spaces $(\operatorname{Co} Y, \|\bullet\|_{\operatorname{Co} Y})$ are independent of the particular choice of the continuous frame Ψ, under a certain equivalence condition on the mixed kernel associated to a pair of continuous Parseval frames.

Proposition 2.15 *If Ψ and $\widetilde{\Psi}$ are continuous Parseval frames for $\boldsymbol{L}^{2,\mathcal{F}}(D)$ such that Assumption 2.11 is satisfied for Ψ and also for $\widetilde{\Psi}$, and if $K_{\Psi,\widetilde{\Psi}}, K_{\widetilde{\Psi},\Psi} \in \mathcal{A}_{m_v} \cap \mathcal{B}_{m_0}$, where*

$K_{\Psi,\widetilde{\Psi}}$ is the mixed *kernel defined by*

$$K_{\Psi,\widetilde{\Psi}}(\lambda, \rho) := \left\langle \widetilde{\psi_\rho}, \psi_\lambda \right\rangle \tag{2.20}$$

then

$$\mathcal{H}_v^1(\Psi) = \mathcal{H}_v^1(\widetilde{\Psi}) \quad \textit{and} \quad \operatorname{Co}(\Psi, Y) = \operatorname{Co}(\widetilde{\Psi}, Y).$$

Proof Assumption 2.11 implies $\mathcal{A}_{m_v} \cap \mathcal{B}_{m_0} \hookrightarrow \mathcal{A}_{m_v,Y}$. Thus, [66, Lemma 2.29] yields the claim. $\square$

2.5 Discretization in Coorbit Spaces

General coorbit theory provides a machinery for constructing Banach spaces $\operatorname{Co} Y$ and associated (Banach) frames and atomic decompositions through sampling of the continuous frame Ψ on Λ. The results summarized here have been developed by Fornasier and Rauhut [46] and extended in [11, 60, 61, 66, 78].

In a nutshell, the idea for discretizing the continuous frame Ψ is to consider a sufficiently fine covering $\mathcal{V} = (V_j)_{j\in J}$ such that the frame $\Psi = (\psi_\lambda)_{\lambda\in\Lambda}$ is *almost constant* (in a suitable sense) on each of the sets V_j. Then, by choosing $\lambda_j \in V_j$, it is intuitively plausible that the discrete family $(\psi_{\lambda_j})_{j\in J}$ behaves similarly to the continuous frame Ψ. The following definition makes this idea of Ψ being almost constant on each of the V_j more precise.

Definition 2.16 Let $\Gamma : \Lambda \times \Lambda \to S^1 \subset \mathbb{C}$ be continuous. The Γ-*oscillation* $\operatorname{osc}_{\mathcal{V},\Gamma} : \Lambda \times \Lambda \to [0, \infty)$ of a continuous Parseval frame $\Psi = (\psi_\lambda)_{\lambda\in\lambda}$ with respect to the topologically admissible covering $\mathcal{V} = (V_j)_{j\in J}$ of Λ is defined as

$$\begin{aligned} \operatorname{osc}_{\mathcal{V},\Gamma}(\lambda, \rho) := \operatorname{osc}_{\Psi,\mathcal{V},\Gamma}(\lambda, \rho) &:= \sup_{\nu\in\mathbf{V}_\varrho} |\langle \psi_\lambda, \psi_\rho - \Gamma(\rho, \nu)\psi_\nu\rangle| \\ &= \sup_{\nu\in\mathbf{V}_\varrho} |K_\Psi(\rho, \lambda) - \overline{\Gamma(\rho, \nu)} K_\Psi(\nu, \lambda)| \\ &= \sup_{\nu\in\mathbf{V}_\varrho} |K_\Psi(\lambda, \rho) - \Gamma(\rho, \nu) K_\Psi(\lambda, \nu)|, \end{aligned} \tag{2.21}$$

where $\mathbf{V}_\varrho := \bigcup_{j\in J \text{ with } \rho\in V_j} V_j$.

Remark 2.17 The oscillation $\operatorname{osc}_{\mathcal{V},\Gamma} : \Lambda \times \Lambda \to [0, \infty)$ is well-defined and lower semicontinuous and hence measurable. Indeed, each set $\mathbf{V}_\rho \subset \Lambda$ is relatively compact as a finite union of relatively compact sets, where finiteness of the union is implied by the remark after Definition 2.3. Next, note that K_Ψ is continuous, since the map $\lambda \mapsto \psi_\lambda$ is (strongly) continuous by Assumption 2.11. Since continuous functions are bounded on relatively compact sets, this shows that $\mathrm{osc}_{\mathcal{V},\Gamma}$ is finite-valued. Now proceed analogous to Remark 2.12.

We further consider specific sequence spaces associated to Y and a collection $\mathcal{W}$ of subsets of Λ.

Definition 2.18 For any family $\mathcal{W} = (W_j)_{j\in J}$ with a countable index set J and consisting of measurable subsets $W_j \subset \Lambda$ with $0 < \mu(W_j) < \infty$ and any sequence $c = (c_j)_{j\in J} \in \mathbb{C}^J$, we define

$$\|c\|_{Y^\flat(\mathcal{W})} := \Big\| \sum_{j\in J} |c_j| \mathbb{1}_{W_j} \Big\|_Y \in [0,\infty] \quad \text{and}$$
$$\|c\|_{Y^\sharp(\mathcal{W})} := \Big\| \sum_{j\in J} \frac{|c_j|}{\mu(W_j)} \mathbb{1}_{W_j} \Big\|_Y \in [0,\infty],$$

and finally

$$\begin{aligned} &Y^\flat(\mathcal{W}) := \{c \in \mathbb{C}^J : \|c\|_{Y^\flat(\mathcal{W})} < \infty\} \\ \text{and}\quad &Y^\sharp(\mathcal{W}) := \{c \in \mathbb{C}^J : \|c\|_{Y^\sharp(\mathcal{W})} < \infty\}. \end{aligned} \tag{2.22}$$

The following set of assumptions summarizes the conditions that ensure applicability of the discretization results from coorbit theory.

Assumption 2.19 In addition to Assumption 2.11, assume the following conditions:

1. $\mathcal{V} = (V_j)_{j\in J}$ is a topologically admissible covering of Λ;
2. $\Gamma : \Lambda \times \Lambda \to S^1$ is continuous;
3. With $m := \max\{m_0, m_v\}$, we have

$$\|\mathrm{osc}_{\mathcal{V},\Gamma}\|_{\mathcal{B}_m} \cdot (2\|K_\Psi\|_{\mathcal{B}_m} + \|\mathrm{osc}_{\mathcal{V},\Gamma}\|_{\mathcal{B}_m}) < 1;$$

Remark 2.20 If $\mathcal{W}$ in Definition 2.18 is identical to the topologically admissible covering $\mathcal{V} = (V_j)_{j\in J}$, we often write $Y^\flat$ and $Y^\sharp$ for $Y^\flat(\mathcal{V})$ or $Y^\sharp(\mathcal{V})$. In fact, it is often possible to choose the product-admissible covering $\mathcal{U}$ from Assumption 2.11 identical to the topologically admissible covering $\mathcal{V}$, and we will indeed do so, but this is not required. However, the oscillation of Ψ provides a useful, straightforward estimate for the maximal kernel associated to K_Ψ:

$$\mathrm{M}_{\mathcal{V}} K_\Psi(\lambda,\rho) \le |K_\Psi(\lambda,\rho)| + \mathrm{osc}^*_{\Psi,\mathcal{V},\Gamma}(\lambda,\rho), \text{ a.e.,} \tag{2.23}$$

for any choice of Γ. Hence, Assumption 2.19(3) implies the second part of Assumption 2.11(7) if $\mathcal{U} = \mathcal{V}$.

Remark 2.21 Note that an appropriate choice of the map $\Gamma : \Lambda \times \Lambda \to S^1$ is crucial to achieve small $\mathcal{B}_m$-norm of the oscillation and, consequently, for satisfying Item 3 above. In this work, we will only consider a single, straightforward choice for Λ and the map $\Gamma : \Lambda \times \Lambda \to S^1$, namely $\Lambda = \mathbb{R}^d \times D$, with $D \subset \mathbb{R}^d$ open, and $\Gamma : \big((y, \omega), (z, \eta)\big) \mapsto e^{-2\pi i \langle y-z,\omega\rangle}$, cf. Theorem 6.1. However, other continuous frames Ψ may require a different choice of Γ.

The following theorem shows that the preceding conditions indeed imply that suitably sampling the continuous frame Ψ produces a Banach frame decomposition of $\mathrm{Co}(Y)$.

Theorem 2.22 *If Assumption 2.19 holds and if for each $j \in J$ some $\lambda_j \in V_j$ is chosen, then the discrete frame $\Psi_d = (\psi_{\lambda_j})_{j\in J}$ forms a Banach frame decomposition for* $\mathrm{Co}(Y) = \mathrm{Co}(\Psi, Y)$*, with the sequence space $Y^\flat$ and $Y^\sharp$ taking the place of $B^\flat$ and $B^\sharp$.*

Proof This follows from [60, Proposition 8.7], by choosing $L := \mathrm{osc}_{\mathcal{V},\Gamma}$ and $\widetilde{\mathcal{U}} = \mathcal{V}$ and by noting that the *topologically admissible covering* $\mathcal{V}$ is admissible in the terminology of [60]. □

One strategy to satisfy the conditions of Theorem 2.22 is the construction of a parametrized family of topologically admissible coverings $\mathcal{V}^\delta$ such that

$$\|\mathrm{osc}_{\mathcal{V}^\delta,\Gamma}\|_{\mathcal{B}_m} \xrightarrow{\delta\to 0} 0. \tag{2.24}$$

Then, $\delta_0 > 0$ can be found such that Theorem 2.22 holds for the fixed frame Ψ and all $\mathcal{V}^\delta$ with $\delta \le \delta_0$.

In [78]—later generalized in [66, Theorem 2.50] —a complementary discretization result is introduced, which allows to derive Banach frame decompositions for all appropriate $\mathrm{Co}\, Y$ directly from (discrete) frames on the Hilbert space $\mathcal{H}$, obtained by sampling a continuous frame. This is an intriguing and important result, given that the explicit construction of frames for $\mathcal{H}$ by sampling a continuous frame is often straightforward, see, e.g., [62]. Although we do not consider this result in detail here, we would like to note that its adjustment to our setting is straightforward.

2.6 Sequence Spaces Associated to Mixed-Norm Lebesgue Spaces

In this subsection, we show for $Y = \mathbf{L}^{p,q}_\kappa(\Lambda)$ and under suitable conditions on the covering $\mathcal{W}$, that the coefficient spaces $Y^\flat(\mathcal{W})$ and $Y^\sharp(\mathcal{W})$ coincide with certain mixed-norm sequence spaces $\ell^{p,q}_{\widetilde{\kappa}}(J)$. Here, given a (countable) index set J of the form $J = J_1 \times J_2$, and any fixed discrete weight $\widetilde{\kappa}\colon J \to \mathbb{R}^+$, the space $\ell^{p,q}_{\widetilde{\kappa}}(J)$ consists of all sequences $c = (c_{\ell,k})_{(\ell,k)\in J} \in \mathbb{C}^J$ for which

$$\|c\|_{\ell^{p,q}_{\widetilde{\kappa}}(J)} := \Big\| k \mapsto \|\widetilde{\kappa}(\bullet, k)\, c_{\bullet,k}\|_{\ell^p(J_1)} \Big\|_{\ell^q(J_2)} < \infty. \tag{2.25}$$

Precisely, our result is as follows:

Lemma 2.23 *Let $J = J_1 \times J_2$ be a countable index set and $\mathcal{Q} = (Q_k)_{k\in J_2}$ an admissible covering of Λ_2. For each $k \in J_2$, let $\mathcal{P}_k = (P_{\ell,k})_{\ell\in J_1}$ be an admissible covering of Λ_1 such that $\sup_{k\in J_2} \mathcal{N}(\mathcal{P}_k) < \infty$. Define $\mathcal{U} = (U_{\ell,k})_{(\ell,k)\in J}$ by*

$$U_{\ell,k} := P_{\ell,k} \times Q_k, \text{ for all } (\ell, k) \in J. \tag{2.26}$$

If the weight function $\kappa : \Lambda \to \mathbb{R}^+$ satisfies

$$\kappa(\lambda_0)/\kappa(\lambda_1) \le C, \text{ for some } C > 0, \text{ all } \lambda_0, \lambda_1 \in U_{\ell,k} \text{ and all } (\ell, k) \in J = J_1 \times J_2, \tag{2.27}$$

then, for all $1 \le p, q \le \infty$,

$$\big(\boldsymbol{L}_\kappa^{p,q}(\Lambda)\big)^\flat(\mathcal{U}) = \ell^{p,q}_{\kappa_{\mathcal{U}}^\flat}(J) \quad \text{and} \quad \big(\boldsymbol{L}_\kappa^{p,q}(\Lambda)\big)^\sharp(\mathcal{U}) = \ell^{p,q}_{\kappa_{\mathcal{U}}^\sharp}(J), \quad \text{with equivalent norms.} \tag{2.28}$$

Here, $\big(\boldsymbol{L}_\kappa^{p,q}(\Lambda)\big)^\flat(\mathcal{U})$ and $\big(\boldsymbol{L}_\kappa^{p,q}(\Lambda)\big)^\sharp(\mathcal{U})$ are the spaces defined in (2.22) and the weights $\kappa_{\mathcal{U}}^\flat$ and $\kappa_{\mathcal{U}}^\sharp$ are given by

$$\begin{aligned}
\kappa_{\mathcal{U}}^\flat(\ell, k) &= [\mu_1(P_{\ell,k})]^{1/p} \cdot [\mu_2(Q_k)]^{1/q} \cdot \kappa_{\ell,k} \quad \text{and} \\
\kappa_{\mathcal{U}}^\sharp(\ell, k) &= [\mu_1(P_{\ell,k})]^{1/p-1} \cdot [\mu_2(Q_k)]^{1/q-1} \cdot \kappa_{\ell,k},
\end{aligned}$$

where $\kappa_{\ell,k} := \sup_{\lambda\in U_{\ell,k}} \kappa(\lambda)$ for all $(\ell, k) \in J$.

Proof We prove the assertion for $p, q < \infty$; the proof for the cases $p = \infty$ or $q = \infty$ is similar and hence omitted.

Note that if $\mathcal{V} = (V_j)_{j\in J}$ is an admissible covering of a set $\mathcal{O}$ and if $(a_j)_{j\in J} \in [0, \infty)^J$, then at most $\mathcal{N}(\mathcal{V})$ summands of the sum $\sum_{j\in J} a_j \mathbb{1}_{V_j}(x)$ are non-zero for each fixed $x \in \mathcal{O}$. Therefore, given any $r \in \mathbb{R}^+$, we have $\big(\sum_{j\in J} a_j \mathbb{1}_{V_j}(x)\big)^r \asymp \sum_{j\in J} a_j^r \mathbb{1}_{V_j}(x)$, where the implied constant only depends on r and on $\mathcal{N}(\mathcal{V})$.

Let $(c_{\ell,k})_{(\ell,k)\in J} \in \mathbb{C}^J$ and set $f_c(\lambda) := \sum_{\ell,k\in J} |c_{\ell,k}| \, \mathbb{1}_{U_{\ell,k}}(\lambda)$. The estimate from the preceding paragraph, first applied to $\mathcal{V} = \mathcal{Q}$, and then applied to $\mathcal{V} = \mathcal{P}_k$ for fixed $k \in J_2$, shows

$$\begin{aligned}
\big(f_c(\lambda)\big)^p &= \Big(\sum_{k\in J_2} \mathbb{1}_{Q_k}(\lambda_2) \sum_{\ell\in J_1} |c_{\ell,k}| \, \mathbb{1}_{P_{\ell,k}}(\lambda_1)\Big)^p \\
&\asymp \sum_{k\in J_2} \Big[\mathbb{1}_{Q_k}(\lambda_2)\Big(\sum_{\ell\in J_1} |c_{\ell,k}| \, \mathbb{1}_{P_{\ell,k}}(\lambda_1)\Big)^p\Big] \\
&\asymp \sum_{k\in J_2} \mathbb{1}_{Q_k}(\lambda_2) \sum_{\ell\in J_1} |c_{\ell,k}|^p \mathbb{1}_{P_{\ell,k}}(\lambda_1).
\end{aligned} \tag{2.29}$$

Furthermore, note that Eq. (2.27) implies $\kappa(\lambda) \asymp \kappa_{\ell,k}$ for $\lambda \in U_{\ell,k} = P_{\ell,k} \times Q_k$. Therefore, integrating the estimate (2.29) over $\lambda_1 \in \Lambda_1$, we see

$$\begin{aligned} g_c(\lambda_2) :=& \int_{\Lambda_1} \big(f_c(\lambda_1,\lambda_2)\cdot\kappa(\lambda_1,\lambda_2)\big)^p \, d\mu_1(\lambda_1) \\ &\asymp \sum_{k\in J_1} \mathbb{1}_{Q_k}(\lambda_2) \sum_{\ell\in J_1} |c_{\ell,k}|^p \int_{\Lambda_1} \big(\kappa(\lambda_1,\lambda_2)\big)^p \cdot \mathbb{1}_{P_{\ell,k}}(\lambda_1)\, d\mu_1(\lambda_1) \\ &\asymp \sum_{k\in J_1} \Big[\mathbb{1}_{Q_k}(\lambda_2) \sum_{\ell\in J_1} |c_{\ell,k}\cdot\kappa_{\ell,k}|^p \cdot \mu_1(P_{\ell,k})\Big]. \end{aligned}$$

Now, we again use the estimate from the beginning of the proof (for $\mathcal{V} = \mathcal{Q}$) to obtain

$$\begin{aligned} [g_c(\lambda_2)]^{q/p} &\asymp \bigg(\sum_{k\in J_1} \Big[\mathbb{1}_{Q_k}(\lambda_2) \sum_{\ell\in J_1} |c_{\ell,k}\cdot\kappa_{\ell,k}|^p \cdot \mu_1(P_{\ell,k})\Big]\bigg)^{\frac{q}{p}} \\ &\asymp \sum_{k\in J_1} \bigg[\mathbb{1}_{Q_k}(\lambda_2) \Big(\sum_{\ell\in J_1} |c_{\ell,k}\cdot\kappa_{\ell,k}|^p \cdot \mu_1(P_{\ell,k})\Big)^{\frac{q}{p}}\bigg]. \end{aligned}$$

Integrating this over $\lambda_2 \in \Lambda_2$, we finally see

$$\begin{aligned} \|c\|^q_{(\mathbf{L}^{p,q}_\kappa(\Lambda))^\flat(\mathcal{U})} &= \|f_c\|^q_{\mathbf{L}^{p,q}_\kappa(\Lambda)} = \int_{\Lambda_2} [g_c(\lambda_2)]^{q/p}\, d\mu_2(\lambda_2) \\ &\asymp \sum_{k\in J_1} \bigg[\mu_2(Q_k)\Big(\sum_{\ell\in J_1} |c_{\ell,k}\cdot\kappa_{\ell,k}|^p \cdot \mu_1(P_{\ell,k})\Big)^{\frac{q}{p}}\bigg] = \|c\|^q_{\ell^{p,q}_{\kappa^\flat_{\mathcal{U}}}}, \end{aligned}$$

which completes the proof for the identification of the space $\big(\mathbf{L}^{p,q}_\kappa(\Lambda)\big)^\flat(\mathcal{U})$.

The identification of $\big(\mathbf{L}^{p,q}_\kappa(\Lambda)\big)^\sharp(\mathcal{U})$ follows by substituting $c_{\ell,k}\mu(U_{\ell,k})^{-1}$ for $c_{\ell,k}$ everywhere in the derivations above. $\square$

Our proof of the above result relies heavily on the product structure of the covering $\mathcal{U}$ in (2.26). Although minor generalizations of the conditions placed on $\mathcal{U}$ are possible without significant complications, one cannot expect to recover a similar result without restrictions on $\mathcal{U}$. However, in our setting of warped time-frequency systems, product coverings as in (2.26) arise quite naturally and the result above is entirely sufficient.

3 Frequency-Adapted Tight Continuous Frames Through Warping

In this section, we define the class of warped time-frequency systems as tools for the analysis and synthesis of functions. The framework presented here generalizes the systems introduced in [61] to arbitrary dimensions. The basic properties presented in this section are proven analogous to the one-dimensional case, such that we only provide references.

As explained in the introduction, a warped time-frequency system generates a joint time-frequency representation in which the trade-off between time- and frequency-resolution at any given frequency position is governed by the associated frequency scale. That frequency scale is generated by the warping function.

Definition 3.1 Let $D \subset \mathbb{R}^d$ be open. A C^1 diffeomorphism $\Phi : D \to \mathbb{R}^d$ is called a *warping function*, if $\det(\mathrm{D}\Phi^{-1}(\tau)) > 0$ for all $\tau \in \mathbb{R}^d$ and if the associated weight function

$$w : \quad \mathbb{R}^d \to \mathbb{R}^+, \quad w(\tau) = \det(\mathrm{D}\Phi^{-1}(\tau)), \tag{3.1}$$

is w_0-moderate for some submultiplicative weight $w_0 : \mathbb{R}^d \to \mathbb{R}^+$.

Remark We note that w_0 is automatically locally bounded, as shown in [57, Theorem 2.1.4] and [94, Theorem 2.2.22].

Let us collect some basic results that are direct consequences of w being w_0-moderate. For the sake of brevity, set

$$A(\tau) := \mathrm{D}\Phi^{-1}(\tau) \qquad \forall\, \tau \in \mathbb{R}^d \tag{3.2}$$

for the remainder of this work. First, note that the chain rule—applied to the identity $\tau = \Phi(\Phi^{-1}(\tau))$ for $\tau \in \mathbb{R}^d$—yields

$$\mathrm{id} = \mathrm{D}\Phi(\Phi^{-1}(\tau)) \cdot A(\tau), \quad \text{i.e.,} \quad A(\tau) = [\mathrm{D}\Phi(\Phi^{-1}(\tau))]^{-1}. \tag{3.3}$$

In particular, we get (for arbitrary $\tau = \Phi(\xi)$) that $w(\Phi(\xi)) = \frac{1}{\det(\mathrm{D}\Phi(\xi))}$. Thus, given any measurable nonnegative $f : \mathbb{R}^d \to [0, \infty)$, a change of variables leads to the frequently useful formulae

$$\begin{aligned} &\int_D f(\Phi(\xi))\, d\xi = \int_{\mathbb{R}^d} w(\tau) \cdot f(\tau)\, d\tau\,, \quad \text{and consequently} \\ &\|f \circ \Phi\|_{\mathbf{L}^p(D)} = \|f\|_{\mathbf{L}^p_{w^{1/p}}}. \end{aligned} \tag{3.4}$$

Finally, we note that submultiplicativity of w_0 and w_0-moderateness of w yields translation invariance of $\mathbf{L}^p_{w^{1/p}}$ and $\mathbf{L}^p_{w_0^{1/p}}$. Indeed, if w is any w_0-moderate weight (not necessarily given by (3.1)), then $w(\tau + \iota) \leq \min\{w(\tau)w_0(\iota), w(\iota)w_0(\tau)\}$, so that (3.4) yields

$$\|\mathbf{T}_\iota f\|^p_{\mathbf{L}^p_{w^{1/p}}} \leq w_0(\iota) \cdot \|f\|^p_{\mathbf{L}^p_{w^{1/p}}} \qquad \text{and} \qquad \|\mathbf{T}_\iota f\|^p_{\mathbf{L}^p_{w^{1/p}}} \leq w(\iota) \cdot \|f\|^p_{\mathbf{L}^p_{w_0^{1/p}}} \tag{3.5}$$

for all measurable $f : \mathbb{R}^d \to \mathbb{C}$ and all w_0-moderate weights w. In particular, one can choose $w = w_0$, since w_0 is submultiplicative and hence w_0-moderate.

Moderateness (and positivity) of the weight function w associated to the warping function Φ ensure that warped time-frequency systems and the associated representations are well-defined and possess some essential properties, as we will see shortly. But first, let us formally introduce warped time-frequency systems.

Definition 3.2 Let Φ be a warping function and $\theta \in \mathbf{L}^2_{\sqrt{w}}(\mathbb{R}^d)$. The *(continuous) warped time-frequency system* generated by θ and Φ is the collection of functions $\mathcal{G}(\theta, \Phi) := (g_{y,\omega})_{(y,\omega)\in\Lambda}$, where

$$g_{y,\omega} := \mathbf{T}_y \widecheck{g_\omega}, \quad \text{with} \quad g_\omega := w(\Phi(\omega))^{-1/2} \cdot (\mathbf{T}_{\Phi(\omega)}\theta) \circ \Phi \text{ for all } y \in \mathbb{R}^d,\ \omega \in D. \tag{3.6}$$

Here, the function $g_\omega : D \to \mathbb{C}$ is extended by zero to a function on all of $\mathbb{R}^d$, so that $\widecheck{g_\omega}$ is well-defined. The *phase space* associated with this family is $\Lambda = \mathbb{R}^d \times D$.

Since w is moderate with respect to w_0, we obtain $g_{y,\omega} \in \mathbf{L}^{2,\mathcal{F}}(D)$. In fact, (3.4) and (3.5) show

$$\|\widehat{g_{y,\omega}}\|^2_{\mathbf{L}^2(D)} \leq \frac{w_0(\Phi(\omega))}{w(\Phi(\omega))} \|\theta\|^2_{\mathbf{L}^2_{\sqrt{w}}(\mathbb{R}^d)} < \infty \quad \text{and} \quad \|\widehat{g_{y,\omega}}\|^2_{\mathbf{L}^2(D)} \leq \|\theta\|^2_{\mathbf{L}^2_{\sqrt{w_0}}} \in [0, \infty]. \tag{3.7}$$

Thus, $\mathcal{G}(\theta, \Phi) \subset \mathbf{L}^{2,\mathcal{F}}(D)$ and the associated analysis operation, i.e., taking inner products with the functions $g_{y,\omega}$, defines a transform on $\mathbf{L}^{2,\mathcal{F}}(D)$.

Definition 3.3 Let Φ be a warping function and $\theta \in \mathbf{L}^2_{\sqrt{w}}(\mathbb{R}^d)$. The Φ-*warped time-frequency transform* of $f \in \mathbf{L}^{2,\mathcal{F}}(D)$ with respect to the *prototype* θ is defined as

$$V_{\theta,\Phi} f : \quad \mathbb{R}^d \times D \to \mathbb{C}, \quad (y, \omega) \mapsto \langle f, g_{y,\omega}\rangle_{\mathbf{L}^2(\mathbb{R}^d)}. \tag{3.8}$$

For $\lambda = (y, \omega) \in \Lambda = \mathbb{R}^d \times D$, we will alternatively use the notations $V_{\theta,\Phi} f(y, \omega) = V_{\theta,\Phi} f(\lambda)$ and $g_{y,\omega} = g_\lambda$, whenever one or the other is more convenient.

By definition and (3.7), we have $V_{\theta,\Phi} f \in \mathbf{L}^\infty(\Lambda)$, whenever $\theta \in \mathbf{L}^2_{\sqrt{w_0}}(\mathbb{R}^d)$. Furthermore, using that $\Phi \in \mathcal{C}^1$ and the translation-invariance of $\mathbf{L}^2_{\sqrt{w}}(\mathbb{R}^d)$, one can also deduce that $V_{\theta,\Phi} f \in \mathcal{C}(\Lambda)$, even under the weaker assumption $\theta \in \mathbf{L}^2_{\sqrt{w}}(\mathbb{R}^d)$.

Proposition 3.4 *Let Φ be a warping function and $\theta \in \boldsymbol{L}^2_{\sqrt{w}}(\mathbb{R}^d)$. Then*

$$V_{\theta,\Phi} f \in \mathcal{C}(\Lambda), \quad \textit{for all } f \in \boldsymbol{L}^{2,\mathcal{F}}(D). \tag{3.9}$$

In fact, the mapping $\mathbb{R}^d \times D \to \boldsymbol{L}^{2,\mathcal{F}}(D)$, $(y, \omega) \mapsto g_{y,\omega}$ is continuous.

Proof Analogous to the proof of [61, Proposition 4.5]. □

The next result provides the crucial property that makes warped time-frequency systems so attractive. Namely, $V_{\bullet,\Phi}$ possesses a norm-preserving property similar to the orthogonality relations (Moyal's formula [53, 72]) for the short-time Fourier transform.

Theorem 3.5 *Let Φ be a warping function and $\theta_1, \theta_2 \in \boldsymbol{L}^2_{\sqrt{w}} \cap \boldsymbol{L}^2(\mathbb{R}^d)$. Then the following holds for all $f_1, f_2 \in \boldsymbol{L}^{2,\mathcal{F}}(D)$:*

$$\int_\Lambda V_{\theta_1,\Phi} f_1(\lambda) \overline{V_{\theta_2,\Phi} f_2(\lambda)}\, d\lambda = \langle f_1, f_2 \rangle_{\boldsymbol{L}^2(\mathbb{R}^d)} \langle \theta_2, \theta_1 \rangle_{\boldsymbol{L}^2(\mathbb{R}^d)}. \tag{3.10}$$

In particular, for any $\theta \in \boldsymbol{L}^2_{\sqrt{w}} \cap \boldsymbol{L}^2(\mathbb{R}^d)$, $\mathcal{G}(\theta, \Phi)$ is a continuous tight frame with frame bound $\|\theta\|^2_{\boldsymbol{L}^2}$.

Proof Analogous to [61, Theorem 4.6]. Note that $\theta_1, \theta_2 \in \mathbf{L}^2(\mathbb{R}^d)$ implies the admissibility condition required there, and moreover serves to justify the application of Plancherel's theorem in the proof. □

As already remarked in [61], $\theta_1, \theta_2 \in \mathbf{L}^2(\mathbb{R}^d)$ is a sort of admissibility condition and, in fact, yields the classical wavelet admissibility, if $d = 1$ and $\Phi = \log$. Besides the tight frame property, Theorem 3.5 shows that the warped time-frequency representations with respect to orthogonal windows, but the same warping function, span orthogonal subspaces of $\mathbf{L}^2(\Lambda)$. Similarly, orthogonal functions f_1, f_2 have orthogonal representations, independent of the prototypes θ_1, θ_2. These additional properties are useful, e.g., for constructing superframes for multiplexing [9, 54] or multitapered representations [33, 87, 97].

The tight frame property itself is a basic requirement for general coorbit theory, and provides a convenient inversion formula:

Corollary 3.6 *Given a warping function Φ and some nonzero $\theta \in \boldsymbol{L}^2_{\sqrt{w}} \cap \boldsymbol{L}^2(\mathbb{R}^d)$. Then any $f \in \boldsymbol{L}^{2,\mathcal{F}}(D)$ can be reconstructed from $V_{\theta,\Phi} f$ by*

$$f = \frac{1}{\|\theta\|^2_{\boldsymbol{L}^2}} \int_\Lambda V_{\theta,\Phi} f(\lambda)\, g_\lambda\, d\lambda. \tag{3.11}$$

The equation holds in the weak sense.

Proof The assertion is a direct consequence of $\mathcal{G}(\theta, \Phi)$ being a tight continuous frame with bound $\|\theta\|^2_{\mathbf{L}^2}$. □

Now that the essential properties of warped time-frequency systems are established, and before proceeding to construct and examine coorbit spaces associated to warped time-frequency systems, we provide some instructive examples of warping functions and the resulting warped time-frequency systems.

3.1 Examples

We present several examples of warping functions. We begin by constructing a d-dimensional function as a separable (coordinate-wise) combination of 1-dimensional warping functions. Examples of such 1-dimensional warping functions can be found in [61].

Separable warping. Fix $\mathcal{C}^1$-diffeomorphisms $\Phi_i : D_i \to \mathbb{R}$, $i \in \underline{d}$, such that $\mathrm{D}\,\Phi_i^{-1}(\tau) = \frac{\partial}{\partial\tau}\Phi_i^{-1}(\tau) > 0$, for all $\tau \in \mathbb{R}$, $i \in \underline{d}$. If each $\mathrm{D}\Phi_i^{-1}$, $i \in \underline{d}$, is $w_{0,i}$-moderate and we take Φ to be defined as

$$\Phi(\xi) = \big(\Phi_1(\xi_1), \dots, \Phi_d(\xi_d)\big)^T, \quad \text{for all} \quad \xi \in D := D_1 \times \cdots \times D_d,$$

then clearly $\Phi : D \to \mathbb{R}^d$ is a diffeomorphism and $\mathrm{D}\Phi^{-1}$ is diagonal, and hence

$$w(\tau) = \det(\mathrm{D}\Phi^{-1}(\tau)) = \prod_{i \in \underline{d}} \mathrm{D}\Phi_i^{-1}(\tau_i) > 0 \qquad \forall\, \tau \in \mathbb{R}^d,$$

and w is w_0-moderate for $w_0(\tau) := \prod_{i \in \underline{d}} w_{0,i}(\tau_i)$.

A family of anisotropic wavelets can be constructed by selecting $\Phi = \mathbf{log}$, where $\mathbf{log} : (\mathbb{R}^+)^d \to \mathbb{R}^d$ denotes the map $\xi \mapsto (\log(\xi_1), \dots, \log(\xi_d))^T$. It follows that Φ^{-1} is the componentwise exponential function and satisfies

$$\mathrm{D}\Phi_i^{-1}(\tau) = \operatorname{diag}(e^{\tau_1}, \dots, e^{\tau_d}) \quad \text{and} \quad w(\tau) = \exp(\tau_1 + \cdots + \tau_d),$$

for all $\tau \in \mathbb{R}^d$. Hence, w is submultiplicative and moderate with respect to itself. Furthermore, writing $\mathbf{d}(\omega) := \operatorname{diag}(\omega_1, \dots, \omega_d) \in \mathbb{R}^{d\times d}$ for $\omega \in (\mathbb{R}^+)^d$, we see that the elements of $\mathcal{G}(\theta, \Phi)$ are given by

$$\begin{aligned} g_{y,\omega} &= w(\Phi(\omega))^{-1/2} \cdot \mathbf{T}_y \mathcal{F}^{-1}\left((\mathbf{T}_{\mathbf{log}(\omega)}\theta) \circ \mathbf{log}\right) \\ &= \det(\mathbf{d}(\omega))^{-1/2} \cdot \mathbf{T}_y \mathcal{F}^{-1}\left(\theta \circ \mathbf{log}([\mathbf{d}(\omega)]^{-1}\langle\cdot\rangle)\right) \\ &= \det(\mathbf{d}(\omega))^{1/2} \cdot \left[\mathcal{F}^{-1}\left(\theta \circ \mathbf{log}\right)\right](\mathbf{d}(\omega)\langle\cdot - y\rangle) \\ &= \det(\mathbf{d}(\omega))^{1/2} \cdot \widetilde{g}(\mathbf{d}(\omega)\langle\cdot - y\rangle), \text{ with } \widetilde{g} := \mathcal{F}^{-1}\big(\theta \circ \mathbf{log}\big). \end{aligned}$$

Thus, $\mathcal{G}(\theta, \Phi)$ is a wavelet system in the sense of [15, 47], with the dilation group given by the diagonal $d \times d$-matrices with entries in $\mathbb{R}^+$. The derivations above do not seem to generalize, however, to a setting that recovers wavelets with respect to general dilation groups. Finally, the expression of $g_{y,\omega}$ through linear operators applied to a single *mother wavelet* $\widetilde{g}$ defined in the time-domain relies on properties of the coordinate-wise logarithm $\mathbf{log}$ and does not generalize to arbitrary warping functions Φ.

Radial warping. By choosing the warping function Φ to be radial, we can construct time-frequency systems with frequency resolution depending on the modulus $|\xi|$ of $\xi \in \mathbb{R}^d$. The deformation is then fixed on any $(d-1)$-sphere of fixed radius, similar to isotropic wavelets (see [32, Section 2.6] and [48, Example 2.30]). Generally, radial warping functions are of the form

$$\Phi_\varrho : \quad \mathbb{R}^d \to \mathbb{R}^d, \quad \xi \mapsto \varrho(|\xi|) \cdot \xi/|\xi|,$$

for a strictly increasing diffeomorphism $\varrho : \mathbb{R} \to \mathbb{R}$. Under suitable additional assumptions on ϱ, it can then be shown that $(\Phi_\varrho)^{-1} = \Phi_{\varrho^{-1}}$ and that Φ_ϱ is a warping function as per Definition 3.1. It will be shown in future work that radial warping does *not* recover isotropic wavelets exactly in dimensions $d > 1$, for any choice of ϱ, but that warped time-frequency systems can be close to isotropic wavelets in a sense that will be made formal in the mentioned follow-up work. An in depth study of radial warping with some specific examples is provided in Sect. 8.

An explicit, exotic example for $d = 2$. To demonstrate that there is potential for warping functions beyond the separable and radial cases, consider the continuous $\mathcal{C}^1$-diffeomorphism

$$\Phi : \quad \mathbb{R}^2 \to \mathbb{R}^2, \quad \xi \mapsto \left(e^{\xi_2}\, \xi_1,\ \xi_2\right)^T.$$

It is straightforward to see that Φ is a diffeomorphism with inverse $\Phi^{-1}(\tau) = (e^{-\tau_2}\, \tau_1,\ \tau_2)^T$, which satisfies

$$\mathrm{D}\Phi^{-1}(\tau) = \begin{pmatrix} e^{-\tau_2} & -e^{-\tau_2}\,\tau_1 \\ 0 & 1 \end{pmatrix}$$

and hence $w(\tau) = \det(\mathrm{D}\Phi^{-1}(\tau)) = e^{-\tau_2} > 0$. Moreover, it is easy to see that w is multiplicative (and in particular submultiplicative) and hence self-moderate. Thus, Φ is a valid warping function that is neither separable nor radial.

4 Membership of the Reproducing Kernel in $\mathcal{B}_m$

As we saw in Sect. 2.4 (see in particular Assumption 2.11), the main challenge in verifying the applicability of coorbit theory for a continuous Parseval frame Ψ lies in showing that (the maximal function of) the reproducing kernel K_Ψ is contained $\mathcal{A}_{m_v}$ or $\mathcal{B}_{m_0}$, for suitable weights $m_v, m_0 : \Lambda \times \Lambda \to \mathbb{R}^+$. We will do so in two steps: (1) In the present section, we will derive verifiable conditions on the warping function Φ and the prototype function θ which ensure that the warped time-frequency system $\Psi = \mathcal{G}(\theta, \Phi)$ satisfies $K_\Psi \in \mathcal{B}_m$, for a weight m satisfying suitable assumptions. (2) In Sect. 6, we do the same for the Γ-oscillation of Ψ and additionally demonstrate that $\|\mathrm{osc}_{\mathcal{V},\Gamma}\|_{\mathcal{B}_m}$ can be made arbitrarily small by choosing an appropriate covering $\mathcal{V}$. Then, the desired properties of the maximal kernel $\mathrm{M}_{\mathcal{V}} K_\Psi$ are a consequence of Remark 2.20.

To prepare for the treatment of the Γ-oscillation, we already consider mixed kernels in the present section. This setting only requires little additional effort. We begin by introducing some notation and conditions that will be used throughout this section.

Notation & Definition 4.1 By Φ, we denote a warping function $\Phi : D \to \mathbb{R}^d$, with associated weights w, w_0 as in Definition 3.1, $A = \mathrm{D}\Phi^{-1}$ and $\Lambda := \mathbb{R}^d \times D$. In all instances, $\theta, \theta_1, \theta_2 \in \mathbf{L}^2_{\sqrt{w_0}}(\mathbb{R}^d)$ and we denote the mixed kernel associated with $\mathcal{G}(\theta_1, \Phi)$ and $\mathcal{G}(\theta_2, \Phi)$ by $K_{\theta_1,\theta_2} := \underbrace{K}_{\mathcal{G}(\theta_1,\Phi),\mathcal{G}(\theta_2,\Phi)}$. Finally, for $\ell \in \{1, 2\}$, we write $\mathcal{G}(\theta_\ell, \Phi) = \big(g^{[\ell]}_{y,\omega}\big)_{y\in\mathbb{R}^d,\omega\in D} = \big(\mathbf{T}_y\, g^{[\ell]}_{\omega}\big)_{y\in\mathbb{R}^d,\omega\in D}$.

1. If there is a continuous function $m^\Phi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^+$ satisfying

$$m\big((x, \Phi^{-1}(\sigma)), (y, \Phi^{-1}(\tau))\big) \leq m^\Phi(x - y, \sigma - \tau) \qquad \forall\, x, y, \sigma, \tau \in \mathbb{R}^d, \tag{4.1}$$

then we say that m is Φ-*convolution-dominated* (by m^Φ). If that is the case, we denote by $M : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^+$ the weight

$$M(x, \tau) := \sup_{y\in\mathbb{R}^d, |y|\leq R_\Phi|x|} \Big[\sqrt{w_0(\tau)}\, m^\Phi(y, \tau)\Big] \quad \text{where}$$
$$R_\Phi := \sup_{\xi\in D} \|\mathrm{D}\Phi(\xi)\| \in \mathbb{R}^+ \cup \{\infty\}. \tag{4.2}$$

2. If there exists an m^Φ as in (1), such that m is Φ-convolution-dominated by m^Φ and

$$R_\Phi < \infty \qquad \text{or} \qquad m^\Phi(x, \sigma) \lesssim m^\Phi(0, \sigma) \text{ for all } x, \sigma \in \mathbb{R}^d, \tag{4.3}$$

then we say that m is Φ-*compatible* (with dominating weight m^Φ).

Furthermore, we require a slightly stricter and more structured notion of regularity for warping functions.

Definition 4.2 Let $\varnothing \neq D \subset \mathbb{R}^d$ be an open set and fix an integer $k \in \mathbb{N}_0$. A map $\Phi : D \to \mathbb{R}^d$ is a *k-admissible warping function* with *control weight* $v_0 : \mathbb{R}^d \to \mathbb{R}^+$, if v_0 is continuous, submultiplicative and radially increasing and Φ satisfies the following assumptions:

- Φ is a C^{k+1}-diffeomorphism.
- $A = \mathrm{D}\Phi^{-1}$ has positive determinant.
- With

$$\phi_\tau\,(\iota) := \Big(A^{-1}(\tau)A(\iota + \tau)\Big)^T = A^T(\iota + \tau) \cdot A^{-T}(\tau), \tag{4.4}$$

we have

$$\big\|\partial^\alpha \phi_\tau\,(\iota)\big\| \leq v_0(\iota) \qquad \text{for all } \tau, \iota \in \mathbb{R}^d \text{ and all multiindices } \alpha \in \mathbb{N}_0^d,\ \ |\alpha| \leq k. \tag{4.5}$$

Remark 4.3 1) The function ϕ_τ describes the regularity of A around τ; its relevance will become clear before long, see Eq. (4.27) below.

2) On the right-hand side of (4.5), one could allow constants C_α and different weights $\tilde{v}_\alpha$ not necessarily being radially increasing, therefore obtaining tighter bounds on $\|\partial^\alpha \phi_\tau(\iota)\|$. However, whenever such C_α, $\tilde{v}_\alpha$ exist, there also exists a weight v_0 satisfying all the requirements of Definition 4.2.
3) We remark that (4.5) generalizes the conditions mentioned in [61], even for the case $d = 1$ considered there.

Theorem 4.4 below shows that smoothness of the prototypes θ_1, θ_2 and decay (or localization) of their partial derivatives implies $K_{\theta_1,\theta_2} \in \mathcal{B}_m$, provided that m is Φ-compatible. In particular, all conditions are surely satisfied for arbitrary $\theta_1, \theta_2 \in \mathcal{C}_c^\infty(\mathbb{R}^d)$. The proof of Theorem 4.4 is deferred to the end of the section.

Theorem 4.4 *Let Φ be a $(d+p+1)$-admissible warping function with control weight v_0, where $p = 0$ if $R_\Phi = \infty$, defined as in (4.2), and $p \in \mathbb{N}_0$ otherwise. Let furthermore $m : \Lambda \times \Lambda \to \mathbb{R}^+$ be a symmetric weight satisfying*

$$m\big((y,\xi),(z,\eta)\big) \leq (1+|y-z|)^p \cdot v_1\big(\Phi(\xi)-\Phi(\eta)\big), \textit{ for all } y,z \in \mathbb{R}^d \textit{ and } \xi,\eta \in D, \tag{4.6}$$

for some continuous and submultiplicative weight $v_1 : \mathbb{R}^d \to \mathbb{R}^+$ satisfying $v_1(\iota) = v_1(-\iota)$ for all $\iota \in \mathbb{R}^d$.

Finally, with

$$w_2 : \quad \mathbb{R}^d \to \mathbb{R}^+, \quad \iota \mapsto (1+|\iota|)^{d+1} \cdot v_1(\iota) \cdot [v_0(\iota)]^{9d/2+3p+3},$$

assume that $\theta_1, \theta_2 \in \mathcal{C}^{d+p+1}(\mathbb{R}^d)$ and

$$\frac{\partial^n}{\partial \iota_j^n}\theta_\ell \in \boldsymbol{L}^2_{w_2}(\mathbb{R}^d), \qquad \textit{for all } j \in \underline{d},\ \ell \in \{1,2\},\ 0 \leq n \leq d+p+1\,,$$

and let

$$C_{\max} := C_{\max}(k,\theta_1,\theta_2) := \prod_{\ell\in\{1,2\}} \left(\max_{j\in\underline{d}} \max_{0\leq n\leq d+p+1} \left\| \frac{\partial^n}{\partial \iota_j^n}\theta_\ell \right\|_{\boldsymbol{L}^2_{w_2}(\mathbb{R}^d)} \right). \tag{4.7}$$

Then, m is Φ-compatible with dominating weight $m^\Phi(x,\tau) = (1+|x|)^p \cdot v_1(\tau)$ and there is a constant $C > 0$, independent of θ_1, θ_2 and m, satisfying

$$\big\|K_{\theta_1,\theta_2}\big\|_{\mathcal{B}_m} \leq C \cdot C_{\max} < \infty.$$

4.1 Bounding $\|K_{\theta_1,\theta_2}\|_{\mathcal{B}_m}$ Via Fourier Integral Operators

Towards an explicit estimate for $\|K_{\theta_1,\theta_2}\|_{\mathcal{B}_m}$, the next result provides an estimate in terms of families of Fourier integral operators [38, 39, 63, 85] dependent on θ_1, θ_2.

Here and in the following, we use e_τ, $\tau \in \mathbb{R}^d$, as short-hand for the (Φ-dependent) map

$$e_\tau : \quad \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}, \quad (x, \iota) \mapsto e^{-2\pi i \langle A^{-T}(\tau)\langle x \rangle, \Phi^{-1}(\iota+\tau)\rangle}. \tag{4.8}$$

Theorem 4.5 *Define*

$$L_{\tau_0}^{(\ell)}(x,\tau) := L_{\tau_0}[\theta_\ell, \theta_{3-\ell}](x,\tau) := \int_{\mathbb{R}^d} \frac{w(\iota+\tau_0)}{w(\tau_0)} \cdot (\theta_{3-\ell} \cdot \overline{\boldsymbol{T}_\tau \theta_\ell})(\iota) \cdot e_{\tau_0}(x,\iota)\, d\iota\,, \tag{4.9}$$

for $\ell \in \{1,2\}$ and $x, \tau, \tau_0 \in \mathbb{R}^d$. If m is Φ-compatible with dominating weight m^Φ, then we have

$$\|K_{\theta_1,\theta_2}\|_{\mathcal{B}_m} \le \max_{\ell\in\{1,2\}} \left[\operatorname*{ess\,sup}_{\eta\in D} \left\| L_{\Phi(\eta)}^{(\ell)} \right\|_{\boldsymbol{L}_M^1(\mathbb{R}^d\times\mathbb{R}^d)} \right], \tag{4.10}$$

with M as in (4.2). In particular, if $\operatorname{ess\,sup}_{\tau_0\in\mathbb{R}^d} \left\| L_{\tau_0}^{(\ell)} \right\|_{\boldsymbol{L}_M^1(\mathbb{R}^d\times\mathbb{R}^d)} < \infty$*, for $\ell \in \{1,2\}$, then $\|K_{\theta_1,\theta_2}\|_{\mathcal{B}_m}$ is finite.*

We prove Theorem 4.5 by means of two intermediate results. First, an (elementary) lemma concerned with the $\mathcal{B}_m$-norm of K_{θ_1,θ_2}.

Lemma 4.6 *If m is Φ-convolution-dominated by m^Φ, we have*

$$\begin{aligned} &\|K_{\theta_1,\theta_2}\|_{\mathcal{B}_m} \\ &\quad \le \max_{\ell\in\{1,2\}} \Bigg[\operatorname*{ess\,sup}_{\eta\in D} \int_D \operatorname*{ess\,sup}_{z\in\mathbb{R}^d} \int_{\mathbb{R}^d} m^\Phi(y-z, \Phi(\omega)-\Phi(\eta)) \\ &\qquad \cdot |K_{\theta_\ell,\theta_{3-\ell}}((y,\omega),(z,\eta))|\, dy\, d\omega \Bigg]. \end{aligned} \tag{4.11}$$

Proof If we define $\widetilde{m}^\Phi(x,\tau) := \min\{m^\Phi(x,\tau), m^\Phi(-x,-\tau)\}$, the symmetry of m easily shows that (4.1) also holds for $\widetilde{m}^\Phi$ instead of m^Φ. Hence, we can assume in what follows that m^Φ satisfies $m^\Phi(-x,-\tau) = m^\Phi(x,\tau)$ for all $x, \tau \in \mathbb{R}^d$.

For $\ell \in \{1,2\}$ and $\omega, \eta \in D$, define

$$B_\ell(\omega,\eta) := \operatorname*{ess\,sup}_{z\in\mathbb{R}^d} \int_{\mathbb{R}^d} m^\Phi\big(y-z, \Phi(\omega)-\Phi(\eta)\big) \cdot \left| K_{\theta_\ell,\theta_{3-\ell}}\big((y,\omega),(z,\eta)\big) \right| dy,$$

and let $C := \max_{\ell\in\{1,2\}} \operatorname{ess\,sup}_{\eta\in D} \int_D B_\ell(\omega,\eta)\, d\omega$, which is precisely the right-hand side of the target inequality. Equation (4.1) yields

$$\operatorname*{ess\,sup}_{z\in\mathbb{R}^d} \int_{\mathbb{R}^d} \left| (m \cdot K_{\theta_1,\theta_2})\big((y,\omega),(z,\eta)\big) \right| dy \le B_1(\omega,\eta). \tag{4.12}$$

Next, note that $\langle \mathbf{T}_y f_1, \mathbf{T}_z f_2 \rangle = \langle \mathbf{T}_{-z} f_1, \mathbf{T}_{-y} f_2 \rangle$ and $\langle f_1, f_2 \rangle = \overline{\langle f_2, f_1 \rangle}$ for all $f_1, f_2 \in \mathbf{L}^2(\mathbb{R}^d)$. Based on these identities and the translation-invariant structure of warped time-frequency systems, we see

$$\begin{aligned} K_{\theta_1,\theta_2}\big((y,\omega),(z,\eta)\big) &= K_{\theta_1,\theta_2}\big((-z,\omega),(-y,\eta)\big) \\ &= \overline{K_{\theta_2,\theta_1}\big((-y,\eta),(-z,\omega)\big)} = \overline{K_{\theta_2,\theta_1}\big((z,\eta),(y,\omega)\big)}. \end{aligned} \tag{4.13}$$

Using these identities and renaming $\widetilde{z} = -y$ and $\widetilde{y} = -z$, we see

$$\begin{aligned} &\operatorname*{ess\,sup}_{y\in\mathbb{R}^d} \int_{\mathbb{R}^d} \big|(m \cdot K_{\theta_1,\theta_2})\big((y,\omega),(z,\eta)\big)\big| \, dz \\ &\leq \operatorname*{ess\,sup}_{y\in\mathbb{R}^d} \int_{\mathbb{R}^d} m^\Phi\big((-z)-(-y), \Phi(\omega)-\Phi(\eta)\big) \cdot \big|K_{\theta_1,\theta_2}\big((-z,\omega),(-y,\eta)\big)\big| \, dz \\ &= \operatorname*{ess\,sup}_{\widetilde{z}\in\mathbb{R}^d} \int_{\mathbb{R}^d} m^\Phi\big(\widetilde{y}-\widetilde{z}, \Phi(\omega)-\Phi(\eta)\big) \cdot \big|K_{\theta_1,\theta_2}\big((\widetilde{y},\omega),(\widetilde{z},\eta)\big)\big| \, d\widetilde{y} \\ &= B_1(\omega,\eta). \end{aligned} \tag{4.14}$$

Combining (4.12) and (4.14), we see with notation as in (2.13) that

$$\big\|(m \cdot K_{\theta_1,\theta_2})^{(\omega,\eta)}\big\|_{\mathcal{A}_1} \leq B_1(\omega,\eta) \qquad \forall\, \omega,\eta \in D.$$

A simple calculation using (4.13) and the symmetry $m^\Phi(-x,-\tau) = m^\Phi(x,\tau)$ proves the identity
$B_1(\omega,\eta) = B_2(\eta,\omega)$. Overall, we thus see

$$\begin{aligned} \|m \cdot K_{\theta_1,\theta_2}\|_{\mathcal{B}_m} = \max\Big\{ &\operatorname*{ess\,sup}_{\eta\in D} \int_D \|(m \cdot K_{\theta_1,\theta_2})^{(\omega,\eta)}\|_{\mathcal{A}_1} \, d\omega, \\ &\operatorname*{ess\,sup}_{\omega\in D} \int_D \|(m \cdot K_{\theta_1,\theta_2})^{(\omega,\eta)}\|_{\mathcal{A}_1} \, d\eta \Big\} \\ \leq \max\Big\{ &\operatorname*{ess\,sup}_{\eta\in D} \int_D B_1(\omega,\eta)\, d\omega, \ \operatorname*{ess\,sup}_{\eta\in D} \int_D B_2(\eta,\omega)\, d\omega \Big\}, \end{aligned}$$

which completes the proof. □

The second intermediate result expresses the integral over D in (4.11) through the Fourier integral operators $L_{\tau_0}^{(\ell)}$.

Lemma 4.7 *Let $L^{(\ell)}_{\tau_0}$, $\ell \in \{1,2\}$, be as in Theorem 4.5. For all $(y,\omega),(z,\eta) \in \Lambda$ and $\ell \in \{1,2\}$, we have*

$$\left|K_{\theta_\ell,\theta_{3-\ell}}\big((y,\omega),(z,\eta)\big)\right| = \sqrt{\frac{w(\Phi(\eta))}{w(\Phi(\omega))}} \cdot \left|L^{(\ell)}_{\Phi(\eta)}\big(A^T(\Phi(\eta))\langle z-y\rangle, \Phi(\omega)-\Phi(\eta)\big)\right|. \tag{4.15}$$

If m is Φ-compatible with dominating weight m^Φ, then we have, for given arbitrary $\ell \in \{1,2\}$ and $\eta \in D$,

$$\begin{aligned}&\int_D \operatorname*{ess\,sup}_{z\in\mathbb{R}^d} \int_{\mathbb{R}^d} m^\Phi\big(y-z, \Phi(\omega)-\Phi(\eta)\big) \cdot |K_{\theta_\ell,\theta_{3-\ell}}\big((y,\omega),(z,\eta)\big)|\, dy\, d\omega \\ &\quad \le \left\|L^{(\ell)}_{\Phi(\eta)}\right\|_{L^1_M(\mathbb{R}^d\times\mathbb{R}^d)},\end{aligned} \tag{4.16}$$

with M as in (4.2).

Proof. We provide the proof for $\ell = 1$; the proof for $\ell = 2$ follows the same steps. First, recall from after Eq. (3.3) the identity $0 < w(\Phi(\xi)) = [\det \mathrm{D}\Phi(\xi)]^{-1}$ for all $\xi \in D$. This identity will be applied repeatedly. To show (4.15), apply Plancherel's theorem and perform the change of variable $\iota = \Phi(\xi) - \Phi(\eta)$ to derive

$$\begin{aligned}\left|K_{\theta_1,\theta_2}\big((y,\omega),(z,\eta)\big)\right| &= \left|\langle g^{[2]}_{z,\eta}, g^{[1]}_{y,\omega}\rangle\right| = \left|\langle \widehat{g^{[2]}_{z,\eta}}, \widehat{g^{[1]}_{y,\omega}}\rangle\right| \\ &= \left|\int_D \frac{\theta_2(\Phi(\xi)-\Phi(\eta))\cdot\overline{\theta_1(\Phi(\xi)-\Phi(\omega))}}{\sqrt{w(\Phi(\eta))\cdot w(\Phi(\omega))}} \cdot e^{-2\pi i\langle z-y,\xi\rangle}\, d\xi\right| \\ &= \left|\int_{\mathbb{R}^d} \theta_2(\iota)\cdot\overline{\theta_1(\iota+\Phi(\eta)-\Phi(\omega))}\cdot \frac{w(\iota+\Phi(\eta))}{\sqrt{w(\Phi(\eta))w(\Phi(\omega))}} \right.\\ &\qquad \left. \cdot e^{-2\pi i\langle z-y, \Phi^{-1}(\iota+\Phi(\eta))\rangle}\, d\iota\right|.\end{aligned} \tag{4.17}$$

This easily implies (4.15).

To prove (4.16), set $\tau_0 := \Phi(\eta)$ and note that (4.15) implies that the left-hand side of (4.16) satisfies

$$\begin{aligned}\circledast &:= \mathrm{LHS}(4.16) \\ &= \int_D \operatorname*{ess\,sup}_{z\in\mathbb{R}^d} \int_{\mathbb{R}^d} m^\Phi(y-z, \Phi(\omega)-\tau_0) \cdot \sqrt{\frac{w(\tau_0)}{w(\Phi(\omega))}} \\ &\qquad \cdot \left|L^{(1)}_{\tau_0}(A^T(\tau_0)\langle z-y\rangle, \Phi(\omega)-\tau_0)\right|\, dy\, d\omega.\end{aligned}$$

Next, perform the change of variable $\tau = \Phi(\omega) - \tau_0$ to obtain

$$\circledast = \int_{\mathbb{R}^d} \operatorname*{ess\,sup}_{z\in\mathbb{R}^d} \int_{\mathbb{R}^d} m^\Phi(y-z,\tau)\sqrt{w(\tau_0)w(\tau+\tau_0)}$$

$$\cdot \left| L_{\tau_0}^{(1)}(A^T(\tau_0)\langle z-y\rangle, \tau) \right| \, dy \, d\tau =: (\dagger).$$

Next, perform the change of variables $x = A^T(\tau_0)\langle z-y\rangle$ in the inner integral and apply the estimate $\sqrt{\frac{w(\tau+\tau_0)}{w(\tau_0)}} \le \sqrt{w_0(\tau)}$ to derive

$$(\dagger) \le \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} m^\Phi(-A^{-T}(\tau_0)\langle x\rangle, \tau) \cdot \sqrt{w_0(\tau)} \cdot \left| L_{\tau_0}^{(1)}(x, \tau) \right| \, dx \, d\tau.$$

Now, in the case where $\mathrm{D}\Phi$ is unbounded, we are done, since in this case Eqs. (4.3) and (4.2) show

$$m^\Phi(-A^{-T}(\tau_0)\langle x\rangle, \tau) \cdot \sqrt{w_0(\tau)} = m^\Phi(0, \tau) \cdot \sqrt{w_0(\tau)} \le M(x, \tau).$$

For the case that $\mathrm{D}\Phi$ is bounded, recall from (3.3) that $A^{-T}(\tau_0) = A^{-T}(\Phi(\eta)) = [\mathrm{D}\Phi]^T(\eta)$ and thus $|A^{-T}(\tau_0)\langle x\rangle| = |\mathrm{D}\Phi^T(\eta)\langle x\rangle| \le R|x|$ by choice of R in (4.2). Therefore, by choice of M, we see

$$m^\Phi(-A^{-T}(\tau_0)\langle x\rangle, \tau) \cdot \sqrt{w_0(\tau)} \le M(x, \tau).$$

□

We now obtain Eq. (4.10) in Theorem 4.5 simply by inserting Eq. (4.16) into Eq. (4.11).

4.2 Uniform Integrability of the Integral Kernels $L_{\tau_0}^{(\ell)}$

To control $\operatorname{ess\,sup}_{\tau_0 \in \mathbb{R}^d} \| L_{\tau_0}^{(\ell)} \|_{\mathbf{L}_M^1}$, we find that k-admissibility of the warping function Φ is crucial. The remainder of this subsection is dedicated to proving Theorem 4.8 below, which will in turn be central to proving Theorem 4.4.

Theorem 4.8 *Let Φ be a k-admissible warping function with control weight v_0. Furthermore, let $w_1 : \mathbb{R}^d \to \mathbb{R}^+$ be continuous and submultiplicative and such that $w_1(-\iota) = w_1(\iota)$ for all $\iota \in \mathbb{R}^d$. Define*

$$w_2 : \quad \mathbb{R}^d \to \mathbb{R}^+, \quad \iota \mapsto w_1(\iota) \cdot [v_0(\iota)]^{d+3k},$$

assume that $\theta_1, \theta_2 \in C^k(\mathbb{R}^d)$ are such that

$$\frac{\partial^n}{\partial \iota_j^n} \theta_\ell \in \boldsymbol{L}_{w_2}^2(\mathbb{R}^d), \qquad \text{for all } j \in \underline{d},\ \ell \in \{1, 2\},\ 0 \le n \le k, \tag{4.18}$$

and recall from Eq. (4.7), that

$$C_{\max} = \prod_{\ell \in \{1,2\}} \left(\max_{j \in \underline{d}} \max_{0 \le n \le k} \left\| \frac{\partial^n}{\partial \iota_j^n} \theta_\ell \right\|_{\boldsymbol{L}_{w_2}^2(\mathbb{R}^d)} \right).$$

Then, with e_τ *as defined in Eq.* (4.8) *and* $L_{\tau_0}^{(\ell)}$ *as in Theorem* 4.5*, there exists a constant* $C = C(d, k, v_0) > 0$ *satisfying for all* $x, \tau, \tau_0 \in \mathbb{R}^d$ *and* $\ell \in \{1, 2\}$ *the estimate*

$$\begin{aligned} \left|L_{\tau_0}^{(\ell)}(x,\tau)\right| &= \left|\int_{\mathbb{R}^d} \frac{w(\iota+\tau_0)}{w(\tau_0)} \big(\theta_{3-\ell}\cdot\overline{\boldsymbol{T}_\tau\theta_\ell}\big)(\iota)\, e_{\tau_0}(x,\iota)\, d\iota\right| \\ &\le C\cdot C_{\max}\cdot(1+|x|)^{-k}\cdot[w_1(\tau)]^{-1}. \end{aligned} \tag{4.19}$$

Remark In Sect. 6, we will apply Theorem 4.8 in a setting in which θ_1, θ_2 depend on x, τ, τ_0. We suggest that the reader keeps this potential dependency in mind.

In a first step, we derive a number of important consequences of Definition 4.2 that will be used repeatedly.

Lemma 4.9 *If* Φ *is a* 0*-admissible warping function with control weight* v_0*, then* Φ *is a warping function in the sense of Definition* 3.1*. In particular,* $w = \det A$ *is* w_0*-moderate with* $w_0 = v_0^d$*, i.e.*

$$w(\iota+\tau) \le w(\iota)\cdot[v_0(\tau)]^d \qquad \forall\,\tau,\iota\in\mathbb{R}^d \tag{4.20}$$

$$\text{and}\quad \|A(\iota+\tau)\| \le \|A(\iota)\|\cdot v_0(\tau) \qquad \forall\,\tau,\iota\in\mathbb{R}^d. \tag{4.21}$$

Additionally, for arbitrary $\gamma \in S^{d-1}$ *and* $\tau, \iota \in \mathbb{R}^d$*, we have*

$$[v_0(\iota-\tau)]^{-1} \le \|A^{-1}(\tau)A(\iota)\|^{-1} \le |\phi_\iota(\tau-\iota)\langle\gamma\rangle| \le \|A^{-1}(\iota)\cdot A(\tau)\| \le v_0(\tau-\iota) \tag{4.22}$$

and

$$\phi_{\tau_0}(\iota) = \phi_{\tau_0+\tau}(\iota-\tau)\cdot\phi_{\tau_0}(\tau). \tag{4.23}$$

Finally, we have

$$[v_0(\tau)]^{-1}\cdot|\gamma| \le |\phi_\iota(\tau)\langle\gamma\rangle| \le v_0(\tau)\cdot|\gamma| \qquad \forall\,\gamma\in\mathbb{R}^d \text{ and } \tau,\iota\in\mathbb{R}^d. \tag{4.24}$$

Proof To show that Φ is a warping function, we need only verify moderateness of $w = \det A$. To prove this moderateness, apply *Hadamard's inequality* $|\det M| \le \|M\|^d = \|M^T\|^d$ (see [79, Chapter 75]) for $M \in \mathbb{R}^{d\times d}$, combined with (4.5) (for $\alpha = 0$) to see that

$$\begin{aligned} \frac{w(\iota+\tau)}{w(\iota)} &= \det\Big([A(\iota)]^{-1}A(\iota+\tau)\Big) \le \big\|[A(\iota)]^{-1}A(\iota+\tau)\big\|^d \\ &= \big\|[\phi_\iota(\tau)]^T\big\|^d \le [v_0(\tau)]^d. \end{aligned}$$

Hence, we obtain (4.20). Moreover,

$$\|A(\iota+\tau)\| = \|A(\iota)A^{-1}(\iota)A(\iota+\tau)\| \le \|A(\iota)\|\cdot\|[\phi_\iota(\tau)]^T\| \overset{(4.5)}{\le} \|A(\iota)\|\cdot v_0(\tau),$$

proving (4.21). To show (4.22), first note for $\gamma \in S^{d-1}$ and any $M \in \mathrm{GL}(\mathbb{R}^d)$ that $|M\gamma| \geq \|M^{-1}\|^{-1}$, and then apply (4.5) twice:

$$\begin{aligned}
\frac{1}{v_0(\iota-\tau)} &\leq \left\|[\phi_\tau(\iota-\tau)]^T\right\|^{-1} = \|A^{-1}(\tau)A(\iota)\|^{-1} = \left\|[A^T(\tau)\cdot A^{-T}(\iota)]^{-1}\right\|^{-1} \\
&= \left\|[\phi_\iota(\tau-\iota)]^{-1}\right\|^{-1} \\
&\leq |\phi_\iota(\tau-\iota)\gamma| \leq \|\phi_\iota(\tau-\iota)\| = \|A^{-1}(\iota)\cdot A(\tau)\| \\
&\leq v_0(\tau-\iota), \quad \text{for all} \quad \iota,\tau \in \mathbb{R}^d, \quad \gamma \in S^{d-1}.
\end{aligned}$$

Finally, assertion (4.23) is easily verified using direct computation, and $[v_0(\tau)]^{-1} \leq |\phi_\iota(\tau)\cdot\gamma| \leq v_0(\tau)$ for $|\gamma| = 1$ is obtained from (4.22) through the bijective map $\tau \mapsto \tau - \iota$ and using that v_0 is radial. This proves (4.24). □

Lemma 4.9 shows that w is v_0^d-moderate. The next result provides v_0^d-moderateness (up to a constant) for the partial derivatives of w.

Lemma 4.10 *Let Φ be a k-admissible warping function with control weight v_0. For every $j \in \underline{d}$ and $n \in \mathbb{N}_0$ with $n \leq k$, we have*

$$\left|\frac{\partial^n}{\partial \iota_j^n} w(\iota+\tau)\right| \leq D_n \cdot [v_0(\iota)]^d \cdot w(\tau), \text{ for all } \iota,\tau \in \mathbb{R}^d, \tag{4.25}$$

with $D_n := D_n(d) := d! \cdot d^n$.

Proof We begin by rewriting $\frac{\partial^n}{\partial \iota_j^n} w(\iota+\tau)$ using some simple properties of determinants:

$$\begin{aligned}
\frac{\partial^n}{\partial \iota_j^n} w(\iota+\tau) &= \frac{\partial^n}{\partial \iota_j^n}\det(A(\iota+\tau)) = \frac{\partial^n}{\partial \iota_j^n}\det(A^T(\iota+\tau)) \\
&= \det(A^T(\tau))\frac{\partial^n}{\partial \iota_j^n}\det(A^T(\iota+\tau)A^{-T}(\tau)) \overset{(4.4)}{=} w(\tau)\frac{\partial^n}{\partial \iota_j^n}\det(\phi_\tau(\iota)).
\end{aligned}$$

Let S_d be the set of permutations on $\underline{d}$. Then, the definition of the determinant yields

$$\begin{aligned}
\frac{\partial^n}{\partial \iota_j^n}\det(\phi_\tau(\iota)) &= \frac{\partial^n}{\partial \iota_j^n}\left[\sum_{\sigma\in S_d}\operatorname{sgn}(\sigma)\prod_{i=1}^d[\phi_\tau(\iota)]_{i,\sigma(i)}\right] \\
&= \sum_{\sigma\in S_d}\operatorname{sgn}(\sigma)\frac{\partial^n}{\partial \iota_j^n}\prod_{i=1}^d[\phi_\tau(\iota)]_{i,\sigma(i)}.
\end{aligned}$$

The general Leibniz rule for products with d terms shows

$$\frac{\partial^n}{\partial \iota_j^n}\prod_{i=1}^d[\phi_\tau(\iota)]_{i,\sigma(i)} = \sum_{\substack{m_1,\ldots,m_d\in\mathbb{N}_0,\\ m_1+\ldots+m_d=n}}\binom{n}{m_1,\ldots,m_d}\prod_{i=1}^d\frac{\partial^{m_i}}{\partial \iota_j^{m_i}}[\phi_\tau(\iota)]_{i,\sigma(i)},$$

where $\binom{n}{m_1,\ldots,m_d} := \frac{n!}{m_1!\cdots m_d!}$ is the usual multinomial coefficient. Moreover, the estimate (4.5) yields

$$\left|\frac{\partial^{m_i}}{\partial \iota_j^{m_i}}[\phi_\tau(\iota)]_{i,\sigma(i)}\right| \leq \left\|\partial^{m_i e_j}\phi_\tau\ (\iota)\right\| \leq v_0(\iota).$$

Altogether, we obtain

$$\begin{aligned}\left|\frac{\partial^n}{\partial \iota_j^n} w(\iota+\tau)\right| &= \left|w(\tau)\cdot\frac{\partial^n}{\partial \iota_j^n}\det(\phi_\tau(\iota))\right| \\ &\leq w(\tau)\cdot[v_0(\iota)]^d\cdot\sum_{\sigma\in S_d}\ \sum_{m_1+\ldots+m_d=n}\binom{n}{m_1,\ldots,m_d} \\ &= d!\cdot d^n\cdot[v_0(\iota)]^d\cdot w(\tau) = D_n\cdot[v_0(\iota)]^d\cdot w(\tau),\end{aligned}$$

where we used $|S_d| = d!$ and the multinomial theorem (see e.g. [45, Exercise 2(a)]), i.e.

$$\sum_{m_1+\ldots+m_d=n}\binom{n}{m_1,\ldots,m_d}\prod_{i=1}^d a_i^{m_i} = (a_1+\ldots+a_d)^n,\ \text{for all } n\in\mathbb{N},\ (a_i)_{i\in\underline{d}}\in\mathbb{R}^d,$$

for $a_1,\ldots,a_d = 1$. Thus, the proof is complete. □

We now turn our attention towards the Fourier integral operators $L_{\tau_0}^{(\ell)}$ defined in (4.9). We will obtain the desired integrability with respect to $x\in\mathbb{R}^d$ by means of an integration by parts argument of the kind well-known for establishing the smoothness-decay duality of a function and its Fourier transform, as well as the asymptotic behavior of oscillatory integrals, cf. [85, Chapter VIII]. An additional complication in our setting is that we require a uniform estimate over all $L_{\tau_0}^{(\ell)}$, $\ell\in\{1,2\}$, $\tau_0\in\mathbb{R}^d$.

For now, we replace $\frac{w(\iota+\tau_0)}{w(\tau_0)}\left(\theta_2\cdot\overline{\mathbf{T}_\tau\theta_1}\right)(\iota)$ in (4.9) by an unspecific, compactly supported function $g\in\mathcal{C}_c^k(\mathbb{R}^d)$, i.e., we consider

$$\int_{\mathbb{R}^d} g(\iota)\cdot e_{\tau_0}(x,\iota)\,d\iota,\ \text{recalling } e_\tau(x,\iota) = e^{-2\pi i\langle A^{-T}(\tau)\langle x\rangle,\Phi^{-1}(\iota+\tau)\rangle},\quad \text{for all } x,\iota,\tau\in\mathbb{R}^d. \tag{4.26}$$

Note that, with $f = e_\tau(x,\bullet)$, we have

$$\begin{aligned}\frac{\partial}{\partial\iota_j}f(\iota) &= -2\pi i\cdot\left\langle A^{-T}(\tau)\langle x\rangle,\frac{\partial}{\partial\iota_j}\Phi^{-1}(\iota+\tau)\right\rangle\cdot e_\tau(x,\iota) \\ &= -2\pi i\cdot(\phi_\tau(\iota)\cdot x)_j\cdot e_\tau(x,\iota).\end{aligned}$$

The final equality can be verified by observing

$$\begin{aligned}\left\langle A^{-T}(\tau)\langle\eta\rangle, \frac{\partial}{\partial\iota_j}\Phi^{-1}(\iota+\tau)\right\rangle &= \langle A^{-T}(\tau)\langle\eta\rangle, \mathrm{D}\Phi^{-1}(\iota+\tau)\langle e_j\rangle\rangle \\ &= \langle A^{-T}(\tau)\langle\eta\rangle, A(\iota+\tau)\langle e_j\rangle\rangle = (\phi_\tau(\iota)\cdot\eta)_j,\end{aligned} \tag{4.27}$$

which motivates the definition of ϕ_τ. Provided $(\phi_\tau(\iota)\cdot x)_j \neq 0$ on the support of g, we obtain, with $\tilde{g}(\iota) = \left(-2\pi i\cdot(\phi_\tau(\iota)\langle x\rangle)_j\right)^{-1}\cdot g(\iota)$,

$$\int g(\iota)f(\iota)d\iota = \int \tilde{g}(\iota)\frac{\partial}{\partial\iota_j}f(\iota)d\iota = -\int\frac{\partial}{\partial\iota_j}\tilde{g}(\iota)f(\iota)d\iota,$$

where the last equality is obtained through integration by parts.

For fixed $x,\tau\in\mathbb{R}^d$ and $j\in\underline{d}$ and all $g\in\mathcal{C}_c^k(\mathbb{R}^d)$ such that $(\phi_\tau(\cdot)\langle x\rangle)_j\neq 0$ on the support of g, we define the differential operator $\square_{j,\tau,x}$ by

$$\left(\square_{j,\tau,x}\,g\right)(\iota) := (2\pi i)^{-1}\frac{\partial}{\partial\iota_j}\left[\frac{g(\iota)}{(\phi_\tau(\iota)\langle x\rangle)_j}\right] = (2\pi i|x|)^{-1}\frac{\partial}{\partial\iota_j}\left[\frac{g(\iota)}{(\phi_\tau(\iota)\langle\rho_x\rangle)_j}\right], \tag{4.28}$$

where $\rho_x\in S^{d-1}$ with $x=|x|\rho_x$. We can rewrite the integral in (4.26) as

$$\begin{aligned}\int_{\mathbb{R}^d} g(\iota)\,e_\tau(x,\iota)d\iota &= \int_{\mathbb{R}^d}\left(\square_{j,\tau,x}\,g\right)(\iota)e_\tau(x,\iota)d\iota \\ &= \int_{\mathbb{R}^d}\left(\square^n_{j,\tau,x}\,g\right)(\iota)\,e_\tau(x,\iota)d\iota, \text{ for } n\le k.\end{aligned} \tag{4.29}$$

where $\square^n_{j,\tau,x}$ denotes n-fold application of $\square_{j,\tau,x}$.

By (4.28), each application of $\square_{j,\tau,x}$ provides additional, linear decay with respect to $|x|$, $x\in\mathbb{R}^d$. For a given pair (Φ,θ) of warping function and prototype, however, we cannot expect the support restriction required for the application of the differential operator $\square_{j,\tau,x}$, i.e., $(\phi_\tau(\cdot)\langle x\rangle)_j\neq 0$ on the support of g, to hold. To account for this, we decompose $g_{\tau,\tau_0}(\iota) := \frac{w(\iota+\tau_0)}{w(\tau_0)}\left(\theta_2\cdot\overline{\mathbf{T}_\tau\theta_1}\right)(\iota)$ into compactly supported functions, such that each of them allows the application of $\square_{j,\tau,x}$, for some $j\in\underline{d}$. Therefore, our next steps are:

Step 1: Find a suitable splitting $g_{\tau,\tau_0} = \sum_{i\in I} g_{i,\tau,\tau_0} = \sum_{i\in I}\varphi_i g_{\tau,\tau_0}$ (with $(\varphi_i)_i$ only depending on τ_0) into compactly supported elements $g_{i,\tau,\tau_0} = \varphi_i g_{\tau,\tau_0}$, such that, for any fixed $\rho_x\in S^{d-1}$, $\tau_0\in\mathbb{R}^d$ and $i\in I$, there is an index $j = j(\rho_x,\tau_0,i)\in\underline{d}$ and a positive function $\tilde{v}$ (*independent* of i,ρ_x,τ_0,τ), such that $|(\phi_{\tau_0}(\iota)\langle\rho_x\rangle)_j|\ge\tilde{v}(\iota)>0$ for $\iota\in\operatorname{supp}\varphi_i\supset\operatorname{supp}(g_{i,\tau,\tau_0})$. Besides being able to apply $\square_{j,\tau_0,x}$, this property lets us control the growth of $\frac{1}{(\phi_{\tau_0}(\iota)\langle\rho_x\rangle)_j}$ independently of the orientation $\rho_x\in S^{d-1}$ and of τ,τ_0.

Step 2: Estimate $\left(\square^n_{j,\tau_0,x}\,g_{i,\tau,\tau_0}\right)(\iota)$, for $x=|x|\cdot\rho_x\neq 0$, independently of i,ρ_x,τ_0. In fact, this estimate will exhibit rapid decay with respect to $|x|$ and depend boundedly on the derivative of g_{i,τ,τ_0}, which can be used to obtain decay with respect to $|\tau|$.

Towards Step 1, we introduce a specific family of coverings in the following lemma. The smooth splitting of g_{τ,τ_0} into the building blocks $g_{i,\tau,\tau_0} = \varphi_i g_{\tau,\tau_0}$, see Lemma 4.13, is provided by a $\mathcal{C}_c^\infty$ partition of unity $(\varphi_i)_i$ with respect to these coverings, introduced in Lemma 4.12. Lemmas 4.14 and 4.15 take care of Step 2.

Lemma 4.11 *Let Φ be a 1-admissible warping function with control weight v_0. For any $\iota_0, \tau_0 \in \mathbb{R}^d$, the following are true:*

1. *The family $\left(U_j^{(\iota_0,\tau_0)}\right)_{j\in\underline{d}}$ defined by*

$$U_j^{(\iota_0,\tau_0)} := \left\{\gamma \in S^{d-1} \;:\; \left|\left(\phi_{\tau_0}(\iota_0)\langle\gamma\rangle\right)_j\right| > \frac{1}{2d}\left|\phi_{\tau_0}(\iota_0)\langle\gamma\rangle\right|\right\}$$

   is a covering of S^{d-1}.
2. *For any $\delta > 0$ satisfying $\delta \cdot v_0(\delta/(4d)\cdot e_1) \le 1/\sqrt{d}$ and arbitrary $\iota \in B_{\delta/(4d)}(\iota_0)$ and $\gamma \in U_j^{(\iota_0,\tau_0)}$, we have*

$$\left|\left(\phi_{\tau_0}(\iota)\langle\gamma\rangle\right)_j\right| \ge C_\delta \cdot [v_0(\iota)]^{-1},$$

   with $C_\delta := C_\delta(d, v_0) := \left[4d \cdot v_0(\delta/(4d)\cdot e_1)\right]^{-1}$.

Remark If $\delta \le \min\{1, 1/(\sqrt{d}\cdot v_0(e_1/(4d)))\}$, then $\delta \cdot v_0(\delta/(4d)\cdot e_1) \le \delta \cdot v_0(e_1/(4d)) \le 1/\sqrt{d}$. Hence, the condition of Part (2) of the lemma is satisfied for all sufficiently small $\delta > 0$.

Proof Part (1) does not use any of the properties of Φ, except that $\phi_{\tau_0}(\iota) \in \mathrm{GL}(\mathbb{R}^d)$: We simply note that any $z \in \mathbb{R}^d \setminus \{0\}$ satisfies

$$|z| \le \sum_{j=1}^{d} |z_j| \le d \cdot \max\{|z_j| : j \in \underline{d}\} < 2d \cdot \max\{|z_j| : j \in \underline{d}\}.$$

Hence, there is some $j \in \underline{d}$ with $|z_j| > \frac{1}{2d}\cdot|z|$. Now apply this to $z = \phi_{\tau_0}(\iota_0)\langle\gamma\rangle$, noting that $z \ne 0$ since $\phi_{\tau_0}(\iota_0) \in \mathrm{GL}(\mathbb{R}^d)$ and $\gamma \in S^{d-1}$.

For part (2), let $\iota \in B_{\delta/(4d)}(\iota_0)$ and $\gamma \in U_j^{(\iota_0,\tau_0)} \subset S^{d-1}$ be arbitrary. The triangle inequality provides

$$\begin{aligned} |(\phi_{\tau_0}(\iota)\langle\gamma\rangle)_j| &\ge |(\phi_{\tau_0}(\iota_0)\langle\gamma\rangle)_j| - |(\phi_{\tau_0}(\iota)\langle\gamma\rangle - \phi_{\tau_0}(\iota_0)\langle\gamma\rangle)_j| \\ (\text{since } \gamma\in U_j^{(\iota_0,\tau_0)}) &\ge \frac{|\phi_{\tau_0}(\iota_0)\langle\gamma\rangle|}{2d} - \left|\left(\phi_{\tau_0}(\iota) - \phi_{\tau_0}(\iota_0)\right)\langle\gamma\rangle\right|. \end{aligned}$$

Note that

$$\phi_{\tau_0}(\iota) - \phi_{\tau_0}(\iota_0) = \left(\phi_{\tau_0+\iota_0}(\iota - \iota_0) - \mathrm{id}\right)\phi_{\tau_0}(\iota_0), \tag{4.30}$$

where we used the identity (4.23) of Lemma 4.9, with $\tau = \iota_0$.

To estimate the first factor on the right-hand side of (4.30), recall that $\phi_{\tau_0+\iota_0}(0) = \mathrm{id}$. Therefore,

$$
\begin{aligned}
\left\|\mathrm{id} - \phi_{\tau_0+\iota_0}(\iota - \iota_0)\right\| &= \left\|\phi_{\tau_0+\iota_0}(0) - \phi_{\tau_0+\iota_0}(\iota - \iota_0)\right\| \\
&= \left\| \int_0^1 \frac{d}{dt}\bigg|_{t=s} \left[\phi_{\tau_0+\iota_0}(t(\iota - \iota_0))\right] ds \right\| \\
&\leq \int_0^1 \sum_{\ell=1}^{d} \left\|(\partial_\ell\, \phi_{\tau_0+\iota_0})(s(\iota - \iota_0))\right\| \cdot \left|(\iota - \iota_0)_\ell\right| ds =: (*).
\end{aligned}
$$

We now rewrite this expression further, recalling that v_0 is radially increasing and applying the Cauchy-Schwarz inequality, and inequality (4.5):

$$
\begin{aligned}
(*) &\overset{\mathrm{CS}}{\leq} |\iota - \iota_0| \cdot \sup_{t\in[0,1]} \left| \begin{pmatrix} \left\|(\partial_1\, \phi_{\tau_0+\iota_0})(t(\iota - \iota_0))\right\| \\ \vdots \\ \left\|(\partial_d\, \phi_{\tau_0+\iota_0})(t(\iota - \iota_0))\right\| \end{pmatrix} \right| \\
&\overset{(4.5)}{\leq} |\iota - \iota_0| \cdot \sqrt{d} \cdot \sup_{t\in[0,1]} v_0(t(\iota - \iota_0)) \\
(\text{since } |\iota-\iota_0|<\delta/(4d)) &\leq \frac{\sqrt{d}\cdot \delta \cdot v_0(\delta/(4d)\cdot e_1)}{4d} \leq \frac{1}{4d}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
|(\phi_{\tau_0}(\iota)\langle\gamma\rangle)_j| &\geq \frac{|\phi_{\tau_0}(\iota_0)\langle\gamma\rangle|}{2d} - \|\phi_{\tau_0+\iota_0}(\iota - \iota_0) - \mathrm{id}\| \cdot |\phi_{\tau_0}(\iota_0)\langle\gamma\rangle| \\
&\geq \frac{|\phi_{\tau_0}(\iota_0)\langle\gamma\rangle|}{4d}.
\end{aligned}
$$

To finish the proof, it remains to show $|\phi_{\tau_0}(\iota_0)\langle\gamma\rangle| \geq 4dC_\delta \cdot [v_0(\iota)]^{-1}$. To see this, note

$$
|\phi_{\tau_0}(\iota_0)\langle\gamma\rangle| \overset{(4.23)}{=} |\phi_{\tau_0+\iota}(\iota_0 - \iota)\cdot \phi_{\tau_0}(\iota)\langle\gamma\rangle| \overset{(4.24)}{\geq} \frac{1}{v_0(\iota - \iota_0)v_0(\iota)} \geq 4dC_\delta \cdot [v_0(\iota)]^{-1},
$$

where we inserted $C_\delta = (4d \cdot v_0(\delta/(4d)\cdot e_1))^{-1}$, using $|\iota - \iota_0| < \delta/(4d)$. □

Lemma 4.12 *Let $\delta' > 0$ be arbitrary. The sequence $\left(B_{\delta'}(\iota_i)\right)_{i\in\mathbb{Z}^d}$, with $(\iota_i)_{i\in\mathbb{Z}^d} = \left(\frac{\delta'}{\sqrt{d}} i\right)_{i\in\mathbb{Z}^d}$, is an open cover of $\mathbb{R}^d$. Moreover, there is a collection of smooth functions $(\varphi_i)_{i\in\mathbb{Z}^d}$, such that*

1. *$\varphi_i \geq 0$ and $\varphi_i \in C^\infty(\mathbb{R}^d)$,*
2. *$\mathrm{supp}(\varphi_i) \subset B_{\delta'}(\iota_i)$,*
3. *$\sum_i \varphi_i \equiv 1$ on $\mathbb{R}^d$, and*
4. *for every multi-index $\alpha \in \mathbb{N}_0^d$, there exists a constant $D_\alpha^{(\delta')} > 0$ such that $|\partial^\alpha \varphi_i(\iota)| \leq D_\alpha^{(\delta')}$ uniformly over $i \in \mathbb{Z}^d$ and $\iota \in \mathbb{R}^d$.*

Proof The result is a direct consequence of standard constructions of smooth partitions of unity; see e.g. [64, Theorem 1.4.6]. □

Lemma 4.13 *Let Φ be a k-admissible warping function with control weight v_0 and $\delta > 0$ be such that $\delta \cdot v_0(\delta/(4d) \cdot e_1) \leq 1/\sqrt{d}$. Set $\delta' = \delta/(4d)$ and let $(\iota_i)_{i\in\mathbb{Z}^d}$, $\left(B_{\delta/(4d)}(\iota_i)\right)_{i\in\mathbb{Z}^d}$ and $(\varphi_i)_{i\in\mathbb{Z}^d}$ be as in Lemma 4.12. Then*

$$\begin{aligned} \sharp\{i \in \mathbb{Z}^d \ : \ B_{\delta'}(\iota_\ell) \cap B_{\delta'}(\iota_i) \neq \varnothing\} &\leq (1{+}4d)^d \quad \text{and} \\ \sharp\{i \in \mathbb{Z}^d \ : \ \iota \in B_{\delta'}(\iota_i)\} &\leq (1{+}4d)^d, \end{aligned} \tag{4.31}$$

for all $\ell \in \mathbb{Z}^d$ and $\iota \in \mathbb{R}^d$. For $\theta_1, \theta_2 \in \mathbf{L}^2_{\sqrt{w_0}}(\mathbb{R}^d) \cap \mathcal{C}^k(\mathbb{R}^d)$, $\iota, \tau, \tau_0 \in \mathbb{R}^d$ and $i \in \mathbb{Z}^d$, define

$$g_{i,\tau,\tau_0}(\iota) := \varphi_i(\iota) g_{\tau,\tau_0}(\iota) \quad \text{with} \quad g_{\tau,\tau_0}(\iota) := \frac{w(\iota+\tau_0)}{w(\tau_0)} \left(\theta_2 \cdot \overline{\boldsymbol{T}_\tau \theta_1}\right)(\iota). \tag{4.32}$$

Then $g_{i,\tau,\tau_0} \in \mathcal{C}_c^k(B_{\delta'}(\iota_i))$ and, for any fixed $\tau_0 \in \mathbb{R}^d$ and $x \in \mathbb{R}^d \setminus \{0\}$, there exists a sequence $(j_i)_{i\in\mathbb{Z}^d}$ with $j_i \in \underline{d}$, such that $\rho_x := x/|x| \in U_{j_i}^{(\iota_i,\tau_0)}$ (where this set is defined is in Lemma 4.11) for all $i \in \mathbb{Z}^d$ and such that

$$\int_{\mathbb{R}^d} g_{\tau,\tau_0}(\iota) e_{\tau_0}(x,\iota)\, d\iota = \sum_{i\in\mathbb{Z}^d} \int_{\mathbb{R}^d} \left(\Box^n_{j_i,\tau_0,x}\, g_{i,\tau,\tau_0}\right)(\iota) \cdot e_{\tau_0}(x,\iota)\, d\iota, \quad \text{for all } n \leq k. \tag{4.33}$$

Proof The first assertion, (4.31), is verified by a straightforward calculation and $g_{i,\tau,\tau_0} \in \mathcal{C}_c^k(B_{\delta'}(\iota_i))$ is a consequence of Lemma 4.12, with k-admissibility of Φ and $\theta_1, \theta_2 \in \mathcal{C}^k(\mathbb{R}^d)$. Lemma 4.11(1) provides the existence of $j_i = j_i(i, \tau_0, \rho_x) \in \underline{d}$ satisfying $\rho_x \in U_{j_i}^{(\iota_i,\tau_0)}$, for arbitrary, fixed τ_0, ρ_x and each $i \in \mathbb{Z}^d$. The elements of the covering $\left(B_{\delta/(4d)}(\iota_i)\right)_{i\in\mathbb{Z}^d}$ are specific instances of the set in Lemma 4.11(2), such that the application of $\Box^n_{j_i,\tau_0,x}$, $n \in \mathbb{N}$, to g_{i,τ,τ_0} is well-defined. Thus, to prove (4.33) it only remains to justify the interchange of integral and summation

$$\int_{\mathbb{R}^d} \sum_{i\in\mathbb{Z}^d} g_{i,\tau,\tau_0}(\iota) e_{\tau_0}(x,\iota)\, d\iota = \sum_{i\in\mathbb{Z}^d} \int_{\mathbb{R}^d} g_{i,\tau,\tau_0}(\iota) e_{\tau_0}(x,\iota)\, d\iota. \tag{4.34}$$

Since

$$\begin{aligned} \int_{\mathbb{R}^d} \sum_{i\in\mathbb{Z}^d} |g_{i,\tau,\tau_0}(\iota) e_{\tau_0}(x,\iota)|\, d\iota &\leq \int_{\mathbb{R}^d} \sum_{i\in\mathbb{Z}^d} \mathbb{1}_{B_{\delta'}(\iota_i)}(\iota) \cdot |g_{\tau,\tau_0}(\iota)|\, d\iota \\ (\text{Eq. } (4.31)) &\leq (1+4d)^d \cdot \|g_{\tau,\tau_0}\|_{\mathbf{L}^1} \\ \left(\text{since } \tfrac{w(\iota+\tau_0)}{w(\tau_0)} \leq w_0(\iota)\right) &\leq (1+4d)^d \cdot \|w_0 \cdot \theta_2 \cdot \mathbf{T}_\tau \theta_1\|_{\mathbf{L}^1} \\ (w_0 \text{ is submultiplicative}) &\leq \sqrt{w_0(\tau)} \cdot (1+4d)^d \cdot \|\theta_2\|_{\mathbf{L}^2_{\sqrt{w_0}}} \cdot \|\mathbf{T}_\tau \theta_1\|_{\mathbf{L}^2_{\sqrt{w_0}}} < \infty, \end{aligned}$$

the dominated convergence theorem justifies (4.34). □

To prepare for an estimate of $\Box^n_{j_i,\tau_0,x}\, g_{i,\tau,\tau_0}$ itself, we consider the partial derivatives of g_{i,τ,τ_0}.

Lemma 4.14 *Let Φ be a k-admissible warping function with control weight v_0, let $\theta_1, \theta_2 \in \boldsymbol{L}^2_{\sqrt{w_0}} \cap \mathcal{C}^k$, and let $(\varphi_i)_{i\in\mathbb{Z}^d}$ be a bounded partition of unity as in Lemma 4.12, for some given $\delta' > 0$. For any fixed $j \in \underline{d}$, $i \in \mathbb{Z}^d$ and $\tau, \tau_0 \in \mathbb{R}^d$, we have*

$$\left| \frac{\partial^n}{\partial \iota_j^n} g_{i,\tau,\tau_0} \right| \le C_n \cdot v_0^d \cdot \sum_{\substack{m_1,m_2\in\mathbb{N}_0 \\ m_1+m_2\le n}} \left| \frac{\partial^{m_1}}{\partial \iota_j^{m_1}} \theta_2 \cdot \frac{\partial^{m_2}}{\partial \iota_j^{m_2}} \overline{\boldsymbol{T}_\tau \theta_1} \right|, \text{ for all } n \le k, \tag{4.35}$$

for some constant $C_n = C_n(\delta', d) > 0$. Here $g_{i,\tau,\tau_0}(\iota) = \frac{w(\iota+\tau_0)}{w(\tau_0)} \varphi_i(\iota) \left(\theta_2 \cdot \overline{\boldsymbol{T}_\tau \theta_1}\right)(\iota)$ is as in (4.32).

***Proof*.** We begin by applying the general Leibniz rule, with 4 terms in this case, to rewrite the partial derivatives of g_{i,τ,τ_0}:

$$\begin{aligned} \frac{\partial^n}{\partial \iota_j^n} g_{i,\tau,\tau_0} &= \frac{\partial^n}{\partial \iota_j^n} \left(\frac{\mathbf{T}_{-\tau_0} w}{w(\tau_0)} \cdot \varphi_i \cdot \theta_2 \cdot \overline{\mathbf{T}_\tau \theta_1} \right) \\ &= \frac{1}{w(\tau_0)} \sum_{\substack{n_1,\dots,n_4\in\mathbb{N}_0 \\ n_1+\dots+n_4=n}} \binom{n}{n_1,\dots,n_4} \cdot \frac{\partial^{n_1}}{\partial \iota_j^{n_1}} \mathbf{T}_{-\tau_0} w \cdot \frac{\partial^{n_2}}{\partial \iota_j^{n_2}} \varphi_i \cdot \frac{\partial^{n_3}}{\partial \iota_j^{n_3}} \theta_2 \cdot \frac{\partial^{n_4}}{\partial \iota_j^{n_4}} \overline{\mathbf{T}_\tau \theta_1}, \end{aligned} \tag{4.36}$$

where $\binom{n}{n_1,\dots,n_4} := \frac{n!}{n_1! n_2! n_3! n_4!}$ is, once more, the usual multinomial coefficient.

We now consider each term appearing in (4.36) individually. Since all the involved sums are finite, there is a finite constant $\widetilde{C}_n > 0$, depending only on $\delta' > 0$, the chosen partition of unity $(\varphi_i)_{i\in\mathbb{Z}^d}$, and (implicitly) $d \in \mathbb{N}$, such that

$$\max_{\substack{n_1,\dots,n_4\in\mathbb{N}_0 \\ n_1+\dots+n_4=n}} \left| \binom{n}{n_1,\dots,n_4} \cdot \frac{\partial^{n_2}}{\partial \iota_j^{n_2}} \varphi_i \right| \le \max_{\substack{n_1,\dots,n_4\in\mathbb{N}_0 \\ n_1+\dots+n_4=n}} \left(\binom{n}{n_1,\dots,n_4} \cdot D^{(\delta')}_{n_2 e_j} \right) \le \widetilde{C}_n \,,$$

where property (4) of $(\varphi_i)_{i\in\mathbb{Z}^d}$ in Lemma 4.12 was used, and $n_2 e_j$ is interpreted as a multi-index.

For the term $[w(\tau_0)]^{-1} \cdot \frac{\partial^{n_1}}{\partial \iota_j^{n_1}} w(\iota + \tau_0)$ on the other hand, we apply the estimate given in Lemma 4.10, i.e.

$$[w(\tau_0)]^{-1} \cdot \left| \frac{\partial^{n_1}}{\partial \iota_j^{n_1}} w(\iota + \tau_0) \right| \leq D_{n_1} \cdot [v_0(\iota)]^d \leq \left(\max_{0 \leq m \leq n} D_m \right) \cdot [v_0(\iota)]^d$$
$$= D_n \cdot [v_0(\iota)]^d,$$

where $D_n = d!d^n$ as in Lemma 4.10. With $C_n := D_n \widetilde{C_n}$, we see that

$$\left| \frac{\partial^n}{\partial \iota_j^n} g_{i,\tau,\tau_0} \right| \leq D_n \widetilde{C}_n \cdot \sum_{\substack{n_1,\ldots,n_4 \in \mathbb{N}_0 \\ n_1 + \cdots + n_4 = n}} \left| v_0^d \cdot \frac{\partial^{n_3}}{\partial \iota_j^{n_3}} \theta_2 \cdot \frac{\partial^{n_4}}{\partial \iota_j^{n_4}} \overline{\mathbf{T}_\tau \theta_1} \right|$$
$$\leq C_n \cdot v_0^d \cdot \sum_{\substack{n_3,n_4 \in \mathbb{N}_0 \\ n_3 + n_4 \leq n}} \left| \frac{\partial^{n_3}}{\partial \iota_j^{n_3}} \theta_2 \cdot \frac{\partial^{n_4}}{\partial \iota_j^{n_4}} \overline{\mathbf{T}_\tau \theta_1} \right|.$$

□

The next lemma provides an estimate of $|\square^n_{j_i,\tau_0,x}\, g|$ in terms of the partial derivatives of g and the weight function v_0 from Definition 4.2.

Lemma 4.15 *Let Φ be a k-admissible warping function with control weight v_0 and choose $\delta > 0$ such that $\delta \cdot v_0(\delta/(4d) \cdot e_1) \leq 1/\sqrt{d}$. Fix $j \in \underline{d}$ and $\iota_0, \tau_0 \in \mathbb{R}^d$, and let $U_j^{(\iota_0,\tau_0)}$ be as in Lemma 4.11(1). If $g \in C_c^k(B_{\delta/(4d)}(\iota_0))$ and if $x \in \mathbb{R}^d \setminus \{0\}$ satisfies $x/|x| \in U_j^{(\iota_0,\tau_0)}$ then, with*

$$\left(\square_{j,\tau_0,x}\, g \right) = (2\pi i |x|)^{-1} \frac{\partial}{\partial \iota_j} \left[\frac{g(\bullet)}{\left(\phi_{\tau_0}(\bullet) \langle x/|x| \rangle \right)_j} \right]$$

as in (4.28), there exists $D_{n,\delta} := D_{n,\delta}(v_0) > 0$, independent of j, x, τ_0, as well as ι_0 and the function $g \in C_c^k(B_{\delta/(4d)}(\iota_0))$, such that

$$\left| \square^n_{j,\tau_0,x}\, g \right| \leq D_{n,\delta} \cdot (2\pi |x|)^{-n} \cdot v_0^{3n} \cdot \sum_{m=0}^{n} \left| \frac{\partial^m}{\partial \iota_j^m} g \right|$$

holds for all $0 \leq n \leq k$.

Proof **Step 1 (Preparation):** Given $j \in \underline{d}$ and a strictly positive (or strictly negative) function $h \in C^1(U)$ defined on an open set $\varnothing \neq U \subset \mathbb{R}^d$, we define the differential operator $\blacksquare_{j,h}$ by $\blacksquare_{j,h}\, g := \frac{\partial}{\partial \iota_j} \left(\frac{g}{h} \right)$. Then the following identity can be derived from the quotient rule by a tedious, but straightforward induction:

$$\blacksquare^n_{j,h}\, g = h^{-2n} \cdot \sum_{m=0}^{n} \left(\frac{\partial^m g}{\partial \iota_j^m} \cdot \sum_{\substack{\alpha \in \mathbb{N}_0^n \\ |\alpha| = n-m}} \left(C^{(m,\alpha)} \cdot \prod_{\ell=1}^{n} \frac{\partial^{\alpha_\ell} h}{\partial \iota_j^{\alpha_\ell}} \right) \right), \quad \text{for all } g \in C^k(U) \text{ and } n \in \underline{k}, \tag{4.37}$$

for suitable constants $C^{(m,\alpha)} \in \mathbb{Z}$ that depend only on $\alpha \in \mathbb{N}_0^n$ and on $m \in \{0, \dots, n\}$. Furthermore, we have the equality

$$\square^n_{i,\tau_0,x}\, g(\iota) = (2\pi i|x|)^{-n} \cdot \blacksquare^n_{j,(\phi_{\tau_0}(\cdot)\langle x/|x|\rangle)_j}\, g(\iota).$$

Step 2 (Completing the proof): For $n = 0$, there is nothing to prove. Hence, we can assume $n \in \underline{k}$. With $U_j^{(\iota_0,\tau_0)} \subset S^{d-1}$ as in Lemma 4.11(1), there is a $j \in \underline{d}$, such that $x/|x| = \rho_x \in U_j^{(\iota_0,\tau_0)}$ and therefore, $\square_{j,\tau_0,x} g$ is well-defined for arbitrary $g \in C_c^k(B_{\delta/(4d)}(\iota_0))$ by Lemma 4.11(2). Now, (4.37) provides

$$\begin{aligned}&\blacksquare^n_{j,(\phi_{\tau_0}(\cdot)\langle\rho_x\rangle)_j}\, g \\ &= \left(\phi_{\tau_0}(\cdot)\langle\rho_x\rangle\right)_j^{-2n} \cdot \sum_{m=0}^{n}\left(\frac{\partial^m g}{\partial\iota_j^m} \cdot \sum_{\substack{\alpha\in\mathbb{N}_0^n\\ |\alpha|=n-m}}\left(C^{(m,\alpha)} \cdot \prod_{\ell=1}^{n}\frac{\partial^{\alpha_\ell}}{\partial\iota_j^{\alpha_\ell}}\left(\phi_{\tau_0}(\cdot)\langle\rho_x\rangle\right)_j\right)\right).\end{aligned} \tag{4.38}$$

We now estimate the modulus of the innermost product by using (4.5):

$$\left|\prod_{\ell=1}^{n}\frac{\partial^{\alpha_\ell}}{\partial\iota_j^{\alpha_\ell}}\left(\phi_{\tau_0}(\iota)\langle\rho_x\rangle\right)_j\right| \overset{|\rho_x|=1}{\leq} \prod_{\ell=1}^{n}\left\|\frac{\partial^{\alpha_\ell}}{\partial\iota_j^{\alpha_\ell}}\phi_{\tau_0}(\iota)\right\| \leq v_0^n(\iota).$$

Insert this estimate into (4.38) to obtain

$$\begin{aligned}\left|\blacksquare^n_{j,(\phi_{\tau_0}(\cdot)\langle\rho_x\rangle)_j}\, g(\iota)\right| &\leq \frac{[v_0(\iota)]^n}{|(\phi_{\tau_0}(\iota)\langle\rho_x\rangle)_j|^{2n}} \cdot \sum_{m=0}^{n}\left(\left|\frac{\partial^m}{\partial\iota_j^m}g(\iota)\right| \cdot \sum_{\substack{\alpha\,\in\,\mathbb{N}_0^n\\ |\alpha|\,=\,n\,-\,m}}\left|C^{(m,\alpha)}\right|\right) \\ (\text{Lemma } 4.11) &\leq \frac{[v_0(\iota)]^{3n}}{C_\delta(d,v_0)^{2n}} \cdot \sum_{m=0}^{n}\left(\left|\frac{\partial^m}{\partial\iota_j^m}g(\iota)\right| \cdot \sum_{\substack{\alpha\,\in\,\mathbb{N}_0^n\\ |\alpha|\,=\,n\,-\,m}}\left|C^{(m,\alpha)}\right|\right) \\ &\leq D_{n,\delta} \cdot [v_0(\iota)]^{3n} \cdot \sum_{m=0}^{n}\left|\frac{\partial^m}{\partial\iota_j^m}g(\iota)\right|,\end{aligned}$$

where $D_{n,\delta}(v_0) := C_\delta(d,v_0)^{-2n} \cdot \max_{m=0,\dots,n}\left(\sum_{|\alpha|=n-m}|C^{(m,\alpha)}|\right)$ only depends on $n \leq k$, $\delta > 0$, and on the control weight v_0. $\square$

We are ready to prove Theorem 4.8, in particular we can now estimate the integral appearing on the right-hand side of (4.9).

Proof of Theorem 4.8 Recall from Lemma 4.9 that $w_0 = v_0^d$. Furthermore, note by submultiplicativity of w_1 that $w_1(0) = w_1(0+0) \leq [w_1(0)]^2$, and hence $w_1(0) \geq 1$. This implies $w_1 \geq 1$: Another application of submultiplicativity yields

$1 \le w_1(0) = w_1(\iota + (-\iota)) \le w_1(\iota) \cdot w_1(-\iota) = [w_1(\iota)]^2$, since $w_1(-\iota) = w_1(\iota)$. By the same arguments, we see $v_0 \ge 1$. Therefore, we conclude that (4.18) implies $\theta_\ell \in \mathbf{L}^2_{v_0^{d/2}}(\mathbb{R}^d) = \mathbf{L}^2_{\sqrt{w_0}}(\mathbb{R}^d)$, i.e., θ_1, θ_2 satisfy the conditions of Lemma 4.13.

In the following, we only consider the case $\ell = 1$; the corresponding estimates for $\ell = 2$ can be obtained simply by swapping θ_1, θ_2; our assumptions, and the definition of $C_{\max}$, are invariant under this operation.

A first estimate for the modulus of $L_{\tau_0}^{(1)}$ (as defined in (4.9))—which is effective for $|x| \le 1$ and which can be obtained using the v_0^d-moderateness of w (see Lemma 4.9) and the submultiplicativity of w_1, v_0—reads as follows:

$$
\begin{aligned}
&\left| \int_{\mathbb{R}^d} \frac{w(\iota + \tau_0)}{w(\tau_0)} \left(\theta_2 \cdot \overline{\mathbf{T}_\tau \theta_1}\right)(\iota)\, e_{\tau_0}(x, \iota)\, d\iota \right| \\
&\le \int_{\mathbb{R}^d} v_0^d(\iota) \cdot |\theta_2(\iota)| \cdot |\theta_1(\iota - \tau)|\ d\iota \\
&= w_1(\tau)^{-1} \cdot w_1(\tau) \cdot \int_{\mathbb{R}^d} v_0^d(\iota) \cdot |\theta_2(\iota)| \cdot |\theta_1(\iota - \tau)|\ d\iota \\
&\le w_1(\tau)^{-1} \int_{\mathbb{R}^d} \left| v_0^d(\iota) w_1(\iota) \theta_2(\iota) \right| |w_1(\tau - \iota)\theta_1(\iota - \tau)|\ d\iota \\
&\text{(w_1 is radial)} \le w_1(\tau)^{-1} \cdot \|\theta_1\|_{\mathbf{L}^2_{w_1}} \cdot \|\theta_2\|_{\mathbf{L}^2_{v_0^d w_1}} \le C_{\max} \cdot [w_1(\tau)]^{-1}.
\end{aligned}
\tag{4.39}
$$

The last step used $v_0 \ge 1$, such that $\|\theta_1\|_{\mathbf{L}^2_{w_1}} \le \|\theta_1\|_{\mathbf{L}^2_{w_2}}$ and likewise $\|\theta_2\|_{\mathbf{L}^2_{v_0^d w_1}} \le \|\theta_2\|_{\mathbf{L}^2_{w_2}}$.

To obtain an estimate which is effective for large $|x|$, we have to work harder: We fix some $\delta = \delta(d, \Phi, v_0) > 0$, such that $\delta v_0(\delta/(4d) \cdot e_1) < 1/\sqrt{d}$. Hence, we can apply Lemma 4.13 to obtain a sequence $(j_i)_{i \in \mathbb{Z}^d}$, with $j_i \in \underline{d}$, such that $\rho_x = x/|x| \in U_{j_i}^{(i)}$ for all $i \in \mathbb{Z}^d$, and

$$
\begin{aligned}
&\left| \int_{\mathbb{R}^d} \frac{w(\iota + \tau_0)}{w(\tau_0)} \left(\theta_2 \cdot \overline{\mathbf{T}_\tau \theta_1}\right)(\iota) \cdot e_{\tau_0}(x, \iota)\, d\iota \right| \\
&= \left| \sum_{i \in \mathbb{Z}^d} \int_{\mathbb{R}^d} \left(\square^k_{j_i, \tau_0, x}\ g_{i,\tau,\tau_0}\right)(\iota) \cdot e_{\tau_0}(x, \iota)\, d\iota \right| \\
&\le \int_{\mathbb{R}^d} \sum_{i \in \mathbb{Z}^d} \left| \left(\square^k_{j_i, \tau_0, x}\ g_{i,\tau,\tau_0}\right)(\iota) \right|\ d\iota \\
&= \sum_{j \in \underline{d}} \int_{\mathbb{R}^d} \sum_{\substack{i \in \mathbb{Z}^d \\ \text{s.t. } j_i = j}} \left| \left(\square^k_{j, \tau_0, x}\ g_{i,\tau,\tau_0}\right)(\iota) \right|\ d\iota,
\end{aligned}
\tag{4.40}
$$

for any $x \in \mathbb{R}^d \backslash \{0\}$, $\tau, \tau_0 \in \mathbb{R}^d$.

For $j_i = j$ (which implies $\rho_x \in U_{j_i}^{(i)} = U_j^{(i)}$) we further see that

$$
\begin{aligned}
&\left|\left(\Box_{j,\tau_0,x}^{k}\, g_{i,\tau,\tau_0}\right)(\iota)\right| \\
&\overset{\text{Lem. 4.15}}{\leq} D_{k,\delta} \cdot (2\pi|x|)^{-k} \cdot v_0^{3k}(\iota) \cdot \sum_{n=0}^{k} \left| \frac{\partial^n}{\partial \iota_j^n} g_{i,\tau,\tau_0}(\iota) \right| \\
&\overset{\text{Lem. 4.14}}{\leq} D_{k,\delta} \cdot \mathbb{1}_{B_{\delta/(4d)}(\iota_i)}(\iota) \cdot (2\pi|x|)^{-k} \cdot v_0^{d+3k}(\iota) \\
&\qquad \cdot \sum_{n=0}^{k} C_n \cdot \sum_{\substack{m_1,m_2 \in \mathbb{N}_0 \\ m_1+m_2 \leq n}} \left| \frac{\partial^{m_1}}{\partial \iota_j^{m_1}} \theta_2(\iota) \cdot \frac{\partial^{m_2}}{\partial \iota_j^{m_2}} \overline{\theta_1(\iota - \tau)} \right|.
\end{aligned}
$$

Note that constants above are independent of $i \in \mathbb{Z}^d$. Next, using the finite overlap property, (4.31), we get

$$
\sum_{\substack{i \in \mathbb{Z}^d \\ \text{s.t. } j_i = j}} \left|\left(\Box_{j,\tau_0,x}^{k}\, g_{i,\tau,\tau_0}\right)(\iota)\right| \leq \widetilde{C} \cdot (2\pi|x|)^{-k} v_0^{d+3k}(\iota) \cdot \sum_{\substack{m_1,m_2 \in \mathbb{N}_0 \\ m_1+m_2 \leq k}} \left| \frac{\partial^{m_1}}{\partial \iota_j^{m_1}} \theta_2(\iota) \cdot \frac{\partial^{m_2}}{\partial \iota_j^{m_2}} \overline{\theta_1(\iota - \tau)} \right|,
$$

where $\widetilde{C} := (k+1) \cdot (1+4d)^d \cdot D_{k,\delta} \cdot \max_{n=0,\ldots,k} C_n$. Insert this estimate into the final line of (4.40), apply the Cauchy-Schwarz inequality, and recall that w_1 is submultiplicative and satisfies $w_1(-\iota) = w_1(\iota)$, whence $1 = [w_1(\tau)]^{-1} \cdot w_1(\iota+\tau-\iota) \leq [w_1(\tau)]^{-1} \cdot w_1(\iota) \cdot w_1(\iota - \tau)$, to obtain

$$
\begin{aligned}
&\sum_{j \in \underline{d}} \int_{\mathbb{R}^d} \sum_{\substack{i \in \mathbb{Z}^d \\ \text{s.t. } j_i = j}} \left|\left(\Box_{j,\tau_0,x}^{k}\, g_{i,\tau,\tau_0}\right)(\iota)\right| \, d\iota \\
&\leq \widetilde{C} \cdot (2\pi|x|)^{-k} \sum_{j \in \underline{d}} \int_{\mathbb{R}^d} v_0^{d+3k}(\iota) \cdot \sum_{\substack{m_1,m_2 \in \mathbb{N}_0 \\ m_1+m_2 \leq k}} \left| \frac{\partial^{m_1}}{\partial \iota_j^{m_1}} \theta_2(\iota) \cdot \frac{\partial^{m_2}}{\partial \iota_j^{m_2}} \overline{\theta_1(\iota - \tau)} \right| \, d\iota \\
&\leq \widetilde{C} \cdot (2\pi|x|)^{-k} [w_1(\tau)]^{-1} \cdot \\
&\sum_{j \in \underline{d}} \sum_{\substack{m_1,m_2 \in \mathbb{N}_0 \\ m_1+m_2 \leq k}} \int_{\mathbb{R}^d} \left| v_0^{d+3k}(\iota) \cdot w_1(\iota) \frac{\partial^{m_1}}{\partial \iota_j^{m_1}} \theta_2(\iota) \cdot w_1(\iota - \tau) \frac{\partial^{m_2}}{\partial \iota_j^{m_2}} \overline{\theta_1(\iota - \tau)} \right| \, d\iota \\
&\leq \widetilde{C} \cdot (2\pi|x|)^{-k} w_1(\tau)^{-1} \sum_{j \in \underline{d}} \sum_{\substack{m_1,m_2 \in \mathbb{N}_0 \\ m_1+m_2 \leq k}} \left\| \frac{\partial^{m_1}}{\partial \iota_j^{m_1}} \theta_2 \right\|_{\mathbf{L}^2_{w_2}} \cdot \left\| \frac{\partial^{m_2}}{\partial \iota_j^{m_2}} \overline{\theta_1} \right\|_{\mathbf{L}^2_{w_1}}.
\end{aligned}
$$

Since all the involved sums are finite, so is the total number of summands. Moreover, the highest order partial derivatives that appear are $\frac{\partial^k}{\partial \iota_j^k}\theta_1$ and $\frac{\partial^k}{\partial \iota_j^k}\theta_2$, for arbitrary $j \in \underline{d}$. Hence, a joint maximization over $j \in \underline{d}$ and the partial derivatives of θ_1, θ_2 yields

$$\begin{aligned}
&\left| \int_{\mathbb{R}^d} \frac{w(\iota + \tau_0)}{w(\tau_0)} \left(\theta_2 \cdot \overline{\mathbf{T}_\tau \theta_1}\right)(\iota)\, e_{\tau_0}(x, \iota)\, d\iota \right| \\
&\quad \leq C' \cdot (2\pi |x|)^{-k} w_1(\tau)^{-1} \\
&\quad \max_{j \in \underline{d}} \left\{ \left(\max_{n=0,\ldots,k} \left\| \frac{\partial^n}{\partial \iota_j^n} \theta_1 \right\|_{\mathbf{L}^2_{w_1}} \right) \cdot \left(\max_{n=0,\ldots,k} \left\| \frac{\partial^n}{\partial \iota_j^n} \theta_2 \right\|_{\mathbf{L}^2_{w_2}} \right) \right\} \\
&\quad \leq C' \cdot (2\pi |x|)^{-k} \cdot [w_1(\tau)]^{-1} \cdot C_{\max},
\end{aligned} \tag{4.41}$$

for a suitable (large) constant $C' > 0$ Here, the last step used again that $w_1 \leq w_2$.

Now, define

$$F(x) := \begin{cases} C_{\max}, & \text{if } |x| < 1 \\ C' \cdot C_{\max} \cdot (2\pi |x|)^{-k}, & \text{else.} \end{cases}$$

It is not hard to see $|F(x)| \leq C'' \cdot C_{\max} \cdot (1 + |x|)^{-k}$ for some constant $C'' > 0$. Combining the inequalities (4.39) and (4.41), we obtain for all $x,\ \tau,\ \tau_0 \in \mathbb{R}^d$ that

$$\begin{aligned}
|L^{(1)}_{\tau_0}(x, \tau)| &= \left| \int_{\mathbb{R}^d} \frac{w(\iota + \tau_0)}{w(\tau_0)} \left(\theta_2 \cdot \overline{\mathbf{T}_\tau \theta_1}\right)(\iota)\, e_{\tau_0}(x, \iota)\, d\iota \right| \\
&\leq [w_1(\tau)]^{-1} \cdot F(x) \\
&\leq C'' \cdot C_{\max} \cdot (1 + |x|)^{-k} \cdot [w_1(\tau)]^{-1}.
\end{aligned}$$

If we collect all the hidden dependencies, then we note that the final constant C'' depends on $D_{k,\delta} = D_{k,\delta}(v_0)$ and $C_n = C_n(\delta', d)$, and also directly on $d,\ k$. However, the support radius δ' of the assumed partition of unity is derived directly from δ, d, where the largest valid choice of δ itself depends only on v_0, d, see Lemma 4.12 for both dependencies. Further, noting that $D_{k,\delta}(v_0)$ is increasing in δ (see proof of Lemma 4.15), we can choose, without loss of generality, the largest possible value of δ. Overall, C'' is a function of $d,\ k$ and v_0, as desired. □

4.3 Proof of Theorem 4.4

Recall that $w = \det A$ is $w_0 := v_0^d$-moderate (Lemma 4.9) and $v_0, v_1 \geq 1$ (see proof of Theorem 4.8), such that $w_2 \geq v_0^{d/2} = \sqrt{w_0}$ and $\theta_1, \theta_2 \in \mathbf{L}^2_{\sqrt{w_0}}(\mathbb{R}^d)$ follows. That m is Φ-compatible with dominating weight m^Φ is an immediate consequence of the inequality (4.6), i.e.,

$$m\big((y,\xi),(z,\eta)\big) \leq (1+|y-z|)^p \cdot v_1\big(\Phi(\xi)-\Phi(\eta)\big), \text{ for all } y,z\in\mathbb{R}^d \text{ and } \xi,\eta\in D,$$

and the choice of $p \in \mathbb{N}_0$ (in particular, $p = 0$ if $R_\Phi = \infty$).

Thus, Lemmas 4.6 and 4.7 can be applied, showing that

$$\|K_{\theta_1,\theta_2}\|_{\mathcal{B}_m} \leq \max_{\ell\in\{1,2\}} \operatorname*{ess\,sup}_{\tau_0\in\mathbb{R}^d} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} M(x,\tau)\cdot |L^{(\ell)}_{\tau_0}(x,\tau)|\,dx\,d\tau,$$

where

$$M(x,\tau) = \sup_{y\in\mathbb{R}^d,|y|\leq R|x|} (1+|y|)^p \cdot v_0^{d/2}(\tau)\cdot v_1(\tau).$$

Note that $M(x,\tau) \leq C_\Phi \cdot (1+|x|)^p \cdot v_0^{d/2}(\tau) \cdot v_1(\tau)$, where $C_\Phi := \max\big\{1, \sup_{\xi\in D}\|\mathrm{D}\Phi(\xi)\|^p\big\}$ if $p > 0$ and $C_\Phi := 1$ otherwise.

Define $w_1 : \mathbb{R}^d \to \mathbb{R}^+$, $\iota \mapsto (1+|\iota|)^{d+1}\cdot v_1(\iota)\cdot[v_0(\iota)]^{d/2}$. Since v_0, v_1 are submultiplicative and satisfy $v_\ell(-\iota) = v_\ell(\iota)$ for $\ell \in \{0,1\}$ and $\iota\in\mathbb{R}^d$, it is easy to see that w_1 satisfies the same two properties. Furthermore, $w_2(\iota) = w_1(\iota)\cdot[v_0(\iota)]^{d+3(d+p+1)}$, so that Theorem 4.8, with $k = d+p+1$, yields a constant $C = C(d, d+p+1, v_0) > 0$ satisfying

$$\begin{aligned}
\|K_{\theta_1,\theta_2}\|_{\mathcal{B}_m} &\leq \max_{\ell\in\{1,2\}} \operatorname*{ess\,sup}_{\tau_0\in\mathbb{R}^d} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} M(x,\tau)\cdot |L^{(\ell)}_{\tau_0}(x,\tau)|\,dx\,d\tau \\
&\leq C_\Phi \cdot \operatorname*{ess\,sup}_{\tau_0\in\mathbb{R}^d} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} (1+|x|)^p \cdot v_0^{d/2}(\tau)\cdot v_1(\tau)\cdot \max_{\ell\in\{1,2\}} |L^{(\ell)}_{\tau_0}(x,\tau)|\,dx\,d\tau \\
(\text{Thm. } 4.8) &\leq C C_\Phi C_{\max} \cdot \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} v_0^{d/2}(\tau)\cdot v_1(\tau)\cdot[w_1(\tau)]^{-1}\cdot(1+|x|)^{-(d+1)}\,d\tau\,dx \\
&\leq C C_\Phi C_{\max} \cdot \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} (1+|\tau|)^{-(d+1)}(1+|x|)^{-(d+1)}\,d\tau\,dx \\
&=: \widetilde{C}\cdot C_{\max} < \infty.
\end{aligned}$$

Here, the final constant $\widetilde{C} = \widetilde{C}(d,p,\Phi,v_0) > 0$ is finite, simply because $(1+|\cdot|)^{-(d+1)} \in \mathbf{L}^1(\mathbb{R}^d)$. Arguably, the dependence of $\widetilde{C}$ on v_0 could be expressed as a consequence of the dependence on Φ, but there may be cases where different choices of v_0 could be of interest, such that we prefer to keep it explicit. This concludes the proof. □

5 The Phase-Space Coverings Induced by the Warping Function Φ

To prepare for the estimation of $\|\mathrm{osc}_{\mathcal{V}^\delta,\Gamma}\|_{\mathcal{B}_m}$ we construct families of coverings $\mathcal{V}^\delta_\Phi = (V^\delta_i)_{i\in I}$ of the phase space Λ, induced by a given warping function Φ and study their properties. In the next section, we will show that $\|\mathrm{osc}_{\mathcal{V}^\delta,\Gamma}\|_{\mathcal{B}_m} \to 0$ as $\delta\to 0$, with $\mathrm{osc}_{\mathcal{V}^\delta,\Gamma}$ as introduced in Definition 2.16.

Definition 5.1 Let $\Phi\colon D \to \mathbb{R}^d$ be a warping function. Define

$$Q_{\Phi,\tau}^{(\delta,r)} := \Phi^{-1}(\delta \cdot B_r(\tau)), \quad \text{for all } r, \delta > 0 \text{ and } \tau \in \mathbb{R}^d. \tag{5.1}$$

We call $\mathcal{V}_\Phi^\delta = (V_{\ell,k}^\delta)_{\ell,k\in\mathbb{Z}^d}$, defined by

$$\begin{aligned} V_{\ell,k}^\delta &:= A^{-T}(\delta k/\sqrt{d})\Big\langle \delta \cdot B_1(\ell/\sqrt{d})\Big\rangle \times Q_k^\delta, \quad \text{with} \\ Q_k^\delta &:= Q_{\Phi,k/\sqrt{d}}^{(\delta,1)} = \Phi^{-1}\big(\delta \cdot B_1(k/\sqrt{d})\big), \end{aligned} \tag{5.2}$$

the Φ-*induced* δ-*fine (phase-space) covering*.

By allowing $r \neq 1$ in (5.1), it is possible to control the overlap of the covering elements. In particular, any radius strictly larger than $1/2$ provides a covering. For proving the feasibility of discretization in coorbit spaces, however, the above choice of $r = 1$ in (5.2) is completely sufficient.

Proposition 5.2 *Let Φ be a 0-admissible warping function with control weight v_0 (see Definition 4.2). Then the Φ-induced δ-fine phase-space covering $\mathcal{V}_\Phi^\delta = (V_{\ell,k}^\delta)_{\ell,k\in\mathbb{Z}^d}$ is a topologically admissible cover of $\Lambda = \mathbb{R}^d \times D$ which is also product-admissible as per Definition 2.4. More precisely, we have the following properties:*

(1) If $k, \ell, k_0, \ell_0 \in \mathbb{Z}^d$ satisfy $|k - k_0| > 2\sqrt{d}$, then $V_{\ell,k}^\delta \cap V_{\ell_0,k_0}^\delta = \varnothing$. Furthermore,

$$\begin{aligned} &\sup_{(\ell,k)\in\mathbb{Z}^{2d}} \#\big\{(\ell_0, k_0) \in \mathbb{Z}^d \times \mathbb{Z}^d \;:\; V_{\ell,k}^\delta \cap V_{\ell_0,k_0}^\delta \neq \varnothing\big\} \\ &\quad \le (1 + 4d)^d \big(1 + 2\sqrt{d} \cdot (1 + v_0(2\delta))\big)^d. \end{aligned}$$

(2) We have $[v_0(\delta)]^{-d} \le \frac{\mu(V_{\ell,k}^\delta)}{[\mu(B_1(0))]^2 \cdot \delta^{2d}} \le [v_0(\delta)]^d$ for all $k, \ell \in \mathbb{Z}^d$.
(3) We have $\mu(V_{\ell,k}^\delta)/\mu(V_{\ell_0,k_0}^\delta) \le [v_0(\delta)]^{2d}$ for arbitrary $\ell, k, \ell_0, k_0 \in \mathbb{Z}^d$.
(4) For each fixed $\delta > 0$, the weight $w_{\mathcal{V}_\Phi^\delta}$ as given in Eq. (2.8) satisfies

$$\begin{aligned} \big(w_{\mathcal{V}_\Phi^\delta}\big)_{\ell,k} &\asymp \min\big\{w(\delta \cdot k/\sqrt{d}), [w(\delta \cdot k/\sqrt{d})]^{-1}\big\} \\ &\asymp \min\big\{w(\Phi(\xi)), [w(\Phi(\xi))]^{-1}\big\} \gtrsim [v_0(\Phi(\xi))]^{-d}, \\ &\quad \textit{for all } \ell, k \in \mathbb{Z}^d,\ \xi \in Q_k^\delta. \end{aligned} \tag{5.3}$$

In particular, there exists a constant $C = C(d, \delta, v_0) > 0$ such that $(w_{\mathcal{V}_\Phi^\delta})_{\ell,k} \big/ (w_{\mathcal{V}_\Phi^\delta})_{\ell_0,k_0} \le C$ for all $\ell, k, \ell_0, k_0 \in \mathbb{Z}^d$ with $V_{\ell,k}^\delta \cap V_{\ell_0,k_0}^\delta \neq \varnothing$. Moreover, (2.9) holds with

$$w_{\mathcal{V}_\Phi^\delta}^c \;:\; \Lambda \to \mathbb{R}^+, \quad (x, \xi) \mapsto \min\big\{w(\Phi(\xi)),\ [w(\Phi(\xi))]^{-1}\big\}.$$

Proof Note that the family $\delta \cdot B_1(\ell/\sqrt{d})$, $\ell \in \mathbb{Z}^d$ forms a covering of $\mathbb{R}^d$, since $\frac{1}{\sqrt{d}}\big(\ell + [0,1)^d\big) \subset B_1(\ell/\sqrt{d})$. Considering that $\Phi : D \to \mathbb{R}^d$ is a diffeomorphism and $A^{-T}(\delta k/\sqrt{d})$, for any $k \in \mathbb{Z}^d$, is an invertible matrix, it follows that $\mathcal{V}_\Phi^\delta$ indeed covers all of Λ.

We first prove part (1). For $k, \ell \in \mathbb{Z}^d$, let

$$J_{\ell,k} := \left\{(\ell_0, k_0) \in \mathbb{Z}^d \times \mathbb{Z}^d \,:\, V_{\ell,k}^\delta \cap V_{\ell_0,k_0}^\delta \neq \varnothing\right\}.$$

If $V_{\ell,k}^\delta \cap V_{\ell_0,k_0}^\delta \neq \varnothing$, then in particular $Q_k^\delta \cap Q_{k_0}^\delta \neq \varnothing$. Straightforward calculations show that the latter implies $|k_0 - k| \leq 2\sqrt{d}$, and then $k_0 \in k + \{-2d, \dots, 2d\}^d$. Moreover, if $(\ell_0, k_0) \in J_{\ell,k}$, then an easy calculation shows that there exist $x_1, x_2 \in B_1(0)$ such that

$$\ell_0 = A^T\left(\delta k_0/\sqrt{d}\right) \cdot A^{-T}\left(\delta k/\sqrt{d}\right)\left\langle \ell + \sqrt{d} \cdot x_1\right\rangle - \sqrt{d} \cdot x_2.$$

Property (4.5) shows that $A_{k,k_0} := A^T\left(\delta k_0/\sqrt{d}\right) \cdot A^{-T}\left(\delta k/\sqrt{d}\right) = \phi_{\delta k/\sqrt{d}}(\delta \cdot (k_0 - k)/\sqrt{d})$ satisfies

$$\left\| A_{k,k_0} \right\| \leq v_0\left(\tfrac{\delta}{\sqrt{d}}(k_0 - k)\right) \leq v_0(2\delta).$$

Here, we used $|k_0 - k| \leq 2\sqrt{d}$ and that v_0 is radially increasing. Since $x_1, x_2 \in B_1(0)$, we thus have $\ell_0 \in A_{k,k_0}\ell + [-C_1, C_1]^d$, where

$$C_1 := \sqrt{d} \cdot \big(1 + v_0(2\delta)\big) \geq \left|\sqrt{d} \cdot A_{k,k_0} x_1 - \sqrt{d} \cdot x_2\right|.$$

Altogether, we have shown

$$J_{\ell,k} \subset \bigcup_{k_0 \in k + \{-2d,\dots,2d\}^d} \left(\left[\mathbb{Z}^d \cap \left(A_{k,k_0}\ell + [-C_1, C_1]^d\right)\right] \times \{k_0\}\right).$$

But we have $\#\big[\mathbb{Z}^d \cap \big(A_{k,k_0}\ell + [-C_1, C_1]^d\big)\big] \leq (1 + 2C_1)^d$ and hence

$$|J_{\ell,k}| \leq \sum_{k_0 \in k + \{-2d,\dots,2d\}^d} (1 + 2C_1)^d = (1 + 4d)^d (1 + 2C_1)^d,$$

completing the proof of part (1). This also shows that $\mathcal{V}_\Phi^\delta$ is an admissible covering. Since each $V_{\ell,k}^\delta$ is open and relatively compact in Λ, we see that $\mathcal{V}_\Phi^\delta$ is topologically admissible.

We proceed to prove Item (2). By the change of variables formula, cf. Eq. (3.4), we get

$$\mu(Q_k^\delta) = \int_D \mathbb{1}_{\delta \cdot B_1(k/\sqrt{d})}(\Phi(\xi))\, d\xi = \int_{\delta \cdot B_1(k/\sqrt{d})} w(\tau)\, d\tau.$$

Recall that w is v_0^d-moderate by Lemma 4.9, where v_0 is submultiplicative and radially increasing. Therefore,

$$[v_0(\delta)]^{-d} \le \frac{w(\tau)}{w(\delta \cdot k/\sqrt{d})} \le [v_0(\delta)]^d, \text{ for all } \tau \in \delta \cdot \overline{B_1}(k/\sqrt{d})$$

In combination, the two preceding displayed equations show that

$$\mu(Q_k^\delta) \in \mu(B_1(0)) \cdot \delta^d \cdot w(\delta \cdot k/\sqrt{d}) \cdot \Big[[v_0(\delta)]^{-d}, [v_0(\delta)]^d\Big]. \tag{5.4}$$

Moreover,

$$\begin{aligned} \mu\left(A^{-T}(\delta k/\sqrt{d})\langle \delta \cdot B_1(\ell/\sqrt{d})\rangle\right) &= \left|\det\left(A^{-T}(\delta \cdot k/\sqrt{d})\right)\right| \cdot \mu\left(\delta \cdot B_1(\ell/\sqrt{d})\right) \\ &= \frac{\mu(B_1(0)) \cdot \delta^d}{w(\delta \cdot k/\sqrt{d})}. \end{aligned} \tag{5.5}$$

Since $\mu(V_{\ell,k}^\delta) = \mu\big(A^{-T}(\delta k/\sqrt{d})\langle \delta \cdot B_1(\ell/\sqrt{d})\rangle\big) \cdot \mu(Q_k^\delta)$, this proves part (2). Finally, part (3) is a direct consequence of part (2).

It remains to prove part (4). Since $\mathcal{V}_\Phi^\delta$ is a covering of $\Lambda = \mathbb{R}^d \times D$ with countable index set and with each set $V_{\ell,k}^\delta$ being a Cartesian product of open sets, this will then imply that $\mathcal{V}_\Phi^\delta$ is product-admissible. First note that $\min\{1, \mu(V_{\ell,k}^\delta)\} \asymp 1$ as a function in $\ell, k \in \mathbb{Z}^d$ and that $V_{\ell,k}^\delta = V_{1,(\ell,k)}^\delta \times V_{2,(\ell,k)}^\delta$ with $V_{1,(\ell,k)}^\delta = A^{-T}(\delta k/\sqrt{d})\langle \delta \cdot B_1(\ell/\sqrt{d})\rangle$ and $V_{2,(\ell,k)}^\delta = Q_k^\delta$. Hence, by (5.5) and (5.4), we have

$$\min\left\{\mu(V_{1,(\ell,k)}^\delta),\ \mu(V_{2,(\ell,k)}^\delta)\right\} \asymp \min\left\{w(\delta k/\sqrt{d}), [w(\delta k/\sqrt{d})]^{-1}\right\}, \quad \text{as a function in } \ell, k \in \mathbb{Z}^d.$$

Together, this yields the first estimate in (5.3). The other two estimates in (5.3) are simple consequences of w being v_0^d-moderate (and thus w^{-1} is as well) and of the identity $\Phi(Q_k^\delta) = \delta \cdot B_1(k/\sqrt{d})$. Note that (5.3) implies (2.9) with the stated choice of $w^c_{\mathcal{V}_\Phi^\delta}$.

To prove that $(w_{\mathcal{V}_\Phi^\delta})_{\ell,k}/(w_{\mathcal{V}_\Phi^\delta})_{\ell_0,k_0} \lesssim 1$ if $V_{\ell,k}^\delta \cap V_{\ell_0,k_0}^\delta \neq \varnothing$, first note that since w is v_0^d-moderate and v_0 is radially increasing (and hence radial). Note that taking reciprocal values, as well as pointwise minima/maxima preserve moderateness relations, see Remark 2.6, such that

$$\frac{\min\left\{w(\delta k/\sqrt{d}),\,[w(\delta k/\sqrt{d})]^{-1}\right\}}{\min\left\{w(\delta k_0/\sqrt{d}),\,[w(\delta k_0/\sqrt{d})]^{-1}\right\}} \le \left[v_0(\delta(k-k_0)/\sqrt{d})\right]^d \qquad \forall\, k, k_0 \in \mathbb{Z}^d.$$

Furthermore, part (1) of the proposition shows that if $V^\delta_{\ell,k} \cap V^\delta_{\ell_0,k_0} \ne \varnothing$, then $|k-k_0| \le 2\sqrt{d}$. Combining these observations with Eq. (5.3) and with the fact that v_0 is radially increasing, we see $(w_{\mathcal{V}^\delta_\Phi})_{\ell,k}/(w_{\mathcal{V}^\delta_\Phi})_{\ell_0,k_0} \lesssim v_0(2\delta)^d \lesssim 1$, where the implied constant depends (only) on d, δ, and v_0. □

The next lemma is concerned with the sets $\mathbf{V}_\lambda = \bigcup_{i\in I \text{ s.t. } \lambda\in V_i} V_i$ defining the oscillation $\operatorname{osc}_{\mathcal{V},\Gamma}$, see Definition 2.16. For the induced coverings $\mathcal{V}^\delta_\Phi$, the set $\mathbf{V}^\delta_\lambda$ can once more be estimated by a convenient product set. Moreover, the lemma implies that if $\lambda = (z,\eta) \in \mathbf{V}^\delta_{\lambda_0}$, with $\lambda_0 = (y,\omega)$, then

$$|A^T(\Phi(\omega))\langle z-y\rangle| \le C_\delta \quad\text{ and }\quad |\Phi(\eta)-\Phi(\omega)| \le C_\delta, \quad\text{with}$$
$$C_\delta > 0 \text{ independent of } \lambda, \lambda_0.$$

In particular, this holds if there exists $(\ell,k) \in \mathbb{Z}^{2d}$ such that $\lambda, \lambda_0 \in V^\delta_{\ell,k}$. These estimates will be crucial for estimating $\|\operatorname{osc}_{\mathcal{V}^\delta,\Gamma}\|_{\mathcal{B}_m}$.

Lemma 5.3 *Let Φ be a warping function, and let $\mathcal{V}^\delta_\Phi$ be the Φ-induced δ-fine covering. For all $(y,\omega) \in \Lambda$ and all $\delta > 0$, we have*

$$\boldsymbol{V}^\delta_{(y,\omega)} = \bigcup_{\substack{(\ell,k) \text{ s.t.}\\ (y,\omega)\in V^\delta_{\ell,k}}} V^\delta_{\ell,k} \subset (y + \boldsymbol{P}^\delta_\omega) \times \boldsymbol{Q}^\delta_\omega, \tag{5.6}$$

where

$$\boldsymbol{Q}^\delta_\omega := \Phi^{-1}\big(\Phi(\omega) + B_{2\delta}(0)\big) \qquad \text{and} \qquad \boldsymbol{P}^\delta_\omega := v_0(\delta)\cdot A^{-T}(\Phi(\omega))\,\langle B_{2\delta}(0)\rangle \tag{5.7}$$

Proof Let $(\ell,k) \in \mathbb{Z}^d \times \mathbb{Z}^d$ be such that $(y,\omega) \in V^\delta_{\ell,k}$. Then $\delta k/\sqrt{d} \in \Phi(\omega) + \delta\cdot B_1(0)$ and by extension of that argument, $Q^\delta_k \subset \Phi^{-1}(\Phi(\omega) + 2\delta B_1(0)) = \mathbf{Q}^\delta_\omega$, which proves the first part of the claim.

Next, for $(x,\xi) \in V^\delta_{\ell,k}$, we have $|A^T(\delta k/\sqrt{d})\,\langle x-y\rangle| < 2\delta$, since $x, y \in A^{-T}(\delta k/\sqrt{d})\left\langle \delta B_1(\ell/\sqrt{d})\right\rangle$. Hence,

$$\begin{aligned}
\left|A^T(\Phi(\omega))\,\langle x-y\rangle\right| &= \left|A^T(\Phi(\omega))A^{-T}(\delta k/\sqrt{d})A^T(\delta k/\sqrt{d})\,\langle x-y\rangle\right| \\
&< 2\delta\cdot\left\|A^T(\Phi(\omega))A^{-T}(\delta k/\sqrt{d})\right\| = 2\delta\cdot\left\|\phi_{\delta k/\sqrt{d}}\big(\Phi(\omega)-\delta k/\sqrt{d}\big)\right\| \\
\text{(cf. Eq. (4.5))} &\le 2\delta\cdot v_0(\Phi(\omega)-\delta k/\sqrt{d}) \le 2\delta\cdot v_0(\delta),
\end{aligned}$$

which shows $x - y \in A^{-T}(\Phi(\omega))\,\langle 2\delta v_0(\delta) B_1(0)\rangle$, and thus $x \in y + \mathbf{P}^\delta_\omega$, as desired. □

Proposition 5.4 *Let Φ be a 0-admissible warping function with control weight v_0. Let further $m_0 : \Lambda \times \Lambda \to \mathbb{R}^+$ be continuous and symmetric, with $1 \leq m_0(\lambda, \rho) \leq C^{(0)} \cdot m_0(\lambda, \nu) \cdot m_0(\nu, \rho)$, for all $\lambda, \rho, \nu \in \Lambda$ and some $C^{(0)} \geq 1$, satisfy*

$$m_0((y, \xi), (z, \eta)) \leq (1 + |y - z|)^p \cdot \zeta_1\big(\Phi(\xi) - \Phi(\eta)\big) \ \ \forall (y, \xi), (z, \eta) \in \Lambda.$$

Here, $p=0$ if $R_\Phi = \sup_{\xi \in D} \|\mathrm{D}\Phi(\xi)\| = \infty$ and $p \in \mathbb{N}_0$ otherwise, and $\zeta_1 : \mathbb{R}^d \to \mathbb{R}^+$ is a continuous function with $\zeta_1(-\tau) = \zeta_1(\tau)$ for all $\tau \in \mathbb{R}^d$. Define, for some arbitrary, fixed $\nu \in \Lambda$,

$$\begin{aligned} u :&\quad \Lambda \to \mathbb{R}^+, \quad \lambda \mapsto m_0(\lambda, \nu) \qquad \textit{and} \\ v :&\quad \Lambda \to \mathbb{R}^+, \quad (y, \xi) \mapsto u(y, \xi) \cdot \max\left\{w(\Phi(\xi)), [w(\Phi(\xi))]^{-1}\right\} \end{aligned}$$

and let m_v be as in Eq. ((2.10)). Then u is $\mathcal{V}_\Phi^\delta$-moderate, for any $\delta > 0$, and m_0 and m_v are Φ-convolution-dominated by $(1 + |\bullet|)^p \cdot \zeta_1(\bullet)$ and $m_v^\Phi := (1 + |\bullet|)^p \cdot \zeta_2(\bullet)$, where $\zeta_2 = v_0^d \cdot \zeta_1$. In particular, items (1)–(3) of Assumption 2.11 are satisfied.

Proof. Proposition 5.2 provides product-admissibility of $\mathcal{V}_\Phi^\delta$, such that item (1) of Assumption 2.11 is satisfied. Item (3) is a direct consequence of the symmetry of m_0:

$$m_0(\lambda, \rho) \leq C^{(0)} \cdot m_0(\lambda, \nu) \cdot m_0(\nu, \rho) = C^{(0)} \cdot u(\lambda) \cdot u(\rho).$$

To show $\mathcal{V}_\Phi^\delta$-moderateness of u (which coincides with item (2) of Assumption 2.11), observe that

$$\frac{u(\lambda)}{u(\rho)} \leq C^{(0)} \frac{m_0(\lambda, \rho) m_0(\rho, \nu)}{m_0(\rho, \nu)} = C^{(0)} m_0(\lambda, \rho). \tag{5.8}$$

If $\lambda = (y, \xi)$ and $\rho = (z, \eta)$ are both contained in $V_{\ell,k}^\delta$, for some $\ell, k \in \mathbb{Z}^d$, then $|\Phi(\xi) - \Phi(\eta)| < \delta$, and

$$|y - z| \leq \delta \cdot \|A^{-T}(\delta k/\sqrt{d})\| \leq \delta \cdot R_\Phi, \quad \text{if } R_\Phi < \infty.$$

Hence, and $\frac{u(y,\xi)}{u(z,\eta)} \leq C^{(0)} m_0((y, \xi), (z, \eta)) \leq C^{(0)} (1 + \delta R_\Phi)^p \cdot \zeta_1(\delta)$, independent of $\ell, k \in \mathbb{Z}^d$. If $R_\Phi = \infty$, then $\frac{u(y,\xi)}{u(z,\eta)} \leq C^{(0)} m_0((y, \xi), (z, \eta)) \leq C^{(0)} \zeta_1(\delta)$ instead.

That m_0 is Φ-convolution-dominated by $(1 + |\bullet|)^p \cdot \zeta_1(\bullet)$ is immediate. To prove that m_v is Φ-convolution-dominated by m_v^Φ, observe

$$\frac{\max\left\{w(\tau), [w(\tau)]^{-1}\right\}}{\max\left\{w(\iota), [w(\iota)]^{-1}\right\}} \leq \max\left\{\frac{w(\tau)}{w(\iota)}, \frac{w(\iota)}{w(\tau)}\right\} \leq v_0^d(\tau - \iota), \text{ for all } \tau, \iota \in \mathbb{R}^d.$$

Combine the above with (5.8), such that

$$\frac{v(y, \Phi(\xi))}{v(z, \Phi(\eta))} \leq C^{(0)} \cdot m_0((y, \Phi(\xi)), (z, \Phi(\eta))) \cdot v_0^d(\Phi(\xi) - \Phi(\eta)). \qquad \square$$

6 Controlling the $\mathcal{B}_m$-Norm of the Oscillation

In this section, we employ the Φ-induced δ-fine phase-space coverings $\mathcal{V}_{\Phi}^{\delta}$, constructed in the previous section, to derive conditions concerning the prototype function θ which ensure that $\|\mathrm{osc}_{\mathcal{V}_{\Phi}^{\delta},\Gamma}\|_{\mathcal{B}_m} < \infty$ with $\|\mathrm{osc}_{\mathcal{V}_{\Phi}^{\delta},\Gamma}\|_{\mathcal{B}_m} \to 0$ as $\delta \to 0$. We will obtain the following result.

Theorem 6.1 *Let Φ be a $(d+p+1)$-admissible warping function with control weight v_0, where $p = 0$ if $R_{\Phi} = \sup_{\xi \in D} \|\mathrm{D}\Phi(\xi)\| = \infty$ and $p \in \mathbb{N}_0$ otherwise. Let furthermore $m : \Lambda \times \Lambda \to \mathbb{R}^+$ be a symmetric weight that satisfies*

$$m((y,\xi),(z,\eta)) \leq (1+|y-z|)^p \cdot v_1(\Phi(\xi)-\Phi(\eta)), \forall\, y,z \in \mathbb{R}^d \text{ and } \xi,\eta \in D, \tag{6.1}$$

for some continuous and submultiplicative weight $v_1 : \mathbb{R}^d \to \mathbb{R}^+$ satisfying $v_1(\iota) = v_1(-\iota)$ for all $\iota \in \mathbb{R}^d$.

Finally, with

$$w_2 : \mathbb{R}^d \to \mathbb{R}^+, \iota \mapsto (1+|\iota|)^{d+1} \cdot v_1(\iota) \cdot [v_0(\iota)]^{9d/2+3p+3},$$

assume that $\theta \in \mathcal{C}^{d+p+1}(\mathbb{R}^d)$ and

$$v_0^n \cdot \frac{\partial^{(d+p+1)-n}}{\partial \iota_j^{(d+p+1)-n}} \theta \in \mathbf{L}^2_{w_2}(\mathbb{R}^d), \qquad \text{for all } j \in \underline{d},\ \ 0 \leq n \leq d+p+1.$$

Then, with $\Gamma : \Lambda \times \Lambda \to \mathbb{C}, \big((y,\omega),(z,\eta)\big) \mapsto e^{-2\pi i \langle y-z,\omega\rangle}$, and $\mathcal{V}_{\Phi}^{\delta} = (V_{\ell,k}^{\delta})_{\ell,k\in\mathbb{Z}^d}$ the Φ-induced δ-fine covering:

$$\|\mathrm{osc}_{\mathcal{V}_{\Phi}^{\delta},\Gamma}\|_{\mathcal{B}_m} < \infty \ \text{ for all } \delta > 0 \qquad \text{and} \qquad \|\mathrm{osc}_{\mathcal{V}_{\Phi}^{\delta},\Gamma}\|_{\mathcal{B}_m} \overset{\delta\to 0}{\to} 0. \tag{6.2}$$

Remark 6.2 The conditions of Theorem 6.1 are largely the same as those for Theorem 4.4. The only difference is the appearance of an additional factor v_0^n, for certain $n \in \mathbb{N}_0$, in the conditions on θ. Since $v_0 \geq v_0(0)$, the conditions of Theorem 6.1 imply those of Theorem 4.4.

To prove Theorem 6.1, we study the second component of the oscillation, i.e., $g_\lambda - \Gamma(\lambda,\rho) g_\rho$, for $\rho \in \mathcal{V}_{\lambda}^{\delta}$. If we can bound certain weighted $\mathbf{L}^2$-norms of this difference and its derivatives uniformly in $\lambda \in \Lambda$ and $\rho \in \mathcal{V}_{\lambda}^{\delta}$, then we can show that $\mathrm{osc}_{\mathcal{V}_{\Phi}^{\delta},\Gamma} \in \mathcal{B}_m$ by a slight variation on Theorem 4.4. In fact, the estimates we obtain converge to 0 for $\delta \to 0$, such that we naturally obtain the second part of Eq. (6.2) as well.

6.1 Local Behavior of the Oscillating Component

In order to rely on the machinery we already developed in Sect. 4, it will be useful to rewrite $g_\lambda - \Gamma(\lambda, \rho) g_\rho$ as the warping of a function $\tilde{\theta}_{\lambda,\rho} \in \mathbf{L}^2_{\sqrt{w_0}}$ (dependent on $\lambda, \rho \in \Lambda$) derived from the prototype θ.

Proposition 6.3 *For $D \subset \mathbb{R}^d$ open, let $\Lambda = \mathbb{R}^d \times D$, and define the phase function Γ via*

$$\Gamma : \Lambda \times \Lambda \to \mathbb{C},\ ((y, \omega), (z, \eta)) \mapsto e^{-2\pi i \langle y-z, \omega\rangle}. \tag{6.3}$$

Let $\Phi : D \to \mathbb{R}^d$ be a warping function, assume $\theta \in \boldsymbol{L}^2_{\sqrt{w_0}}(\mathbb{R}^d)$ and denote $(g_{y,\omega})_{(y,\omega)\in\Lambda} = \mathcal{G}(\theta, \Phi)$ as usual. Then the identity

$$\widehat{g_{y,\omega}} - \Gamma((y, \omega), (z, \eta))\widehat{g_{z,\eta}} = e^{-2\pi i \langle y, \cdot\rangle} \cdot \left(w(\Phi(\omega))^{-1/2} \cdot \left(\boldsymbol{T}_{\Phi(\omega)} \tilde{\theta}_{(y,\omega),(z,\eta)} \right) \circ \Phi \right), \tag{6.4}$$

holds for all $(y, \omega), (z, \eta) \in \Lambda$, with

$$\tilde{\theta}_{(y,\omega),(z,\eta)} := \left(\theta - \sqrt{\frac{w(\Phi(\omega))}{w(\Phi(\eta))}} \cdot \mathrm{E}_{\Phi(\omega), A^T(\Phi(\omega))\langle y-z\rangle} \left(\boldsymbol{T}_{\Phi(\eta)-\Phi(\omega)} \theta \right) \right) \in \boldsymbol{L}^2_{\sqrt{w_0}}. \tag{6.5}$$

The operator $\mathrm{E}_{\tau,\varepsilon}$ in Eq. (6.5) is a multiplication operator defined by

$$\mathrm{E}_{\tau,\varepsilon} f := e^{2\pi i \langle A^{-T}(\tau)\langle\varepsilon\rangle, \Phi^{-1}(\cdot+\tau) - \Phi^{-1}(\tau)\rangle} \cdot f \quad \text{for all} \quad f : \mathbb{R}^d \to \mathbb{C} \text{ and } \tau, \varepsilon \in \mathbb{R}^d. \tag{6.6}$$

Proof. To see that $\tilde{\theta}_{(y,\omega),(z,\eta)} \in \mathbf{L}^2_{\sqrt{w_0}}$, note that $\mathbf{T}_{\Phi(\eta)-\Phi(\omega)}$ and $\mathrm{E}_{\tau,\varepsilon}$ are bounded operators on $\mathbf{L}^2_{\sqrt{w_0}}$ and that $\sqrt{w(\Phi(\omega))/w(\Phi(\eta))}$ is finite for all $\omega, \eta \in D$. Here, boundedness of $\mathbf{T}_x$ on $\mathbf{L}^2_{\sqrt{w_0}}$ is a consequence of (3.5), since w_0 is submultiplicative. To prove (6.4), note that, by definition,

$$\begin{aligned}
\left(\widehat{g_{y,\omega}} - \Gamma((y, \omega), (z, \eta))\widehat{g_{z,\eta}} \right)(\xi) &= e^{-2\pi i \langle y, \xi\rangle} g_\omega(\xi) - e^{-2\pi i \langle y-z, \omega\rangle} e^{-2\pi i \langle z, \xi\rangle} g_\eta(\xi) \\
&= e^{-2\pi i \langle y, \xi\rangle} \left(g_\omega - e^{-2\pi i \langle y-z, \omega-\cdot\rangle} g_\eta \right)(\xi),
\end{aligned}$$

and furthermore

$$
\begin{aligned}
& g_\omega - e^{-2\pi i\langle y-z,\omega-\cdot\rangle} g_\eta \\
& \quad = w(\Phi(\omega))^{-1/2} \left(\mathbf{T}_{\Phi(\omega)}\theta\right) \circ \Phi - e^{-2\pi i\langle y-z,\omega-\cdot\rangle} w(\Phi(\eta))^{-1/2} \left(\mathbf{T}_{\Phi(\eta)}\theta\right) \circ \Phi \\
& \quad = w(\Phi(\omega))^{-1/2} \left(\mathbf{T}_{\Phi(\omega)}\theta - \sqrt{\frac{w(\Phi(\omega))}{w(\Phi(\eta))}} \cdot e^{2\pi i\langle y-z,\Phi^{-1}(\cdot)-\omega\rangle} \mathbf{T}_{\Phi(\eta)}\theta\right) \circ \Phi \\
& \quad = w(\Phi(\omega))^{-1/2} \\
& \qquad \cdot \left(\mathbf{T}_{\Phi(\omega)} \left(\theta - \sqrt{\frac{w(\Phi(\omega))}{w(\Phi(\eta))}} \cdot e^{2\pi i\langle y-z,\Phi^{-1}(\cdot+\Phi(\omega))-\omega\rangle} \mathbf{T}_{\Phi(\eta)-\Phi(\omega)}\theta\right)\right) \circ \Phi \\
& \quad = w(\Phi(\omega))^{-1/2} \left(\mathbf{T}_{\Phi(\omega)}\tilde{\theta}_{(y,\omega),(z,\eta)}\right) \circ \Phi. \qquad \square
\end{aligned}
$$

Now that we can express $g_\lambda - \Gamma(\lambda,\rho) g_\rho$ through $\tilde{\theta}(\lambda,\rho)$, we aim to derive conditions on θ, such that Lemma 4.8 can be applied with $\theta_1 = \theta, \theta_2 = \tilde{\theta}(\lambda,\rho)$. In particular, we investigate the (uniform) continuity of the map $(\tau,\varepsilon) \mapsto \mathrm{E}_{\tau,\varepsilon}$, in the next lemma. Here, $\mathrm{E}_{\tau,\varepsilon}$ is considered as an operator on $\mathbf{L}^q_{\tilde{w}}(\mathbb{R}^d)$, for suitable weights $\tilde{w}$.

Lemma 6.4 *Let $q \in [1,\infty)$ and let $\tilde{w} : \mathbb{R}^d \to \mathbb{R}^+$ be a continuous weight function. Furthermore, assume that Φ is a k-admissible warping function with control weight v_0.*

The operator $\mathrm{E}_{\tau,\varepsilon} : \boldsymbol{L}^q_{\tilde{w}}(\mathbb{R}^d) \to \boldsymbol{L}^q_{\tilde{w}}(\mathbb{R}^d)$, $\tau,\varepsilon \in \mathbb{R}^d$, given by (6.6), is well-defined and has the following properties:

(1) *If $\vartheta \in \boldsymbol{L}^q_{\tilde{w}}(\mathbb{R}^d)$ with $\operatorname{supp}(\vartheta) \subset \overline{B_\delta}(0)$ for some $\delta > 0$, then*

$$
\|\vartheta - \mathrm{E}_{\tau,\varepsilon}\vartheta\|_{\boldsymbol{L}^q_{\tilde{w}}} \leq \sqrt{2[1-\cos\left(\pi \cdot \min\{1, 2|\varepsilon|\delta v_0(\delta)\}\right)]} \cdot \|\vartheta\|_{\boldsymbol{L}^q_{\tilde{w}}}. \tag{6.7}
$$

(2) *If $\vartheta \in \boldsymbol{L}^q_{\tilde{w}}(\mathbb{R}^d)$, then $\sup_{\tau\in\mathbb{R}^d} \|\vartheta - \mathrm{E}_{\tau,\varepsilon}\vartheta\|_{\boldsymbol{L}^q_{\tilde{w}}} \overset{\varepsilon\to 0}{\to} 0$.*

(3) *If $\vartheta \in \mathcal{C}^m(\mathbb{R}^d)$ for some $0 \leq m \leq k+1$, and if $j \in \underline{d}$ with*

$$
v_0^n \cdot \frac{\partial^{m-n}}{\partial \iota_j^{m-n}} \vartheta \in \boldsymbol{L}^q_{\tilde{w}}(\mathbb{R}^d) \text{ for all } 0 \leq n \leq m, \tag{6.8}
$$

then $\frac{\partial^\ell}{\partial \iota_j^\ell}\vartheta \in \boldsymbol{L}^q_{\tilde{w}}(\mathbb{R}^d)$ for $0 \leq \ell \leq m$, $\frac{\partial^m}{\partial \iota_j^m}(\mathrm{E}_{\tau,\varepsilon}\vartheta) \in \boldsymbol{L}^q_{\tilde{w}}(\mathbb{R}^d)$ for all $\tau,\varepsilon \in \mathbb{R}^d$, and

$$
\sup_{\tau\in\mathbb{R}^d} \left\| \frac{\partial^m}{\partial \iota_j^m}(\vartheta - \mathrm{E}_{\tau,\varepsilon}\vartheta) \right\|_{\boldsymbol{L}^q_{\tilde{w}}} \overset{\varepsilon\to 0}{\to} 0.
$$

Furthermore, for each $\varepsilon_0 > 0$*, there is a constant* $C_{m,\varepsilon_0} > 0$ *satisfying for all* $|\varepsilon| \leq \varepsilon_0$ *that*

$$\sup_{\tau \in \mathbb{R}^d} \left\| \frac{\partial^m}{\partial \iota_j^m} (\mathrm{E}_{\tau,\varepsilon} \vartheta) \right\|_{\boldsymbol{L}^q_{\tilde{w}}} \leq \left\| \frac{\partial^m}{\partial \iota_j^m} \vartheta \right\|_{\boldsymbol{L}^q_{\tilde{w}}} + C_{m,\varepsilon_0} \cdot |\varepsilon| \cdot \sum_{n=1}^{m} \left\| v_0^n \cdot \frac{\partial^{m-n}}{\partial \iota_j^{m-n}} \vartheta \right\|_{\boldsymbol{L}^q_{\tilde{w}}} < \infty. \tag{6.9}$$

Proof Assumption (6.8) implies $\frac{\partial^\ell}{\partial \iota_j^\ell} \vartheta \in \mathbf{L}^q_{\tilde{w}}(\mathbb{R}^d)$ for all $0 \leq \ell \leq n$, since v_0 is radially increasing. Now, to prove (1), note for arbitrary $\iota \in \mathbb{R}^d$ that

$$|\vartheta(\iota) - (\mathrm{E}_{\tau,\varepsilon} \vartheta)(\iota)| = \left| 1 - e^{2\pi i \langle \varepsilon, A^{-1}(\tau)(\Phi^{-1}(\iota+\tau) - \Phi^{-1}(\tau)) \rangle} \right| \cdot |\vartheta(\iota)|,$$

where $\operatorname{supp} \vartheta \subset \overline{B_\delta}(0)$, such that it suffices to estimate this expression for $|\iota| \leq \delta$. We begin by expressing the difference $\Phi^{-1}(\iota + \tau) - \Phi^{-1}(\tau)$ through the Jacobian $A = \mathrm{D}\Phi^{-1}$ of Φ^{-1} by using the directional derivative. This furnishes the following estimate:

$$\begin{aligned} |A^{-1}(\tau)(\Phi^{-1}(\iota + \tau) - \Phi^{-1}(\tau))| &= \left| \int_0^1 A^{-1}(\tau) A(\tau + r\iota) \langle \iota \rangle \, dr \right| \\ &\leq |\iota| \cdot \max_{r \in [0,1]} \|A^{-1}(\tau) A(\tau + r\iota)\| \overset{\iota \in \overline{B_\delta}(0)}{\leq} \delta \cdot v_0(\delta), \end{aligned}$$

where we used (4.5) in the last step. Therefore, $|\langle \varepsilon, A^{-1}(\tau)(\Phi^{-1}(\iota+\tau) - \Phi^{-1}(\tau)) \rangle| \leq |\varepsilon| \cdot \delta \cdot v_0(\delta)$.

Next, a simple calculation shows that $|1 - e^{\pi i r}| = \sqrt{2[1 - \cos(\pi r)]}$, which is an even function that is increasing on $[0, 1]$ and converges to 0 for $r \to 0$. Thus, we obtain

$$|(\vartheta - \mathrm{E}_{\tau,\varepsilon} \vartheta)(\iota)| \leq \sqrt{2\,[1 - \cos(\pi \cdot 2|\varepsilon| \delta v_0(\delta))]} \cdot |\vartheta(\iota)|$$

for all $0 \leq |\varepsilon| \leq \frac{1}{2\delta v_0(\delta)}$. For $|\varepsilon| > (2\delta v_0(\delta))^{-1}$ apply the trivial estimate $|1 - e^{\pi i r}| \leq 2 = \sqrt{2[1 - \cos(\pi)]}$ instead. This easily yields (6.7), in fact for any solid Banach space X, and not only for $\mathbf{L}^q_{\tilde{w}}$.

To prove (2), note that for a given $\vartheta \in \mathbf{L}^q_{\tilde{w}}(\mathbb{R}^d)$, we have $\|\vartheta - \vartheta_n\|_{\mathbf{L}^q_{\tilde{w}}} \to 0$ as $n \to \infty$ for the sequence $\vartheta_n = \vartheta \cdot \mathbb{1}_{\overline{B_n}(0)}$, by the dominated convergence theorem. Furthermore, for every $n \in \mathbb{N}$,

$$\begin{aligned} \sup_{\tau \in \mathbb{R}^d} \|\vartheta - \mathrm{E}_{\tau,\varepsilon} \vartheta\|_{\mathbf{L}^q_{\tilde{w}}} &\leq \|\vartheta - \vartheta_n\|_{\mathbf{L}^q_{\tilde{w}}} + \sup_{\tau \in \mathbb{R}^d} \left(\|\vartheta_n - \mathrm{E}_{\tau,\varepsilon} \vartheta_n\|_{\mathbf{L}^q_{\tilde{w}}} + \|\mathrm{E}_{\tau,\varepsilon} \vartheta_n - \mathrm{E}_{\tau,\varepsilon} \vartheta\|_{\mathbf{L}^q_{\tilde{w}}} \right) \\ &= 2\|\vartheta - \vartheta_n\|_{\mathbf{L}^q_{\tilde{w}}} + \sup_{\tau \in \mathbb{R}^d} \|\vartheta_n - \mathrm{E}_{\tau,\varepsilon} \vartheta_n\|_{\mathbf{L}^q_{\tilde{w}}}. \end{aligned}$$

For any $n \in \mathbb{N}$ and any $\varepsilon_0 > 0$, we can choose $\varepsilon_n > 0$ such that

$$3\|\vartheta\|_{\mathbf{L}^q_{\tilde{w}}} \cdot \sqrt{2[1 - \cos(\pi \cdot \min\{1, 2|\varepsilon_n| n v_0(n))\}} < \varepsilon_0.$$

Noting that $\operatorname{supp}(\vartheta_n) \subset \overline{B_n}(0)$ by definition, we can now apply (6.7) with $\delta = n$ and any $\varepsilon \in \overline{B_{\varepsilon_n}}(0)$ to obtain $\|\vartheta_n - \mathrm{E}_{\tau,\varepsilon}\vartheta_n\|_{\mathbf{L}^q_{\tilde{w}}} < \varepsilon_0/3$, for all $\tau \in \mathbb{R}^d$. If additionally, $n \in \mathbb{N}$ is such that $\|\vartheta - \vartheta_n\|_{\mathbf{L}^q_{\tilde{w}}} < \varepsilon_0/3$, then $\|\vartheta - \mathrm{E}_{\tau,\varepsilon}\vartheta\|_{\mathbf{L}^q_{\tilde{w}}} < \varepsilon_0$. Since $\varepsilon_0 > 0$ was arbitrary, we obtain

$$\forall\, \varepsilon_0 > 0\ \exists\, n \in \mathbb{N} \text{ and } \varepsilon_n > 0, \text{ such that } \varepsilon \in \overline{B_{\varepsilon_n}}(0) \text{ implies } \sup_{\tau \in \mathbb{R}^d} \|\vartheta - \mathrm{E}_{\tau,\varepsilon}\vartheta\|_{\mathbf{L}^q_{\tilde{w}}} < \varepsilon_0.$$

To prove (3), we first note that for $m = 0$, all claims in this part are easy consequences of the definitions and of item (2). Therefore, we can assume $m \in \underline{k+1}$. Apply Leibniz's rule to obtain

$$\frac{\partial^m}{\partial \iota_j^m}(\mathrm{E}_{\tau,\varepsilon}\vartheta)(\iota) = \sum_{n=0}^{m} \left(\binom{m}{n} \frac{\partial^n}{\partial \iota_j^n} \left(e^{2\pi i \langle A^{-T}(\tau)\langle\varepsilon\rangle, \Phi^{-1}(\cdot+\tau) - \Phi^{-1}(\tau)\rangle} \right)(\iota) \cdot \frac{\partial^{m-n}}{\partial \iota_j^{m-n}} \vartheta(\iota) \right). \tag{6.10}$$

Moreover, Faa Di Bruno's formula [28, Corollary 2.10]— a form of the chain rule for higher derivatives—yields for $n \in \underline{m}$ that

$$\begin{aligned} \frac{\partial^n}{\partial \iota_j^n} \left(e^{2\pi i \langle A^{-T}(\tau)\langle\varepsilon\rangle, \Phi^{-1}(\cdot+\tau) - \Phi^{-1}(\tau)\rangle} \right)(\iota) &= e^{2\pi i \langle A^{-T}(\tau)\langle\varepsilon\rangle, \Phi^{-1}(\iota+\tau) - \Phi^{-1}(\tau)\rangle} \cdot P_{n,\tau,\varepsilon}(\iota) \\ &= \mathrm{E}_{\tau,\varepsilon} P_{n,\tau,\varepsilon}(\iota), \end{aligned}$$

where

$$\begin{aligned} &P_{n,\tau,\varepsilon}(\iota) \\ &= \sum_{\ell=1}^{n} \left((2\pi i)^\ell \cdot \sum_{\sigma \in (\underline{n-\ell+1})^\ell} \left(C_\sigma \cdot \prod_{i=1}^{\ell} \frac{\partial^{\sigma_i}}{\partial \iota_j^{\sigma_i}} \left\langle A^{-T}(\tau)\langle\varepsilon\rangle, \Phi^{-1}(\iota+\tau) - \Phi^{-1}(\tau) \right\rangle \right) \right) \\ &= \sum_{\ell=1}^{n} \left((2\pi i)^\ell \cdot \sum_{\sigma \in (\underline{n-\ell+1})^\ell} \left(C_\sigma \cdot \prod_{i=1}^{\ell} \frac{\partial^{\sigma_i}}{\partial \iota_j^{\sigma_i}} \left\langle \varepsilon, A^{-1}(\tau)\langle\Phi^{-1}(\iota+\tau)\rangle \right\rangle \right) \right), \end{aligned} \tag{6.11}$$

for suitable constants $C_\sigma \geq 0$. For the second equality, note that $\sigma_i \geq 1$ for all i, so that the term $\langle A^{-T}(\tau)\langle\varepsilon\rangle, \Phi^{-1}(\tau)\rangle$ —which is constant with respect to ι—can be ignored. In fact, the main statement of Faa Di Bruno's formula is exactly which C_σ are nonzero and what value they attain, see also Lemma 8.6, but these details are not required here. Similar to (4.27), we have that

$$\begin{aligned}\frac{\partial^{\sigma_i}}{\partial \iota_j^{\sigma_i}}\left\langle \varepsilon, A^{-1}(\tau)\langle \Phi^{-1}(\iota+\tau)\rangle\right\rangle &= \frac{\partial^{\sigma_i-1}}{\partial \iota_j^{\sigma_i-1}}\left\langle \varepsilon, A^{-1}(\tau)A(\iota+\tau)\langle e_i\rangle\right\rangle \\ &= \frac{\partial^{\sigma_i-1}}{\partial \iota_j^{\sigma_i-1}}\left(\left[A^{-1}(\tau)A(\iota+\tau)\right]^T \varepsilon\right)_i \\ &= \left(\frac{\partial^{\sigma_i-1}}{\partial \iota_j^{\sigma_i-1}}\phi_\tau(\iota)\langle\varepsilon\rangle\right)_i ,\end{aligned}$$

where $\phi_\tau = \left[A^{-1}(\tau)A(\cdot+\tau)\right]^T$ is as in (4.4). By (4.5), we can estimate

$$\left|\left(\frac{\partial^{\sigma_i-1}}{\partial \iota_j^{\sigma_i-1}}\phi_\tau(\iota)\langle\varepsilon\rangle\right)_i\right| \le \left\|\frac{\partial^{\sigma_i-1}}{\partial \iota_j^{\sigma_i-1}}\phi_\tau(\iota)\right\| \cdot |\varepsilon| \le v_0(\iota)\cdot|\varepsilon|$$

and inserting this into (6.11),

$$\begin{aligned}|P_{n,\tau,\varepsilon}(\iota)| &\le \sum_{\ell=1}^{n}\left((2\pi\cdot v_0(\iota)\cdot|\varepsilon|)^\ell \cdot \sum_{\sigma\in\underline{(n-\ell+1)}^\ell} C_\sigma\right) \\ &\le \tilde{C}\cdot|\varepsilon|\cdot\sum_{\ell=1}^{n}\left(v_0(\iota)^\ell\cdot|\varepsilon|^{\ell-1}\right),\end{aligned}$$

for a suitably large $\tilde{C} = \tilde{C}(n) > 0$. Since we only consider $n \in \underline{m}$, we can in fact choose the same constant $\tilde{C}$ for all values of n. Moreover, $1 \le v_0^\ell \le v_0^n$ for all $\ell \le n$.

By assembling all the pieces and by separating the term $n = 0$ in (6.10), we thus get

$$\begin{aligned}&\left|\frac{\partial^m}{\partial \iota_j^m}(\mathrm{E}_{\tau,\varepsilon}\vartheta)(\iota) - \mathrm{E}_{\tau,\varepsilon}\left(\frac{\partial^m}{\partial \iota_j^m}\vartheta\right)(\iota)\right| \\ &\le \sum_{n=1}^{m}\binom{m}{n}\left|(\mathrm{E}_{\tau,\varepsilon}P_{n,\tau,\varepsilon})(\iota)\cdot\left(\frac{\partial^{m-n}}{\partial \iota_j^{m-n}}\vartheta\right)(\iota)\right| \\ &= \sum_{n=1}^{m}\binom{m}{n}\left|P_{n,\tau,\varepsilon}(\iota)\cdot\left(\frac{\partial^{m-n}}{\partial \iota_j^{m-n}}\vartheta\right)(\iota)\right| \\ &\le |\varepsilon|\cdot\sum_{n=1}^{m}\left(\left|\frac{\partial^{m-n}}{\partial \iota_j^{m-n}}\vartheta(\iota)\right|\cdot\sum_{\ell=1}^{n}\left(\tilde{C}\binom{m}{n}v_0(\iota)^\ell\cdot|\varepsilon|^{\ell-1}\right)\right) \\ &\le |\varepsilon|\cdot\sum_{n=1}^{m}\left(\left|\left(v_0^n\cdot\frac{\partial^{m-n}}{\partial \iota_j^{m-n}}\vartheta(\iota)\right)\right|\cdot\sum_{\ell=1}^{n}\left(\tilde{C}\binom{m}{n}\cdot|\varepsilon|^{\ell-1}\right)\right).\end{aligned}$$

Let $0 \leq C_{m,\varepsilon} := \max_{n \in \underline{m}} \left(\sum_{\ell=1}^{n} \binom{m}{n} \cdot \tilde{C} |\varepsilon|^{\ell-1} \right) < \infty$ to obtain the estimate

$$\left| \frac{\partial^m}{\partial \iota_j^m} (\mathrm{E}_{\tau,\varepsilon} \vartheta)(\iota) - \mathrm{E}_{\tau,\varepsilon} \left(\frac{\partial^m}{\partial \iota_j^m} \vartheta \right)(\iota) \right| \leq C_{m,\varepsilon} \cdot |\varepsilon| \cdot \sum_{n=1}^{m} \left| v_0^n \cdot \frac{\partial^{m-n}}{\partial \iota_j^{m-n}} \vartheta(\iota) \right|.$$

Since $\mathbf{L}^q_{\tilde{w}}$ is solid, we conclude

$$\left\| \frac{\partial^m}{\partial \iota_j^m} (\mathrm{E}_{\tau,\varepsilon} \vartheta) - \mathrm{E}_{\tau,\varepsilon} \left(\frac{\partial^m}{\partial \iota_j^m} \vartheta \right) \right\|_{\mathbf{L}^q_{\tilde{w}}} \leq C_{m,\varepsilon} \cdot |\varepsilon| \cdot \sum_{n=1}^{m} \left\| v_0^n \cdot \frac{\partial^{m-n}}{\partial \iota_j^{m-n}} \vartheta \right\|_{\mathbf{L}^q_{\tilde{w}}} < \infty. \tag{6.12}$$

Finally, with $C_{m,\varepsilon} \leq C_{m,\varepsilon_0 \cdot e_1} =: C_{m,\varepsilon_0}$ for $|\varepsilon| \leq \varepsilon_0$, we obtain

$$\begin{aligned}
&\sup_{\tau \in \mathbb{R}^d} \left\| \frac{\partial^m}{\partial \iota_j^m} (\vartheta - \mathrm{E}_{\tau,\varepsilon} \vartheta) \right\|_{\mathbf{L}^q_{\tilde{w}}} \\
&\quad \leq \sup_{\tau \in \mathbb{R}^d} \left(\left\| \frac{\partial^m}{\partial \iota_j^m} \vartheta - \mathrm{E}_{\tau,\varepsilon} \left(\frac{\partial^m}{\partial \iota_j^m} \vartheta \right) \right\|_{\mathbf{L}^q_{\tilde{w}}} \right. \\
&\qquad \left. + \left\| \mathrm{E}_{\tau,\varepsilon} \left(\frac{\partial^m}{\partial \iota_j^m} \vartheta \right) - \frac{\partial^m}{\partial \iota_j^m} (\mathrm{E}_{\tau,\varepsilon} \vartheta) \right\|_{\mathbf{L}^q_{\tilde{w}}} \right) \xrightarrow[|\varepsilon| \to 0]{} 0,
\end{aligned}$$

as a consequence of part (2), and (6.12).

To prove (6.9) (and thus also $\frac{\partial^m}{\partial \iota_j^m} (\mathrm{E}_{\tau,\varepsilon} \vartheta) \in \mathbf{L}^q_{\tilde{w}}$), observe $\| \mathrm{E}_{\tau,\varepsilon} f \|_{\mathbf{L}^q_{\tilde{w}}} = \| f \|_{\mathbf{L}^q_{\tilde{w}}}$ for all $f \in \mathbf{L}^q_{\tilde{w}}(\mathbb{R}^d)$. By Eq. (6.12) the triangle inequality for norms yields

$$\begin{aligned}
\sup_{\tau \in \mathbb{R}^d} \left\| \frac{\partial^m}{\partial \iota_j^m} (\mathrm{E}_{\tau,\varepsilon} \vartheta) \right\|_{\mathbf{L}^q_{\tilde{w}}} &\leq \sup_{\tau \in \mathbb{R}^d} \left\| \mathrm{E}_{\tau,\varepsilon} \left(\frac{\partial^m}{\partial \iota_j^m} \vartheta \right) \right\|_{\mathbf{L}^q_{\tilde{w}}} + C_{m,\varepsilon} \cdot |\varepsilon| \cdot \sum_{n=1}^{m} \left\| v_0^n \cdot \frac{\partial^{m-n}}{\partial \iota_j^{m-n}} \vartheta \right\|_{\mathbf{L}^q_{\tilde{w}}} \\
&= \left\| \frac{\partial^m}{\partial \iota_j^m} \vartheta \right\|_{\mathbf{L}^q_{\tilde{w}}} + C_{m,\varepsilon} \cdot |\varepsilon| \cdot \sum_{n=1}^{m} \left\| v_0^n \cdot \frac{\partial^{m-n}}{\partial \iota_j^{m-n}} \vartheta \right\|_{\mathbf{L}^q_{\tilde{w}}}.
\end{aligned}$$

This proves (6.9), since $C_{m,\varepsilon} \leq C_{m,\varepsilon_0}$ as noted before. □

We now show that $\tilde{\theta}_{(y,\omega),(z,\eta)}$ uniformly converges to 0 as $\delta \to 0$, for $(y, \omega) \in \Lambda$ and $(z, \eta) \in (y + \mathbf{P}^\delta_\omega) \times \mathbf{Q}^\delta_\omega$. Recall that $(y + \mathbf{P}^\delta_\omega) \times \mathbf{Q}^\delta_\omega$ was introduced in Lemma 5.3 as a simple superset to $\mathbf{V}^\delta_{(y,\omega)} = \bigcup_{V_{\ell,k} \ni (y,\omega)} V_{\ell,k}$, appearing in the oscillation. The considered notion of convergence is in terms of the $\mathbf{L}^q_{\tilde{w}}$-norm of certain derivatives of $\tilde{\theta}_{(y,\omega),(z,\eta)}$. With Lemma 6.4, obtaining the desired estimates for $\tilde{\theta}_{(y,\omega),(z,\eta)}$ amounts to little more than an application of the triangle inequality and a somewhat elaborate three-ε-argument.

Lemma 6.5 *Let $q \in [1,\infty)$ and let $\tilde{w} : \mathbb{R}^d \to \mathbb{R}^+$ be a continuous, submultiplicative weight function. Furthermore, assume that Φ is a k-admissible warping function with control weight v_0. If*

$$\theta \in \mathcal{C}^m(\mathbb{R}^d) \text{ for some } 0 \le m \le k+1, \quad \text{and}$$
$$v_0^n \cdot \frac{\partial^{m-n}}{\partial \iota_j^{m-n}} \theta \in \boldsymbol{L}^q_{\tilde{w}}(\mathbb{R}^d) \text{ for all } 0 \le n \le m,\ j \in \underline{d},$$

then

$$\frac{\partial^m}{\partial \iota_j^m} \tilde{\theta}_{(y,\omega),(z,\eta)} = \frac{\partial^m}{\partial \iota_j^m} \left(\theta - \sqrt{\frac{w(\Phi(\omega))}{w(\Phi(\eta))}} \mathrm{E}_{\Phi(\omega), A^T(\Phi(\omega))\langle y-z\rangle} \left(\boldsymbol{T}_{\Phi(\eta)-\Phi(\omega)} \theta \right) \right) \in \boldsymbol{L}^q_{\tilde{w}}(\mathbb{R}^d) \tag{6.13}$$

for all $(y,\omega), (z,\eta) \in \Lambda$, and $j \in \underline{d}$. Furthermore, with

$$F_{j,m}(\delta;\theta,q,\tilde{w}) := \sup_{(y,\omega)\in\Lambda} \ \sup_{z\in(y+\boldsymbol{P}^\delta_\omega),\eta\in\boldsymbol{Q}^\delta_\omega} \left\| \frac{\partial^m}{\partial \iota_j^m} \tilde{\theta}_{(y,\omega),(z,\eta)} \right\|_{\boldsymbol{L}^q_{\tilde{w}}}, \tag{6.14}$$

where $\boldsymbol{Q}^\delta_\omega$ and $\boldsymbol{P}^\delta_\omega$ are as in Lemma 5.3, we have

$$F_{j,m}(\delta;\theta,q,\tilde{w}) < \infty \quad \text{for all } \delta > 0, \qquad \text{and} \quad F_{j,m}(\delta;\theta,q,\tilde{w}) \xrightarrow[\delta\to 0]{} 0. \tag{6.15}$$

Proof Since v_0 and $\tilde{w}$ are submultiplicative, so is $v_0^n\tilde{w}$, and $\mathbf{L}^q_{v_0^n\tilde{w}}(\mathbb{R}^d)$ is translation-invariant, see (3.5). Hence, since $\frac{\partial^{m-n}}{\partial \iota_j^{m-n}}\theta \in \mathbf{L}^q_{v_0^n\tilde{w}}(\mathbb{R}^d)$, $0 \le n \le m$ and $i \in \underline{d}$, the same holds for arbitrary translates. Thus, Lemma 6.4(3) shows $\frac{\partial^m}{\partial \iota_j^m}\theta$, $\frac{\partial^m}{\partial \iota_j^m}\mathrm{E}_{\tau,\varepsilon}(\mathbf{T}_{\tau_0}\theta) \in \mathbf{L}^q_{\tilde{w}}(\mathbb{R}^d)$ for all $\tau_0, \tau, \varepsilon \in \mathbb{R}^d$. This establishes (6.13), since $\frac{w(\Phi(\omega))}{w(\Phi(\eta))} < \infty$.

Fix $\delta > 0$ and $(y,\omega) \in \Lambda$ and $(z,\eta) \in (y+\mathbf{P}^\delta_\omega)\times\mathbf{Q}^\delta_\omega$. For brevity, set $\tau := \Phi(\omega) - \Phi(\eta)$ and $\varepsilon := A^T(\Phi(\omega))\langle y-z\rangle$, noting that $\tau \in B_{2\delta}(0)$ and $\varepsilon \in A^T(\Phi(\omega))\langle \mathbf{P}^\delta_\omega\rangle = B_{2\delta v_0(\delta)}(0) =: B_{\varepsilon_\delta}(0)$. In particular, $\varepsilon_\delta \le \varepsilon_{\delta_0}$, for all $\delta \le \delta_0$, and $\varepsilon_\delta \to 0$ as $\delta \to 0$. Recall the definition of $\tilde{\theta}_{(y,\omega),(z,\eta)}$ (given in (6.5)), and apply the triangle inequality twice to obtain the estimate

$$\begin{aligned}\left\| \frac{\partial^m}{\partial \iota_j^m} \tilde{\theta}_{(y,\omega),(z,\eta)} \right\|_{\mathbf{L}^q_{\tilde{w}}} &\leq \left| 1 - \sqrt{\frac{w(\Phi(\omega))}{w(\Phi(\eta))}} \right| \cdot \left\| \frac{\partial^m}{\partial \iota_j^m} \theta \right\|_{\mathbf{L}^q_{\tilde{w}}} \\ &+ \sqrt{\frac{w(\Phi(\omega))}{w(\Phi(\eta))}} \cdot \left\| \frac{\partial^m}{\partial \iota_j^m} \left(\theta - \mathrm{E}_{\Phi(\omega),\varepsilon} \theta\right) \right\|_{\mathbf{L}^q_{\tilde{w}}} \\ &+ \sqrt{\frac{w(\Phi(\omega))}{w(\Phi(\eta))}} \cdot \left\| \frac{\partial^m}{\partial \iota_j^m} \mathrm{E}_{\Phi(\omega),\varepsilon} (\theta - \mathbf{T}_{\Phi(\eta)-\Phi(\omega)} \theta) \right\|_{\mathbf{L}^q_{\tilde{w}}}. \end{aligned} \tag{6.16}$$

Next, Lemma 6.4(3) yields

$$\begin{aligned} E_\delta &:= \sup_{|\varepsilon| \leq \varepsilon_\delta} \sup_{\omega \in D} \left\| \frac{\partial^m}{\partial \iota_j^m} \left(\theta - \mathrm{E}_{\Phi(\omega),\varepsilon} \theta\right) \right\|_{\mathbf{L}^q_{\tilde{w}}} \leq \infty, \text{ for all } \delta > 0, \\ &\text{with } E_\delta \to 0 \text{ as } \delta \to 0, \text{ and} \end{aligned} \tag{6.17}$$

$$\begin{aligned} F_\delta &:= \sup_{|\varepsilon| \leq \varepsilon_\delta} \sup_{\omega \in D} \left\| \frac{\partial^m}{\partial \iota_j^m} \mathrm{E}_{\Phi(\omega),\varepsilon} (\theta - \mathbf{T}_{-\tau} \theta) \right\|_{\mathbf{L}^q_{\tilde{w}}} \\ &\leq \left\| \frac{\partial^m}{\partial \iota_j^m} \theta - \mathbf{T}_{-\tau} \left(\frac{\partial^m}{\partial \iota_j^m} \theta \right) \right\|_{\mathbf{L}^q_{\tilde{w}}} + C_{m,\varepsilon_{\delta_0}} \cdot \varepsilon_\delta \cdot \sum_{n=1}^{m} \left\| v_0^n \cdot \frac{\partial^{m-n}}{\partial \iota_j^{m-n}} (\theta - \mathbf{T}_{-\tau} \theta) \right\|_{\mathbf{L}^q_{\tilde{w}}}. \end{aligned} \tag{6.18}$$

Note that the first term of the right-hand side of (6.18) converges to 0 for $\delta \to 0$, since $|\tau| \leq 2\delta$ and translation is continuous in $\mathbf{L}^q_{\tilde{w}}$, since $\tilde{w}$ is continuous and submultiplicative. Furthermore, the sum over n in the right-hand side of (6.18) is finite, since $\mathbf{L}^q_{\tilde{w}}$ is translation-invariant and hence, all summands are finite by assumption. Therefore, F_δ vanishes for $\delta \to 0$. In fact, since $|\varepsilon| \leq \varepsilon_\delta$ and w is v_0^d-moderate with radially increasing v_0 (cf. Lemma 4.9), $\frac{w(\Phi(\omega))}{w(\Phi(\eta))} \leq v_0^d(\varepsilon_\delta)$, which settles the desired convergence of the second and third term in (6.16).

To settle convergence of the first term, we need to show that $\frac{w(\Phi(\omega))}{w(\Phi(\eta))} \overset{\delta \to 0}{\to} 1$, uniformly with respect to $\omega \in D$, $\eta \in \mathbf{Q}^\delta_\omega$. To this end, note that

$$\begin{aligned} &\frac{w(\Phi(\omega))}{w(\Phi(\eta))} \\ &\quad = \frac{w(\Phi(\eta)) + \int_0^1 \frac{d}{dt}\big|_{t=s} [w(\Phi(\eta) + s\tau)] \, ds}{w(\Phi(\eta))} \leq 1 + \frac{\sup_{\iota \in B_{2\delta}(\Phi(\eta))} \nabla_\tau w(\iota)}{w(\Phi(\eta))}, \end{aligned}$$

where ∇_τ denotes the derivative in direction $\tau \in \mathbb{R}^d$. We now use Jacobi's formula

$$\frac{d}{dt} \det A(t) = \det A(t) \cdot \operatorname{trace}([A(t)]^{-1} \cdot A'(t)),$$

valid for the derivative of the determinant (see [70, Section 8.3, Equation (2)]) of any differentiable function $M\colon I \to \mathrm{GL}(\mathbb{R}^d)$, where $I \subset \mathbb{R}$, to obtain

$$\begin{aligned}
\nabla_\tau w(\iota) &= \sum_{j\in\underline{d}} \tau_j \frac{\partial}{\partial \iota_j} \det(A(\iota)) = \det(A(\iota)) \cdot \sum_{j\in\underline{d}} \tau_j \cdot \operatorname{trace}\left(A^{-1}(\iota) \frac{\partial}{\partial \iota_j} A(\iota) \right) \\
&= w(\iota) \cdot \sum_{j\in\underline{d}} \tau_j \cdot \operatorname{trace}\left(\left(\frac{\partial}{\partial \eta_i}\bigg|_{\eta=0} A^{-1}(\iota) A(\iota+\eta) \right)^T \right) \\
&= w(\iota) \cdot \sum_{j\in\underline{d}} \tau_j \cdot \operatorname{trace}\left(\frac{\partial}{\partial \eta_i}\bigg|_{\eta=0} \phi_\iota(\eta) \right),
\end{aligned}$$

with ϕ_ι as in (4.4). Note that $\phi_\iota(0) = \mathrm{id}$ for all $\iota \in \mathbb{R}^d$, so that (4.5) yields $\|(\partial_i \phi_\iota)(0)\| \leq v_0(0)$. Additionally, the trace of a matrix $M \in \mathbb{R}^{d\times d}$ can be (coarsely) estimated by $|\operatorname{trace}(M)| \leq d\|M\|$, such that

$$|\nabla_\tau w(\iota)| \leq d \cdot w(\iota) \cdot \sum_{j\in\underline{d}} v_0(0) \cdot |\tau_j| \leq d \cdot w(\iota) \cdot \|\tau\|_1 \cdot v_0(0) \leq d^{3/2} \cdot w(\iota) \cdot |\tau| \cdot v_0(0).$$

Therefore, with $|\tau| \leq 2\delta$ and v_0^d-moderateness of w,

$$\begin{aligned}
\left| 1 - \frac{w(\Phi(\omega))}{w(\Phi(\eta))} \right| &\leq |\tau| \cdot d^{3/2} \cdot v_0(0) \cdot \max_{r\in[0,1]} \frac{w(\Phi(\eta) + r\tau)}{w(\Phi(\eta))} \\
&\leq 2\delta \cdot d^{3/2} \cdot v_0^d(2\delta) \cdot v_0(0) =: C^\delta < \infty. \qquad (6.19)
\end{aligned}$$

The final estimate is independent of $\omega \in D$, and of $\eta \in \mathbf{Q}_\omega^\delta$, and $C^\delta \to 0$ as $\delta \to 0$. □

We are now ready to prove Theorem 6.1.

6.2 Proof of Theorem 6.1

Recall that, by Remark 2.17, $\operatorname{osc}_{\mathcal{V}_\Phi^\delta,\Gamma}$ is continuous. Using Proposition 6.3 and Parseval's formula, we can rewrite the oscillation at $((y,\omega),(z,\eta)) \in \Lambda \times \Lambda$, as follows:

$$\begin{aligned}
\operatorname{osc}_{\mathcal{V}_\Phi^\delta,\Gamma}((y,\omega),(z,\eta)) &= \sup_{(z_0,\eta_0)\in\mathbf{V}_{(z,\eta)}^\delta} \left| \left\langle \widehat{g_{y,\omega}}, \widehat{g_{z,\eta}} - \Gamma((z,\eta),(z_0,\eta_0)) \cdot \widehat{g_{z_0,\eta_0}} \right\rangle \right| \\
&= \sup_{(z_0,\eta_0)\in\mathbf{V}_{(z,\eta)}^\delta} \left| K_{\theta,\tilde{\theta}_{(z,\eta),(z_0,\eta_0)},\Phi}((y,\omega),(z,\eta)) \right|. \qquad (6.20)
\end{aligned}$$

Based on (6.20), Lemma 4.7 provides

$$\begin{aligned}&\left|K_{\theta,\tilde{\theta}_{(z,\eta),(z_0,\eta_0)},\Phi}((y,\omega),(z,\eta))\right|\\ &\quad=\sqrt{\frac{w(\Phi(\eta))}{w(\Phi(\omega))}}\cdot\left|L_{\Phi(\eta)}[\theta,\tilde{\theta}_{(z,\eta),(z_0,\eta_0)}](A^T(\Phi(\eta))\langle z-y\rangle,\Phi(\omega)-\Phi(\eta))\right|,\end{aligned}$$

where $L_{\Phi(\eta)}$ is as in (4.9). If we define $\mathcal{L}_{\tau_0}:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}_0^+$, $\tau_0\in\mathbb{R}^d$, by

$$\mathcal{L}_{\tau_0}(x,\tau):=\sup_{z\in\mathbb{R}^d}\sup_{(z_0,\eta_0)\in\mathbf{V}^\delta_{(z,\Phi^{-1}(\tau_0))}}\left|L_{\tau_0}[\theta,\tilde{\theta}_{(z,\Phi^{-1}(\tau_0)),(z_0,\eta_0)}](x,\tau)\right|, \tag{6.21}$$

then, for all $(y,\omega),(z,\eta)\in\Lambda$,

$$\operatorname{osc}_{\mathcal{V}^\delta_\Phi,\Gamma}((y,\omega),(z,\eta))\leq\sqrt{\frac{w(\Phi(\eta))}{w(\Phi(\omega))}}\mathcal{L}_{\Phi(\eta)}(A^T(\Phi(\eta))\langle z-y\rangle,\Phi(\omega)-\Phi(\eta)).$$

Via a tedious, but straightforward derivation involving several changes of variable in a manner similar to the proof of Lemma 4.7, we obtain in particular that

$$\|\operatorname{osc}_{\mathcal{V}^\delta_\Phi,\Gamma}\|_{\mathcal{B}_m}\leq\operatorname*{ess\,sup}_{\tau_0\in\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}M(x,\tau)\mathcal{L}_{\tau_0}(x,\tau)\,dx\,d\tau, \tag{6.22}$$

where M is defined as in (4.2) and we used that m is Φ-compatible with the (symmetric) dominating weight $m^\Phi(x,\tau)=(1+|x|)^p\cdot v_1(\tau)$.

By Lemma 6.5, with $\tilde{w}=w_2$, all functions $\tilde{\theta}_{(z,\eta),(z_0,\eta_0)}$ with $(z,\eta)\in\Lambda$ and $(z_0,\eta_0)\in\mathbf{V}^\delta_{z,\eta}\subset(z+\mathbf{P}^\delta_\eta)\times\mathbf{Q}^\delta_\eta$ satisfy the conditions of Theorem 4.8, as does θ. Hence, for any $z,\tau_0\in\mathbb{R}^d$ and $(z_0,\eta_0)\in\mathbf{V}^\delta_{z,\Phi^{-1}(\tau_0)}$, Theorem 4.8 yields

$$|L_{\tau_0}[\theta,\tilde{\theta}_{(z,\Phi^{-1}(\tau_0)),(z_0,\eta_0)}](x,\tau)|\leq C\cdot C_{\max}\cdot(1+|x|)^{-(d+p+1)}\cdot v_0^{4d+3p+3}(\tau)\cdot[w_2(\tau)]^{-1}, \tag{6.23}$$

where $C>0$ depends only on d, k and the control weight v_0, and furthermore

$$\begin{aligned}C_{\max}&=C_{\max}\left(d+p+1,\theta,\tilde{\theta}_{(z,\Phi^{-1}(\tau_0)),(z_0,\eta_0)}\right)\\ &=\max_{\substack{j\in\underline{d}\\0\leq m\leq d+p+1}}\left\|\frac{\partial^m}{\partial\iota_j^m}\theta\right\|_{\mathbf{L}^2_{w_2}(\mathbb{R}^d)}\cdot\max_{\substack{j\in\underline{d}\\0\leq m\leq d+p+1}}\left\|\frac{\partial^m}{\partial\iota_j^m}\tilde{\theta}_{(z,\Phi^{-1}(\tau_0)),(z_0,\eta_0)}\right\|_{\mathbf{L}^2_{w_2}(\mathbb{R}^d)}\\ \text{(Lem. 6.5)}&\leq\max_{\substack{j\in\underline{d}\\0\leq m\leq d+p+1}}\left\|\frac{\partial^m}{\partial\iota_j^m}\theta\right\|_{\mathbf{L}^2_{w_2}(\mathbb{R}^d)}\cdot\max_{\substack{j\in\underline{d}\\0\leq m\leq d+p+1}}F_{j,m}(\delta;\theta,2,w_2)=:D^\delta_{\max}<\infty.\end{aligned}$$

Note that the estimate $D^{\delta}_{\max}$ is independent of $\tau_0 \in \mathbb{R}^D$, $z \in \mathbb{R}^d$, and $(z_0, \eta_0) \in \mathbf{V}^{\delta}_{z,\Phi^{-1}(\tau_0)}$, such that taking $D^{\delta}_{\max}$ instead of $C_{\max}$ in (6.23) produces a valid upper estimate for $\mathcal{L}_{\tau_0}(x, \tau)$. Moreover, note that Lemma 6.5 implies $D^{\delta}_{\max} \to 0$ as $\delta \to 0$.

Proving $\|\mathrm{osc}_{\mathcal{V}^{\delta}_{\Phi},\Gamma}\|_{\mathcal{B}_m} < \infty$ is now analogous to the proof of Theorem 4.4, and $\|\mathrm{osc}_{\mathcal{V}^{\delta}_{\Phi},\Gamma}\|_{\mathcal{B}_m} \to 0$ as $\delta \to 0$ follows directly from $D^{\delta}_{\max} \to 0$. □

7 Coorbit Space Theory of Warped Time-Frequency Systems

We have now developed explicit sufficient conditions that ensure $K_{\theta,\Phi}, \mathrm{osc}_{\mathcal{V}^{\delta}_{\Phi},\Gamma} \in \mathcal{B}_m$ and hence, by Eq. (2.23), $\mathrm{M}_{\mathcal{V}^{\delta}_{\Phi}} K_{\theta,\Phi} \in \mathcal{B}_m$, since $\mathcal{B}_m$ is solid. These are the crucial ingredients for applying coorbit theory in the setting of warped time-frequency representations.

Theorem 7.1 *Let Φ be a $(d + p + 1)$-admissible warping function with control weight v_0, where $p = 0$ if $R_{\Phi} = \sup_{\xi \in D} \|\mathrm{D}\Phi(\xi)\| = \infty$ and $p \in \mathbb{N}_0$ otherwise. Let furthermore $m_0 : \Lambda \times \Lambda \to \mathbb{R}^+$ be a symmetric weight that satisfies $1 \le m_0(\lambda, \rho) \le C^{(0)} m_0(\lambda, \nu) m_0(\nu, \rho)$ for all $\lambda, \rho, \nu \in \Lambda$ and*

$$m_0((y, \xi), (z, \eta)) \le (1 + |y - z|)^p \cdot v_1(\Phi(\xi) - \Phi(\eta)), \qquad \text{for all } y, z \in \mathbb{R}^d \text{ and } \xi, \eta \in D,\ \tau, \iota \in \mathbb{R}^d, \tag{7.1}$$

for some continuous and submultiplicative weight $v_1 : \mathbb{R}^d \to \mathbb{R}^+$ with $v_1(\iota) = v_1(-\iota)$ for all $\iota \in \mathbb{R}^d$.

Then there exist nonzero $\theta \in \boldsymbol{L}^2_{v_0^{d/2}}(\mathbb{R}^d)$, such that for any rich, solid Banach space $Y \hookrightarrow \boldsymbol{L}^1_{\mathrm{loc}}(\Lambda)$ with $\mathcal{B}_{m_0}(Y) \hookrightarrow Y$,

1. *$\mathrm{Co}(\mathcal{G}(\theta, \Phi), Y)$ is a well-defined Banach function space.*
2. *There is a $\delta_0 = \delta_0(\theta, \Phi, m) > 0$ independent of Y, such that*

$$(g_{y_{\ell,k},\omega_{\ell,k}})_{\ell,k\in\mathbb{Z}^d} \subset \mathcal{G}(\theta, \Phi)$$

is a Banach frame decomposition for $\mathrm{Co}(\mathcal{G}(\theta, \Phi), Y)$, whenever the points $\big((y_{\ell,k}, \omega_{\ell,k})\big)_{\ell,k\in\mathbb{Z}^d} \subset \Lambda$ satisfy $(y_{\ell,k}, \omega_{\ell,k}) \in V^{\delta}_{\ell,k}$, where $\mathcal{V}^{\delta}_{\Phi} = (V^{\delta}_{\ell,k})_{\ell,k\in\mathbb{Z}^d}$ is the Φ-induced δ-fine covering and $\delta \le \delta_0$.

In particular, items (1) and (2) above hold for $Y = \boldsymbol{L}^{p,q}_{\kappa}(\Lambda)$, with $1 \le p, q \le \infty$ and any weight $\kappa : \Lambda \to [1, \infty)$ that satisfies $m_{\kappa} \lesssim m_0$.

Proof By Propositions 5.2 and 5.4, the Φ-induced δ-fine covering $\mathcal{V}^{\delta}_{\Phi}$ is a topologically admissible, product-admissible covering that satisfies items (1)–(3) of Assumption 2.11 and item (1) of Assumption 2.19. Moreover, item (6) of Assumption 2.11 is satisfied, by the assumptions of this theorem.

Next, choose $\theta \in \mathbf{L}^2_{\sqrt{w_0}}(\mathbb{R}^d)$, such that $\|\theta\|_{\mathbf{L}^2(\mathbb{R}^d)} = 1$ and the assumptions of Theorem 6.1 are satisfied with $m = m_v$ defined by

$$m_v((y,\omega),(z,\eta)) = \max\left\{\frac{v((y,\omega))}{v((z,\eta))}, \frac{v((z,\eta))}{v((y,\omega))}\right\},$$
$$\text{with} \quad v((y,\omega)) := m_0((y,\omega),(x,\xi)) \cdot \max\{w(\Phi(\omega)), [w(\Phi(\omega))]^{-1}\},$$

for all $(y,\omega),(z,\eta) \in \Lambda$ and some fixed, arbitrary $(x,\xi) \in \Lambda$. This is always possible, since any function $\theta \in \mathcal{C}_c^\infty(\mathbb{R}^d) \subset \mathbf{L}^2(\mathbb{R}^d)$ with unit $\mathbf{L}^2$-norm satisfies these assumptions. In particular, the assumptions of Theorem 6.1 are also satisfied for $m = m_0 \leq m_v$. By Proposition 3.4, the map $(y,\omega) \mapsto g_{y,\omega}$ is continuous and by Corollary 3.6, the warped time-frequency system $\mathcal{G}(\theta,\Phi)$ is a tight Parseval frame, such that item (4) of Assumption 2.11 is satisfied. In particular, by Eq. (3.7), $\sup_{(y,\omega)\in\Lambda}\|g_{y,\omega}\|_2 \leq \|\theta\|_{\mathbf{L}^2_{\sqrt{w_0}}} < \infty$. Hence, with $w^c_{\mathcal{V}^\delta_\Phi} = \max\{w(\Phi(\omega)), [w(\Phi(\omega))]^{-1}\}$ as in Proposition 5.2 and $u(\lambda) := m_0(\lambda,(x,\xi))$ with the same choice of $(x,\xi) \in \Lambda$ as above, item (5) of Assumption 2.11 is satisfied as well.

Moreover, by choice of θ, and with Γ as in Theorem 6.1, we have

$$\|K_{\theta,\Phi}\|_{\mathcal{B}_{m_v}} < \infty \quad \text{and} \quad \|\mathrm{M}_{\mathcal{V}^\delta_\Phi} K_{\theta,\Phi}\|_{\mathcal{B}_{m_0}} \leq \|K_{\theta,\Phi}\|_{\mathcal{B}_{m_0}} + \|\mathrm{osc}_{\mathcal{V}^\delta_\Phi,\Gamma}\|_{\mathcal{B}_{m_0}} < \infty,$$

showing that the final item (7) of Assumption 2.11 is satisfied. Hence, Assumption 2.11 is fully satisfied and we can apply Theorem 2.14 to show that $\mathrm{Co}(\mathcal{G}(\theta,\Phi), Y)$ is a well-defined Banach function space.

Finally, note that Γ as in Theorem 6.1 is continuous, to verify that item (2) of Assumption 2.19 is satisfied. By the same theorem, we can choose $\delta_0 > 0$, such that

$$\|\mathrm{osc}_{\mathcal{V}^\delta_\Phi,\Gamma}\|_{\mathcal{B}_{m_v}} \cdot (2\|K_\Psi\|_{\mathcal{B}_{m_v}} + \|\mathrm{osc}_{\mathcal{V}^\delta_\Phi,\Gamma}\|_{\mathcal{B}_{m_v}}) < 1$$

for all $\delta \leq \delta_0$, proving the second assertion. The proof is completed by observing that the statement about weighted, mixed-norm Lebesgue spaces is a direct consequence of (2.14). □

By definition, the coorbit space $\mathrm{Co}(\mathcal{G}(\theta,\Phi), Y)$ depends on both the prototype function θ and the warping function Φ. The dependence on the warping function Φ is an essential consequence of (sufficiently) different warping functions inducing time-frequency representations with vastly different properties. Relations between coorbit spaces associated to different warping functions are studied in the framework of decomposition spaces [19, 42, 95] in a follow-up contribution. Here, we will show that the dependence on the generating prototype θ can be weakened, i.e., under certain conditions on θ_1, θ_2, the coorbit spaces $\mathrm{Co}(\mathcal{G}(\theta_1,\Phi), Y)$ and $\mathrm{Co}(\mathcal{G}(\theta_2,\Phi), Y)$ are equal, similar to modulation spaces for the STFT. Before we do so, however, we show that the mixed kernel associated with two warped time-frequency systems inherits the membership in $\mathcal{B}_m$ (or $\mathcal{A}_m$) from the kernels of the individual systems.

Lemma 7.2 *Let $X \in \{\mathcal{A}_m, \mathcal{B}_m\}$, with a symmetric weight m satisfying $m(\lambda, \rho) \le C^{(0)} m(\lambda, \nu) m(\nu, \rho)$, for some $C^{(0)}$ and all $\lambda, \rho, \nu \in \Lambda$. If $\theta_1, \theta_2 \in \boldsymbol{L}^2_{\sqrt{w_0}} \cap \boldsymbol{L}^2(\mathbb{R}^d)$ are nonzero and such that $K_{\theta_1,\Phi}, K_{\theta_2,\Phi} \in X$, then*

$$K_{\theta_1,\theta_2,\Phi} := K_{\mathcal{G}(\theta_1,\Phi),\mathcal{G}(\theta_2,\Phi)} \in X. \tag{7.2}$$

Proof. We first consider the case $\langle \theta_1, \theta_2 \rangle \neq 0$. In that case, the orthogonality relations, Theorem 3.5, applied to the kernel $K_{\theta_1,\Phi} \cdot K_{\theta_2,\Phi}$ yield, for all $(y, \omega), (z, \eta) \in \Lambda$,

$$\begin{aligned} K_{\theta_1,\Phi} \cdot K_{\theta_2,\Phi}((y,\omega),(z,\eta)) &= \int_\Lambda K_{\theta_1,\Phi}((y,\omega),(x,\xi)) K_{\theta_2,\Phi}((x,\xi),(z,\eta))\, d(x,\xi) \\ &= \int_\Lambda \overline{\langle g^{(1)}_{y,\omega}, g^{(1)}_{x,\xi}\rangle} \langle g^{(2)}_{z,\eta}, g^{(2)}_{x,\xi}\rangle\, d(x,\xi) \\ \text{(Def. of } V_{\bullet,\Phi}) &= \int_\Lambda V_{\theta_2,\Phi} g^{(2)}_{z,\eta}(x,\xi) \overline{V_{\theta_1,\Phi} g^{(1)}_{y,\omega}(x,\xi)}\, d(x,\xi) \\ &= \langle V_{\theta_2,\Phi} g^{(2)}_{z,\eta}, V_{\theta_1,\Phi} g^{(1)}_{y,\omega}\rangle \\ \text{(orth. rel.)} &= \langle g^{(2)}_{z,\eta}, g^{(1)}_{y,\omega}\rangle \langle \theta_1, \theta_2\rangle = \langle \theta_1, \theta_2\rangle \cdot K_{\theta_1,\theta_2,\Phi}((y,\omega),(z,\eta)). \end{aligned}$$

Since, under the conditions on m, $\mathcal{A}_m, \mathcal{B}_m$ are algebrae, this establishes (7.2).

If $\langle \theta_1, \theta_2 \rangle = 0$, then we need an auxiliary function θ_3, which may be any function in $\mathbf{L}^2_{\sqrt{w_0}} \cap \mathbf{L}^2(\mathbb{R}^d)$ such that $K_{\theta_3,\Phi} \in X$ and that is neither orthogonal to θ_1 nor to θ_2. For example, θ_3 could satisfy the conditions of Theorem. 4.4. By the first part of the proof, we obtain

$$(K_{\theta_1,\Phi} \cdot K_{\theta_3,\Phi}) \cdot (K_{\theta_3,\Phi} \cdot K_{\theta_2,\Phi}) = \langle \theta_1, \theta_3\rangle \overline{\langle \theta_2, \theta_3\rangle} \cdot\ K_{\theta_1,\theta_3,\Phi} \cdot K_{\theta_3,\theta_2,\Phi} \in X.$$

Now, apply the argument in the first part of the proof again to obtain that

$$K_{\theta_1,\theta_2,\Phi} = C^{-1} (K_{\theta_1,\Phi} \cdot K_{\theta_3,\Phi}) \cdot (K_{\theta_3,\Phi} \cdot K_{\theta_2,\Phi}) \quad \text{with} \quad C = \|\theta_3\|^2 \langle \theta_1, \theta_3\rangle \overline{\langle \theta_2, \theta_3\rangle}. \quad \square$$

Remark 7.3 If θ_1, θ_2 satisfy the conditions of Theorem 4.4, then the assumptions of Lemma 7.2 can be verified by applying that theorem. However, since Theorem 4.4 only provides *sufficient* conditions, there might be θ_1, θ_2 with $K_{\theta_1,\Phi}, K_{\theta_2,\Phi} \in \mathcal{B}_m$ that do not satisfy those conditions, for which Lemma 7.2 remains valid.

Theorem 7.4 *Assume that Φ, m_0 and both $\theta_1 \in \boldsymbol{L}^2_{\sqrt{w_0}}$ and $\theta_2 \in \boldsymbol{L}^2_{\sqrt{w_0}}$ jointly satisfy the conditions of Theorem 6.1.*

Then, for any rich, solid Banach space $Y \hookrightarrow \boldsymbol{L}^1_{\mathrm{loc}}(\Lambda)$ with $\mathcal{B}_{m_0}(Y) \hookrightarrow Y$, we have

$$\mathrm{Co}(\mathcal{G}(\theta_1, \Phi), Y) = \mathrm{Co}(\mathcal{G}(\theta_2, \Phi), Y).$$

In particular, the statement holds for $Y = \boldsymbol{L}^{p,q}_\kappa(\mu)$, with $1 \le p, q \le \infty$ and any weight $\kappa : \Lambda \to [1, \infty)$ that satisfies $m_\kappa \lesssim m_0$.

Proof The same derivations as in the proof of Theorem 7.1 show that Assumptions 2.11 and 2.19 are fully satisfied and consequently, by Theorem 2.14, $\mathrm{Co}(\mathcal{G}(\theta_1, \Phi), Y)$ and $\mathrm{Co}(\mathcal{G}(\theta_2, \Phi), Y)$ are well-defined Banach spaces. By Lemma 7.2, the mixed kernel K_{θ_1,θ_2} is contained in $\mathcal{B}_{m_v} \subset \mathcal{B}_{m_0}$, with v as in the proof of Theorem 7.1. Hence, we can apply Proposition 2.15 to obtain the desired result. The statement about weighted, mixed-norm Lebesgue spaces is, once more, a direct consequence of (2.14). □

8 Radial Warping

In this section, we consider warped time-frequency representations for which the warping of frequency space depends only on the modulus in the frequency domain, i.e., we study maps of the form

$$\Phi_\varrho : \mathbb{R}^d \to \mathbb{R}^d, \xi \mapsto \xi/|\xi| \cdot \varrho(|\xi|)\,,$$

which we call the **radial warping function** associated to the **radial component** $\varrho : [0, \infty) \to [0, \infty)$. More precisely, we will provide conditions on the radial component ϱ which ensure that Φ_ϱ is a (k-admissible) warping function, as introduced in Definitions 3.1 and 4.2. In particular, we will see that if ϱ is a strictly increasing C^{k+1} diffeomorphism which is also linear on a neighborhood of the origin, then Φ_ϱ is a C^{k+1} diffeomorphism, with inverse $\Phi_\varrho^{-1} = \Phi_{\varrho^{-1}}$. Finally, under additional "moderateness assumptions" on the derivatives of ϱ^{-1}, we will show that the diffeomorphism Φ_ϱ is a k-admissible warping function. These claims will be established in Sect. 8.1.

Section 8.2 is concerned with circumventing the somewhat unnatural restriction that ϱ is linear in a neighborhood of the origin. Using the so-called **slow-start construction**, one can associate to a "sufficiently well-behaved" function $\varsigma : [0, \infty) \to [0, \infty)$ a k-admissible radial component $\varrho : [0, \infty) \to [0, \infty)$, which equals ς outside an arbitrarily small neighborhood of the origin.

Finally, we discuss several examples of radial warping functions in Sect. 8.3.

8.1 General Properties of Radial Warping Functions

To enable a more compact notation, we will from now on denote by $\varrho_* := \varrho^{-1}$ the inverse of a bijection $\varrho : \mathbb{R} \to \mathbb{R}$.

Definition 8.1 Let $k \in \mathbb{N}_0$. A function $\varrho : \mathbb{R} \to \mathbb{R}$ is called a k-**admissible radial component with control weight** $v : \mathbb{R} \to \mathbb{R}^+$, if the following hold:

1. ϱ is a strictly increasing C^{k+1}-diffeomorphism with inverse $\varrho_* = \varrho^{-1}$.
2. ϱ is antisymmetric, that is, $\varrho(-\xi) = -\varrho(\xi)$ for all $\xi \in \mathbb{R}$. In particular, $\varrho(0) = 0$.
3. There are $\varepsilon > 0$ and $c > 0$ with $\varrho(\xi) = c \cdot \xi$ for all $\xi \in (-\varepsilon, \varepsilon)$.
4. The weight v is continuous, submultiplicative, and radially increasing. Additionally, ϱ_*' and

$$\widetilde{\varrho_*} : \mathbb{R} \to \mathbb{R}^+, \quad \text{defined by} \quad \widetilde{\varrho_*}(\xi) := \varrho_*(\xi)/\xi, \quad \text{for} \quad \xi \neq 0, \quad \text{and} \quad \widetilde{\varrho_*}(0) := c^{-1} \tag{8.1}$$

are v-moderate.

5. There are constants $C_0, C_1 > 0$ with

$$C_0 \cdot \widetilde{\varrho_*}(\xi) \leq \varrho_*'(\xi) \leq C_1 \cdot (1+\xi) \cdot \widetilde{\varrho_*}(\xi) \qquad \forall \xi \in \mathbb{R}^+. \tag{8.2}$$

6. We have

$$|\varrho_*^{(\ell)}(\xi)| \leq v(\xi - \eta) \cdot \varrho_*'(\eta) \qquad \forall \eta, \xi \in [0, \infty) \text{ and } \ell \in \underline{k+1}. \tag{8.3}$$

Note that the property (8.3) can equivalently be exchanged by the simpler $|\varrho_*^{(\ell)}| \leq C\varrho_*'$, for all $\ell \in \underline{k+1}$ (using that ϱ_*' is v-moderate and v is submultiplicative), at the cost of introducing a multiplicative constant $Cv(0)$ on the right-hand side of (8.3).

Remark 8.2 The reader may wonder why Definition 8.1 prescribes properties of ϱ on the negative half-axis at all. These requirements are not strictly necessary, but neither are they an actual restriction: The existence of an odd extension of regularity $C^{k+1}(\mathbb{R})$ of a function $\varrho_0 \colon [0, \infty) \to [0, \infty)$ is, in fact, necessary for the radial warping function Φ_{ϱ_0} induced by ϱ_0 to be in $C^{k+1}(\mathbb{R}^d)$. This is easily seen by considering the case $d = 1$.

On the other hand, the third condition in Definition 8.1 could indeed be slightly weakened, as long as $\widetilde{\varrho_*}(\xi) = \varrho_*(\xi)/\xi$ has a positive, finite limit for $\xi \to 0$ (and sufficiently many of its derivatives have a finite limit at 0), and none of the other conditions are violated. However, the behavior of ϱ in a small neighborhood of zero has comparably little effect on the induced warped time-frequency system. The slow-start construction discussed in Sect. 8.2 provides a method to modify functions satisfying a weaker variant of Definition 8.1 in a small neighborhood of zero, resulting in a k-admissible radial component. Concerning the (lack of) impact of the slow-start construction on the resulting coorbit spaces, cf. Remark 8.10.

Remark 8.3 (1) An important consequence of these assumptions is that there exists a constant $C_2 = C_2(\varrho, v) > 0$ with

$$|\varrho_*^{(\ell)}(\xi)| \leq C_2 \cdot (1+\xi) \cdot \widetilde{\varrho_*}(\xi) \qquad \forall \xi \in \mathbb{R}^+ \text{ and } \ell \in \{0\} \cup \underline{k+1}. \tag{8.4}$$

Indeed, for $\ell = 0$ (8.4) is always satisfied as long as $C_2 \geq 1$, since ϱ_* is increasing with $\varrho_*(0) = 0$, whence $|\varrho_*^{(0)}(\xi)| = \varrho_*(\xi) = \xi \cdot \widetilde{\varrho_*}(\xi) \leq (1+\xi) \cdot \widetilde{\varrho_*}(\xi)$ for $\xi \in \mathbb{R}^+$. Thus, it remains to verify Eq. (8.4) for $\ell \in \underline{k+1}$. But for this case, applying (8.3) with $\eta = \xi$, we see that

$$|\varrho_*^{(\ell)}(\xi)| \leq v(\xi - \xi) \cdot \varrho_*'(\xi) = v(0) \cdot \varrho_*'(\xi),$$

so that (8.2) yields $|\varrho_*^{(\ell)}(\xi)| \le v(0) \cdot \varrho_*'(\xi) \le C_1 \cdot v(0) \cdot (1+\xi) \cdot \widetilde{\varrho_*}(\xi)$. Setting $C_2 := \max\{1, C_1 \cdot v(0)\}$ and C_1 only depends on the radial component ϱ, we have thus established (8.4).

(2) To indicate that being an admissible radial component is a nontrivial restriction on ϱ, we observe that condition (8.2) entails certain growth restrictions on the function $\varrho_* = \varrho^{-1}$. Indeed, for arbitrary $\varepsilon > 0$ and $\xi \ge 1/\varepsilon$, Eq. (8.2) shows $\varrho_*'(\xi) \le C_1 \cdot (1+\xi) \cdot \varrho_*(\xi)/\xi \le (1+\varepsilon) C_1 \cdot \varrho_*(\xi)$. This implies

$$\begin{aligned}\frac{d}{d\xi}\left(e^{-(1+\varepsilon)C_1\xi} \cdot \varrho_*(\xi)\right) &= -(1+\varepsilon)C_1 \cdot e^{-(1+\varepsilon)C_1\xi} \cdot \varrho_*(\xi) + e^{-(1+\varepsilon)C_1\xi} \cdot \varrho_*'(\xi) \\ &\le -(1+\varepsilon)C_1 \cdot e^{-(1+\varepsilon)C_1\xi} \cdot \varrho_*(\xi) + (1+\varepsilon)C_1 \\ &\quad \cdot e^{-(1+\varepsilon)C_1\xi} \cdot \varrho_*(\xi) = 0\end{aligned}$$

for all $\xi \ge 1/\varepsilon$. For any $\xi \ge a \ge 1/\varepsilon$, this implies $e^{-(1+\varepsilon)C_1\xi} \cdot \varrho_*(\xi) \le e^{-(1+\varepsilon)C_1 a} \cdot \varrho_*(a)$, and hence

$$\varrho_*(\xi) \le \frac{\varrho_*(a)}{e^{(1+\varepsilon)C_1 a}} \cdot e^{(1+\varepsilon)C_1\xi} \qquad \forall\, \xi \ge a \ge \varepsilon^{-1}, \text{ for any } \varepsilon > 0. \tag{8.5}$$

Likewise, the lower bound in (8.2) implies

$$\begin{aligned}\frac{d}{d\xi}\left(\xi^{-C_0} \cdot \varrho_*(\xi)\right) &= (-C_0)\xi^{-C_0-1} \cdot \varrho_*(\xi) + \xi^{-C_0} \cdot \varrho_*'(\xi) \\ &\ge (-C_0)\xi^{-C_0} \cdot \frac{\varrho_*(\xi)}{\xi} + C_0 \cdot \xi^{-C_0} \cdot \frac{\varrho_*(\xi)}{\xi} = 0\end{aligned}$$

for all $\xi \in \mathbb{R}^+$. Thus, for $\xi \ge a > 0$, we get $\xi^{-C_0} \cdot \varrho_*(\xi) \ge a^{-C_0} \cdot \varrho_*(a)$, and thus

$$\varrho_*(\xi) \ge \frac{\varrho_*(a)}{a^{C_0}} \cdot \xi^{C_0} \qquad \forall\, \xi \ge a > 0. \tag{8.6}$$

In words, Eqs. (8.5) and (8.6) show that *the inverse of an admissible radial component ϱ can grow at most exponentially, and has to grow at least like a positive (not necessarily integer) power of ξ.*

We define (for a larger class of radial components) the radial warping function associated with ϱ.

Definition 8.4 For a diffeomorphism $\varrho : \mathbb{R} \to \mathbb{R}$ with $\varrho(\xi) = c\xi$ for all $\xi \in (-\varepsilon, \varepsilon)$ and suitable $\varepsilon, c > 0$, the **associated radial warping function** is given by

$$\begin{aligned}\Phi_\varrho : \mathbb{R}^d \to \mathbb{R}^d, \xi \mapsto \widetilde{\varrho}(|\xi|) \cdot \xi, \quad &\text{with} \quad \widetilde{\varrho}(t) := \varrho(t)/t \quad \text{for } t \in \mathbb{R} \setminus \{0\}, \\ &\text{and} \quad \widetilde{\varrho}(0) := c.\end{aligned} \tag{8.7}$$

Clearly, if $\varrho \in C^k(\mathbb{R})$, then $\widetilde{\varrho} \in C^k(\mathbb{R})$. Our goal in this section is to show that Φ_ϱ is a k-admissible warping function as per Definition 4.2, provided that ϱ is a k-admissible radial component. To this end, we first show that the inverse Φ_ϱ^{-1} of Φ_ϱ is given by $\Phi_\varrho^{-1} = \Phi_{\varrho^{-1}}$, and provide a convenient expression of the Jacobian $\mathrm{D}\Phi_\varrho^{-1}$. The following notation will be helpful for that purpose: For $\xi \in \mathbb{R}^d \setminus \{0\}$, we define

$$\xi_\circ := \xi/|\xi|, \qquad \pi_\xi : \mathbb{R}^d \to \mathbb{R}^d, \tau \mapsto \langle \tau, \xi_\circ \rangle \cdot \xi_\circ\,, \qquad \text{and} \qquad \pi_\xi^\perp := \mathrm{id}_{\mathbb{R}^d} - \pi_\xi\,, \tag{8.8}$$

so that π_ξ is the orthogonal projection on the space spanned by ξ, while $\pi_\xi^\perp$ is the orthogonal projection on the orthogonal complement of this space. With these notations, the derivative of Φ_ϱ and Φ_ϱ^{-1} can be described as follows:

Lemma 8.5 *Let $\varrho : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^k$-diffeomorphism with $\varrho(t) = ct$ for all $t \in (-\varepsilon, \varepsilon)$ and suitable $\varepsilon, c > 0$. Then Φ_ϱ is $\mathcal{C}^k$, and for $\xi \in \mathbb{R}^d \setminus \{0\}$, we have*

$$\begin{aligned} \mathrm{D}\Phi_\varrho(\xi) &= \widetilde{\varrho}(|\xi|) \cdot \pi_\xi^\perp + \varrho'(|\xi|) \cdot \pi_\xi\,, \quad \textit{and} \\ [\mathrm{D}\Phi_\varrho(\xi)]^{-1} &= [\widetilde{\varrho}(|\xi|)]^{-1} \cdot \pi_\xi^\perp + [\varrho'(|\xi|)]^{-1} \cdot \pi_\xi\,. \end{aligned} \tag{8.9}$$

Furthermore, Φ_ϱ is a $\mathcal{C}^k$-diffeomorphism, with inverse
$\Phi_\varrho^{-1} = \Phi_{\varrho_*}$ and satisfies $\varrho_*(t) = t/c$ for $t \in (-c\varepsilon, c\varepsilon)$.
Finally, if ϱ is a 0-admissible radial component, then we have

$$\|[\mathrm{D}\Phi_{\varrho_*}(\xi)]^{-1}\| \lesssim 1/\widetilde{\varrho_*}(|\xi|) \qquad \forall\, \xi \in \mathbb{R}^d, \quad \textit{with } \widetilde{\varrho_*} \textit{ as in (8.1)}, \tag{8.10}$$

where the implied constant only depends on the constant C_0 in (8.2).

Proof Recall that $\widetilde{\varrho} \in \mathcal{C}^k(\mathbb{R})$, with $\widetilde{\varrho} \equiv c$ on $(-\varepsilon, \varepsilon)$, and hence $\Phi_\varrho \in \mathcal{C}^k(\mathbb{R}^d)$.

Now, a direct computation using the identity $\partial_j |\xi| = \xi_j / |\xi|$ shows for $\xi \in \mathbb{R}^d \setminus \{0\}$ that

$$\partial_j (\Phi_\varrho)_i(\xi) = \partial_j \left(\xi_i \cdot \frac{\varrho(|\xi|)}{|\xi|} \right) = \widetilde{\varrho}(|\xi|) \cdot \delta_{i,j} + \frac{\varrho'(|\xi|) - \widetilde{\varrho}(|\xi|)}{|\xi|^2} \cdot \xi_i \xi_j\,.$$

In vector notation, and with $\xi_\circ = \xi/|\xi|$ as in (8.8), this means

$$\mathrm{D}\Phi_\varrho(\xi) = \widetilde{\varrho}(|\xi|) \cdot \mathrm{id} + \big(\varrho'(|\xi|) - \widetilde{\varrho}(|\xi|)\big) \cdot \xi_\circ \xi_\circ^T\,.$$

Now, recall that $\xi_\circ \xi_\circ^T$ is the matrix representing the linear map π_ξ, and that $\mathrm{id} = \pi_\xi + \pi_\xi^\perp$. Inserting these identities into the previous displayed equation establishes the claimed formula for $\mathrm{D}\Phi_\varrho(\xi)$. In particular, each $\eta \in \mathbb{R}^d$ with $\eta \perp \xi$ is mapped to $\widetilde{\varrho}(|\xi|) \cdot \eta$ by $\mathrm{D}\Phi_\varrho(\xi)$, while each $\eta \in \mathrm{span}(\xi)$ is mapped to $\varrho'(|\xi|) \cdot \eta$. Since $\mathbb{R}^d = \xi^\perp \oplus \mathrm{span}(\xi)$, the stated formula for $[\mathrm{D}\Phi_\varrho(\xi)]^{-1}$ follows.

Linearity of $\varrho_*(t) = t/c$ for $t \in (-c\varepsilon, c\varepsilon)$ is clear, such that Φ_{ϱ_*} is a radial warping function as per Definition 8.4. Note $|\Phi_\varrho(\xi)| = \varrho(|\xi|)$ for $\xi \in \mathbb{R}^d \setminus \{0\}$, such that $\varrho_*(|\Phi_\varrho(\xi)|) = \varrho_*(\varrho(|\xi|)) = |\xi|$ and $\Phi_\varrho(\xi)/|\Phi_\varrho(\xi)| = \xi/|\xi|$. Together, this implies $\Phi_{\varrho_*}(\Phi_\varrho(\xi)) = \xi$, for all $\xi \in \mathbb{R}^d \setminus \{0\}$ and thus, by continuity, for $\xi = 0$ as well. Repeating this argument after interchanging ϱ_* and ϱ yields $\Phi_\varrho \circ \Phi_{\varrho_*} = \mathrm{id}$.

To prove (8.10), consider $\xi \in \mathbb{R}^d \setminus \{0\}$ and observe that $\|[\mathrm{D}\Phi_{\varrho_*}(\xi)]^{-1}\| = \max\left\{[\widetilde{\varrho_*}(|\xi|)]^{-1}, [\varrho_*'(|\xi|)]^{-1}\right\}$, by (8.9). Applying the lower inequality in (8.2), we get

$$\|[\mathrm{D}\Phi_{\varrho_*}(\xi)]^{-1}\| \leq \max\{1, C_0^{-1}\} \cdot 1/\widetilde{\varrho_*}(|\xi|).$$

For $\xi = 0$ the result follows by continuity. □

To verify Property (4.5) of Definition 4.2, i.e., $\|\partial^\alpha \phi_\tau(\iota)\| \leq v_0(\iota)$, for all $\tau, \iota \in \mathbb{R}^d$ and all $\alpha \in \mathbb{N}_0^d$, $|\alpha| \leq k$, we need to control certain derivatives of the (matrix-valued) function

$$\phi_\tau(\iota) = \Big(A^{-1}(\tau) \cdot A(\iota + \tau)\Big)^T \quad \text{with} \quad A(\tau) = \mathrm{D}\Phi_\varrho^{-1}(\tau) \tag{8.11}$$

from (4.4). To this end, we will frequently use *Faa di Bruno's formula*, a chain rule for higher derivatives. Precisely, we will use the following form of the formula, which is a slightly simplified (but less precise) version of [28, Corollary 2.10]. Note that, for a nonnegative multiindex α, i.e., $\alpha \in \mathbb{N}_0^d$, we denote the sum of its components by $|\alpha| \geq 0$ and by $\alpha = 0$ we refer to the unique multiindex with $|\alpha| = 0$.

Lemma 8.6 *For $\alpha \in \mathbb{N}_0^d \setminus \{0\}$ and $n \in \underline{|\alpha|}$, set*

$$\Gamma_{\alpha,n} := \left\{ \gamma = (\gamma_1, \ldots, \gamma_n) \in \left[\mathbb{N}_0^d \setminus \{0\}\right]^n : \sum_{j=1}^n \gamma_j = \alpha \right\}.$$

Furthermore, set $\Gamma := \bigcup_{\alpha \in \mathbb{N}_0^d \setminus \{0\}} \bigcup_{n=1}^{|\alpha|} \Gamma_{\alpha,n}$.

Then, for each $\gamma \in \Gamma$, there is a constant $D_\gamma \in \mathbb{R}$ such that for any open sets $U \subset \mathbb{R}^d$ and $V \subset \mathbb{R}$, and any C^k functions $f : V \to \mathbb{R}$ and $g : U \to V$, the following holds for any $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \in \underline{k}$:

$$\partial^\alpha (f \circ g)(x) = \sum_{n=1}^{|\alpha|} \left[f^{(n)}(g(x)) \cdot \sum_{\gamma \in \Gamma_{\alpha,n}} \left(D_\gamma \cdot \prod_{j=1}^n (\partial^{\gamma_j} g)(x) \right) \right] \quad \forall\, x \in U\,,$$

where $f^{(n)}$ denotes the n-th derivative of f.

Remark From the statement of [28, Corollary 2.10], it might appear that the constants D_γ also depend on α, n, d, in addition to γ. But these parameters are determined by γ: On the one hand, we have $\gamma \in [\mathbb{N}_0^d]^n$, which uniquely determines n and d. On the other hand, $\alpha = \sum_{j=1}^n \gamma_j$ for $\gamma \in \Gamma_{\alpha,n}$.

With these preparations, we can now prove that the radial warping function Φ_ϱ associated to a k-admissible radial component ϱ is indeed a k-admissible warping function. Most significantly, the following proposition proves that Property (4.5), cf. Definition 4.2 or the discussion preceding the above lemma, is satisfied.

Proposition 8.7 *Let $\varrho : \mathbb{R} \to \mathbb{R}$ be a k-admissible radial component with control weight $v : \mathbb{R} \to \mathbb{R}^+$.*

Then there is a constant $C \geq 1$, dependent on ϱ, v, d, and k, such that with

$$v_0 : \mathbb{R}^d \to \mathbb{R}^+, \tau \mapsto C \cdot (1 + |\tau|) \cdot v(|\tau|),$$

the function Φ_ϱ satisfies (4.5) for all $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \leq k$.

Proof It is easy to see that v_0 is submultiplicative and radially increasing as the product of submultiplicative and radially increasing weights C, $(1 + |\bullet|)$ and $v(|\bullet|)$.

The proof is divided into five steps. As a preparation for these, recall from Lemma 8.5 that $\Phi_\varrho^{-1} = \Phi_{\varrho_*} = (\bullet) \cdot \widetilde{\varrho_*}(|\bullet|)$, with $\varrho_* = \varrho^{-1}$ and $\widetilde{\varrho_*}$ as defined in (8.1). By Lemma 8.5, $\widetilde{\varrho_*} \in C^{k+1}(\mathbb{R})$. Our main goal is to estimate the derivatives of Φ_{ϱ_*}.

Step 1 - Estimate the derivatives of $\widetilde{\varrho_*}$: A trivial induction shows $\frac{d^\ell}{dt^\ell} t^{-1} = (-1)^\ell \cdot \ell! \cdot t^{-(1+\ell)}$. With this, Leibniz's rule shows for any $n \in \underline{k+1}$ and any $t \in [c\varepsilon, \infty)$ that

$$\begin{aligned}
\left| \widetilde{\varrho_*}^{(n)}(t) - \frac{\varrho_*^{(n)}(t)}{t} \right| &= \left| \sum_{\ell=1}^{n} \binom{n}{\ell} \cdot \frac{d^\ell t^{-1}}{dt^\ell} \cdot \varrho_*^{(n-\ell)}(t) \right| \\
&\leq C(k) \cdot \sum_{\ell=1}^{n} t^{-(1+\ell)} \cdot |\varrho_*^{(n-\ell)}(t)| \\
&\overset{(8.4)}{\leq} C(k) C_2 \cdot \sum_{\ell=1}^{n} t^{-\ell+1} \cdot \left(\frac{1+t}{t} \right)^2 \cdot \frac{\widetilde{\varrho_*}(t)}{1+t} \\
{\scriptstyle (t^{-1} \leq (c\varepsilon)^{-1} \text{ and } (1+t)/t = t^{-1}+1 \leq (c\varepsilon)^{-1}+1)} &\leq C^{(1)} \cdot \widetilde{\varrho_*}(t)/(1+t) ,
\end{aligned} \tag{8.12}$$

where the constant $C(k) > 0$ in the first inequality only depends on k, C_2 is as in (8.4), and $C^{(1)}$ is given by $C^{(1)} = C(k) \cdot C_2((c\varepsilon)^{-1} + 1)^2 \cdot (k+1) \cdot \max\{1, (c\varepsilon)^{-k}\}$.

In particular, $|\widetilde{\varrho_*}^{(n)}(t)| \leq C^{(1)} \cdot \widetilde{\varrho_*}(t)/(1+t) + |\varrho_*^{(n)}(t)/t|$, such that

$$|\widetilde{\varrho_*}^{(n)}(t)| \leq C^{(1)} \cdot \frac{\widetilde{\varrho_*}(t)}{1+t} + \left| \frac{\varrho_*^{(n)}(t)}{t} \right| \overset{(8.2),(8.3)}{\leq} (C^{(1)} C_0^{-1} + v(0)) \cdot \frac{\varrho_*'(t)}{t}. \tag{8.13}$$

Furthermore, the same estimate yields, with $(1+t)/t \leq (c\varepsilon)^{-1} + 1$ and (8.4),

$$|\widetilde{\varrho_*}^{(n)}(t)| \leq C^{(1)} \cdot \widetilde{\varrho_*}(t) + C_2 \cdot ((c\varepsilon)^{-1} + 1) \widetilde{\varrho_*}(t) \leq 2C^{(1)} \cdot \widetilde{\varrho_*}(t), \tag{8.14}$$

where both (8.13) and (8.14) hold for all $t \in [c\varepsilon, \infty)$ and $n \in \underline{k+1}$.

Step 2 - Estimate the partial derivatives of $\tau \mapsto |\tau|$ for $\tau \neq 0$: It is well known that the derivative of order $n \in \mathbb{N}$ of the square root function $t \mapsto t^{1/2}$ has the form $t \mapsto c_n \cdot t^{-n+1/2}$, for all $t \in \mathbb{R}^+$ and some constant $c_n \neq 0$. Further, noting that $\partial^\alpha |\tau|^2 = 0$ unless $\alpha = i e_j$ for $i \in \{0,1,2\}$ and $j \in \underline{d}$, it is easy to see that $\left| \partial^\alpha |\tau|^2 \right| \le 2 \cdot |\tau|^{2-|\alpha|}$ for all $\tau \in \mathbb{R}^d \setminus \{0\}$ and $\alpha \in \mathbb{N}_0^d$.

Since $\tau \mapsto |\tau|$ equals the composition $|\bullet| = (\bullet)^{1/2} \circ |\bullet|^2$, Faa di Bruno's formula (see Lemma 8.6) yields

$$\left|\partial^\alpha |\tau|\right| = \left| \sum_{n=1}^{|\alpha|} c_n |\tau|^{1-2n} \cdot \sum_{\gamma \in \Gamma_{\alpha,n}} \left(D_\gamma \cdot \prod_{j=1}^{n} (\partial^{\gamma_j} |\bullet|^2)(\tau) \right) \right|,$$

for $\alpha \in \mathbb{N}_0^d \setminus \{0\}$ and $\tau \in \mathbb{R}^d \setminus \{0\}$. But we have $\sum_{j=1}^n \gamma_j = \alpha$ for $\gamma \in \Gamma_{\alpha,n}$, and hence, using $n \le |\alpha|$, we have $\left| \prod_{j=1}^n (\partial^{\gamma_j} |\bullet|^2)(\tau) \right| \le 2^{|\alpha|} \cdot |\tau|^{2n-|\alpha|}$. Overall, we obtain

$$\left|\partial^\alpha |\tau|\right| \le C_\alpha \cdot |\tau|^{1-|\alpha|} \qquad \forall\, \tau \in \mathbb{R}^d \setminus \{0\} \text{ and } \alpha \in \mathbb{N}_0^d, \tag{8.15}$$

for some constants C_α that may also depend on d.

The estimate is trivial in case of $\alpha = 0$.

Step 3 - Estimate the partial derivatives of $\zeta : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}, \tau \mapsto \widetilde{\varrho_*}(|\tau|)$: Note that this map is just the composition of $\widetilde{\varrho_*}$ with the map $\tau \mapsto |\tau|$ analyzed in the preceding step. Thus, Faa di Bruno's formula (see Lemma 8.6) shows for any $\alpha \in \mathbb{N}_0^d \setminus \{0\}$ with $|\alpha| \le k+1$ and $\tau \in \mathbb{R}^d \setminus \{0\}$ that

$$\partial^\alpha \zeta(\tau) = \sum_{n=1}^{|\alpha|} \left[\widetilde{\varrho_*}^{(n)}(|\tau|) \cdot \sum_{\gamma \in \Gamma_{\alpha,n}} \left(D_\gamma \cdot \prod_{j=1}^{n} \partial^{\gamma_j} |\tau| \right) \right]. \tag{8.16}$$

In the previous step, we saw $|\partial^{\gamma_j} |\tau|| \le C_{\gamma_j} \cdot |\tau|^{1-|\gamma_j|}$. Since $\sum_{j=1}^n \gamma_j = \alpha$ for $\gamma = (\gamma_1, \ldots, \gamma_n) \in \Gamma_{\alpha,n}$, there thus exists a constant $C_\gamma > 0$ that may additionally depend on d, such that

$$\left| \prod_{j=1}^{n} \partial^{\gamma_j} |\tau| \right| \le C_\gamma \cdot |\tau|^{n-|\alpha|} \qquad \forall\, \tau \in \mathbb{R}^d \setminus \{0\}, \quad \alpha \in \mathbb{N}_0^d \setminus \{0\}, \quad n \in \underline{|\alpha|}, \quad \text{and} \quad \gamma \in \Gamma_{\alpha,n}. \tag{8.17}$$

Now, let us focus on the case $|\tau| \ge c\varepsilon$. Then, if $n \in \underline{|\alpha| - 1}$, the estimate (8.14) yields with $|\tau|^{n-|\alpha|} \le (c\varepsilon)^{n+1-|\alpha|} \cdot |\tau|^{-1} \le (c\varepsilon)^{n+1-|\alpha|} (1 + (c\varepsilon)^{-1})/(1+|\tau|)$ that

$$\left| \widetilde{\varrho_*}^{(n)}(|\tau|) \cdot \sum_{\gamma \in \Gamma_{\alpha,n}} \left(D_\gamma \cdot \prod_{j=1}^{n} \partial^{\gamma_j} |\tau| \right) \right| \lesssim \frac{\widetilde{\varrho_*}(|\tau|)}{1+|\tau|} \quad \forall\, \tau \in \mathbb{R}^d \setminus \overline{B_{c\varepsilon}}(0) \text{ and } n \in \underline{|\alpha| - 1}. \tag{8.18}$$

Here, $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \in \underline{k+1}$.

Overall, by combining (8.16)– (8.18) (and noting that the case $n = |\alpha|$ is not covered by (8.18)), we get

$$|\partial^\alpha \zeta(\tau)| = \left|\partial^\alpha \left(\widetilde{\varrho}_*(|\tau|)\right)\right| \lesssim \left|\widetilde{\varrho}_*{}^{(|\alpha|)}(|\tau|)\right| + \frac{\widetilde{\varrho}_*(|\tau|)}{1+|\tau|} \tag{8.19}$$

$$\overset{(8.14)}{\lesssim} \widetilde{\varrho}_*(|\tau|) \quad \forall \alpha \in \mathbb{N}_0^d$$

$$\text{with } |\alpha| \in \underline{k+1} \text{ and } \tau \in \mathbb{R}^d \setminus B_{c\varepsilon}(0). \tag{8.20}$$

We note that the total implied constant between the left and right hand sides of Eq. (8.20) depends on ϱ, v, d, and k. Note that the quantities c, ε, C_0, C_1 and C_2 that are more explicitly present in the dependencies are themselves directly derived from ϱ and v. Finally, the case $\alpha = 0$ is trivial.

Step 4 - Estimate $\partial^\alpha \phi_\tau(\iota)$ for $0 \leq |\alpha| \leq k$ and $\iota \in \mathbb{R}^d \setminus B_{c\varepsilon}(-\tau)$: Recall from (8.11) the definition of $\phi_\tau(\iota) = \left(A^{-1}(\tau) \cdot A(\tau+\iota)\right)^T$. Since $\|M\| = \|M^T\|$ for all $M \in \mathbb{R}^{d\times d}$, and since $\partial^\alpha[M(\tau)]^T = [\partial^\alpha M(\tau)]^T$ for any sufficiently smooth matrix-valued function $M : \mathbb{R}^d \to \mathbb{R}^{d\times d}$, it is sufficient to estimate $\|\partial^\alpha \varphi_\tau(\iota)\|$ with $\varphi_\tau(\iota) := A^{-1}(\tau) \cdot A(\tau+\iota)$, where $A(\tau) = \mathrm{D}\Phi_{\varrho_*}(\tau)$.

Furthermore, $\|M\| \leq \sum_{j=1}^d |M_{\bullet,j}|$ for all $M \in \mathbb{R}^{d\times d}$, such that it is sufficient to estimate the columns of M individually. In the following, we denote, for $\alpha \in \mathbb{N}_0^d$ and $\iota \in \mathbb{R}^d$, $\partial_\iota^\alpha = \frac{\partial^{|\alpha|}}{\partial_{\iota_1}^{\alpha_1} \cdots \partial_{\iota_d}^{\alpha_d}}$. Let us fix $\tau \in \mathbb{R}^d \setminus \{0\}$. Then, $\partial^\alpha \varphi_\tau(\iota) = A^{-1}(\tau) \cdot (\partial^\alpha A)(\tau+\iota)$. We see that the j-th column of $\partial^\alpha \varphi_\tau(\iota)$ is simply

$$\left[\partial^\alpha \varphi_\tau(\iota)\right]_{\bullet,j} = A^{-1}(\tau) \cdot \partial_\iota^\alpha [A(\tau+\iota)]_{\bullet,j} = A^{-1}(\tau) \cdot (\partial_\iota^{\alpha+e_j} \Phi_{\varrho_*})(\tau+\iota). \tag{8.21}$$

Now fix $j \in \underline{d}$, and set $\sigma := \alpha + e_j$ for brevity. Note $\sigma \in \mathbb{N}_0^d \setminus \{0\}$ with $|\sigma| \in \underline{k+1}$.

By definition of Φ_{ϱ_*}, the i-th entry of $\Phi_{\varrho_*}(\tau)$ is $[\Phi_{\varrho_*}(\tau)]_i = \tau_i \cdot \widetilde{\varrho}_*(|\tau|)$. Let α_+, for $\alpha \in \mathbb{Z}_0^d$, be the elementwise positive part, i.e., $(\alpha_+)_i = \max\{0, \alpha_i\}$, $i \in \underline{d}$. The Leibniz rule, with $\partial^\beta \tau_i = 0$ for $\beta \notin \{0, e_i\}$ and $\partial_i \tau_i = 1$, yields

$$[\partial^\sigma \Phi_{\varrho_*}(\tau)]_i = \tau_i \cdot \left[\partial^\sigma\left(\widetilde{\varrho}_*(|\tau|)\right)\right] + \sigma_i \cdot \partial_\tau^{(\sigma-e_i)_+}\left(\widetilde{\varrho}_*(|\tau|)\right) \qquad \forall i \in \underline{d},$$

or in other words,

$$\partial^\sigma \Phi_{\varrho_*}(\tau) = \left(\partial_\tau^\sigma\left[\widetilde{\varrho}_*(|\tau|)\right]\right) \cdot \tau + v_{\sigma,\tau}, \quad \text{with} \quad v_{\sigma,\tau} := \left[\sigma_i \cdot \partial_\tau^{(\sigma-e_i)_+}\left(\widetilde{\varrho}_*(|\tau|)\right)\right]_{i=1,\ldots,d}. \tag{8.22}$$

Now, by (8.20), we have $|v_{\sigma,\tau}| \lesssim \widetilde{\varrho}_*(|\tau|)$, for all $\tau \in \mathbb{R}^d \setminus B_{c\varepsilon}(0)$. Furthermore, Lemma 8.5 provides the estimate $\|A^{-1}(\tau)\| = \|[\mathrm{D}\Phi_{\varrho_*}(\tau)]^{-1}\| \leq \max\{1, C_0^{-1}\}/\widetilde{\varrho}_*(|\tau|)$ with C_0 as in (8.2). Note that we inserted the explicit constant derived in the proof of Lemma 8.5 above. Since $\widetilde{\varrho}_*$ is v-moderate and v is radially increasing, this implies

$$
\begin{aligned}
|A^{-1}(\tau) \cdot v_{\sigma,\tau+\iota}| &\lesssim \max\{1, C_0^{-1}\} \cdot \frac{\widetilde{\varrho}_*(|\tau+\iota|)}{\widetilde{\varrho}_*(|\tau|)} \leq \max\{1, C_0^{-1}\} \cdot v(|\iota|) \\
&= \max\{1, C_0^{-1}\} \cdot v_0(\iota)
\end{aligned}
\tag{8.23}
$$

for all $\iota \in \mathbb{R}^d \setminus B_{c\varepsilon}(-\tau)$. Thus, in view of (8.22), it remains to estimate $\left(\partial^{\sigma}_{(\tau+\iota)}\left[\widetilde{\varrho}_*(|\tau+\iota|)\right]\right) \cdot A^{-1}(\tau)\langle\tau+\iota\rangle$ for $\tau+\iota \in \mathbb{R}^d \setminus B_{c\varepsilon}(0)$.

Lemma 8.5 implies

$$
A^{-1}(\tau) = [\mathrm{D}\Phi_{\varrho_*}(\tau)]^{-1} = [\widetilde{\varrho}_*(|\tau|)]^{-1} \cdot \pi_\tau^{\perp} + [\varrho_*'(|\tau|)]^{-1} \cdot \pi_\tau . \tag{8.24}
$$

Now, we apply (8.20), and v-moderateness of $\widetilde{\varrho}_*$ for radially increasing v, to derive

$$
\begin{aligned}
&\left|\partial_\iota^{\sigma}\left(\widetilde{\varrho}_*(|\tau+\iota|)\right)\right| \cdot \left[\widetilde{\varrho}_*(|\tau|)\right]^{-1} \cdot |\pi_\tau^{\perp}(\tau+\iota)| \\
&\lesssim \left[\widetilde{\varrho}_*(|\tau|)\right]^{-1} \cdot \widetilde{\varrho}_*(|\tau+\iota|) \cdot |\iota| \\
&\leq v(|\iota|) \cdot |\iota| \\
{\scriptstyle (v_0 \geq (1+|\bullet|)\cdot v(|\bullet|))}\; &\leq v_0(\iota).
\end{aligned}
\tag{8.25}
$$

Here, we additionally used the straightforward estimate $|\pi_\tau^{\perp}(\tau+\iota)| = |\pi_\tau^{\perp}(\tau) + \pi_\tau^{\perp}(\iota)| = |\pi_\tau^{\perp}(\iota)| \leq |\iota|$.

Finally, with the elementary estimate $|\pi_\tau(\iota+\tau)| \leq |\iota+\tau|$, we get

$$
\begin{aligned}
&\left|\partial_\iota^{\sigma}\left(\widetilde{\varrho}_*(|\iota+\tau|)\right)\right| \cdot \left[\varrho_*'(|\tau|)\right]^{-1} \cdot |\pi_\tau(\iota+\tau)| \\
{\scriptstyle ((8.19),\ (8.2),\ \text{and } (8.13))}\; &\lesssim |\iota+\tau| \cdot \frac{\varrho_*'(|\iota+\tau|)}{|\iota+\tau| \cdot \varrho_*'(|\tau|)} \\
{\scriptstyle ((8.3))}\; &\lesssim v(|\iota|) \leq C^{(11)} \cdot v_0(\iota).
\end{aligned}
\tag{8.26}
$$

Overall, combining (8.21)– (8.26), we finally see

$$
\begin{aligned}
\|\partial^\alpha \phi_\tau(\iota)\| = \|\partial^\alpha \varphi_\tau(\iota)\| &\leq d \cdot \max_{j \in \underline{d}} \left|\left[\partial^\alpha \varphi_\tau(\iota)\right]_{\bullet,j}\right| \\
{\scriptstyle ((8.21))}\; &\leq d \cdot \max_{j \in \underline{d}} \left|A^{-1}(\tau) \cdot (\partial^{\alpha+e_j}\Phi_{\varrho_*})(\tau+\iota)\right| \\
{\scriptstyle ((8.22))}\; &\leq d \cdot \max_{j \in \underline{d}} \Big(\left|\partial_\iota^{\alpha+e_j}\left(\widetilde{\varrho}_*(|\iota+\tau|)\right)\right| \cdot |A^{-1}(\tau)\,(\iota+\tau)| \\
&\qquad + |A^{-1}(\tau)\, v_{\alpha+e_j,\tau+\iota}|\Big) \\
{\scriptstyle ((8.23)\text{–}(8.26))}\; &\lesssim v_0(\iota) \quad \text{for all } \iota \in \mathbb{R}^d \setminus B_{c\varepsilon}(-\tau) \text{ and } |\alpha| \leq k\,,
\end{aligned}
$$

where the implied constant between left and right hand side depends on ϱ, v, d, and k.

Step 5 - Estimate $\partial^\alpha \phi_\tau(\iota)$ for $0 \le |\alpha| \le k$ and $\iota \in B_{c\varepsilon}(-\tau)$: By Lemma 8.5, $\varrho_*(t) = t/c$, and thus $\widetilde{\varrho_*}(t) = c^{-1}$ for $t \in (-c\varepsilon, c\varepsilon)$. Hence, $\Phi_{\varrho_*}(\tau) = c^{-1} \cdot \tau$ for all $\tau \in B_{c\varepsilon}(0)$, so that $A(\tau) = \mathrm{D}\Phi_{\varrho_*}(\tau) = c^{-1} \cdot \mathrm{id}_{\mathbb{R}^d}$ for $\tau \in B_{c\varepsilon}(0)$.

Hence, $\phi_\tau(\iota) = A^T(\tau + \iota) \cdot A^{-T}(\tau) = c^{-1} \cdot A^{-T}(\tau)$, whence $\|\partial^\alpha \phi_\tau(\iota)\| = 0 \le v_0(\iota)$ for $\iota \in B_{c\varepsilon}(-\tau)$ and $\alpha \in \mathbb{N}_0^d$ with $|\alpha| \in \underline{k}$. For $\alpha = 0$, Eq. (8.10) in Lemma 8.5 shows

$$\|\phi_\tau(\iota)\| = c^{-1} \cdot \|A^{-T}(\tau)\| = c^{-1} \cdot \|[\mathrm{D}\Phi_{\varrho_*}(\tau)]^{-1}\| \le \max\{1, C_0^{-1}\} \cdot c^{-1} / \widetilde{\varrho_*}(|\tau|)\,.$$

But since $\widetilde{\varrho_*}$ is v-moderate, we have $c^{-1} = \widetilde{\varrho_*}(0) \le \widetilde{\varrho_*}(|\tau|) \cdot v(|\tau|)$, and finally $|\tau| \le c\varepsilon + |\iota|$, such that $v(|\tau|) \le v(c\varepsilon) \cdot v(|\iota|)$. Altogether, $\|\phi_\tau(\iota)\| \le \max\{1, C_0^{-1}\} \cdot v(c\varepsilon) \cdot v(|\iota|) \lesssim v_0(\iota)$, for all $\tau \in \mathbb{R}^d$ and $\iota \in B_{c\varepsilon}(-\tau)$. □

That every radial warping function associated to a k-admissible radial component ϱ is indeed a k-admissible warping function is now a straightforward corollary.

Corollary 8.8 *Let $\varrho : \mathbb{R} \to \mathbb{R}$ be a k-admissible radial component with control weight v, for some $k \in \mathbb{N}$ with $k \ge d + 1$. Then there is a constant $C \ge 1$, dependent on ϱ, v, d, and k, such that with*

$$v_0 : \quad \mathbb{R}^d \to \mathbb{R}^+, \quad \tau \mapsto C \cdot (1 + |\tau|) \cdot v(|\tau|),$$

the associated radial warping function $\Phi_\varrho : \mathbb{R}^d \to \mathbb{R}^d$ is a k-admissible warping function, with control weight v_0. Furthermore, the weight $w = \det(\mathrm{D}\Phi_\varrho^{-1})$ is given by

$$w(\tau) = \varrho_*'(|\tau|) \cdot [\widetilde{\varrho_*}(|\tau|)]^{d-1}. \tag{8.27}$$

Proof Lemma 8.5 shows that $\Phi_\varrho : \mathbb{R}^d \to \mathbb{R}^d$ is a C^{k+1} diffeomorphism with $\Phi_\varrho^{-1} = \Phi_{\varrho_*}$, and (8.9) implies that $w(\tau) = \det \mathrm{D}\Phi_{\varrho_*}(\tau) = \varrho_*'(|\tau|) \cdot [\widetilde{\varrho_*}(|\tau|)]^{d-1} > 0$, for all $\tau \in \mathbb{R}^d \backslash \{0\}$. By continuity, and since $\varrho_*'(0) = \widetilde{\varrho_*}(0) = c^{-1}$ is positive, the above formula remains true for $\tau = 0$. The remaining properties required in Definition 8.1 follow from Proposition 8.7. □

8.2 The Slow Start Construction for Radial Components

So far, see Definition 8.1, we assumed that a k-admissible radial component ϱ has to be linear on a neighborhood of the origin. Our goal in this section is to show that if a given function ς satisfies (slightly modified versions of) all the other conditions from Definition 8.1, then one can modify ς in a neighborhood of the origin so that it becomes linear there, but all other properties are retained. We call this the **slow start construction**.

Definition 8.9 Fix some $\varepsilon > 0$, and let $\varsigma : [0, \infty) \to [0, \infty)$ be continuous and strictly increasing with $\varsigma(0) = 0$. Furthermore, fix $c \in \big(0, \varsigma(\varepsilon)/(2\varepsilon)\big)$, and an even function $\Omega \in C_c^\infty(\mathbb{R})$ that satisfies $\Omega(\xi) = 1$ for $x \in B_\varepsilon(0)$, $\Omega(\xi) = 0$ for $x \notin B_{2\varepsilon}(0)$, and $\Omega'(\xi) \le 0$ for $\xi \in [0, \infty)$. Then the function

$$\varrho : \quad \mathbb{R} \to \mathbb{R}, \quad \xi \mapsto \begin{cases} c\xi \cdot \Omega(\xi) + \operatorname{sgn}(\xi) \cdot (1 - \Omega(\xi)) \cdot \varsigma(|\xi|), & \text{if } \xi \neq 0\,, \\ 0, & \text{if } \xi = 0 \end{cases} \tag{8.28}$$

is called a **slow start version** of ς.

Remark 8.10 The intent of the slow start construction is to establish a k-admissible warping function that only differs from a radial function derived directly from ς in a small neighborhood of zero. This raises the question whether different slow start versions of ς, obtained, e.g., by choosing different values of ε in Definition 8.9, are equivalent in the sense that they generate the same coorbit spaces. Although we suspect that this can be shown directly by verifying the conditions of Proposition 2.15, instead, under fairly general conditions, we will obtain this equivalence as a consequence of identifying the respective coorbit spaces with certain decomposition spaces [19, 42, 94, 95] in a follow-up contribution.

The following lemma summarizes the main *elementary* properties of this construction.

Lemma 8.11 *Let $\varsigma : [0,\infty) \to [0,\infty)$ be continuous and strictly increasing with $\varsigma(0) = 0$. Let $\varepsilon > 0$ be arbitrary, and $c \in \big(0, \varsigma(\varepsilon)/(2\varepsilon)\big)$. Then, the function ϱ defined in (8.28) has the following properties:*

1. *We have $\varrho(\xi) = \varsigma(\xi)$ for all $\xi \in [2\varepsilon, \infty)$.*
2. *ϱ is antisymmetric.*
3. *$\varrho(\xi) = c\xi$ for all $\xi \in (-\varepsilon, \varepsilon)$.*
4. *If $\varsigma|_{\mathbb{R}^+}$ is C^k for some $k \in \mathbb{N}_0$, then ϱ is C^k.*
5. *If $\varsigma|_{\mathbb{R}^+}$ is C^1 with $\varsigma'(\xi) > 0$ for all $\xi \in (\varepsilon, \infty)$, then $\varrho'(\xi) > 0$ for all $\xi \in \mathbb{R}$.*
6. *If $\varsigma|_{\mathbb{R}^+}$ is C^k with $\varsigma'(\xi) > 0$ for all $\xi \in (\varepsilon, \infty)$, and if furthermore $\varsigma(\xi) \to \infty$ as $\xi \to \infty$, then $\varrho : \mathbb{R} \to \mathbb{R}$ is a C^k-diffeomorphism and $\varsigma : [0,\infty) \to [0,\infty)$ is a homeomorphism. Finally, we have*

$$\varrho^{-1}(\xi) = \varsigma^{-1}(\xi) \qquad \forall \xi \in [\varsigma(2\varepsilon), \infty).$$

Remark Item (6) above is particularly interesting, since it is often more important to know the properties of the *inverse* of the warping function ($\Phi_\varrho^{-1} = \Phi_{\varrho^{-1}}$ by Lemma 8.5) than those of the warping function itself.

Proof **Ad (1):** For $\xi \in [2\varepsilon, \infty)$, we have $\Omega(\xi) = 0$. Therefore, $\varrho(\xi) = \operatorname{sgn}(\xi) \cdot \varsigma(|\xi|) = \varsigma(\xi)$.

Ad (2): Ω is symmetric, i.e., $\Omega(-\xi) = \Omega(\xi)$ for all $\xi \in \mathbb{R}$. For $\xi \neq 0$, this implies

$$\begin{aligned} \varrho(-\xi) &= c \cdot (-\xi) \cdot \Omega(-\xi) + \operatorname{sgn}(-\xi) \cdot (1 - \Omega(-\xi)) \cdot \varsigma(|-\xi|) \\ &= -\Big(c\xi \cdot \Omega(\xi) + \operatorname{sgn}(\xi) \cdot (1 - \Omega(\xi)) \cdot \varsigma(|\xi|)\Big) = -\varrho(\xi)\,. \end{aligned}$$

For $\xi = 0$, we trivially have $\varrho(-\xi) = 0 = -\varrho(\xi)$.

Ad (3): By choice of Ω, we have $\Omega(\xi) = 1$ for $\xi \in (-\varepsilon, \varepsilon)$. For $\xi \neq 0$, this immediately yields $\varrho(\xi) = c\xi$, which clearly also holds for $\xi = 0$.

Ad (4): Since Ω is smooth, and since the functions $\xi \mapsto \mathrm{sgn}(\xi)$ and $\xi \mapsto |\xi|$ are smooth on $\mathbb{R}\backslash\{0\}$, it is clear that ϱ is $\mathcal{C}^k$ on $\mathbb{R} \setminus \{0\}$. But in the preceding point we saw that ϱ is linear (and hence smooth) in a neighborhood of zero. Hence, ϱ is $\mathcal{C}^k$.

Ad (5): On $(-\varepsilon, \varepsilon)$, we have $\varrho(\xi) = c\xi$, and thus $\varrho'(\xi) = c > 0$ on $[-\varepsilon, \varepsilon]$. Also, on $(-\infty, -2\varepsilon) \cup (2\varepsilon, \infty)$, we have $\Omega(\xi) = 0$, and hence $\varrho(\xi) = \mathrm{sgn}(\xi) \cdot \varsigma(|\xi|)$. Since $\xi \mapsto |\xi|$ is smooth away from zero, with $\frac{d}{d\xi}|\xi| = \mathrm{sgn}(\xi)$, this implies $\varrho'(\xi) = (\mathrm{sgn}(\xi))^2 \cdot \varsigma'(|\xi|) > 0$ for $\xi \in \mathbb{R}$ with $|\xi| \geq 2\varepsilon$.

For $\xi \in (\varepsilon, 2\varepsilon)$, we have $\varrho'(\xi) = \big[\Omega(\xi) \cdot c + (1 - \Omega(\xi)) \cdot \varsigma'(\xi)\big] + (-\Omega'(\xi)) \cdot (\varsigma(\xi) - c\xi) > 0$, since all three terms are nonnegative and they cannot vanish simultaneously. To see this, note that $\Omega'(\xi) \leq 0$ for $\xi \in [0, \infty)$, $\varsigma'(\xi) > 0$ for $\xi \in (\varepsilon, \infty)$, and $\varsigma(\xi) \geq \varsigma(\varepsilon) > 2c\varepsilon > c\xi$ for $\xi \in (\varepsilon, 2\varepsilon)$. For the last inequality, recall $c \in (0, \varsigma(\varepsilon)/(2\varepsilon))$. Positivity of ϱ' on $(-2\varepsilon, -\varepsilon)$ follows from ϱ being antisymmetric.

Ad (6): We have $\varrho(0) = 0$ and $\varrho(\xi) = \varsigma(\xi)$ for $\xi \geq 2\varepsilon$, such that $\varrho([0, \infty)) \supset [0, \infty)$ by the intermediate value theorem. Hence, ϱ is surjective by (2) and with $\varrho' > 0$ by (5) even bijective. As a strictly increasing bijective $\mathcal{C}^k$ map with positive derivative, ϱ is a $\mathcal{C}^k$-diffeomorphism by the inverse function theorem.

Similar arguments show that ς is a homeomorphism. The remaining property $\varrho_*(\xi) = \varsigma_*(\xi)$ for all $\xi \in [\varsigma(2\varepsilon), \infty)$ is now a straightforward consequence of $\varrho(\xi) = \varsigma(\xi)$ for all $\xi \in [2\varepsilon, \infty)$. □

Our final goal in this subsection is to state convenient criteria on ς which ensure that ϱ is a k-admissible radial component. For this, the following general lemma will be helpful.

Lemma 8.12 *Let $\delta > 0$, and let $\theta_1, \theta_2 : [\delta, \infty) \to [0, \infty)$ and $u : [0, \infty) \to \mathbb{R}^+$ be continuous and increasing with $u(\xi + \eta) \leq u(\xi) \cdot u(\eta)$ for all $\xi, \eta \in [0, \infty)$. Furthermore, assume that there is some $D > 0$ such that*

$$D \leq \theta_2(\eta) \cdot u(\eta) \quad \text{and} \quad \theta_1(\xi) \leq \theta_2(\eta) \cdot u(|\xi - \eta|) \quad \forall\, \xi, \eta \in [\delta, \infty)\,. \tag{8.29}$$

If $\beta_1 : \mathbb{R} \to [0, \infty)$ and $\beta_2 : \mathbb{R} \to \mathbb{R}^+$ are continuous with $\beta_j(\xi) = \theta_j(|\xi|)$ for all $\xi \in \mathbb{R}$ with $|\xi| \geq \delta$ and all $j \in \{1, 2\}$, then there is a constant $C \geq 1$ with

$$\beta_1(\xi) \leq C \cdot \beta_2(\eta) \cdot u(|\xi - \eta|) \qquad \forall\, \xi, \eta \in \mathbb{R}.$$

Proof. By continuity of $\beta_1 : \mathbb{R} \to [0, \infty)$ and $\beta_2 : \mathbb{R} \to \mathbb{R}^+$, there are constants $c_1, c_2 > 0$ with $\beta_1(\xi) \leq c_1$ and $\beta_2(\xi) \geq c_2$ for all $\xi \in [-\delta, \delta]$. Further, note that the conditions on u imply $u(0) \geq 1$ and that $u(|\bullet|)$ is submultiplicative and radially increasing. We distinguish four cases:

Case 1 ($|\xi| < \delta$ **and** $|\eta| < \delta$): $\beta_1(\xi) \leq c_1 \leq \frac{c_1}{c_2 \cdot u(0)} \cdot \beta_2(\eta) \cdot u(|\xi - \eta|)$.

Case 2 ($|\xi| \geq \delta$ **and** $|\eta| \geq \delta$): $\beta_1(\xi) = \theta_1(|\xi|) \leq \theta_2(|\eta|) \cdot u\big(\big|\,|\xi| - |\eta|\,\big|\big) \leq \beta_2(\eta) \cdot u(|\xi - \eta|)$.

Case 3 ($|\xi| < \delta$ **and** $|\eta| \geq \delta$): We have $D \leq \theta_2(|\eta|) \cdot u(|\eta|) \leq \theta_2(|\eta|) \cdot u(|\eta - \xi|) \cdot u(\delta)$, since $u(|\xi|) \leq u(\delta)$. Hence, $\beta_1(\xi) \leq c_1 \leq \frac{c_1 \cdot u(\delta)}{D} \cdot \theta_2(|\eta|) \cdot u(|\eta - \xi|) \leq \frac{c_1 \cdot u(\delta)}{D} \cdot \beta_2(\eta) \cdot u(|\eta - \xi|)$.

Case 4 ($|\xi| \geq \delta$ **and** $|\eta| < \delta$): We have $\big| |\xi| - \delta \big| \leq |\xi| \leq |\xi - \eta| + |\eta| < |\xi - \eta| + \delta$. Hence, $\beta_1(\xi) = \theta_1(|\xi|) \leq c_2^{-1} \cdot \beta_2(\eta) \cdot \theta_1(|\xi|) \leq \frac{\theta_2(\delta) \cdot u(\delta)}{c_2} \cdot \beta_2(\eta) \cdot u(|\xi - \eta|)$.

Altogether, we have shown $\beta_1(\xi) \leq C \beta_2(\eta) \cdot u(|\xi - \eta|)$ for all $\xi, \eta \in \mathbb{R}$, with

$$C := \max \left\{ 1, \quad \frac{c_1}{c_2 \cdot u(0)}, \quad \frac{c_1 \cdot u(\delta)}{D}, \quad \frac{\theta_2(\delta) \cdot u(\delta)}{c_2} \right\}.$$

□

We now formally introduce a class of functions $\varsigma : [0, \infty) \to [0, \infty)$ for which the slow-start construction produces a k-admissible radial component. This will be proven in Proposition 8.15 below.

Definition 8.13 Let $k \in \mathbb{N}_0$. A continuous function $\varsigma : [0, \infty) \to [0, \infty)$ is called a **weakly k-admissible radial component with control weight** $u : [0, \infty) \to \mathbb{R}^+$, if it satisfies the following conditions:

1. ς is C^{k+1} on $\mathbb{R}^+$, with $\varsigma'(\xi) > 0$ for all $\xi \in \mathbb{R}^+$.
2. $\varsigma(0) = 0$ and $\varsigma(\xi) \to \infty$ as $\xi \to \infty$.
3. The control weight u is continuous and increasing with $u(\xi + \eta) \leq u(\xi) \cdot u(\eta)$ for all $\xi, \eta \in [0, \infty)$. Furthermore, there are $\delta > 0$ and $C_0, C_1 > 0$ with the following properties:

$$C_0 \cdot \frac{\varsigma_*(\xi)}{\xi} \leq \varsigma_*'(\xi) \leq C_1 \cdot \varsigma_*(\xi) \qquad \forall \xi \in [\delta, \infty), \tag{8.30}$$

$$\frac{\varsigma_*(\xi)}{\xi} \leq \frac{\varsigma_*(\eta)}{\eta} \cdot u(|\xi - \eta|) \qquad \forall \xi, \eta \in [\delta, \infty), \tag{8.31}$$

$$|\varsigma_*^{(m)}(\xi)| \leq \varsigma_*'(\eta) \cdot u(|\xi - \eta|) \qquad \forall \xi, \eta \in [\delta, \infty) \text{ and } m \in \underline{k+1}. \tag{8.32}$$

Remark 8.14 Properties (1) and (2) imply that $\varsigma : [0, \infty) \to [0, \infty)$ is a homeomorphism, with inverse $\varsigma_* := \varsigma^{-1}$.

In many cases, one even has the stronger condition $\varsigma_*'(\xi) \asymp \varsigma_*(\xi)/\xi$ for all $\xi \in [\delta, \infty)$ instead of (8.30). In this case, it is not necessary to verify condition (8.31), since—after possibly replacing u by $C \cdot u$ for some $C \geq 1$—this condition is implied by (8.32) for $m = 1$. Indeed, if (8.32) holds, then

$$\frac{\varsigma_*(\xi)}{\xi} \asymp \varsigma_*'(\xi) \leq \varsigma_*'(\eta) \cdot u(|\xi - \eta|) \lesssim \frac{\varsigma_*(\eta)}{\eta} \cdot u(|\xi - \eta|) \qquad \text{for } \xi, \eta \in [\delta, \infty).$$

Overall, if $\varsigma_*'(\xi) \asymp \varsigma_*(\xi)/\xi$ for $\xi \in [\delta, \infty)$, then ς is a weakly k-admissible radial component, if ς is C^{k+1} with $\varsigma'(\xi) > 0$, $\varsigma(0) = 0$ and with $\varsigma(\xi) \to \infty$ as $\xi \to \infty$ and ς satisfies (8.32).

Our final result in this subsection shows that the slow-start construction, applied to a weakly k-admissible radial component, yields a k-admissible radial component.

Proposition 8.15 *Let $k \in \mathbb{N}_0$, and let $\varsigma : [0, \infty) \to [0, \infty)$ be a weakly k-admissible radial component with control weight $u : [0, \infty) \to \mathbb{R}^+$. Furthermore, let ϱ be a "slow-start version" of ς as in (8.28). Then there exists a constant $C := C(k) \geq 1$, such that ϱ is a k-admissible radial component with control weight*

$$v : \mathbb{R} \to \mathbb{R}^+, \xi \mapsto C \cdot u(|\xi|).$$

Proof Lemma 8.11 shows that $\varrho : \mathbb{R} \to \mathbb{R}$ satisfies conditions (1)–(3) of Definition 8.1. As already observed in the proof of Lemma 8.11, the conditions on u imply that $u(|\bullet|)$ is submultiplicative, such that the same holds for v, since $C \geq 1$. Note furthermore, that $\varrho_*(\xi) = \varsigma_*(\xi)$ for all $\xi \geq \delta' := \max\{\varsigma(2\varepsilon), \delta\}$, with $\varepsilon > 0$ as in Lemma 8.11 and $\delta > 0$ as in Definition 8.13.

We proceed to prove condition (5) of Definition 8.1: For $|\xi| \geq \delta'$, the inequality (8.2) (with some constants $\tilde{C}_1, \tilde{C}_2$ in place of C_1, C_2) is a direct consequence of (8.30) and (8.28).

For $|\xi| \leq \varsigma(\varepsilon)$, $\widetilde{\varrho_*}(\xi) = \varsigma_*(\xi)/\xi = c^{-1}$, such that $\widetilde{\varrho_*}$ is continuous and there are $c_1, c_2, c_3, c_4 > 0$, such that for all $\xi \in [-\delta', \delta']$, $c_1 \leq \widetilde{\varrho_*}(\xi) \leq c_2$ and $c_3 \leq \varrho_*'(\xi) \leq c_4$. Thus, with $C_1 = \min\{\tilde{C}_1, c_3/c_2\}$ and $C_2 = \max\{\tilde{C}_2, c_4/c_1\}$, (8.2) is satisfied for all $\xi \in \mathbb{R}$.

To prove condition (6) of Definition 8.1, consider the following: For $|\xi| \geq \delta'$, the antisymmetry of ϱ implies that
$\varrho_*(\xi) = \operatorname{sgn}(\xi) \cdot \varsigma_*(\operatorname{sgn}(\xi) \cdot \xi)$. A straightforward induction therefore shows

$$|\varrho_*^{(m)}(\xi)| = \left|\varsigma_*^{(m)}(\operatorname{sgn}(\xi) \cdot \xi)\right| = \left|\varsigma_*^{(m)}(|\xi|)\right| \quad \text{for all} \quad m \in \underline{k+1} \quad \text{and} \quad |\xi| \geq \delta'.$$

Furthermore, note that (8.32) with $m = 1$ and $\xi = \delta$ and $\varsigma'(\xi) > 0$ for all $\xi \in (\varepsilon, \infty)$ implies $0 < \varsigma_*'(\delta)/u(\delta) \leq \varsigma_*'(\eta)u(\eta)$, since u is increasing.

Fix some $\ell \in \underline{k+1}$. In view of (8.32), we can apply Lemma 8.12 (with δ' instead of δ), with $\theta_1 = \left|\varsigma_*^{(\ell)}|_{[\delta',\infty)}\right|$, $\theta_2 = \varsigma_*'|_{[\delta',\infty)}$, and with $\beta_1 = |\varrho_*^{(\ell)}|$, $\beta_2 = \varrho_*'$. Consequently, there is a constant $G_\ell \geq 1$ such that

$$|\varrho_*^{(\ell)}(\xi)| = \beta_1(\xi) \leq G_\ell \cdot \beta_2(\eta) \cdot u(|\eta - \xi|) = G_\ell \cdot \varrho_*'(\eta) \cdot u(|\eta - \xi|) \qquad \forall\, \eta, \xi \in \mathbb{R}. \tag{8.33}$$

Since $\ell \in \underline{k+1}$ was arbitrary, (8.3) is satisfied with $C \geq \max\{G_1, \ldots, G_{k+1}\}$.

In particular, if we set $\ell = 1$, then (8.33) implies that ϱ_*' is v-moderate with $v = Cu(|\bullet|)$ and any $C \geq G_1$. Hence, for condition (4) in Definition 8.1 it only remains to prove that $\widetilde{\varrho_*}$ is v-moderate.

With $\theta_1 = \theta_2 = \varsigma_*/|\bullet|$ (and δ' instead of δ) the inequality (8.29) is implied by (8.31). Therefore, we can invoke Lemma 8.12 with this choice of θ_1, θ_2 and $\beta_1 = \beta_2 = \widetilde{\varrho_*}$. Note that $\beta_j(\xi) = \theta_j(|\xi|)$ for $|\xi| \geq \delta'$. We obtain a constant $G \geq 1$, such that $\widetilde{\varrho_*}$ is v-moderate with $v = Cu(|\bullet|)$ and any $C \geq G$. Altogether, condition (4) in Definition 8.1 is satisfied with $v = Cu(|\bullet|)$, for any $C \geq \max\{G_1, G\}$. □

8.3 Examples of Radial Warping Functions

We now present two examples of radial components $\varsigma : [0, \infty) \to [0, \infty)$. We show that they are *weakly k-admissible* as per Definition 8.13. By Proposition 8.15 and Corollary 8.8, any slow start version ϱ of ς yields a radial, k-admissible warping function Φ_ϱ. Additionally, we provide in each case a control weight v_0 for Φ_ϱ.

Example 8.16 Let $p \geq 1$, and consider the function

$$\varsigma : [0, \infty) \to [0, \infty), \xi \mapsto (1 + \xi)^{1/p} - 1.$$

Conditions (1)–(2) of Definition 8.13 are clear. For $p = 1$, Condition (3) is easily verified with $u \equiv 1$. To verify Condition (3) for $p > 1$, we first show that $\varsigma_*' \asymp \varsigma_*/(\bullet)$. By Remark 8.14, it is then sufficient to verify only (8.32).

Note that $\varsigma_*(\xi) = (1+\xi)^p - 1$. For $\xi > \delta := 1$, it is easy to see that $(1+\xi)^r - 1 \asymp (1 + \xi)^r$, for any $r > 0$. In particular, with $r = p - 1$, we obtain

$$\varsigma_*'(\xi) = p \cdot (1 + \xi)^{p-1} \asymp \frac{(1 + \xi)^p}{1 + \xi} \asymp \frac{(1 + \xi)^p - 1}{\xi} = \frac{\varsigma_*(\xi)}{\xi} \quad \text{for} \quad \xi \geq 1.$$

Note the inequality $1 + \xi \leq 1 + \eta + |\xi - \eta| \leq (1 + \eta) \cdot (1 + |\xi - \eta|)$, which holds for $\eta, \xi \geq 0$. As a direct consequence, we obtain for all $\eta, \xi \geq 0$ and $\alpha, \beta \in \mathbb{R}$ with $\alpha \leq \beta$ that

$$(1 + \xi)^\alpha \leq (1 + \xi)^\beta \leq (1 + \eta)^\beta \cdot (1 + |\eta - \xi|)^{|\beta|}. \tag{8.34}$$

Define $\tilde{u} = (1+(\bullet))^{|p-1|}$ and note that $\varsigma_*^{(m)}(\xi) = C_m \cdot (1+\xi)^{p-m}$ for all $m \in \underline{k+1}$, for suitable constants $C_m = C_m(m, p) \in \mathbb{R}$, in particular, $C_1 = p > 0$. Therefore,

$$\begin{aligned} |\varsigma_*^{(m)}(\xi)| \leq |C_m| \cdot (1 + \xi)^{p-m} &\overset{(8.34)}{\leq} |C_m| \cdot (1 + \eta)^{p-1} \tilde{u}(|\xi - \eta|) \\ &= \frac{|C_m|}{p} \cdot \varsigma_*'(\eta) \cdot \tilde{u}(|\xi - \eta|), \end{aligned}$$

for all $\eta, \xi \geq 1$. This proves (8.32) with $u = \max_{m \in \underline{k+1}} \{|C_m|/p\} \cdot \tilde{u}$.

Hence, ς is a weakly k-admissible radial component with control weight $u : [0, \infty) \to \mathbb{R}^+, u(\xi) = C \cdot (1+\xi)^{|p-1|}$, for any $k \in \mathbb{N}_0$ and some appropriate constant $C \geq 1$, depending on p and k. By Proposition 8.15 any "slow start" version ϱ of ς is k-admissible, with control weight $v = C' \cdot u(|\bullet|)$, for some $C' \geq 1$. Therefore, Corollary 8.8 shows that the associated radial warping function Φ_ϱ is indeed a k-admissible warping function with control weight $v_0 = C'' \cdot (1+|\bullet|) \cdot u(|\bullet|) = C''(1+|\bullet|)^{1+|p-1|}$, for constant $C'' \geq 1$.

At this point, we conjecture that the coorbit spaces $\mathrm{Co}(\mathcal{G}(\theta, \Phi_\varrho), \mathbf{L}_\kappa^{p,q})$ that are associated to the warping function Φ_ϱ constructed here coincide with certain α-**modulation spaces** , specifically with $\alpha = p^{-1}(p - 1) \in [0, 1)$, for a proper choice of the weight κ. In future work, we will verify this by identifying $\mathrm{Co}(\mathcal{G}(\theta, \Phi_\varrho), \mathbf{L}_\kappa^{p,q})$ with certain decomposition spaces, cf. [19, 42], and considering embeddings between the resulting decomposition spaces and α-**modulation spaces** [30, 41, 51, 55] using the theory developed in [94, 95].

Example 8.17 Consider the function $\varsigma : [0,\infty) \to [0,\infty)$, $\xi \mapsto \ln(1+\xi)$. It is easy to see that conditions (1)–(2) of Definition 8.13 are satisfied and that $\varsigma_*(\xi) = \varsigma^{-1}(\xi) = e^{\xi} - 1$.

We now verify condition (3) of Definition 8.13 by proving that the inequalities (8.30)– (8.32) hold with $\delta = 1$ and $u : [0,\infty) \to [1,\infty)$, $\xi \mapsto e^{\xi}$. Note that $\varsigma_*^{(\ell)} = u$ for all $\ell \in \mathbb{N}$, such that (8.32) clearly holds, even for all $\xi \in \mathbb{R}^+$.

Ad (8.30): For $\xi \geq \delta = 1$, we have $1 \leq e^{\xi}/e$, and thus $\varsigma_*(\xi) = e^{\xi} - 1 \geq e^{\xi} \cdot (1 - e^{-1})$. Therefore,

$$\frac{e^{\xi} - 1}{\xi} \leq e^{\xi} \leq (1 - e^{-1})^{-1} \cdot \varsigma_*(\xi)\,,$$

so that (8.30) is fulfilled with $C_0 = 1$ and $C_1 = (1 - e^{-1})^{-1} > 0$.

Ad (8.31): Let $\widetilde{\varsigma}_*(\xi) := \frac{\varsigma_*(\xi)}{\xi} = \frac{e^{\xi}-1}{\xi}$ for $\xi \in \mathbb{R}^+$, and note that $\widetilde{\varsigma}_*$ has the power series expansion

$$\widetilde{\varsigma}_*(\xi) = \frac{1}{\xi} \cdot \left(\sum_{n=0}^{\infty} \frac{\xi^n}{n!} - 1 \right) = \sum_{n=1}^{\infty} \frac{\xi^{n-1}}{n!} = \sum_{\ell=0}^{\infty} \frac{\xi^{\ell}}{(\ell+1)!},$$

which shows that $\widetilde{\varsigma}_*$ is increasing, since each term of the series is increasing on $\mathbb{R}^+$. Therefore, $\xi \leq \eta$ implies $\varsigma_*(\xi) \leq \varsigma_*(\eta) \leq \varsigma_*(\eta) e^{|\xi-\eta|}$.

If $0 < \eta < \xi$, then

$$\frac{e^{\eta} - 1}{\eta} \cdot e^{|\xi-\eta|} = \frac{e^{\eta} - 1}{\eta} \cdot e^{\xi-\eta} = \frac{1 - e^{-\eta}}{\eta} \cdot e^{\xi} \geq \frac{e^{\xi} - 1}{\xi}.$$

Here, the final inequality uses that $\xi \mapsto \frac{1-e^{-\xi}}{\xi}$ is decreasing on $\mathbb{R}^+$. Therefore, (8.31) even holds for all $\xi, \eta \in \mathbb{R}^+$.

In other words, ς is a weakly k-admissible radial component with control weight $u : [0,\infty) \to \mathbb{R}^+$, $u(\xi) = e^{\xi}$ (for any $k \in \mathbb{N}_0$). By Proposition 8.15, any "slow start version" ϱ of ς as per (8.28), is a k-admissible radial component with control weight $v : \mathbb{R} \to \mathbb{R}^+$, $\xi \mapsto C \cdot e^{|\xi|}$, for some $C \geq 1$. By Corollary 8.8, the associated radial warping function Φ_{ϱ} is a k-admissible warping function with control weight $v_0 : \mathbb{R}^d \to \mathbb{R}^+$, $\tau \mapsto C' \cdot (1 + |\tau|) \cdot e^{|\tau|}$, for a suitable $C' \geq 1$.

It is likely that the coorbit spaces $\mathrm{Co}(\mathcal{G}(\theta, \Phi_{\varrho}), \mathbf{L}_{\kappa}^{p,q})$ associated with the warping function Φ_{ϱ} constructed can be embedded into certain **inhomogeneous Besov spaces** [88–90], if the weight κ is chosen properly. If such an embedding exists, we expect the converse to be true as well, possibly with a different weight $\tilde{\kappa}$ instead of κ. Similar to the previous examples, the interpretation of $\mathrm{Co}(\mathcal{G}(\theta, \Phi_{\varrho}), \mathbf{L}_{\kappa}^{p,q})$ as decomposition space will be the first step towards verifying such embeddings.

9 Conclusion

We developed a theory of warped time-frequency systems for functions of arbitrary dimensionality. These systems, defined by a prototype function θ and a diffeomorphism Φ, form tight continuous frames and admit the construction of coorbit spaces $\mathrm{Co}_\Phi(Y)$, which we have shown to be well-defined Banach spaces, provided that Φ is a k-admissible warping function and Y is a suitable, solid Banach space. We have further shown that stable discretization, in the sense of Banach frame decompositions, of the continuous system $\mathcal{G}(\theta, \Phi)$ is achieved across said coorbit spaces, simply by sampling densely enough. In all cases, the results are realized by choosing the prototype θ from a class of smooth, localized functions that includes $\mathcal{C}_c^\infty(\mathbb{R}^d)$. Moreover, the results can be invoked simultaneously for a large class of spaces Y including, but not limited to, weighted mixed-norm Lebesgue spaces $\mathbf{L}_\kappa^{p,q}$, $1 \leq p, q \leq \infty$. Finally, we considered radial warping functions as an important special case, showed how they can be constructed from (weakly) admissible radial components, and provided examples of radial warping functions for which we expect a relation to well-known smoothness spaces. Altogether, we have demonstrated that warped time-frequency systems, a vast class of translation-invariant time-frequency systems that enable the adaptation to a specific frequency-bandwidth relationship, can be analyzed with a unified, and surprisingly deep mathematical theory.

There is an abundance of opportunities for further generalization, of which we mention only two: (1) That the weight m may only depend on the time variable if $\sup_{\xi \in D} \|D\Phi(\xi)\| < \infty$ (in Theorems 4.4 and 6.1) remains an irritating and somewhat unnatural condition, but cannot be dropped if m is to be majorized by the product of a time-dependent and another frequency-dependent weight. If the latter requirement is relaxed and a more general weight is considered, it may be possible to consider time-dependent weights if $\|D\Phi(\xi)\|$ is unbounded. (2) The construction analyzed in this work does not accommodate frames with arbitrary directional sensitivity. In particular, the degree of anisotropy is determined directly by the warping function and cannot be chosen freely. For example, without further modification, it cannot currently mimic popular directional frames like curvelets or shearlets, or even isotropic wavelets; see below for the last point.

While the present article shows that the coorbit spaces $\mathrm{Co}_\Phi(Y)$ are well-defined Banach spaces admitting a rich discretization theory, it does not answer all open questions regarding the structure of $\mathrm{Co}_\Phi(Y)$ as *smoothness spaces*. These questions concern, e.g., the description of $\mathrm{Co}_\Phi(Y)$ purely in terms of Fourier analysis, as well as the existence of embeddings between the spaces $\mathrm{Co}_\Phi(Y)$ for different choices of the warping function Φ and the space Y, or between $\mathrm{Co}_\Phi(Y)$ and established smoothness spaces, such as Besov spaces, Sobolev spaces, α-modulation spaces, or spaces of dominating mixed smoothness [74, 75, 96]. In a follow-up article, we will study these questions in the context of decomposition spaces, a common generalization of Besov- and modulation spaces. Specifically, we will show that the spaces $\mathrm{Co}_\Phi(Y)$ are special decomposition spaces, so that the rich theory of these spaces can be employed to answer the questions posed above. In that work, we will confirm the conjectured relation to α-modulation spaces (see Example 8.16) and prove that equality between (inhomogeneous) Besov spaces and the coorbit spaces related to warped time-frequency systems can only be achieved in the one-dimensional case, thereby making the statement about isotropic wavelets in the previous paragraph formal.

Appendix A Formal Details for Making Sense of the Intersection $B \cap B'$

To make sense of the intersection $B \cap B'$ that appears in Definition 2.1, we assume that *B is compatible with a suitable* ***Gelfand triple*** in the following sense: We assume that there exists a topological vector space V of "test functions" which satisfies $V \hookrightarrow \mathcal{H}$, with dense image. For instance, in the case $\mathcal{H} = \mathbf{L}^2(\mathbb{R}^d)$ one could choose $V = C_c^\infty(\mathbb{R}^d)$ or $V = \mathcal{S}(\mathbb{R}^d)$. One can then identify each $h \in \mathcal{H}$ with the anti-linear functional (or "generalized distribution")

$$\varphi_h : \quad V \to \mathbb{C}, \quad v \mapsto \langle h, v \rangle_{\mathcal{H}},$$

since it is easy to see that the map $\mathcal{H} \to V^{\urcorner}, h \mapsto \varphi_h$ is linear and injective. Since $V \hookrightarrow \mathcal{H}$, we can thus consider V as a subset of $V^{\urcorner}$, by virtue of the dual pairing coming from $\mathcal{H}$.

Then, we say that a Banach space B is compatible with the Gelfand triple $(V, \mathcal{H}, V^{\urcorner})$, if B satisfies the following properties:

(i) $(B, \|\cdot\|_B)$ is a Banach space,
(ii) $B \subset V^{\urcorner}$ as sets (here, one potentially has to make some (canonical) identifications, such as considering $\mathbf{L}^p(\mathbb{R}^d)$ as a subset of $\mathcal{S}^{\urcorner}(\mathbb{R}^d)$),
(iii) the inclusion $B \hookrightarrow V^{\urcorner}$ is continuous (with respect to the weak-$*$-topology on $V^{\urcorner}$),
(iv) $V \subseteq B$ is dense (this rules out spaces such as $B = \mathbf{L}^\infty(\mathbb{R}^d)$ for $\mathcal{H} = \mathbf{L}^2(\mathbb{R}^d)$ and $V = \mathcal{S}(\mathbb{R}^d)$, but one can then instead use the closure of V in B, which in this case would be equal to $C_0(\mathbb{R}^d)$, the space of continuous functions "vanishing at infinity").

For such a compatible Banach space, we then say that $b \in B \subset V^{\urcorner}$ satisfies $b \in B \cap B'$, if there exists a constant $C > 0$ such that

$$|b(v)| \leq C \cdot \|v\|_B \qquad \forall\, v \in V.$$

Since $V \subset B$ is dense, this implies that $\overline{b} \in V'$ (given by $\langle \overline{b}, v \rangle_{V,V'} = \overline{b(v)}$) uniquely extends to a continuous linear functional on B; we then identify b with this functional.

Note that if $h \in B \cap \mathcal{H}$, then since we are identifying h with the functional φ_h, we have for $v \in V$ that

$$\overline{h}(v) = \overline{\varphi_h(v)} = \langle v, h \rangle_{\mathcal{H}},$$

so that this interpretation of elements of B as elements of B' is again consistent with the duality pairing coming from $\mathcal{H}$.

Acknowledgements The authors wish to thank the anonymous reviewers for their comments and suggestions and the role they played in improving the final manuscript. N.H. is grateful for the hospitality and support of the Katholische Universität Eichstätt-Ingolstadt during his visit. F.V. would like to thank the

Acoustics Research Institute for the hospitality during several visits, which were partially supported by the Austrian Science Fund (FWF): 31225–N32.

Author Contributions N.H. and F.V. collaborated on all aspects of the presented work. N.H. and F.V. jointly drafted the manuscript, and contributed to and approved the submitted version. We declare that both authors contributed equally to this work.

Funding Open access funding provided by Österreichische Akademie der Wissenschaften. N.H. was supported by the Austrian Science Fund (FWF): I 3067–N30; F.V. acknowledges support by the German Science Foundation (DFG) in the context of the Emmy Noether junior research group: VO 2594/1–1.

Declarations

Conflict of interest None

Ethical standards The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

References

1. Abreu, L.D.: Sampling and interpolation in Bargmann–Fock spaces of polyanalytic functions. Appl. Comput. Harmon. Anal. **29**(3), 287–302 (2010)
2. Abreu, L.D.: Super-wavelets versus poly-Bergman spaces. Integr. Equ. Oper. Theory **73**(2), 177–193 (2012)
3. Adams, R.A., Fournier, J.J.: Sobolev Spaces, vol. 140. Academic Press, Cambridge (2003)
4. Adler, A., Emiya, V., Jafari, M.G., Elad, M., Gribonval, R., Plumbley, M.D.: Audio inpainting. IEEE Trans. Audio Speech Lang. Process. **20**(3), 922–932 (2012)
5. Akan, A., Chaparro, L.: Multi-window Gabor expansion for evolutionary spectral analysis. Signal Process. **63**(3), 249–262 (1997)
6. Ali, S., Antoine, J.-P., Gazeau, J.-P.: Continuous frames in Hilbert space. Ann. Phys. **222**(1), 1–37 (1993)
7. Ali, S., Antoine, J.-P., Gazeau, J.-P.: Coherent States,Wavelets and their Generalizations. Graduate Texts in Contemporary Physics. Springer, New York (2000)
8. Balan, R., Christensen, J., Krishtal, I.A., Okoudjou, K.A., Romero, J.L.: Multi-window Gabor frames in amalgam spaces. Math. Res. Lett. **21**(1), 55–69 (2014)
9. Balan, R.V.: Multiplexing of signals using superframes. In: Wavelet Applications in Signal and Image Processing VIII, volume 4119, pp. 118–130. International Society for Optics and Photonics (2000)
10. Balazs, P., Dörfler, M., Jaillet, F., Holighaus, N., Velasco, G.A.: Theory, implementation and applications of nonstationary Gabor frames. J. Comput. Appl. Math. **236**(6), 1481–1496 (2011)
11. Balazs, P., Holighaus, N.: Discretization in generalized coorbit spaces: extensions, annotations and errata for "Continuous frames, function spaces and the discretization problem" by M. Fornasier and H. Rauhut. https://www.univie.ac.at/nonstatgab/warping/baho15.pdf and arXiv:1702.06485 (2015)
12. Balk, M.B.: Polyanalytic functions. Mathematical Research (1991)
13. Bargmann, V.: On a Hilbert space of analytic functions and an associated integral transform. Commun. Pure Appl. Math. **14**, 187–214 (1961)

14. Bargmann, V.: On a Hilbert space of analytic functions and an associated integral transform. Part II. A family of related function spaces. Application to distribution theory. Commun. Pure Appl. Math. **20**, 1–101 (1967)
15. Bernier, D., Taylor, K.F.: Wavelets from square-integrable representations. SIAM J. Math. Anal. **27**(2), 594–608 (1996)
16. Besov, O.V.: On a certain family of functional spaces. Embedding and extension theorems. Dokl. Akad. Nauk SSSR **126**, 1163–1165 (1959)
17. Boas, R.P., Jr.: Entire Functions. Academic Press Inc., New York (1954)
18. Boashash, B.: Time-Frequency Signal Analysis and Processing: A Comprehensive Reference. Academic Press, Cambridge (2015)
19. Borup, L., Nielsen, M.: Frame decomposition of decomposition spaces. J. Fourier Anal. Appl. **13**(1), 39–70 (2007)
20. Cai, J.-F., Chan, R.H., Shen, Z.: A framelet-based image inpainting algorithm. Appl. Comput. Harmon. Anal. **24**(2), 131–149 (2008). (**Special Issue on Mathematical Imaging – Part II**)
21. Candès, E., Donoho, D.: New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities. Commun. Pure Appl. Math. **57**(2), 219–266 (2004)
22. Candès, E.J., Donoho, D.L.: Ridgelets: A key to higher-dimensional intermittency? Philos. Trans. R. Soc. Lond. A **357**(1760), 2495–2509 (1999)
23. Candès, E.J., Donoho, D.L.: Continuous curvelet transform: I. Resolution of the wavefront set. Appl. Comput. Harmon. Anal. **19**(2), 162–197 (2005)
24. Chan, T.F., Shen, J., Zhou, H.-M.: Total variation wavelet inpainting. J. Math. Imaging Vis. **25**(1), 107–125 (2006)
25. Christensen, O.: An Introduction to Frames and Riesz Bases. Springer International Publishing, Berlin (2016)
26. Cohen, A., Dahmen, W., DeVore, R.: Adaptive wavelet techniques in numerical simulation. Encyclopedia of Computational Mechanics (2004)
27. Cohn, D.L.: Measure theory. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], second edition Birkhäuser/Springer, New York (2013)
28. Constantine, G., Savits, T.: A multivariate Faa di Bruno formula with applications. Trans. Am. Math. Soc. **348**(2), 503–520 (1996)
29. Cordoba, A., Fefferman, C.: Wave packets and Fourier integral operators. Commun. Partial Differ. Equ. **3**, 979–1005 (1978)
30. Dahlke, S., Fornasier, M., Rauhut, H., Steidl, G., Teschke, G.: Generalized coorbit theory, Banach frames, and the relation to α-modulation spaces. Proc. Lond. Math. Soc. **96**(2), 464–506 (2008)
31. Dahlke, S., Kutyniok, G., Steidl, G., Teschke, G.: Shearlet coorbit spaces and associated Banach frames. Appl. Comput. Harmon. Anal. **27**(2), 195–214 (2009)
32. Daubechies, I.: Ten Lectures on Wavelets, volume 61 of CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia, PA (1992)
33. Daubechies, I., Wang, Y.G., Wu, H.: Conceft: concentration of frequency and time via a multitapered synchrosqueezed transform. Philos. Trans. R. Soc. A **374**(2065), 20150193 (2016)
34. Dörfler, M.: Quilted Gabor frames: a new concept for adaptive time-frequency representation. Adv. Appl. Math. **47**(4), 668–687 (2011)
35. Dörfler, M., Matusiak, E.: Nonstationary Gabor frames: Approximately dual frames and reconstruction errors. Adv. Comput. Math. **41**, 293–316 (2014)
36. Dörfler, M., Matusiak, E.: Nonstationary Gabor frames: Existence and construction. Int. J. Wavelets Multiresolut. Inf. Process. **12**(03), 1450032 (2014)
37. Dörfler, M., Romero, J.L.: Frames adapted to a phase-space cover. Constr. Approx. **39**(3), 445–484 (2014)
38. Duistermaat, J.J.: Fourier Integral Operators. Birkhäuser, Basel (2011)
39. Duistermaat, J.J., Hörmander, L.: Fourier integral operators. II. Acta Math. **128**(1), 183–269 (1972)
40. Feichtinger, H.: Modulation Spaces on Locally Compact Abelian Groups. Technical Report. University of Vienna, Vienna (1983)
41. Feichtinger, H.G., Fornasier, M.: Flexible Gabor-wavelet atomic decompositions for L_2 Sobolev spaces. Ann. Mat. Pura Appl. **185**(1), 105–131 (2006)
42. Feichtinger, H.G., Gröbner, P.: Banach spaces of distributions defined by decomposition methods. I. Math. Nachr. **123**, 97–120 (1985)

43. Feichtinger, H.G., Gröchenig, K.: Banach spaces related to integrable group representations and their atomic decompositions, I. J. Funct. Anal. **86**(2), 307–340 (1989)
44. Feichtinger, H.G., Gröchenig, K.: Banach spaces related to integrable group representations and their atomic decompositions, II. Monatsh. Math. **108**(2–3), 129–148 (1989)
45. Folland, G.: Real Analysis: Modern Techniques and Their Applications. Pure and Applied Mathematics, second edition Wiley, New York (1999)
46. Fornasier, M., Rauhut, H.: Continuous frames, function spaces, and the discretization problem. J. Fourier Anal. Appl. **11**(3), 245–287 (2005)
47. Führ, H.: Wavelet frames and admissibility in higher dimensions. J. Math. Phys. **37**(12), 6353–6366 (1996)
48. Führ, H.: Abstract Harmonic Analysis of Continuous Wavelet Transforms. Springer, New York (2005)
49. Führ, H., Lemvig, J.: System bandwidth and the existence of generalized shift-invariant frames. J. Funct. Anal. **276**(2), 563–601 (2019)
50. Gabor, D.: Theory of communication. J. IEE **93**(26), 429–457 (1946)
51. Gröbner, P.: Banachräume glatter Funktionen und Zerlegungsmethoden. PhD thesis, University of Vienna (1992)
52. Gröchenig, K.: Describing functions: atomic decompositions versus frames. Monatsh. Math. **112**(3), 1–41 (1991)
53. Gröchenig, K.: Foundations of Time-Frequency Analysis. Appl. Numer. Harmon. Anal, Birkhäuser, Boston, MA (2001)
54. Gröchenig, K., Lyubarskii, Y.: Gabor (super) frames with Hermite functions. Math. Ann. **345**(2), 267 (2009)
55. Han, J., Wang, B.: α-modulation spaces (i) scaling, embedding and algebraic properties. J. Math. Soc. Jpn. **66**(4), 1315–1373 (2014)
56. Harbrecht, H., Utzinger, M.: On adaptive wavelet boundary element methods. J. Comput. Math. **36**(1), 90–109 (2018)
57. Heil, C.: Wiener amalgam spaces in generalized harmonic analysis and wavelet theory. ProQuest LLC, Ann Arbor, MI, (1990). Thesis (Ph.D.)–University of Maryland, College Park, https://drum.lib.umd.edu/handle/1903/24738
58. Hernández, E., Labate, D., Weiss, G.: A unified characterization of reproducing systems generated by a finite family. II. J. Geom. Anal. **12**(4), 615–662 (2002)
59. Holighaus, N.: Structure of nonstationary Gabor frames and their dual systems. Appl. Comput. Harmon. Anal. **37**(3), 442–463 (2014)
60. Holighaus, N., Voigtlaender, F.: Schur-type Banach modules of integral kernels acting on mixed-norm Lebesgue spaces. J. Funct. Anal. **281**(9), 109197 (2021)
61. Holighaus, N., Wiesmeyr, C., Balazs, P.: Continuous warped time-frequency representations-coorbit spaces and discretization. Appl. Comput. Harmon. Anal. **47**(3), 975–1013 (2019)
62. Holighaus, N., Wiesmeyr, C., Průša, Z.: A class of warped filter bank frames tailored to non-linear frequency scales. J. Fourier Anal. Appl. **26**, 22 (2020)
63. Hörmander, L.: Fourier integral operators. I. Acta Math. **127**(1), 79–183 (1971)
64. Hörmander, L.: The Analysis of Linear Partial Differential Operators I: Distribution Theory and Fourier Analysis. Classics in Mathematics. Springer, Berlin Heidelberg (2015)
65. Jakobsen, M.S., Lemvig, J.: Reproducing formulas for generalized translation invariant systems on locally compact abelian groups. Trans. Am. Math. Soc. **368**(12), 8447–8480 (2016)
66. Kempka, H., Schäfer, M., Ullrich, T.: General coorbit space theory for quasi-Banach spaces and inhomogeneous function spaces with variable smoothness and integrability. J. Fourier Anal. Appl. **23**(6), 1348–1407 (2017)
67. Kutyniok, G., Labate, D.: Shearlets: Multiscale Analysis for Multivariate Data. Springer, Berlin (2012)
68. Leoni, G.: A First Course in Sobolev Spaces. Graduate Studies in Mathematics, vol. 181, 2nd edn. American Mathematical Society, Providence, RI (2017)
69. Li, H., Manjunath, B., Mitra, S.K.: Multisensor image fusion using the wavelet transform. Graph. Models Image Process. **57**(3), 235–245 (1995)
70. Magnus, J.R., Neudecker, H.: Matrix differential calculus with applications in statistics and econometrics. Wiley Series in Probability and Statistics. Wiley, Chichester (1999). Revised reprint of the 1988 original
71. Mallat, S.: A Wavelet Tour of Signal Processing: The Sparse Way, 3rd edn. Elsevier/Academic Press, Amsterdam (2008)

72. Moyal, J.E.: Quantum mechanics as a statistical theory. Proc. Camb. Philos. Soc. **45**, 99–124 (1949)
73. Nazaret, B., Holschneider, M.: An interpolation family between Gabor and wavelet transformations: application to differential calculus and construction of anisotropic Banach spaces. In Nonlinear hyperbolic equations, spectral theory, and wavelet transformations, volume 145 of Oper. Theory Adv. Appl., pp. 363–394. Birkhäuser, Basel (2003)
74. Nikol'skij, S.M.: On boundary properties of differentiable functions of several variables. Dokl. Akad. Nauk SSSR **146**, 542–545 (1962)
75. Nikol'skij, S.M.: On stable boundary values of differentiable functions of several variables. Mat. Sb. **61**, 224–252 (1963)
76. Perelomov, A.: Generalized Coherent States and Their Applications. Springer Science & Business Media, Berlin (2012)
77. Portnoff, M.: Time-scale modification of speech based on short-time Fourier analysis. IEEE Trans. Acoust. Speech Signal Process. **29**(3), 374–390 (1981)
78. Rauhut, H., Ullrich, T.: Generalized coorbit space theory and inhomogeneous function spaces of Besov–Lizorkin–Triebel type. J. Funct. Anal. **260**(11), 3299–3362 (2011)
79. Riesz, F., Sz.-Nagy, B.: Functional Analysis. Dover Books on Advanced Mathematics. Dover Publications Inc, New York (1990)
80. Romero, J.L.: Surgery of spline-type and molecular frames. J. Fourier Anal. Appl. **17**, 135–174 (2011)
81. Ron, A., Shen, Z.: Generalized shift-invariant systems. Constr. Approx., pp. OF1–OF45, (2004)
82. Schwartz, L.: Théorie des Distributions. (Distribution Theory). Nouveau Tirage. Vols. 1. Paris: Hermann (1957)
83. Schwartz, L.: Théorie des Distributions. (Distribution Theory). Nouveau Tirage. Vols. 2. Paris: Hermann (1957)
84. Speckbacher, M., Bayer, D., Dahlke, S., Balazs, P.: The α-modulation transform: admissibility, coorbit theory and frames of compactly supported functions. Monatsh. Math. **184**(1), 133–169 (2017)
85. Stein, E.M.: Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals. Princeton University Press, Princeton (1993)
86. Taubman, D., Marcellin, M.: JPEG2000 Image Compression Fundamentals, Standards and Practice, vol. 642. Springer Science & Business Media, Berlin (2012)
87. Thomson, D.J.: Spectrum estimation and harmonic analysis. Proc. IEEE **70**(9), 1055–1096 (1982)
88. Triebel, H.: Theory of Function Spaces I. Number 1 in Monographs in Mathematics. Birkhäuser Verlag, Boston (1983)
89. Triebel, H.: Characterizations of Besov-Hardy-Sobolev spaces: A unified approach. J. Approx. Theory **52**(2), 162–203 (1988)
90. Triebel, H.: Theory of Function Spaces III, Monographs in Mathematics, vol. 100. Birkhäuser, Boston (2006)
91. Triebel, H.: Theory of Function Spaces II. Modern Birkhäuser Classics, Springer, Basel (2010)
92. Ullrich, T.: Continuous characterizations of Besov–Lizorkin–Triebel spaces and new interpretations as coorbits. J. Funct. Spaces Appl. **47**, 163213 (2012)
93. van Velthoven, J.T.: On the local integrability condition for generalised translation-invariant systems. Collect. Math. **70**(3), 407–429 (2019)
94. Voigtlaender, F.: Embedding Theorems for Decomposition Spaces with Applications to Wavelet Coorbit Spaces. PhD thesis, RWTH Aachen University (2015). http://publications.rwth-aachen.de/record/564979
95. Voigtlaender, F.: Embeddings of decomposition spaces. Mem. Am. Math. Soc. **287**, 1426 (2023)
96. Vybiral, J.: Function spaces with dominating mixed smoothness. PhD thesis, PhD thesis. Friedrich-Schiller Universitäat Jena (2006)
97. Xiao, J., Flandrin, P.: Multitaper time-frequency reassignment for nonstationary spectrum estimation and chirp enhancement. IEEE Trans. Signal Process. **55**(6), 2851–2860 (2007)
98. Zhao, S., Robeltson, D., Wang, G., Whiting, B., Bae, K.T.: X-ray CT metal artifact reduction using wavelets: an application for imaging total hip prostheses. IEEE Trans. Med. Imaging **19**(12), 1238–1247 (2000)
99. Zibulski, M., Zeevi, Y.Y.: Analysis of multiwindow Gabor-type schemes by frame methods. Appl. Comput. Harmon. Anal. **4**(2), 188–221 (1997)