Materials & Design 265 (2026) 115888

Contents lists available at ScienceDirect

Materials & Design

journal homepage: www.elsevier.com/locate/matdes

Explainable zero-shot transfer learning for cross-domain In-Situ acoustic monitoring in laser powder bed fusion process using learnable wavelet scattering

Vigneashwara Pandiyan [a], Rafał Wróbel [b,c], Sergey Shevchik [c], Christian Leinenbach [c,d,*]

[a] Digital Manufacturing and Surface Engineering (DMS), Department of Mechanical and Materials Engineering, University of Turku, FI-20014, Turun Yliopisto, Finland
[b] ETH Zürich, Department of Materials, Laboratory for Nanometallurgy, Vladimir-Prelog-Weg 1-5/10, Zürich 8093, Switzerland
[c] Laboratory for Advanced Materials Processing (LAMP), Swiss Federal Laboratories for Materials Science and Technology (Empa)-CH-3602 Thun, Switzerland
[d] Laboratory for Photonic Materials and Characterization, Ecole Polytechnique Fédérale de Lausanne 1015 Lausanne, Switzerland

ARTICLE INFO

Keywords:
Laser Powder Bed Fusion
In-situ Monitoring
Wavelet Scattering
Domain Adaptation
Acoustic Emission
Explainable AI (XAI)

ABSTRACT

Reliable in-situ monitoring of Laser Powder Bed Fusion (LPBF) remains challenging because acoustic emission (AE) signals exhibit data drift under changes in material composition, scan parameters, and sensor conditions. We propose an explainable Learnable Wavelet Scattering (LWS) framework that learns physics-consistent time–frequency representations and enables cross-domain generalization via zero-shot transfer learning. A trainable Morlet wavelet bank adaptively refines its center frequencies and bandwidths to capture process-specific spectral patterns. Multi-scale scattering coefficients are projected into a compact latent space and classified into melt-pool regimes: lack of fusion (LoF), conduction, and keyhole. Bayesian optimization selects an effective parameter configuration, achieving ~97% validation accuracy with stable convergence. Model-level causal influence quantifies band-wise contributions, showing that keyhole dynamics are dominated by low-frequency bands, whereas conduction and LoF rely on mid-to-higher frequencies. The learned filters converge toward physically meaningful bands most responsive to melt-pool transitions, providing actionable guidance for sensor bandwidth selection and tuning. Zero-shot transfer to an unseen dataset maintains high performance without retraining, indicating domain-invariant embeddings. Overall, LWS delivers an interpretable and robust AE-based monitoring approach for LPBF under realistic process drift. The framework is lightweight, requires only AE waveforms, and can be integrated into digital-twin workflows for scalable transferable process-state recognition.

1. Introduction

Additive Manufacturing (AM) represents a transformative approach to production, enabling the fabrication of complex geometries with unparalleled design freedom, minimized material waste, and enhanced resource efficiency. Its layer-wise fabrication principle facilitates the creation of lightweight, topology-optimized structures that are often unattainable through conventional subtractive methods [1]. This capability has spurred innovation across high-value sectors such as aerospace, automotive, biomedical, and energy, where functional integration and mass reduction are paramount. Furthermore, the rapid prototyping capacity of AM significantly shortens development cycles, accelerating the transition from conceptual design to final deployment. For low-volume or customized production, AM offers distinct economic advantages by eliminating the need for specialized tooling and enabling on-demand manufacturing. Its near-net-shape nature and absence of part-specific molds drastically reduce material waste, supporting both cost-effectiveness and environmental sustainability. Within this broad spectrum of AM technologies, metal-based processes have garnered significant prominence due to their ability to produce fully dense, structural components directly from digital models. Among these, Laser Powder Bed Fusion (LPBF) has emerged as the most mature and widely adopted technique for producing high-precision metallic parts with controlled microstructural attributes [2]. The pre-eminence of LPBF within metal AM is distinguished by its capacity to fabricate intricate metallic components with sub-millimeter dimensional accuracy. A

* Corresponding author at: Laboratory for Advanced Materials Processing (LAMP), Swiss Federal Laboratories for Materials Science and Technology (Empa)-CH-3602 Thun, Switzerland.
E-mail address: christian.leinenbach@empa.ch (C. Leinenbach).

https://doi.org/10.1016/j.matdes.2026.115888
Received 24 January 2026; Received in revised form 28 February 2026; Accepted 19 March 2026
Available online 21 March 2026

further critical advantage is the substantial improvement in material utilization efficiency, as LPBF drastically reduces the buy-to-fly ratio compared to conventional subtractive machining [3]. Despite these advantages, the LPBF process is inherently challenged by its complex, multi-physics nature, which involves tightly coupled thermo-fluid-mechanical dynamics. The process initiates as a high-energy laser source irradiates the powder bed, leading to rapid heating, melting, and the formation of a transient melt pool governed by Marangoni convection, capillary forces, and vapor recoil pressure [4–6]. The interplay between energy input, powder-bed characteristics, and melt-pool hydrodynamics is highly nonlinear. Consequently, even minor deviations in critical process parameters, such as volumetric energy density, can rapidly shift the local energy balance, destabilizing the melt pool and inducing transitions between stable conduction, unstable keyhole (driven by vapor-cavity oscillation and collapse), and *LoF* regimes [7–9]. These detrimental mode transitions directly generate defects—including gas-entrapped porosity and *LoF*, balling instabilities, and spatter ejection—which can propagate through successive layers [10–12]. This defect inheritance ultimately compromises the mechanical integrity, fatigue life, and dimensional fidelity of the fabricated component.

A comprehensive understanding of these process instabilities necessitates a detailed examination of the governing factors and their controllability. In LPBF, process reliability is not dictated by a single variable but emerges from the complex interplay of preset conditions, actively controlled parameters, and stochastic in-situ phenomena [13,14]. The primary controllable inputs are established during process planning and include the core processing parameters (laser power, scan speed, hatch spacing, and layer thickness) and part design decisions (orientation and support structures). These inputs define the intended energy density and thermal management strategy. However, these inputs interact with inherent or quasi-static conditions that are typically fixed for a given build. These include material characteristics (powder morphology, size distribution, and chemical composition) and the fundamental calibration of the equipment (optical alignment and laser spot characteristics) [15]. The central challenge arises from the dynamic and stochastic physical phenomena that occur during fabrication [9]. These are not directly controllable and include local variations in powder packing density, the complex thermo-fluid dynamics of the melt pool (e.g., Marangoni convection, vapor plume interactions, and spatter ejection), and the resultant ensuing solidification [16,17]. Process stability exists within a narrow window, determined by the complex interplay of machine settings, material properties, and the inherent randomness of the melting process. This variability makes it difficult to consistently achieve critical quality attributes like density and stress, acting as a major barrier to qualification [18]. Ultimately, the unpredictable nature of this multi-parameter system remains a significant barrier to the qualification of high-value components, resulting in increased rework, intensive non-destructive inspection, and extended time-to-market.

The inherent coupling of controllable inputs with stochastic variables renders the assurance of repeatable quality through parameter optimization alone exceptionally challenging. Consequently, achieving robust process qualification is increasingly dependent on multi-scale, in-situ monitoring to capture the transient physical signatures of the process as it unfolds [19–21]. Traditional qualification methods, reliant on post-build destructive testing and non-destructive evaluation (e.g., X-ray computed tomography), are fundamentally post-mortem and inefficient [22]. These approaches only detect defects after fabrication is complete, resulting in significant scrap loss and providing no insight into the dynamic phenomena that initiated the flaw. This limitation is critically amplified by the sequential nature of layer-wise manufacturing, where a single undetected defect can propagate through subsequent layers, compromising an entire build. To address these shortcomings, the AM community is shifting towards integrated sensor systems designed to capture critical physical signatures, such as thermal evolution and melt-pool dynamics, across distinct spatiotemporal scales [23,24]. This strategy encompasses monitoring powder-bed uniformity at the macroscale, melt-pool morphology and thermal gradients at the mesoscale, and plume and spatter dynamics at the microscale. The correlation of data from off-axial and co-axial photodiodes, high-speed cameras, and Acoustic Emission (AE) sensors yields multi-modal signals that encode rich information on energy-coupling efficiency, melt-pool stability, and defect nucleation [25–27]. This integrated, multi-scale sensing framework provides the foundational data necessary to transition from post-mortem inspection to predictive quality assurance via in-situ monitoring [28–31]. Optical and thermal monitoring systems yield critical information on surface emissions during LPBF but remain susceptible to occlusion and optical-path distortions. Air-borne AE sensing, by contrast, offers a versatile and economical means to resolve the rapid transient dynamics of the melt pool [32–36]. In addition to air-borne AE, structure-borne AE has been widely investigated for LPBF monitoring, typically by mounting piezoceramic transducers beneath the build plate or on the machine frame to capture elastic waves propagating through the solid structure [37,38]. Structure-borne AE is particularly effective for fast, bulk-driven events such as cracking and delamination, which often generate broadband emissions with dominant peaks above 100 kHz and are captured more reliably in solids than in air due to reduced attenuation [39–41]. However, structure-borne AE responses are strongly shaped by sensor coupling, structural transfer paths, and platform-specific vibration propagation, which can limit reproducibility and complicate direct comparability across machines without careful calibration and transfer-function modelling [37,38]. In contrast, air-borne AE uses non-contact microphones to measure pressure waves in the chamber atmosphere and are well suited for capturing plume and surface melt-pool dynamics, which are closely coupled to regime transitions and are commonly exploited for regime classification [37,38]. From a deployment standpoint, air-borne AE offers simpler integration, avoids coupling variability, and enables more standardized sensor placement across builds and machines, which aligns directly with the goal of learning stable embeddings under domain shift.

Machine Learning (ML) has become an indispensable tool for AM process monitoring, addressing the profound complexity and high dimensionality of in-situ sensor data that eludes conventional rule-based analysis [42–44]. The stochastic nature of LPBF generates intricate, nonlinear relationships between process signals and defect formation, which ML is uniquely suited to decode [26]. Standard monitoring setups produce heterogeneous, multi-modal data streams: one-dimensional temporal signals from photodiodes and acoustic sensors, two-dimensional spatial data from high-speed melt-pool imaging, and higher-order reconstructions from thermography. Uncovering physically meaningful correlations across these disparate modalities requires algorithms capable of capturing both long-range temporal dependencies and complex spatial contexts—a task that exceeds the capabilities of classical statistical methods [45,46]. By training on large datasets from co-axial and off-axial sensors, ML models learn the latent mappings between process signatures and melt-pool stability regimes, enabling real-time defect classification, anomaly detection, and predictive quality assessment [47,48]. The choice of learning paradigm is guided by data availability and labelling fidelity [49]. Supervised learning frameworks, including convolutional and recurrent neural networks, leverage fully annotated datasets to classify specific defect types or process states [50]. Where annotation is costly or ambiguous, unsupervised methods such as clustering and autoencoders discover hidden patterns and regime transitions without explicit labels. Semi-supervised and self-supervised approaches bridge this gap, leveraging limited labels alongside abundant unlabelled data to enhance model generalization and mitigate the manual labelling burden [51–53]. Collectively, these paradigms construct a hierarchical understanding of process dynamics, wherein models extract discriminative features, infer probabilistic class boundaries, and feed results into decision-level control. In practice, the ML workflow encompasses sensor data acquisition, preprocessing and normalization, feature extraction across temporal, spectral, and spatial

domains, followed by model training, validation, and deployment for real-time inference. Through this structured progression, ML transforms monitoring from passive observation to active process intelligence, an evolution that signifies a decisive leap toward autonomous qualification and robust closed-loop control in metal additive manufacturing [54–57]. Data-driven frameworks now exploit convolutional and interpretable architectures to infer melt-pool states directly from high-frequency AE data [34,40,58,59]. Further, physics-embedded neural networks have shown promise in capturing spatial coupling and heat-transfer relationships [60].

Despite the significant progress enabled by ML, the increasing diversity of machine architectures, process parameters, and sensor configurations in LPBF continues to pose a persistent challenge to model generalization [61]. Models trained within a specific parameter space or sensor configuration often fail when deployed under new conditions because of data drift, that is, systematic shifts in the statistical distribution of sensor signals induced by variations in material composition, optical alignment, chamber acoustics, or melt-pool dynamics [62–64]. This issue is particularly evident in AE –based monitoring, where even subtle changes in build-chamber geometry, powder morphology, or acoustic coupling can alter signal amplitude and spectral content, producing domain-specific artifacts that confound otherwise accurate classifiers. AE signals, while rich in process information, are inherently high-dimensional, non-stationary time series whose frequency content evolves dynamically with melt-pool transitions and plume activity [65,66]. Consequently, models that rely purely on statistical fitting often overfit to these dataset-specific characteristics and fail to generalize across machines or powder batches [63,67]. Conventional data-driven approaches, although effective within fixed experimental domains, lack physical awareness and thus exhibit limited transferability across varying process conditions. Addressing this limitation requires embedding physics-consistent priors into the learning pipeline so that models interpret sensor data not merely as abstract numerical sequences but as measurable signatures of the underlying laser–material and laser–powder interactions. While transfer learning provides a pragmatic route to mitigate domain shifts by reusing knowledge from related process spaces, a deeper question arises: *can a model inherently capture the governing physics of the process without explicit retraining*? A physics-aware framework that learns invariant representations of melt-pool behavior, particularly those encoded in the time–frequency structure of AE signals, should in principle remain generalizable across machines, parameter spaces, and sensor configurations [60]. Cross-machine transfer frameworks, semi-supervised few-shot defect detectors, and transfer-learning-based acoustic monitoring illustrate early progress toward domain-general monitoring [32,68,69]. Prior work has addressed AE domain shift in LPBF by applying deep learning based unsupervised domain adaptation to bridge dissimilar process maps and mitigate distribution offsets across powder and parameter spaces [63]. In the proposed study, parts of the same experimental AE dataset are reused to provide a controlled benchmark for a different question than domain adaptation, namely whether a supervised, physics-aware representation with frequency-localized structure can achieve zero-shot transfer across powder-induced domain shifts without retraining or explicit domain alignment. This shifts the emphasis from distribution matching to physically grounded encodings of melt-pool dynamics. Achieving such interoperability represents a paradigm shift toward physics-informed monitoring representations that capture causal links between process emissions and melting regimes. Ultimately, certifiable repeatability in LPBF requires monitoring frameworks that are both informative and transferable. The intrinsic non-stationarity of AE signals, coupled with variations in machine optics, powder size, and scanning parameters, underscores the necessity for representations that preserve the time–frequency structure of AEs with explicit ties to melt-pool physics [70]. Such physically grounded encodings not only enhance interpretability but also ensure robustness across diverse process windows and hardware platforms, laying the foundation for explainable, domain-invariant monitoring frameworks.

To advance toward this goal, the present study introduces an explainable Learnable Wavelet Scattering (LWS) framework that enhances the information richness of AE signals and preserves robustness under data drift. The proposed approach unites learnable frequency-selective feature extraction with distribution-invariant representation learning and causal frequency attribution, thereby linking spectral descriptors directly to the underlying melt-pool physics. Unlike prior studies that apply generic ML algorithms to AE data without considering sensor bandwidth or the physical semantics of the signal, the LWS framework captures frequency-resolved patterns that remain stable across changing process conditions. This enables a unified, machine-agnostic monitoring model that combines the adaptability of deep learning with the interpretability of physics-based representations offering a viable pathway toward generalizable, explainable, and transferable process monitoring in LPBF. This paper is structured into seven sections. Section 1 introduces the background of the LPBF process and delineates the research gap motivating this work. Section 2 outlines the methodological framework founded on LWS, which enables interpretable, frequency-resolved representation learning from AE signals. Section 3 describes the experimental configuration, including material preparation, process parameters, monitoring instrumentation, and data-acquisition procedures. Section 4 presents the domain-drift characterization, quantifying spectral and distributional deviations in AE features induced by variations in process-parameter spaces. Section 5 details the network architecture, Bayesian hyperparameter optimization, and training strategy underpinning the LWS model. Section 6 reports the inference and explainability results, encompassing model convergence, wavelet-filter evolution, causal frequency attribution, and cross-domain transfer evaluation. Finally, Section 7 consolidates the principal findings, emphasizing the interpretability, robustness, and domain-invariant nature of the proposed framework, and discusses future directions toward adaptive, physics-guided in-situ monitoring for LPBF.

2. Methodology

2.1. Learnable wavelet scattering

Wavelet scattering offers a principled framework for constructing stable, interpretable, and translation-invariant signal representations through cascaded convolutions with analytic wavelets, followed by nonlinear modulus and local averaging operations [71]. This approach unites the rigor of classical time–frequency analysis with the flexibility of neural optimization, enabling the network to preserve physical interpretability while adapting its spectral representation to the data. Each analytic Morlet wavelet [72] acts as a localized band-pass filter and is mathematically expressed as in Eq (1)

$$\psi f_o(t) = e^{2\pi f_o t} e^{-t^2/(2\sigma^2)} \tag{1}$$

where f_0 denotes the center frequency and σ controls the temporal spread and bandwidth. Shorter kernels (small σ) provide excellent temporal resolution but span a broader frequency range, whereas longer kernels capture steady oscillations with higher frequency selectivity. To analyze the signal across multiple resolutions, a family of such wavelets is distributed on a logarithmic frequency axis according to Eq (2)

$$f_{j,q} = 2^{-j/Q} f_{max},\ j = 0, \cdots, J-1,\ q = 1, \cdots, Q, \tag{2}$$

where J specifies the number of octaves and Q the number of wavelets per octave, jointly defining the depth and density of the spectral decomposition. These two parameters determine how densely and how deeply the frequency spectrum is partitioned, balancing spectral coverage with computational cost. The concept is illustrated in Fig. 1, which shows the normalized magnitude responses of a Morlet filter bank with $J = 3$ octaves and $Q = 3$ wavelets per octave. Each color represents

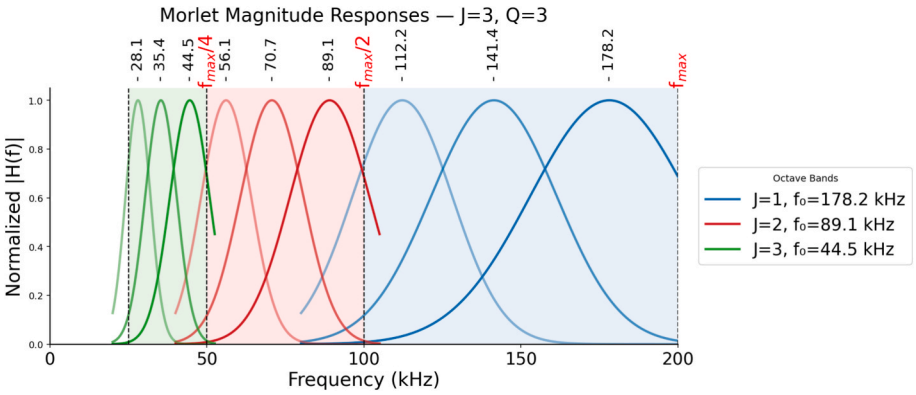

Fig. 1. Normalized Morlet magnitude responses for $J = 3$ octaves and $Q = 3$ wavelets per octave. The filters are distributed logarithmically across frequency, illustrating how each octave band (green, red, blue) covers progressively lower frequency ranges with distinct bandwidths. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

a distinct octave group, with the corresponding center frequency marked by vertical dashed lines. The highest octave (blue) covers the upper end of the frequency range near $f_{\max}$, while subsequent octaves (red and green) progressively capture mid- and low-frequency components. This logarithmic tiling ensures that lower frequencies are analyzed with narrower relative bandwidths (high spectral precision), whereas higher frequencies are represented with broader filters (high temporal precision). The parameter pair (J, Q) thus defines the overall spectral geometry of the wavelet bank, determining how information is distributed across scales.

After decomposing the signal across logarithmically spaced frequency bands, each wavelet acts as a localized "listener" that responds only to oscillations within its own spectral neighborhood. Through this multi-resolution decomposition, a signal $x(t)$ is transformed into a hierarchy of sub-band representations that jointly encode its temporal and spectral characteristics. The complex convolution of the signal with each wavelet is expressed as in Eq (3),

$$\mathscr{Y}_{j,q}(t) = \left(x^* \psi_{j,q}(t)\right)(t) \tag{3}$$

where $\psi_{j,q}(t)$ denotes the Morlet kernel corresponding to a specific scale j and sub-band index q.

Fig. 2 illustrates how the kernel length k_{size} influences the time–frequency resolution of the Morlet wavelet for a fixed center frequency of 10 kHz. In the time domain (left column), increasing k_{size} expands the duration of the wavelet, allowing it to "listen" over a longer time window. In the frequency domain (middle column), this produces a narrower bandwidth (smaller –3 dB width), indicating improved frequency selectivity. Conversely, shorter kernels offer sharper temporal localization but broader spectral coverage. The right column of Fig. 2 visualizes these kernels mapped over a 12 ms window in non-sliding setting, highlighting how the oscillatory structure becomes more slowly varying as k_{size} increases. This trade-off between temporal precision and

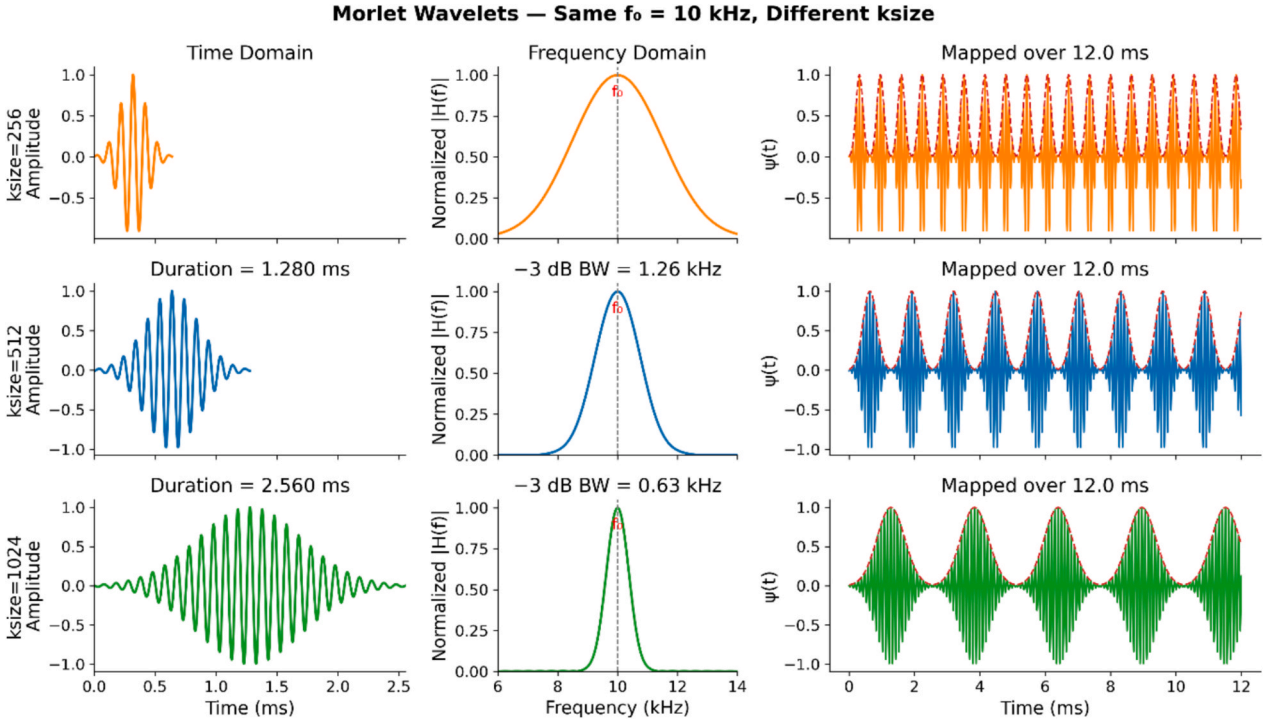

Fig. 2. Effect of kernel size (k_{size}) on the Morlet wavelet at fixed center frequency $f_0 = 10$kHz. Left: time-domain amplitude showing longer temporal support for larger kernels. Middle: frequency-domain responses illustrating narrowing bandwidths. Right: Non-overlapping kernels mapped over a 12 ms window, emphasizing the trade-off between time and frequency resolution that governs wavelet design.

frequency resolution lies at the core of wavelet analysis: short kernels are ideal for detecting transient or rapidly changing components, whereas longer kernels capture smoother, more stationary oscillations. In learnable implementations, k_{size} becomes a tunable hyperparameter controlling the temporal support of each wavelet filter, directly influencing how sensitively the model responds to short- or long-duration signal patterns. Together with the octave depth J, number of wavelets per octave Q, and averaging window avg_{size}, the kernel length defines the overall balance between localization and stability in the learned time–frequency representation.

The modulus of this response is then taken to extract its instantaneous amplitude. This operation converts the oscillatory convolution output into a measure of local energy, emphasizing how strongly the signal oscillates around the wavelet's center frequency while discarding phase information as shown in Fig. 3. It provides a phase-invariant descriptor of signal intensity that varies smoothly over time and reveals transient structures hidden in raw oscillations. The modulus is defined as in Eq (4)

$$E_{j,q}(t) = \left| \mathscr{Y}_{j,q}(t) \right| \tag{4}$$

where $E_{j,q}(t)$ denotes the instantaneous amplitude envelope of the signal filtered through the wavelet $\psi_{j,q}(t)$. The first panel in Fig. 3 shows a non-stationary input signal composed of three Gaussian bursts centered around 10 kHz. The second panel displays the real part of the Morlet wavelet ($k_{\text{size}} = 256, f_0 = 10.0$ kHz) used as the analyzing kernel. When the signal is convolved with this wavelet, as shown in the third panel, the resulting response highlights temporal segments where oscillatory energy near 10 kHz is most dominant. Applying the modulus operation yields the envelope shown in the bottom panel, which traces the instantaneous energy evolution of the signal. The two distinct peaks correspond precisely to the temporal locations of the bursts, demonstrating how the wavelets localize energy in both time and frequency while producing smooth, interpretable amplitude profiles.

This modulus-based envelope extraction represents a key stage in wavelet scattering: it bridges the transition from oscillatory components to energy-based representations, enabling phase-invariant, temporally resolved characterization of signals. It provides a physically meaningful intermediate feature that captures the strength and timing of dynamic events before subsequent averaging introduces further stability and invariance. To achieve robustness against small temporal shifts and local deformations, the extracted envelopes are subsequently smoothed using a low pass averaging filter $\phi_{\text{avg}}(t)$. This operation aggregates short-term variations into slowly varying trends, producing coefficients that are both translation-invariant and spectrally localized. The averaging stage is mathematically expressed as in Eq (5),

$$S_{j,q}(t) = E_{j,q}(t) * \phi_{\text{avg}}(t). \tag{5}$$

where $\phi_{\text{avg}}(t)$ is a smoothing kernel, typically flat or Hann-shaped. Its length, defined by avg_{ksize}, controls how much temporal information is averaged. Smaller avg_{ksize} values preserve fine temporal details but are more sensitive to short-term fluctuations, whereas larger values enhance stability by averaging over longer durations. Fig. 4 demonstrates this effect. The upper panel shows the instantaneous envelope $|y(t)|$ obtained from the modulus stage, which oscillates rapidly due to fine-scale variations in signal energy. The lower panel displays the results after applying moving-average smoothing with progressively larger window lengths. As the averaging window increases, the envelope becomes smoother and more stable, suppressing high-frequency fluctuations while retaining the primary energy peaks corresponding to the main signal events. This transformation yields a compact and robust representation that emphasizes structural trends rather than transient noise.

Together, these three operations convolution, magnitude extraction, and averaging transform a raw signal into a structured hierarchy of time–frequency descriptors that are both localized and invariant [73]. The parameters J and Q define the spectral tiling, while k_{size} and avg_{ksize} set the temporal precision and degree of invariance. The parameters jointly determine the resolution and stability of this representation. When balanced appropriately, these parameters yield compact, interpretable features that preserve the essential dynamics of the signal while filtering out irrelevant variability. In summary, LWS extends the classical scattering framework by replacing fixed analytical filters with trainable parameters that adapt to data through end-to-end optimization. The core difference in LWS is that elements such as filter parameters or thresholding operators inside the scattering network become

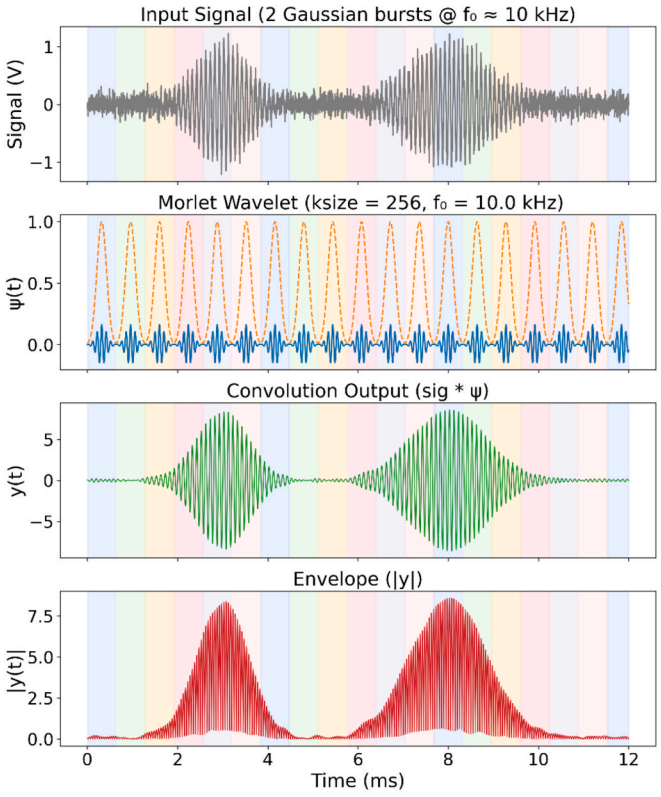

Fig. 3. Illustration of convolution and envelope extraction using a Morlet wavelet. (Top) Input signal with two Gaussian bursts at $f_0 \approx 10$ kHz. (Second) Real part of the Morlet wavelet ($k_{\text{size}} = 256, f_0 = 10$ kHz). (Third) Convolution output showing temporal localization of oscillatory activity. (Bottom) Envelope $|y(t)|$ depicting instantaneous energy variations across time.

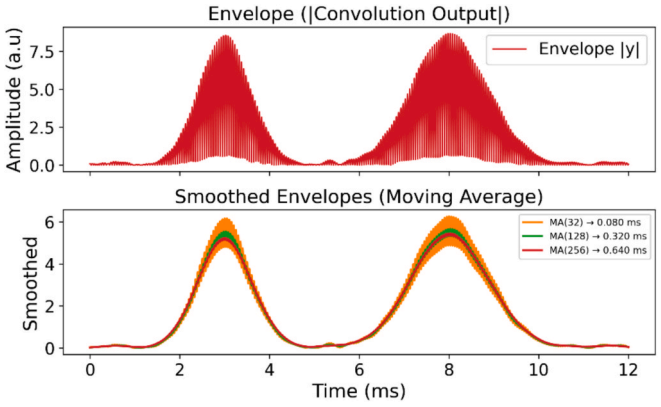

Fig. 4. Averaging stage of the wavelet feature extraction process. (Top) Instantaneous envelope $|y(t)|$ obtained from the modulus stage. (Bottom) Smoothed envelopes obtained using different averaging window lengths, showing progressive suppression of high-frequency fluctuations and increased temporal stability.

learnable, allowing optimization for specific tasks via backpropagation. This flexibility enables the model to discover frequency partitions that are maximally informative for the problem while retaining the interpretability and mathematical rigor of traditional wavelet theory. By combining the stability and invariance of handcrafted transforms with the adaptability of neural architectures, LWS achieves a principled balance between physics-based structure and data-driven learning.

2.2. Methodology

The proposed framework processes AE signals acquired during LPBF process to generate frequency-resolved representations that are both learnable and physically interpretable. In LPBF, process emissions comprising transient acoustic and optical signals encode rich information about melt-pool dynamics, defect formation, and stability of the laser–material interaction. Harnessing these signals requires representations that are sensitive to localized temporal variations yet robust to stochastic noise and process-space variability. At the core of the monitoring framework lies the LWS front-end, which extends the traditional scattering transformation by allowing wavelet parameters to adapt through backpropagation. The module, implemented as a LearnableMorletBank1D, parameterizes a family of Morlet filters with trainable logarithmic center frequencies (log$f$) and bandwidths (logbw), initialized across a multi-resolution grid defined by the parameters J and Q. These parameters determine the spectral coverage and resolution across the AE bandwidth, typically spanning 0–200 kHz which is the sensor range. During training, the network learns to reposition its wavelet centers and adjust bandwidths, emphasizing frequency regions that most effectively capture process-related variations such as *LoF* events, *conduction*-mode stability, or *keyhole* oscillations. In this way, the LWS transforms raw acoustic signals into structured spectral encodings that adaptively highlight the most informative bands for monitoring the laser–material interaction. The complete pipeline of the proposed monitoring framework is illustrated in Fig. 5, which summarizes the sequence from wavelet initialization and feature extraction to optimization, classification, and explainability. Each normalized signal $x(t)$ is convolved with the trainable Morlet filters to obtain subband responses $y_{j,q}(t)$, whose magnitudes $|y_{j,q}(t)|$ represent localized energy envelopes corresponding to distinct frequency bands. These envelopes are further smoothed using Hann low-pass filters of kernel length avg_{ksize}, ensuring temporal stability while preserving the underlying oscillatory structure. The resulting coefficients constitute first-order scattering features, which capture the sub-band energy distribution of AE signals with built-in translation invariance. These scattering coefficients are concatenated with a global average pooling term to form a compact feature vector representing the signal's hierarchical frequency structure. The resulting representation is frequency-aware, physically interpretable, and directly linked to underlying melt-pool dynamics, enabling visualization through wavelet magnitude responses and Δ-spectral maps before and after training. To organize and compress the extracted features, the LWS outputs are projected into a latent embedding space using a batch-normalized projection head with Rectified Linear Unit (ReLU) activation. This stage reduces the high-dimensional scattering representation to 128-dimensional embedding, preserving class separability while minimizing redundant variability across LPBF build domains.

A lightweight linear classifier then maps the latent embeddings to melt-pool regime labels—*Lack of Fusion (LoF)*, *conduction*, and *keyhole*—corresponding to distinct modes of LPBF process stability. The entire network, including the learnable wavelet bank, projection head, and classifier, is trained end-to-end using the Adam optimizer with a cosine-annealed learning rate (lr) schedule, ensuring stable convergence while preventing overfitting to specific emission patterns. The main hyperparameters— lr, J, Q, k_{size}, and avg_{ksize}—are tuned through Bayesian Optimization (BO) using the Optuna framework with a Tree-structured Parzen Estimator (TPE) sampler [74]. Early pruning strategies based on the validation macro-F1 score ensure computational efficiency and focus the search on high-performing regions of parameter space. Overall, the proposed LWS-based monitoring framework transforms raw AE waveforms into a hierarchically organized, physically grounded representation that captures the fundamental time–frequency structure of LPBF process dynamics. By learning from real process emissions rather than relying solely on predefined analytical wavelets, the network achieves a balance between interpretability and

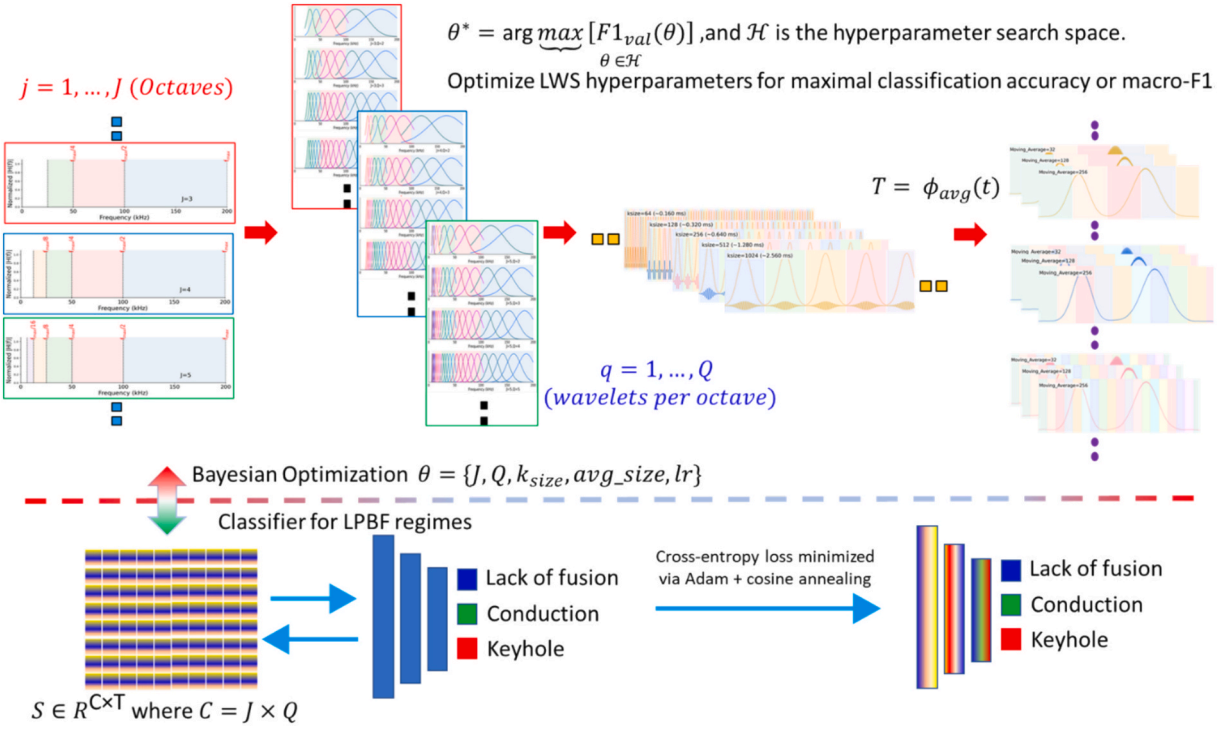

Fig. 5. Workflow of the LWS front-end and feature pipeline for LPBF process monitoring.

adaptability—providing an explainable, domain-transferable foundation for in-situ monitoring of AM processes.

3. Experimental setup, materials and data acquisition

3.1. Experimental setup and materials

All experiments were performed on a SISMA MySint 100 LPBF system equipped with a continuous-wave fiber laser (1070 nm, max. 200 W) producing a Gaussian beam with a 55 μm $1/e^2$ spot diameter. Gas-atomized 316 L stainless steel powder (Oerlikon Metco) was classified using a 325-mesh sieve (45 μm aperture) into a coarse fraction (P1, >45 μm) and a fine fraction (P2, <45 μm) to induce a controlled, physically grounded domain shift. The powder characteristics have been described in our previous studies [59,70]; for completeness, the corresponding particle size statistics (D10, D50, and D90) are reported again in Table 1. This particle-size shift modifies powder-bed packing and effective heat transport, perturbing melt-pool wetting, stability, and spatter activity, which in turn alters the amplitude, spectral content, and temporal structure of airborne AE. Such powder-induced shifts can degrade conventional CNN performance without retraining or explicit domain adaptation. Here, we test whether a physics-aware Learnable Wavelet Scattering (LWS) front-end enables zero-shot transfer across powder domains via frequency-localized, interpretable representations. The experimental platform, sensor mounting, and trigger-synchronized acquisition pipeline are unchanged from our prior publications [63]; the same validated LPBF framework and AE dataset are reused solely to address this distinct objective, and only essential details are summarized in this section.

3.2. Process parameters

The objective of this study is to evaluate a lightweight, physically interpretable, and transferable framework for classifying LPBF melt-pool regimes (*LoF, conduction, keyhole*) and enabling causal, frequency-resolved interpretation of AE under process drift. Two datasets were collected using identical procedures but different powder size distributions: D_1 employed the coarse fraction (>45 μm), while D_2 employed the fine fraction (<45 μm). The three regimes were produced by systematically varying laser power and scan speed within each powder domain, as summarized in Table 2. For the same nominal regime, the laser power was intentionally offset between the two powder distributions (while maintaining the corresponding scan speed) to introduce a measurable yet physically consistent cross-domain shift. Further experimental details are provided in our prior publications [63].

The formation of the three distinct melt-pool regimes was validated through optical microscopy on cross-sections taken perpendicular to the scanning direction. These samples were sectioned, ground, and polished according to metallographic standards. Characteristic differences in porosity, void morphology, and relative density, shown in Fig. 6 (see [63] for full metallographic details), confirmed that the selected process parameters successfully produced the intended regimes for both powder distributions.

3.3. Monitoring setup, data acquisition pipeline and pre-processing

A CM16/CMPA airborne AE sensor (Avisoft Bioacoustics), featuring a flat frequency response from 2 to 200 kHz, was integrated into the LPBF chamber to capture acoustic-wave activity generated during laser–material interaction (Fig. 7). An off-axis optical trigger was used to time-align AE with laser exposure: optical emissions were collected using a fixed-focus collimator (F810SMA-1064, Thorlabs) coupled via a 550 μm fiber to a photodiode (PDA20CS2, Thorlabs; 800–1700 nm), which detects near-infrared back-reflection and melt-pool radiation. The photodiode gain was set to saturate at 5 V during active irradiation, and each saturation interval (one scan vector) served as the temporal reference for extracting the corresponding AE segment.

Both optical and AE signals were digitized using a PCIe DAQ card (Advantech 1840) at ± 5 V and 400 kHz under custom C# control, with acquisition triggered when the photodiode exceeded 0.5 V. AE tracks were segmented into fixed 12.5 ms windows (5,000 samples) and filtered offline using a fourth-order Butterworth low-pass filter with a 150 kHz cutoff to suppress high-frequency electronic noise and match the effective sensor bandwidth. This acquisition and preprocessing pipeline is identical to that already reported in our prior work [63]. and was kept unchanged across both datasets to ensure that observed spectral differences reflect powder- and process-induced variations rather than inconsistencies in signal handling. Key dataset parameters are summarized in Table 3.

Table 1
Statistics of mean cumulative particle size distribution in terms of D10, D50 and D90.

Category	*D10*	*D50*	*D90*
Particle distribution above 45 μm [P1]	30.9 μm	42.3 μm	57.9 μm
Particle distribution below 45 μm [P2]	19.1 μm	29.6 μm	45.3 μm

Table 2
LPBF process parameters on the two different powder distributions.

	Laser Power (W)	*Scan speed (mm/s)*	Particle distribution	Dataset name
LoF	110	800	Particle distribution above 45 μm [P1]	D_1
Conduction	160	400		
Keyhole	160	75		
LoF	90	800	Particle distribution below 45 μm [P2]	D_2
Conduction	140	400		
Keyhole	180	75		

4. Domain drift characterization

To quantify the spectral characteristics of the AE signals and assess the degree of drift between the two datasets (D_1: coarse > 45 μm; D_2: fine < 45 μm), each 12.5 ms AE segment was transformed into the frequency domain using the Welch power-spectral-density (PSD) estimator. The resulting PSD for each signal was normalized to remove amplitude bias, enabling scale-independent comparison of spectral structures. From each normalized spectrum, two key frequency-domain features were extracted to describe the energy distribution and spectral complexity of the AE signals. The spectral centroid (f_c) represents the "center of mass" of the signal's spectral energy distribution and provides a measure of where most of the acoustic power is concentrated along the frequency axis as in Eq (6)

$$f_c = \sum_{i=1}^{M} f_i P(f_i) \tag{6}$$

A higher f_c indicates that the signal energy is concentrated at higher frequencies, typically associated with rapid melt-pool oscillations and fine-scale process dynamics. The spectral entropy (H_s) quantifies the degree of disorder or flatness in the frequency spectrum, reflecting how evenly the energy is distributed across frequencies as in Eq (7)

$$H_s = -\sum_{i=1}^{M} P(f_i)\log P(f_i) \tag{7}$$

Higher H_s values correspond to broader and more stochastic spectral content, while lower values indicate tonal or periodic signals. Fig. 8 shows the distribution of f_c and H_s across the three melt-pool regimes—*LoF*, *conduction*, and *keyhole*—for both D_1 and D_2, combining box and scatter representations. The annotated p-values are derived from two-sample Kolmogorov–Smirnov (KS) tests [75] comparing D_1 and D_2 within each regime.

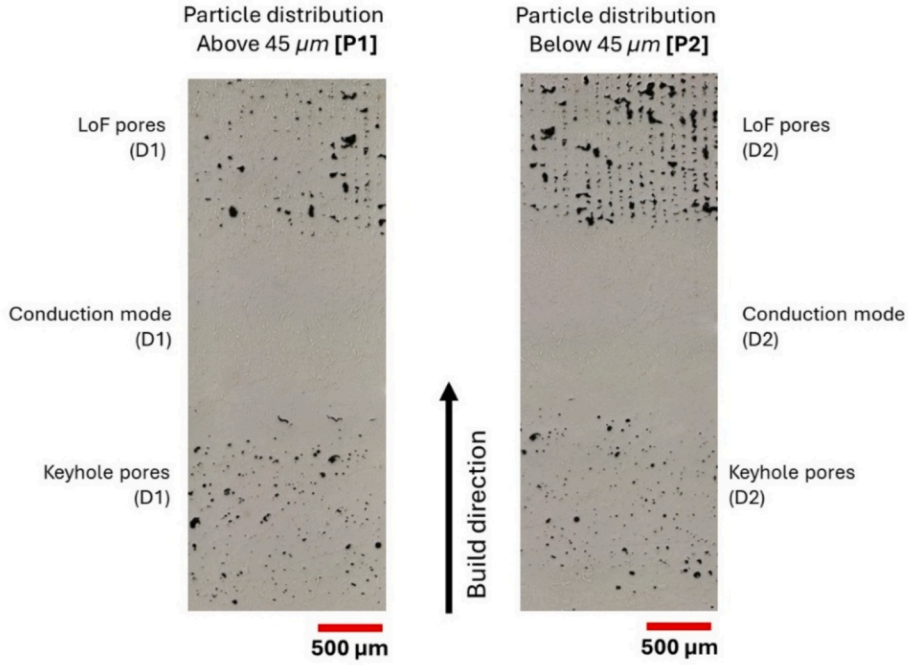

Fig. 6. Optical micrograph for each regime across two 316 L powder distributions *P1* and *P2*.

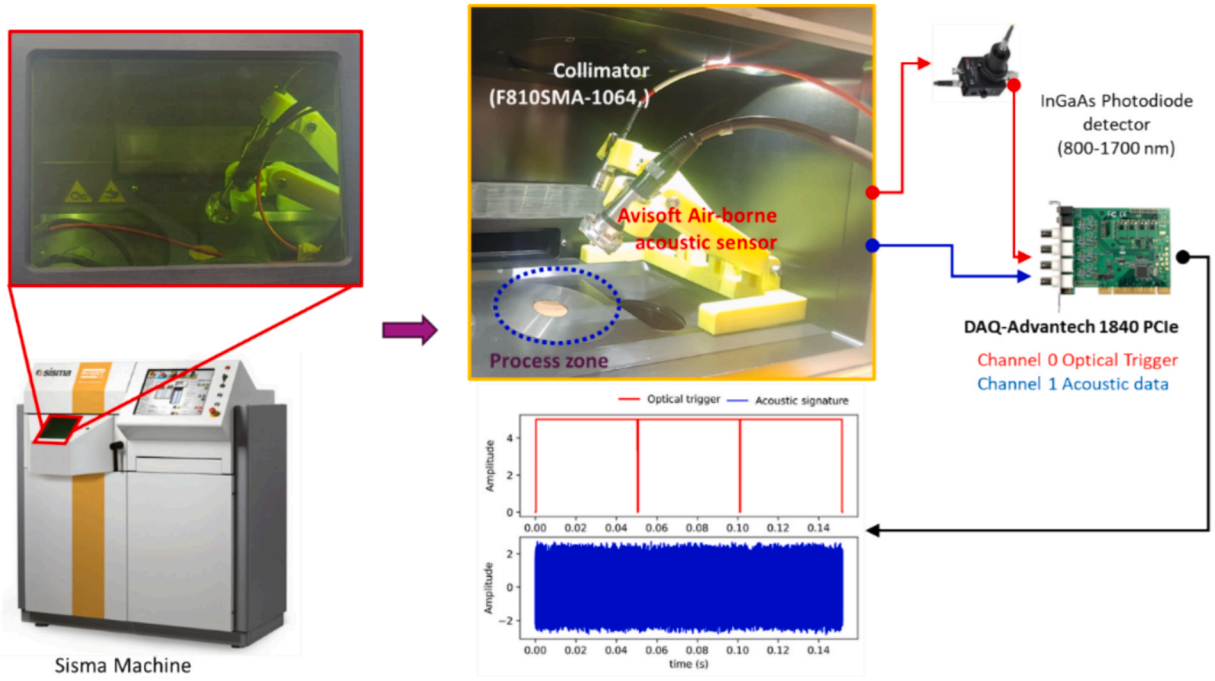

Fig. 7. Sisma MYSINT 100 augmented with air-borne acoustic sensor and photodiode trigger for LPBF process.

Table 3
Composition of datasets D_1 and D_2 formed from AE segments (12.5 ms, 5000 samples) captured under particle distributions P1 and P2 for zero-shot transfer evaluation.

Particle distribution	*LoF pores*	*Conduction*	*Keyhole*
Dataset [D_1]	10,450	10,450	10,450
Dataset [D_2]	10,100	10,100	10,100

Across all regimes, AE signals from the fine-powder domain (D_2) exhibit systematically higher centroid frequencies and entropies compared to the coarse-powder domain (D_1), with all differences being statistically significant ($p < 1 \times 10^{-4}$). The upward shift in spectral centroid indicates enhanced high-frequency activity in D_2, consistent with finer particle morphology, increased absorptivity, and more rapid melt-pool oscillations, while the elevated entropy values signify greater spectral complexity and broader energy distribution. The *keyhole* regime displays the largest spread in both descriptors, reflecting its inherently unstable and broadband nature associated with vapor-plume pulsations and melt-pool collapse events. These consistent shifts in the distributions of spectral centroid and entropy confirm that the AE feature space

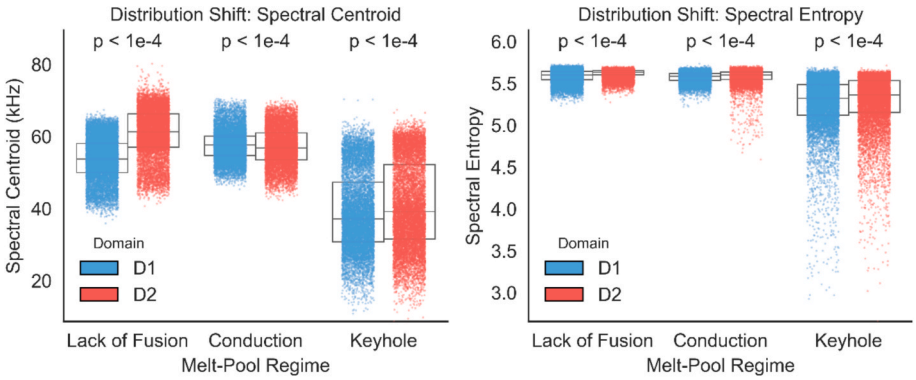

Fig. 8. Distributional shift of spectral centroid and entropy between D_1 (coarse) and D_2 (fine) domains.

undergoes a systematic domain drift between D_1 and D_2, arising from physical variations in powder morphology and process-energy coupling rather than from labeling or sampling bias. Consequently, a model trained solely on D_1 would encounter a shifted spectral landscape when evaluated on D_2, motivating the development of the LWS framework described in section 5 to learn domain-invariant, physics-consistent representations that generalize robustly across such process drifts.

5. Training architecture

5.1. Input representation

The proposed model operates directly on one-dimensional AE waveforms acquired during the LPBF process. Each waveform corresponds to a single laser scan vector, segmented from the continuous AE stream such that each segment contains $T = 5000 samples$, representing the full acoustic response over a temporal window of 0.0125 s (12.5 ms) at a sampling rate of 400 kHz. Formally, each segment is expressed as a discrete time series as in Eq (8).

$$x = \{x_t | t = 1, 2, \cdots, T\}, x_t \in R \tag{8}$$

where x_t denotes the instantaneous acoustic amplitude at time index t. To eliminate baseline drift and amplitude bias, each AE segment is normalized on a per-instance basis as in Eq (9).

$$\widetilde{x}_t = \frac{x_t - \mu_x}{\sigma_x}, \mu_x = \frac{1}{T}\sum_{t=1}^{T} x_t, \sigma_x = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(x_t - \mu_x)^2} \tag{9}$$

This zero-mean, unit-variance normalization ensures that subsequent convolutional operations emphasize the relative temporal–spectral structure of the signal rather than absolute energy magnitudes affected by sensor coupling or gain variations. The normalized signal $\widetilde{x} \in \mathbb{R}^{1\times T}$ is reshaped into a tensor $\mathrm{X} = \mathrm{reshape}(\widetilde{x}) \in \mathbb{R}^{1\times 5000}$,

preserving temporal continuity for one-dimensional convolutional filtering. The dataset $\mathscr{D} = \{(\mathrm{X}_i, y_i)\}_{i=1}^{N}$ is randomly partitioned into 70% training and 30% testing subsets, maintaining balanced class distributions across the three melting regimes $y_i \in \{0,1,2\} \Rightarrow \{LoF, conduction, keyhole\}$. This stratified sampling ensures unbiased evaluation of the model's performance and generalization across distinct process states.

5.2. Network architecture, latent embedding and classifier

The overall architecture consists of three main stages: a LWS front-end for physics-informed feature encoding, a projection head for latent embedding, and a linear classifier for melt-pool regime identification. The LWS front-end, previously described in Section 2.1, performs hierarchical time–frequency decomposition using a trainable Morlet filter bank whose center frequencies and bandwidths are optimized through gradient descent. For each scale j and sub-band q, the convolutional response of the normalized AE signal $\widetilde{x}(t)$ is given by Eq (10),

$$y_{j,q}(t) = (\widetilde{x} * \psi_{j,q})(t), S_{j,q}(t) = \left|y_{j,q}(t)\right| * \phi_{\mathrm{avg}}(t), \tag{10}$$

where $\psi_{j,q}(t)$ denotes the learnable Morlet kernel and $\phi_{\mathrm{avg}}(t)$ is a Hann-shaped averaging window of length avg_{ksize}. The resulting coefficients $S_{j,q}(t)$ form a structured feature tensor $\mathrm{S} \in \mathbb{R}^{J\times Q\times T'}$ that captures the multiscale energy distribution across the acoustic spectrum. These scattering features are vectorized and passed to a projection head that performs dimensionality reduction and normalization according to Eq (11),

$$\mathrm{z} = \mathrm{ReLU}\ \left(\mathrm{Batch\ Norm}\left(\mathrm{W}_p\,\mathrm{vec}(\mathrm{S}) + \mathrm{b}_p\right)\right), \widehat{\mathrm{z}} = \frac{\mathrm{z}}{\|\mathrm{z}\|_2}, \tag{11}$$

where $\mathrm{z} \in \mathbb{R}^{128}$ represents the latent embedding, Batch Norm$(\cdot)$ denotes batch normalization, and W_p and b_p are learnable parameters. This embedding space compresses the high-dimensional scattering representation into a 128-dimensional manifold that retains discriminative spectral features while mitigating overfitting. Finally, a linear classifier maps the normalized embedding $\widehat{\mathrm{z}}$ to the posterior probability distribution over the three melt-pool regimes— *LoF*, *conduction*, and *keyhole* via a temperature-scaled softmax function as in Eq (12)

$$\mathrm{p} = \mathrm{softmax}\left(\frac{\mathrm{W}_c\,\widehat{\mathrm{z}} + \mathrm{b}_c}{\tau}\right), \tag{12}$$

where τ is the temperature parameter. The complete network, comprising the LWS front-end, projection head, and classifier, is trained end-to-end using the Adam optimizer with a cosine-annealed learning-rate schedule, ensuring stable convergence and physically interpretable feature learning. This design harmonizes the stability and interpretability of wavelet-based analysis with the adaptability of deep learning, yielding robust spectral representations that encode melt-pool dynamics with high discriminative fidelity.

5.3. Bayesian optimization and parameter selection

To identify the optimal architectural and training hyperparameters, a BO framework was implemented using the Optuna library with a TPE sampler. The search space encompassed the $lr \in [1\times 10^{-4}, 3\times 10^{-3}]$ (log-scaled), the number of scales $J \in [4,7]$, the number of wavelets per octave $Q \in [1, 4]$, the wavelet kernel length $k_{\mathrm{size}} \in [129, 513]$(odd-enforced), and the averaging kernel length $avg_{\mathrm{ksize}} \in [65,513]$. Each trial trained the LWS-based network for 40–50 epochs using fixed optimizer settings and batch size of 64, and the validation macro-F1 score served as the objective function $\mathscr{L}_{\mathrm{BO}}$ as in Eq (13), (Fig. 9)

$$\mathscr{L}_{\mathrm{BO}} = 1 - \mathrm{F1}_{\mathrm{val}}. \tag{13}$$

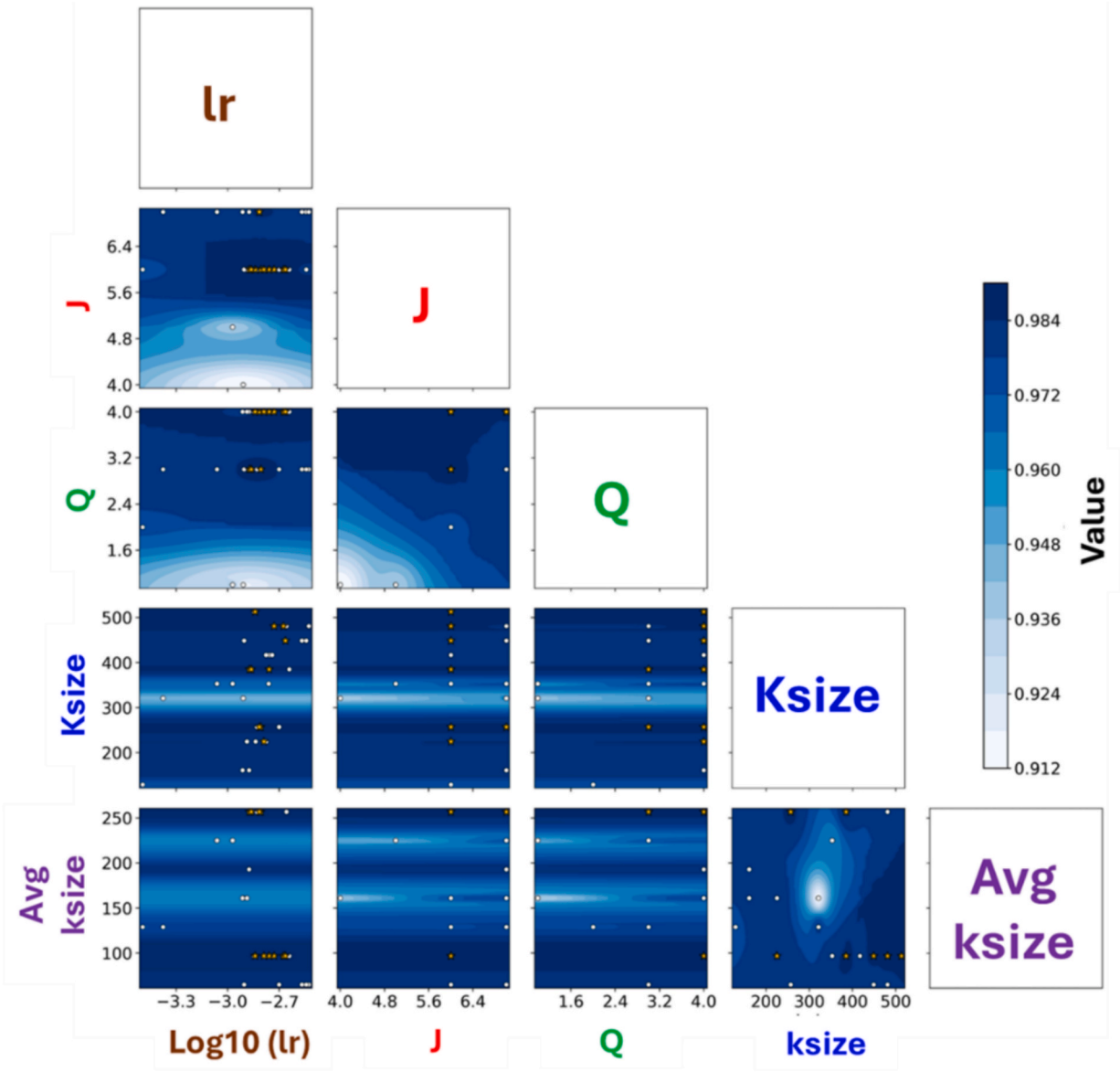

Fig. 9. Pairwise parameter sensitivity maps from BO.

Early-stopping based on median pruning was employed to terminate underperforming trials and expedite convergence toward high-performing regions of the hyperparameter space. The Fig. 10 illustrates the distribution of validation macro-F1 scores across combinations of the five principal hyperparameters—*lr*, *J*, *Q*, k_{size}, and averaging kernel size (avg_{ksize}).

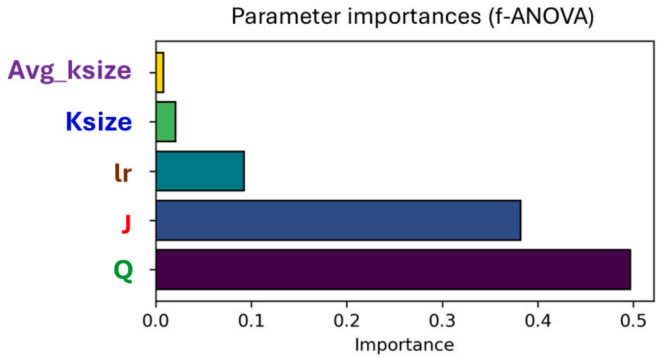

Fig. 10. Relative parameter importances from f-ANOVA analysis during BO.

Each cell represents the interpolated performance landscape derived from BO trials, where darker regions correspond to higher macro-F1 values. A well-defined and stable optimum emerges near $J = 6$ and $Q = 3$, confirming that the spectral decomposition depth and sub-band density exert the most significant influence on discriminative performance. In contrast, the nearly horizontal gradients observed along k_{size} and avg_{ksize} indicate low sensitivity to moderate kernel variations, suggesting that the learned representations remain robust to small temporal-resolution changes within the specified parameter range. After conducting 75 completed trials, F-test within Analysis of Variance (f-ANOVA) decomposition was performed to further quantify the contribution of each hyperparameter to the overall performance variance (validation macro-F1). The analysis confirms that *Q* and *J* dominate the sensitivity landscape, while the *lr* exerts a moderate effect as shown in Fig. 10. In contrast, wavelet kernel size (k_{size}) and averaging window length (avg_{ksize}) contribute minimally, indicating that spectral resolution and decomposition depth primarily govern the discriminative capability of the LWS front-end. The optimized configuration identified through BO was $lr^* = 1.56 \times 10^{-3}, J^* = 6, Q^* = 3, k^*_{size} = 257, avg_{ksize}{}^* = 257$.

This configuration provided a balanced trade-off between spectral coverage and model compactness, resulting in smooth convergence behavior and high classification fidelity. These hyperparameters were subsequently fixed for all training, inference, and zero-shot transfer learning experiments to ensure reproducibility and consistent cross-domain evaluation throughout the study.

5.4. Training strategy

Once the optimal set of hyperparameters was determined through BO, the network training and evaluation were performed on an NVIDIA RTX 4000 GPU using the PyTorch 2.0 environment [76]. The model was trained in a supervised learning setting with a cross-entropy loss, employing the Adam optimization algorithm in conjunction with a cosine-annealed learning-rate scheduler to maintain gradual and stable convergence while avoiding oscillatory gradients. Each experiment consisted of 100 epochs, a batch size of 64, and a weight-decay coefficient of 1×10^{-4} to regulate parameter growth and mitigate overfitting, ensuring a smooth optimization trajectory. Throughout training, the framework continuously tracked the mean and variance of batch-level losses together with epoch-wise accuracy and macro-F1 statistics for both training and validation datasets. These diagnostic measures provided quantitative insight into convergence dynamics and generalization capability. Within the learnable Morlet wavelet bank, the center frequencies and bandwidths of individual filters were adaptively refined through backpropagation, enabling the network to emphasize frequency bands most representative of distinct melt-pool behaviors. At regular checkpoints, the evolving wavelet parameters were stored and examined through frequency–response trajectories and Δ-spectral maps, illustrating the progressive specialization of the spectral filters during learning.

Model performance was further validated using a suite of complementary analyses, including confusion-matrix evaluation, *t*-Distributed Stochastic Neighbor Embedding (*t*-SNE) visualizations of latent-space separation, and class-wise precision and F-score measurements. The optimal checkpoint, identified by the highest validation macro-F1 score, was used in subsequent analyses focused on explainability, frequency attribution, and zero-shot transfer learning. Fig. 11 illustrates the complete workflow of the LWS framework used for LPBF process monitoring. The model converts raw AE signals into physically grounded, data-adaptive spectral representations through a hierarchy of trainable components. The LWS front-end employs a parameterized Morlet filter bank whose spectral parameters are iteratively optimized to capture the most informative frequency regions. The resulting scattering coefficients are processed by a projection head composed of batch-normalized linear layers with learnable weights and biases, which compress the high-dimensional spectral features into a 128-dimensional latent manifold that enhances separability among melt-pool regimes. The classifier head, implemented as a temperature-scaled linear layer, transforms these latent embeddings into class logits representing the *LoF*, *conduction*, and *keyhole* modes. All parameters—including those of the LWS filters, projection head, and classifier—are jointly optimized end-to-end via backpropagation of the cross-entropy loss, ensuring consistent gradient flow and unified adaptation across the network.

The full configuration of the training setup is summarized in Table 4. The model was trained under a supervised classification framework using the Adam optimizer with a *lr* of 1.56×10^{-3} for 100 epochs. A batch size of 64 and a dropout rate of 10% were applied to enhance generalization, while data were reshuffled each epoch and split into 70% training and 30% testing subsets. The final network comprised 7991 trainable parameters, offering a compact yet expressive representation

Table 4
Training parameters on the LWS model trained after BO.

Training parameters	LWS-CNN
Type of analysis	Classification
Solver name	Adam
Learning rate	1.56×10^{-3}
Training epochs	100
Batch size	64
Dropout	10%
Loss	Cross Entropy
Shuffle	Every-epoch
Training set	70%
Testing set	30%
Trainable parameter	7991

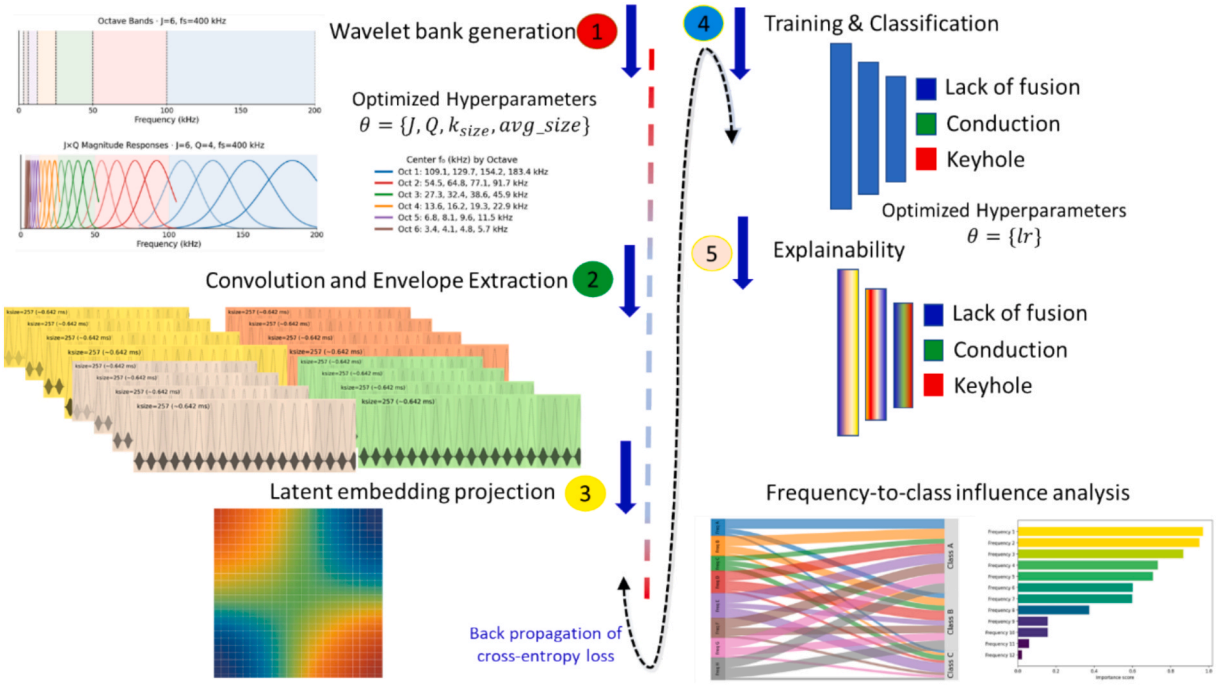

Fig. 11. Workflow of the LWS framework for LPBF based on AE monitoring.

of AE dynamics. This training strategy resulted in stable convergence, high classification accuracy, and physically interpretable filter evolution, establishing a robust foundation for the inference and zero-shot transfer analyses described in the following section.

6. Inference and explainable frequency attribution

6.1. Training convergence and model stability

The LWS network demonstrated stable convergence and strong generalization across 100 training epochs. As shown in Fig. 12, the training loss exhibited a rapid decline during the initial epochs and stabilized near the twentieth epoch, indicating efficient gradient propagation and effective feature learning. The cosine-annealed learning-rate schedule facilitated gradual optimization, preventing oscillatory updates and ensuring smooth convergence toward the minimum loss. The corresponding accuracy profile revealed that both training and validation accuracies exceeded 95% after approximately 60 epochs and progressively saturated near 98–99%. This trend reflects the robustness of the learned representations and the absence of overfitting. The narrow separation between the training and validation curves highlights the strong inductive bias imparted by the learnable Morlet filter bank, which constrains learning to physically meaningful spectral domains and enhances stability. The optimized configuration—$J = 6$, $Q = 3$, $k_{\text{size}} = 257$, and $avg_{\text{ksize}} = 257$ yielded a smooth optimization trajectory and consistent classification fidelity across all three melt-pool regimes.

During inference, the trained model maintained excellent generalization on the held-out test dataset. The normalized confusion matrix in Table 5 demonstrates well-balanced performance across all classes, achieving accuracies of 98% for *LoF*, 97% for *conduction*, and 99% for *keyhole*. The minimal off-diagonal components confirm that the learned spectral embeddings effectively discriminate among *LoF*, *conduction*, and *keyhole* dominated LPBF regimes.

To further examine the organization of learned features, the 128-dimensional latent embeddings from the projection head were projected into two dimensions using *t*-SNE, as shown in Fig. 13. Distinct, non-overlapping clusters corresponding to the *LoF*, *conduction*, and *keyhole* regimes are clearly observed. The tight intra-class cohesion and wide inter-class separation indicate that the latent manifold preserves well-defined decision boundaries and captures discriminative, process-dependent frequency patterns. These findings confirm that the LWS framework not only achieves high predictive accuracy but also provides interpretable and physically consistent representations of melt-pool acoustic dynamics.

6.2. Cross-domain transfer evaluation

To evaluate the robustness and domain invariance of the proposed framework, the LWS model trained on the reference dataset (D_1) was directly applied to an unseen target domain (D_2). This zero-shot transfer configuration assesses the model's ability to generalize under variations in material composition, powder particle size, and signal distribution. During inference, the trained weights obtained from D_1 were frozen, and the same wavelet parameterization was maintained ($J = 6$, $Q = 3$, $k_{\text{size}} = 257$, and $avg_{\text{ksize}} = 257$). Despite the distributional shift between domains, the model maintained excellent predictive performance, achieving a retention rate exceeding 96% relative to its source-domain accuracy. This result demonstrates that the learned representations are largely domain-invariant and encode physically consistent acoustic features rather than dataset-specific artifacts. The normalized confusion matrix for the target domain (Table 6) indicates balanced and consistent classification performance, with accuracies of 95% for *LoF*, 94% for *conduction*, and 99% for *keyhole*. The minimal off-diagonal entries confirm that the latent spectral embeddings remain stable and discriminative across domains, effectively capturing the essential

Table 5
Class-wise prediction accuracy of the LWS model.

Ground truth Classification (accuracy [%])	*LoF*	*Conduction*	*Keyhole*
LoF	98.5	1.5	0.0
Conduction	1.5	97.5	1.0
Keyhole	0.0	0.3	99.7

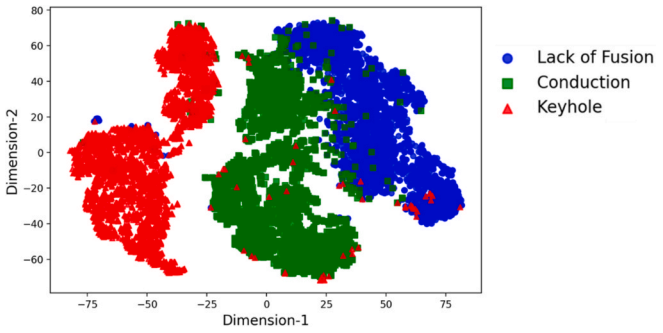

Fig. 13. *T*-SNE visualization of the 128-dimensional latent embeddings obtained from the LWS model, illustrating distinct clustering of melt-pool regimes (*LoF*, *conduction*, and *keyhole*).

Table 6
Class-wise prediction accuracy of the LWS model under zero-shot domain transfer on D_2.

Ground truth Classification (accuracy [%])	*LoF*	*Conduction*	*Keyhole*
LoF	95.0	5.0	0.0
Conduction	4.0	94.0	2.0
Keyhole	0.5	0.5	99.0

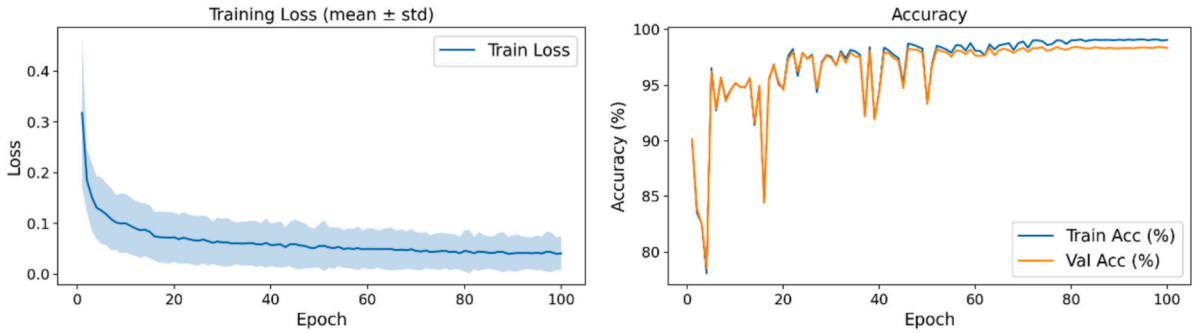

Fig. 12. Training convergence behavior of the LWS model.

time–frequency cues of melt-pool dynamics.

To further verify feature stability, the 128-dimensional latent embeddings obtained from D_2 were projected into two dimensions using *t*-SNE. The resulting clusters (Fig. 14) exhibit clear, non-overlapping separations corresponding to the three melt-pool regimes, confirming that the latent space learned from D_1 preserves class boundaries even under distributional shift. These findings demonstrate that the LWS front-end captures domain-invariant acoustic representations governed by melt-pool physics rather than experimental variability.

6.3. Wavelet evolution and explainability analysis

To elucidate the spectral learning dynamics of the LWS network, the evolution of the learnable Morlet filters was analyzed between the initial and final training epochs. The frequency response of each filter $H_j(f)$ was computed as the magnitude of its Fourier transform in Eq (14):

$$H_j(f) = \left|\mathscr{F}\{\psi_j(t)\}\right|, \tag{14}$$

where $\psi_j(t)$ denotes the time-domain Morlet kernel corresponding to the j^{th} filter. The relative change in spectral magnitude between the final ($H_j^{(\text{final})}$) and initial ($H_j^{(\text{init})}$) training epochs was then expressed as in Eq (15)

$$\Delta H_j(f) = 20\log_{10}\left(\frac{H_j^{(\text{final})}(f)}{H_j^{(\text{init})}(f)}\right), \tag{15}$$

providing a signed measure of gain or attenuation across the frequency spectrum.

The resulting spectral adaptation of the learnable wavelet bank is visualized in Fig. 15. The left panel presents the *ternary sign map*, indicating the direction of spectral change for each filter and frequency bin—red denotes positive gain (+1), blue represents attenuation (−1), and gray indicates negligible change (0) within ± 0.25 dB. Lower-order filters (indices 1–6), centered below 40 kHz, exhibit localized positive gains that correspond to enhanced representation of *conduction* – and partially fused melt-pool behavior. Conversely, higher-order filters show dominant negative shifts, reflecting suppression of high-frequency components that contribute limited physical information. The right panel displays the *signed Δ-magnitude map*, which quantifies both the magnitude and extent of spectral adaptation. The color scale encodes the signed difference in decibels between the final and initial filter magnitudes, with red regions denoting amplified bands and blue regions representing attenuation. Collectively, these visualizations demonstrate that the LWS network adaptively redistributes its spectral emphasis—reinforcing low- and mid-frequency filters most responsive to melt-pool dynamics while suppressing high-frequency bands dominated by stochastic noise. This adaptive spectral tuning confirms the model's capacity to self-organize its frequency response toward physically meaningful and discriminative regions.

Fig. 16 illustrates the frequency-domain magnitude responses of the 18 learnable Morlet filters before and after training, arranged by ascending center frequency from approximately 1 Hz to 160 kHz. The gray curves denote the initial (untrained) responses, while the colored curves show the final optimized filters. The shaded gray regions indicate octave boundaries, ensuring logarithmic spacing across the 0–200 kHz spectrum. During optimization, the low-order filters (1–6) preserved stable center frequencies with minor bandwidth adjustments, effectively capturing *conduction*-mode and partially fused melt-pool behavior. Mid-frequency filters (7–12) exhibited moderate passband sharpening and a gradual downward shift in their center frequencies. High-order filters (13–18) showed a pronounced migration of their center frequencies toward lower ranges, accompanied by bandwidth narrowing and selective amplification. This collective downward movement of the wavelet centers indicates a redistribution of spectral focus from high-frequency transients to more physically informative low- and mid-frequency regions, emphasizing the predominance of sub-40 kHz components in representing melt-pool dynamics and demonstrating the LWS model's convergence toward the most causally relevant acoustic bands.

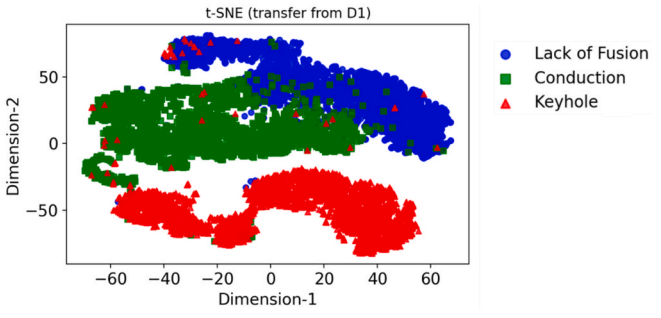

Fig. 14. *T*-SNE projection of 128-dimensional embeddings from the LWS model under zero-shot transfer ($D_1 \rightarrow D_2$).

To assess whether the observed spectral reorganization depends on the initial placement of the wavelet filters, the Morlet bank was initialized using three schemes: random initialization within the sensor bandwidth, linear frequency spacing, and the logarithmic spacing used in the primary experiments. All other architectural and training hyperparameters were kept identical. As shown in Fig. 17, despite substantially different initial frequency distributions, the learned filter responses after 100 training epochs converge to comparable frequency allocations. In particular, the concentration of filters toward the low-to-mid frequency range remains consistent across all initialization strategies. These findings indicate that the final spectral organization is not dictated by the logarithmic initialization prior but emerges from the data-driven optimization process. With 18 wavelets spanning the full sensor bandwidth (~0–200 kHz), each initialization provides sufficient initial spectral coverage to support effective gradient-based refinement across the relevant frequency regions. During training, the optimization consistently reallocates filters toward frequency bands that maximize class separability, indicating the presence of a stable, data-driven optimum. Classification performance and convergence behavior were also consistent across the three initialization schemes, showing no meaningful dependence on the choice of initialization. Overall, this robustness supports the conclusion that the observed spectral organization arises from the underlying AE signal structure rather than from a specific initialization strategy.

6.4. Causal and frequency-resolved explainability

To quantify the contribution of individual frequency bands to class discrimination, a model-level causal analysis was performed on the trained LWS model. Each wavelet sub-band was selectively attenuated within short temporal windows, and the resulting change in class probability was measured. The perturbed forward pass was implemented by applying an attenuation factor (α) between 0 and 1 to the *jth* frequency band and recomputing the class posterior probability. The causal effect of each band on a specific class was then calculated as the reduction in posterior confidence relative to the unperturbed baseline. A positive difference indicated that suppression of the frequency band reduced classification confidence, thereby revealing that the band carried causal importance for that class. Each wavelet sub band j was selectively attenuated within short temporal windows, and the resulting changes in class probability were measured. The perturbed forward pass was defined as in Eq (16)

$$p_c^{(\alpha j)} = \text{softmax}\left(\frac{W_c\, z^{(\alpha j)} + b_c}{\tau}\right), \tag{16}$$

where $\alpha \in [0, 1]$ is the attenuation factor applied to the j^{th} band, $z^{(\alpha j)}$ denotes the latent embedding computed under perturbation, and τ is the

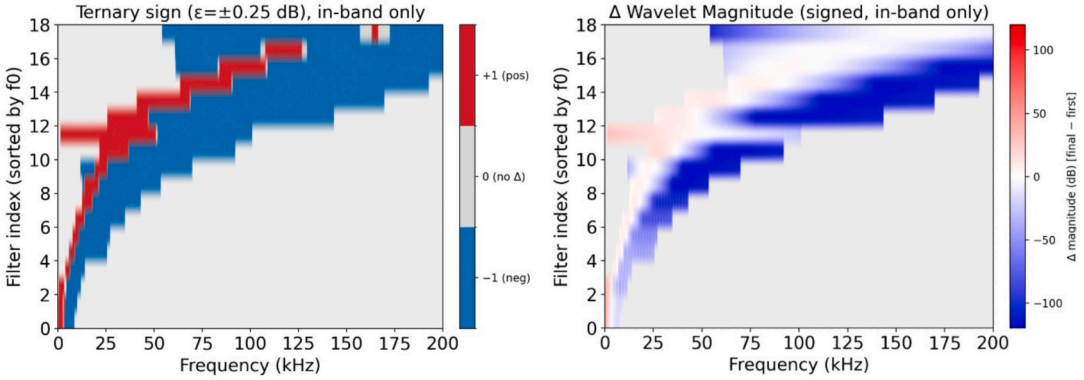

Fig. 15. Ternary sign and Δ-Magnitude maps depicting spectral adaptation of the learnable wavelet bank.

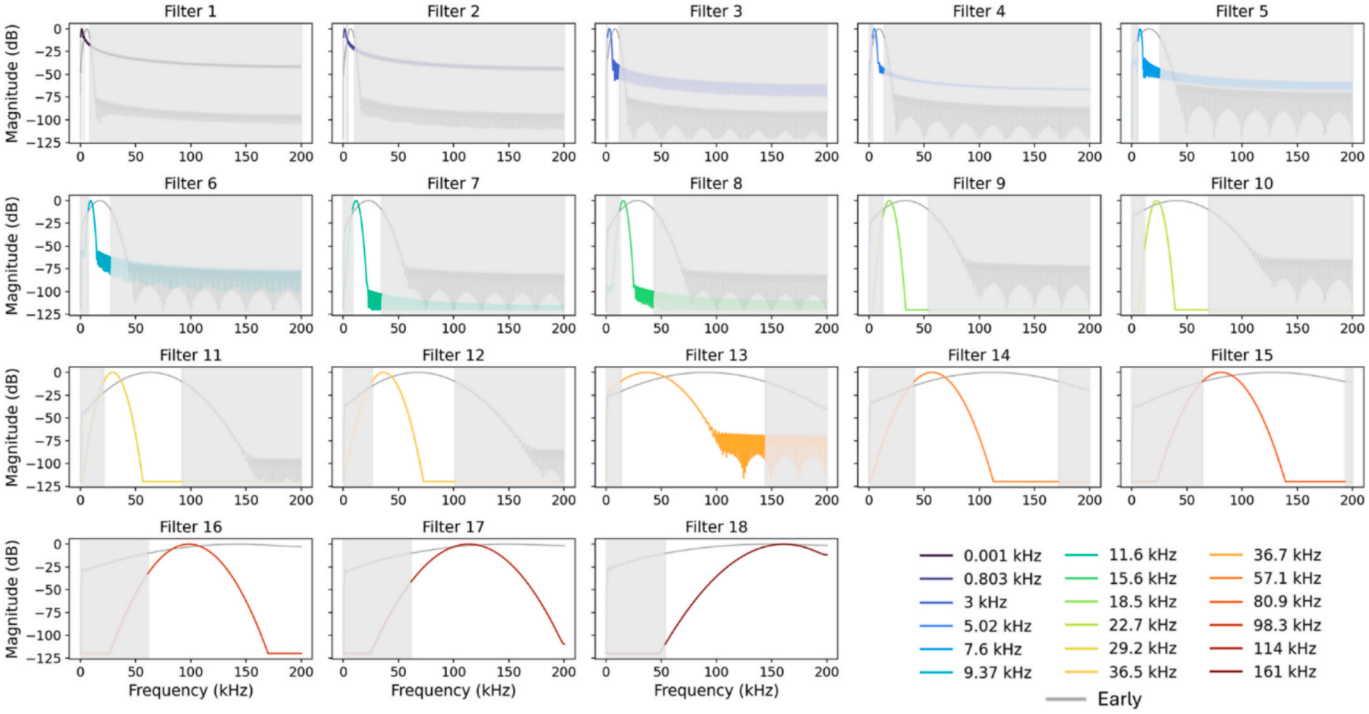

Fig. 16. Frequency-domain magnitude responses of all 18 learnable Morlet filters before and after training, ordered by ascending center frequency.

temperature scaling parameter. The causal effect of each band on *classc* was quantified as the mean reduction in posterior probability as in Eq (17)

$$\Delta P_c^{(j)} = P_c^{(\text{base})} - P_c^{(\alpha j)}, \tag{17}$$

where $P_c^{(\text{base})}$ is the unperturbed prediction probability. A positive $\Delta P_c^{(j)}$ indicates that attenuation of the j^{th} frequency band decreases confidence in class c, thereby revealing that the corresponding spectral band is causally important for that regime. The aggregated causal responses across all filters were visualized through two complementary representations. Fig. 18 illustrates the dominant inter-class causal dependencies across four characteristic frequency bands: 0–20 kHz, 20–40 kHz, 40–80 kHz, and 80–160 kHz. The edge thickness represents the strength of causal influence, while the color denotes the corresponding frequency range. In the low-frequency domain (0–40 kHz), all three regimes— *LoF*, *conduction*, and *keyhole* exhibit strong bidirectional coupling, indicating that these lower bands encode broadband stress-wave interactions common to stable and transitional melting. Between 40 kHz and 80 kHz, the *keyhole* connections largely diminish, and only weak *conduction – LoF* links remain, suggesting that the sensor data within this frequency range is insufficient or unreliable for capturing meaningful melt-pool information. At higher frequencies above 80 kHz, the causal activity concentrates primarily between *conduction* and *LoF*, while the *keyhole* regime becomes nearly inactive. This transition toward selective low-frequency dominance highlights that the key physical information governing melt-pool dynamics resides mainly below 40 kHz, with higher bands contributing minimally to process discrimination.

The Fig. 19 provides a detailed visualization of the frequency-to-class causal Sankey mapping derived from the perturbation analysis. Each link represents the cumulative causal contribution $\Delta P_c^{(j)}$ of a specific wavelet band to a target class, and the link width is proportional to its magnitude, allowing direct comparison of spectral relevance across regimes. The frequency bands are arranged vertically on the left, progressing from the lowest to the highest acoustic frequencies, while the three melt-pool regimes— *LoF*, *conduction*, and *keyhole*—are shown on the right. The causal importance of each frequency band was quantified by systematically attenuating individual wavelet sub-bands and measuring the resulting reduction in class confidence; a larger reduction indicated stronger causal relevance for that class. Formally, the total

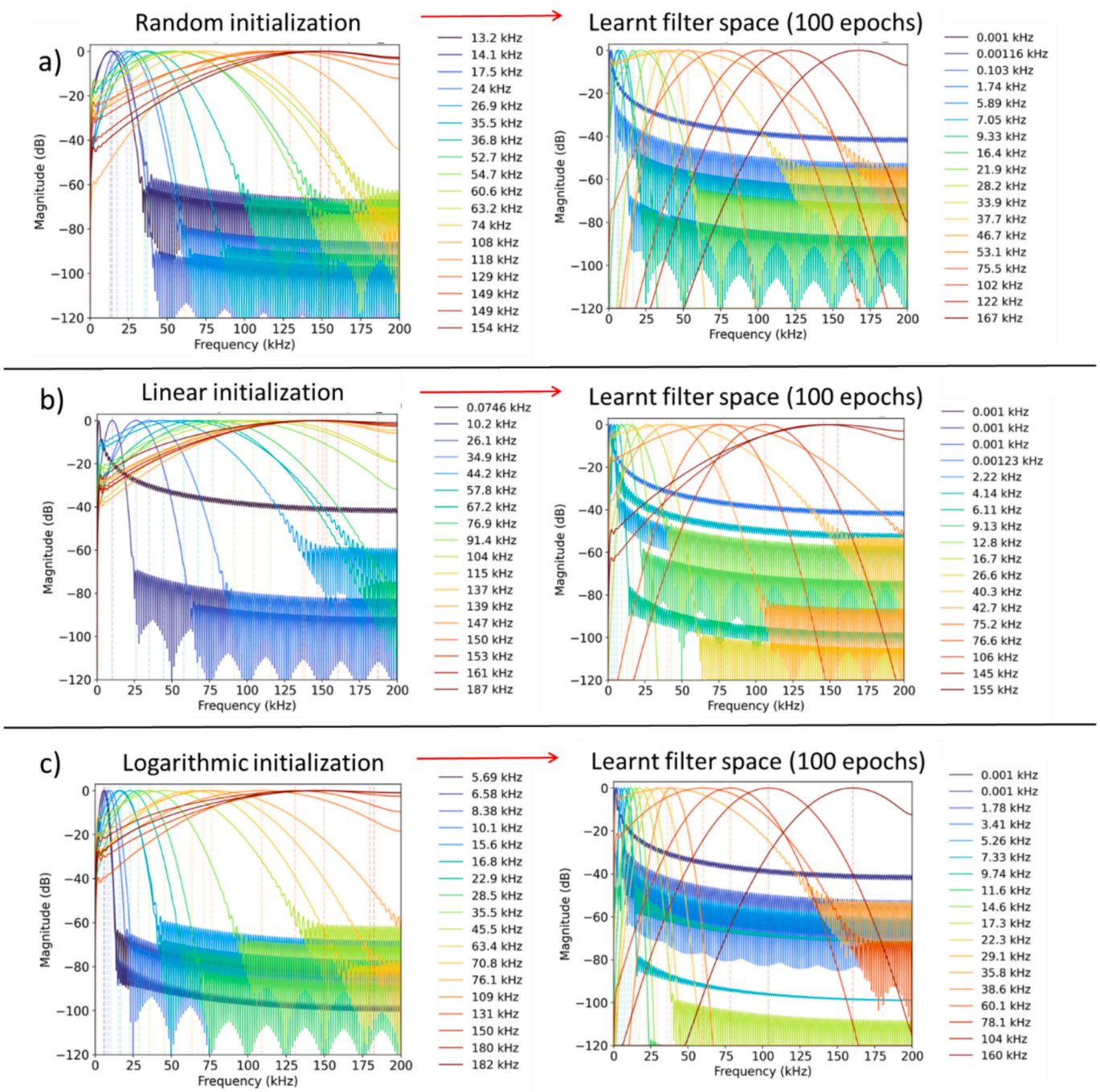

Fig. 17. Initialization sensitivity of the 18 learnable Morlet filters before and after training. *Left*: initial magnitude responses for (a) random, (b) linear, and (c) logarithmic frequency initialization. *Right*: corresponding learned filter spaces after 100 training epochs, showing convergence to comparable frequency allocations across initializations.

band-level causal importance integrated over all classes is expressed as in Eq (18)

$$I_j = \sum_{c=1}^{C} \left| \Delta P_c^{(j)} \right|, \tag{18}$$

and the class-specific importance ratio is computed as in Eq (19)

$$R_c^{(j)} = \frac{\left| \Delta P_c^{(j)} \right|}{\sum_{c'} \left| \Delta P_{c'}^{(j)} \right|}, \tag{19}$$

The thickness of each link in the Fig. 19 is proportional to I_j, while the color encodes its spectral position, enabling direct comparison of causal relevance across frequency bands and melt-pool regimes. The Fig. 19 reveals that most of the causal energy is concentrated in the low-frequency range below 40 kHz, which contributes dominantly to both *conduction* and *LoF* classifications. These frequencies correspond to broadband mechanical stress waves and thermoelastic oscillations associated with stable or partially fused melt pools. In contrast, the mid-frequency region (40–80 kHz) shows very limited or scattered connectivity, indicating that the sensor response in this range is insufficient to capture reliable melt-pool information. Beyond 80 kHz, only weak and sporadic causal flows are observed, primarily linking *conduction* and *LoF*, while the *keyhole* regime receives comparatively minor input from these high-frequency components. Overall, the Sankey representation

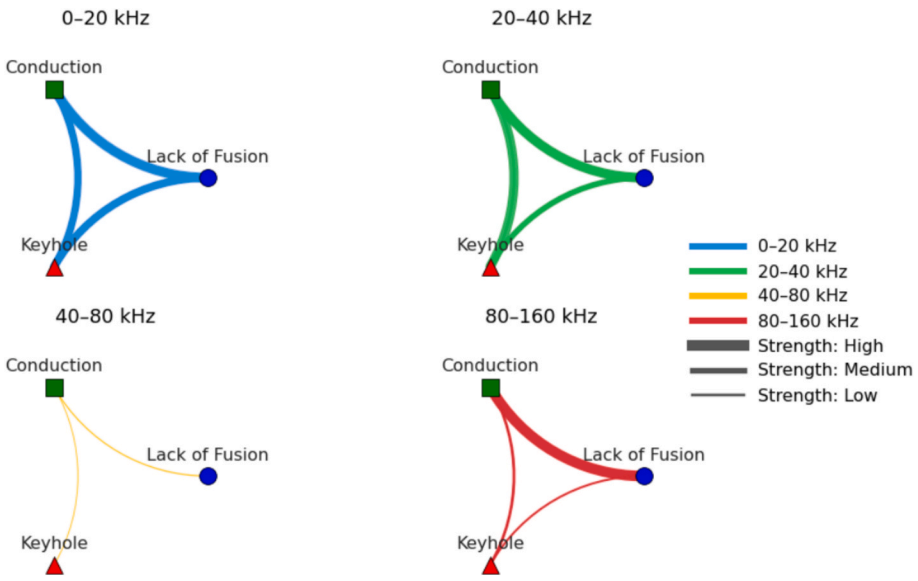

Fig. 18. Causal transition diagrams illustrate frequency-dependent inter-class connectivity.

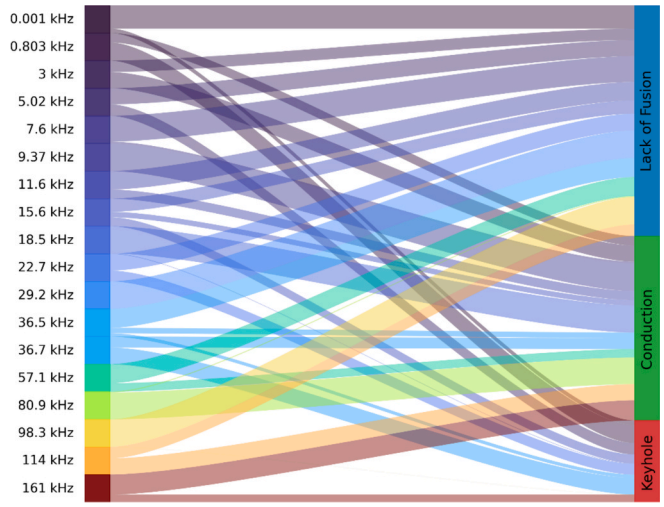

Fig. 19. Sankey diagram visualizing band-to-class causal contributions.

emphasizes that the most informative spectral content for process discrimination lies below 40 kHz, confirming the low-frequency dominance of melt-pool acoustics and the limited physical relevance of the 40–80 kHz band for sensing purposes.

The differing frequency dependence across regimes can be interpreted in terms of the characteristic spatial and temporal scales of melt-pool dynamics. In the *keyhole* regime, acoustic emission is governed by large-scale vapor cavity oscillations driven by recoil pressure and surface tension, which occur at longer characteristic time scales ($\tau \sim 1/f$) and therefore concentrate energy in low-frequency bands, typically below 30 to 40 kHz [77]. In contrast, the *conduction* and *LoF* regimes involve shallower melt pools and more localized, transient thermo-mechanical events such as capillary surface perturbations, intermittent bonding, rapid solid–liquid transitions, and localized stress release. These processes occur over shorter time scales and smaller spatial domains, which naturally contribute to broader and higher-frequency acoustic components. This provides a mechanistic explanation for why *keyhole* influence diminishes above approximately 80 kHz in Fig. 19, while *conduction* and *LoF* retain measurable sensitivity in higher bands. This interpretation is consistent with prior explainable machine learning analysis on airborne AE in 316 L, which identified decision-relevant information predominantly below 40 kHz while highlighting the importance of finer low-to-mid frequency structure for separating closely related regimes [70]. The model-level causal analysis reveals that the LWS framework offers frequency-resolved interpretability by directly linking learned representations to physically meaningful process signatures. By quantifying the direction and magnitude of class transitions across frequency scales, the method transforms model explainability into a physics-grounded spectral attribution framework. The analysis confirms that the LWS model learns stable, process-consistent acoustic features that generalize across domains, capturing invariant frequency bands governing melt-pool transitions. These physically robust embeddings provide a reliable foundation for assessing the model's transferability, as discussed in the following section on cross-domain analysis.

6.5. Benchmarking

To contextualize zero-shot transfer performance, we include a standard lightweight 1D-CNN baseline trained on the source domain D_1 using the same AE segments, preprocessing, optimization settings, and classifier head as the proposed LWS pipeline. The CNN *(CNN-1)* is then evaluated on the target domain D_2 without retraining or fine-tuning, serving as a benchmark reference for how a conventional time-domain feature learner degrades under cross-domain drift. To ensure a fair and literature-consistent baseline, *CNN-1* adopts a standard 1D convolutional design inspired by previously reported AE-based LPBF monitoring frameworks that did not investigate zero-shot transfer learning [63]. Fig. 20 (a) illustrates the *CNN-1* architecture employed in this study. Each AE segment of 12.5 ms (≈ 5 000 samples at 400 kHz) was processed through five one-dimensional convolutional layers (kernel size = 16), followed by a Fully Connected layer *(FC)* and a final classification layer. Each convolutional block incorporated batch normalization and a ReLU activation, with a 5% dropout applied to mitigate overfitting. The output of *FC* served as a compact latent embedding for subsequent visualization. Overall, *CNN-1* consisted of 46 455 trainable parameters. The model was trained with a multi-class classification objective using the cross-entropy loss function. Optimization was performed using the Adam optimizer with a *lr* of 0.001 over 100 epochs. Each epoch comprised mini-batches of 512 AE samples, with the dataset shuffled at the start of every epoch to ensure randomized training. A dropout rate of 0.05 was consistently applied across all convolutional layers, and the data were split into 70% for training and

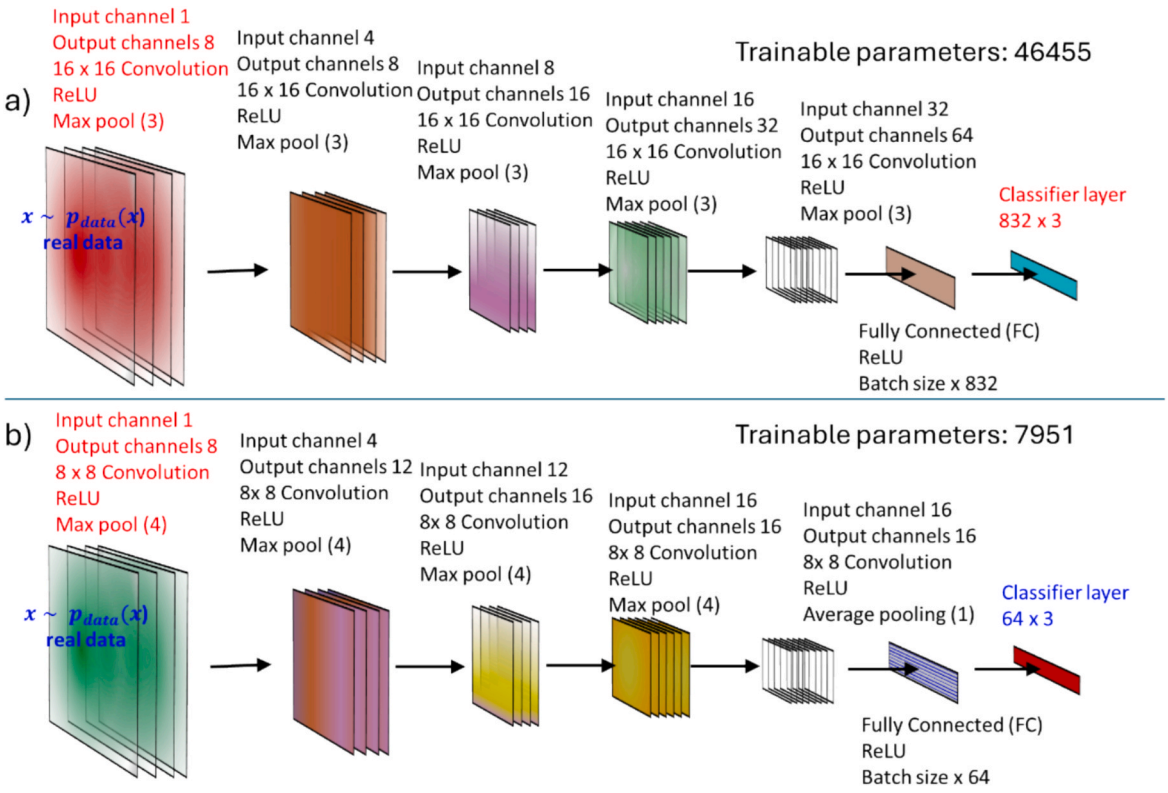

Fig. 20. Architectures of the CNN baselines used for benchmarking: (a) over-parameterized *CNN-1* (46,455 parameters) and (b) parameter-matched *CNN-2* (7,951 parameters) implemented to isolate the effect of model capacity from architectural inductive bias.

30% for testing. The total of 46,455 tunable parameters was optimized to classify the AE signals into the three melt-pool regimes: *LoF*, *conduction*, and *keyhole*.

On the source domain (D_1), CNN-1 achieved a mean classification accuracy of 98.1%, clearly distinguishing the three melt-pool regimes (Table 7). However, when the trained model was directly applied to the target domain (D_2) without retraining, its performance dropped markedly to 77.8%, with most misclassifications occurring between the *LoF* and *conduction* regimes. To further examine whether the observed transfer degradation is attributable to model over-parameterization rather than architectural inductive bias, we implemented an additional compact convolutional baseline (*CNN-2*) with approximately 7,951 trainable parameters, closely matching the LWS model (~7,991 parameters). The architecture of *CNN-2* is illustrated in Fig. 20(b). It follows the same preprocessing pipeline, data split (70%/30%), optimizer (Adam), learning rate, number of epochs, and loss function as *CNN-1*, while employing reduced channel widths to control model capacity. No additional hyperparameter tuning was performed beyond capacity matching to ensure a fair comparison. On the source domain (D_1), *CNN-2* achieved a mean classification accuracy of 97%, indicating that the reduced parameter count still permits strong in-domain discrimination (Table 7). However, under zero-shot transfer ($D_1 \rightarrow D_2$), its accuracy decreased to 76.9%, closely mirroring the degradation observed for *CNN-1*. The comparable cross-domain performance of *CNN-1* (46,455 parameters) and *CNN-2* (~8k parameters) indicates that the transfer gap is not primarily explained by model size or implicit regularization effects. Instead, it reflects a structural limitation of purely time-domain convolutional architectures under spectral distribution shift. While *CNN-1* and *CNN-2* exhibit substantial degradation under cross-domain drift, this effect is not uniform across regimes. In particular, *keyhole* classification remains highly robust on D_2, whereas most errors arise from increased confusion between *LoF* and *conduction*. This indicates that the principal benefit of LWS lies in stabilizing the learned representation and decision boundary for the more spectrally adjacent *LoF* and *conduction modes* under drift, rather than improving an already well-separated *keyhole* regime. The strong transferability of the *keyhole* class for both CNN and LWS further suggests that *keyhole* acoustics encode a comparatively distinctive signature, whereas *LoF* and *conduction* require finer, frequency-resolved separation that is more sensitive to domain-dependent spectral shifts.

In addition to quantitative evaluation, the organization of the learned latent embeddings under zero-shot transfer was examined using *t*-SNE projection. Fig. 21 shows the 2D embedding space for the fixed-scattering front-end when trained on D_1 and evaluated on D_2. Compared to the full LWS model (Section 6.2), the clusters exhibit increased overlap between *LoF* and *conduction* regimes, indicating

Table 7
Class-wise classification accuracy (%) of *CNN-1* (46,455 parameters) and parameter-matched *CNN-2* (7,951 parameters) across source (D_1) and zero-shot target (D_2) domains.

Ground truth Classification (accuracy [%])	*LoF*	*Conduction*	*Keyhole*
LoF	**98.3**	**1.7**	**0.0**
	57.3	*42.7*	*0.0*
	98.0	**2.0**	**0.0**
	56.8	*43.2*	*0.0*
Conduction	**2.2**	**97.1**	**0.7**
	15.6	*78.2*	*6.2*
	3.2	**95.0**	**1.8**
	17.0	*76.0*	*7.0*
Keyhole	**0.2**	**0.7**	**99.1**
	1.2	*0.8*	*98.0*
	0.9	**1.1**	**98.0**
	1.2	*0.8*	*98.0*

Each cell's classification outcomes are arranged in the descending order listed below: *CNN-1* (**Bold- D_1**), *CNN-1 (Italics- D_2)*, *CNN-2* (**Bold- D_1**) and *CNN-2 (Italics- D_2)*. All values are in %.

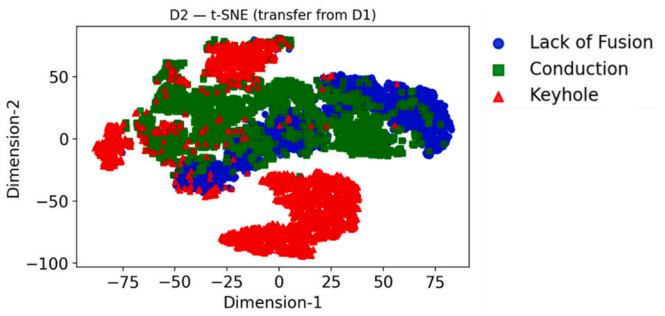

Fig. 21. *T*-SNE projection of latent embeddings obtained from the fixed-scattering front-end under zero-shot transfer ($D_1 \rightarrow D_2$), showing increased inter-class overlap compared to the learnable LWS configuration.

reduced class separability under distribution shift. This reduced separability also translates into degraded classification performance in the target domain, resulting in lower prediction accuracy for the fixed-scattering configuration. Overall, these results indicate that while fixed wavelet scattering provides structured time–frequency representations, adaptive refinement of wavelet center frequencies and bandwidth parameters improves spectral alignment with melt-pool dynamics, yielding more stable cross-domain embeddings. The increased *LoF–conduction* overlap suggests that a fixed wavelet tiling cannot adequately accommodate domain-shifted AE spectral patterns; in contrast, learnable adaptation of wavelet center frequencies and bandwidths mitigates this overlap by re-centering filters on the most discriminative, physically informative bands, confirming that the "learnable" component is critical for robustness under drift.

6.6. Discussion and insights

By replacing fixed analytical wavelets with a trainable Morlet filter bank, the proposed LWS model dynamically adapts its spectral decomposition to the process-specific acoustic landscape while preserving interpretability through explicit frequency localization. The entire framework followed a systematic pipeline, beginning with process parameter selection via BO to identify the optimal configuration of spectral hyperparameters. Eighteen learnable wavelets were obtained through the combination of $J = 6$ octaves and $Q = 3$ wavelets per octave, ensuring dense spectral coverage of the AE bandwidth while maintaining computational efficiency. This parameterization offered an effective balance between frequency resolution, model compactness, and discriminative capability, allowing the model to autonomously identify sub-bands corresponding to physically meaningful melt-pool phenomena such as *LoF*, *conduction*, and *keyhole* oscillations. The evolution of the learnable filters during training provided insight into how discriminative spectral regions are allocated across regimes, rather than serving only as qualitative validation. Lower-frequency filters captured mechanical stress-wave activity associated with *keyhole*-mode defects and bulk AE waveforms, while higher-frequency filters responded to transient vapor-plume oscillations and fine-scale melt-pool instabilities. Importantly, the learned redistribution of spectral emphasis clarifies regime separability under domain drift, particularly along the spectrally adjacent *LoF–conduction* boundary, where conventional CNN baselines exhibited greater confusion. The *keyhole*-mode regimes exhibited dominant low-frequency bursts, typically below 30 kHz, linked to vapor cavity oscillations, plume pulsations, and periodic *keyhole* collapse events. The learnable wavelet filters adaptively sharpened their sensitivity in these regions, revealing the spectral fingerprints of melt-pool instability and the transition toward the *keyhole* regime. As training progressed, the center frequencies of several wavelets migrated toward lower spectral ranges, emphasizing the dominance of low-frequency content in encoding process-relevant acoustic dynamics. From a sensor-design perspective, this observation is particularly significant because most of the learned wavelet centers lie within the sub-40 kHz range, suggesting that effective acoustic monitoring can be achieved at moderate sampling rates. This provides actionable guidance for optimizing data-acquisition systems, reducing sampling frequency, storage requirements, and signal-processing overhead, thereby bridging model interpretability and monitoring strategy.

The Δ-magnitude and ternary-sign analyses demonstrated that the learned spectral adaptations align closely with melt-pool physics rather than arbitrary statistical fitting, reinforcing the framework's ability to maintain interpretability while achieving data-driven optimization. The systematic downward drift of the wavelet center frequencies provides further evidence that the model converges toward physically discriminative and stable acoustic bands instead of high-frequency noise. The causal and frequency-resolved attribution analyses were used to quantify model-level causal influence of specific frequency bands on regime predictions, clarifying how the network links acoustic frequencies to melt-pool states. Perturbation-based causal mapping quantitatively identified the frequency bands that most strongly influence classification outcomes, clarifying how the network links acoustic frequencies to physical melt-pool states. These relationships, visualized through circular transition and Sankey diagrams, provide a mechanistic interpretation of the model's decision process, showing that classification boundaries arise from physically grounded spectral cues. The zero-shot transfer evaluation confirmed that the learned acoustic embeddings remain robust to variations in powder size distribution, material composition, and process parameters. The model's consistent cross-domain performance validates that it learns invariant acoustic principles of laser–material interaction rather than dataset-specific artifacts, establishing the LWS framework as a physics-consistent, domain-independent foundation for process monitoring in LPBF.

In the context of our previously reported domain adaptation and self-supervised representation learning approaches [59,63], the LWS framework introduces a distinct perspective. Whereas prior domain adaptation strategies require exposure to target-domain data to align embeddings across distributions, the present approach learns frequency-structured embeddings exclusively from the source domain and preserves their stability under domain shift without additional alignment. Moreover, the compact parameterization (~8k trainable parameters) and the explicitly inspectable Morlet filter bank provide both computational efficiency and direct interpretability of the learned spectral basis. The ability to visualize the evolution of center frequencies and bandwidths offers a physically grounded understanding of how embedding reorganizes under training, which is less transparent in purely alignment-driven or high-capacity CNN representations. In this sense, LWS complements prior adaptation-based methods by providing a zero-shot, compact, and frequency-interpretable embedding framework for LPBF monitoring.

It should be noted that the cross-domain robustness demonstrated here pertains to intra-material distribution shifts (powder size variation within 316 L) under fixed machine architecture. Transfer across fundamentally different alloy systems may exhibit altered spectral perturbation patterns due to changes in thermophysical properties and melt-pool dynamics, as indicated in prior studies [58,78]. While the physics-aligned wavelet parameterization provides a structured embedding, full inter-material zero-shot generalization cannot be assumed and requires dedicated validation. Similarly, machine-specific acoustic transfer functions may introduce spectral bias; however, sensor-position standardization and calibration protocols can mitigate such effects. In scenarios involving more pronounced material or hardware variability, integrating acoustic embeddings with complementary sensing modalities (e.g., optical plume monitoring, melt-pool imaging, or thermal signals) through structured sensor-fusion strategies may further enhance robustness. Multimodal fusion can help disentangle material-dependent melt-pool dynamics from machine-induced acoustic transfer characteristics, enabling more stable cross-platform representations. These broader generalization scenarios, including multimodal

fusion for cross-material and cross-machine transfer, remain important directions for future work.

7. Conclusion

This study introduced a physics-informed deep learning framework based on LWS for interpretable and transferable in-situ process monitoring in LPBF. The proposed approach addresses one of the fundamental challenges in additive manufacturing the limited generalization of data-driven monitoring models under domain drift arising from variations in material composition, process parameters, and sensor conditions. By integrating trainable Morlet filters with hierarchical time–frequency decomposition, the LWS front-end autonomously adapts its spectral representation to the acoustic characteristics of the melt-pool dynamics. The framework effectively captures the discriminative frequency structures corresponding to the *LoF*, *conduction*, and *keyhole* regimes, achieving high classification fidelity and smooth convergence across training and validation datasets. The spectral evolution analysis revealed that the learned filters self-organize toward frequency bands most responsive to melt-pool physics, while the model-level causal analysis established frequency-resolved interpretability identifying spectral regions that causally influence classification outcomes. These findings demonstrate that the LWS model does not operate as a black box but instead builds transparent relationships between learned representations and the underlying physical processes. Crucially, the framework achieved zero-shot transfer across distinct powder distributions and process conditions without retraining, validating the domain invariance of the learned spectral embeddings. This capability confirms that the network encodes stable, process-driven acoustic features rather than dataset-specific noise, establishing a pathway toward generalizable, sensor-informed monitoring systems for industrial LPBF platforms. Key contributions of this work include:

- A fully trainable LWS architecture featuring a Morlet wavelet bank that autonomously adapts its spectral filters to process-specific acoustic regimes, integrating feature extraction, latent embedding, and causal explainability into a unified, end-to-end framework that achieves high classification accuracy while maintaining strong interpretability of melt-pool states.
- A frequency-resolved attribution framework that quantitatively links specific spectral regions to distinct melting modes. The *keyhole* regime was predominantly concentrated in the lower-frequency band (<40 kHz), and a sensor bandwidth up to 40 kHz was found sufficient to distinguish among the three LPBF modes.
- Demonstration of strong cross-domain and zero-shot transferability, confirming that the learned embeddings remain robust, physically meaningful, and invariant under data drift. The physics-driven design of the LWS model renders it agnostic to transient variations, enabling generalization across different powder distributions and process conditions.

While the proposed framework demonstrates strong interpretability and transferability, several limitations remain. The current study focused exclusively on AE data, which, although informative, captures only one modality of the complex melt-pool dynamics. The dataset was also restricted to a specific machine and alloy system, and further validation across multiple hardware configurations and process scales is required to confirm generalization. Future research will extend this framework toward multimodal sensor fusion, integrating acoustic, optical, and thermal signals to provide a holistic understanding of melt-pool dynamics. Additionally, coupling the LWS model with reinforcement learning and adaptive control strategies could enable real-time, closed-loop stabilization of the melting process. The incorporation of uncertainty quantification and physics-informed priors will further enhance reliability, paving the way for intelligent, self-calibrating monitoring systems in next-generation AM system. For those interested, the data and codes for this study can be accessed in the following repositories (https://gitlab.utu.fi/vpsora/additive-manufacturing-learnable-explainable-wavelet-scattering.git).

CRediT authorship contribution statement

Vigneashwara Pandiyan: Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Rafał Wróbel:** Writing – review & editing, Methodology, Investigation. **Sergey Shevchik:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Christian Leinenbach:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank Marc Leparoux for his support with the powder distribution characterization. We also wish to acknowledge Roland Axel Ritchner and Antonios Baganis for their contributions to the experimental setup, data-collection planning, and for their valuable input during scientific discussions.

Data availability

Data will be made available on request.

References

[1] S. Patel, Y. Liu, Z. Siddique, I. Ghamarian, Metal additive manufacturing: principles and applications, J. Manuf. Process. 131 (2024) 1179–1201.
[2] X. Qi, X. Liang, J. Wang, H. Zhang, X. Wang, Z. Liu, Microstructure tailoring in laser powder bed fusion (L-PBF): strategies, challenges, and future outlooks, J. Alloy. Compd. 970 (2024) 172564.
[3] N. Rizk, M. Nazzal, B.M. Darras, I. Deiab, Comparative sustainability assessment of laser powder bed fusion additive manufacturing and conventional machining for Ti-6Al-4V: a multi-criteria decision-making approach, J. Clean. Prod. 525 (2025) 146569.
[4] S. Chowdhury, N. Yadaiah, C. Prakash, S. Ramakrishna, S. Dixit, L.R. Gupta, D. Buddhi, Laser powder bed fusion: a state-of-the-art review of the technology, materials, properties & defects, and numerical modelling, J. Mater. Res. Technol. 20 (2022) 2109–2172.
[5] S. Patel, M. Vlasea, Melting modes in laser powder bed fusion, Materialia 9 (2020) 100591.
[6] C.L.A. Leung, S. Marussi, R.C. Atwood, M. Towrie, P.J. Withers, P.D. Lee, In situ X-ray imaging of defect and molten pool dynamics in laser additive manufacturing, Nat. Commun. 9 (2018) 1355.
[7] E. Li, H. Shen, L. Wang, G. Wang, Z. Zhou, Laser shape variation influence on melt pool dynamics and solidification microstructure in laser powder bed fusion, Addit. Manuf. Lett. 6 (2023) 100141.
[8] X. Yang, Y. Li, B. Li, Formation mechanisms of lack of fusion and keyhole-induced pore defects in laser powder bed fusion process: a numerical study, Int. J. Therm. Sci. 188 (2023) 108221.
[9] A.A. Martin, N.P. Calta, S.A. Khairallah, J. Wang, P.J. Depond, A.Y. Fong, V. Thampy, G.M. Guss, A.M. Kiss, K.H. Stone, Dynamics of pore formation during laser powder bed fusion additive manufacturing, Nat. Commun. 10 (2019) 1987.
[10] Z. Snow, A.R. Nassar, E.W. Reutzel, Invited Review Article: review of the formation and impact of flaws in powder bed fusion additive manufacturing, Addit. Manuf. 36 (2020) 101457.
[11] Y. Huang, T.G. Fleming, S.J. Clark, S. Marussi, K. Fezzaa, J. Thiyagalingam, C.L. A. Leung, P.D. Lee, Keyhole fluctuation and pore formation mechanisms during laser powder bed fusion additive manufacturing, Nat. Commun. 13 (2022) 1170.
[12] T. DebRoy, H.L. Wei, J.S. Zuback, T. Mukherjee, J.W. Elmer, J.O. Milewski, A.M. Beese, A.d. Wilson-Heid, A. De, W. Zhang, Additive manufacturing of metallic components–process, structure and properties, Progress in materials science, 92 (2018) 112-224.
[13] J.P. Oliveira, A. LaLonde, J. Ma, Processing parameters in laser powder bed fusion metal additive manufacturing, Mater. Des. 193 (2020) 108762.
[14] J. Wang, R. Zhu, Y. Liu, L. Zhang, Understanding melt pool characteristics in laser powder bed fusion: an overview of single-and multi-track melt pools for process optimization, Adv. Powder Mater. 2 (2023) 100137.

[15] P. Bidare, I. Bitharas, R. Ward, M. Attallah, A.J. Moore, Fluid and particle dynamics in laser powder bed fusion, Acta Mater. 142 (2018) 107–120.
[16] I. Bitharas, N. Parab, C. Zhao, T. Sun, A. Rollett, A. Moore, The interplay between vapour, liquid, and solid phases in laser powder bed fusion, Nat. Commun. 13 (2022) 2959.
[17] W. Yuan, H. Chen, S. Ruan, R. Lupoi, S. Qin, E. Guo, J. Wang, S. Yin, Oscillatory nature in melt-gas-powder interactions during laser powder bed fusion process revealed by CFD-DEM coupled modelling, Virt. Phys. Prototyp. 20 (2025) e2446619.
[18] A. Montelione, R. Schur, R. Schleusener, M. Ramulu, D. Arola, Powder reuse and variability in laser powder bed fusion additive manufacturing of Ti6Al4V: a round robin study, J. Mater. Res. Technol. 35 (2025) 4231–4248.
[19] Z.Y. Chua, I.H. Ahn, S.K. Moon, Process monitoring and inspection systems in metal additive manufacturing: status and applications, Int. J. Precis. Eng. Manuf. Green Technol. 4 (2017) 235–245.
[20] G. Tapia, A. Elwany, A review on process monitoring and control in metal-based additive manufacturing, J. Manuf. Sci. Eng. 136 (2014) 060801.
[21] C. Hole, Cost and practicality of in-process monitoring for metal Additive Manufacturing, Met Addit. Manuf. 2 (2016) 63–69.
[22] P. Charalampous, I. Kostavelis, D. Tzovaras, Non-destructive quality control methods in additive manufacturing: a survey, Rapid Prototyp. J. 26 (2020) 777–790.
[23] W. Yang, Y. Qiu, W. Liu, X. Qiu, Q. Bai, Defect prediction in laser powder bed fusion with the combination of simulated melt pool images and thermal images, J. Manuf. Process. 106 (2023) 214–222.
[24] R. Wang, D. Garcia, R.R. Kamath, C. Dou, X. Ma, B. Shen, H. Choo, K. Fezzaa, H. Z. Yu, Z. Kong, In situ melt pool measurements for laser powder bed fusion using multi sensing and correlation analysis, Sci. Rep. 12 (2022) 13716.
[25] H. Balhara, A. Karthikeyan, A. Hanchate, T.G. Nakkina, S.T. Bukkapatnam, Imaging systems and techniques for fusion-based metal additive manufacturing: a review, Front. Manuf. Technol. 3 (2023) 1271190.
[26] T. Özel, A review on in-situ process sensing and monitoring systems for fusion-based additive manufacturing, Int. J. Mech. Manuf. Syst. 16 (2023) 115–154.
[27] K. Taherkhani, E. Sheydaeian, C. Eischer, M. Otto, E. Toyserkani, Development of a defect-detection platform using photodiode signals collected from the melt pool of laser powder-bed fusion, Addit. Manuf. 46 (2021) 102152.
[28] C.K.P. Vallabh, S. Sridar, W. Xiong, X. Zhao, Predicting melt pool depth and grain length using multiple signatures from in-situ single camera two-wavelength imaging pyrometry for laser powder bed fusion, J. Mater. Process. Technol. 308 (2022) 117724.
[29] D. Mahmoud, M. Magolon, J. Boer, M. Elbestawi, M.G. Mohammadi, Applications of machine learning in process monitoring and controls of L-PBF additive manufacturing: a review, Appl. Sci. 11 (2021) 11910.
[30] D. Birtalan, W. Nunley, Optoelectronics: infrared-visable-ultraviolet devices and applications, CRC Press, 2018.
[31] A. Papatheodorou, N. Papadimitriou, E. Stathatos, P. Benardos, G.-C. Vosniakos, Recent advances in sensor fusion monitoring and control strategies in laser powder bed fusion: a review, Machines 13 (2025) 820.
[32] Z. Li, Z. Zhang, S. Zhang, Z. Bai, R. Qin, J. Huang, J. Wang, K. Huang, Q. Zhang, G. Wen, A novel approach of online monitoring for laser powder bed fusion defects: air-borne acoustic emission and deep transfer learning, J. Manuf. Process. 102 (2023) 579–592.
[33] Z. Li, Z. Zhang, J. Wang, Z. Du, S. Zhang, K. Huang, Q. Zhang, Y. Su, G. Wen, X. Chen, In-situ porosity monitoring in laser powder bed fusion through acoustic signal and interpretable multi-sensor fusion, Mech. Syst. Sig. Process. 241 (2025) 113482.
[34] S. Zhang, Z. Zhang, R. Qin, J. Wang, J. Huang, Z. Li, Y. Su, G. Wen, Q. Zhang, X. Chen, Variable-scale acoustic texture image and interpretable filter convolutional networks for defect monitoring in laser powder bed fusion, Expert Syst. Appl. 129221 (2025).
[35] D. Kouprianoff, N. Luwes, I. Yadroitsava, I. Yadroitsev, Acoustic emission technique for online detection of fusion defects for single tracks during metal laser powder bed fusion, Solid Freeform Fabricat. (2018).
[36] J.R. Tempelman, A.J. Wachtor, E.B. Flynn, P.J. Depond, J.-B. Forien, G.M. Guss, N. P. Calta, M.J. Matthews, Sensor fusion of pyrometry and acoustic measurements for localized keyhole pore identification in laser powder bed fusion, J. Mater. Process. Technol. 308 (2022) 117656.
[37] N. Eschner, L. Weiser, B. Häfner, G. Lanza, Classification of specimen density in laser powder bed fusion (L-PBF) using in-process structure-borne acoustic process emissions, Addit. Manuf. 34 (2020) 101324.
[38] R. Esmaeilzadeh, V. Pandiyan, S. Van Petegem, M. Van der Meer, M.H. Nasab, C. de Formanoir, J. Jhabvala, C. Navarre, L. Schlenger, R. Richter, Acoustic emission signature of martensitic transformation in laser powder bed fusion of Ti6Al4V-Fe, supported by operando X-ray diffraction, Addit. Manuf. 96 (2024) 104562.
[39] M. Seleznev, T. Gustmann, J.M. Friebel, U.A. Peuker, U. Kühn, J.K. Hufenbach, H. Biermann, A. Weidner, In situ detection of cracks during laser powder bed fusion using acoustic emission monitoring, Addit. Manuf. Lett. 3 (2022) 100099.
[40] R.A. Richter, J. Jhabvala, N. Mari, R. Logé, E. Iseli, Fine detection of process state and secondary process emissions in laser powder-bed fusion (LPBF) by dual acoustic sensor acquisition and machine learning, in, IOP Conf. Ser. Mater. Sci. Eng. IOP Publ. (2025) 012032.
[41] J.Y. Song, A. Dass, A. Moridi, G.C. McLaskey, Detection of defects during laser-powder interaction by acoustic emission sensors and signal characteristics, Addit. Manuf. 82 (2024) 104035.
[42] L. Chen, G. Bi, X. Yao, J. Su, C. Tan, W. Feng, M. Benakis, Y. Chew, S.K. Moon, In-situ process monitoring and adaptive quality enhancement in laser additive manufacturing: a critical review, J. Manuf. Syst. 74 (2024) 527–574.
[43] S.K. Everton, M. Hirsch, P. Stravroulakis, R.K. Leach, A.T. Clare, Review of in-situ process monitoring and in-situ metrology for metal additive manufacturing, Mater. Des. 95 (2016) 431–445.
[44] Y. Du, T. Mukherjee, R. Li, Z. Hou, S. Dutta, C.B. Arnold, A. Elwany, S. Kung, J. Tang, T. DebRoy, A review of deep learning in metal additive manufacturing: impact on process, structure, and properties, Prog. Mater Sci. 101587 (2025).
[45] Z. Snow, L. Scime, A. Ziabari, B. Fisher, V. Paquit, Scalable in situ non-destructive evaluation of additively manufactured components using process monitoring, sensor fusion, and machine learning, Addit. Manuf. 78 (2023) 103817.
[46] J. Kim, Z. Yang, Y. Lu, A physics-guided deep generative model for predicting melt pool behavior in laser powder bed fusion additive manufacturing, J. Intell. Manuf. (2024) 1–21.
[47] V. Pandiyan, G. Masinelli, N. Claire, T. Le-Quang, M. Hamidi-Nasab, C. de Formanoir, R. Esmaeilzadeh, S. Goel, F. Marone, R. Logé, Deep learning-based monitoring of laser powder bed fusion process on variable time-scales using heterogeneous sensing and operando X-ray radiography guidance, Addit. Manuf. 58 (2022) 103007.
[48] A. Gaikwad, R.J. Williams, H. de Winton, B.D. Bevans, Z. Smoqi, P. Rao, P. A. Hooper, Multi phenomena melt pool sensor data fusion for enhanced process monitoring of laser powder bed fusion additive manufacturing, Mater. Des. 221 (2022) 110919.
[49] Y. Zhang, W. Yan, Applications of machine learning in metal powder-bed fusion in-process monitoring and control: status and challenges, J. Intell. Manuf. 34 (2023) 2557–2580.
[50] K.-H. Lee, H.W. Lee, G.J. Yun, A defect detection framework using three-dimensional convolutional neural network (3D-CNN) with in-situ monitoring data in laser powder bed fusion process, Opt. Laser Technol. 165 (2023) 109571.
[51] S. Larsen, P.A. Hooper, Deep semi-supervised learning of dynamics for anomaly detection in laser powder bed fusion, J. Intell. Manuf. 33 (2022) 457–471.
[52] V. Pandiyan, R. Drissi-Daoudi, S. Shevchik, G. Masinelli, T. Le-Quang, R. Logé, K. Wasmer, Semi-supervised monitoring of Laser powder bed fusion process based on acoustic emissions, Virt. Phys. Prototyp. 16 (2021) 481–497.
[53] B. Kavas, R.A. Richter, M.R. Tucker, V. Pandiyan, Adaptive In-Situ monitoring for laser powder bed Fusion: self-Supervised learning for layer thickness monitoring across scan lengths based on pyrometry, Opt. Laser Technol. 192 (2025) 114070.
[54] A. Riensche, K. Deshmukh, A. Carrington Jr, M. Darji, B. Bevans, K. Snyder, J. Sions, Y. Plotnikov, K.D. Cole, P. Rao, Rapid, autonomous, and shape-agnostic physics-guided thermal history control to improve part quality in laser powder bed fusion additive manufacturing, J. Manuf. Sci. Eng. 147 (2025) 081005.
[55] J. Sousa, B. Brandau, R. Hemschik, R. Darabi, A. Sousa, L.P. Reis, F. Brueckner, A. Reis, Jema-sindyc: end-to-end control using joint embedding multimodal alignment in directed energy deposition, Addit. Manuf. 104888 (2025).
[56] B. Kavas, E.C. Balta, M.R. Tucker, R. Krishnadas, A. Rupenyan, J. Lygeros, M. Bambach, In-situ controller autotuning by Bayesian optimization for closed-loop feedback control of laser powder bed fusion process, Addit. Manuf. 99 (2025) 104641.
[57] B. Kavas, E.C. Balta, M. Tucker, A. Rupenyan, J. Lygeros, M. Bambach, Layer-to-layer closed-loop feedback control application for inter-layer temperature stabilization in laser powder bed fusion, Addit. Manuf. 78 (2023) 103847.
[58] V. Pandiyan, A. Baganis, R. Axel Richter, R. Wróbel, C. Leinenbach, Qualify-as-you-go: sensor fusion of optical and acoustic signatures with contrastive deep learning for multi-material composition monitoring in laser powder bed fusion process, Virt. Phys. Prototyp. 19 (2024) e2356080.
[59] V. Pandiyan, R. Wróbel, R.A. Richter, M. Leparoux, C. Leinenbach, S. Shevchik, Self-Supervised Bayesian representation learning of acoustic emissions from laser powder bed Fusion process for in-situ monitoring, Mater. Des. 235 (2023) 112458.
[60] S. Zhang, Z. Zhang, J. Wang, J. Huang, R. Qin, H. Qin, Z. Li, G. Wen, Q. Zhang, X. Chen, A graph neural network integrating physical prior knowledge for defect monitoring in laser powder bed fusion, J. Manuf. Process. 153 (2025) 516–530.
[61] F.G. Fischer, M.G. Zimmermann, N. Praetzsch, C. Knaak, Monitoring of the powder bed quality in metal additive manufacturing using deep transfer learning, Mater. Des. 222 (2022) 111029.
[62] H. Zhang, Z. Zhao, C. Wang, X. Zhang, X. Chen, Mitigating domain shift in online process monitoring for material extrusion additive manufacturing via transfer learning, Addit. Manuf. 94 (2024) 104467.
[63] V. Pandiyan, R. Wróbel, R.A. Richter, M. Leparoux, C. Leinenbach, S. Shevchik, Monitoring of Laser Powder Bed Fusion process by bridging dissimilar process maps using deep learning-based domain adaptation on acoustic emissions, Addit. Manuf. 80 (2024) 103974.
[64] V. Pandiyan, R. Drissi-Daoudi, S. Shevchik, G. Masinelli, T. Le-Quang, R. Loge, K. Wasmer, Deep transfer learning of additive manufacturing mechanisms across materials in metal-based laser powder bed fusion process, J. Mater. Process. Technol. 303 (2022) 117531.
[65] R. Drissi-Daoudi, G. Masinelli, C. de Formanoir, K. Wasmer, J. Jhabvala, R.E. Logé, Acoustic emission for the prediction of processing regimes in Laser Powder Bed Fusion, and the generation of processing maps, Addit. Manuf. 67 (2023) 103484.
[66] M. Hamidi Nasab, G. Masinelli, C. de Formanoir, L. Schlenger, S. Van Petegem, R. Esmaeilzadeh, K. Wasmer, A. Ganvir, A. Salminen, F. Aymanns, Harmonizing sound and light: X-ray imaging unveils acoustic signatures of stochastic inter-regime instabilities during laser melting, Nat. Commun. 14 (2023) 8008.
[67] Q. Luo, J.D. Shimanek, T.W. Simpson, A.M. Beese, An image-based transfer learning approach for using in situ processing data to predict laser powder bed

fusion additively manufactured Ti-6Al-4V mechanical properties, 3D Print. Addit. Manuf. 12 (2025) 48–60.
[68] S. Liu, A.P. Stebner, B.B. Kappes, X. Zhang, Machine learning for knowledge transfer across multiple metals additive manufacturing printers, Addit. Manuf. 39 (2021) 101877.
[69] M.A. Ilani, Y.M. Banad, TransMatch: A Transfer-Learning Framework for Defect Detection in Laser Powder Bed Fusion Additive Manufacturing, arXiv preprint arXiv:2509.01754, (2025).
[70] V. Pandiyan, R. Wróbel, C. Leinenbach, S. Shevchik, Optimizing in-situ monitoring for laser powder bed fusion process: deciphering acoustic emission and sensor sensitivity with explainable machine learning, J. Mater. Process. Technol. 321 (2023) 118144.
[71] N. Aydin, Wavelet scattering transform based doppler signal classification, Comput. Biol. Med. 167 (2023) 107611.
[72] S. Mallat, A wavelet tour of signal processing, Academic Press, 1999.
[73] C. Pedersen, M. Eickenberg, S. Ho, Learnable wavelet neural networks for cosmological inference, arXiv preprint arXiv:2307.14362, (2023).
[74] T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, Optuna: A next-generation hyperparameter optimization framework, in: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, 2019, pp. 2623-2631.
[75] V.W. Berger, Y. Zhou, Kolmogorov–smirnov test: Overview, Statistics reference online, Wiley statsref, 2014.
[76] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, Pytorch: an imperative style, high-performance deep learning library, Adv. Neural Inf. Proces. Syst. 32 (2019).
[77] S.A. Khairallah, T. Sun, B.J. Simonds, Onset of periodic oscillations as a precursor of a transition to pore-generating turbulence in laser melting, Addit. Manuf. Lett. 1 (2021) 100002.
[78] R. Drissi-Daoudi, V. Pandiyan, R. Logé, S. Shevchik, G. Masinelli, H. Ghasemi-Tabasi, A. Parrilli, K. Wasmer, Differentiation of materials and laser powder bed fusion processing regimes from airborne acoustic emission combined with machine learning, Virt. Phys. Prototyp. 17 (2022) 181–204.