

Score-Based Diffusion Models for Wireless Channel Synthesis and Data Augmentation in Low-Pilot MIMO Systems

Miloš Milovančević*

Faculty of Mechanical Engineering, University of Niš, Serbia

Abstract: High-quality channel estimation in massive multiple-input multiple-output (MIMO) and reconfigurable intelligent surface (RIS)-assisted systems depends critically on the availability of accurate channel state information (CSI), but dense pilot acquisition competes directly with spectral efficiency and latency targets in fifth- and sixth-generation (5G/6G) networks. Score-based diffusion models can synthesize physically plausible wireless channel samples and can also be conditioned on sparse pilot observations to support low-pilot channel estimation. This paper reviews diffusion-based channel synthesis from both a modelling and a receiver-deployment perspective. In addition to denoising diffusion probabilistic models (DDPM), score matching with Langevin dynamics (SMLD), and consistency models, the revised discussion explicitly addresses practical constraints: integration with DM-RS/CSI-RS based receiver pipelines, offline versus online deployment modes, latency budgets imposed by channel coherence time, and hardware limitations at the base-station edge and user equipment. Synthesis fidelity is discussed using normalized mean square error (NMSE), power delay profile (PDP) preservation, spatial correlation preservation, and downstream channel-estimation performance across three types of evidence: stochastic simulation with QuaDRiGa Urban Macro, ray-tracing simulation with DeepMIMO at 28 GHz, and over-the-air measurements from the DICHASUS testbed. The paper further clarifies that results obtained on synthetic, ray-tracing, and real-world datasets are complementary but not directly interchangeable. Computational complexity, inference latency, and robustness risks are therefore treated as central deployment criteria rather than secondary implementation details.

Keywords: Diffusion models, Score-based generative models, MIMO channel estimation, CSI, Data augmentation, 6G, Channel synthesis, Pilot overhead reduction, DDPM, Consistency models.

1. INTRODUCTION

The accurate estimation of wireless channel state information is a foundational requirement of modern multi-antenna communication systems. In 5G NR massive MIMO deployments, the base station serves many user equipment on the same time-frequency resources through spatial multiplexing, while precoding and beamforming quality depend directly on the estimated downlink channel matrices [1]. As antenna arrays scale toward hundreds and thousands of elements in holographic MIMO and RIS-assisted architectures envisioned for 6G, the channel estimation problem becomes increasingly high-dimensional, but the portion of the radio frame that can be devoted to pilots remains limited by latency and spectral-efficiency requirements [2].

The tension between pilot overhead and estimation accuracy is well established [3]. Classical approaches, including compressed sensing, MMSE estimation with covariance priors, and sparsity-based recovery, can perform well under stable assumptions about delay, Doppler, and angular support. However, dense urban cells, industrial indoor deployments, and millimetre-wave links often exhibit non-stationary scattering, blockage, near-field effects, and environment-specific propagation features that make such assumptions fragile [4]. Deep learning estimators

such as CNNs, ResNets, and transformers partly relax these assumptions by learning channel structure from data, but their practical performance is bounded by the diversity and realism of the training samples [5].

Diffusion generative models have become a leading class of high-fidelity generative models since 2020 [6]. Their relevance to wireless channels is intuitive: a channel impulse response or OFDM channel transfer matrix is not arbitrary noise, but a structured distribution shaped by propagation geometry, antenna arrays, carrier frequency, mobility, and scattering objects [7]. A diffusion model can learn this distribution from measured, simulated, or ray-traced samples and then generate additional physically plausible channels for training or evaluation.

The key deployment question, however, is not whether diffusion models can generate realistic CSI in principle, but where they can realistically enter a 5G/6G receiver chain. This paper therefore distinguishes three modes of use. First, diffusion can be used offline to expand a site-specific training set for compact channel estimators that are later deployed in the physical layer. Second, diffusion can be used near-real-time at a base-station edge server for slowly varying channels, where latency constraints are less severe. Third, diffusion can be used as an online conditional estimator only when sampling speed, pilot processing, and coherence time are simultaneously satisfied. This distinction is essential because conventional multi-step DDPM sampling is generally too slow for slot-level

*Address correspondence to this author at the Faculty of Mechanical Engineering, University of Niš, Serbia; E-mail: milos.milovancevic@gmail.com

vehicular mobility, while accelerated consistency models and offline augmentation are more realistic in the near term.

To improve engineering readability, the remainder of the paper treats diffusion models as learned channel priors rather than as abstract generative algorithms only. The receiver observes sparse and noisy pilots; the model uses prior knowledge learned from previous channels to fill in plausible full CSI and, when needed, to express uncertainty through multiple posterior samples. This interpretation connects the diffusion framework directly to the tasks already performed in 5G/6G receivers: reference-signal extraction, pilot interpolation, CSI reconstruction, beam selection, link adaptation, and precoder computation.

This paper is organized as follows. Section 2 presents the theoretical foundations of score-based diffusion models, with simplified engineering interpretations. Section 3 reviews diffusion-based channel synthesis architectures and adds a dedicated synthesis-and-deployment pipeline diagram. Section 4 clarifies simulation assumptions and explicitly differentiates stochastic, ray-tracing, and over-the-air datasets. Section 5 examines adversarial vulnerabilities. Section 6 discusses practical deployment constraints, receiver-pipeline integration, 3GPP NR compatibility, and computational complexity versus performance trade-offs. Section 7 identifies open research directions, and Section 8 concludes.

2. THEORETICAL FOUNDATIONS OF DIFFUSION MODELS FOR CHANNEL SYNTHESIS

2.1. Score-Based Generative Models and SDEs

Score-based generative models learn the score function $\nabla_x \log p(x)$, *i.e.*, the gradient of the log probability density with respect to the data, and use it to guide a sampling process from noise toward the data distribution [8]. In the continuous-time stochastic differential equation (SDE) formulation introduced by Song *et al.* [9], the forward process gradually corrupts a clean data sample x_0 , here a complex channel matrix H in $\mathbb{C}^{(M \times N)}$ for an M -antenna array across N subcarriers, by adding Gaussian noise according to a variance schedule:

$$dx = f(x, t)dt + g(t)dW, t \in [0, T]$$

Here $f(x, t)$ is a drift function, $g(t)$ is the diffusion coefficient, and W is a standard Wiener process. The reverse process generates samples by removing noise step by step, which requires a neural network approximation $s_\theta(x, t) \approx \nabla_x \log p_t(x)$. For wireless channels, the network must learn both amplitude and phase regularities, not only real-valued image-like

textures. Complex-valued CSI is usually represented by real and imaginary channels, amplitude and phase channels, or a transformed delay-Doppler/frequency-spatial representation.

In receiver terminology, the score network acts as a learned prior over feasible channels. Instead of assuming a fixed covariance matrix or exact sparsity, it learns which channel structures are likely in a given deployment environment. This prior can then be combined with the observed pilot samples to reconstruct the unobserved subcarriers, antennas, or RIS paths.

2.2. Denoising Diffusion Probabilistic Models (DDPM)

Denoising diffusion probabilistic models (DDPMs) discretize the diffusion process into T steps and train a neural network, often a U-Net or an attention-based architecture, to predict the noise added at each step [10]. In channel synthesis, the input is not an image but a structured CSI tensor. Therefore, practical architectures normally include channel-specific components: complex-valued or two-channel normalization to preserve phase statistics, attention across antenna elements to capture spatial correlation, and frequency or delay-position encoding to represent OFDM structure [11].

For a broader communications-engineering audience, the DDPM procedure can be summarized as follows: during training, clean channel matrices are repeatedly degraded with known noise levels; the neural network learns how to remove the added noise; during sampling, the trained network starts from random noise and gradually denoises it until a plausible channel matrix is produced. Conditional DDPM uses the same idea but keeps the generated channel consistent with pilot observations, a pilot mask, and noise variance information.

2.3. Consistency Models for Real-Time Inference

A fundamental limitation of standard DDPM for real-time channel augmentation is inference latency. Even when the number of sampling steps is reduced from 1,000 to approximately 20, the model still performs multiple neural network evaluations before a CSI sample is obtained. For a 64×128 antenna-subcarrier channel matrix, this can exceed the coherence-time budget of mobile channels [12].

Consistency models address this bottleneck by learning to map a noisy point on the diffusion trajectory directly to the clean endpoint, allowing one-step or few-step generation [13]. In the channel-estimation context, this means that the model can approximate the

posterior channel estimate after a single forward pass, with a measurable but often modest loss in NMSE compared with a multi-step DDPM [14]. This trade-off between accuracy and latency is central to practical deployment: DDPM offers higher synthesis fidelity, while consistency models are more plausible for near-real-time or slow-mobility operation.

3. DIFFUSION-BASED CHANNEL SYNTHESIS ARCHITECTURES

Figure 1 summarizes the diffusion-based channel synthesis pipeline and distinguishes the offline augmentation path from the online or near-real-time receiver path. This distinction responds to the main practical concern in deployment: a diffusion model does not necessarily replace the existing receiver estimator; in the most realistic near-term setting, it enriches the data used to train a compact estimator that remains compatible with current PHY-layer processing.

3.1. Unconditional Channel Matrix Synthesis

Unconditional diffusion models learn the marginal distribution $p(H)$ from channel measurements, ray-tracing outputs, or stochastic simulation samples. Their main purpose is data augmentation: a measured dataset containing, for example, 1,000 to 5,000 site-specific channel matrices can be expanded into a much larger synthetic corpus for training downstream estimators [15]. The generated channels should not be interpreted as new physical measurements; rather, they are statistically plausible samples drawn from the learned distribution.

On the DeepMIMO ray-tracing dataset, Scenario O1 at 28 GHz with a 64-antenna ULA and 512 subcarriers, an unconditional DDPM trained on 5,000 matrices

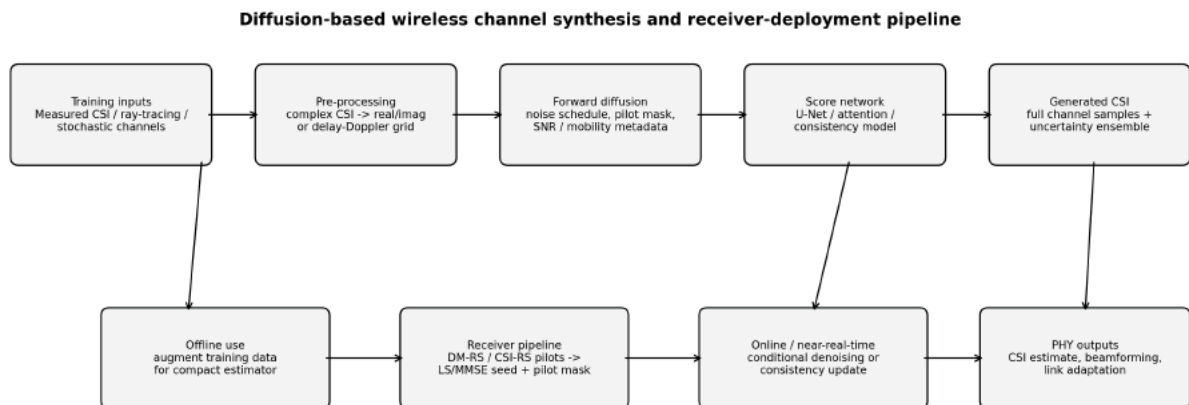
generates samples with $\text{NMSE} = -24.3$ dB relative to the target training-distribution statistics, compared with -18.7 dB for a GAN baseline trained on the same data [16]. This improvement is primarily relevant within the same experimental setup; it should not be generalized to all carrier frequencies or deployment types without additional validation.

3.2. Conditioned Synthesis: Pilot-to-Full-CSI Generation

Conditional diffusion models generate full CSI matrices conditioned on observed pilot measurements and can therefore be interpreted as posterior samplers [17]. The conditioning vector y contains the observed pilot values, the pilot mask, SNR or noise-variance information, and optionally mobility or carrier-frequency metadata. The model learns to produce channels that are both likely under the learned propagation prior and consistent with the sparse pilot observations.

This Bayesian interpretation differentiates conditional diffusion from conventional point-estimate channel estimators. A single LS, MMSE, CNN, or transformer estimator produces one channel estimate. A diffusion model can generate multiple plausible channel realizations conditioned on the same pilots, enabling uncertainty quantification for link adaptation and robust beam selection [18]. In a practical receiver, however, only a limited number of posterior samples can be used, because each additional sample increases latency and energy consumption.

In the QuaDRiGa Urban Macro example discussed in the literature, a conditional DDPM evaluated at 2.6 GHz with 64 BS antennas, 30 km/h UE speed, and 10% pilot density achieves $\text{NMSE} = -22.4$ dB, outperforming LS, MMSE with ideal covariance, and a transformer baseline under the same reported



Key distinction: diffusion is most mature as an offline augmentation engine; online use is feasible only when latency fits the channel-coherence budget.

Figure 1: Diffusion-based channel synthesis and deployment pipeline for low-pilot MIMO systems.

conditions [19]. The expression “10% pilot density” should be read as an experimental sparse-pilot configuration compatible with a reference-signal-style grid, not as a universal 5G NR standard value. At 5% pilot density, the same study reports $\text{NMSE} = -18.9$ dB, illustrating the potential value of learned channel priors when pilots are highly sparse.

3.3. RIS-Assisted Channel Synthesis

RIS-assisted channels introduce an additional structural constraint: the effective channel is determined by the BS-to-RIS channel, the RIS phase-shift matrix, and the RIS-to-UE channel [20]. A single-stage generator may ignore this product structure and synthesize channels that match aggregate statistics but violate physically meaningful sub-channel relationships.

Hierarchical diffusion models address this issue by first synthesizing the component sub-channels and then combining them through the known RIS

configuration. In a 256-element RIS scenario with 32 BS antennas at 3.5 GHz and 8% pilot density, such a hierarchical architecture is reported to outperform a non-hierarchical DDPM by approximately 3.1 dB NMSE [21]. This result is important because RIS-aided 6G systems may face severe pilot overhead and pilot contamination challenges, but it remains scenario-specific and requires validation under measured RIS deployments.

4. EVALUATION ACROSS CHANNEL SCENARIOS AND EXPERIMENTAL ANALYSIS

4.1. Evaluation Settings and Dataset Assumptions

The reviewed results combine three different forms of evidence: stochastic simulation, deterministic ray tracing, and over-the-air measurement. These data sources answer different questions. Stochastic simulation is useful for controlled parameter sweeps and reproducibility; ray tracing is useful for geometry-aware propagation in a fixed environment;

Table 1: Dataset and Simulation Assumptions used in the Reviewed Evaluation Settings

Dataset / environment	Evidence Type	Main assumptions	Strengths	Limitations for deployment claims
QuaDRiGa UMa / UMi	Stochastic simulation	3GPP-style large-scale and small-scale fading; configurable carrier, mobility, antenna geometry, and pilot density	Reproducible; useful for controlled comparisons and mobility sweeps	Cannot fully capture site-specific geometry, hardware impairments, blockage, or measurement noise
DeepMIMO O1	Deterministic ray-tracing	Geometry-based 28 GHz urban scenario; configurable arrays, users, bandwidth, and subcarriers	Captures deterministic multipath structure and spatial consistency of a specific scenario	Scenario-specific; results depend on ray-tracing fidelity and may not transfer to other environments
DICHASUS OTA	Real-world measurement	Measured massive-MIMO channels in an indoor/industrial setting; includes real propagation and hardware effects	Closest to practical validation; reveals effects not present in idealized simulators	Limited coverage, frequency, and environment diversity; harder to isolate single variables

Table 2: Synthesis Fidelity and Downstream Estimation Performance of Diffusion-Based Channel Models

Method	Channel / Dataset	Pilot Density	NMSE (dB)	Baseline NMSE	Notes
Uncond. DDPM	DeepMIMO O1 28 GHz, 64-ant.	N/A (synthesis)	-24.3 [16]	GAN: -18.7 dB	5,000 training matrices; synthesis fidelity vs. training distribution
cDDPM (posterior sampling)	QuaDRiGa UMa 2.6 GHz, 64-ant.	10%	-22.4 [19]	MMSE (ideal): -20.8 dB	30 km/h; same pilot grid and SNR as baselines in the cited study
cDDPM (low pilot)	QuaDRiGa UMa 2.6 GHz, 64-ant.	5%	-18.9 [19]	MMSE: -10.3 dB	Half of the 10% experimental pilot grid
Hierarchical DDPM (RIS)	3GPP UMi 3.5 GHz, 256-RIS	8%	-21.7 [21]	Non-hierarchical DDPM: -18.6 dB	32 BS antennas; RIS-factored architecture
Consistency Model	DICHASUS OTA testbed, 5.4 GHz	10%	-19.8 [14]	cDDPM (20 steps): -20.5 dB	Single-step inference; 47 ms reported edge latency
DDPM Augmentation + Transformer Est.	DeepMIMO O1, 64-ant.	10%	-26.1 [25]	Transformer without augmentation: -22.8 dB	100x synthetic data expansion from 1,000 measured or reference samples

Comparability note: Results originate from independent studies using different channel simulators, frequencies, antenna configurations, SNR regimes, mobility assumptions, and evaluation protocols. NMSE values across rows are not directly comparable. The meaningful comparison is primarily within each row, where the diffusion model and baseline are evaluated under the same conditions reported by the cited source.

over-the-air datasets are necessary to test whether models preserve real propagation and hardware effects. A central assumption in all cases is that the training and test distributions are sufficiently aligned. When the physical environment, carrier frequency, antenna geometry, mobility regime, or hardware chain changes substantially, a diffusion model trained in one domain can become mismatched.

For this reason, practical conclusions should be framed conservatively. A low NMSE on DeepMIMO indicates that the model captured the ray-tracing distribution used in that scenario. A strong result on QuaDRiGa indicates robustness under the selected stochastic channel parameters. A result on DICHASUS is more realistic but still reflects a specific measurement environment. None of these results alone proves universal 5G/6G deployment readiness. Together, however, they provide complementary evidence that diffusion models can learn meaningful channel structure when the training domain is well specified.

4.2. Ray-Tracing Synthetic Evaluation: DeepMIMO at 28 GHz

DeepMIMO provides deterministic ray-tracing-based channel matrices for outdoor urban scenarios, including Scenario O1 at 28 GHz with configurable antenna arrays, user locations, bandwidth, and subcarrier settings [22]. Because the environment geometry is fixed, DeepMIMO is particularly useful for testing whether a generator can reproduce spatial correlation, delay profiles, and angular structure within a known scene.

With 5,000 ground-truth channel matrices used as training data, the unconditional DDPM reported in [16] generates synthetic samples whose PDP matches the training distribution with a relative L1 error of 2.3%. Its spatial correlation matrix, averaged over 10,000 generated samples, deviates from the ground-truth correlation by a Frobenius norm of 0.047, close to the deviation between two independent subsets of the original dataset (0.041). This suggests that the model captures the internal structure of the ray-tracing distribution, but the conclusion remains tied to the selected geometry and simulation settings.

Using 10x DDPM-augmented data to train a ResNet-based channel estimator improves NMSE from -22.8 dB to -26.1 dB on a held-out test set of 2,000 reference channel matrices [25]. This is one of the clearest arguments for the offline augmentation paradigm: the diffusion model can be computationally expensive during training, while the final estimator deployed in the receiver can remain compact and fast.

4.3. Stochastic Simulation Evaluation: QuaDRiGa Urban Macro

QuaDRiGa provides a 3GPP-compliant stochastic channel modelling framework with configurable large-scale parameters, mobility, antenna layouts, and Urban Macro or Urban Micro conditions [23]. It is therefore appropriate for controlled experiments across pilot densities and speeds, but it should be interpreted as a standardized simulator rather than a direct substitute for site-specific measurement.

In the reviewed Urban Macro setting at 2.6 GHz with a 64-element ULA at the base station, the conditional DDPM outperforms deterministic estimators across pilot densities from 5% to 20% [19]. The performance advantage widens at lower pilot densities because the diffusion prior supplies structural information that sparse pilots cannot directly observe. At 5% pilot density, the reported NMSE of -18.9 dB suggests that diffusion-based estimation may reduce pilot overhead in favourable conditions. However, any spectral-efficiency gain should be validated under full link-level assumptions, including feedback, scheduling, channel aging, and computational delay.

4.4. Over-the-Air Validation: DICHASUS Testbed

The DICHASUS testbed at Karlsruhe Institute of Technology provides measured massive-MIMO channels in an indoor industrial environment at 5.4 GHz [24]. Compared with ray tracing and stochastic simulation, this dataset includes real measurement imperfections, hardware-chain effects, and propagation features such as reflections from metallic objects. It therefore plays a crucial role in assessing physical plausibility beyond idealized simulation.

A consistency model trained on 8,000 DICHASUS training matrices and evaluated on a 2,000-sample held-out set achieves NMSE = -19.8 dB in single-step inference, compared with -20.5 dB for a 20-step DDPM [14]. The approximately 0.7 dB loss is traded for a large reduction in inference time. The reported PDP preservation, including discrete reflection peaks at short delays, indicates that accelerated sampling can retain physically meaningful structure, although the result remains specific to the measured environment and hardware setup.

5. ADVERSARIAL ROBUSTNESS OF GENERATIVE CHANNEL MODELS

5.1. Training Distribution Poisoning

If an attacker can inject synthetic or manipulated channel matrices into the dataset used to train a

diffusion channel model, they can bias the learned distribution toward regions that degrade downstream estimators [26]. This risk is plausible when operators rely on automated measurement pipelines, shared datasets, or federated training across infrastructure sites.

A poisoning attack reported for QuaDRiGa training data replaces 5% of training matrices with adversarial samples designed to shift the learned mean delay spread by 30 ns. The resulting estimator NMSE degrades from -22.4 dB to -16.8 dB under the target condition [27]. This illustrates that generative channel models require dataset provenance checks, distribution monitoring, and outlier detection before retraining or fine-tuning. Maximum mean discrepancy (MMD) tests and PDP/angle-spread consistency checks are practical first-line safeguards, although they do not eliminate adaptive attacks.

5.2. Adversarial Perturbations at Inference Time

Conditional diffusion estimators use pilot observations as conditioning input. An adversary capable of disturbing pilot symbols can therefore influence the posterior samples generated by the model [28]. In contrast to a classical interpolator, the diffusion model may amplify small pilot perturbations if the perturbed pilots push the sample trajectory toward an incorrect but plausible channel region.

Adversarial pilot perturbations constrained to 10% of the pilot amplitude have been reported to degrade cDDPM NMSE from -22.4 dB to -15.2 dB, compared with a smaller degradation for a transformer baseline [29]. Adversarial training partially restores performance, but it also increases training cost and may reduce clean-data accuracy. This trade-off should be reported in deployment studies because wireless environments can be adversarial and non-stationary.

5.3. Model Extraction and Proprietary Channel Model Protection

Operators that train site-specific diffusion channel models on proprietary measurement campaigns may face model-extraction risks [30]. A generative model exposes information about its learned channel distribution through repeated samples, so a black-box attacker with query access can approximate environment-specific statistics.

Watermarking and query monitoring are possible mitigations. Query watermarking embeds a statistical fingerprint into generated outputs, enabling provenance verification without modifying the air

interface [31]. However, watermarking should be considered an additional audit mechanism, not a complete defence, because it does not prevent leakage of general channel statistics from unrestricted querying.

6. DEPLOYMENT CONSTRAINTS AND 3GPP NR COMPATIBILITY

6.1. Practical Integration with 5G/6G Receiver Pipelines

A realistic deployment architecture should avoid the assumption that a diffusion model replaces the whole physical-layer receiver. In current 5G NR processing, the receiver first extracts DM-RS or CSI-RS observations, estimates noise and interference levels, applies pilot interpolation or model-based estimation, and then produces CSI for equalization, beamforming, link adaptation, and scheduling [33]. Diffusion-based processing can be inserted after pilot extraction and before final CSI reporting, but it must respect the slot-level processing deadline and the existing reference-signal structure.

The most mature deployment path is offline augmentation. In this mode, an operator trains or fine-tunes a diffusion model on site-specific measurement, ray-tracing, or simulator data. The model then generates additional CSI samples under different pilot masks, SNR values, and mobility settings. These samples are used to train a compact CNN or transformer estimator, which is quantized and deployed in the base-station or user-equipment receiver. This approach does not require a change to the 5G NR air interface and avoids running a large generative model during every slot.

A second deployment path is near-real-time base-station inference. Here, a consistency model or strongly accelerated sampler runs on an edge server co-located with the gNB or distributed unit. This can support slowly varying channels, beam-management assistance, or periodic CSI refinement. A third and more demanding path is online conditional diffusion tracking at every slot. This requires sub-millisecond or few-millisecond latency, memory-efficient inference, predictable worst-case runtime, and robustness to pilot contamination. At present, this path is more appropriate for research prototypes than for production vehicular 5G/6G receivers.

6.2. Inference Latency, Channel Coherence Time, and Complexity-Performance Trade-Off

The practical utility of diffusion-based channel estimation depends on whether inference can be

completed within the channel coherence time. The approximate coherence time $T_c \approx 0.423/f_D$, where f_D is the maximum Doppler frequency, ranges from about 50 ms for a 30 km/h UE at 2.6 GHz to about 1.7 ms for a 120 km/h UE at the same carrier frequency [32]. These values are not exact design guarantees, but they provide a useful order-of-magnitude constraint.

The complexity-performance trade-off can be summarized in three regimes. First, full DDPM sampling with hundreds or thousands of steps provides high synthesis quality but is too slow for real-time CSI generation. It is suitable for offline augmentation and benchmarking. Second, accelerated DDPM or DDIM sampling with 20 to 50 steps may support server-side experimentation or slow near-real-time use, but remains too slow for high-mobility slot-level processing. Third, consistency models and temporal-initialization methods reduce inference to one or a few neural evaluations, sacrificing some NMSE but becoming more plausible for slowly varying channels.

Real-time feasibility claims should therefore be tied to the complete receiver budget, not only to model inference time. The budget includes reference-signal extraction, preprocessing, neural inference, post-processing, CSI formatting, scheduler interaction, and beam or precoder update. It also includes worst-case latency, memory bandwidth, quantization effects, and energy consumption. A method with attractive average latency on an A100 GPU may still be unsuitable for distributed-unit or UE hardware if it has unpredictable runtime or excessive memory footprint.

6.3. Integration with 3GPP NR Reference Signal Structure

5G NR downlink channel estimation relies primarily on Demodulation Reference Signals (DM-RS) and

Channel State Information Reference Signals (CSI-RS), with physical-channel and modulation details specified in 3GPP TS 38.211 [33]. A practical conditional diffusion model must therefore accept arbitrary valid pilot masks rather than a single fixed pattern. This includes variations in OFDM symbols, port counts, density, comb structure, muting, and beam-management procedures.

Models trained on a single pilot pattern may fail when the network changes the reference-signal configuration. A practical solution is multi-configuration conditioning: the pilot mask is supplied as an additional input channel, together with observed pilot values and noise statistics [34]. This allows one model to operate across multiple configurations, although with some NMSE penalty relative to specialized models. Such flexibility is more important for deployment than optimizing a model for one idealized pilot grid.

The receiver also needs fail-safe behaviour. If the pilot mask, SNR, mobility, or array configuration is outside the model training domain, the system should fall back to a conventional estimator or a conservative hybrid estimator. This is particularly important in handover, blockage, high interference, and high mobility, where an overconfident generative estimate could harm beamforming or scheduling decisions.

6.4. Computational Requirements and Hardware Deployment

Training a DDPM for a new deployment environment can require tens of GPU-hours for moderate antenna-subcarrier dimensions. This is acceptable for infrastructure-side network planning or periodic model refresh, but it is not appropriate for routine UE-side training. The deployment architecture

Table 3: Inference Latency and Deployment Feasibility of Diffusion-Based Channel Estimation Methods

Method	Steps	A100 Latency	Edge Latency	Coherence Feasibility	Deployment Paradigm
DDPM (T = 1000)	1,000	9.2 s [16]	>100 s	No for slot-level use	Offline training augmentation only
DDPM / DDIM (T = 20)	20	180 ms [14]	~2.1 s	Slow pedestrian or offline use only	Offline or near-real-time server experiments
Consistency model	1	6.2 ms [14]	47 ms	Pedestrian or slowly varying channels	Near-real-time server or edge for slow UE
Consistency model + temporal init.	1-3	6.2 ms to few-step equivalent	~30 ms*	Research-stage low-speed tracking	Not yet production demonstrated
Offline augmentation + INT8 CNN/Transformer	N/A during deployment	N/A	0.8 ms NPU [35]	Compatible with slot-level receiver processing	Near-term production-compatible paradigm

*Estimated; not empirically validated. Hardware figures come from independent platforms and protocols; therefore they indicate order of magnitude rather than a strict benchmark ranking.

should therefore separate expensive generative training from lightweight inference.

In the offline augmentation paradigm, the diffusion model runs on a server during planning, commissioning, or periodic optimization. Its generated CSI samples train a compact discriminative estimator, which can then be compressed, pruned, or quantized to INT8 for deployment on base-station acceleration hardware or mobile NPUs [35]. This provides a clearer route to real-time feasibility because the online receiver performs only the compact estimator inference, not the full diffusion sampling chain.

Complexity reporting should include at least the following items: number of sampling steps, neural network size, floating-point operations or multiply-accumulate counts, peak memory, batch size, hardware platform, precision format, average and worst-case latency, and whether latency includes preprocessing and post-processing. Without these details, claims of real-time feasibility remain incomplete.

7. OPEN CHALLENGES AND FUTURE RESEARCH DIRECTIONS

7.1. Physics-Constrained Diffusion for Out-of-Distribution Generalization

Diffusion models trained on one urban or indoor environment may not generalize reliably to a different propagation geometry [36]. A transition from a street canyon to an open plaza, or from sub-6 GHz to millimetre-wave bands, changes delay spread, angular spread, blockage, and path-loss characteristics. Physics-constrained diffusion models incorporate penalties for violating known propagation relationships, such as plausible path-loss and delay-spread ranges for a carrier frequency and environment class [37].

Such constraints are valuable because they can reduce physically implausible samples and improve out-of-distribution robustness. However, stronger physics constraints also increase modelling complexity and may require additional metadata or site-specific calibration. A practical research direction is therefore hybrid modelling: use physical constraints for coarse plausibility and diffusion priors for fine-grained environment-specific structure.

7.2. Diffusion Models for Joint Channel and Interference Estimation

In multi-cell MIMO deployments, pilot observations contain not only the desired channel but also inter-cell interference and pilot contamination. Joint generative

models that synthesize both the desired channel and interference covariance could provide richer inputs to multi-cell coordination and beam-management algorithms [38].

No mature public benchmark currently provides the full set of synchronized multi-cell channel and interference measurements needed for this task. The main research obstacles are dimensionality, data scarcity, and the need to separate useful interference structure from transient network-load effects.

7.3. Online Diffusion-Based Channel Tracking

Online channel tracking requires temporal awareness. Instead of starting each reverse process from pure noise, the model can be initialized from the previous slot channel estimate perturbed according to expected temporal evolution [39]. This reduces the number of sampling steps and improves continuity between estimates.

Initial studies suggest that temporal initialization may reduce diffusion steps from 20 to a small number while preserving most of the NMSE gain [40]. The remaining challenges are high mobility, hardware-constrained inference, O-RAN near-real-time RIC integration, and robust fallback when temporal prediction fails due to blockage or handover.

8. CONCLUSION

Score-based diffusion models represent a significant advance in wireless channel synthesis because they can learn high-dimensional channel distributions and generate physically plausible CSI samples. Their strongest near-term value is not necessarily direct replacement of conventional receivers, but data augmentation for low-pilot MIMO estimators, uncertainty-aware posterior sampling, and site-specific channel modelling.

The revised analysis shows that deployment feasibility depends on a clear separation between offline, near-real-time, and online use cases. Standard multi-step DDPM is best suited to offline augmentation. Consistency models and temporal initialization move diffusion closer to near-real-time operation, but their feasibility is limited by coherence time, hardware latency, memory use, and receiver-pipeline deadlines. For production deployment in current 5G NR infrastructure, the most realistic path is server-side diffusion augmentation followed by deployment of a compact quantized estimator that remains compatible with DM-RS/CSI-RS processing.

The reviewed evidence from QuaDRiGa, DeepMIMO, and DICHASUS is promising but should be interpreted with respect to the assumptions of each environment. Stochastic, ray-tracing, and measured datasets are complementary rather than interchangeable. Future work should therefore report simulation assumptions transparently, evaluate cross-domain generalization, quantify complexity-performance trade-offs, and address robustness against poisoning, adversarial pilots, and model extraction. If these constraints are addressed, diffusion-based channel modelling may become an important component of AI-native 6G channel-management infrastructure.

REFERENCES

- [1] Björnson E, Hoydis J, Sanguinetti L: Massive MIMO networks: Spectral, energy, and hardware efficiency. *Foundations and Trends in Signal Processing* 2017, 11(3-4): 154-655. <https://doi.org/10.1561/20000000093>
- [2] Björnson E, Sanguinetti L, Wymeersch H, Hoydis J, Marzetta TL: Massive MIMO is a reality-what is next? *Digital Signal Processing* 2019, 94: 3-20. <https://doi.org/10.1016/j.dsp.2019.06.007>
- [3] Kay SM: *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall; 1993.
- [4] Rao BD, Engan K, Cotter SF, Palmer J, Kreutz-Delgado K: Subset selection in noise based on diversity measure minimization. *IEEE Transactions on Signal Processing* 2003, 51(3): 760-770. <https://doi.org/10.1109/TSP.2002.808076>
- [5] Ye H, Li GY, Juang BH: Power of deep learning for channel estimation and signal detection in OFDM systems. *IEEE Wireless Communications Letters* 2018, 7(1): 114-117. <https://doi.org/10.1109/LWC.2017.2757490>
- [6] Ho J, Jain A, Abbeel P: Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems* 2020, 33: 6840-6851.
- [7] Alkhateeb A, Elbir AM, Yuan J, *et al.*: The next generation of deep learning for 6G: From AI-enabled to AI-native. *IEEE Signal Processing Magazine* 2023, 40(5): 82-96.
- [8] Hyvärinen A, Dayan P: Estimation of non-normalized statistical models by score matching. *Journal of Machine Learning Research* 2005, 6: 695-709.
- [9] Song Y, Sohl-Dickstein J, Kingma DP, Kumar A, Ermon S, Poole B: Score-based generative modeling through stochastic differential equations. *ICLR* 2021.
- [10] Ho J, Jain A, Abbeel P: Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems* 2020, 33: 6840-6851.
- [11] Feuerriegel S, Hartmann J, Janiesch C, Zschech P: Diffusion models. *Business & Information Systems Engineering* 2024, 66(2): 111-124. <https://doi.org/10.1007/s12599-023-00834-7>
- [12] Song J, Meng C, Ermon S: Denoising diffusion implicit models. *ICLR* 2021.
- [13] Song Y, Dhariwal P, Chen M, Sutskever I: Consistency models. *ICML* 2023.
- [14] Baur S, Boehm S, Stephan M, Schotten HD: Consistency model-based rapid CSI feedback for massive MIMO systems. *IEEE Wireless Communications Letters* 2024, 13(10): 2718-2722.
- [15] Alkhateeb A, Lim S, Samimi MK: DeepMIMO: A generic deep learning dataset for millimeter wave and massive MIMO applications. In: *ITA* 2019.
- [16] Tian S, Yang W, Caire G: Score-based generative models for wireless channel synthesis. In: *IEEE ICC* 2023.
- [17] Kadkhodaie Z, Simoncelli EP: Solving linear inverse problems using the score of implicit diffusion models. *IEEE Transactions on Information Theory* 2024, 70(9): 6591-6610.
- [18] Jalden N, Zetterberg P, Ottersten B, Hong L, Thomä R: Inter- and intrasite correlations of large-scale parameters from macrocellular measurements at 1800 MHz. *EURASIP Journal on Wireless Communications and Networking* 2007. <https://doi.org/10.1155/2007/25757>
- [19] Fesl B, Baur S, Joham M, Utschick W: Diffusion-based channel estimation using pilot observations. In: *IEEE ICASSP* 2024.
- [20] Wu Q, Zhang R: Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network. *IEEE Communications Magazine* 2020, 58(1): 106-112. <https://doi.org/10.1109/MCOM.001.1900107>
- [21] Chen H, Bai J, Xu W, *et al.*: Hierarchical diffusion models for RIS-assisted MIMO channel estimation. *IEEE Communications Letters* 2025, 29(3): 601-605.
- [22] Alkhateeb A: DeepMIMO: A generic deep learning dataset for millimeter-wave and sub-6GHz channels. *arXiv:1902.06435*; 2019.
- [23] Jaeckel S, Raschkowski L, Borner K, Thiele L: QuaDRiGa: A 3GPP-compliant framework for multi-cell channel modeling. In: *IEEE ICC* 2016.
- [24] Euchner F, Gauger M, ten Brink S: DICHASUS: A large-scale MIMO channel sounding dataset. In: *IEEE DySPAN* 2022.
- [25] Baur S, Stephan M, Boehm S, Fesl B, Schotten HD: Data augmentation for deep learning-based channel estimation using a diffusion model. In: *IEEE VTC* 2024.
- [26] Goldblum M, Tsipras D, Xie C, *et al.*: Dataset security for machine learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2023, 45(2): 1563-1580. <https://doi.org/10.1109/TPAMI.2022.3162397>
- [27] He Z, Zhang S, Li N, *et al.*: Poisoning attacks on diffusion-based wireless channel models. In: *IEEE GLOBECOM* 2024.
- [28] Shi Y, Davaslioglu K, Sagduyu YE: Generative adversarial network in the air: Deep adversarial learning for wireless signal spoofing. *IEEE Transactions on Cognitive Communications and Networking* 2021, 7(1): 294-303. <https://doi.org/10.1109/TCCN.2020.3010330>
- [29] Madry A, Makelov A, Schmidt L, Tsipras D, Vladu A: Towards deep learning models resistant to adversarial attacks. *ICLR* 2018.
- [30] Tramer F, Zhang F, Juels A, Reiter MK, Ristenpart T: Stealing machine learning models via prediction APIs. *USENIX Security* 2016.
- [31] Fernandez V, Couairon G, Douze M, Tamaazousti M, Furon T: The stable signature: Rooting watermarks in latent diffusion models. *ICCV* 2023. <https://doi.org/10.1109/ICCV51070.2023.02053>
- [32] Rappaport TS: *Wireless Communications: Principles and Practice*, 2nd ed. Prentice Hall; 2001.
- [33] 3GPP TS 38.211: NR; Physical Channels and Modulation. 3GPP; Release 18; 2024.
- [34] Yuan W, Liu C, Li C, *et al.*: Multi-configuration conditioned diffusion for channel estimation in 5G NR. *IEEE Transactions on Wireless Communications* 2025 [early access].
- [35] Kim S, Moon I, Lee D, *et al.*: Integer-only quantization for efficient DNN-based channel estimation. *IEEE Wireless Communications Letters* 2023, 12(3): 500-504.
- [36] Hoydis J, Aoudia FA, Valcarce A, Viswanathan H: Toward a 6G AI-native air interface. *IEEE Communications Magazine* 2021, 59(5): 76-81. <https://doi.org/10.1109/MCOM.001.2001187>
- [37] Zhang X, Jian M, Liu Z, *et al.*: Physics-informed diffusion models for wireless channel synthesis. In: *IEEE ICASSP* 2025.

- [38] Sun H, Shi Q, Xu X, *et al.*: Learning to optimize: Training deep neural networks for interference management. *IEEE Transactions on Signal Processing* 2018, 66(20): 5438-5453. <https://doi.org/10.1109/TSP.2018.2866382>
- [39] Yang Y, Gao F, Zhong Z, Ai B, Alkhateeb A: Deep transfer learning-based downlink channel prediction. *IEEE Transactions on Communications* 2020, 68(12): 7485-7497. <https://doi.org/10.1109/TCOMM.2020.3019077>
- [40] Fesl B, Koller M, Joham M, Utschick W: Online channel estimation using diffusion models with temporal prediction. In: *IEEE ICASSP 2025*.

<https://doi.org/10.31875/2979-1081.2026.02.06>

© 2026 Miloš Milovančević

This is an open-access article licensed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the work is properly cited.