

ACCEPTED MANUSCRIPT

Digital Twin Entanglement: Enhanced measurement & reference benchmarking

R. Morandi, M. Picone, N. Bicocchi, M. Fogli and C. Giannelli

Future Generation Computer Systems, 2027

Cite this as

R. Morandi, M. Picone, N. Bicocchi, M. Fogli and C. Giannelli, “Digital Twin Entanglement: Enhanced measurement & reference benchmarking,” in *Future Generation Computer Systems*, vol. 188, March 2027, Art. no. 108808, doi: 10.1016/j.future.2026.108808.

Notice

© 2026. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Highlights

Digital Twin Entanglement: Enhanced Measurement & Reference Benchmarking

Riccardo Morandi, Marco Picone, Nicola Bicocchi, Mattia Fogli, Carlo Giannelli

- Formulates the Enhanced ODTE metric for granular, per-property synchronization in CPS.
- Introduces a Data Continuity dimension to detect and penalize out-of-sequence telemetry.
- Aligns digital evaluation with physical constraints by mapping RTOS priorities to weights.
- Proposes the deterministic Digital Twin Entanglement Benchmark (DTEB) as a reference baseline.

Digital Twin Entanglement: Enhanced Measurement & Reference Benchmarking

Riccardo Morandi^{a,*}, Marco Picone^b, Nicola Bicocchi^a, Mattia Fogli^c and Carlo Giannelli^c

^aDepartment of Engineering Enzo Ferrari, University of Modena and Reggio Emilia, Modena, 41125, Italy

^bDepartment of Science and Method for Engineering, University of Modena and Reggio Emilia, Reggio Emilia, 42121, Italy

^cDepartment of Mathematics and Computer Science, University of Ferrara, Ferrara, 44122, Italy

ARTICLE INFO

Keywords:

Digital Twin
Real-Time Systems
Synchronization Metrics
Telemetry Granularity
Digital Twin Entanglement

ABSTRACT

Digital Twins (DTs) are a fundamental paradigm for monitoring Cyber-Physical Systems (CPS). However, accurately quantifying the entanglement between a DT and its Physical Twin (PT) remains a critical challenge. In heterogeneous CPS, maintaining this alignment is highly complex: telemetry streams are generated at varying frequencies, reflect differing criticality levels, and face constant disruption from network latency and computational interference. Existing monolithic entanglement metrics, such as ODTE, fail to jointly capture these dynamics; by treating all data loss uniformly and ignoring ordering violations, they mask subtle state desynchronization. To resolve these limitations, we propose the Enhanced Overall Digital Twin Entanglement (E-ODTE) metric. E-ODTE explicitly models the fine-grained structure of CPS telemetry by decomposing the evaluation into a granular, per-property representation. Furthermore, it introduces a data continuity component to penalize out-of-sequence updates and incorporates semantic weighting to encode operational criticality. For rigorous validation, we introduce the Digital Twin Entanglement Benchmark (DTEB), a deterministic reference baseline founded on strict update alignment. E-ODTE is evaluated within a real-time stress-testing framework, mapping native Real-Time Operating System (RTOS) scheduling priorities directly onto its weighting structure. Experimental validation utilizing a Franka Emika Panda robotic setup controlled by a PREEMPT_RT kernel demonstrates that E-ODTE consistently and faithfully tracks the DTEB baseline, even under severe workload preemption. Beyond its effectiveness as a reliable surrogate metric, E-ODTE acts as a powerful diagnostic instrument, successfully exposing latent telemetry buffer anomalies that remain invisible to conventional evaluation approaches.

1. Introduction

Digital Twins (DTs) have emerged as a fundamental paradigm for the monitoring, analysis, and control of Cyber-Physical Systems (CPSs) [1–3]. More recently, research efforts have increasingly focused on the integration of DTs into complex and heterogeneous system architectures, including hard Real-Time (RT) environments [4]. In such settings, maintaining and quantifying a reliable operational coupling — formally defined as *entanglement* [5] — between the Physical Twin (PT) and its digital counterpart represents a critical challenge. Entanglement describes the degree to which the digital representation faithfully reflects the temporal evolution of the physical system. Because the PT is exposed to network variability, dynamic workloads, and stringent scheduling constraints [6], its telemetry generation processes are directly affected. Consequently, without a robust measure of entanglement, a DT may operate on stale or corrupted state representations, potentially leading to suboptimal or unsafe decision-making [7,8].

Section 2 surveys the relevant literature on DT synchronization and fidelity assessment. The Overall Digital Twin

Entanglement (ODTE) metric, detailed in Section 3, laid the conceptual groundwork for quantifying the quality of this entanglement [9]. However, the monolithic paradigm it relies upon is not well suited to the complex requirements of decoupled CPS architectures. First, traditional approaches like ODTE assume that telemetry is transmitted as a single, homogeneous unit, thereby neglecting the asynchronous nature of multi-property systems and the distinct frequencies of RT telemetry streams. Second, they lack an explicit mechanism to enforce temporal continuity: out-of-order deliveries, which are common in preemptive RT environments due to scheduling-induced jitter [10], may silently degrade the DT state without being detected by such metrics. Finally, this foundational framework assigns uniform significance to all data loss events, failing to differentiate between a delayed safety-critical update and a low-priority diagnostic signal. This approach assumes data loss is purely stochastic, ignoring that in hard RT environments, it is typically a deterministic consequence of resource-constrained scheduling. We note that multi-rate data handling, out-of-order detection, and criticality-aware weighting are, individually, well-established techniques in distributed systems and real-time scheduling; the contribution lies not in these mechanisms themselves, but in reconciling them — a non-trivial task in this context — into a single, per-property, RT-scheduling-aware metric, validated against a deterministic ground truth unmatched by existing DT-PT synchronization approaches.

To overcome these limitations, Section 4 introduces the Enhanced Overall Digital Twin Entanglement (E-ODTE)

*Corresponding author

 riccardo.morandi@unimore.it (R. Morandi);

marco.picone@unimore.it (M. Picone); nicola.bicocchi@unimore.it (N. Bicocchi); mattia.fogli@unife.it (M. Fogli); carlo.giannelli@unife.it (C. Giannelli)

ORCID(S): 0009-0001-3272-9600 (R. Morandi); 0000-0001-8902-6909 (M. Picone); 0000-0003-4182-1887 (N. Bicocchi); 0000-0002-9875-4550 (M. Fogli); 0000-0002-2394-1191 (C. Giannelli)

metric, specifically designed to bridge the gap between DT architectures and hard RT environments through property-level granularity, temporal continuity enforcement, and priority-weighted aggregation, directly mapping native Real-Time Operating System (RTOS) scheduling policies onto the metric's weighting structure. As a result, the DT effectively *inherits* the PT's operational hierarchy: the entanglement score is strictly penalized for the loss of high-priority updates, whereas the degradation of low-priority telemetry is treated as acceptable and consistent with the physical system's design intent.

While several works in the literature propose metrics to quantify DT timeliness or completeness [11–13], there is currently no standardized benchmark specifically designed to measure DT-PT entanglement under controlled experimental conditions. To address this gap, Section 5 introduces the Digital Twin Entanglement Benchmark (DTEB), a deterministic, update-alignment-based reference benchmark, which Section 6 then uses to validate E-ODTE through a real-world stress-testing campaign. Finally, Section 7 concludes the paper.

2. Related Work

DTs have been extensively studied as a fundamental paradigm for modeling, monitoring, and optimizing CPSs across domains such as industrial automation, robotics, and smart manufacturing [14,15]. Most existing works primarily focus on architectural design, data integration pipelines, cloud-to-edge offloading strategies, and application-driven use cases [16,17]. In these conventional settings, the quality of synchronization between the PT and its digital counterpart is often implicitly assumed rather than explicitly quantified. Domain-specific frameworks do define explicit synchronization steps — e.g., Shen et al. [18] propose an adaptive digital twin for power systems with an explicit initialization/synchronization phase. However, such approaches remain domain-specific and do not offer a general, quantitative, per-property, priority-aware entanglement metric applicable across heterogeneous CPS, which is the gap this paper addresses. The prevailing literature relies on best-effort communication, assuming network and control infrastructures can maintain accurate, timely representations of the physical system while neglecting scheduling effects and data heterogeneity inherent in modern distributed CPSs.

Network theory concepts like Age of Information (AoI) [11] and Quality of Service (QoS) are commonly used to evaluate DT fidelity, data freshness, and transmission latency [19]. Although AoI formally measures the time since the last update's generation, it remains inherently network-centric: it characterizes channel performance but misses the semantic coupling—or entanglement—between physical and digital states. Similarly, Quality of Data (QoD) frameworks assess sensor data accuracy and completeness [20,21], but predominantly operate on historical datasets, limiting their applicability to RT synchronization assessment. Consequently,

existing entanglement-evaluation approaches remain largely qualitative, theoretical, or scenario-specific.

This limitation is reflected in how DT literature treats data transmission, typically aggregating telemetry into monolithic payloads. While ensuring temporal consistency within a single update, this approach introduces significant serialization overhead and demands bandwidth that constrained CPS environments may lack. Conversely, multi-stream architectures alleviate payload bottlenecks but introduce challenges like network jitter, asynchronous delivery, and out-of-order reception [22]. Although stream processing literature acknowledges temporal disorder risks, existing DT synchronization models lack formal mechanisms to detect and penalize inconsistencies that silently degrade the reconstructed digital state.

Most DT frameworks also neglect data criticality, treating delayed or lost packets uniformly regardless of whether they carry safety-critical control signals or low-impact environmental measurements — a fundamental limitation when integrating DTs with RT systems. In contrast, RT literature emphasizes deterministic execution, schedulability analysis, and priority-driven preemption [23]: under transient overload, RT schedulers intentionally delay or drop lower-priority tasks to preserve critical control loops, introducing heterogeneous degradation patterns across properties.

A first significant attempt to formalize DT-PT entanglement beyond generic QoS metrics was the ODTE metric [9], a scalar synchronization measure effective for homogeneous telemetry. However, its formulation assumes all properties are equally relevant and synchronously updated as a unified dataset — an assumption that breaks down in complex, RT-constrained CPS environments with heterogeneous rates, criticality, and temporal constraints (detailed in Section 3).

To address this gap, recent work introduced a granular telemetry paradigm [24] in which PT properties are decoupled, transmitted independently, and assigned distinct priorities and frequencies; although this improves observability under RT stress, a formal quantitative framework for evaluating the resulting decoupled, priority-aware behavior is still lacking.

Finally, DT synchronization evaluation is further hindered by the absence of standardized benchmarking methodologies: existing studies rely on idealized simulations that miss real-world timing anomalies, or ad-hoc visual inspections that limit reproducibility and cross-system comparability [25,26], leaving no widely accepted data-driven benchmark for objectively assessing DT synchronization under realistic conditions.

None of the works discussed above jointly addresses the three structural gaps targeted by E-ODTE: as noted, AoI/AoII and QoD remain network-centric or historical-data-bound, offering no per-property granularity, out-of-order detection, or priority notion. Tan and Matta [12] model synchronization at a system level without per-property decomposition or RTOS-derived criticality, while Muñoz et al. [13] compare DT/PT traces offline without a granular, RT-computable score or continuity dimension. Finally,

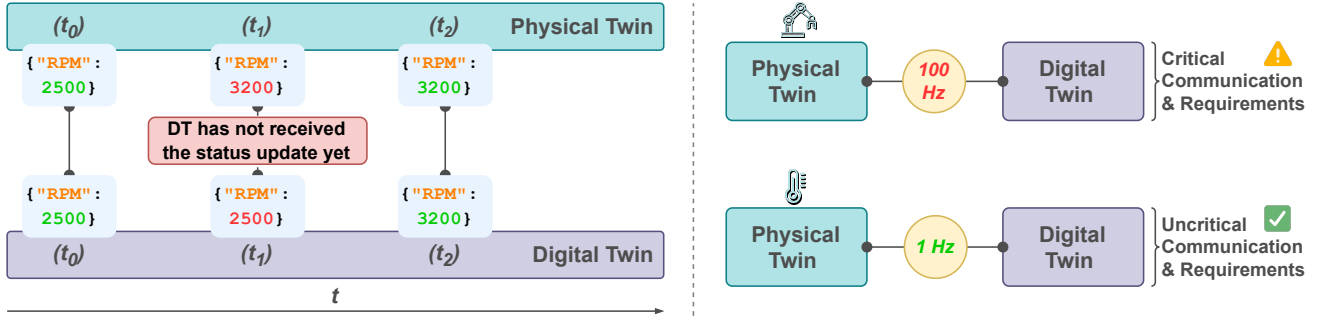


Figure 1: High-level representation of entanglement (left) and comparison between critical and non-critical scenarios (right).

Loubany et al. [27] confirm that a metric jointly combining these three dimensions is still missing from the literature.

This paper addresses these limitations by introducing a framework, inspired by the conceptual foundations of the ODTE metric [9], which incorporates property-level granularity, temporal continuity constraints, and semantic weighting grounded in RT scheduling priorities to enable quantitative evaluation across heterogeneous, RT-constrained CPSs. Furthermore, to mitigate the lack of standardized evaluation tools, we introduce a deterministic benchmark (DTEB) providing a reproducible reference baseline for consistent, mathematically grounded DT entanglement assessment in complex, resource-constrained environments.

3. ODTE

3.1. Overview

ODTE is a previously proposed metric [9] for quantitatively measuring the entanglement between a PT and its DT; we summarize its formulation below before introducing our enhancements in Section 4.

The left side of Figure 1 illustrates this concept. At time t_0 , the DT accurately mirrors the PT state. At t_1 , a delayed update causes temporary desynchronization. At t_2 , the delayed update arrives, restoring entanglement. Since external observers cannot directly perceive this internal synchronization state, ODTE provides a necessary quantitative evaluation.

The right side of Figure 1 shows varying telemetry regimes, emphasizing that entanglement quality depends on application-specific requirements, not just raw update frequency. Safety-critical systems demand strict temporal bounds, whereas less critical applications tolerate sporadic updates without degrading entanglement.

The original ODTE [9] accommodates these distinctions via a bidirectional approach over a discrete observation window. It relies on three core components: *Timeliness* (freshness of data/actions against thresholds), *Reliability* (completeness of communication versus expected volumes), and *Availability* (system up-time).

For the physical-to-digital direction, *Timeliness* (T) and *Reliability* (R) are defined as follows:

$$T^{PT \rightarrow DT} = \begin{cases} \frac{1}{\mu_{meas}} \sum_{i=1}^{\mu_{meas}} 1_{\{t_i^{DT} - t_i^{PT} + \Delta t_i^{exec} \leq T_d\}} & \text{if } \mu_{meas} > 0 \\ 0 & \text{if } \mu_{meas} = 0 \end{cases} \quad (1)$$

$$R^{PT \rightarrow DT} = \begin{cases} \frac{\mu_{meas}}{\mu_{exp}} & \text{if } \mu_{exp} > 0 \\ 0 & \text{if } \mu_{exp} = 0 \end{cases} \quad (2)$$

Equation 1 explicitly accounts for the DT's internal processing delay (Δt_i^{exec}). To ensure mathematical stability, both formulas safely evaluate to zero in degenerate cases (e.g., zero measurable or expected data points).

The digital-to-physical direction similarly models action execution and heartbeat-based *Availability*:

$$T^{DT \rightarrow PT} = \begin{cases} \frac{1}{\mu_{comp}} \sum_{i=1}^{\mu_{comp}} 1_{\{t_{i,a}^{DT} - t_{i,a}^{PT} + o_{i,a}^{PT \rightarrow DT} \leq T_{d,a}\}} & \text{if } \mu_{comp} > 0 \\ 0 & \text{if } \mu_{comp} = 0 \end{cases} \quad (3)$$

$$R^{DT \rightarrow PT} = \begin{cases} \frac{\mu_{comp}}{\mu_{req}} & \text{if } \mu_{req} > 0 \\ 0 & \text{if } \mu_{req} = 0 \end{cases} \quad (4)$$

$$A^{DT \rightarrow PT} = \begin{cases} \frac{h_{meas}}{h_{exp}} & \text{if } h_{exp} > 0 \\ 0 & \text{if } h_{exp} = 0 \end{cases} \quad (5)$$

As before, these ratios conservatively evaluate to zero when denominators are null. Aggregating these components yields the final holistic ODTE score.

3.2. Limitations

While foundational, the ODTE formulation faces three major challenges in hard RT environments.

First, it assumes uniform state processing against a single global *Timeliness* threshold. In reality, CPSs and Industrial Internet of Things (IIoT) applications involve independent sensor streams with vastly different sampling frequencies and latency requirements (e.g., millisecond-level control signals versus infrequent environmental updates).

Enhanced Feature	General CPS Challenge	Hard RT Challenge	E-ODTE Solution
Per-Property & Per-Action Granularity	Monolithic telemetry payloads mask delays in heterogeneous, asynchronous multi-sensor streams.	Rigid multi-rate task scheduling requires strictly enforced, distinct execution frequencies.	Deconstructs evaluation into independent streams, allowing specific constraints (e.g., T_d) for each property/action.
Data Continuity	Network jitter and distributed topologies frequently result in out-of-sequence data reception.	Workload-dependent preemption deliberately delays lower-priority tasks, causing sequence violations.	Introduces a dedicated dimension using incremental identifiers to actively detect and penalize out-of-order data processing.
Priority Weighting	Egalitarian evaluation treats all data loss equally, ignoring the varying semantic criticality of data.	RTOS schedulers perform graceful degradation, intentionally dropping lower-priority tasks under load.	Maps raw physical priorities directly into normalized metric weights, ensuring the DT inherits the PT operational hierarchy.

Table 1

Summary of how E-ODTE addresses general CPS and hard RT challenges through each enhanced feature.

Second, network jitter and RT task preemption frequently cause out-of-order data delivery. Preserving temporal continuity is essential; processing data non-chronologically, even within acceptable timing bounds, leads to incorrect state reconstruction and silent decoupling between the DT and PT.

Third, the original ODTE lacks a notion of data criticality, treating all losses equally. In hard RT systems, preemption is a structural necessity; schedulers deliberately drop lower-priority telemetry under high load to preserve critical control loops. Consequently, an egalitarian metric fails to reflect the actual severity of these disparate failure modes [27,28].

Consequently, overcoming these limitations demands a framework that supports a multi-property architecture with stream-specific constraints, guarantees temporal continuity, and incorporates priority-weighted evaluations to accurately reflect the PT's semantic hierarchy, combining established mechanisms into a single, DT-specific entanglement metric.

4. Enhanced ODTE

To address the limitations highlighted in Section 3.2, we propose the E-ODTE metric, whose goal is to achieve granular, property-aware telemetry. To this end, we introduce three key advancements: Section 4.1 defines the Per-Property, Per-Action dimension, enabling evaluation of independent telemetry streams under property-specific timing constraints; Section 4.2 introduces the Data Continuity dimension, enforcing temporal ordering and integrity of incoming sequences to prevent state corruption; and Section 4.3 presents the Priority-Weighted dimension, which accounts for the PT operational hierarchy by penalizing critical data loss more heavily than non-critical updates.

Table 1 summarizes these contributions and their mapping to CPS bottlenecks, highlighting their role in enabling hard RT applicability. Collectively, these enhancements ensure that the DT state remains a faithful and timely representation of the physical system under asynchronous, delayed, and out-of-order telemetry.

4.1. Per-Property & Per-Action Granularity

Figure 2 illustrates the dichotomy between traditional and heterogeneous CPS architectures. On the left, where

Symbol	Meaning	First used in
$T_{d,p}, T_{d,a}$	Per-property / per-action Timeliness threshold	Sec. 4.1
$\mu_{meas,p}, \mu_{exp,p}$	Measured / expected sample counts for property p	Sec. 4.1
$\mu_{comp,a}, \mu_{req,a}$	Completed / requested counts for action a	Sec. 4.1
$v_{i,p}, v_{i,a}$	Sample Integrity Score (physical-to-digital / digital-to-physical)	Sec. 4.1
R_{state}	State Reliability coefficient	Sec. 4.1
$\phi_{i,a}^{PT \rightarrow DT}$	Transmission time of the last property update completing action a	Sec. 4.1
N_p, N_a	Number of valid properties / actions for a given component	Sec. 4.1
$I_{(\cdot)}, L_{(\cdot)}$	Validity indicator / Loss term for a given component	Sec. 4.1
$\eta_{meas,p}, \eta_{meas,a}$	Sample / action count within the Continuity observation window	Sec. 4.2
id_i	Incremental identifier attached to each sample	Sec. 4.2
$w_{(\cdot)}, P_p, V$	Normalized priority weight, raw RTOS priority, set of valid properties/actions	Sec. 4.3

Table 2

Nomenclature of the new symbols introduced by E-ODTE relative to the original ODTE formulation.

multiple properties are bundled into a single monolithic frame, the original ODTE formulation behaves as intended. Conversely, on the right, where properties are decoupled into independent streams with heterogeneous frequencies, the original formulation breaks down. Lacking the granularity to evaluate independent data flows, a rigid global metric can obscure critical delays in individual variables behind the timely arrival of other data. To address this limitation, E-ODTE introduces a granular formulation, enabling per-property and per-action evaluation.

Table 1 summarizes these contributions and their mapping to CPS bottlenecks, highlighting their role in enabling

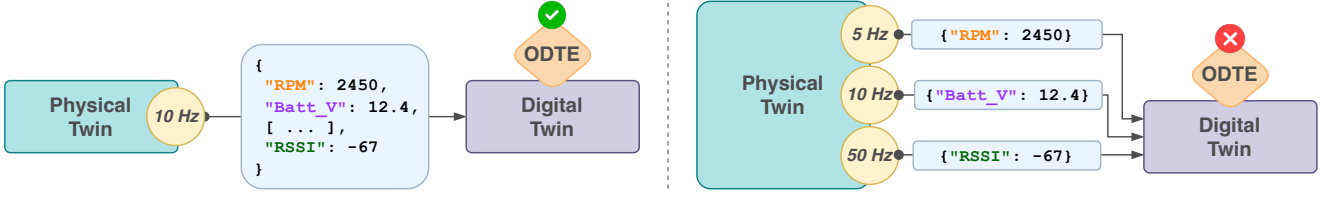


Figure 2: Monolithic data stream (left) compared to granular, multi-frequency transmission (right).

hard RT applicability. Collectively, these enhancements ensure that the DT state remains a faithful and timely representation of the physical system under asynchronous, delayed, and out-of-order telemetry. Table 2 summarizes the new notation introduced throughout this section for ease of reference.

4.1.1. Physical-to-Digital Granularity

Starting with *Timeliness*, the proposed formulation assigns each property its own threshold T_d , enabling support for heterogeneous timing requirements across different data streams. This is essential, as individual properties may exhibit distinct transmission delays, while state computation time can also vary significantly depending on the specific variable being processed. Consequently, *Timeliness* is redefined on a per-property basis as:

$$T_p^{PT \rightarrow DT} = \begin{cases} \frac{1}{\mu_{meas,p}} \sum_{i=1}^{\mu_{meas,p}} 1_{\{t_{i,p}^{DT} - t_{i,p}^{PT} + \Delta t_{i,p}^{exec} \leq T_{d,p}\}} & \text{if } \mu_{meas,p} > 0 \\ \text{undef} & \text{if } \mu_{meas,p} = 0 \end{cases} \quad (6)$$

By observing the proposed architecture of per-property *Timeliness*, we note a fundamental departure from the foundational monolithic approach presented in Equation 1. In particular, the computation is now performed on individual samples associated with a specific property, rather than on the aggregate set of all samples within the time window. Specifically, the summation of ones used to count the number of samples satisfying the *Timeliness* constraint is computed over a measurement window of size $\mu_{meas,p}$, which represents all data received within the time window for a given property. Furthermore, the threshold is now property-specific ($T_{d,p}$). This enables the physical-to-digital *Timeliness* to be evaluated strictly on a per-property basis, rather than yielding a single aggregated value for the overall E-ODTE. Furthermore, unlike the original formulation, if no messages are received for a given property within the time window, the metric is not assigned a value of zero. Instead, the per-property *Timeliness* is treated as undefined. This distinction is essential for the subsequent computation of the overall *Timeliness*: the absence of data for a single property within the time window does not necessarily imply that the physical-to-digital *Timeliness* should be penalized or reduced as a consequence.

Reliability undergoes a similar transformation: since different properties in real-world systems may exhibit distinct update frequencies, *Reliability* must be computed using property-specific expectations of updates within the time window. However, a purely quantitative approach based on

counting received packets is insufficient for capturing complex entanglement. A property update is only meaningful if the contained data is valid. Therefore, we incorporate a *sample integrity check* into the metric: a sample is considered valid if it satisfies predefined syntactic and semantic constraints, as formally defined in Equation Appendix A.1. The resulting per-property *Reliability* is defined as:

$$R_p^{PT \rightarrow DT} = \begin{cases} \frac{\sum_{i=1}^{\mu_{meas,p}} v_{i,p}}{\mu_{exp,p}} & \text{if } \mu_{exp,p} > 0 \\ \text{undef} & \text{if } \mu_{exp,p} = 0 \end{cases} \quad (7)$$

Here, $\mu_{meas,p}$ denotes the received samples for property p , and $v_{i,p}$ is the Sample Integrity Score. Thus, the numerator reflects valid, rather than merely received, samples. If expected samples are zero, per-property *Reliability* is undefined, though global *Reliability* remains computable. *Availability* is maintained at 1 for the physical-to-digital direction, requiring no per-property decomposition.

To reconstruct the comprehensive physical-to-digital E-ODTE, we must handle properties with undefined *Timeliness* or *Reliability*. We introduce indicator functions (Equations Appendix A.2 and Appendix A.3) that evaluate to 1 if a property is valid—i.e., if measured samples > 0 for *Timeliness*, and expected samples > 0 for *Reliability*. Summing these indicators yields the total valid properties for each component: $N_{T^{PT \rightarrow DT}}$ (Equation Appendix A.4) and $N_{R^{PT \rightarrow DT}}$ (Equation Appendix A.5).

Finally, to aggregate these heterogeneous values, we define per-property *Timeliness* Loss and *Reliability* Loss. These metrics quantify the degree to which a given property deviates from its optimal behavior (maximum achievable value):

$$L_{T_p^{PT \rightarrow DT}} := 1 - T_p^{PT \rightarrow DT} \quad (8)$$

$$L_{R_p^{PT \rightarrow DT}} := 1 - R_p^{PT \rightarrow DT} \quad (9)$$

Now we can define the global physical-to-digital *Timeliness* and *Reliability*:

$$T^{PT \rightarrow DT} = \begin{cases} 1 - \frac{\sum_{p=1}^{N_p} L_{T_p^{PT \rightarrow DT}}}{N_{T^{PT \rightarrow DT}}} & \text{if } N_{T^{PT \rightarrow DT}} > 0 \\ 0 & \text{if } N_{T^{PT \rightarrow DT}} = 0 \end{cases} \quad (10)$$

Quantity	Physical-to-Digital	Digital-to-Physical	Substitution
Per-unit Timeliness	$T_p^{PT \rightarrow DT}$ (Eq. 6)	$T_a^{DT \rightarrow PT}$ (Eq. 12)	$p \rightarrow a, \mu_{meas,p} \rightarrow \mu_{comp,a}, T_{d,p} \rightarrow T_{d,a}$
Per-unit Reliability	$R_p^{PT \rightarrow DT}$ (Eq. 7)	$R_a^{DT \rightarrow PT}$ (Eq. 13)	$\mu_{meas,p} \rightarrow \mu_{comp,a}, \mu_{exp,p} \rightarrow \mu_{req,a}, v_{i,p} \rightarrow v_{i,a}$
Validity indicators	Eq. Appendix A.2, Appendix A.3	Eq. Appendix A.8, Appendix A.9 (App. Appendix A)	same rule
Valid-unit totals	Eq. Appendix A.4, Appendix A.5	Eq. Appendix A.10, Appendix A.11 (App. Appendix A)	same rule
Loss terms	Eq. 8, 9	Eq. Appendix A.12, Appendix A.13 (App. Appendix A)	same rule
Global Timeliness / Reliability	$T^{PT \rightarrow DT}, R^{PT \rightarrow DT}$ (Eq. 10, 11)	$T^{DT \rightarrow PT}, R^{DT \rightarrow PT}$ (Eq. 14, 15)	same rule
Per-unit Continuity	$C_p^{PT \rightarrow DT}$ (Eq. 16)	$C_a^{DT \rightarrow PT}$ (Eq. 19)	$p \rightarrow a, \eta_{meas,p} \rightarrow \eta_{meas,a}$
Continuity validity, totals, Loss	Eq. Appendix B.1, Appendix B.2, Appendix B.3	Eq. Appendix B.4, Appendix B.5, Appendix B.6 (App. Appendix B)	same rule
Global Continuity	$C^{PT \rightarrow DT}$ (Eq. 17)	$C^{DT \rightarrow PT}$ (Eq. 20)	same rule
Weight normalization (T, R, C)	Eq. 22, 23, 24	Eq. Appendix C.1, Appendix C.2, Appendix C.3 (App. Appendix C)	$p \rightarrow a$
Weighted aggregation (T, R, C)	Eq. 25, 26, 27	Eq. Appendix C.4, Appendix C.5, Appendix C.6 (App. Appendix C)	same rule
Priority-to-weight mapping	Eq. 28 (shared formula, applies to both directions)		—

Table 3

Correspondence between physical-to-digital and digital-to-physical quantities across the Enhanced ODTE formulation (Section 4).

$$R^{PT \rightarrow DT} = \begin{cases} R_{state} \cdot \left(1 - \frac{\sum_{p=1}^{N_p} I_{R_p^{PT \rightarrow DT}}, L_{R_p^{PT \rightarrow DT}}}{N_{R^{PT \rightarrow DT}}} \right) & \text{if } N_{R^{PT \rightarrow DT}} > 0 \\ 0 & \text{if } N_{R^{PT \rightarrow DT}} = 0 \end{cases} \quad (11)$$

For *Timeliness*, the formula remains an average of the individual property Losses, considering only properties with valid indicators. For *Reliability*, however, we introduce the State *Reliability* coefficient (R_{state} , defined in Equation Appendix A.6), defined as the ratio between the number of successfully computed states and the number of expected state updates within the time window.

This formulation ensures that even if all individual properties are perfectly reliable, the global *Reliability* will be penalized if the DT model fails to synchronize them into a valid global state. Conversely, if the state computation is efficient but the incoming data is corrupted, the global metric will decrease accordingly. In both cases, if the number of valid properties is zero, the resulting *Timeliness* or *Reliability* is set to zero. Therefore, in the physical-to-digital scenario, while the E-ODTE framework preserves the conceptual spirit of holistic entanglement evaluation, it establishes a fundamentally distinct, fine-grained architecture driven by the adoption of per-property *Timeliness* and *Reliability*.

4.1.2. Digital-to-Physical Granularity

Table 3 summarizes the correspondence between the physical-to-digital quantities introduced above and their digital-to-physical counterparts, together with the substitution rule linking them; this correspondence is used throughout the remainder of this section to avoid repeating derivations that follow an identical structure.

Following this correspondence, the per-action *Timeliness* $T_a^{DT \rightarrow PT}$ mirrors $T_p^{PT \rightarrow DT}$ (Equation 6); unlike its physical-to-digital counterpart, the action-specific threshold $T_{d,a}$ was already present in the original ODTE formulation (Equation 3), though never evaluated at the per-action level. It is defined as:

$$T_a^{DT \rightarrow PT} = \begin{cases} \frac{1}{\mu_{comp,a}} \sum_{i=1}^{\mu_{comp,a}} 1_{\{t_{i,a}^{DT} - t_{i,a}^{PT} + o_{i,a}^{PT \rightarrow DT} \leq T_{d,a}\}} & \text{if } \mu_{comp,a} > 0 \\ \text{undef} & \text{if } \mu_{comp,a} = 0 \end{cases} \quad (12)$$

The term $o_{i,a}^{PT \rightarrow DT}$, with no physical-to-digital counterpart, denotes the transmission time of the final property update required to satisfy the completion criteria for action a : an action is considered *complete* only once the full set of mandatory property-change notifications has been dispatched to the DT.

Per-action *Reliability* $R_a^{DT \rightarrow PT}$ follows the same substitution (Table 3), integrating a Sample Integrity check into the command flow (Equation Appendix A.7) so that a completed action is credited only if it also passes physical safety validation at the PT:

$$R_a^{DT \rightarrow PT} = \begin{cases} \frac{\sum_{i=1}^{\mu_{comp,a}} v_{i,a}}{\mu_{req,a}} & \text{if } \mu_{req,a} > 0 \\ \text{undef} & \text{if } \mu_{req,a} = 0 \end{cases} \quad (13)$$

Here, $\mu_{req,a}$ is the number of times action a was requested by the DT, and $v_{i,a}$ is the Sample Integrity Score for the i -th completion.

Regarding digital-to-physical *Availability*, no modification is required. Its definition is based on a PT-level heartbeat signal; therefore, decomposing it into finer-grained components is not meaningful, as its semantics are inherently tied to the availability of the physical system as a whole.

The auxiliary validity indicators, valid-action totals, and Loss terms needed to aggregate *Timeliness* and *Reliability* across actions follow the same construction as their physical-to-digital counterparts (Table 3), and are given in full in Appendix Appendix A. The resulting global components are:

$$T^{DT \rightarrow PT} = \begin{cases} 1 - \frac{\sum_{a=1}^{N_a} I_{T_a^{DT \rightarrow PT}} L_{T_a^{DT \rightarrow PT}}}{N_{T^{DT \rightarrow PT}}} & \text{if } N_{T^{DT \rightarrow PT}} > 0 \\ 0 & \text{if } N_{T^{DT \rightarrow PT}} = 0 \end{cases} \quad (14)$$

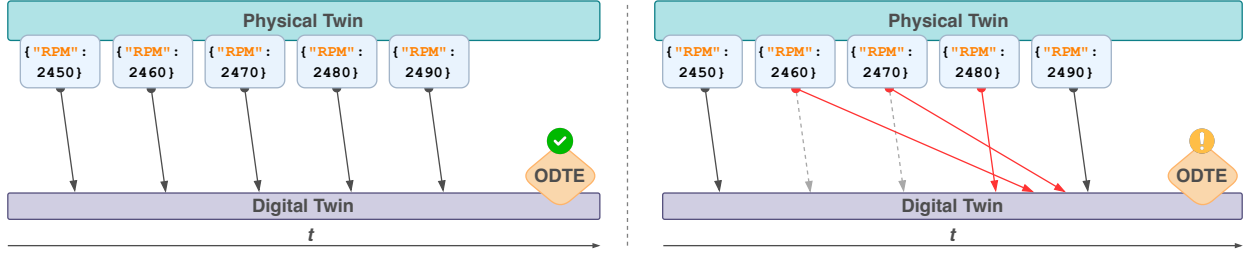


Figure 3: Data transmission and reception in the correct temporal order (left) vs. data received out of sequence compared to transmission (right).

$$R_{DT \rightarrow PT} = \begin{cases} R_{state} \cdot \left(1 - \frac{\sum_{a=1}^{N_a} I_{R_a^{DT \rightarrow PT}}, L_{R_a^{DT \rightarrow PT}}}{N_{R^{DT \rightarrow PT}}} \right) & \text{if } N_{R^{DT \rightarrow PT}} > 0 \\ 0 & \text{if } N_{R^{DT \rightarrow PT}} = 0 \end{cases} \quad (15)$$

As in the physical-to-digital direction, R_{state} (Equation Appendix A.6) acts as a common scaling factor across both directions, penalizing digital-to-physical *Reliability* whenever the DT fails to generate valid internal states.

4.2. Data Continuity Formulation

The second challenge concerns the temporal continuity of telemetry data, illustrated in Figure 3. While the left scenario depicts a standard, ordered transmission that enables the DT to maintain an accurate internal state, the right scenario shows the disruptive effect of out-of-sequence telemetry. As established, processing temporally disordered data causes the DT to compute an incorrect internal state, thereby silently decoupling it from the PT. Since the original formulation lacks a mechanism to detect such chronological violations, it fails to capture this anomaly and may incorrectly report a stable entanglement. To address this limitation, we introduce a dedicated dimension: the *Continuity* component. This metric is designed to monitor the temporal ordering of incoming data and adjust the overall E-ODTE whenever out-of-order delivery compromises the entanglement between DT and PT. Similarly to the other components, it represents a value ranging between 0 and 1, allowing for seamless integration with the existing metrics to calculate the total E-ODTE value. As we have remodeled the metric to function with individual properties and actions, Continuity has been modeled in the same manner, analyzing the continuity of data transmitted by a single property. Specifically, it is calculated over a discrete data window which is defined for each property.

4.2.1. Physical-to-Digital Continuity

The Data Continuity of a property, specifically in the physical-to-digital context, is defined as follows:

$$C_p^{PT \rightarrow DT} = \begin{cases} \frac{1}{\eta_{meas,p}} \sum_{i=1}^{\eta_{meas,p}} 1_{\{id_i=id_1+(i-1)\}} & \text{if } \eta_{meas,p} > 0 \\ \text{undef} & \text{if } \eta_{meas,p} = 0 \end{cases} \quad (16)$$

As observed, the structure of the formulation closely mirrors that of *Timeliness*. Here, $\eta_{meas,p}$ denotes the number of property updates collected for property p within the Continuity observation window; unlike the time-based window used for *Timeliness* and *Reliability* ($\mu_{meas,p}$), the Continuity window spans a fixed number of consecutive samples, which is why a distinct symbol is used. It is based on a summation of unitary contributions (i.e., ones) representing the number of property updates that satisfy the continuity constraint. To assess whether a property stream is continuous, each data sample must include an identifier, denoted as id , transmitted alongside the payload. This identifier is strictly incremental, enabling the comparison of consecutive samples to verify temporal ordering.

The identifier associated with the first property update within the observation window (i.e., the earliest datum in the discrete window) is assumed to be valid. Subsequent samples are then evaluated with respect to this reference, and any violation of the expected sequential ordering is excluded from the count of valid observations for the corresponding property, thereby reducing the resulting Continuity value. In the absence of any data within the discrete window, the Continuity for that property is defined as undefined. As with the previously introduced components, a validity indicator, a valid-property total, and a Continuity Loss term are required to aggregate Continuity across properties (Table 3, Equations Appendix B.1, Appendix B.2, Appendix B.3).

Now that all required components have been established, we proceed to calculate the global physical-to-digital Continuity, defined as per the following formula:

$$C^{PT \rightarrow DT} = \begin{cases} 1 - \frac{\sum_{p=1}^{N_p} I_{C_p^{PT \rightarrow DT}} L_{C_p^{PT \rightarrow DT}}}{N_{C^{PT \rightarrow DT}}} & \text{if } N_{C^{PT \rightarrow DT}} > 0 \\ 0 & \text{if } N_{C^{PT \rightarrow DT}} = 0 \end{cases} \quad (17)$$

Thanks to the introduction of Continuity, it is now possible to verify whether data arrives at the DT following the same temporal sequence in which it was generated on the

PT, an aspect that was previously unobservable within the original formulation. Consequently, the revised physical-to-digital E-ODTE, now incorporating the Continuity component, is defined as:

$$ODTE^{PT \rightarrow DT} = T^{PT \rightarrow DT} \times R^{PT \rightarrow DT} \times C^{PT \rightarrow DT} \times A^{PT \rightarrow DT} \quad (18)$$

4.2.2. Digital-to-Physical Continuity

Digital-to-Physical Continuity mirrors the physical-to-digital formulation above (Table 3), substituting properties with actions: an out-of-order completion acknowledgment can similarly lead the DT to infer an incorrect system state and generate suboptimal or erroneous decisions. Continuity is therefore defined on a per-action basis as:

$$C_a^{DT \rightarrow PT} = \begin{cases} \frac{1}{\eta_{meas,a}} \sum_{i=1}^{\eta_{meas,a}} 1_{\{id_i=id_1+(i-1)\}} & \text{if } \eta_{meas,a} > 0 \\ \text{undef} & \text{if } \eta_{meas,a} = 0 \end{cases} \quad (19)$$

The corresponding validity indicator, valid-action total, and Continuity Loss (Table 3, Appendix Appendix B) follow the same construction as their physical-to-digital counterparts. The global Continuity is:

$$C^{DT \rightarrow PT} = \begin{cases} 1 - \frac{\sum_{a=1}^{N_a} I_{C_a^{DT \rightarrow PT}} L_{C_a^{DT \rightarrow PT}}}{N_{C^{DT \rightarrow PT}}} & \text{if } N_{C^{DT \rightarrow PT}} > 0 \\ 0 & \text{if } N_{C^{DT \rightarrow PT}} = 0 \end{cases} \quad (20)$$

Finally, we redefine the calculation of the digital-to-physical E-ODTE as follows:

$$ODTE^{DT \rightarrow PT} = T^{DT \rightarrow PT} \times R^{DT \rightarrow PT} \times C^{DT \rightarrow PT} \times A^{DT \rightarrow PT} \quad (21)$$

Thanks to the integration of this new component into the metric, it is now possible to assess the continuity of data exchanged between the PT and DT, encompassing both the properties transmitted by the PT and the actions triggered by the DT.

4.3. Priority Weighting

The final critical dimension of CPS telemetry evaluation concerns data criticality and task priority. As illustrated in Figure 4, the inherently egalitarian nature of the original metric introduces a significant operational limitation. In its current form, the loss of a high-priority telemetry property (e.g., the reporting stream carrying critical motor temperature readings) reduces the total E-ODTE by the same amount as the loss of a low-priority telemetry property (e.g., the reporting stream for ambient humidity). We stress that this

priority ordering concerns telemetry/reporting tasks exclusively, i.e., the relative importance of different data streams feeding the DT; it does not extend to, nor compete with, the PT's own internal, safety-critical control loops. In our target architecture, control loops retain the highest RTOS scheduling priority unconditionally, while telemetry tasks — regardless of their assigned E-ODTE weight — are confined to a separate, strictly lower-priority band reserved for DT reporting. The weights introduced below are therefore only meaningful for comparisons within that telemetry band (e.g., a critical temperature reading versus a less relevant humidity reading), and never imply that a telemetry task should outrank a time-critical internal control loop. Given that the E-ODTE framework is inherently decomposed at the per-property level, integrating a weighting mechanism to penalize the loss of critical data more severely serves as a core architectural pillar. This weighted formulation more accurately reflects the operational and scheduling hierarchies of complex CPS environments, enabling the explicit prioritization of mission-critical telemetry properties relative to one another. We now define how the existing metric components are extended to incorporate priority information, beginning with the physical-to-digital context.

Given that each per-property component of the E-ODTE, as currently formulated, may yield an undefined value, and relying on the previously introduced notion of validity for inclusion in the computation, it is necessary to incorporate weighting factors during the aggregation of the various Loss components. Accordingly, we define a set of weights which, for all valid properties and for each specific metric component, are normalized such that their sum equals one. This normalization yields the following formulation:

$$\sum_{p=1}^{N_p} I_{T_p^{PT \rightarrow DT}} w_{T_p^{PT \rightarrow DT}} = 1 \quad (22)$$

$$\sum_{p=1}^{N_p} I_{R_p^{PT \rightarrow DT}} w_{R_p^{PT \rightarrow DT}} = 1 \quad (23)$$

$$\sum_{p=1}^{N_p} I_{C_p^{PT \rightarrow DT}} w_{C_p^{PT \rightarrow DT}} = 1 \quad (24)$$

A specific weight w is introduced for each component, since a property valid for *Timeliness* is not necessarily valid for *Reliability* or *Continuity*.

Having defined the weighting structure, we now proceed to specify how the individual components are aggregated:

$$T^{PT \rightarrow DT} = \begin{cases} 1 - \sum_{p=1}^{N_p} I_{T_p^{PT \rightarrow DT}} w_{T_p^{PT \rightarrow DT}} L_{T_p^{PT \rightarrow DT}} & \text{if } N_{T^{PT \rightarrow DT}} > 0 \\ 0 & \text{if } N_{T^{PT \rightarrow DT}} = 0 \end{cases} \quad (25)$$

$$R^{PT \rightarrow DT} = \begin{cases} 1 - \sum_{p=1}^{N_p} I_{R_p^{PT \rightarrow DT}} w_{R_p^{PT \rightarrow DT}} L_{R_p^{PT \rightarrow DT}} & \text{if } N_{R^{PT \rightarrow DT}} > 0 \\ 0 & \text{if } N_{R^{PT \rightarrow DT}} = 0 \end{cases}$$

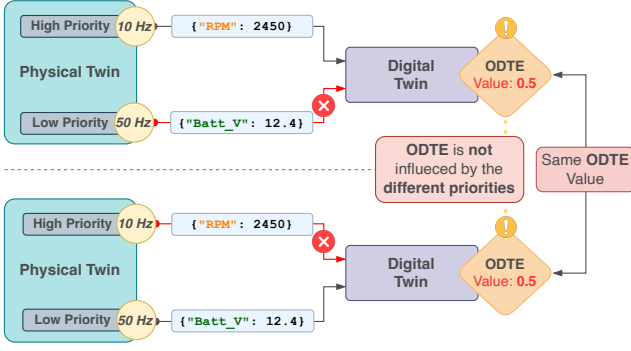


Figure 4: Visual demonstration of the ODTE metric inability to weight data loss based on property priority (high vs. low).

(26)

$$C_{PT \rightarrow DT} = \begin{cases} 1 - \sum_{p=1}^{N_p} I_{C_p^{PT \rightarrow DT}} w_{C_p^{PT \rightarrow DT}} L_{C_p^{PT \rightarrow DT}} & \text{if } N_{C^{PT \rightarrow DT}} > 0 \\ 0 & \text{if } N_{C^{PT \rightarrow DT}} = 0 \end{cases} \quad (27)$$

Each Loss term is now scaled by its corresponding normalized weight w , replacing the plain average used in the unweighted formulation with a weighted sum, so that each property contributes to the final E-ODTE score in proportion to its assigned relevance.

The digital-to-physical direction follows the same substitution (Table 3): actions executed by distinct RTOS-scheduled tasks are weighted analogously, with the corresponding weight-normalization (Equations Appendix C.1–Appendix C.3) and weighted-aggregation (Equations Appendix C.4–Appendix C.6) equations given in Appendix Appendix C. To operationalize this priority-aware evaluation, we define a direct mathematical mapping between task priorities and normalized E-ODTE weights. Assuming a standard control configuration, such as in an RTOS, where higher numerical values correspond to higher priorities, the raw priority P_p of a given property (or action) p is transformed into its normalized weight w_p as follows:

$$w_p = \frac{P_p}{\sum_{k \in V} P_k} \quad (28)$$

This objective mapping avoids arbitrary weight assignment and establishes a direct semantic alignment between the physical control logic and the digital evaluation model, ensuring the DT inherits the PT’s operational hierarchy: the loss of high-priority tasks is treated as a critical failure, while the loss of low-priority tasks is treated as acceptable degradation.

Through the combination of these enhancements, it is now possible to assess the entanglement between the PT and the DT with significantly greater fidelity. The E-ODTE metric now supports multiple independent properties and

actions, enforces strict temporal continuity of transmitted data, and inherently incorporates the priority hierarchy of the underlying physical architecture, ensuring that evaluation is both structurally consistent and operationally meaningful.

5. Digital Twin Entanglement Benchmark

The evaluation of entanglement between a DT and its PT is a fundamental challenge for RT-constrained DTs: without a deterministic, reproducible ground truth, a metric’s performance and robustness cannot be objectively verified or cross-compared. To address this gap, this paper introduces DTEB, a deterministic benchmark that produces a time-series representation of entanglement independently of the E-ODTE metric, enabling controlled experimentation and objective comparison. DTEB serves as an absolute ground-truth reference for calibration and validation, whereas E-ODTE is designed as a practical, deployable metric; the objective of defining DTEB is to establish this rigorous baseline for systematically validating each structural enhancement introduced in E-ODTE.

5.1. Design Principles

The DTEB requires the PT and DT to either run on the same physical machine or be strictly time-synchronized via hardware-level protocols, constraints that are often impossible to meet in distributed industrial CPSs. Thus, we use the DTEB strictly as a *calibration tool* to ensure that our deployable metric E-ODTE remains faithful to the physical reality even when such ideal conditions are absent.

To compute the entanglement level between the DT and the PT, state samples are collected simultaneously from both entities at a predetermined frequency (1 millisecond by default, to ensure fine-grained and consistent capture of all state transitions), including all properties exchanged between the two. For each property, an *Incremental Identifier* is generated on the PT side and mirrored by the DT, yielding two distinct identifiers $ID_{PT}^{(i)}(t)$ and $ID_{DT}^{(i)}(t)$, both functions of the sampling instant t and the specific property p_i to which they refer, and enabling precise verification of alignment between the PT and DT at a given time instant. Once the PT and DT datasets, containing their respective state trajectories over time, have been collected, the benchmark performs an intersection of the two sets, retaining only those samples that share the same timestamp (i.e., the same epoch). This operation aligns the datasets for subsequent comparison by removing non-matching data points. Given the high sampling frequency, this filtering step does not compromise the fidelity of the analysis.

To compare the states across the two datasets, it is necessary to define the condition under which one state is considered entangled with another. To this end, we first specify the alignment criterion at the level of individual properties, i.e., we define when a single property within a state is considered aligned with its corresponding counterpart in another state. This definition is formalized in the following expression:

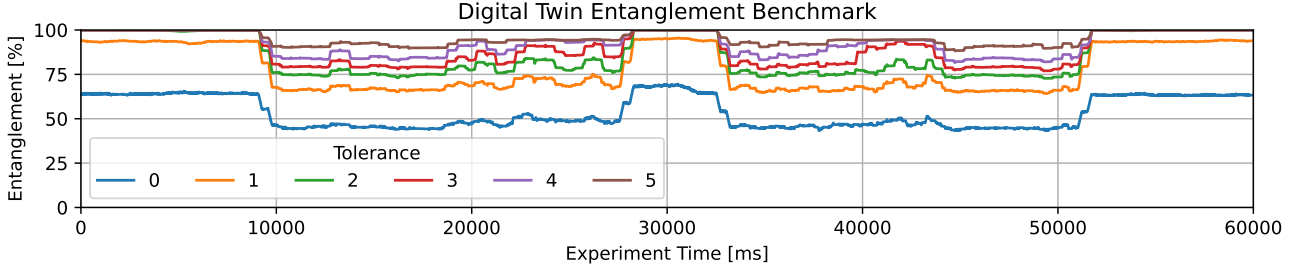


Figure 5: DTEB score [%] versus experiment time [ms] for different tolerance values.

$$\delta_i(t) = \begin{cases} 1 & \text{if } |ID_{PT}^{(i)}(t) - ID_{DT}^{(i)}(t)| \leq \tau, \\ 0 & \text{otherwise.} \end{cases} \quad (29)$$

The formula defines whether a single property in the PT is aligned with its corresponding property in the DT. This is evaluated by computing the difference between the PT identifier and the DT identifier. If the absolute difference is less than or equal to a predefined tolerance τ , the property is considered aligned (assigned a value of 1); otherwise, it is considered unaligned (assigned a value of 0). The parameter τ represents the alignment tolerance and directly determines the strictness of the benchmark. When $\tau = 0$, the formulation enforces strict alignment, fully disregarding any inherent temporal offset between the PT and DT; conversely, increasing τ relaxes this constraint, allowing for partial temporal misalignment to still be considered acceptable.

To extend this notion to an entire PT state at a given time instant t , the following formulation is used:

$$E(t) = \frac{1}{N} \sum_{i=1}^N \delta_i(t) \times 100. \quad (30)$$

The above expression computes the instantaneous entanglement at time t by aggregating the alignment indicators of all individual properties within the PT and DT states at that timestamp. The sum is then normalized by the total number of properties N shared between the PT and DT states. Finally, the result is multiplied by 100 to express the value as a percentage.

Having defined the instantaneous entanglement between a single PT state and its corresponding DT state at a given timestamp, we now introduce the formulation of the overall benchmark entanglement value:

$$E_{bench}(W) = \frac{1}{N_W} \sum_{t \in W} E(t) \quad (31)$$

As shown in the formulation, this value is computed over a defined time window W , enabling the aggregation and smoothing of instantaneous entanglement values across a specified observation period. The metric is obtained as the mean of the instantaneous values $E(t)$ for all samples within

the window. The window size is an arbitrarily configurable number of samples; by default, a window size of 1000 samples is used, which, given the 1-millisecond sampling frequency, corresponds to a 1-second time window. To complete the benchmark execution, once the window size has been defined, it is applied as a sliding window over the dataset. This procedure yields a time series representing the evolution of E_{bench} throughout the duration of the experiment.

5.2. Execution

Figure 5 presents the DTEB applied to the experimental architecture originally proposed in Morandi et al. [24]. While the full technical parameters of this infrastructure are extensively detailed in the upcoming Section 6, the setup briefly comprises a simulated 7-DOF Franka Emika Panda robot controlled by a strict PREEMPT_RT kernel via ROS 2, following the Software-in-the-Loop (SiL) paradigm detailed in Section 6, in which this simulated robot fulfills the role of the PT. During the task execution, synthetic computational workloads are intentionally injected into the RT controller to artificially stress the RT scheduler.

As expected, when the tolerance is set to zero, the entanglement value does not reach its maximum (100%) due to the inherent communication latency between the PT and the DT. Conversely, increasing the tolerance reveals operational regions in which maximum entanglement is achieved.

Across all curves, distinct intervals emerge where the entanglement sharply decreases. This behavior perfectly mirrors the actual conditions of the experiment: these drops correspond precisely to the high-load regions deliberately injected into the RT system, demonstrating the benchmark's sensitivity to task preemption and scheduling-induced delays under severe stress conditions.

The close match between E-ODTE and DTEB confirms the former as a reliable surrogate for the reference baseline, ensuring that, in real-world deployments where the deterministic DTEB cannot be executed, E-ODTE can still provide a high-fidelity estimate of entanglement.

6. Experimental Evaluation

This section validates the proposed metric by evaluating PT-to-DT synchronization in a simulated RT scenario. Consistent with the established SiL testing paradigm, the

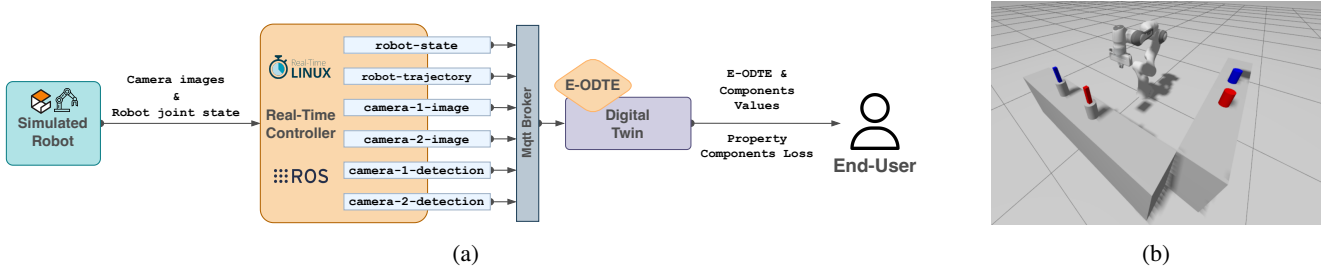


Figure 6: (a) Architectural representation and data flow of the experimental setup used for the physical-to-digital E-ODTE calculation. (b) Franka Emika Robot Simulation on Gazebo.

real-time control stack and RTOS scheduler are unmodified and real, while the physical process they control is provided by a physics-based simulation; hereafter, we refer to this simulated process as the simulated PT, consistent with its role within this SiL configuration. As illustrated in Figure 6, the PT comprises a 7-DOF Franka Emika Panda robot in Gazebo [29], communicating with the RT controller via ROS 2 [30].

The setup uses two identical workstations (Intel Core i7 4.2 GHz, 16 GB RAM) connected via dedicated Ethernet, assuming negligible packet loss. One runs the simulation, while the other runs Ubuntu with a PREEMPT_RT kernel [10], ensuring strict decoupling of the simulation environment and RT control logic.

The negligible-packet-loss assumption is a deliberate methodological choice: isolating degradation to RT scheduling effects allows the metric to be validated against the DTEB ground truth without conflating scheduling- and network-induced causes. This does not limit E-ODTE’s applicability to network-induced degradation, since from the DT’s perspective a lost or delayed update is indistinguishable whether caused by network loss or CPU preemption – both appear as a missing sample or broken identifier sequence, so Reliability and Continuity are, by construction, equally sensitive to both, even though the present experiments isolate only the scheduling-induced case. Instrumenting a PT with the telemetry needed to feed a DT also introduces computational load absent from the system’s original RT budget; the experiments quantify exactly this DT-induced overhead, a new interference source introduced by DT adoption itself rather than a pre-existing design concern.

During a pick-and-place task involving red and blue objects (Figure 6), we extract robot joint telemetry (position, velocity, effort) and video streams from two overhead cameras used for object localization.

To evaluate the metric under progressive performance degradation, a synthetic periodic workload is injected into the RT controller to artificially inflate CPU utilization. This workload is deliberately designed as a targeted stress test that exercises the specific structural mechanisms introduced by E-ODTE — per-property granularity, data continuity, and priority weighting — rather than as a general-purpose validation of DT robustness under arbitrary operational conditions. Without this controlled stress, the system would

property-name	priority (uniform)	priority (granular)	T_d [ms]	μ_{exp} [p/s]	γ [p]
robot-state	10	70	20	null	10
robot-trajectory	10	10	10	100	50
camera-1-image	10	20	20	30	10
camera-2-image	10	20	20	30	10
camera-1-detection	10	60	15	30	10
camera-2-detection	10	60	15	30	10

Table 4

Experimental parameters for E-ODTE calculation per property, under both the uniform-priority and granular-priority configurations.

remain nominal, and the metric would trivially stay at its maximum, preventing a meaningful assessment.

The RT controller generates multiple heterogeneous telemetry streams via ROS nodes (Figure 6). Each stream uses a dedicated thread, allowing independent configuration of execution priorities and frequencies.

These correspond to six heterogeneous telemetry topics: *robot-state*, an aperiodic JSON message (payload below 10 KB) reporting the current task phase; *robot-trajectory*, reporting per-joint position, velocity, and effort at 100 packets/s (payload below 10 KB); *camera-1-image/camera-2-image*, raw frames of the two stations at 30 packets/s (payload of a few tens of KB); and *camera-1-detection/camera-2-detection*, reporting detected object position/color at 30 packets/s (payload below 10 KB). All streams use JSON over MQTT, consistent with the parameters in Table 4.

Telemetry is transmitted via MQTT [31] to an EMQX broker. The Java-based DT, built with the WLDL library [32], processes these topics to reconstruct the simulated PT’s state and compute the E-ODTE metric. This allows fine-grained monitoring of overall scores, individual components, and per-property loss values. To assess metric robustness, a domain expert systematically calibrated and varied the following parameters across the experiments:

- Simulated PT task priority levels enforced by the RT controller;
- The threshold T_d for each property, used in the *Timeliness* computation;

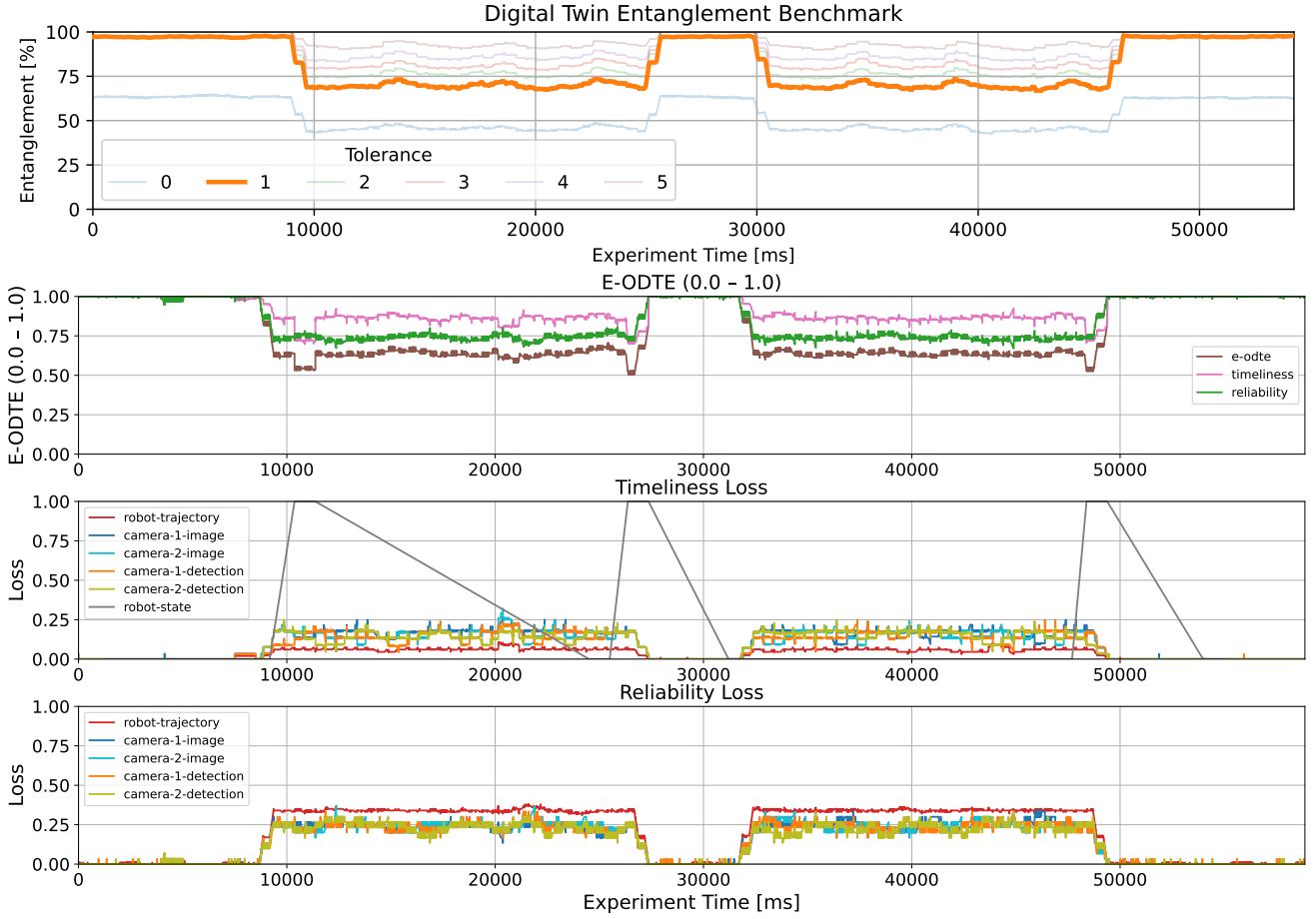


Figure 7: Temporal results in the absence of granular priorities for different properties. (a) DTEB [%] with 6 different levels of tolerance. (b) Top: E-ODTE scores. Middle: *Timeliness* Loss per property. Bottom: *Reliability* Loss per property.

- The expected mean value μ_{exp} for each property, used in the *Reliability* computation;
- The sliding window size γ for each property, used in the *Data Continuity* computation.

We note that the original ODTE metric is not directly benchmarked as an experimental baseline in this section. As illustrated in Figure 6, our architecture deliberately decomposes PT telemetry into six independent, heterogeneous-rate MQTT topics, whereas the original ODTE formulation presented in Section 3 assumes a single, homogeneous state update per observation. Applying ODTE to this testbed would require artificially re-aggregating all six streams into one synthetic monolithic message at each evaluation instant, which would both alter the system under test and reintroduce the very coarse-graining that E-ODTE is designed to overcome, making any resulting comparison self-serving rather than illuminating. ODTE is therefore used only as the conceptual and mathematical starting point for E-ODTE (Sections 3-4); its architectural incompatibility with a decoupled, multi-rate telemetry setup is itself additional evidence of the gap that E-ODTE addresses.

6.1. E-ODTE: Per-Property Granularity

The initial experiment was conducted with a specific focus on the physical-to-digital computation and property-level granularity. In this first experimental phase, the configuration of the individual properties was set according to the values reported in Table 4. Uniform priority values were assigned to all properties in this initial experimental phase. This configuration enables the evaluation of an RT system in which no differentiation is made among telemetry-related tasks, thereby allowing an assessment of whether the proposed metric is capable of capturing such homogeneous scheduling behavior.

Subsequently, distinct *Timeliness* thresholds (T_d , expressed in milliseconds) were defined for each property, depending on the nature and characteristics of the transmitted data. In particular, properties associated with higher payloads, such as *camera-1-image*, are assigned larger T_d values compared to lighter data streams, such as *robot-trajectory*, in order to account for their inherently higher transmission and processing latency requirements.

Furthermore, μ_{exp} was defined for each property, representing the expected number of packets within the observation time window. Since the time window is fixed at 1

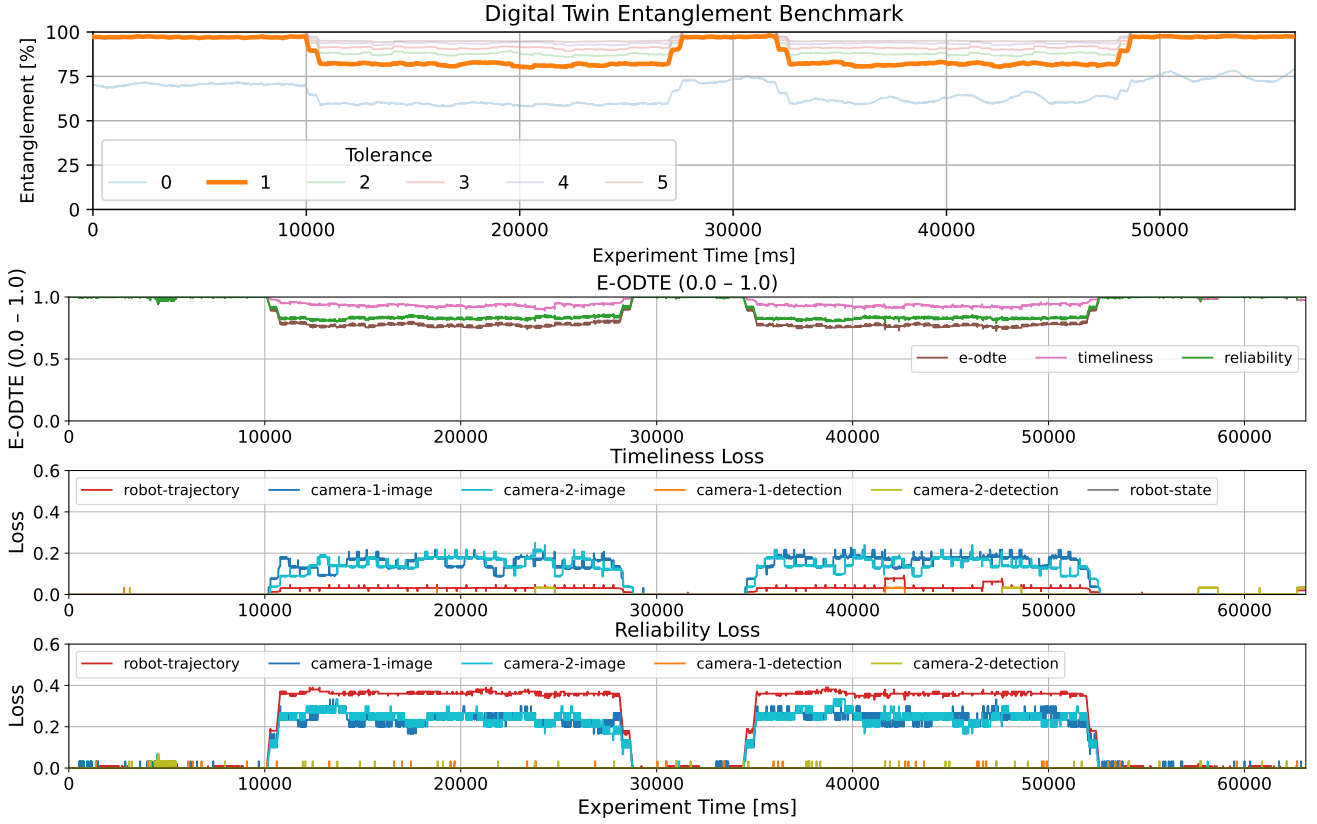


Figure 8: Temporal results following the assignment of granular priorities to distinct properties. (a) DTEB [%] with 6 different levels of tolerance. (b) Top: E-ODTE scores. Middle: *Timeliness* Loss per property. Bottom: *Reliability* Loss per property.

second across all experiments, μ_{exp} is expressed in packets per second (p/s), thereby reflecting the heterogeneous transmission rates associated with different data types. It is important to note that no μ_{exp} value was defined for the *robot-state* property, as this constitutes an aperiodic event. Consequently, the computation of *Reliability* based on a fixed expected packet rate is not analytically meaningful in this specific case.

Finally, the parameter γ was defined as the number of packets considered within the sliding window. In order to ensure an accurate reconstruction of the robot trajectory, and given that the *robot-trajectory* property exhibits a significantly higher packet generation rate compared to other streams, a larger γ value was assigned to this property. However, it is important to emphasize that this initial experimental phase adopts the E-ODTE formulation prior to the introduction of the *Continuity* component; therefore, the γ parameter does not affect the metric computation at this stage.

6.1.1. Flat vs. Granular Telemetry Task Prioritization

We analyze the metric using the uniform priority configuration from Table 4, where all RT tasks share equal criticality. Figure 7a shows the renewed DTEB results alongside the corresponding E-ODTE evaluation (Figure 7b), which includes *Timeliness* and *Reliability* losses.

As seen in the upper plot of Figure 7b, the E-ODTE curves align with the DTEB ground truth while yielding a more stable metric under the periodic workload. Between load injections, the E-ODTE fully recovers to 1, reflecting optimal nominal performance.

During load phases, uniform priorities cause all data streams to be preempted, dropping the total E-ODTE below 0.6. This degradation is primarily driven by a sharp decline in *Reliability*, as data transmission effectively ceases.

The lower plots break down these losses per property. While all properties suffer transmission delays, *Timeliness* loss peaks for the *robot-state* stream. Because *robot-state* is sporadic, its 1-second evaluation window often contains only a single sample; if this sample exceeds the T_d threshold, the resulting loss becomes maximal.

Conversely, *Reliability* degradation is dominated by the *robot-trajectory* stream. Its high transmission frequency leads to the largest absolute number of missed packets during preemption.

Finally, the metric enables quantitative evaluation of entanglement improvements when applying differentiated priorities. The granular-priority column of Table 4 assigns higher priorities to critical tasks like *camera-detection* and *robot-state*, while T_d , μ_{exp} , and γ remain unchanged.

Figure 8a shows the DTEB behavior under granular priority assignments. The overall entanglement increases

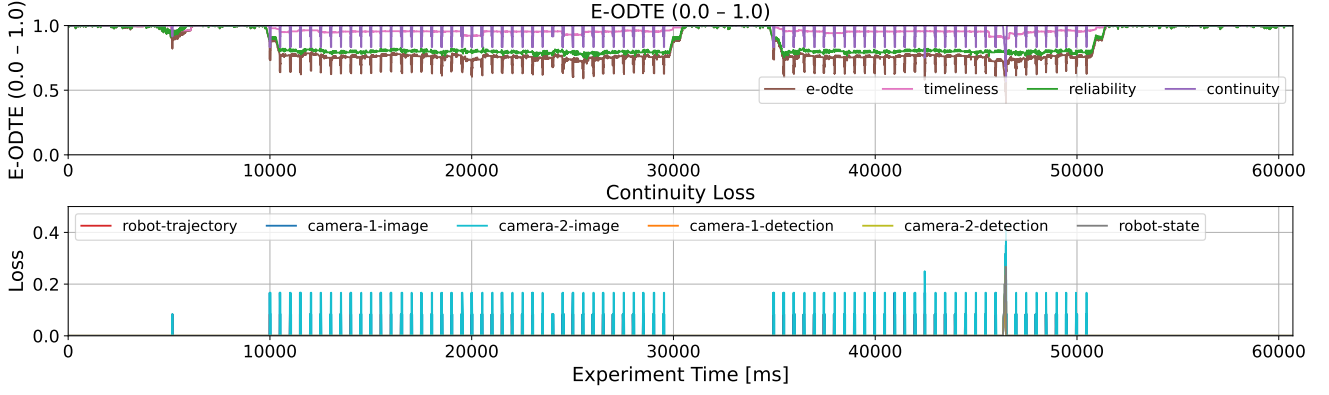


Figure 9: Temporal results following the introduction of the Continuity component. Top: E-ODTE scores (including Continuity). Bottom: Continuity Loss per property.

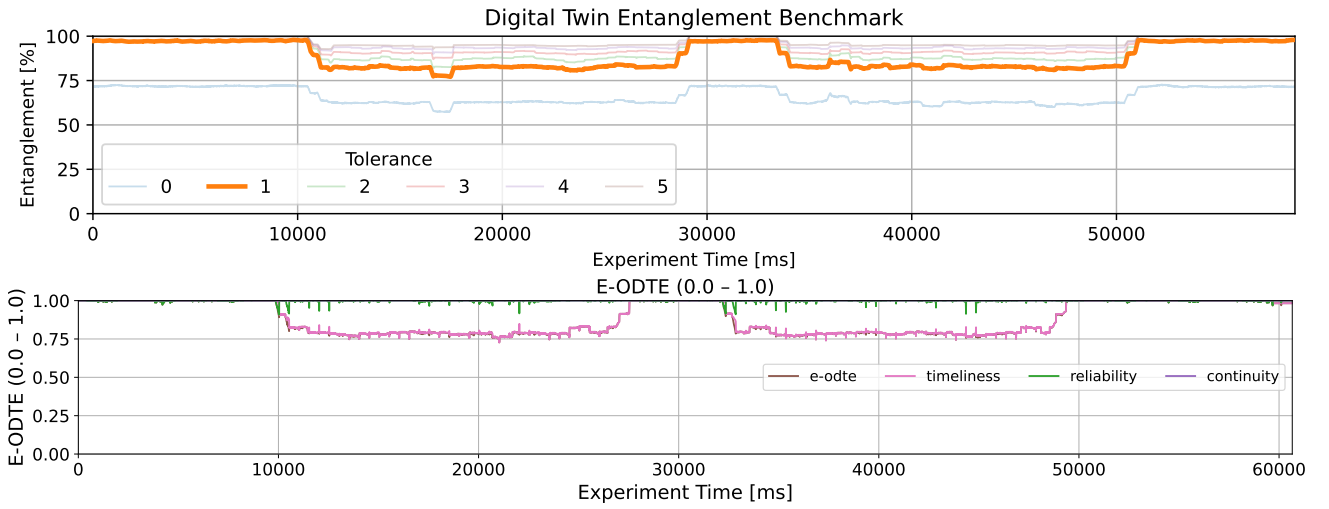


Figure 10: Temporal results following the resolution of the temporal continuity issue. (a) DTEB [%] with 6 different levels of tolerance. (b) E-ODTE scores.

significantly compared to the previous setup, demonstrating that differentiated priorities better maintain DT-PT synchronization.

Figure 8b plots the corresponding E-ODTE metric alongside *Timeliness* and *Reliability* losses. The metric closely tracks the benchmark, confirming E-ODTE reliably captures true entanglement across different operational conditions.

The top plot reveals that E-ODTE improves under load, stabilizing near 0.8—a significant recovery from the severe degradation observed in the uniform-priority scenario.

The lower plots detail per-property loss contributions. *Timeliness* degradation is now driven by the lower-priority *camera-image* and *robot-trajectory* tasks. Conversely, the highest-priority *robot-state* property no longer suffers significant *Timeliness* loss, proving that the RT scheduling policy effectively enforces its temporal constraints.

Similarly, *Reliability* losses are concentrated in the *camera-image* and *robot-trajectory* streams. Due to its high transmission frequency, *robot-trajectory* drops the highest absolute number of packets under load, dominating the overall *Reliability* degradation.

Ultimately, these results highlight how the proposed E-ODTE framework provides fine-grained visibility into individual property behaviors. Furthermore, experiments confirm that RT priority differentiation significantly improves PT-DT entanglement by safeguarding critical telemetry during computational overloads.

6.2. E-ODTE: Data Continuity

Next, we evaluate the *Continuity* component (Section 4.2) using the identical setup and parameters (Table 4). *Reliability* and *Timeliness* graphs are omitted here, as they remain identical to the previous iteration.

Figure 9 reveals anomalous Continuity fluctuations during load injections. Specifically, the Continuity Loss graph

demonstrates that all properties lose temporal ordering under high computational load on the simulated PT.

This anomaly, which would remain undetected without the E-ODTE Continuity metric, exposed a critical flaw in the RT controller's buffer logic. During preemption, unsent data accumulated in the buffer; upon resuming, the task transmitted the oldest data point but flushed the remainder, permanently breaking the temporal sequence.

Modifying the controller logic to transmit the complete buffered sequence resolved this issue. Figure 10a shows the DTEB after the fix. The total E-ODTE (Figure 10b) perfectly tracks the benchmark, validating the Continuity component.

While the overall E-ODTE behavior remains stable, the Continuity score is now fixed at 1, and the Continuity Loss is zero. This confirms that all properties strictly preserve their temporal ordering upon DT reception.

Ultimately, Continuity is a fundamental metric component; without it, silent temporal sequencing errors would cause flawed DT state calculations.

6.3. E-ODTE: Priority Weighting

The final experimental phase evaluates the definitive, priority-weighted E-ODTE formulation (Section 4.3). This iteration distinguishes critical high-priority failures from acceptable low-priority degradation, a crucial requirement for faithfully monitoring RT systems.

Component-specific weights are derived from the RT scheduling priorities in Table 4 using Equation 28, normalizing only across properties valid for each respective metric component.

Figure 11a presents the DTEB executed at a higher tolerance (level 2), necessary to fairly compare the egalitarian benchmark with the inherently higher scores of the priority-weighted E-ODTE (Figure 11b); under these conditions, the two remain perfectly aligned.

As shown in the top graph of Figure 11b, the overall E-ODTE score significantly improves, maintaining values around 0.9 even during critical loads. By heavily penalizing critical failures while scaling down the impact of preempted low-priority tasks (e.g., *robot-trajectory*), this weighted formulation provides a superior, realistic representation of the simulated RT PT.

The underlying component graphs confirm this behavior:

- **Timeliness Loss:** Reduced significantly, as delays predominantly affect low-priority properties with minimized weights.
- **Reliability Loss:** Improved overall; following the buffer fix, delayed data is rapidly transmitted post-preemption, eliminating previous loss spikes.
- **Continuity Loss:** Remains consistently zero, confirming strict temporal sequence preservation.

Ultimately, monitoring the E-ODTE metric enables a faithful, priority-aware evaluation of RT DT synchronization, successfully exposing and resolving underlying physical system anomalies.

6.4. Discussion & Open Challenges

While experimental results confirm the diagnostic power and operational fidelity of the E-ODTE, transitioning to this highly granular framework introduces specific architectural requirements and boundaries that must be considered before deployment.

The primary architectural constraint lies in the telemetry payload structure: per-property *Timeliness* requires the PT to embed an accurate generation timestamp into every update, while Continuity requires an incremental identifier (sequence number) within each packet. Consequently, E-ODTE cannot be seamlessly retrofitted as a “plug-and-play” observer for legacy, black-box CPSs that prohibit modifications to their data schemas. Relying on generation timestamps further necessitates strict clock synchronization between the PT and DT, since any uncompensated drift directly injects systemic errors into the physical-to-digital *Timeliness* score.

Per-property and per-action granularity also increases computational overhead on the DT, which must independently track, verify, sequence, and weight many distinct streams in RT rather than evaluating a single monolithic payload as in the original ODTE; scaling to massive Industrial IoT environments therefore requires careful optimization to avoid a system bottleneck. Our current WLDT-based Java implementation processes six MQTT topics per evaluation cycle with negligible overhead relative to the 1 ms DTEB sampling rate, requiring no dedicated optimization at this scale; however, this cost grows with the number of monitored properties and actions, and scaling to deployments with hundreds of properties would require batching or parallelizing these per-property computations to preserve real-time responsiveness.

Beyond these requirements, further open challenges remain. Measurement noise could impact the Sample Integrity checks underlying Reliability and Continuity, misclassifying valid updates, while a malicious actor could spoof or replay identifiers to defeat Continuity or tamper with timestamps to manipulate Timeliness, without distinguishing genuine network or scheduling effects; addressing this requires authentication and integrity-verification mechanisms, outside the present scope. Similarly, validation scope is limited: relying solely on synthetic CPU-bound preemption to stress the RT controller enables precise, reproducible control over the targeted failure modes, but not general operational relevance. Likewise, this work offers no direct quantitative comparison against the original ODTE, given the testbed's multi-rate architecture's incompatibility with its single-message formulation. Evaluating additional degradation modes — network-level packet loss/jitter, sensor/actuator faults, concurrent multi-modal failures — and constructing a monolithic-transmission testbed variant for a direct ODTE comparison are both left as concrete directions for future work; together, these three points detail in full the boundaries of the present experimental validation.

Ultimately, these constraints do not invalidate the metric but rather define its operational scope. They represent the architectural investments necessary to transition from a

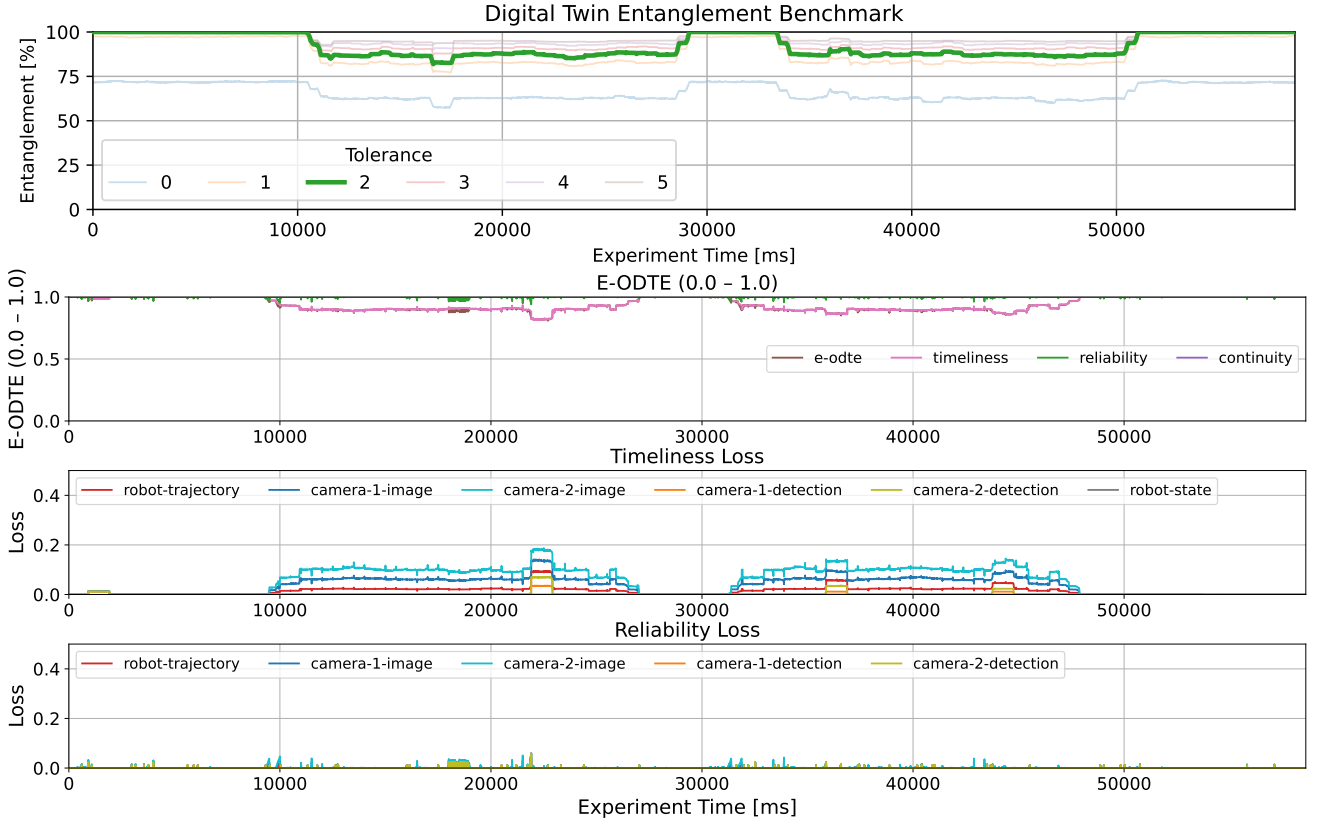


Figure 11: Temporal results adopting the priority-weighted E-ODTE calculation. (a) DTEB [%] with 6 different levels of tolerance. (b) Top: E-ODTE scores. Below (descending): *Timeliness* Loss and *Reliability* Loss per property.

generic, best-effort synchronization estimate to a rigorous, high-fidelity entanglement evaluation suitable for demanding RT architectures.

7. Conclusions & Future Works

In this paper, we introduced the E-ODTE metric, departing from traditional monolithic paradigms to establish a novel, granular framework specifically designed for modern cyber-physical systems. By incorporating property-level synchronization for independent multi-rate data streams, introducing a Temporal Continuity dimension to penalize out-of-order execution, and defining a semantic weighting mechanism grounded in physical criticality, the proposed metric effectively addresses the inherent complexities of heterogeneous telemetry.

These contributions were rigorously validated by explicitly mapping semantic weights to RTOS priority levels and by employing the newly introduced DTEB as a deterministic reference baseline. Experimental evaluations conducted on a PREEMPT_RT robotic platform demonstrate that the proposed metric behaves as a highly faithful surrogate of system entanglement and serves as a powerful diagnostic tool. In particular, it is capable of revealing subtle system-level anomalies, including buffering inefficiencies, while ensuring that digital *Reliability* consistently reflects the underlying physical behavior.

Future work will investigate phase-dependent evaluation strategies to dynamically adapt metric constraints over time, automated closed-loop policy mechanisms driven by RT E-ODTE feedback, and the broader applicability of the proposed framework in large-scale and highly distributed IIoT architectures.

CRedit authorship contribution statement

Riccardo Morandi: Writing – original draft, Validation, Software, Formal analysis, Conceptualization. **Marco Piconi:** Writing – review & editing, Validation, Formal analysis, Conceptualization. **Nicola Bicocchi:** Writing – review & editing, Validation, Formal analysis, Conceptualization. **Mattia Fogli:** Writing – review & editing, Validation, Formal analysis, Conceptualization. **Carlo Giannelli:** Writing – review & editing, Validation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Corresponding Author and Co-Authors know Dr. Marcello Pietri. Given his role as Executive Guest Editor, had no involvement in the peer review of this article and had no access to information regarding its peer review. Full

responsibility for the editorial process for this article was delegated to another journal editor – R.M., M.P., N.B.

Corresponding Author and Co-Authors know Dr. Matteo Martinelli. Given his role as Editor, had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor – R.M., M.P., N.B.

Co-Author knows Prof. Andreas Wortmann. Given his role as Editor, had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor – M.P.

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

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Per-Property & Per-Action Granularity

We define the Sample Integrity Score $v_{i,p}$ for the i -th sample of property p as:

$$v_{i,p} = \begin{cases} 1 & \text{if } \text{val}(i, p) \in \mathcal{D}_p \wedge \text{chk}(i, p) = \text{true} \\ 0 & \text{otherwise} \end{cases} \quad (\text{Appendix A.1})$$

Where $\text{val}(i, p)$ is the value contained in the i -th update of property p , $\mathcal{D}_p$ is the admissible domain (range) of the property, and $\text{chk}(i, p)$ represents a set of auxiliary integrity checks, such as checksum verification or data format consistency.

Validity indicators for *Timeliness* and *Reliability* Per-Property from Physical to Digital:

$$I_{T_p^{PT \rightarrow DT}} := 1_{\{\mu_{\text{meas},p} > 0\}} \quad (\text{Appendix A.2})$$

$$I_{R_p^{PT \rightarrow DT}} := 1_{\{\mu_{\text{exp},p} > 0\}} \quad (\text{Appendix A.3})$$

Total number of valid properties for *Timeliness* and *Reliability* from Physical to Digital:

$$N_{T^{PT \rightarrow DT}} := \sum_{p=1}^{N_p} I_{T_p^{PT \rightarrow DT}} \quad (\text{Appendix A.4})$$

$$N_{R^{PT \rightarrow DT}} := \sum_{p=1}^{N_p} I_{R_p^{PT \rightarrow DT}} \quad (\text{Appendix A.5})$$

The State *Reliability* coefficient (R_{state}) measures the DT modeling engine's efficiency in generating synchronized global states, an internal metric of the DT computational logic, unlike property *Reliability*, which is tied to the physical-to-digital data flow. It is defined as:

$$R_{\text{state}} = \begin{cases} \frac{\mu_{\text{state,comp}}}{\mu_{\text{state,exp}}} & \text{if } \mu_{\text{state,exp}} > 0 \\ \text{undef} & \text{if } \mu_{\text{state,exp}} = 0 \end{cases} \quad (\text{Appendix A.6})$$

Where:

- $\mu_{\text{state,comp}}$ is the number of global states successfully computed and updated by the DT within the time window.
- $\mu_{\text{state,exp}}$ is the number of expected state updates, determined by the DT synchronization frequency or the arrival of mandatory property sets.

Similarly, for the digital-to-physical flow, the Sample Integrity Score $v_{i,a}$ for the i -th completion of action a is defined as:

$$v_{i,a} = \begin{cases} 1 & \text{if } \text{exec}(i, a) = \text{success} \wedge \text{auth}(i, a) = \text{true} \\ 0 & \text{otherwise} \end{cases} \quad (\text{Appendix A.7})$$

Where $\text{exec}(i, a)$ verifies the action's execution within the PT's operational parameters, and $\text{auth}(i, a)$ confirms command authenticity and protocol compliance.

Validity indicators for *Timeliness* and *Reliability* Per-Action from Digital to Physical:

$$I_{T_a^{DT \rightarrow PT}} := 1_{\{\mu_{\text{comp},a} > 0\}} \quad (\text{Appendix A.8})$$

$$I_{R_a^{DT \rightarrow PT}} := 1_{\{\mu_{\text{req},a} > 0\}} \quad (\text{Appendix A.9})$$

Total number of valid properties for *Timeliness* and *Reliability* from Digital to Physical:

$$N_{T^{DT \rightarrow PT}} := \sum_{a=1}^{N_a} I_{T_a^{DT \rightarrow PT}} \quad (\text{Appendix A.10})$$

$$N_{R^{DT \rightarrow PT}} := \sum_{a=1}^{N_a} I_{R_a^{DT \rightarrow PT}} \quad (\text{Appendix A.11})$$

Loss values for *Timeliness* and *Reliability* Per-Action from Digital to Physical:

$$L_{T_a^{DT \rightarrow PT}} := 1 - T_a^{DT \rightarrow PT} \quad (\text{Appendix A.12})$$

$$L_{R_a^{DT \rightarrow PT}} := 1 - R_a^{DT \rightarrow PT} \quad (\text{Appendix A.13})$$

Appendix B. Data Continuity

Validity indicator for Continuity Per-Property from Physical to Digital:

$$I_{C_p^{PT \rightarrow DT}} := 1_{\{\eta_{meas,p} > 0\}} \quad (\text{Appendix B.1})$$

Total number of valid properties for Continuity from Physical to Digital:

$$N_{C^{PT \rightarrow DT}} := \sum_{p=1}^{N_p} I_{C_p^{PT \rightarrow DT}} \quad (\text{Appendix B.2})$$

Loss value for Continuity Per-Property from Physical to Digital:

$$L_{C_p^{PT \rightarrow DT}} := 1 - C_p^{PT \rightarrow DT} \quad (\text{Appendix B.3})$$

Validity indicator Per-Action, Total number of valid Actions and Loss value for Continuity from Digital to Physical:

$$I_{C_a^{DT \rightarrow PT}} := 1_{\{\eta_{meas,a} > 0\}} \quad (\text{Appendix B.4})$$

$$N_{C^{DT \rightarrow PT}} := \sum_{a=1}^{N_a} I_{C_a^{DT \rightarrow PT}} \quad (\text{Appendix B.5})$$

$$L_{C_a^{DT \rightarrow PT}} := 1 - C_a^{DT \rightarrow PT} \quad (\text{Appendix B.6})$$

Appendix C. Priority Weighting

Sum of the Weights for valid Actions for *Timeliness*, *Reliability* and Continuity from Digital to Physical:

$$\sum_{a=1}^{N_a} I_{T_a^{DT \rightarrow PT}} w_{T_a^{DT \rightarrow PT}} = 1 \quad (\text{Appendix C.1})$$

$$\sum_{a=1}^{N_a} I_{R_a^{DT \rightarrow PT}} w_{R_a^{DT \rightarrow PT}} = 1 \quad (\text{Appendix C.2})$$

$$\sum_{a=1}^{N_a} I_{C_a^{DT \rightarrow PT}} w_{C_a^{DT \rightarrow PT}} = 1 \quad (\text{Appendix C.3})$$

Aggregation of the Losses for *Timeliness*, *Reliability* and Continuity from Digital to Physical:

$$T^{DT \rightarrow PT} = \begin{cases} 1 - \sum_{a=1}^{N_a} I_{T_a^{DT \rightarrow PT}} w_{T_a^{DT \rightarrow PT}} L_{T_a^{DT \rightarrow PT}} & \text{if } N_{T^{DT \rightarrow PT}} > 0 \\ 0 & \text{if } N_{T^{DT \rightarrow PT}} = 0 \end{cases} \quad (\text{Appendix C.4})$$

$$R^{DT \rightarrow PT} = \begin{cases} 1 - \sum_{a=1}^{N_a} I_{R_a^{DT \rightarrow PT}} w_{R_a^{DT \rightarrow PT}} L_{R_a^{DT \rightarrow PT}} & \text{if } N_{R^{DT \rightarrow PT}} > 0 \\ 0 & \text{if } N_{R^{DT \rightarrow PT}} = 0 \end{cases} \quad (\text{Appendix C.5})$$

$$C^{DT \rightarrow PT} = \begin{cases} 1 - \sum_{a=1}^{N_a} I_{C_a^{DT \rightarrow PT}} w_{C_a^{DT \rightarrow PT}} L_{C_a^{DT \rightarrow PT}} & \text{if } N_{C^{DT \rightarrow PT}} > 0 \\ 0 & \text{if } N_{C^{DT \rightarrow PT}} = 0 \end{cases} \quad (\text{Appendix C.6})$$

Data availability

No data was used for the research described in the article.

References

- [1] V. Damjanovic-Behrendt, W. Behrendt, An open source approach to the design and implementation of digital twins for smart manufacturing, *International Journal of Computer Integrated Manufacturing* 32 (2019) 366–384. doi:10.1080/0951192X.2019.1599436.
- [2] D. Lehner, J. Pfeiffer, E.-F. Tinsel, M. M. Strlic, S. Sint, M. Vierhauser, A. Wortmann, M. Wimmer, Digital twin platforms: Requirements, capabilities, and future prospects, *IEEE Software* 39 (2) (2022) 53–61. doi:10.1109/MS.2021.3133795.
- [3] N. Mohammadi, J. E. Taylor, Smart city digital twins, in: 2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017, Honolulu, HI, USA, November 27 - Dec. 1, 2017, IEEE, 2017, pp. 1–5. doi:10.1109/SSCI.2017.8285439.
- [4] A. Alkhateeb, S. Jiang, G. Charan, Real-time digital twins: Vision and research directions for 6g and beyond, *IEEE Communications Magazine* 61 (11) (2023) 128–134. doi:10.1109/MCOM.001.2200866.
- [5] R. Minerva, G. M. Lee, N. Crespi, Digital twin in the iot context: a survey on technical features, scenarios, and architectural models, *Proceedings of the IEEE* 108 (10) (2020) 1785–1824.
- [6] J. C. Eidson, E. A. Lee, S. Matic, S. A. Seshia, J. Zou, Distributed real-time software for cyber-physical systems, *Proceedings of the IEEE* 100 (1) (2012) 45–59. doi:10.1109/JPROC.2011.2161237.
- [7] A. Maatouk, S. Kriouile, M. Assaad, A. Ephremides, The age of incorrect information: A new performance metric for status updates, *IEEE/ACM Transactions on Networking* PP (2020) 1–14. doi:10.1109/TNET.2020.3005549.
- [8] C. Schiavo, M. Favero, A. Buratto, L. Badia, Bidirectional age of incorrect information: A performance metric for status updates in virtual dynamic environments, 2025, pp. 365–372. doi:10.1109/MetaCom65502.2025.00062.
- [9] P. Bellavista, N. Bicocchi, M. Fogli, C. Giannelli, M. Mamei, M. Picone, Ode: A metric for digital twin entanglement, *IEEE Open Journal of the Communications Society* 5 (2024) 2377–2390. doi:10.1109/OJCOMS.2024.3385659.
- [10] W. Dewit, A. Paolillo, J. Goossens, A preliminary assessment of the real-time capabilities of real-time linux on raspberry pi 5, Working-paper, The 18th annual workshop on Operating Systems Platforms for Embedded Real-Time applications (2024).
- [11] R. D. Yates, Y. Sun, D. R. Brown, S. K. Kaul, E. Modiano, S. Ulukus, Age of information: An introduction and survey, *IEEE Journal on Selected Areas in Communications* 39 (5) (2021) 1183–1210.
- [12] B. Tan, A. Matta, The digital twin synchronization problem: Framework, formulations, and analysis, *IIEE Transactions* 56 (6) (2024) 652–665. doi:10.1080/24725854.2023.2253869.
- [13] P. Muñoz, M. Wimmer, J. Troya, A. Vallecillo, Measuring the fidelity of a physical and a digital twin using trace alignments, *IEEE Transactions on Software Engineering* 50 (12) (2024) 3122–3145. doi:10.1109/TSE.2024.3462978.
- [14] J. Leng, D. Wang, W. Shen, X. Li, Q. Liu, X. Chen, Digital twins-based smart manufacturing system design in industry 4.0: A review, *Journal of Manufacturing Systems* 60 (2021) 119–137.
- [15] Q. Qi, F. Tao, T. Hu, N. Anwer, A. Liu, Y. Wei, L. Wang, A. Nee, Enabling technologies and tools for digital twin, *Journal of Manufacturing Systems* 58 (2021) 3–21. doi:10.1016/j.jmsy.2019.10.001.
- [16] K. M. Alam, A. El Saddik, C2ps: A digital twin architecture reference model for the cloud-based cyber-physical systems, *IEEE access* 5 (2017) 2050–2062.
- [17] S. Ahelerooff, X. Xu, R. Zhong, Y. Lu, Digital twin as a service (dtaas) in industry 4.0: An architecture reference model, *Advanced Engineering Informatics* 47 (2021) 101225. doi:10.1016/j.aei.2020.101225.
- [18] Z. Shen, F. Arraño-Vargas, G. Konstantinou, Development of adaptive model-based digital twins for evolving power systems, *Computers*

- and Electrical Engineering 125 (2025) 110418. doi:10.1016/j.compeleceng.2025.110418.
- [19] M. A. Abd-Elmagid, N. Pappas, H. S. Dhillon, On the role of age of information in the internet of things, IEEE Communications Magazine 57 (12) (2019) 72–77. doi:10.1109/MCOM.001.1900041.
- [20] T. Peixoto, B. Oliveira, O. Oliveira, F. Ribeiro, Data quality assessment in smart manufacturing: A review, Systems 13 (4) (2025). doi:10.3390/systems13040243.
URL <https://www.mdpi.com/2079-8954/13/4/243>
- [21] M. Gómez-Omella, B. Sierra, S. Ferreiro, On the evaluation, management and improvement of data quality in streaming time series, IEEE Access 10 (2022) 1–1. doi:10.1109/ACCESS.2022.3195338.
- [22] S. Schöning, T. Küfner, R. Jasinski, A. Ermer, Vertical data continuity with lean edge analytics for industry 4.0 production, Computers in Industry 125 (12 2020). doi:10.1016/j.compind.2020.103389.
- [23] G. Buttazzo, Hard Real-Time Computing Systems: Predictable Scheduling Algorithms and Applications - Second Edition, Springer 2005, Vol. 24, 2011. doi:10.1007/978-1-4614-0676-1.
- [24] R. Morandi, M. Picone, N. Bicocchi, Real-time systems & digital twins: Exploring integration challenges and requirements, in: 2025 33rd International Conference on Enabling Technologies (WETICE), 2025, pp. 1–6. doi:10.1109/WETICE67341.2025.11091961.
- [25] S. Gil, P. Mikkelsen, C. Gomes, P. Larsen, Survey on open-source digital twin frameworks—a case study approach, Software: Practice and Experience 54 (2024) 929–960. doi:10.1002/spe.3305.
- [26] B. Lima, R. Pinto, Current challenges and future perspectives in testing iot systems: A comprehensive review, IEEE Sensors Reviews PP (2025) 1–27. doi:10.1109/SR.2025.3628264.
- [27] A. Loubany, M. Itani, S. Sharafeddine, From age of information to age of digital twin: A review on synchronization metrics for iot networks, IEEE Access PP (2025) 1–1. doi:10.1109/ACCESS.2025.3591589.
- [28] H. Farag, C. Stefanovic, Toward efficient deployment and synchronization in digital twins-empowered networks (2026). arXiv:2604.00566.
URL <https://arxiv.org/abs/2604.00566>
- [29] N. Koenig, A. Howard, Design and use paradigms for gazebo, an open-source multi-robot simulator, in: 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566), Vol. 3, 2004, pp. 2149–2154 vol.3. doi:10.1109/IROS.2004.1389727.
- [30] S. Macenski, T. Foote, B. Gerkey, C. Lalancette, W. Woodall, Robot operating system 2: Design, architecture, and uses in the wild, Science Robotics 7 (66) (May 2022). doi:10.1126/scirobotics.abm6074.
URL <http://dx.doi.org/10.1126/scirobotics.abm6074>
- [31] OASIS Standard, Mqtt version 5.0, <https://docs.oasis-open.org/mqtt/mqtt/v5.0/os/mqtt-v5.0-os.html> (2019).
- [32] M. Picone, M. Mamei, F. Zambonelli, Wldt: A general purpose library to build iot digital twins, SoftwareX 13 (2021) 100661. doi:<https://doi.org/10.1016/j.softx.2021.100661>.