

A Proposed Lossless ECG Compression Method Using Second-Order Delta Encoding and Arithmetic Coding on a Raspberry Pi-Based Embedded System

Jalu Wicaksana¹, Ahmad Sugiana², and Junartha Halomoan²

¹ Master's Program in Electrical Engineering, School of Electrical Engineering, Telkom University, Bandung, Indonesia

² School of Electrical Engineering, Telkom University, Bandung, Indonesia

Corresponding authors: Jalu Wicaksana (e-mail: jaluwicaksana@student.telkomuniversity.ac.id), Junartha Halomoan (e-mail: junartha@telkomuniversity.ac.id), **Author(s) Email:** Ahmad Sugiana (e-mail: sugianaa@telkomuniversity.ac.id)

Abstract Portable electrocardiogram (ECG) recorders have limited memory, yet their recordings must be stored without any loss of diagnostic information. This study aims to develop and validate a lossless ECG compression method simple enough for a low-cost embedded device. The main contribution is a compression method that combines second-order delta encoding with arithmetic coding, together with a mathematical rule that predicts when second-order delta encoding compresses better than first-order delta encoding. The method was implemented on a Raspberry Pi Zero 2 W with an AD8232 ECG sensor and an ADS1115 analog-to-digital converter, and it was compared with Huffman, Lempel-Ziv-Welch, and run-length coding, with and without delta encoding. The evaluation used hardware recordings from a single volunteer and 687 records from three public databases (MIT-BIH Arrhythmia, European ST-T, and PTB Diagnostic). All 8,244 tested combinations were reconstructed exactly, bit for bit. The proposed method achieved a mean compression ratio of 3.248 on the MIT-BIH database, whereas first-order delta encoding performed better on the European ST-T (2.969) and PTB (4.964) databases, as predicted by the derived rule. On hardware recordings, the compression ratio was about 2.0, and a standing posture raised 50 Hz power-line noise and lowered the compression ratio by up to 23%. Heartbeat detection on 745 beats showed 100% sensitivity and 100% positive predictivity, with no significant timing difference caused by 8-bit quantization ($p = 0.361$). In conclusion, the proposed method can roughly double the recording time of a memory-limited ECG device without additional hardware, which makes it a practical option for long-term wearable heart monitoring. Future work will extend the validation to multiple subjects and multi-lead recordings.

Keywords Lossless ECG compression, Second-order delta encoding, Arithmetic coding, Embedded system, Raspberry Pi

1. Introduction

Cardiovascular disease (CVD) remains the leading cause of death worldwide, with the World Health Organization reporting an estimated 19.8 million deaths attributable to CVD in 2022, approximately 32% of all global deaths [1]. Because many cardiac events occur suddenly and intermittently, continuous, long-term electrocardiogram (ECG) monitoring is essential for early detection, driving the development of portable, low-power ECG monitoring devices that operate outside conventional clinical settings [2]. Contemporary ECG monitoring architectures increasingly combine wearable sensing with Internet-of-Things connectivity,

which further increases the volume of data that must be stored or transmitted [3].

Such devices typically operate under significant resource constraints, most notably limited onboard storage capacity and battery life. Portable Holter-type monitors have historically limited continuous recordings to approximately 24-48 hours, after which a lower-fidelity looping-memory mode is often used, risking the loss of clinically significant event data [4], [5]. Data compression provides a practical means of extending effective storage capacity without added hardware cost. The role of compression in the large-scale transfer and storage of typical biomedical signals

Manuscript received 31 July 2026; Revised 16 September 2026; Accepted 30 September 2026; Available online 20 October 2026

DOI: <https://doi.org/10.35882/ijeeemi.v8i4.382>

Copyright © 2026 by the authors. Published by Jurusan Teknik Elektromedik, Politeknik Kesehatan Kemenkes Surabaya Indonesia. This work is an open-access article and licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

has also been examined in the hospital setting [6]. Because ECG signals carry direct diagnostic value, the method proposed in this paper is restricted to lossless compression, in which the reconstructed signal is guaranteed to be bit-exact -- a stronger guarantee than required by even the strictest diagnostic-grade accuracy standard, ANSI/AAMI EC11 [7].

Prior work has established that delta (differential) encoding combined with entropy coding is an effective lossless approach for ECG signals, and the SCP-ECG interchange standard mandates delta+Huffman encoding specifically for resource-constrained mobile devices [8]. Tiwari and Falk surveyed 31 lossless ECG compression studies, finding that the large majority relied on the MIT-BIH Arrhythmia Database for evaluation and that predictive (delta-based) pre-processing combined with entropy coding remains among the most computationally efficient approaches for resource-constrained devices [9]. Bannajak et al. proposed an adaptive multi-order linear prediction scheme combined with modified Golomb-Rice coding, evaluated across three databases with different acquisition characteristics, achieving compression ratios (CR) of 3.16-3.75 [10]. Directly relevant to the hardware platform used in this paper, Banerjee and Singh proposed a real-time lossless compression scheme combining second-order delta encoding with a buffer-array-based Run-Length Encoding variant, implemented and tested on an AD8232 sensor and Raspberry Pi hardware platform nearly identical to the one used in this study, achieving a compression ratio of 2.34 on self-collected single-channel data and up to 3.82-6.52 on public databases [11]. Tsai and Hussain presented a VLSI implementation of a lossless ECG compression algorithm targeting low-power devices [12], while Tsai and Kuo proposed an adaptive linear prediction scheme with content-adaptive Golomb-Rice coding specifically for embedded telemedicine platforms [13], and Deepu et al. integrated QRS detection with lossless compression directly on a 3-lead ECG-on-chip design for wireless sensor applications [14]. Campobello et al. proposed a low-memory lossless algorithm using a zero-order predictor and RAKE encoding suitable for resource-constrained microcontrollers, reporting compression ratios exceeding 2.67 with minimal computational overhead [15]. Rzepka proposed a low-complexity lossless multichannel ECG compression scheme based on selective linear prediction [16], and Mitra et al. recently combined an adaptive autoregressive model with a deep autoencoder, reporting very high compression ratios at a small but non-zero reconstruction error (PRD of 0.209%) [17]. Beyond predictive coding, wavelet-based approaches have also been extensively studied for ECG compression: Manikandan and Dandapat reviewed wavelet-based ECG compression methods

and their performance trade-offs [18], and Pal et al. proposed a tunable-Q wavelet transform combined with metaheuristic optimization to improve compression ratio, though both of these approaches are lossy rather than lossless [19]. Similarly, Devindi et al. reported a low-complexity real-time algorithm for low-power monitoring based on pulse-width-modulation-integrated quantization, which is also lossy [20]. At the systems level, Wasimuddin et al. surveyed ECG signal analysis approaches spanning traditional signal processing to machine learning [21], and Cuevas-Gonzalez et al. reviewed ECG data standards and formats for interoperability between mobile-health and clinical information systems, underscoring the continued relevance of compact, standards-compatible data representations such as the one proposed in this paper [22]. Jaleldine et al. established the foundational compression-ratio and distortion metrics still used throughout the ECG compression literature, including in this study [23]. Koski demonstrated an early lossless ECG encoding scheme combining linear prediction with entropy coding, establishing many of the design principles this paper builds upon [24]. Kerdjidi et al. [25] and Zaman and Morshed [26] further demonstrated compressed-sensing and low-power scanner architectures, respectively, for resource-constrained e-Health monitoring platforms, while an earlier ECG-on-chip design by Deepu et al. integrated a lossless data compressor consuming only 535 nW per channel, underscoring the practical importance of low computational and power overhead [27].

Several limitations nevertheless recur across this body of work. Adaptive, multi-order predictive schemes such as Bannajak et al.'s [10] achieve high compression ratios but require per-sample side-information overhead that is not directly comparable to a fixed-order predictor's cost. Hardware-validated implementations such as Banerjee and Singh's [11] re-engineer their entropy stage for, and report results on, a single fixed acquisition configuration, without deriving or empirically validating a criterion for when a given delta order will compress best across different sampling rates or correlation regimes. Custom-silicon implementations [12], [13], [14], [27] achieve very low power consumption but are evaluated only on a single database at a fixed configuration, and do not report cross-database or cross-sampling-rate validation. Most studies in this body of work, including [10], [11], [12], [13], [14], [15], do not isolate the pre-processing stage's own contribution to compression ratio via a controlled ablation, nor do they include an independent, QRS-detection-based diagnostic validation of the compressed and reconstructed signal. Wavelet-based approaches [18], [19], while competitive on compression ratio, are predominantly lossy, making

them unsuitable where a bit-exact diagnostic guarantee is required.

The research gap addressed here is therefore specific: no prior study combines a physically validated embedded hardware implementation, a mathematically derived and empirically validated criterion for selecting the delta order, a controlled ablation isolating the delta pre-processing stage's own contribution to compression ratio, and an independent, QRS-detection-based diagnostic validation of the reconstructed signal within a single evaluation framework. This study therefore aims to close that gap by proposing and comprehensively validating a simpler, fixed-order alternative to adaptive multi-order prediction second-order delta encoding combined with Arithmetic Coding on resource-constrained embedded hardware. The proposed method is validated systematically through a controlled ablation that isolates the pre-processing stage's contribution, a mathematically derived and empirically validated criterion for delta-order selection, cross-database evaluation under matched acquisition conditions, independent QRS-detection-based diagnostic validation, and a complete physical hardware implementation.

This paper makes four contributions: (1) the proposed Delta2+Arithmetic compression method, along with a mathematical rationale for each of its components; (2) an ablation study isolating the compression-ratio contribution of the delta pre-processing stage; (3) validation of the proposed method against the MIT-BIH Arrhythmia Database, the European ST-T Database, and the PTB Diagnostic ECG Database, including independent QRS-detection-based diagnostic validation; and (4) a complete physical hardware implementation on a Raspberry Pi Zero 2 W, confirming real-world deployability, together with a comparison against established Holter-monitoring storage practice.

The remainder of this paper is organized as follows. Section II describes the hardware, the proposed compression method, and the evaluation procedure. Section III presents the compression, diagnostic-validation, and computational-efficiency results. Section IV discusses these results, compares them with prior studies, and states the limitations. Section V concludes the paper and outlines future work.

II. Materials and Method

A. Hardware and Signal Acquisition

ECG signals were acquired using an AD8232 single-lead ECG front-end and an ADS1115 16-bit analog-to-digital converter, interfaced over I2C with a Raspberry Pi Zero 2 W. The ADS1115 was configured in continuous conversion mode rather than the default

single-shot mode, bringing the achieved sampling rate to within 0.1 Hz of the configured target. A three-electrode configuration (RA, LA, right-leg-drive) was used for common-mode noise rejection. The AD8232 front-end provides an on-chip second-order analog filter, whose output quality relative to higher-order digital filtering has been characterized previously [28].

Recordings were acquired under three configurations: a 200 Hz baseline across five sessions (1-15 minutes) for cross-session consistency analysis, and two acquisition parameter-matched single-session recordings -- 360 Hz/30 minutes (matching MIT-BIH) and 250 Hz/120 minutes (matching the European ST-T Database) -- for direct external comparison. A configuration matching the PTB Diagnostic ECG Database (1000 Hz) could not be acquired: the ADS1115 supports only fixed native data rates up to 860 SPS, and 1000 Hz exceeds this by a factor of $1000/860 = 1.163$ (16.3%), a hardware ceiling rather than a configuration limitation. PTBDB is therefore evaluated only at the external-database level.

Table 1. Key specifications of the hardware components used in the proposed ECG acquisition and compression system.

Component	Key Specification
AD8232 ECG front-end	Single-lead instrumentation amplifier, integrated right-leg-drive (RLD) circuit, lead-off detection
ADS1115 ADC	16-bit resolution, I2C interface, native data rates: 8/16/32/64/128/250/475/860 SPS
Raspberry Pi Zero 2 W	Quad-core ARM Cortex-A53 @ 1 GHz, 512 MB RAM, I2C host controller

Table 1 summarizes the key specifications of the three hardware components used to implement the proposed method. The ADS1115's fixed set of native data rates directly explains why the PTBDB-matched (1000 Hz) configuration could not be acquired: 1000 SPS falls between no two of the available native rates and exceeds the fastest one (860 SPS) outright.

B. Proposed Compression Method

Fig. 1 shows the complete system pipeline implemented in this study, from ECG acquisition through quantization, delta pre-processing, entropy coding and storage to decoding, signal reconstruction and evaluation; each stage is specified below. The acquired signal is linearly quantized to 8 bits per sample using min-max scaling. The proposed method applies second-order delta encoding, $d(n) = x(n) - [2x(n-1) - x(n-2)]$, computed using modulo-256 arithmetic to guarantee exact reversibility, followed by Arithmetic Coding [29]. The choice between first- and second-order delta pre-processing is governed by their relative residual variances. For a wide-sense-stationary signal with variance σ^2 , lag-1 autocorrelation ρ_1 , and lag-2 autocorrelation ρ_2 , the first-order residual

$D_1(n) = X(n) - X(n-1)$ has variance $Var(D_1) = 2\sigma^2(1 - \rho_1)$, while the second-order residual $D_2(n) = X(n) - 2X(n-1) + X(n-2)$ has variance $Var(D_2) = 2\sigma^2(3 - 4\rho_1 + \rho_2)$. Comparing these, $Var(D_2) < Var(D_1)$ -- i.e., second-order prediction reduces residual variance further -- exactly when $\rho_2 < 3\rho_1 - 2$.

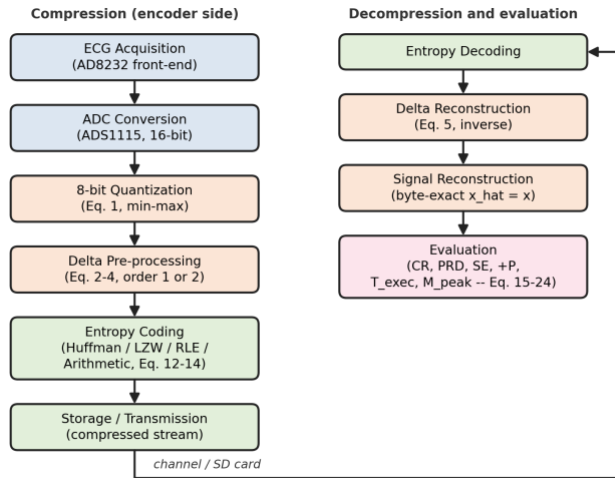


Fig. 1. Block diagram of the proposed ECG acquisition, compression, reconstruction, and evaluation system.

This signal-variance criterion is necessary but not alone sufficient to predict the better-compressing order in practice: mod-256 delta encoding also amplifies quantization noise, by a factor of 2 for first-order coefficients (1,-1) and by a factor of 6 for second-order coefficients (1,-2,1). Consequently, second-order delta pre-processing is expected to outperform first-order whenever its signal-variance advantage is large enough to outweigh this larger quantization-noise penalty, and to lose to first-order when the variance advantage is marginal or reversed.

Section III.B reports full-database autocorrelation measurements confirming this combined criterion correctly predicts the empirically observed winner

Algorithm 1. Proposed encoding procedure.

- (1) **Input:** quantized signal $x[0..N-1]$ (8-bit unsigned)
- (2) **Output:** header h , compressed stream c , compressed size
- (3) $h \leftarrow (x[0], x[1])$ $\triangleright$ stored raw
- (4) **FOR** $n = 2$ **TO** $N-1$ **DO**
- (5) $\text{pred} \leftarrow (2 \cdot x[n-1] - x[n-2]) \bmod 256$, Eq. (3)
- (6) $r[n-2] \leftarrow (x[n] - \text{pred}) \bmod 256$, Eq. (4)
- (7) **END FOR**
- (8) $c \leftarrow \text{ArithmeticEncode}(r)$ $\triangleright$ Section II.B
- (9) **RETURN** $h, c, \text{size}(c)$

across all three evaluation databases. Arithmetic Coding is selected over Huffman coding because it is not restricted to an integer number of bits per symbol: representing a sequence s as a sub-interval of width $P(s)$ requires $-\log_2 P(s)$ bits in total (plus a small constant, sequence-level overhead), so that the average code length per symbol converges to the source entropy rate $H(X)$ as sequence length grows, without the up-to-1-bit-per-symbol gap that bounds Huffman coding.

Algorithm 2. Proposed decoding procedure.

- (1) **Input:** header h , compressed stream c
- (2) **Output:** reconstructed signal $\hat{x}[0..N-1]$
- (3) $r \leftarrow \text{ArithmeticDecode}(c)$
- (4) $\hat{x}[0], \hat{x}[1] \leftarrow h$
- (5) **FOR** $n = 2$ **TO** $N-1$ **DO**
- (6) $\text{pred} \leftarrow (2 \cdot \hat{x}[n-1] - \hat{x}[n-2]) \bmod 256$
- (7) $\hat{x}[n] \leftarrow (r[n-2] + \text{pred}) \bmod 256$, Eq. (5)
- (8) **END FOR**
- (9) **RETURN** $\hat{x}$ $\triangleright$ byte-exact to x by construction

All arithmetic in Algorithms 1 and 2 is performed modulo 256, which guarantees exact reversibility regardless of the numeric sign or magnitude of the intermediate prediction error, since decoding simply reverses each modular addition with the identical predictor computed from previously reconstructed (and, by construction, identical to original) samples.

B. Mathematical Rationale for the Proposed Design

The choice of second-order delta encoding is grounded in the variance relations derived below. Writing $X(n)$ for the quantized sample sequence, σ^2 for its variance and ρ_1 for its lag-1 autocorrelation coefficient, the variance of the first-order residual $D_1(n) = X(n) - X(n-1)$ is given by Eq. (7). This shows that $Var(D_1) < \sigma^2$ whenever $\rho_1 > 0.5$, and $Var(D_1) \rightarrow 0$ as $\rho_1 \rightarrow 1$. Because the diagnostically relevant content of an ECG signal is concentrated below approximately 40 Hz while the signal is sampled at 200-1000 Hz in this study, adjacent samples are strongly correlated (ρ_1 close to 1), so delta pre-processing substantially reduces the residual variance -- and, since entropy for a fixed alphabet is maximized by the distribution of a given variance closest to Gaussian, this variance reduction directly bounds the achievable reduction in Shannon entropy prior to entropy coding. The full quantitative derivation underlying the proposed method is formalized below as Eqs. (1)–(14). Eqs. (7)–(11) are derived in this work from the variance of a linear combination of correlated samples; the remaining relations are standard results, cited where a source is available. The acquired signal is first linearly quantized to an 8-bit unsigned representation using min-max scaling over each recording, as given in Eq. (1):

$$q(n) = \text{round} \left[\frac{x(n) - x_{\min}}{x_{\max} - x_{\min}} (255) \right] \quad (1)$$

The first- and second-order delta residuals used in this study, and their corresponding predictors, are given in Eqs. (2)–(4) [8], [11], [24]:

$$d_1(n) = [X(n) - X(n-1)] \bmod 256 \quad (2)$$

$$\text{pred}_2(n) = 2x(n-1) - x(n-2) \quad (3)$$

$$d_2(n) = [x(n) - \text{pred}_2(n)] \bmod 256 \quad (4)$$

Because all arithmetic is performed modulo 256, the decoder reconstructs the signal exactly by inverting each modular addition using the identical predictor computed from previously reconstructed samples, as given in Eq. (5):

$$\hat{x}(n) = [d_k(n) + \text{pred}_k(n)] \bmod 256, \quad k \in \{1, 2\} \quad (5)$$

For a wide-sense-stationary source, the lag- k autocorrelation coefficient is defined in Eq. (6):

$$\rho(k) = \frac{\text{Cov}[X(n), X(n-k)]}{\text{Var}[X(n)]} \quad (6)$$

with $\rho_1 = \rho(1)$ and $\rho_2 = \rho(2)$ denoting the lag-1 and lag-2 coefficients used throughout this derivation. Using this notation, the first-order residual variance derived above is restated in Eq. (7):

$$\begin{aligned} \text{Var}(D_1) &= \text{Var}[X(n) - X(n-1)] \\ &= 2\sigma^2(1 - \rho_1) \end{aligned} \quad (7)$$

Because modulo-256 delta encoding also propagates quantization noise, and the first-order residual is a linear combination of samples with filter coefficients (1, -1), its quantization-noise amplification factor is the sum of the squared coefficients, as given in Eq. (8):

$$A_1 = 1^2 + [-1]^2 = 2 \quad (8)$$

The same variance-decomposition approach applied to the second-order residual $D_2(n) = X(n) - 2X(n-1) + X(n-2)$, whose filter coefficients are (1, -2, 1), gives, as derived in this work, Eq. (9):

$$\begin{aligned} \text{Var}(D_2) &= \text{Var}[X(n)] + 4\text{Var}[X(n-1)] \\ &\quad + \text{Var}[X(n-2)] - 4\text{Cov}[X(n), X(n-1)] \\ &\quad + 2\text{Cov}[X(n), X(n-2)] - 4\text{Cov}[X(n-1), X(n-2)] \\ &= \sigma^2(6 - 8\rho_1 + 2\rho_2) = 2\sigma^2(3 - 4\rho_1 + \rho_2) \end{aligned} \quad (9)$$

with quantization-noise amplification factor given by the sum of the squared second-order filter coefficients, as given in Eq. (10):

$$A_2 = 1^2 + (-2)^2 + 1^2 = 6 \quad (10)$$

Eq. (9) is consistent with two limiting cases: for an uncorrelated source ($\rho_1 = \rho_2 = 0$), $\text{Var}(D_2) = 6\sigma^2 = A_2\sigma^2$, matching the quantization-noise-only regime; for a perfectly linear local trend ($\rho_1 = \rho_2 = 1$), $\text{Var}(D_2) = 0$, confirming exact second-order predictability. Comparing Eqs. (7) and (9), the second-order residual has lower signal variance than the first-

order residual, $\text{Var}(D_2) < \text{Var}(D_1)$, exactly when the criterion in Eq. (11) is satisfied:

$$\rho_2 < 3\rho_1 - 2 \quad (11)$$

Eq. (11) is a purely theoretical criterion on the signal's own autocorrelation structure; this study does not separately estimate ρ_2 from the acquired recordings, and instead relies on the combined criterion below, which also accounts for the quantization-noise amplification factors A_1 and A_2 (Eqs. (8) and (10)), to interpret the empirically observed winner between first- and second-order delta pre-processing in Section III.B. The Shannon entropy of the (delta-encoded) residual stream, treated as a discrete source with alphabet $\{0, \dots, 255\}$ and empirical symbol probabilities $p(0), \dots, p(255)$, is given in Eq. (12) [30]:

$$H(X) = -\sum_i p(x_i) \log_2 p(x_i) \quad (12)$$

The entropy is reported throughout this study in bits per symbol (bit/symbol). A canonical Huffman code achieves an average codeword length L bounded as given in Eq. (13) [30], [31]:

$$H(X) \leq L < H(X) + 1 \quad (13)$$

whereas Arithmetic Coding represents an entire sequence as a sub-interval of (0,1), requiring $-\log_2 P(s)$ bits for a sequence of probability $P(s)$, so that by the Asymptotic Equipartition Property for a stationary source the average code length per symbol converges to the entropy, as given in Eq. (14) [29], [30]:

$$\lim_{n \rightarrow \infty} \frac{L_{\text{code}}(n)}{n} = H(X) \quad (14)$$

without the up-to-1-bit-per-symbol gap of Eq. (13), which is the mathematical basis for selecting Arithmetic Coding as the entropy coder in the proposed method. To isolate and quantify the contribution of the proposed method's two components, three configurations are compared: no pre-processing (entropy coding applied directly to the quantized stream), first-order delta encoding, and the proposed second-order delta encoding, each combined with Huffman coding, LZW coding [32] [33], Run-Length Encoding, and Arithmetic Coding in turn. Losslessness is verified on a byte-exact basis for every combination.

C. External Database Comparison and Diagnostic Validation

The proposed method was applied to the complete MIT-BIH Arrhythmia Database (MITDB, 48 of 48 records) [34], European ST-T Database (EDB, 90 of 90 records) [35], and PTB Diagnostic ECG Database (PTBDB, 549 of 549 records) [36] -- full-database evaluation rather than a fixed-size subset, eliminating subset-dependent conclusions. Records were accessed via the PhysioNet [37] wfdb package; EDB records were truncated to 30 minutes for computational tractability. The full-database scope above applies to

the compression-ratio evaluation. QRS-detection validation is computationally far more expensive per record and was therefore performed on a randomly selected 10-record subset of MITDB and of EDB (23,343 and 20,903 annotated beats respectively), which is stated explicitly wherever those results are reported. QRS complex detection was performed using the Pan-Tompkins algorithm [38] (wfdb.processing.XQRS implementation) on both the original and the reconstructed signal, and Sensitivity (SE) and Positive Predictivity (+P) were computed relative to each database's reference beat annotations using a ± 150 ms tolerance window. Because comparing the original signal against the reconstructed signal is trivially identical under a verified-lossless codec, a second, more diagnostic comparison was additionally performed directly on hardware-acquired recordings: QRS detection on the original, pre-quantization signal against the 8-bit quantized signal -- the one processing step in the pipeline that is not bit-exact reversible -- reported in Section III.E below with confidence intervals and paired statistical tests.

III. Results

A. Contribution of Delta Pre-processing

Throughout Section III, compression ratio is defined in Eq. (15) [23] as

$$CR = \frac{S_{original}}{S_{compressed}} \quad (15)$$

computed on a byte-exact basis for every combination evaluated (Section II.B). Because every evaluated combination is lossless, the percent root-mean-square difference between original and reconstructed signal, given by Eq. (16) [23],

$$PRD = \sqrt{\frac{\sum_n [x(n) - \hat{x}(n)]^2}{\sum_n x(n)^2}} (100\%) \quad (16)$$

Table 2. Compression-ratio improvement obtained using second-order delta preprocessing compared with no preprocessing for Huffman and Arithmetic Coding.

Data source	NoDelta → D2+Huffman	NoDelta → D2+Arithmetic
Hardware, 360 Hz	1.231 → 1.975	1.239 → 2.011
MITDB	1.416 → 3.181	1.492 → 3.248
Hardware, 250 Hz	1.234 → 1.786	1.241 → 1.942
EDB	1.499 → 2.722	1.571 → 2.779

is identically $PRD = 0\%$ by construction ($\hat{x}(n) = x(n)$ for all n), which is precisely what the byte-exact verification reported throughout Section III confirms empirically. Table 2 compares the compression ratio achieved with no pre-processing against the proposed second-order delta encoding, for the Huffman and Arithmetic coders. Database rows are full-database means (48/48 MITDB and 90/90 EDB records); hardware rows are the matched single-session

recordings. Each cell lists the compression ratio without and with pre-processing. PTBDB is not listed because no matching hardware recording exists. The proposed pre-processing stage provided a substantial gain in every case (+45% to +125%), with a markedly larger gain on the clean external databases than on the hardware-acquired recordings, indicating that its effectiveness is sensitive to acquisition noise.

B. Validation Against External Reference Databases

Table 3. Mean compression ratio achieved by the proposed Delta2+Arithmetic method on hardware-acquired ECG signals and external ECG databases.

Data source	Mean CR
Hardware, 360 Hz (matched to MITDB)	2.011
MITDB (external)	3.248
Hardware, 250 Hz (matched to EDB)	1.942
EDB (external)	2.779
PTBDB (external only)	4.231

Table 3 summarizes the mean compression ratio obtained with the proposed Delta2+Arithmetic configuration on each data source, computed over the complete databases. The highest value is 4.231 on PTBDB and the lowest 1.942 on the 250 Hz hardware recording, a 2.18-fold spread across sources. All combinations evaluated were independently verified to be byte-exact lossless. Full-database re-evaluation (48/48 MITDB, 90/90 EDB, 549/549 PTBDB records; 8,244 method-record combinations) shows that the best-performing delta order is database-dependent rather than uniformly second-order: Delta2+Arithmetic achieved the highest mean CR on MITDB (3.248, SD 0.365), while Delta1+Arithmetic outperformed Delta2+Arithmetic on both EDB (2.969 vs. 2.779) and PTBDB (4.964 vs. 4.231). Section II.B derives the criterion governing this pattern: Delta2 reduces residual variance relative to Delta1 when $\rho_2 < 3\rho_1 - 2$, but second-order delta pre-processing also amplifies quantization noise three times more than first-order (coefficient sum of squares 6 vs. 2). Measured full-database autocorrelation confirms the combined criterion predicts the observed winner in all three databases: MITDB ($\rho_1 = 0.985, \rho_2 = 0.943$, criterion satisfied with wide margin) sees Delta2 win; EDB ($\rho_1 = 0.969, \rho_2 = 0.896$, criterion satisfied only marginally) and PTBDB ($\rho_1 = 0.996, \rho_2 = 0.993$, criterion not satisfied) both see the quantization-noise penalty outweigh Delta2's marginal or absent variance advantage, so Delta1 wins. This is reported as a substantive finding: the optimal delta order is signal-dependent and predictable from measured autocorrelation, rather than a fixed universal choice. In practice, the delta order can be selected per recording.

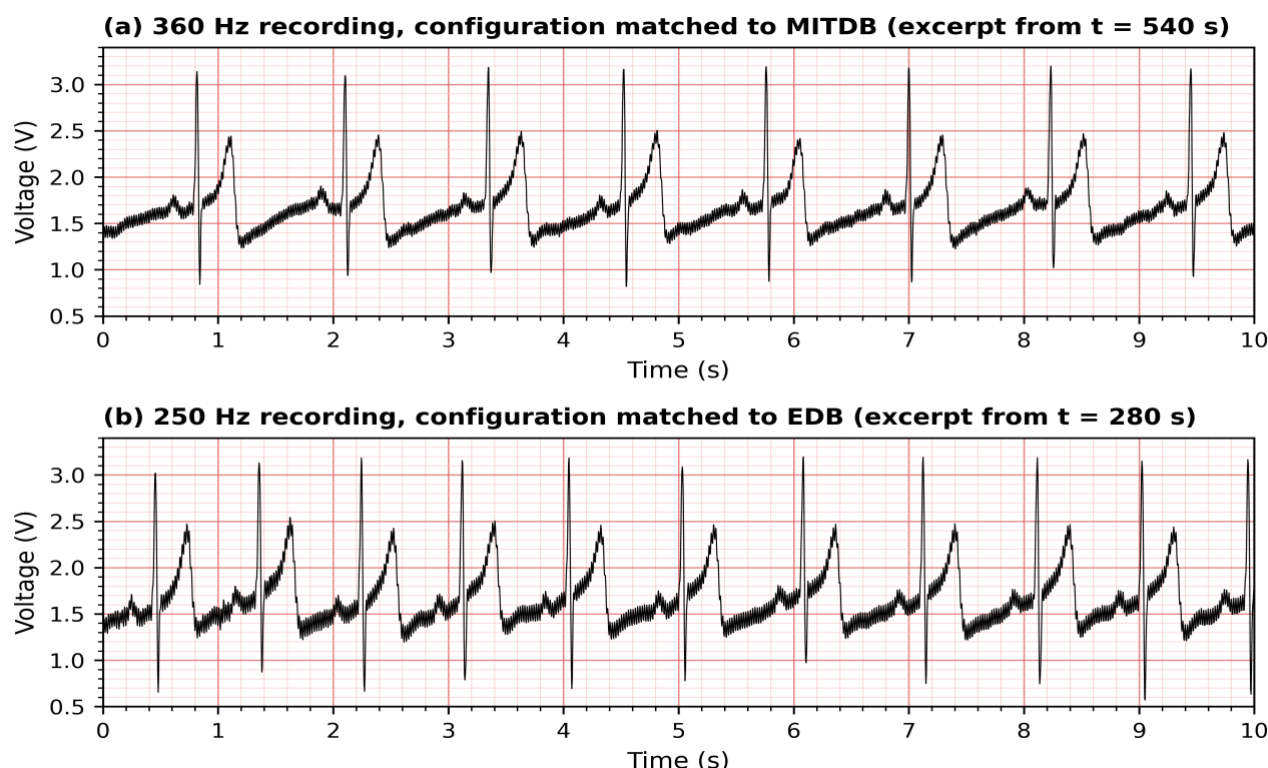


Fig. 2. Representative 10-second ECG signals acquired using the proposed hardware system: (a) 360 Hz configuration matched to MITDB and (b) 250 Hz configuration matched to EDB. The baseline oscillation indicates the residual 50 Hz mains-interference component analyzed in Section III.B.

This ordering across data sources is consistent with the derivation in Section II.B: residual variance reduction from delta pre-processing scales with lag-1 and lag-2 autocorrelation, which are influenced by sampling rate relative to signal bandwidth. The two hardware-acquired configurations achieved a lower CR than their respective matched external databases (2.011 vs. 3.248 for the 360 Hz configuration; 1.942 vs. 2.779 for the 250 Hz configuration) despite using an identical processing pipeline and matched sampling rate and duration. To characterize this gap directly, a posture ablation was performed: 18 new 30-second recordings (sitting-still, standing-still, and sitting-with-mild-movement, three trials each at both 250 Hz and 360 Hz). Welch's-method power spectral density analysis [39] confirmed a sharp 50.000 Hz peak in every condition, consistent with AC mains interference, but its power ratio varied 7.2-7.4-fold by posture: 5.6-5.7% of total signal power when standing, versus 0.75-1.75% when sitting (both still and with mild movement), at both sampling rates. This indicates that mains-hum coupling in this acquisition setup is posture-dependent -- likely reflecting differences in capacitive coupling between the body and nearby AC wiring when standing versus seated -- rather than a fixed environmental noise floor. The elevated noise directly reduced

compression performance: the mean CR for Delta2+Arithmetic decreased from 2.124 (sitting) to 1.632 (standing) at 360 Hz (-23.2%), and from 1.782 to 1.559 (-12.5%) at 250 Hz. Thus, the 360 Hz configuration lost nearly twice the reduction in CR. These results confirm that residual mains hum, rather than motion artifact per se, is the dominant driver of the hardware-vs-external CR gap and identify acquisition posture as a practically controllable factor for improving on-device compression performance. In practice, this suggests that recording while seated, or adding a digital 50 Hz notch filter before quantization, could narrow this gap; however, a notch filter would alter the stored signal and was therefore deliberately not applied in this study. This observation suggests that compression ratios reported on curated public databases should be interpreted as an upper bound on achievable performance in field deployment, where acquisition noise is typically higher.

C. Comparison with a Directly Comparable Prior Method

Table 4 compares the proposed method with Bannajak et al. [10] on the same three public databases. Using the best-performing delta order per database (Section III.B: Delta2 on MITDB, Delta1 on EDB and PTBDB), the proposed method's full-database mean CR falls

Table 4. Compression-ratio comparison between the proposed method and Bannajak et al. [10] on the MITDB, EDB, and PTBDB datasets.

Database	Proposed method	Bannajak et al. [10]	Difference
MITDB	3.248	3.52	-7.7%
EDB	2.969	3.75	-20.8%
PTBDB	4.964	3.16	+57.1%

below Bannajak et al.'s adaptive, per-sample predictor on MITDB (3.248 vs. 3.52, -7.7%) and EDB (2.969 vs. 3.75, -20.8%), but exceeds it substantially on PTBDB (4.964 vs. 3.16, +57.1%), consistent with first-order delta's lower quantization-noise amplification being particularly advantageous at PTBDB's high sampling rate (Section II.B). This comparison is now full-database on both sides (this study: all 48/90/549 records; Bannajak et al.: complete databases as reported in their study), removing the subset-size asymmetry present in earlier drafts of this comparison; it should nonetheless still be read as indicative rather than a controlled equivalence test, since Bannajak et al.'s per-sample adaptive predictor carries side-information overhead not directly comparable to this study's fixed-order approach, and percentage differences should not be interpreted as a general superiority or inferiority claim. Fig. 2 shows representative 10-second excerpts of the hardware-acquired signal at both matched sampling rates; the low-amplitude baseline oscillation visible in both traces is the residual 50 Hz mains component quantified above.

D. Diagnostic Validation and Comparison with Holter-Monitoring Practice

QRS-detection performance is reported using Sensitivity and Positive Predictivity, defined in terms of true-positive (TP), false-negative (FN), and false-positive (FP) beat detections relative to the reference annotations of each database, given by Eqs. (17) and (18) [38]:

$$SE = \frac{TP}{(TP + FN)} (100\%) \quad (17)$$

$$+P = \frac{TP}{(TP + FP)} (100\%) \quad (18)$$

Table 5. Pooled QRS-detection performance of the original and reconstructed ECG signals on 10-record subsets of MITDB and EDB.

Database	Ref. QRS	SE-Orig (%)	+P-Orig (%)	SE-Recon (%)	+P-Recon (%)
MITDB	23343	99.61	99.79	99.61	99.79
EDB	20903	99.89	98.33	99.88	98.30

Table 5 reports pooled QRS-detection performance before and after the complete compression-decompression pipeline, pooled over a 10-record subset of each database (23,343 annotated beats for MITDB and 20,903 for EDB). The largest change in either metric is a 0.03 percentage-point decrease in EDB positive predictivity (98.33% to 98.30%), while the MITDB values are identical to two decimal places. Pooled SE and +P were essentially unchanged after the complete compression-decompression pipeline, confirming that the proposed method preserves diagnostic usability. However, because this comparison is trivially near-identical whenever the underlying codec is verified lossless, it does not by itself characterize the one genuinely lossy step in the pipeline: 8-bit min-max quantization of the native ADC reading. Section III.E reports a statistically rigorous, dedicated comparison against that step.

E. Statistical Validation: Original Signal versus 8-Bit Quantized Signal

To directly assess the diagnostic impact of the one non-bit-exact step in the pipeline, QRS detection was compared between the original, pre-quantization signal and its 8-bit quantized counterpart across 18 hardware recordings (3 postures x 2 sampling rates x 3 trials, 30 s each). Detected beats were matched by nearest-neighbor within a ± 150 ms tolerance window. Across all 745 matched beats, Sensitivity and Positive Predictivity both remained at 100%. Beat-level timing error (quantized-signal detection minus original-signal detection) averaged 0.045 ms (SD = 1.330 ms, 95% CI [-0.051, 0.140] ms via t-distribution; median = 0.000 ms). Neither a paired t-test ($t = 0.915$, $p = 0.361$) nor a Wilcoxon signed-rank test [40] ($W = 4.0$, $p = 0.705$) rejected the null hypothesis of zero median timing difference. These results provide statistically robust evidence that 8-bit min-max quantization introduces no detectable degradation in QRS-detection timing accuracy at this sampling resolution, while the lossless guarantee itself, as clarified in Section II, continues to apply only to the quantized representation and not to the original ADC samples.

The reported mean, standard deviation, 95% confidence interval, and paired t-test statistic for the $n = 745$ matched-beat timing differences d_i are defined in Eqs. (19)–(22) as

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i \quad (19)$$

$$s_d = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (d_i - \bar{d})^2} \quad (20)$$

$$CI_{95\%} = \bar{d} \pm t_{0.975, n-1} \frac{s_d}{\sqrt{n}} \quad (21)$$

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \quad (22)$$

Compared against established Holter-monitoring references -- the Mortara H12+ [41] and the ISHNE Standard Output Format [42] -- all of which store ECG data without any compression, or with compression explicitly disallowed to preserve random access, the proposed method offers a directly applicable extension of recording duration (or reduction of storage footprint) proportional to its compression ratio, without violating the diagnostic accuracy requirements of any of these references. Quantitatively, single-lead 8-bit acquisition at 360 Hz produces 1.296 MB per hour, or 31.1 MB per 24-hour Holter-equivalent record, so a 1 GB (10^9 bytes) storage partition holds approximately 32 days of continuous uncompressed recording. At the compression ratio measured on hardware (CR = 2.011) the same partition holds approximately 65 days, a 50.3% reduction in stored volume, and the 24-hour record shrinks from 31.1 MB to 15.5 MB. At database-level ratios of 3.248 (MITDB) and 4.964 (PTBDB), the same partition would hold approximately 104 and 160 days, corresponding to space savings of 69.2% and 79.9%, respectively. These figures make the practical engineering significance of the reported compression ratios explicit: on a storage-constrained wearable recorder, the proposed method roughly doubles achievable recording duration under real acquisition noise, and more than triples it on clean signals, at no additional hardware cost.

F. Computational Efficiency

Execution time and peak memory are defined and measured as given in Eqs. (23) and (24):

$$T_{exec} = t_{end} - t_{start} \text{ (seconds)} \quad (23)$$

$$M_{peak} = \max_{t_{start} \leq t \leq t_{end}} m(t) \quad (24)$$

where t_{end} and t_{start} are wall-clock timestamps captured immediately before and after a single compression call using Python's `time.perf_counter()`, and M_{peak} is measured using Python's `tracemalloc` module, with $m(t)$ denoting the memory allocated by the compression call at time t , reported in KB or MB.

Wall-clock execution time (`time.perf_counter()`) and peak memory usage (`tracemalloc`) were measured directly on the Raspberry Pi Zero 2 W for Huffman, LZW, and RLE applied to the second-order delta residual of both hardware-acquired recordings. Huffman required 0.254 s (360 Hz recording) / 0.740 s (250 Hz recording) with 5.1-14.1 MB peak memory; LZW was the most expensive by a wide margin at 8.9-25.5 s and 37-87 MB; RLE required 1.1-3.9 s with negligible (<1 KB) memory overhead. We disclose transparently that the corresponding Arithmetic Coding timing figure in our internal records was an entropy-bound theoretical estimate rather than a direct wall-clock measurement of the adaptive coder described in Section II.B, and we do not present it here as a

measured value to avoid overstating precision; a direct measurement of the real coder's wall-clock time and throughput on-device is identified as an immediate follow-up. CPU utilization and energy/power consumption were not measured in this study: introducing current-measurement instrumentation into the shared power rail of the only available acquisition unit was judged an unacceptable equipment risk within the revision window and is noted as a limitation (Section IV) rather than omitted silently. LZW's cost identifies it as the least suitable method for tightly resource-constrained or real-time deployment despite its competitive compression ratio, supporting the record-then-compress (rather than streaming) architecture adopted in this study's hardware implementation.

IV. Discussion

The results demonstrate that the proposed method's two components each contribute for identifiable, mathematically grounded reasons: delta encoding reduces the residual variance (and thus the Shannon entropy) whenever adjacent-sample correlation is high, with the better-performing order (first vs. second) determined by the balance between signal-variance reduction and quantization-noise amplification derived in Section II.B and empirically confirmed in Section III.B; and Arithmetic Coding approaches the entropy bound more closely than Huffman coding by avoiding a fixed per-symbol bit-length constraint. An adaptive multi-order extension of the proposed predictor, following Bannajak et al. [10] was also evaluated but underperformed the proposed fixed-order method by 7-9%, since the side-information required to communicate the selected predictor order per sample outweighed its finer prediction accuracy under general-purpose entropy coders; this indicates that the proposed fixed-order design is preferable unless a dedicated side-information coding scheme is adopted.

A. Limitations

This study's hardware validation remains limited to single-lead ECG acquisition and a single human subject; broader claims of deployment readiness would require validation across multiple subjects, which is identified as a priority direction for future work rather than resolved in this revision. Within this single-subject scope, a posture ablation (sitting-still, standing-still, sitting-with-mild-movement; three trials each at both 250 Hz and 360 Hz) was added to provide the first repeated-session, multi-condition evidence in this study (Section III.B), which revealed that acquisition posture measurably affects mains-hum coupling and compression ratio -- a finding we believe strengthens the paper's practical contribution despite the remaining single-subject scope. Energy and power consumption

Table 6. Quantitative comparison of the proposed method with previous lossless ECG compression methods in terms of platform, sampling frequency, compression ratio, and losslessness.

Study	Method	Platform	fs (Hz)	Mean CR	Lossless
Bannajak et al. [10]	Adaptive multi-order linear prediction + modified Golomb-Rice coding	Offline software (no hardware)	DB-native	3.16–3.75	Yes
Banerjee & Singh [11]	Second-order delta + buffer-array run-length encoding	AD8232 + Raspberry Pi	DB-native	2.34 (HW) / 3.82–6.52 (DB)	Yes
Tsai & Hussain [12]	Lossless ECG compression algorithm in dedicated VLSI	VLSI ASIC (0.18 μm)	MITDB-native	2.77	Yes
Tsai & Kuo [13]	Adaptive linear prediction + content-adaptive Golomb-Rice coding	Embedded platform	MITDB-native	2.77–2.83	Yes
Deepu et al. [14]	QRS detection integrated with lossless compression	3-lead ECG-on-chip ASIC	256	2.15	Yes
Campobello et al. [15]	Zero-order predictor + RAKE encoding	Microcontroller	DB-native	> 2.67	Yes
This study	First/second-order delta (mod 256) + Arithmetic Coding	Raspberry Pi Zero 2 W + AD8232 / ADS1115	250 / 360 (HW); DB-native	1.942–2.011 (HW) / 2.969–4.964 (DB)	Yes (byte-exact)

on the Raspberry Pi Zero 2 W were not measured in this study: introducing series/parallel current-measurement instrumentation into the shared power rail of the only available acquisition unit was judged an unacceptable equipment risk within the revision window, and this is noted explicitly as a remaining direction for future work rather than omitted silently. The PTB Diagnostic ECG Database could not be replicated on the proposed hardware platform, as its 1000 Hz specification exceeds the ADS1115's maximum native sampling rate (860 SPS).

B. Generalizability and Future Work

Beyond the specific hardware and databases evaluated in this paper, the proposed method's design principles are expected to generalize to other single-lead, resource-constrained biomedical acquisition scenarios sharing similar sample-to-sample correlation characteristics, such as photoplethysmogram (PPG) monitoring, since the mathematical justification for delta pre-processing in Section II.B depends only on the signal exhibiting high lag-1 autocorrelation relative to its sampling rate, a property common to many slowly varying physiological signals. Future work could extend the proposed method to multi-lead ECG acquisition, evaluate its performance across a larger and more demographically diverse subject pool for the hardware-acquired comparison, and investigate a dedicated entropy-coding scheme for adaptive-order side information, which this study found to be a limiting factor when a per-sample adaptive predictor was evaluated as a supplementary extension beyond the proposed fixed-order design.

Relative to the closely related hardware-validated approach of Banerjee and Singh [11], which re-engineered the entropy stage itself (a buffer-array RLE variant) to maximize compression ratio on similar AD8232/Raspberry Pi hardware, the proposed method

instead uses an unmodified, general-purpose Arithmetic Coding stage. This design choice trades some achievable compression ratio for a clearer, more directly interpretable separation between the contribution of the proposed delta pre-processing stage and the contribution of the entropy coding stage itself (Table 2), and avoids introducing an entropy-coder-specific optimization that would be harder to disentangle from the delta-encoding contribution this paper aims to isolate. Similarly, whereas wavelet-based approaches such as [18], [19] can achieve competitive compression ratios, they are predominantly lossy, making them unsuitable for the diagnostic-fidelity guarantee that motivates the lossless design adopted in this paper.

Table 6 situates the proposed method quantitatively against six prior lossless ECG compression studies spanning offline algorithms, custom VLSI/ASIC hardware, and general-purpose embedded platforms. In terms of computational requirements, the custom-silicon implementations [12], [13], [14] achieve sub-microwatt power consumption unattainable on a general-purpose embedded platform, but at the cost of fixed-function hardware that cannot be reconfigured or reused for other signal-processing tasks; the present study instead targets a commodity, field-reprogrammable platform (Raspberry Pi Zero 2 W), trading power efficiency for deployment flexibility and a substantially lower non-recurring engineering cost. Relative to sampling frequency, [12], [13], [14] report results at a single fixed rate (typically MIT-BIH-native or 256 Hz), whereas this study additionally evaluates and explains -- via the analytical criterion of Section II.B -- how compression ratio shifts across three different sampling rates (200, 250, and 360 Hz) and correspondingly different correlation regimes, a dimension largely unaddressed in [12], [13], [14]. Each comparator study's principal limitation, drawn from its

own reporting, is summarized as follows: Bannajak et al. [10] does not report a physical hardware implementation; Banerjee & Singh [11] re-engineers its entropy stage for a single fixed configuration without a delta-order-selection criterion; Tsai & Hussain [12], Tsai & Kuo [13], and Deepu et al. [14] are evaluated only on MIT-BIH and do not report cross-database or cross-sampling-rate validation; and Campobello et al. [15] targets minimal memory footprint at the expense of achievable compression ratio. The proposed method's compression ratio on hardware-acquired data (1.942-2.011) is comparable to or below several prior studies, as expected because this study evaluates unmodified, general-purpose entropy coders to isolate the delta pre-processing contribution (Section III.A) rather than maximize compression ratio through coder-specific engineering.

V. Conclusion

This paper proposed a lossless ECG compression method that combines delta encoding with arithmetic coding and was implemented on a Raspberry Pi Zero 2 W with an AD8232 front-end and ADS1115 converter. Full-database evaluation (687 records, 8,244 method-record combinations, all verified byte-exact lossless) showed that the best-performing delta order is signal-dependent: second-order delta pre-processing achieved the highest mean compression ratio on MITDB (3.248), while first-order delta pre-processing performed best on EDB (2.969) and PTBDB (4.964) – a pattern explained by an analytically derived criterion balancing signal-variance reduction against quantization-noise amplification (Section II.B), and confirmed empirically via full-database autocorrelation measurements. A posture ablation additionally showed that residual 50 Hz mains-hum coupling, not motion artifact, is the dominant driver of the gap between hardware-acquired and public-database compression ratios, with acquisition posture identified as a practically controllable factor. QRS-based statistical validation comparing the original pre-quantization signal against its 8-bit quantized counterpart found no statistically significant timing difference (745 matched beats, paired t-test $p=0.361$, Wilcoxon $p=0.705$), with Sensitivity and Positive Predictivity both at 100%. Compared with established Holter-monitoring practice, which typically applies no compression at all, the proposed method offers a promising means of extending recording duration or reducing storage footprint, with the important caveat that the lossless guarantee applies to the 8-bit quantized representation rather than the original ADC samples, and that hardware validation to date remains limited to a single subject.

Future work will prioritize four directions: (1) multi-subject validation to establish inter-subject variability in ECG morphology and compressibility beyond the

single-subject scope of this study; (2) extension to multi-lead ECG acquisition; (3) direct measurement of energy and power consumption on the embedded platform, which equipment constraints prevented in this revision; and (4) direct, on-device wall-clock measurement of the Arithmetic Coding stage's execution time and throughput, complementing the entropy-bound theoretical estimate reported in Section III.F.

Ethics Approval and Consent to Participate

The study involved non-invasive, single-lead ECG acquisition from one adult volunteer for engineering validation purposes. Written informed consent was obtained from the participant before data collection. No clinical intervention, diagnosis, or treatment decision was performed as part of the study. Formal institutional ethics committee approval was not obtained. The participant was informed of the purpose of the study, the voluntary nature of participation, and the right to withdraw at any time without consequence.

Consent for Publication

Written informed consent for publication of the anonymized ECG recordings and derived results was obtained from the participant. No personally identifiable information was collected or reported in the manuscript.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Competing Interests

The authors declare that they have no financial or non-financial competing interests that could have influenced the conduct, analysis, interpretation, or reporting of this study.

Author Contributions

Jalu Wicaksana conceived and designed the study, developed the proposed compression method, implemented the hardware acquisition and compression pipeline, performed the experiments and data analysis, and prepared the original manuscript. Ahmad Sugiana and Junartha Halomoan contributed to supervision, methodological guidance, interpretation of the results, and critical revision of the manuscript. All authors reviewed and approved the final version of the manuscript.

Data Availability

The public datasets used in this study, including the MIT-BIH Arrhythmia Database, European ST-T Database, and PTB Diagnostic ECG Database, are publicly available through PhysioNet. The hardware-acquired ECG recordings generated during the study

are available from the corresponding authors upon reasonable request.

Code Availability

The Python code used for ECG acquisition, preprocessing, compression, decompression, entropy coding, and QRS-based validation is available from the corresponding authors upon reasonable request.

Acknowledgments

The authors thank the reviewers for their constructive comments and suggestions, which contributed to improving the quality and presentation of the manuscript.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During preparation of this manuscript, the authors used Claude (Anthropic) to assist with language editing, rewording and restructuring text in response to reviewer comments, typesetting equations, redrawing Fig. 1, and document formatting. The authors subsequently reviewed and edited the content and take full responsibility for the final manuscript.

References

- [1] World Health Organization, "Cardiovascular diseases (CVDs)." Accessed: Sep. 21, 2026. [Online]. Available: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
- [2] A. Němcová, R. Smíšek, L. Maršánová, L. Smital, and M. Vitek, "A comparative analysis of methods for evaluation of ECG signal quality after compression," *Biomed Res. Int.*, vol. 2018, 2018, doi: 10.1155/2018/1868519.
- [3] M. A. Serhani, H. T. El Kassabi, H. Ismail, and A. Nujum Navaz, "ECG monitoring systems: Review, architecture, processes, and key challenges," *Sensors (Switzerland)*, vol. 20, no. 6, Mar. 2020, doi: 10.3390/s20061796.
- [4] N. C. Smith, "Method and apparatus for data compression in an ECG monitoring system," US 4,947,858 A, Aug. 14, 1990 Accessed: Sep. 29, 2026. [Online]. Available: patents.google.com/patent/US4947858
- [5] G. H. Bardy, J. Mikalson Bishay, and J. Felix, "Electrocardiography and respiratory monitor," US 10,602,977 B2, Mar. 31, 2020 Accessed: Sep. 29, 2026. [Online]. Available: patents.google.com/patent/US10602977B2
- [6] M. Jacobsson, F. Seoane, and F. Abtahi, "The role of compression in large scale data transfer and storage of typical biomedical signals at hospitals," *Health Informatics J.*, vol. 29, no. 4, Oct. 2023, doi: 10.1177/14604582231213846.
- [7] Association for the Advancement of Medical Instrumentation, "ANSI/AAMI EC11:1991 Diagnostic electrocardiographic devices," Arlington, VA, 1991.
- [8] P. Rubel, J. Fayn, P. W. Macfarlane, D. Pani, A. Schlögl, and A. Värri, "The History and Challenges of SCP-ECG: The Standard Communication Protocol for Computer-Assisted Electrocardiography," *Hearts*, vol. 2, no. 3, pp. 384–409, Aug. 2021, doi: 10.3390/hearts2030031.
- [9] A. Tiwari and T. H. Falk, "Lossless electrocardiogram signal compression: A review of existing methods," *Biomed. Signal Process. Control*, vol. 51, pp. 338–346, May 2019, doi: 10.1016/j.bspc.2019.03.004.
- [10] K. Bannajak, N. Theera-Umpon, and S. Auephanwiriyakul, "Signal Acquisition-Independent Lossless Electrocardiogram Compression Using Adaptive Linear Prediction," *Int. J. Environ. Res. Public Health*, vol. 20, no. 3, Feb. 2023, doi: 10.3390/ijerph20032753.
- [11] S. Banerjee and G. K. Singh, "A New Real-time Lossless Data Compression Algorithm for ECG and PPG Signals," *Biomed. Signal Process. Control*, vol. 79, p. 104127, 2023, doi: 10.1016/j.bspc.2022.104127.
- [12] T. H. Tsai and M. A. Hussain, "VLSI Implementation of Lossless ECG Compression Algorithm for Low Power Devices," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 12, pp. 3317–3321, Dec. 2020, doi: 10.1109/TCSII.2020.2978554.
- [13] T. H. Tsai and W. T. Kuo, "An Efficient ECG Lossless Compression System for Embedded Platforms with Telemedicine Applications," *IEEE Access*, vol. 6, pp. 42207–42215, Jul. 2018, doi: 10.1109/ACCESS.2018.2858857.
- [14] C. J. Deepu, X. Zhang, C. H. Heng, and Y. Lian, "A 3-Lead ECG-on-chip with QRS detection and lossless compression for wireless sensors," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 63, no. 12, pp. 1151–1155, Dec. 2016, doi: 10.1109/TCSII.2016.2613564.
- [15] G. Campobello, A. Segreto, S. Zanafi, and S. Serrano, "An Efficient Lossless Compression Algorithm for Electrocardiogram Signals," in *2018 26th European Signal Processing Conference (EUSIPCO)*, Rome, Italy: IEEE, Sep. 2018, pp. 777–781. doi: 10.23919/EUSIPCO.2018.8553597.
- [16] D. Rzepka, "Low-complexity lossless multichannel ECG compression based on selective linear prediction," *Biomed. Signal Process. Control*, vol. 57, Mar. 2020, doi: 10.1016/j.bspc.2019.101705.

- [17] A. Mitra, P. Kundu, and R. Gupta, "Deep Learning Based Approach for Lossless ECG Compression," *Cardiovasc. Eng. Technol.*, vol. 17, no. 2, pp. 155–174, 2026, doi: 10.1007/s13239-026-00821-5.
- [18] M. S. Manikandan and S. Dandapat, "Wavelet-based electrocardiogram signal compression methods and their performances: A prospective review," *Biomed. Signal Process. Control*, vol. 14, pp. 73–107, 2014, doi: 10.1016/j.bspc.2014.07.002.
- [19] H. S. Pal, A. Kumar, A. Vishwakarma, and M. K. Ahirwal, "Electrocardiogram signal compression using tunable-Q wavelet transform and meta-heuristic optimization techniques," *Biomed. Signal Process. Control*, vol. 78, Sep. 2022, doi: 10.1016/j.bspc.2022.103932.
- [20] I. Devindi, S. Liyanage, T. Jayarathna, J. Alawatugoda, and R. Ragel, "A novel ECG compression algorithm using Pulse-Width Modulation integrated quantization for low-power real-time monitoring," *Sci. Rep.*, vol. 14, p. 17162, 2024, doi: 10.1038/s41598-024-68022-5.
- [21] M. Wasimuddin, K. Elleithy, A.-S. Abuzneid, M. Faezipour, and O. Abuzagheh, "Stages-Based ECG Signal Analysis From Traditional Signal Processing to Machine Learning Approaches: A Survey," *IEEE Access*, vol. 8, pp. 177782–177803, 2020, doi: 10.1109/ACCESS.2020.3026968.
- [22] D. Cuevas-González *et al.*, "ECG Standards and Formats for Interoperability between mHealth and Healthcare Information Systems: A Scoping Review," *Int. J. Environ. Res. Public Health*, vol. 19, no. 19, Oct. 2022, doi: 10.3390/ijerph191911941.
- [23] S. M. S. Jalaaliddine, C. G. Hutchens, R. D. Strattan, and W. A. Coberly, "ECG Data Compression Techniques-A Unified Approach," *IEEE Trans. Biomed. Eng.*, vol. 37, no. 4, pp. 329–343, Apr. 1990, doi: 10.1109/10.52340.
- [24] A. Koski, "Lossless ECG encoding," *Comput. Methods Programs Biomed.*, vol. 52, no. 1, pp. 23–33, 1997, doi: 10.1016/S0169-2607(96)01779-8.
- [25] O. Kerdjadj, A. Amira, K. Ghanem, N. Ramzan, S. Katsigiannis, and F. Chouireb, "An FPGA implementation of the matching pursuit algorithm for a compressed sensing enabled e-Health monitoring platform," *Microprocess. Microsyst.*, vol. 67, pp. 131–139, Jun. 2019, doi: 10.1016/j.micpro.2019.03.007.
- [26] M. S. Zaman and B. I. Morshed, "A low-power portable scanner for body-worn Wireless Resistive Analog Passive (WRAP) sensors for mHealth applications," *Measurement*, vol. 177, p. 109214, Jun. 2021, doi: 10.1016/j.measurement.2021.109214.
- [27] C. J. Deepu, X. Zhang, W. S. Liew, D. L. T. Wong, and Y. Lian, "An ECG-on-Chip with 535 nW/channel integrated lossless data compressor for wireless sensors," *IEEE J. Solid-State Circuits*, vol. 49, no. 11, pp. 2435–2448, Nov. 2014, doi: 10.1109/JSSC.2014.2349994.
- [28] N. R. Saputri *et al.*, "Enhancing the Electrocardiogram Signal Quality by Applying Butterworth Infinite Impulse Response Filter 8th Order," *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 4, no. 4, pp. 235–242, Oct. 2022, doi: 10.35882/ijeemi.v4i4.259.
- [29] I. H. Witten, R. M. Neal, and J. G. Cleary, "Arithmetic coding for data compression," *Commun. ACM*, vol. 30, no. 6, pp. 520–540, Jun. 1987, doi: 10.1145/214762.214771.
- [30] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, Jul. 1948, doi: 10.1002/j.1538-7305.1948.tb01338.x.
- [31] D. A. Huffman, "A Method for the Construction of Minimum-Redundancy Codes," *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098–1101, Sep. 1952, doi: 10.1109/JRPROC.1952.273898.
- [32] J. Ziv and A. Lempel, "Compression of Individual Sequences via Variable-Rate Coding," *IEEE Trans. Inf. Theory*, vol. 24, no. 5, pp. 530–536, Sep. 1978, doi: 10.1109/TIT.1978.1055934.
- [33] T. A. Welch, "A Technique for High-Performance Data Compression," *IEEE Computer*, vol. 17, no. 6, pp. 8–19, Jun. 1984, doi: 10.1109/MC.1984.1659158.
- [34] G. B. Moody and R. G. Mark, "The Impact of the MIT-BIH Arrhythmia Database History, Lessons Learned, and Its Influence on Current and Future Databases," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, 2001, doi: 10.1109/51.932724.
- [35] A. Taddei *et al.*, "The European ST-T database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography," *Eur. Heart J.*, vol. 13, no. 9, pp. 1164–1172, Sep. 1992, doi: 10.1093/oxfordjournals.eurheartj.a060332.
- [36] R. Bousseljot, D. Kreiseler, and A. Schnabel, "Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet," *Biomedizinische Technik*, vol. 40, no. s1, pp. 317–318, 1995, doi: 10.1515/bmte.1995.40.s1.317.
- [37] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," *Circulation*, vol. 101, no. 23,

- pp. e215–e220, Jun. 2000, doi: 10.1161/01.CIR.101.23.e215.
- [38] J. Pan and W. J. Tompkins, "A Real-Time QRS Detection Algorithm," *IEEE Trans. Biomed. Eng.*, vol. 32, no. 3, pp. 230–236, Mar. 1985, doi: 10.1109/TBME.1985.325532.
- [39] P. D. Welch, "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms," *IEEE Transactions on Audio and Electroacoustics*, vol. 15, no. 2, pp. 70–73, Jun. 1967, doi: 10.1109/TAU.1967.1161901.
- [40] F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945, doi: 10.2307/3001968.
- [41] Mortara Instrument Inc., "H12+ Digital Holter Recorder User Manual," 2018. [Online]. Available: <http://www.mortara.com>
- [42] F. Badilini, "The ISHNE Holter Standard Output File Format," *Ann. Noninvasive Electrocardiol.*, vol. 3, no. 3, pp. 263–266, 1998, doi: 10.1111/j.1542-474X.1998.tb00353.x.

Author Biography



Jalu Wicaksana is currently pursuing a Master's degree (S2) in Electrical Engineering at Telkom University, Bandung, Indonesia. His graduate research addresses lossless data compression for resource-constrained wearable healthcare monitoring devices, with particular emphasis on electrocardiogram acquisition and on-device storage. For the present study, he designed and built a complete single-lead ECG acquisition system based on an AD8232 analog front-end, an ADS1115 sigma-delta converter, and a Raspberry Pi Zero 2 W host, and implemented and benchmarked four entropy coding schemes Huffman, LZW, Run-Length Encoding, and Arithmetic Coding in combination with first- and second-order delta pre-processing, validating every configuration against the MIT-BIH Arrhythmia, European ST-T and PTB Diagnostic ECG Databases. His broader research interests include embedded systems, biomedical signal processing, low-power wireless communication for Internet-of-Things applications, and the validation of signal-processing pipelines against public clinical reference databases.



Ahmad Sugiana received a B.E. degree in computer science from the University of Padjadjaran, Bandung, Indonesia in 2002, and an M.Eng. degree in electrical engineering from the Bandung Institute of Technology (BIT), Indonesia in 2009. In 2017, he received Dr. Eng.

degree in control and instrumentation engineering from Kwangwoon University, Seoul, South Korea. During Ph.D. program, he was involved in some research about railway signaling engineering such as interlocking, etc. From 1999 to 2013, he joined the Department of Signaling, Telecommunication and Electricity, Indonesia Railway Company. Since 2017, he has joined the Department of Electronics Engineering, at Telkom University. He joined the Railway Convergence Technology Management Institute, Woosong University as a postdoctoral researcher and assistant professor from 2020 to 2022. Since 2024, he has been assigned as head of the Center of Excellence (CoE) for Smart Transportation and Robotics. He is also a lecturer in the Master's (S2) Electrical Engineering program at Telkom University, Bandung, Indonesia, where he supervises graduate research in embedded systems and biomedical signal processing.



Junartha Halomoan is a Senior Lecturer in the Department of Electrical Engineering at Telkom University, Indonesia. He earned his Bachelor's and Doctoral degrees in Electrical Engineering from Universitas Indonesia, and his Master's degree from Telkom University. His primary research interests encompass biomedical engineering, industrial automation, and embedded systems. His pioneering doctoral research focuses on physiological signal processing, specifically developing highly accurate driving fatigue detection frameworks. His extensively published research also spans diverse technical domains, including advanced image processing, wood fiber detection algorithms, balancing systems for waiter robots, and smart home automation utilizing wireless sensor networks. These innovative findings have been widely published in peer-reviewed international journals, reflecting a profound dedication to leveraging machine learning and electronics to enhance human safety, smart automation, and practical engineering solutions.