Quantizing for Noisy Flash Memory Channels

Juyun Oh*, Taewoo Park*, Jiwoong Im*, Yuval Cassuto†, and Yongjune Kim*
*Department of Electrical Engineering, POSTECH, Pohang, South Korea
Email: {juyunoh, parktaewoo, jw3562, yongjune}@postech.ac.kr
†Viterbi Department of Electrical and Computer Engineering, Technion - Israel Institute of Technology, Haifa, Israel
Email: ycassuto@ee.technion.ac.il

***Abstract*—Flash memory-based processing-in-memory (flash-based PIM) offers high storage capacity and computational efficiency but faces significant reliability challenges due to noise in high-density multi-level cell (MLC) flash memories. Existing verify level optimization methods are designed for general storage scenarios and fail to address the unique requirements of flash-based PIM systems, where metrics such as mean squared error (MSE) and peak signal-to-noise ratio (PSNR) are critical. This paper introduces an integrated framework that jointly optimizes quantization and verify levels to minimize the MSE, considering both quantization and flash memory channel errors. We develop an iterative algorithm to solve the joint optimization problem. Experimental results on quantized images and SwinIR model parameters stored in flash memory show that the proposed method significantly improves the reliability of flash-based PIM systems.**

I. Introduction

Processing-in-memory (PIM), also known as in-memory computing, has emerged as a promising alternative to the traditional von Neumann architecture, providing reduced data transfer overhead and improved computational efficiency [1]–[3]. In particular, flash memory-based PIM (flash-based PIM) offers a significantly larger storage capacity compared to SRAM- and DRAM-based PIM, making it an attractive solution for storing vast datasets and model parameters [4], [5]. However, the continued scaling of the flash memory technology introduces significant reliability challenges, particularly for high-density flash memory cells. In N-bit per cell multi-level cell (MLC) flash memories, which represent 2^N distinct states within a constrained threshold voltage window, noise significantly impairs the accuracy of the resulting computations [6], [7].

Several techniques have been proposed to minimize the impact of flash-memory channel noise by optimizing the distances between states [8]–[10]. This optimization problem is commonly referred to as *write voltage* optimization [8], [9] or *verify levels* optimization [10], as the distances between states (or write voltages) can be adjusted using the verify levels in incremental step pulse programming (ISPP) [11]. However, these approaches are primarily designed for general data-storage scenarios, and thus focus on minimizing the bit error rate (BER), rather than optimizing measures more relevant to PIM such as mean squared error (MSE) or peak signal-to-noise ratio (PSNR). These performance measures fit better the requirements of data-intensive applications in domains such as machine learning and computer vision. They also do not rely on conventional channel coding schemes that introduce high latency in computing environments with stringent latency requirements [12]. Therefore, new strategies for optimizing verify levels should be developed to ensure the efficient and reliable operation of flash-based PIM systems.

Optimizing the MSE of the in-memory representation opens the opportunity to optimize jointly with another factor affecting the MSE performance: *quantization*. Since the data stored in memory is itself a quantized version of the application data, the overall MSE performance depends on both the quantization levels (source optimization) and the memory verify levels (channel optimization). Quantization is typically optimized without regards to channel errors, employing classical algorithms such as the Lloyd-Max algorithm [13], [14]. The importance of channel errors to the quantization performance has been observed before in [15], where the authors introduce the concept of *quantizing for noisy channels*. However, this method assumes fixed channel transition probabilities that cannot be controlled, thus not offering the potential benefits of verify-level optimization that were demonstrated in other applications [8], [10].

In this paper, we propose an integrated framework that optimizes both quantization levels and verify levels to effectively enhance the reliability of flash-based PIM systems. We formulate a joint optimization problem to minimize the MSE, taking into account errors introduced by quantization and the flash memory channel. In this framework, the optimization variables include the quantization levels and verify levels. Since the impacts of these two sets of variables on MSE are interdependent, we develop an iterative algorithm to solve the formulated optimization problem. Unlike conventional approaches that focus separately on verify-level optimization [8]–[10] and quantization level optimization [13]–[15], our approach jointly optimizes both the quantization levels and verify levels in an iterative manner. To the best of our knowledge, this is the first work to jointly optimize the quantization and verify levels for flash-based PIM systems.

To comprehensively evaluate the reliability of data and model parameters stored in the flash-based PIM systems, we conduct two distinct experiments. First, we compare the PSNRs of images stored using the proposed method to that of a conventional approach. Second, we evaluate the PSNRs of the SwinIR model [16], assuming its model parameters

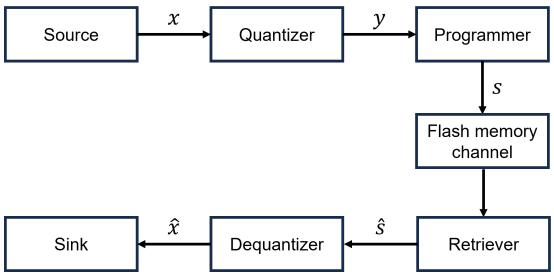


Fig. 1. System model for flash-based PIM systems.

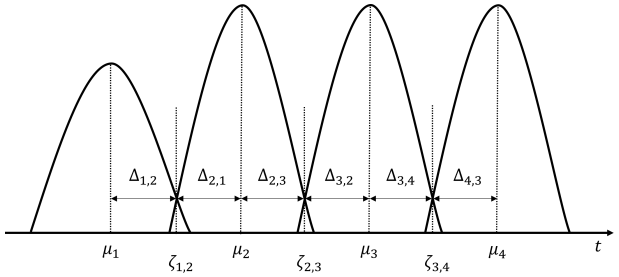


Fig. 2. Threshold voltage distribution for 2-bit per cell flash memories with the constrained window $W = \mu_4 - \mu_1$.

are quantized to 4 bits and stored in quad-level cell (QLC) flash memory. We then compare the proposed method to the conventional method under these conditions.

The rest of this paper is organized as follows. Section II introduces the system model of flash-based PIM and reviews quantization for noisy channels. Section III derives the MSE for flash-based PIM systems, formulates the joint optimization problem, and presents an iterative algorithm to solve this problem. Section IV provides experimental results, and Section V concludes the paper.

II. System Model and Channel-Aware Quantization

In this section, we first provide the system model and then briefly review the scheme of channel-aware quantization from [15].

A. System Model

The source data is modeled as a realization of a continuous random variable X with the probability density function (PDF) $f_X(x)$. The quantizer maps x into y, a value from a finite set of real numbers $\{v_j\} := \{v_1, \ldots, v_M\}$, following the mapping rule: $y = v_j$ if $u_j < x \leq u_{j+1}$. The set $\{u_j\} := \{u_2, \ldots, u_M\}$ denotes the quantization thresholds, with u_1 and u_{M+1} defined as the greatest lower and least upper bound of the input data, respectively. Subsequently, the encoder assigns a binary codeword to the quantizer output y, and this codeword is programmed into flash memory cells. The output of the noisy memory channel is denoted as $\widehat{s}$. The output of the dequantizer $\widehat{x}$ is the real number v_j mapped back from $\widehat{s}$. The noisy flash memory channel is characterized by a channel transition matrix $P = [P_{i,j}]$, where $P_{i,j} = P(\widehat{x} = v_j | y = v_i)$. The system model is described in Fig. 1.

In N-bit per cell flash memories, the threshold voltage distribution of memory cells consists of 2^N distinct states, ranging from s_1 (the erase state) to s_{2^N} (the highest state). This distribution can be approximated as a mixture of Gaussian distributions [10]. Consequently, the threshold voltage distribution $f_T(t)$ is given by

$$f_T(t) = \sum_{i=1}^{2^N} p(s_i) f_i(t) = \sum_{i=1}^{2^N} \frac{p(s_i)}{\sqrt{2\pi}\sigma_i} e^{-\frac{(t-\mu_i)^2}{2\sigma_i^2}}, \qquad (1)$$

where t denotes the threshold voltage, $f_i(t)$ is the Gaussian PDF with mean μ_i and variance σ_i^2 for state s_i, and $p(s_i)$ represents the probability of the state s_i.

To simplify the architecture of flash-based PIM systems, we set $M = 2^N$, such that each flash memory cell stores one N-bit quantized value. Then, the quantized value v_i is programmed into the state s_i for $i = 1, \ldots, M$. If the retrieved state is s_j, then the decoder outputs $\widehat{x} = v_j$. Hence, the channel transition probability is given by

$$P_{i,j} = P(\widehat{x} = v_j | y = v_i) = P(\widehat{s} = s_j | s = s_i). \qquad (2)$$

In the Gaussian mixture model, the dominant source of errors arises from the misreading of adjacent states. Hence, flash memory channel errors can be expressed as the following two types of transition probabilities:

$$P_{i,i-1} = P(\widehat{s} = s_{i-1} | s = s_i) \approx Q\left(\frac{\Delta_{i,i-1}}{\sigma_i}\right), \qquad (3)$$

$$P_{i,i+1} = P(\widehat{s} = s_{i+1} | s = s_i) \approx Q\left(\frac{\Delta_{i,i+1}}{\sigma_i}\right), \qquad (4)$$

where $\Delta_{i,i-1} = \mu_i - \zeta_{i-1,i}$ and $\Delta_{i,i+1} = \zeta_{i,i+1} - \mu_i$. Here, $\zeta_{i,i+1}$ is the transition level between s_i and s_{i+1}. The tail probability function is given by $Q(z) = \int_z^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt$. Hence, the channel transition matrix P is represented as a tridiagonal matrix.

Fig. 2 shows the threshold voltage distribution for 2-bit per cell flash memories. Within the constrained window W, there are four distinct states ranging from s_1 (erase state) to s_4 (highest state). The window W is defined as the distance between the mean of the erase state and the mean of the highest state, i.e., $W = \mu_{2^N} - \mu_1$. Once we determine the size-$(2^{N+1} - 2)$ set $\{\Delta_{i,j}\}$, the distances between states are uniquely determined, and since the $\{\sigma_i\}$ parameters are assumed given, the $\{f_i(t)\}$ distributions are also determined.

B. Channel-Aware Quantization

While the Lloyd-Max algorithm [13], [14] aims to minimize the MSE, it does not account for channel errors. In [15], the authors proposed a channel-aware quantization technique, referred to as *quantizing for noisy channels*, which improves the MSE by explicitly considering channel errors during the

quantization process. By accounting for channel errors, the MSE can be expressed as follows [15]:

$$\begin{aligned}&\mathbb{E}[(X-\widehat{X})^2]\\&=\int_{-\infty}^{\infty}x^2 f_X(x)dx-2\sum_{j=1}^{M}v_j\sum_{i=1}^{M}P_{i,j}\cdot\int_{u_i}^{u_{i+1}}xf_X(x)dx\\&+\sum_{j=1}^{M}v_j^2\sum_{i=1}^{M}P_{i,j}\cdot\int_{u_i}^{u_{i+1}}f_X(x)dx.\end{aligned}\tag{5}$$

The channel-aware quantization aims to minimize (5) by jointly optimizing $\{v_j\}$ and $\{u_j\}$, as given in [15]:

$$v_j=\frac{\sum_{i=1}^{M}P_{i,j}\int_{u_i}^{u_{i+1}}xf_X(x)dx}{\sum_{i=1}^{M}P_{i,j}\int_{u_i}^{u_{i+1}}f_X(x)dx},\qquad j=1,\ldots,M,\tag{6}$$

$$u_j=\frac{1}{2}\cdot\frac{\sum_{k=1}^{M}v_k^2(P_{j,k}-P_{j-1,k})}{\sum_{k=1}^{M}v_k(P_{j,k}-P_{j-1,k})},\qquad j=2,\ldots,M.\tag{7}$$

As in the Lloyd-Max algorithm, the dependence between (6) and (7) requires some iterative algorithm to alternate between (6) and (7) until convergence. For a noiseless channel, the channel matrix simplifies to the identity matrix, i.e., $P_{i,j}=\delta_{i,j}$ (Kroneker delta). Then, (6) and (7) reduce to the expressions for the Lloyd-Max algorithm.

III. Integrated Optimization of Quantization Levels and Verify Levels

In this section, we propose an integrated framework that jointly optimizes quantization levels and verify levels to minimize the MSE of (5). The optimization variables include $\{v_j\}$ and $\{u_j\}$ defining the quantization levels, and $\{\Delta_{i,j}\}$ determining the verify levels. For simplicity, we set $\mathbf{u}:=\{u_j\}_{j=2}^{M}$, $\mathbf{v}:=\{v_j\}_{j=1}^{M}$, and $\mathbf{\Delta}:=\{\Delta_{i,i+1},\Delta_{i+1,i}\}_{i=1}^{M-1}$. Then, the joint optimization problem can be formulated as:

$$\begin{aligned}\underset{\mathbf{u},\mathbf{v},\mathbf{\Delta}}{\text{minimize}}\quad&\mathrm{MSE}(\mathbf{u},\mathbf{v},\mathbf{\Delta})\\\text{subject to}\quad&\|\mathbf{\Delta}\|_1=W\\&\mathbf{\Delta}\succeq 0,\end{aligned}\tag{8}$$

where $\|\cdot\|_1$ denotes the ℓ_1-norm, and $\mathbf{\Delta}\succeq 0$ implies $\Delta_{i,j}\geq 0$ for all i and j.

This problem is non-convex, making it challenging to directly obtain the optimal solutions for $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{\Delta}$. Furthermore, the impacts of $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{\Delta}$ on the MSE are interdependent. Given a channel transition probability matrix, the optimal $\mathbf{u}$ and $\mathbf{v}$ can be determined iteratively using (6) and (7), as described in Sec. II-B. However, $\mathbf{\Delta}$ determines the channel transition matrix, which in turn affects $\mathbf{u}$ and $\mathbf{v}$. Conversely, the optimal $\mathbf{\Delta}$ also depends on $\mathbf{u}$ and $\mathbf{v}$ through the distribution $p(s_i)$ in (1).

To address this interdependency, we propose an iterative algorithm to solve (8). First, we initialize a starting point $\mathbf{\Delta}^{(0)}$ that satisfies the constraints of (8). Next, the corresponding channel transition matrix $P=[P_{i,j}]$ is computed, which can be approximated as a tridiagonal matrix using (3) and (4). Given $P=[P_{i,j}]$, the corresponding $\mathbf{u}$ and $\mathbf{v}$ are iteratively updated using (6) and (7). Subsequently, $\mathbf{\Delta}$ is optimized for the updated $\mathbf{u}$ and $\mathbf{v}$. This process is then repeated iteratively.

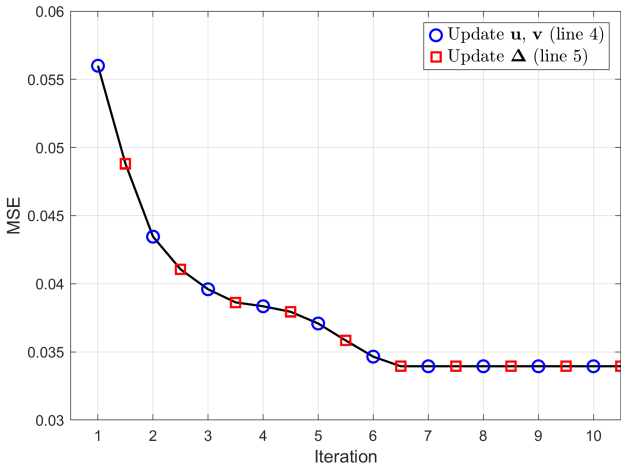


Fig. 3. Iterative joint optimization for $X\sim\mathcal{N}(0,1)$, $\sigma=0.2$, $W=5$, and $M=16$.

The following proposition describes how to update $\mathbf{\Delta}$ for given $\mathbf{u}$ and $\mathbf{v}$.

Proposition 1: For given $\mathbf{u}$ and $\mathbf{v}$, $\mathbf{\Delta}$ is updated by solving the following *convex* optimization problem:

$$\begin{aligned}\underset{\mathbf{\Delta}}{\text{minimize}}\quad&\sum_{i=1}^{M-1}\left\{\gamma_{i,i+1}\cdot p(s_i)Q\left(\frac{\Delta_{i,i+1}}{\sigma_i}\right)\right.\\&\left.+\gamma_{i+1,i}\cdot p(s_{i+1})Q\left(\frac{\Delta_{i+1,i}}{\sigma_{i+1}}\right)\right\}\\\text{subject to}\quad&\|\mathbf{\Delta}\|_1=W\\&\mathbf{\Delta}\succeq 0,\end{aligned}\tag{9}$$

where $\gamma_{i,j}=(\widetilde{v}_i-v_j)^2$, $p(s_i)=\int_{u_i}^{u_{i+1}}f_X(x)dx$, and

$$\widetilde{v}_i=\frac{\int_{u_i}^{u_{i+1}}xf_X(x)dx}{\int_{u_i}^{u_{i+1}}f_X(x)dx}.\tag{10}$$

Proof: The proof is given in Appendix A. ■

Note that $p(s_i)$ and $\gamma_{i,j}$ are determined by the given $\mathbf{u}$ and $\mathbf{v}$. Since the optimization problem of (9) is convex, it can be solved using standard methods [17].

The proposed iterative algorithm to jointly optimize $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{\Delta}$ is summarized in Algorithm 1.

Algorithm 1 Joint Optimization Algorithm for Quantization and Verify Levels

1: Choose a starting point $\mathbf{\Delta}^{(0)}$ satisfying the constraints of (8) and set $k=0$.
2: **repeat**
3: Compute $P=[P_{i,j}]$ by using $\mathbf{\Delta}^{(k)}$.
4: Iteratively update $\mathbf{u}^{(k)}$ and $\mathbf{v}^{(k)}$ by (6) and (7).
5: Update $\mathbf{\Delta}^{(k+1)}$ for given $\mathbf{u}^{(k)}$ and $\mathbf{v}^{(k)}$. ▷ Prop. 1
6: $k:=k+1$.
7: **until** Stopping criterion is satisfied.

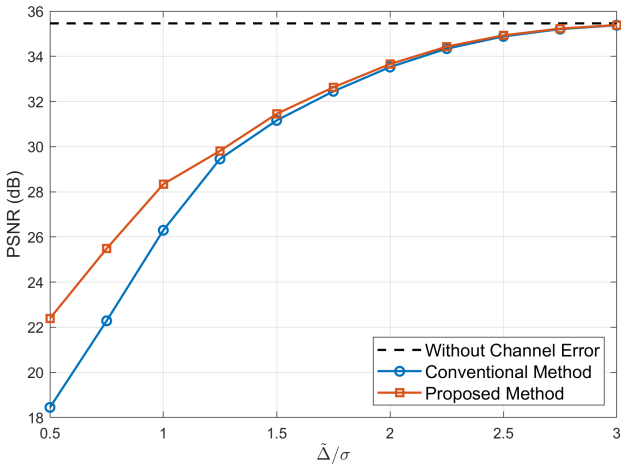


Fig. 4. Comparison of PSNRs for image retrieval.

(a) (b) (c)

Fig. 5. Visual comparison for image retrieval on the image "*test068*" from BSD68 at $\widetilde{\Delta}/\sigma = 0.75$: (a) original image, (b) conventional method (17.92 dB), and (c) proposed method (23.13 dB).

Remark 1: Since (8) is non-convex, a starting point can affect the final solutions. To address this, we begin by considering a noiseless channel where $P_{i,j} = \delta_{i,j}$, and determine $\mathbf{u}$ and $\mathbf{v}$ using the Lloyd-Max algorithm. Next, we initialize a starting point $\mathbf{\Delta}^{(0)}$ that minimizes the overall BER. Subsequently, we update $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{\Delta}$ using Algorithm 1.

Remark 2 (Stopping Criterion): The stopping criterion can be defined in several ways. For example, it may be based on the difference between $\mathrm{MSE}(\mathbf{u}^{(k)}, \mathbf{v}^{(k)}, \mathbf{\Delta}^{(k)})$ and $\mathrm{MSE}(\mathbf{u}^{(k+1)}, \mathbf{v}^{(k+1)}, \mathbf{\Delta}^{(k+1)})$, or the difference between $\mathbf{\Delta}^{(k)}$ and $\mathbf{\Delta}^{(k+1)}$. Alternatively, a maximum number of iterations can be used as the stopping criterion.

Fig. 3 shows the iterative reduction of the MSE over 10 iterations, eventually converging.

IV. Numerical Results

To evaluate the effectiveness of our proposed method, we assess its performance in image retrieval and image super-resolution tasks by considering 4-bit per cell quad-level cell (QLC) flash memory. We compare the results of our proposed method with those of a conventional approach. The conventional method consists of two separate steps: 1) quantization is performed without accounting for channel errors, which reduces to the solutions of the Lloyd-Max algorithm, 2) verify levels are optimized to minimize the overall BER as in [10] using the state probabilities obtained from the first step.

We define $\widetilde{\Delta} = \frac{W}{2(M-1)}$, representing the average value of Δ. When solving (9), we assume uniform standard deviations across all states for simplicity. However, the proposed algorithm is general and can accommodate arbitrary noise variances and window sizes.

A. Quantized Image Retrieval

In the image retrieval experiments, we utilize the Berkeley Segmentation Dataset 68 (BSD68), a simplified version of the BSD100 [18] containing gray scale images. Each pixel of a grayscale image, represented as an 8-bit unsigned integer, is quantized to 4 bits and subsequently stored in an individual QLC flash cell.

Fig. 4 compares the PSNRs of retrieved images using the proposed method with those restored by the conventional method. The numerical results indicate that the proposed method significantly outperforms the conventional approach under higher noise variance, i.e., lower $\widetilde{\Delta}/\sigma$. For low noise variance, which is close to a noise-free channel, the PSNRs of the conventional method and the proposed method are nearly identical. Additionally, the visual comparisons in Fig. 5 demonstrate that the proposed method produces noticeably fewer artifacts, resulting in improved overall perceptual quality.

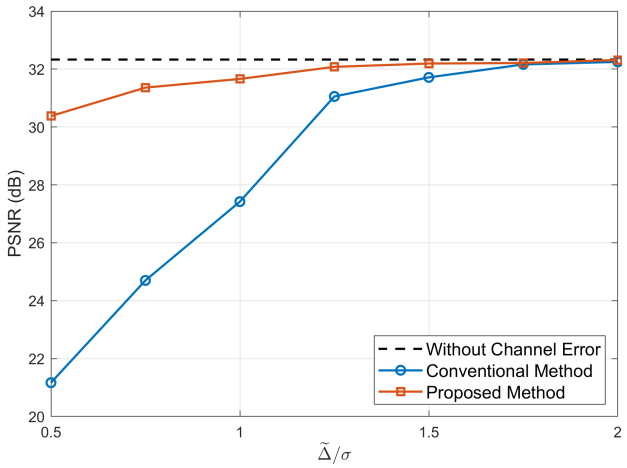
B. Image Super-Resolution

We employ the SwinIR-S($\times 3$) model [16] for the image super-resolution task on two datasets, Set5 [19] and Set14 [20], commonly used for image quality evaluation. In the SwinIR model, the parameters are selectively quantized to 4 bits, focusing on the multi-head self-attention (MSA) and multi-layer perception (MLP) sublayers. Specifically, the quantization is applied to the QKV projection, multi-head concatenation projection, and the two fully connected layers' parameters, as these modules account for the majority of the computational workload and parameter count in transformer-based architectures. Each quantized parameter is stored in a single QLC cell, with identical verify levels applied across all parameters.

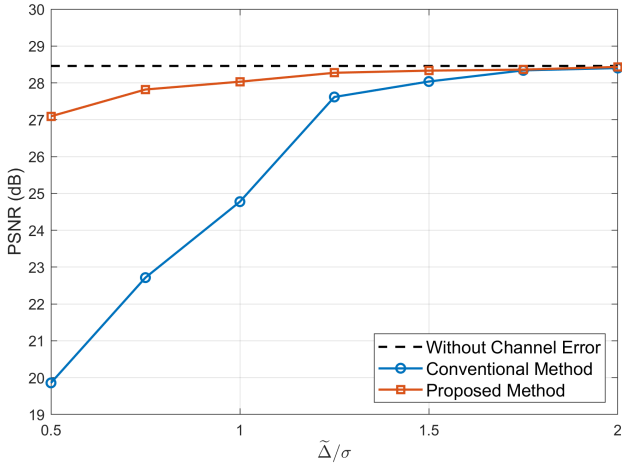
Fig. 6 compares the PSNRs of the SwinIR model for the conventional and proposed methods. The proposed method consistently outperforms the conventional approach, especially in the high noise variance range. This indicates that the proposed method is robust to channel errors particularly in the high-error range. Fig. 7 presents a visual comparison of SwinIR for image super-resolution. The results clearly show that the proposed method delivers sharper details, reduced artifacts, and higher-quality reconstructions, significantly outperforming the conventional method in super-resolution tasks.

V. Conclusion

We proposed a novel integrated optimization of quantization and verify levels for flash-based PIM systems. After formulating the joint optimization problem, we developed an iterative algorithm to improve the MSE under a given constraint. Experimental results show that the proposed method effectively improves MSE performance for both image quantization and image super-resolution tasks.

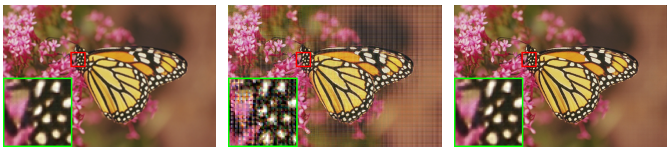


(a)



(b)

Fig. 6. Comparison of PSNRs for image super-resolution by SwinIR: (a) Set5 and (b) Set14.



(a) (b) (c)

Fig. 7. Visual comparison for the image super-resolution task on the "*monarch*" image from Set14 at $\widetilde{\Delta}/\sigma = 0.5$: (a) low resolution, (b) conventional method (16.12 dB), and (c) proposed method (31.62 dB).

Appendix A
Proof of Proposition 1

The MSE can be expressed as follows:

$$\begin{aligned}
&\mathbb{E}_{X,\widehat{X}}[(X-\widehat{X})^2] \\
&= \mathbb{E}_{Y,\widehat{X}}[\mathbb{E}_X[(X-\widehat{X})^2|Y,\widehat{X}]] \\
&= \mathbb{E}_{Y,\widehat{X}}\left[\mathbb{E}_X[(X-\mathbb{E}[X|Y])^2|Y] + (\mathbb{E}[X|Y]-\widehat{X})^2\right] \\
&= \mathbb{E}_{X,Y}\left[(X-\mathbb{E}[X|Y])^2\right] \\
&\quad + \sum_{i=1}^{M}\sum_{j=1}^{M} p_{Y,\widehat{X}}(v_i,v_j)\left(\mathbb{E}[X|Y=v_i]-v_j\right)^2, \qquad (11)
\end{aligned}$$

where $\mathbb{E}[X|Y=v_i]=\widetilde{v}_i$ by (10). Hence, $(\mathbb{E}[X|Y=v_i]-v_j)^2$ can be replaced with $\gamma_{i,j} = (\widetilde{v}_i - v_j)^2$ for $i,j \geq 1$. In (11), the first term is independent of optimizing the verify levels. Consequently, we focus on the second term, which explicitly accounts for the channel errors. The second term can be reformulated as follows:

$$\begin{aligned}
&\sum_{i=1}^{M}\sum_{j=1}^{M} p_{Y,\widehat{X}}(v_i,v_j)\left(\widetilde{v}_i - v_j\right)^2 \\
&= \sum_{i=1}^{M} p_Y(v_i)\sum_{j=1}^{M} p_{\widehat{X}|Y}(v_j|v_i)\left(\widetilde{v}_i - v_j\right)^2 \qquad (12) \\
&\approx \sum_{i=1}^{M-1} p_Y(v_i)P_{i,i+1}\cdot\gamma_{i,i+1} + p_Y(v_{i+1})P_{i+1,i}\cdot\gamma_{i+1,i} \qquad (13) \\
&\approx \sum_{i=1}^{M-1}\left\{p(s_i)Q\left(\frac{\Delta_{i,i+1}}{\sigma_i}\right)\gamma_{i,i+1}\right. \\
&\quad \left. + p(s_{i+1})Q\left(\frac{\Delta_{i+1,i}}{\sigma_{i+1}}\right)\gamma_{i+1,i}\right\}, \qquad (14)
\end{aligned}$$

where (13) follows from the tridiagonal channel transition matrix, with the assumption that $\gamma_{i,i} \ll 1$. Finally, (14) follows from $p_Y(v_i) = p_S(s_i)$, (3), and (4).

Note that (14) is convex with respect to $\boldsymbol{\Delta}$ since $Q(\cdot)$ is a convex function and the coefficients are nonnegative.

Acknowledgment

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea Government (MSIT) (RS-2023-00229849, MPU/Connectivity/TinyML SoC solution for IoT intelligence with foundry-based eFLASH) and the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. RS-2023-00212103).

References

[1] E. Dupraz, F. Leduc-Primeau, K. Cai, and L. Dolecek, "Turning to information theory to bring in-memory computing into practice," *IEEE BITS Inform. Theory Mag.*, vol. 3, no. 3, pp. 64–77, Sep. 2023.

[2] M. Kang, Y. Kim, A. D. Patil, and N. R. Shanbhag, "Deep in-memory architectures for machine learning–accuracy versus efficiency trade-offs," *IEEE Trans. Circuits Syst. I*, vol. 67, no. 5, pp. 1627–1639, May 2020.

[3] N. R. Shanbhag, N. Verma, Y. Kim, A. D. Patil, and L. R. Varshney, "Shannon-inspired statistical computing for the nanoscale era," *Proc. IEEE*, vol. 107, no. 1, pp. 90–107, Jan. 2019.

[4] M. M. Hasan and B. Ray, "Reliability of NAND flash memory as a weight storage device of artificial neural network," *IEEE Trans. Device Mat. Rel.*, vol. 20, no. 3, pp. 596–603, Sep. 2020.

[5] S. K. Gonugondla, M. Kang, Y. Kim, M. Helm, S. Eilert, and N. Shanbhag, "Energy-efficient deep in-memory architecture for NAND flash memories," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2018, pp. 1–5.

[6] Y. Cai, S. Ghose, E. F. Haratsch, Y. Luo, and O. Mutlu, "Error characterization, mitigation, and recovery in flash-memory-based solid-state drives," *Proc. IEEE*, vol. 105, no. 9, pp. 1666–1704, Sep. 2017.

[7] Y. Kim, E. Hwang, and B. V. K. Vijaya Kumar, "Writing on dirty flash memory: Combating inter-cell interference via coding with side information," *J. Commun. Netw.*, vol. 24, no. 6, pp. 633–644, Dec. 2022.

[8] C. A. Aslam, Y. L. Guan, and K. Cai, "Read and write voltage signal optimization for multi-level-cell (MLC) NAND flash memory," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1613–1623, Apr. 2016.

[9] A. Asmani, S. Galijasevic, and R. D. Wesel, "Write voltage optimization to increase flash lifetime in a two-variance Gaussian channel," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2024, pp. 1143–1148.

[10] Y. Kim, J. Kim, J. J. Kong, B. V. K Vijaya Kumar, and X. Li, "Verify level control criteria for multi-level cell flash memories and their applications," *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 1, p. 196, Sep. 2012.

[11] K.-D. Suh, B.-H. Suh, Y.-H. Lim, J.-K. Kim, Y.-J. Choi, Y.-N. Koh, S.-S. Lee, S.-C. Kwon, B.-S. Choi, J.-S. Yum, J.-H. Choi, J.-R. Kim, and H.-K. Lim, "A 3.3 V 32 mb NAND flash memory with incremental step pulse programming scheme," *IEEE J. Solid-State Circuits*, vol. 30, no. 11, pp. 1149–1156, Nov. 1995.

[12] H. Cılasun, S. Resch, Z. I. Chowdhury, M. Zabihi, Y. Lv, B. Zink, J.-P. Wang, S. S. Sapatnekar, and U. R. Karpuzcu, "On error correction for nonvolatile processing-in-memory," in *Proc. ACM/IEEE Annu. Int. Symp. Comput. Archit. (ISCA)*, Jun. 2024, pp. 678–692.

[13] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.

[14] J. Max, "Quantizing for minimum distortion," *IEEE Trans. Inf. Theory*, vol. 6, no. 1, pp. 7–12, Mar. 1960.

[15] A. Kurtenbach and P. Wintz, "Quantizing for noisy channels," *IRE Trans. Commun. Syst.*, vol. 17, no. 2, pp. 291–302, Apr. 1969.

[16] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, "SwinIR: Image restoration using swin transformer," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2021, pp. 1833–1844.

[17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, Mar. 2004.

[18] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 2, Aug. 2002, pp. 416–423.

[19] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. A. Morel, "Low-complexity single-image super-resolution based on nonnegative neighbor embedding," in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, Sep. 2012, pp. 1–10.

[20] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Proc Int. Conf. Curves Surfaces*, Jun. 2012, pp. 711–730.