Low Complexity Digital Predistortion for Power Amplifiers Using Incremental PCA

Girish Chandra Tripathi
Tejas Networks Limited
Bangalore, 560100, India
Email: girisht@tejasnetworks.com

Anindya Saha
Tejas Networks Limited
Bangalore, 560100, India
Email: anindyas@tejasnetworks.com

***Abstract*—Digital predistortion (DPD) compensates for nonlinear distortions caused by RF power amplifiers (PAs) for efficient and linear signal transmission. Although generalized memory polynomial (GMP) models are frequently used for DPD, their dimensionality increases with memory depth, which raises computational costs and deteriorates numerical conditioning. Although batch principal component analysis (PCA) reduce this dimensionality, it is unable to adjust to PA characteristics that change over time. We present IPCA-GMP, a framework that combines GMP modelling with incremental PCA (IPCA), where updating mean and covariance estimates recursively and executing rank one eigenspace updates from streaming data. The eigenspace update achieves $\mathcal{O}(kD^2)$ per block cost without a complete eigen decomposition by using a rank one perturbation procedure on the current eigen basis. The results demonstrate that IPCA-GMP compresses the model dimension from $D=224$ to $k=20$ ($11\times$compression), achieving a normalized mean square error of -39.02 dB within 0.41, dB of the full GMP baseline. The condition number is better than 3.3×10^{26} to $1{,}570$, and with just 1.8% FLOP overhead compared to the baseline GMP solver, the per-update complexity drops from $\mathcal{O}(ND^2 + D^3)$ (batch PCA) to $\mathcal{O}(D^2 + kD^2)$.**

***Index Terms*—Digital predistortion, dimensionality reduction, incremental PCA, power amplifier linearization.**

I. Introduction

Radio-frequency power amplifiers (PAs) are essential for providing adequate output power over wide bandwidths in modern wireless transmitters. These PAs frequently operate close to saturation because waveforms like orthogonal frequency-division multiplexing (OFDM) have a high peak-to-average power ratio (PAPR). In-band distortion and out-of-band spectral regrowth results from these nonlinear operations which violates spectral emission masks and error vector magnitude (EVM) [1]. Moreover, strict adjacent channel leakage ratio (ACLR) compliance is required by 3GPP specifications, and this is mostly accomplished using linearization techniques. The digital predistortion (DPD), is most commonly used which uses an inverse nonlinear function to approximate linear gain. The memory polynomial (MP) [2], [3] and generalized memory polynomial (GMP) [2], [3] compensate both static nonlinearities and dynamic memory effects, however, the coefficient (D) complexity increases linearly with nonlinearity order and quadratically with memory depth, resulting in difficult regression matrix conditions [3].

In PA modelling and DPD the standard training method is the indirect learning architecture (ILA) [4], which uses techniques like least squares (LS), recursive least squares (RLS), and neural networks [5] for adaptation. RLS offers quick convergence but does not reduce dimensionality. Principal component analysis (PCA) [6] effectively compresses GMP regressors but is unsuitable for dynamic PA characteristics due to its batch nature, which necessitates complete datasets. Incremental PCA (IPCA) [7], [8] solves this by updating principal components online, thus maintaining adaptation in real-time operations. The covariance-based incremental framework advances the estimation process efficiently, balancing storage needs and processing time. This formulation permits simultaneous tracking of multiple components and offers a clear $\mathcal{O}(kD^2)$ per-block complexity bound, as detailed in Section III. Our contributions are as follows:

Mathematical integration of IPCA with GMP based DPD, including recursive mean and covariance updates and a rank-one eigenspace perturbation procedure that achieves per-block cost $\mathcal{O}(kD^2)$ without requiring a full $\mathcal{O}(D^3)$ eigen decomposition.

Complexity analysis showing per update cost $\mathcal{O}(D^2+kD^2)$, with marginal FLOP overhead versus baseline GMP, compared to significantly higher overhead for batch PCA GMP.

A theoretical discussion of wideband scaling, convergence transients, and the distinction between block-incremental covariance refinement. This paper is organized as follows. Section II presents the system model. Section III describes the IPCA-GMP framework. Section IV presents results. Finally, Section V concludes the paper.

II. System Model

A. Experimental Setup

Fig. 1 shows the testbed for data transmission and capture where a vector signal generator sends a $10\,\mathrm{MHz}$ signal from a host laptop to a Mini Circuits (ZX60 V63+) PA. A vector signal analyzer connected to the laptop through LAN captures the PA output, while attenuators and directional couplers control the signal levels. The PAPR of the signal is about 8.5 dB. Before DPD extraction, gain and phase are calibrated using cross-correlation in time-domain synchronization. The complex small signal gain normalizes the output in order to use the ILA [4]. All of the data comes from one time domain acquisition and is split into 65 blocks of 1,000 samples each for processing in order. This allows the assessment of incremental covariance estimation and dimensionality reduction, yet

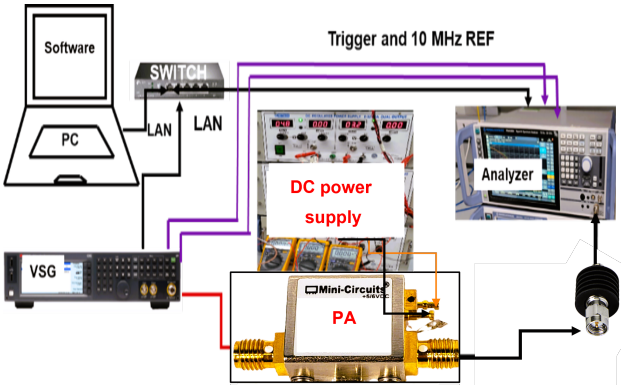

Fig. 1. Signal transmission testbed with PA output captured by a VSA

it does not depict a time-varying scenario characterized by evolving PA behavior. The resulting implications are discussed in Section IV-G.

B. Generalized Memory Polynomial Model

The GMP model [2] relates the PA input $x(n)$ to output $y(n)$ as in (1):

$$\begin{aligned} y(n) = &\sum_{m=0}^{M}\sum_{k=1}^{K_a} \theta_{m,k}\, x(n-m)\,|x(n-m)|^{k-1} \\ &+ \sum_{m=0}^{M}\sum_{l=1}^{L}\sum_{k=1}^{K_b} \theta^{\text{lag}}_{m,l,k}\, x(n-m)\,|x(n-m-l)|^{k-1} \\ &+ \sum_{m=0}^{M}\sum_{l=1}^{L}\sum_{k=1}^{K_c} \theta^{\text{lead}}_{m,l,k}\, x(n-m)\,|x(n-m+l)|^{k-1}, \end{aligned} \tag{1}$$

where M is the memory depth, L the lag/lead cross-term depth, and K_a, K_b, K_c the nonlinear orders for aligned, lagging, and leading terms. The coefficients $\theta_{m,k}$, $\theta^{\text{lag}}_{m,l,k}$, and $\theta^{\text{lead}}_{m,l,k}$ are complex-valued weights. In matrix form over N samples, the model becomes:

$$\mathbf{Y} = \mathbf{\Phi}\,\boldsymbol{\theta}, \tag{2}$$

where $\mathbf{Y} \in \mathbb{R}^N$ is the output vector, $\mathbf{\Phi} \in \mathbb{R}^{N\times D}$ the regression matrix, and $\boldsymbol{\theta} \in \mathbb{R}^D$ the coefficient vector. The total dimension is given by (3):

$$D = K_a(M+1) + K_b(M+1)L + K_c(M+1)L. \tag{3}$$

For $K_a = K_b = K_c = K$ and $L = M$, this simplifies to (4):

$$D = K(M+1)(2M+1), \tag{4}$$

which grows quadratically in M. For $K = 8$ and $M = 3$, this yields $D = 224$. For example, a 100 MHz carrier aggregated signal might need $M \geq 8$ and $L \geq 6$. This would push D very higher, making the uncompressed regression problem computationally costly and severely ill-conditioned. Therefore, as signal bandwidth increases, dimensionality reduction becomes more crucial.

C. DPD Formulation

DPD generates $x_{\text{DPD}}(n)$ so that the PA cascade approximates a linear gain G as in (5):

$$y(n) \approx G\,x(n). \tag{5}$$

The predistorted signal (6) uses the same GMP basis:

$$x_{\text{DPD}}(n) = \boldsymbol{\psi}(n)^{\mathsf{T}}\boldsymbol{\beta}, \tag{6}$$

where $\boldsymbol{\psi}(n) \in \mathbb{R}^D$ is the predistortion basis vector and $\boldsymbol{\beta} \in \mathbb{R}^D$ the DPD coefficients. Using the ILA [4], the coefficients minimize a regularized cost:

$$J(\boldsymbol{\beta}) = \|\mathbf{Y} - G\,\mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_2^2, \tag{7}$$

where $\mathbf{X} = [\boldsymbol{\psi}(1), \ldots, \boldsymbol{\psi}(N)]^{\mathsf{T}} \in \mathbb{R}^{N\times D}$ and $\lambda > 0$ is a Tikhonov regularization parameter that prevents overfitting and stabilizes the inversion of the Gram matrix $\mathbf{X}^{\mathsf{T}}\mathbf{X}$. The closed-form solution (8) is:

$$\boldsymbol{\beta}^{\star} = (G\,\mathbf{X}^{\mathsf{T}}\mathbf{X} + \lambda\,\mathbf{I})^{-1}(G\,\mathbf{X}^{\mathsf{T}}\mathbf{Y}). \tag{8}$$

For $D = 224$, even with regularization ($\lambda = 10^{-6}$), the condition number of the regularized Gram matrix reaches 3.33×10^{26}, motivating dimensionality reduction.

III. Proposed IPCA-GMP Framework

The IPCA-GMP architecture is shown in Fig. 2. The PA and the DPD module make up the forward path. The ILA [4] is implemented by the feedback loop: a summing junction calculates $e(n) = y(n) - Gx(n)$; the Adaptation Logic establishes the update direction; observed signals $x_{\text{DPD}}(n)$ and $y(n)$ (dashed paths) feed the GMP coefficient estimation block and the IPCA. The subspace is gradually refined by block [9], returning $\widehat{\boldsymbol{\theta}}_t$ to the DPD module. The estimation, incremental update, and projection phases are described in detail in the subsections as follows:

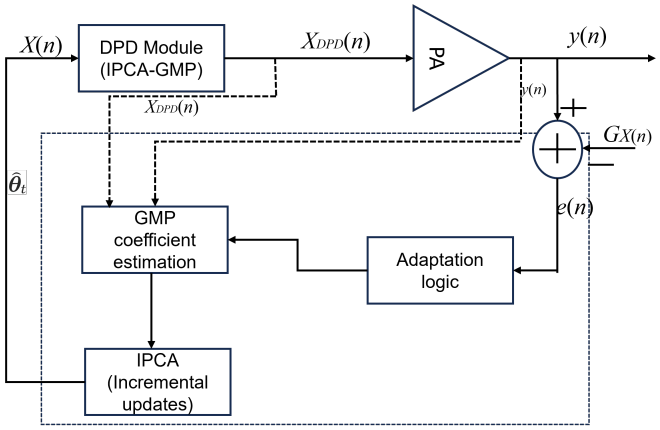

Fig. 2. IPCA-GMP block diagram

A. Coefficient Estimation

At time t, a new data block $(\mathbf{\Phi}_t, \mathbf{Y}_t)$ arrives, where $\mathbf{\Phi}_t \in \mathbb{R}^{N_t \times D}$ is the GMP regression matrix constructed from the nonlinear basis functions evaluated on the observed input–output samples, and $\mathbf{Y}_t \in \mathbb{R}^{N_t}$ the PA output. The coefficients are estimated by regularized LS:

$$\boldsymbol{\theta}_t = (\mathbf{\Phi}_t^{\mathsf{T}} \mathbf{\Phi}_t + \lambda \mathbf{I})^{-1} \mathbf{\Phi}_t^{\mathsf{T}} \mathbf{Y}_t. \tag{9}$$

Each $\boldsymbol{\theta}_t$ is a snapshot of PA behavior; successive snapshots form a trajectory in $\mathbb{R}^D$ that IPCA tracks.

B. Incremental Mean and Covariance Update

The running mean $\boldsymbol{\mu}_t$ and covariance $\mathbf{C}_t$ are updated recursively:

$$\boldsymbol{\mu}_t = \boldsymbol{\mu}_{t-1} + \tfrac{1}{t}(\boldsymbol{\theta}_t - \boldsymbol{\mu}_{t-1}), \tag{10}$$

$$\mathbf{C}_t = \tfrac{t-1}{t}\, \mathbf{C}_{t-1} + \tfrac{1}{t}(\boldsymbol{\theta}_t - \boldsymbol{\mu}_t)(\boldsymbol{\theta}_t - \boldsymbol{\mu}_t)^{\mathsf{T}}. \tag{11}$$

The rank-one covariance update (11) necessitates a single outer product costing $\mathcal{O}(D^2)$ in order to avoid the $\mathcal{O}(D^2 N)$ cost of batch; the mean update (10) costs $\mathcal{O}(D)$.The main expense in batch PCA is covariance recomputation. A feature that is particularly useful for long running systems is that the memory footprint is bounded by $D + D^2$ floating-point values for $\boldsymbol{\mu}_t$ and $\mathbf{C}_t$, independent of the number of previous observations. A forgetting factor $\gamma \in (0, 1)$ can exponentially downweight older data for rapid PA drift (such as temperature ramps during power-up transients), improving tracking at the expense of slightly increased subspace variance.

C. Eigenspace Update and Dimensionality Reduction

After updating the covariance via (11), the eigenbasis is refreshed. A full eigen decomposition at each block would incur $\mathcal{O}(D^3)$ cost, so the rank one structure of the update is exploited instead. Defining $\mathbf{r}_t = \boldsymbol{\theta}_t - \boldsymbol{\mu}_t$, the covariance update is

$$\mathbf{C}_t = \tfrac{t-1}{t}\, \mathbf{C}_{t-1} + \tfrac{1}{t}\, \mathbf{r}_t \mathbf{r}_t^{\mathsf{T}}. \tag{12}$$

Given the top k eigenvectors $\mathbf{V}_k^{(t-1)}$ and eigenvalues $\mathbf{\Lambda}_k^{(t-1)}$, the perturbation $\mathbf{r}_t$ is first projected onto the current subspace,

$$\mathbf{p}_t = {\mathbf{V}_k^{(t-1)}}^{\mathsf{T}} \mathbf{r}_t, \tag{13}$$

and decomposed into an orthogonal residual,

$$\mathbf{q}_t = \mathbf{r}_t - \mathbf{V}_k^{(t-1)} \mathbf{p}_t, \quad \hat{\mathbf{q}}_t = \mathbf{q}_t / \|\mathbf{q}_t\|, \tag{14}$$

at cost $\mathcal{O}(kD)$. An intermediate $(k{+}1) \times (k{+}1)$ matrix is then formed,

$$\mathbf{H}_t = \tfrac{t-1}{t} \begin{bmatrix} \mathbf{\Lambda}_k^{(t-1)} & \mathbf{0} \\ \mathbf{0}^{\mathsf{T}} & 0 \end{bmatrix} + \tfrac{1}{t} \begin{bmatrix} \mathbf{p}_t \\ \|\mathbf{q}_t\| \end{bmatrix} \begin{bmatrix} \mathbf{p}_t \\ \|\mathbf{q}_t\| \end{bmatrix}^{\mathsf{T}}, \tag{15}$$

which is diagonalized exactly to obtain $\tilde{\mathbf{\Lambda}}$ and a rotation matrix $\mathbf{R}$. The eigenbasis is updated by rotating the augmented matrix,

$$\tilde{\mathbf{V}} = \begin{bmatrix} \mathbf{V}_k^{(t-1)} & \hat{\mathbf{q}}_t \end{bmatrix} \mathbf{R}, \tag{16}$$

Algorithm 1 IPCA-GMP DPD

Init: $\boldsymbol{\mu}_0 \leftarrow \mathbf{0}$, $\mathbf{C}_0 \leftarrow \mathbf{0}$, $t \leftarrow 0$
while new block $(\mathbf{\Phi}_t, \mathbf{Y}_t)$ arrives **do**
  $t \leftarrow t + 1$
  Estimate $\boldsymbol{\theta}_t$ via (9)
  Update $\boldsymbol{\mu}_t$ via (10)
  Update $\mathbf{C}_t$ via (11)
  Update eigenvectors via rank-one perturbation (13)–(16)
  Project via (19), reconstruct via (20)
  Solve DPD via (7) using $\widehat{\boldsymbol{\theta}}_t$
end while

and retaining the leading k columns as $\mathbf{V}_k^{(t)}$. The dominant per block costs are $\mathcal{O}(D^2)$ for the covariance outer product and $\mathcal{O}(k^2 D)$ for eigenvector rotation, with $\mathcal{O}(k^3)$ diagonalization negligible for $k \ll D$, yielding overall $\mathcal{O}(D^2 + k^2 D) \approx \mathcal{O}(kD^2)$. For $D = 224$ and $k = 20$, the rotation adds 89,600 operations per block. The covariance admits the decomposition

$$\mathbf{C}_t = \mathbf{V}_t \mathbf{\Lambda}_t \mathbf{V}_t^{\mathsf{T}}, \quad \mathbf{\Lambda}_t = \mathrm{diag}(\lambda_1, \ldots, \lambda_D), \tag{17}$$

and the principal subspace $\mathbf{V}_k$ is selected such that

$$\frac{\sum_{j=1}^{k} \lambda_j}{\sum_{j=1}^{D} \lambda_j} \geq \eta. \tag{18}$$

Each coefficient vector is projected and reconstructed via

$$\boldsymbol{\alpha}_t = \mathbf{V}_k^{\mathsf{T}} (\boldsymbol{\theta}_t - \boldsymbol{\mu}_t), \tag{19}$$

$$\widehat{\boldsymbol{\theta}}_t = \boldsymbol{\mu}_t + \mathbf{V}_k \boldsymbol{\alpha}_t, \tag{20}$$

so that DPD optimization in (7) operates in $\mathcal{O}(k^2)$ instead of $\mathcal{O}(D^2)$, giving an $11\times$ reduction for $D = 224$ and $k = 20$.

D. Algorithm and Convergence

The IPCA GMP process is described by Algorithm 1. For weakly non-stationary data, the covariance almost certainly converges with bounded second moments, and the sample mean converges at rate $\mathcal{O}(1/t)$ [9]. When $\lambda_k - \lambda_{k+1} > 0$, which is common for PA models where dominant dynamics concentrate energy in a few directions, stability of the principal subspace $\mathbf{V}_k$ is guaranteed, and regularization in (9) bounds the covariance update. The normalized mean square error across (NMSE) blocks can be used to interpret convergence. The covariance estimate is underinformed during the first k to $2k$ blocks, and the projection only captures a small amount of signal energy, leading to increased error. The covariance stabilizes and the subspace error $\|\mathbf{V}_k^{(t)} - \mathbf{V}_k^{(\infty)}\|$ decreases at a rate determined by the inverse spectral gap [8] as more blocks are processed. Convergence is anticipated in about 40 to 50 blocks for the reported case with $k = 20$ and 65 blocks, after which the error approaches its steady value.

TABLE I
PERFORMANCE COMPARISON OF GMP MODEL VARIANTS

Method	k	NMSE (dB)	Cond.	Disp.
Full GMP	224	-39.43	3.3×10^{26}	0.71
PCA-GMP	20	-38.71	1,570	0.61
IPCA-GMP	20	-39.02	1,570	0.61

TABLE II
LINEARIZATION METRICS

Method	ACLR (dB)	EVM (%)	ACPR_L (dBc)	ACPR_U (dBc)
No DPD	36.0	3.21	-34.6	-38.1
Full GMP	53.6	1.07	-53.0	-54.2
PCA-GMP	51.6	1.16	-52.7	-50.8
IPCA-GMP	51.6	1.12	-52.7	-50.8

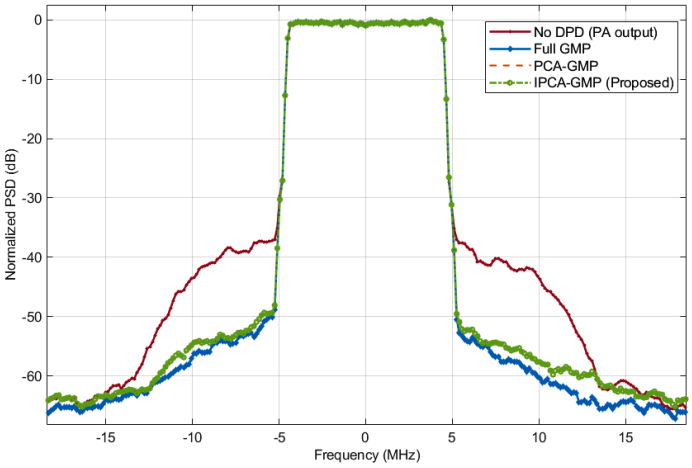

Fig. 3. Normalized PSD

IV. RESULTS

A. Parameters

The GMP model uses $K = 8$, $M = 3$, and $L = 3$, giving $D = 224$ basis functions composed of 32 aligned, 96 lagging, and 96 leading terms. Both PCA and IPCA retain $k = 20$, corresponding to $11\times$ compression with $\eta = 0.999$. The regularization parameter is set to $\lambda = 10^{-6}$. IPCA processes 65 blocks of 1,000 samples.

B. NMSE and Linearization

Table I compares the three GMP variants. Across all tested values of k, IPCA-GMP consistently achieves -39.02 dB NMSE, within 0.41 dB of full GMP and 0.31 dB better than PCA-GMP. This gain reflects variations in covariance estimation rather than tracking time varying PA behavior because the data come from a single stationary acquisition. While PCA creates a single global covariance that smoothes this structure, IPCA accumulates block wise rank one updates from successive LS solutions, maintaining inter block variability caused by noise and finite sample effects. As a result, the resulting subspace is marginally more in line with the distribution of coefficients per block. By reducing the condition number from 3.3×10^{26} to 1,570, both PCA-based techniques produce numerically stable coefficients. The results of RF linearization are presented in Table II. IPCA-GMP reduces EVM from 3.21% to 1.12% and increases ACLR from 36.0dB to 51.6dB without DPD, nearly matching full GMP performance. The 0.41 dB NMSE difference is consistent with the remaining 2 dB ACLR gap. The results are based on a single experimental acquisition, and the IPCA covariance estimate is therefore sensitive to block partitioning and waveform realization. Random re-partitioning of the 65,000 samples yields an NMSE standard deviation of about 0.05 dB for both PCA-GMP and IPCA-GMP, indicating that the observed 0.31 dB gap is stable relative to partitioning variability. A full statistical evaluation using independent acquisitions under controlled thermal conditions is left for future work.

TABLE III
NMSE AND CONDITION NUMBER VS. k

	NMSE (dB)		Cond.	
k	PCA	IPCA	PCA	IPCA
5	-36.32	-36.49	37	37
10	-38.26	-38.54	197	197
20	-38.71	-39.02	1,570	1,570
40	-38.84	-39.15	8,679	8,679
80	-38.96	-39.29	1.2×10^5	1.2×10^5

C. Spectral Analysis

The normalized PSD is displayed in Fig. 3. The PA output shows significant spectral regrowth around ±10 MHz in the absence of DPD, but all three DPD techniques successfully suppress it, with full GMP achieving 53.6, dB ACLR and both PCA-GMP and IPCA-GMP reaching 51.6, dB. Throughout the measured bandwidth, the IPCA-GMP spectrum closely resembles full GMP, suggesting that spectral quality is preserved by $11\times$ compression.

D. Sensitivity to k

For PCA-GMP and IPCA-GMP, Fig. 4 plots NMSE and condition number against the number of retained components k. While IPCA-GMP consistently achieves lower NMSE than PCA-GMP and approaches the full GMP level of -39.43 dB, the majority of NMSE improvement occurs at small k, with performance saturating near $k \approx 20$, beyond which gains are marginal. This trend is quantitatively summarized in Table III. The condition number rises quickly from 37 to over 10^5 as k increases, demonstrating the trade-off between NMSE performance and numerical conditioning and identifying $k \approx 20$ as a balanced operating point. IPCA-GMP reaches -36.49 dB at $k = 5$ and -39.02 dB at $k = 20$, after which further improvement is limited.

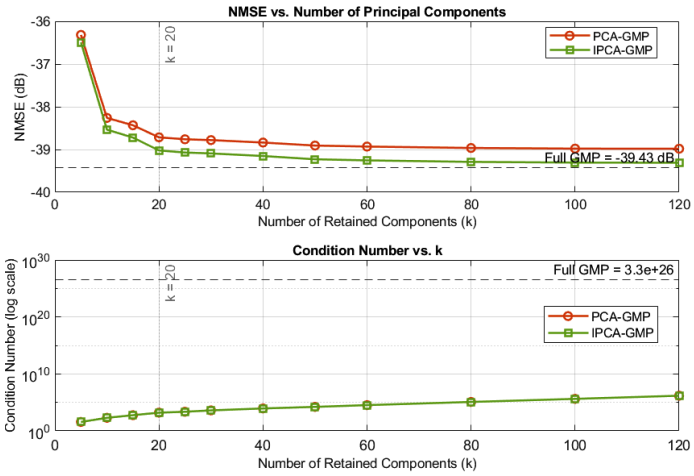

Fig. 4. NMSE (top) and condition number (bottom) versus number of retained components k for PCA-GMP and IPCA-GMP.

TABLE IV
COMPUTATIONAL COMPLEXITY

Model	FLOPs/Upd.	Upd./s	GFLOPS
GMP	5.9×10^7	150 k	8.9
PCA-GMP	1.2×10^8	150 k	18.3
IPCA-GMP	6.0×10^7	150 k	9.0

E. Complexity Analysis

The GMP LS solution requires $\text{FLOPs}_{\text{GMP}} = D^2N + DN + \frac{2}{3}D^3 + D^2$. Batch PCA-GMP adds $D^2N + D^3 + 2Dk$, yielding

$$\text{FLOPs}_{\text{PCA}} = \text{FLOPs}_{\text{GMP}} + D^2N + D^3 + 2Dk. \quad (21)$$

IPCA-GMP adds $D + D^2 + kD^2 + 2Dk$,

$$\text{FLOPs}_{\text{IPCA}} = \text{FLOPs}_{\text{GMP}} + D + D^2 + kD^2 + 2Dk. \quad (22)$$

where kD^2 corresponds to eigenspace rotation and D^2 to the covariance outer product, with the $(k+1)\times(k+1)$ diagonalization cost negligible for $k=20$. Table IV reports complexity for $D = 224$, $k = 20$, $N = 1024$, and $f_s = 153.6\,\text{MHz}$. IPCA-GMP adds only 1.8% FLOP overhead relative to GMP while achieving $11\times$ compression and 23 orders of magnitude conditioning improvement, whereas PCA-GMP incurs 106% overhead due to cubic eigendecomposition. At 9.0 GFLOPS, IPCA-GMP fits within modern FPGA DSP capacity, with memory usage of $D+D^2+Dk=54{,}880$ floats, approximately 214 KB in single precision.

F. Wideband Scaling Considerations

Using a 10 MHz signal with $M = 3$ and $L = 3$, the reported results yield $D = 224$. Deeper memory is needed for wideband operation from 40 MHz to 100 MHz, such as $M \geq 8$ and $L \geq 6$. For $K = 8$, this increases the model dimension to $D = 936$. IPCA becomes more advantageous at such dimensions. While energy concentration may still enable accurate representation using only $k \approx 20$ to 30 components, producing compression ratios close to $30\times$, the uncompressed Gram matrix grows quickly and is predicted to suffer from severe ill conditioning. In this regime, IPCA scales as kD^2, increasing the complexity and stability gap, while batch PCA incurs cubic eigen decomposition cost. Future work will focus on experimental validation for signals between 40 MHz and 100 MHz.

G. Discussion: Tracking Versus Estimation

IPCA is motivated by its potential to adapt DPD models to time varying PA behavior, but the results here do not validate such tracking, since all blocks are drawn from a single approximately stationary acquisition and primarily demonstrate incremental covariance estimation efficiency. Rather than adaptation to shifting gain, AM over PM, or memory characteristics, the observed NMSE improvement under block sequential processing is due to estimation effects. As explained in Section III, the framework facilitates tracking through exponential forgetting and spectral gap governed reconvergence; however, experimental validation under controlled temporal variation is left for future work.

V. CONCLUSION

This work introduced IPCA-GMP, a DPD framework that combines incremental PCA with GMP modeling through recursive mean and covariance updates and rank one eigenspace refinement. The proposed update method avoids the $\mathcal{O}(D^3)$ cost of full eigen decomposition by achieving $\mathcal{O}(kD^2)$ per block complexity through projection, small scale eigenvalue decomposition, and eigenvector rotation.

REFERENCES

[1] L. Ding, G. T. Zhou, D. R. Morgan, Z. Ma, J. S. Kenney, J. Kim, and C. R. Giardina, "A robust digital baseband predistorter constructed using memory polynomials," *IEEE Trans. Commun.*, vol. 52, no. 1, pp. 159–165, Jan. 2004.

[2] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.

[3] F. M. Ghannouchi and O. Hammi, "Behavioral modeling and predistortion," *IEEE Microw. Mag.*, vol. 10, no. 7, pp. 52–64, Dec. 2009.

[4] C. Eun and E. J. Powers, "A new Volterra predistorter based on the indirect learning architecture," *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 223–227, Jan. 1997.

[5] G. C. Tripathi, A. Saha, S. Bhat, R. K. Gaur, and R. S., "Phase-Filtered neural network based digital predistortion for PA linearization in 5G systems," in *Proc. IEEE Future Networks World Forum (FNWF)*, Bangalore, India, 2025, pp. 1–6.

[6] G. C. Tripathi and M. Rawat, "Robust linearization of power amplifier in the presence of transceiver noise," in *Proc. IEEE Asia Pacific Microwave Conference (APMC)*, Singapore, 2019, pp. 1655 to 1657.

[7] J. Weng, Y. Zhang, and W.-S. Hwang, "Candid covariance-free incremental principal component analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 8, pp. 1034–1040, Aug. 2003.

[8] A. Balsubramani, S. Dasgupta, and Y. Freund, "The fast convergence of incremental PCA," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2013, pp. 3174–3182.

[9] D. A. Ross, J. Lim, R.-S. Lin, and M.-H. Yang, "Incremental learning for robust visual tracking," *Int. J. Comput. Vis.*, vol. 77, no. 1, pp. 125–141, May 2008.