Blind Bilinear Approach for Nonlinear-Based System Identification

Abdulmajid Lawal, Karim Abed-Meraim, Azzedine Zerguine, and Ali Muqaibel

Abstract—In this paper, we develop an efficient nonlinear channel identification method for single input multiple output finite impulse response (FIR) channels. The developed algorithm utilizes the structure embedded in the columns and rows subspaces of the received signal matrix. Both the Toeplitz structure available in the signal matrix and the block Sylvester structure present in the channel matrix is used to develop a criterion that can be minimized to establish the optimal solution of the channel estimates. With nonlinearity in the system, the proposed bilinear nonlinear approach produces some extremely intriguing channel estimation findings.

***Keywords*—Blind identification, nonlinear SIMO channel, bilinear, subspace method.**

NOMENCLATURE

This section includes the abbreviations used in the manuscript:

FIR	Finite Impulse Response.
SIMO	Single Input Multiple Output.
SCS	Structured Channel Subspace
BSCS	Nonlinear Biliear Structured Channel Subspace.
LBSCS	Linear Biliear Structured Channel Subspace.

I. INTRODUCTION

Many practical systems have inherent nonlinear behaviors which necessitate the use of dedicated processing especially when such nonlinearities can significantly impact the input signal restoration [1, 2]. The analysis and solutions to nonlinear problems have attracted different specialties such as engineers, mathematicians, and physicists. In particular, the transmission and reception of signals in communication systems involve the use of nonlinear devices such as power amplifiers and optical equipment [3]. Hence, communication channels may be corrupted as a result of nonlinear distortions caused by nonlinear multiple access interference, intersymbol interference, and inter-carrier interference just to mention a few. The signal obtained at the receiving end may deteriorate significantly as a result of these distortions. To tackle such problems, nonlinear models are deployed to accurately represent the channels and enhance the development of dedicated signal processing techniques that can effectively mitigate nonlinear distortions.

In system identification, the 'linear in parameter' nonlinear models are widely adopted. While the relationship between the input and output of the system is nonlinear, the identification problem is linear concerning the coefficients of the channel. Some typical examples are the Volterra filters [4] and polynomial filters which have been employed widely in several fields such as electrical engineering, mechanical engineering, and control engineering to mention a few [5]. These filters excellently model the real-life behavior of nonlinear phenomena and their memory effect. In nonlinear system identification, different Volterra filter-based approaches have been proposed. Some of these approaches are adaptive and use training symbols [6] and are typically based on least mean squares algorithms, recursive least mean square algorithms and affine projection algorithm [7]. Others are fully blind such as higher-order output cumulant-based method [8], and the subspace-based method [9, 10].

Moreover, it is worth pointing out here that the techniques presented in the recent studies of [11] and [12] were originally developed for, and successfully applied to, linear systems for which they exhibited an excellent performance. This has therefore provided us with ample encouragement to extend, in this work, these techniques to nonlinear systems so that their performance could be used as a baseline against which the performance of our newly-developed algorithm could be directly compared, thus providing us with a fair and reliable way to assess the level of improvement achieved by our proposed algorithm.

To emphasize more, the work in [13] is designed for linear systems and employs a linear model while the proposed work is designed for nonlinear systems and employs a nonlinear model. While this difference in approaches makes the comparison between both algorithms infeasible, it emphasizes the generality of our approach in that the nonlinear approach presented in our paper actually subsumes that of the linear approach presented in [13]. As such, our paper ought therefore to represent a useful and important extension of our previous work [13], which offers wider practical applications than can be afforded with the linear approach.

In this work, we propose a bilinear structure subspace method for blind channel estimation of a nonlinear SIMO system. The approach uses the intrinsic Toeplitz structure present in the signal matrix and the block Sylvester structure present in the channel matrix to construct a criterion that is minimized to estab-

Manuscript received December 16, 2024; revised June 12, 2025.

A. Lawal, A. Zerguine and A. Muqaibel are with the Department of Electrical Engineering and the Center for Communication Systems and Sensing, King Fahd University of Petroleum & Minerals, Dhahran, 31261, Saudi Arabia, (e-mail: abdulmajid.lawal@kfupm.edu.sa, azzedine@kfupm.edu.sa, and muqaibel@kfupm.edu.sa)

K. Abed-Meraim is a member of IUF affiliated with University of Orléans, PRISME Lab., France, (e-mail: karim.abed-meraim@univ-orleans.fr)

Digital Object Identifier (DOI): 10.53907/enpesj.v5i1.307

2716-912X © 2025 Ecole Nationale Polytechnique

lish the estimates of the desired channel parameters. Simulation results reported in this work reveal the excellent performance of the proposed algorithm in a nonlinear environment.

Notations: The symbols $()^T$, $()^*$, $()^H$, $()^{-1}$, and $Tr()$, stand for the transpose, conjugate, conjugate transpose, inverse, and the trace operations respectively. A scalar is denoted by a, a vector by $\boldsymbol{a}$, and a matrix by $\boldsymbol{A}$. $\| \, . \, \|_{\mathrm{F}}^2$ represents the Frobenius norm operation. An $a \times a$ dimensional identity matrix is represented by $\boldsymbol{I}_a$, $a \times b$ dimensional zero matrix matrix is represented by $\boldsymbol{0}_{a,b}$. The entry of $\boldsymbol{A}$ at position (i, j) is denoted by $\boldsymbol{A}(i, j)$.

II. The System Model

Let us consider a SIMO nonlinear system consisting of a single transmit antenna, transmitting a signal $s(n)$, and multiple receiving antennas of size N_r as illustrated in Fig. 1. The received signal vector $\boldsymbol{y}(n)$, is given as follows [14]:

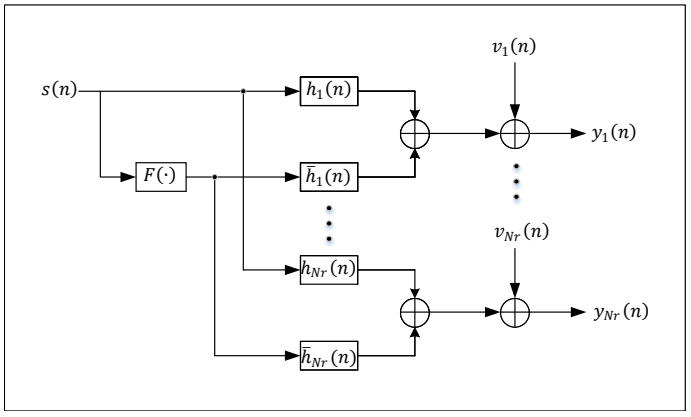


Fig. 1: The proposed SIMO system's block diagram.

$$\begin{aligned} \boldsymbol{y}(n) &= \sum_{k=0}^{L_L} \boldsymbol{h}(k) s(n-k) + \sum_{k=0}^{L_{NL}} \tilde{\boldsymbol{h}}(k) \tilde{s}(n-k) \\ &+ \boldsymbol{v}(n), \end{aligned} \tag{1}$$

where $\boldsymbol{y}(n) = \left[y_1(n) \cdots y_{N_r}(n)\right]^T$, $\boldsymbol{h}(n) = \left[h_1(n) \cdots h_{N_r}(n)\right]^T$ and $\tilde{\boldsymbol{h}}(n) = \left[\tilde{h}_1(n) \cdots \tilde{h}_{N_r}(n)\right]^T$ denote the output signal vector, the channel vector of linear part and the channel vector of the nonlinear part of size $N_r \times 1$, respectively, $h_i(n)$ and $\tilde{h}_i(n)$ are the channel taps of the i^{th} receiving antenna and $\boldsymbol{v}(n) = \left[v_1(n) \cdots v_{N_r}(n)\right]^T$ is an additive white noise of covariance $\sigma_v^2 \boldsymbol{I}_{N_r}$, assumed to be independent of the transmitted signal. L_L and L_{NL}, respectively, represent the channel orders of the linear channel $\boldsymbol{h}(n)$ and the nonlinear channels $\tilde{h}(n)$ parts. The transmitted linear portion of the signal input, which is considered to be an i.i.d. complex random variable, is denoted by $s(n)$, $\tilde{s}(n)$ represents the nonlinear portion and F represents the nonlinear function so that $\tilde{s}(n) = F(s(n), s(n-1), ...)$ [6, 15]. In this work a quadratic nonlinearity is considered i.e., $F(s(n)) = s^2(n)$ due to the fact that many real-life applications have been modeled in this manner, amongst the popular example is the power amplifier and the optical devices [14, 16, 17]. Since the proposed model is linear with respect to the coefficients of the channel, a MIMO model that has two inputs, i.e., $\bar{\boldsymbol{s}}(n) = [s(n) \;\; \tilde{s}(n)]^T$ can be used to represent the proposed SIMO model [14]. As a result, the equivalent MIMO model representation is given as:

$$\boldsymbol{y}(n) = \sum_{k=0}^{L} \boldsymbol{H}(k) \bar{\boldsymbol{s}}(n-k) + \boldsymbol{v}(n), \;\; n = 0, \cdots, N-1, \tag{2}$$

where N is the signal size, the k^{th} channel matrix tap is denoted as $\boldsymbol{H}(k)$ can be expressed as:

$$\boldsymbol{H}(k) = \begin{bmatrix} h_1(k) & \tilde{h}_1(k) \\ \vdots & \vdots \\ h_{N_r}(k) & \tilde{h}_{N_r}(k) \end{bmatrix}. \tag{3}$$

and $L = max\{L_L, L_{NL}\}$. Assuming that N_w samples are successfully stacked into a $\boldsymbol{y}_{N_w}(n)$ vector of $M = N_w N_r$ dimension given as

$$\boldsymbol{y}_{N_w}(n) = [\boldsymbol{y}^T(n) \;\; \boldsymbol{y}^T(n-1) ... \boldsymbol{y}^T(n-N_w+1)]^T, \tag{4}$$

$$\boldsymbol{y}_{N_w}(n) = \boldsymbol{H}_{N_w} \bar{\boldsymbol{s}}_K(n) + \boldsymbol{v}_{N_w}(n), \tag{5}$$

$$\boldsymbol{H}_{N_w} = \begin{bmatrix} \boldsymbol{H}(0) & \cdots & \boldsymbol{H}(L) & & \boldsymbol{0} \\ & \ddots & & \ddots & \\ \boldsymbol{0} & & \boldsymbol{H}(0) & \cdots & \boldsymbol{H}(L) \end{bmatrix}, \tag{6}$$

where $\bar{\boldsymbol{s}}_K(n) = [\bar{\boldsymbol{s}}^T(n) \; \bar{\boldsymbol{s}}^T(n-1) ... \bar{\boldsymbol{s}}^T(n-K+1)]^T$ and $K = N_w + L$ and $\boldsymbol{H}_{N_w}$ represent the block Sylvester channel matrix. Finally, one can set up the data matrix as follows:

$$\begin{aligned} \boldsymbol{Y} &= [\boldsymbol{y}_{N_w}(N_w - 1) \;\; \boldsymbol{y}_{N_w}(N_w), \cdots, \boldsymbol{y}_{N_w}(N-1)] \\ &= \boldsymbol{H}_{N_w} \bar{\boldsymbol{S}}_K + \boldsymbol{V}_{N_w}, \end{aligned} \tag{7}$$

where

$$\bar{\boldsymbol{S}}_K = \begin{bmatrix} \bar{\boldsymbol{s}}(N_w - 1) & \bar{\boldsymbol{s}}(N_w) & \cdots & \bar{\boldsymbol{s}}(N-1) \\ \bar{\boldsymbol{s}}(N_w - 2) & \bar{\boldsymbol{s}}(N_w - 1) & \cdots & \bar{\boldsymbol{s}}(N-2) \\ \vdots & \vdots & \vdots & \\ \bar{\boldsymbol{s}}(-L) & \bar{\boldsymbol{s}}(-L+1) & \cdots & \bar{\boldsymbol{s}}(N-K) \end{bmatrix} \tag{8}$$

In the following sections, the manuscript maintains the following assumptions: The input symbols are sufficiently complex to ensure that matrix $\bar{\boldsymbol{S}}_K$ has a full row rank. The block-Toeplitz matrix $\boldsymbol{H}_{N_w}$ has full column rank.

III. Bilinear-SIMO Nonlinear Channel Estimation Approach

This section derives the proposed nonlinear SIMO bilinear method. The method exploits the information from both the column and row subspaces of matrix $\boldsymbol{Y}$ which are used to develop a criterion that estimates the channel matrix $\boldsymbol{H}_{N_w}$ in an iterative manner. To implement the proposed method, the singular value decomposition (SVD) of the $\boldsymbol{Y}$ matrix is considered:

$$\boldsymbol{Y} = \boldsymbol{U} \boldsymbol{\Sigma} \boldsymbol{V}^H. \tag{9}$$

To start, let $\boldsymbol{U}_s$ represents a sub matrix of $\boldsymbol{U}$ that contains the first $2K$ columns of $\boldsymbol{U}$ with dimension $M \times 2K$, $\boldsymbol{\Sigma}_s$ is a sub matrix of $\boldsymbol{\Sigma}$ that has a dimension of $2K \times 2K$, and $\boldsymbol{V}_s$ is also

Note that in our study F(.) can be considered as a general nonlinear function. Only to test the validity of the proposed work in the sequel, we considered a second order nonlinearity.

formed by the first $2K$ columns of $\boldsymbol{V}$ and has a dimension of $(N - N_w + 1) \times 2K$. The proposed bilinear method for nonlinear system is built by exploiting the structures of both the columns and row subspaces of matrix $\boldsymbol{Y}$ to estimate the channel matrix $\boldsymbol{H}_{N_w}$ (or equivalently $\boldsymbol{S}_K$ the signal matrix) in an iterative manner. In a noiseless case, the data matrix can be expressed exactly as follows:

$$\begin{aligned} \boldsymbol{Y} &= \boldsymbol{U}_s \boldsymbol{\Sigma}_s \boldsymbol{V}_s^H \\ &= \boldsymbol{U}_s \tilde{\boldsymbol{V}}_s \\ &= \boldsymbol{H}_{N_w} \bar{\boldsymbol{S}}_K, \end{aligned} \tag{10}$$

where $\tilde{\boldsymbol{V}}_s = \boldsymbol{\Sigma}_s \boldsymbol{V}_s^H$. It is obvious that the expression $\boldsymbol{H}_{N_w} \bar{\boldsymbol{S}}_K = (\boldsymbol{U}_s \boldsymbol{\mathcal{Q}}) \left(\boldsymbol{\mathcal{Q}}^{-1} \tilde{\boldsymbol{V}}_s \right)$ can be satisfied by any non-singular matrix $\boldsymbol{\mathcal{Q}}$. Therefore, the aim is to search for the matrix $\boldsymbol{\mathcal{Q}}$ such that $\boldsymbol{H}_{N_w} = \boldsymbol{U}_s \boldsymbol{\mathcal{Q}}$ and $\bar{\boldsymbol{S}}_K = \boldsymbol{\mathcal{Q}}^{-1} \tilde{\boldsymbol{V}}_s$ (to the level of ambiguities inherent in blind processing).

The latter equalities are only approximately satisfied in the presence of noise by minimizing a composite criterion relative to the $\boldsymbol{H}_{N_w}$ and $\bar{\boldsymbol{S}}_K$ Sylvester and Toeplitz structure respectively as well as the nonlinear relationship between $\tilde{s}(n)$ and $s(n)$. A nonlinear matrix inversion (i.e., $\boldsymbol{\mathcal{Q}}^{-1}$) is required in such criterion. Hence, an iterative approach paired with a suitable linear approximation of matrix inverse update is proposed as

$$\boldsymbol{\mathcal{Q}}_{new} = \boldsymbol{\mathcal{Q}}_{old}(\boldsymbol{I} + \boldsymbol{\mathcal{E}}) \tag{11}$$

$$\boldsymbol{\mathcal{Q}}_{new}^{-1} \approx (\boldsymbol{I} - \boldsymbol{\mathcal{E}}) \boldsymbol{\mathcal{Q}}_{old}^{-1}, \tag{12}$$

where $\boldsymbol{\mathcal{Q}}_{old}$ and $\boldsymbol{\mathcal{Q}}_{new}$ represent the current and updated value of $\boldsymbol{\mathcal{Q}}$, respectively. Here, $\boldsymbol{\mathcal{E}}$ is used to represent correction matrix, the matrix elements have values that are so small to permit the linear approximation considered. The expression for the combined cost function using(11) and (12) is given as follows:

$$J(\boldsymbol{\mathcal{E}}) = J_1 + J_2 + J_3 + J_4, \tag{13}$$

where J_1 represents the cost function responsible for minimizing the non zero portion of the Toeplitz structure of $\boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{new}$, J_2 minimizes the Toeplitz structure present in $\boldsymbol{\mathcal{Q}}_{new}^{-1} \tilde{\boldsymbol{V}}_s$, J_3 is the cost function responsible for minimizing the zero terms present in the first row and column blocks of $\boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{new}$ and J_4 is to enforce the nonlinear relation between entries $\tilde{s}(n)$ and $s(n)$ in $\boldsymbol{\mathcal{Q}}_{new}^{-1} \tilde{\boldsymbol{V}}_s$.

In the ensuing, the newly proposed cost function details are provided, starting with:

$$\begin{aligned} J_1 &= \| \boldsymbol{J}_a \boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{old}(\boldsymbol{I} + \boldsymbol{\mathcal{E}}) \tilde{\boldsymbol{J}}_a - \boldsymbol{J}_b \boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{old}(\boldsymbol{I} + \boldsymbol{\mathcal{E}}) \tilde{\boldsymbol{J}}_b \|_{\mathrm{F}}^2 \\ &= \| \boldsymbol{A} + \boldsymbol{A}_1 \boldsymbol{\mathcal{E}} \tilde{\boldsymbol{J}}_a - \boldsymbol{A}_2 \boldsymbol{\mathcal{E}} \tilde{\boldsymbol{J}}_b \|_{\mathrm{F}}^2, \end{aligned}$$

where $\boldsymbol{J}_a$, $\tilde{\boldsymbol{J}}_a$, $\boldsymbol{J}_b$, and $\tilde{\boldsymbol{J}}_b$ are all selection matrices that contain ones and zeros and are used to pick the desired portion. These are, respectively, defined as $\boldsymbol{J}_a = [\boldsymbol{I}_{M-N_r} \ \boldsymbol{0}_{M-N_r,N_r}]$, $\tilde{\boldsymbol{J}}_a = [\boldsymbol{I}_{2K-1} \ \boldsymbol{0}_{2(K-1),2}]^T$, $\boldsymbol{J}_b = [\boldsymbol{0}_{M-N_r,N_r} \ \boldsymbol{I}_{M-N_r}]$, $\tilde{\boldsymbol{J}}_b = [\boldsymbol{0}_{2(K-1),2} \ \boldsymbol{I}_{2(K-1)}]^T$. Also, the matrices $\boldsymbol{A}$, $\boldsymbol{A}_1$, and $\boldsymbol{A}_2$ are found to be expressed as $\boldsymbol{A} = \boldsymbol{J}_a \boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{old} \tilde{\boldsymbol{J}}_a - \boldsymbol{J}_b \boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{old} \tilde{\boldsymbol{J}}_b$, $\boldsymbol{A}_1 = \boldsymbol{J}_a \boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{old}$, and $\boldsymbol{A}_2 = \boldsymbol{J}_b \boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{old}$.

A first-order approximation of J_1 can be shown to be:

$$J_1 \approx \| \boldsymbol{A} \|_{\mathrm{F}}^2 + 2Re \left\{ Tr \left((\tilde{\boldsymbol{J}}_a \boldsymbol{A}^H \boldsymbol{A}_1 - \tilde{\boldsymbol{J}}_b \boldsymbol{A}^H \boldsymbol{A}_2) \boldsymbol{\mathcal{E}} \right) \right\},$$

Similarly, the second part of the criterion (13) is given as follows

$$\begin{aligned} J_2 &\approx \| \boldsymbol{J}_c (\boldsymbol{I} - \boldsymbol{\mathcal{E}}) \boldsymbol{\mathcal{Q}}_{old}^{-1} \tilde{\boldsymbol{V}}_s \tilde{\boldsymbol{J}}_c - \boldsymbol{J}_d (\boldsymbol{I} - \boldsymbol{\mathcal{E}}) \boldsymbol{\mathcal{Q}}_{old}^{-1} \tilde{\boldsymbol{V}}_s \tilde{\boldsymbol{J}}_d \|_{\mathrm{F}}^2 \\ &\approx \| \boldsymbol{B} \|_{\mathrm{F}}^2 + 2Re \left\{ Tr \left((\boldsymbol{B}_2 \boldsymbol{B}^H \boldsymbol{J}_d - \boldsymbol{B}_1 \boldsymbol{B}^H \boldsymbol{J}_c) \boldsymbol{\mathcal{E}} \right) \right\}, \end{aligned}$$

where $\boldsymbol{J}_c = \left[\boldsymbol{I}_{2(K-1)} \ \boldsymbol{0}_{2(K-1),2} \right]$, $\tilde{\boldsymbol{J}}_c = \left[\boldsymbol{I}_{N-N_w} \ \boldsymbol{0}_{1,(N-N_w)} \right]^T$, $\boldsymbol{J}_d = \left[\boldsymbol{0}_{2(K-1),2} \ \boldsymbol{I}_{2(K-1)} \right]$, and $\tilde{\boldsymbol{J}}_d = \left[\boldsymbol{0}_{(N-N_w),1} \ \boldsymbol{I}_{N-N_w} \right]^T$, with $\boldsymbol{B} = \boldsymbol{J}_c \boldsymbol{\mathcal{Q}}_{old}^{-1} \tilde{\boldsymbol{V}}_s \tilde{\boldsymbol{J}}_c - \boldsymbol{J}_d \boldsymbol{\mathcal{Q}}_{old}^{-1} \tilde{\boldsymbol{V}}_s \tilde{\boldsymbol{J}}_d$, $\boldsymbol{B}_1 = \boldsymbol{\mathcal{Q}}_{old}^{-1} \tilde{\boldsymbol{V}}_s \tilde{\boldsymbol{J}}_c$ and $\boldsymbol{B}_2 = \boldsymbol{\mathcal{Q}}_{old}^{-1} \tilde{\boldsymbol{V}}_s \tilde{\boldsymbol{J}}_d$.

Finally, J_3 can be expressed as follows:

$$\begin{aligned} J_3 &= \| \boldsymbol{J}_{rw} \boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{old}(\boldsymbol{I} + \boldsymbol{\mathcal{E}}) \tilde{\boldsymbol{J}}_{rw} \|_{\mathrm{F}}^2 \\ &\quad + \| \boldsymbol{J}_{cl} \boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{old}(\boldsymbol{I} + \boldsymbol{\mathcal{E}}) \tilde{\boldsymbol{J}}_{cl} \|_{\mathrm{F}}^2 \\ &\approx \| \boldsymbol{C} \|_{\mathrm{F}}^2 + \| \boldsymbol{D} \|_{\mathrm{F}}^2 \\ &\quad + 2Re \left\{ Tr \left((\tilde{\boldsymbol{J}}_{rw} \boldsymbol{C}^H \boldsymbol{C}_1 + \tilde{\boldsymbol{J}}_{cl} \boldsymbol{D}^H \boldsymbol{D}_1) \boldsymbol{\mathcal{E}} \right) \right\}, \end{aligned}$$

where $\boldsymbol{J}_{rw} = [\boldsymbol{I}_{N_r} \ \boldsymbol{0}_{N_r,N_w N_r - N_r}]$, $\tilde{\boldsymbol{J}}_{rw} = [\boldsymbol{0}_{2(N_w-1),2(L+1)} \ \boldsymbol{I}_{2(N_w-1)}]^T$, $\boldsymbol{J}_{cl} = [\boldsymbol{0}_{M-N_r,N_r} \ \boldsymbol{I}_{M-N_r}]$ and $\tilde{\boldsymbol{J}}_{cl} = [\boldsymbol{I}_2 \ \boldsymbol{0}_{2,2(K-1)}]^T$, $\boldsymbol{C} = \boldsymbol{J}_{rw} \boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{old} \tilde{\boldsymbol{J}}_{rw}$, $\boldsymbol{C}_1 = \boldsymbol{J}_{rw} \boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{old}$, $\boldsymbol{D} = \boldsymbol{J}_{cl} \boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{old} \tilde{\boldsymbol{J}}_{cl}$, and $\boldsymbol{D}_1 = \boldsymbol{J}_{cl} \boldsymbol{U}_s \boldsymbol{\mathcal{Q}}_{old}$.

As for the last part J_4, since $\bar{\boldsymbol{S}}_K \approx \boldsymbol{\mathcal{Q}}_{new}^{-1} \tilde{\boldsymbol{V}}_s$, the linear and nonlinear parts of the matrix can be extracted using appropriate selection matrix $\boldsymbol{J}_e$ and $\boldsymbol{J}_f$. $\boldsymbol{J}_e$ is a matrix of size $K \times 2K$ formed from the odd indexed rows of an identity matrix of size $2K \times 2K$, while $\boldsymbol{J}_f$ is formed from the even indexed rows of the same matrix. Let $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ represent the linear and nonlinear signal matrices, respectively. Hence, they can be written as follows:

$$\begin{aligned} \boldsymbol{S}_1 &= \boldsymbol{J}_e \boldsymbol{\mathcal{Q}}_{new}^{-1} \tilde{\boldsymbol{V}}_s \approx \boldsymbol{J}_e (\boldsymbol{I} - \boldsymbol{\mathcal{E}}) \boldsymbol{\mathcal{Q}}_{old}^{-1} \tilde{\boldsymbol{V}}_s \\ \boldsymbol{S}_2 &= \boldsymbol{J}_f \boldsymbol{\mathcal{Q}}_{new}^{-1} \tilde{\boldsymbol{V}}_s \approx \boldsymbol{J}_f (\boldsymbol{I} - \boldsymbol{\mathcal{E}}) \boldsymbol{\mathcal{Q}}_{old}^{-1} \tilde{\boldsymbol{V}}_s \end{aligned} \tag{14}$$

Hence, J_4 can be expressed element-wise as follows:

$$\begin{aligned} J_4 &= \sum_{i,j} |F(\boldsymbol{S}_1(i,j)) - \boldsymbol{S}_2(i,j)|^2 \\ &= \sum_{i,j} |\boldsymbol{S}_1^2(i,j) - \boldsymbol{S}_2(i,j)|^2 \\ &\approx \sum_{i,j} |(\boldsymbol{J}_e (\boldsymbol{I} - \boldsymbol{\mathcal{E}}) \bar{\boldsymbol{S}}_{old}(i,j))^2 - \boldsymbol{J}_f (\boldsymbol{I} - \boldsymbol{\mathcal{E}}) \bar{\boldsymbol{S}}_{old}(i,j)|^2 \end{aligned} \tag{15}$$

The tightness of the proposed approximation is because the subspace method in [13] (we used for initialization) provides already a good channel estimate that is further refined by our proposed method using (typically) just few iterations.

where $\bar{\boldsymbol{S}}_{old} = \boldsymbol{\mathcal{Q}}_{old}^{-1}\tilde{\boldsymbol{V}}_s$. After straightforward derivations, the first order approximation of J_4, can be written in matrix form as:

$$J_4 \approx J_{4,old} + 2Re\left\{Tr((\bar{\boldsymbol{S}}_{old}\boldsymbol{F}\boldsymbol{J}_e + \bar{\boldsymbol{S}}_{old}\boldsymbol{E}\boldsymbol{J}_f)\boldsymbol{\mathcal{E}})\right\} \quad (16)$$

where $J_{4,old} = |(\boldsymbol{J}_e\bar{\boldsymbol{S}}_{old}(i,j))^2 - \boldsymbol{J}_f\bar{\boldsymbol{S}}_{old}(i,j)|^2$ and

$$\boldsymbol{E}(j,i) = [(\boldsymbol{J}_e\bar{\boldsymbol{S}}_{old}(i,j))^2 - \boldsymbol{J}_f\bar{\boldsymbol{S}}_{old}(i,j)]^*$$

$$\boldsymbol{F}(j,i) = -2\boldsymbol{E}(j,i)(\boldsymbol{J}_e\bar{\boldsymbol{S}}_{old}(i,j)).$$

Finally, the first order expansion of $J(\boldsymbol{\mathcal{E}})$ is expressed as follows:

$$\begin{aligned} J(\boldsymbol{\mathcal{E}}) \approx &\| \boldsymbol{A} \|_{\mathrm{F}}^2 + \| \boldsymbol{B} \|_{\mathrm{F}}^2 + \| \boldsymbol{C} \|_{\mathrm{F}}^2 + \| \boldsymbol{D} \|_{\mathrm{F}}^2 \\ &+J_{4,old} + 2Re\left\{Tr\left((\boldsymbol{M}_a + \boldsymbol{M}_b + \boldsymbol{M}_c + \boldsymbol{M}_d)\boldsymbol{\mathcal{E}}\right)\right\}, \end{aligned} \quad (17)$$

where $\boldsymbol{M}_a = \tilde{\boldsymbol{J}}_a\boldsymbol{A}^H\boldsymbol{A}_1 - \tilde{\boldsymbol{J}}_b\boldsymbol{A}^H\boldsymbol{A}_2$, $\boldsymbol{M}_b = \boldsymbol{B}_2\boldsymbol{B}^H\boldsymbol{J}_d - \boldsymbol{B}_1\boldsymbol{B}^H\boldsymbol{J}_c$, $\boldsymbol{M}_c = \tilde{\boldsymbol{J}}_{rw}\boldsymbol{C}^H\boldsymbol{C}_1 + \tilde{\boldsymbol{J}}_{cl}\boldsymbol{D}^H\boldsymbol{D}_1$ and $\boldsymbol{M}_d = \boldsymbol{S}_{old}\boldsymbol{F}\boldsymbol{J}_e + \boldsymbol{S}_{old}\boldsymbol{E}\boldsymbol{J}_f$. Here, to ensure that the cost function decreases, $\boldsymbol{\mathcal{E}}$ is selected such that it moves in a direction opposite to that of the gradient, according to:

$$\boldsymbol{\mathcal{E}} = -\rho\left(\boldsymbol{M}_a + \boldsymbol{M}_b + \boldsymbol{M}_c + \boldsymbol{M}_d\right)^H, \quad (18)$$

where ρ represents a small positive constant.

The proposed bilinear method for nonlinear systems uses the structured channel subspace method in [13] for initialization and few iterations are used for channel and signal matrix refinement in order to improve estimation quality. The step-by-step procedure involved in the proposed bilinear method is described in Algorithm 1.

Algorithm 1: Summary of the proposed Bilinear method.

initialization;
$\boldsymbol{\mathcal{Q}}_{old}$, $\bar{\boldsymbol{S}}_{old}$
while *Number of iterations* $\leq N_{max}$ **do**
  $\boldsymbol{\mathcal{Q}}_{old} = \boldsymbol{\mathcal{Q}}_{new}$
  $\boldsymbol{M}_a = \tilde{\boldsymbol{J}}_a\boldsymbol{A}^H\boldsymbol{A}_1 - \tilde{\boldsymbol{J}}_b\boldsymbol{A}^H\boldsymbol{A}_2$
  $\boldsymbol{M}_b = \boldsymbol{B}_2\boldsymbol{B}^H\boldsymbol{J}_d - \boldsymbol{B}_1\boldsymbol{B}^H\boldsymbol{J}_c$
  $\boldsymbol{M}_c = \tilde{\boldsymbol{J}}_{rw}\boldsymbol{C}^H\boldsymbol{C}_1 + \tilde{\boldsymbol{J}}_{cl}\boldsymbol{D}^H\boldsymbol{D}_1$
  $\boldsymbol{M}_d = \bar{\boldsymbol{S}}_{old}\boldsymbol{F}\boldsymbol{J}_e + \bar{\boldsymbol{S}}_{old}\boldsymbol{E}\boldsymbol{J}_f$.
  $\boldsymbol{\mathcal{E}} = -\rho\left(\boldsymbol{M}_a + \boldsymbol{M}_b + \boldsymbol{M}_c + \boldsymbol{M}_d\right)^H$
  $\boldsymbol{\mathcal{Q}}_{new} = \boldsymbol{\mathcal{Q}}_{old}(\boldsymbol{I} + \boldsymbol{\mathcal{E}})$
  $\boldsymbol{\mathcal{Q}}_{new}^{-1} \approx (\boldsymbol{I} - \boldsymbol{\mathcal{E}})\boldsymbol{\mathcal{Q}}_{old}^{-1}$
  $\bar{\boldsymbol{S}}_{new} \approx (\boldsymbol{I} - \boldsymbol{\mathcal{E}})\bar{\boldsymbol{S}}_{old}$
  $\boldsymbol{\mathcal{H}}_{N_w} = \boldsymbol{U}_s\boldsymbol{\mathcal{Q}}_{\boldsymbol{new}}$
end

IV. Computational Complexity

The proposed method's computation complexity is compared to the complexity of the subspace (SS) method [11] and the structured channel subspace (SCS) method [12]. The proposed bilinear method has the heaviest computational cost with a total complexity of $O((N_rN_w)(N-N_w)^2) + O((K(N-N_w))^2) + O((N_rN_wK)^2)$ due to the data matrix SVD and the fact that it is initialized with $\boldsymbol{\mathcal{Q}}$ from the SCS method. The next in terms of computational cost is the SCS method with a complexity of $O((N_rN_w)^2(N-N_w)) + O((N_rN_wK)^2)$. Finally, the SS method has the least computational cost with $O((N_rN_w)^2(N-N_w)) + O((N_r(L+1))^2)$. Here, O represents the order of complexity.

V. Simulation Results

In this section, the normalized mean squared error (NMSE) is used as performance metric for the developed nonlinear bilinear method (BSCS), the SS method and the SCS method are investigated via simulations experiments. In fact, the SCS and SS methods provide partial channel estimate of the $N_r \times 2$ MIMO system in (2). More precisely, the estimated channel taps $\hat{\boldsymbol{H}}(k)$, $k = 0, \cdots, L$ satisfy

$$\hat{\boldsymbol{H}}(k) \approx \boldsymbol{H}(k)\mathbf{D}$$

where $\mathbf{D}$ is a 2×2 unknown matrix (see details in [18]). In comparison, our bilinear method fully estimates the SIMO nonlinear channel up to a scalar factor (which is the inherent ambiguity of the blind processing methods for SIMO systems) [19]- [20]. In the sequel, for comparison purpose, we remove these ambiguities from the channel estimates, using a least squares fitting criterion with the exact channel matrix, before their use in criterion (19).

The effectiveness of the proposed channel estimator is firstly verified by the normalized mean squares error (NMSE) of channel estimation, which is given in dB as:

$$\text{NMSE} = 20\log_{10}\left(\frac{\sqrt{\frac{1}{N_{mc}}\sum_{i=1}^{N_{mc}} \| \hat{\boldsymbol{H}}_i - \boldsymbol{H} \|_{\mathrm{F}}^2}}{\| \boldsymbol{H} \|_{\mathrm{F}}^2}\right), \quad (19)$$

where $N_{mc} = 100$ is the number of Monte Carlo runs, $\boldsymbol{H} = [\boldsymbol{H}(0) \cdots \boldsymbol{H}(L)]$ represents the true channel employed in the simulations, and $\hat{\boldsymbol{H}}_i$ represents the estimated channel at the i^{th} run.

The input stream $s(n)$ is drawn from a QAM 16 constellation, the nonlinear part is obtained by passing the input stream through a nonlinear function $\tilde{s}(n) = F(s(n)) = s^2(n)$ and the additive noise is white Gaussian with zero mean which is generated for each Monte Carlo run. Throughout the simulations, the input signal length of $N = 100$ is considered, $N_r = 4$ receiver antennas, a window size of $N_w = 5$, and the channel is generated randomly with an order of $L_L = L_{NL} = L = 3$ in all experiments except when otherwise specified. It is crucial to note that the nonlinear bilinear technique is iterative, with the step size set to $\rho = 1 \times 10^{-6}$ in all experiments.Here, in this work our aim was to test the superiority of our algorithm in fair comparison with the rest of the algorithms.

The proposed bilinear method significantly outperforms the SS method, the SCS method, and the Bayesian method as shown in Fig. 2, with a gain of over 2dB in our algorithm's favor.

In the next experiment, the performance of the algorithms, including the previous linear-based algorithm (LBSCS) [13], and the Bayesian method is investigated in terms of the symbol error rate (SER) a more discerning and a more reliable metric

This method estimates the $N_r \times 2$ MIMO channel in (3) up to a 2×2 unknown matrix (see [13] for details).

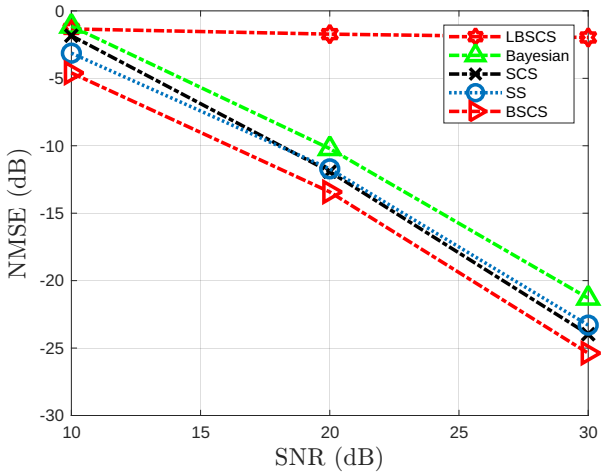


Fig. 2: The NMSE performance versus SNR.

for this type of steady-state analysis. The results of this comparison are all shown in Fig. 3, and clearly demonstrate the strong consistency between the performance of our proposed algorithm and other previously published and related algorithms. As shown in Fig 3, the proposed bilinear method outperforms all the other algorithms, including our algorithm in [13] with the worst performance, and gives it an edge in steady-state performance over the other algorithms. For a 6×10^{-3} SER, a gain of 2 dB in favor of the proposed algorithm against the SCS method is achieved. Consistency in performance of the proposed algorithm is maintained in all the experiments performed in this study even the Bayesian method was outperformed by the proposed algorithm. The proposed method requires the longest computation time among the evaluated techniques, primarily due to the joint estimation of signal and channel subspaces and the iterative refinement process. The Bayesian method follows, exhibiting moderate computational time. In contrast, both the SCS and SS methods are computationally efficient, with execution times significantly shorter than those of the proposed and Bayesian approaches.

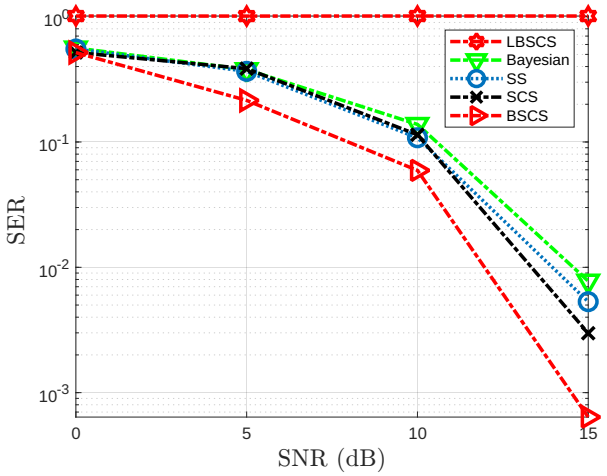


Fig. 3: The SER performance versus the SNR.

VI. Conclusion

In this paper, an iterative and efficient bilinear structure subspace method for blind channel identification is developed which exploits the column and row subspaces. In addition, the algorithm was designed to take care of any nonlinearity in the system. This was achieved by introducing a new term in the minimization procedure as to tackle the system's nonliearity. The algorithm performs better than the considered methods in terms of NMSE and SER. To examine how well the proposed algorithm performed, several scenarios were examined.

Overall, the proposed method outperformed other methods at the expense of a moderate computational load which is mainly due to the involvement of the iterative procedure.

References

[1] V. J. Mathews, "Adaptive polynomial filters," *Jul*, vol. 1, pp. 10–26, 1991.

[2] J. Schoukens and L. Ljung, "Nonlinear system identification: A user-oriented road map," *IEEE Control Syst. Mag.*, vol. 39, no. 6, pp. 28–99, 2019.

[3] A. Amari, O.A. Dobre, R. Venkatesan, O.S. Kumar, P. Ciblat, Y. Jaouen, "A survey on fiber nonlinearity compensation for 400 Gb/s and beyond optical communication systems," *IEEE Commun. Surv. Tutor.*, vol. 19, no. 4, pp. 3097–3113, 2017.

[4] T. Ogunfunmi, *Adaptive nonlinear system identification: The Volterra and Wiener model approaches*. Springer Science & Business Media, 2007.

[5] C. Cheng, Z. Peng, W. Zhang, and G. Meng, "Volterra-series-based nonlinear system modeling and its engineering applications: A state-of-the-art review," *Mechanical Systems and Signal Processing*, vol. 87, pp. 340–364, 2017.

[6] T. Ogunfunmi and S.-L. Chang, "Second-order adaptive volterra system identification based on discrete nonlinear wiener model," *IEE Proceedings-Vision, Image and Signal Processing*, vol. 148, no. 1, pp. 21–29, 2001.

[7] Y. V. Zakharov, G. P. White, and J. Liu, "Low-complexity RLS algorithms using dichotomous coordinate descent iterations," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3150–3161, 2008.

[8] N. Kalouptsidis and P. Koukoulas, "Blind identification of volterra-hammerstein systems," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2777–2787, 2005.

[9] J. C. Gómez and E. Baeyens, "Subspace-based identification algorithms for hammerstein and wiener models," *European Journal of Control*, vol. 11, no. 2, pp. 127–136, 2005.

[10] W. Qiu and Y. Hua, "Performance analysis of the subspace method for blind channel identification," *Signal Processing*, vol. 50, no. 1-2, pp. 71–81, 1996.

[11] C. Yu, L. Xie, M. Verhaegen, and J. Chen, "Blind identification of MIMO systems," in *Blind Identification of Structured Dynamic Systems*, pp. 167–198, Springer, 2022.

[12] Q. Mayyala, K. Abed-Meraim, and A. Zerguine, "Structure-based subspace method for multichannel blind system identification," *IEEE Signal Process. Lett.*, vol. 24, no. 8, pp. 1183–1187, 2017.

[13] A. Lawal, Q. Mayyala, K. Abed-Meraim, N. Iqbal, and A. Zerguine, "Toeplitz structured subspace for multi-channel blind identification methods," *Signal Processing*, vol. 188, p. 108152, 2021.

[14] O. Rekik, K. Abed-Meraim, M. Nait-Meziane, A. Mokraoui, and N. L. Trung, "Maximum likelihood based identification for nonlinear multichannel communications systems," *Signal Processing*, vol. 189, p. 108297, 2021.

[15] L. Tao, H. Tan, C. Fang, and N. Chi, "Volterra series based blind equalization for nonlinear distortions in short reach optical cap system," *Optics Communications*, vol. 381, pp. 240–243, 2016.

2716-912X © 2025 Ecole Nationale Polytechnique

[16] A. H. Abdelhafiz, O. Hammi, A. Zerguine, A. T. Al-Awami, and F. M. Ghannouchi, "A pso based memory polynomial predistorter with embedded dimension estimation," *IEEE Transactions on Broadcasting*, vol. 59, no. 4, pp. 665–673, 2013.

[17] A. E. Abdelrahman, O. Hammi, A. K. Kwan, A. Zerguine, and F. M. Ghannouchi, "A novel weighted memory polynomial for behavioral modeling and digital predistortion of nonlinear wireless transmitters," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 3, pp. 1745–1753, 2015.

[18] K. Abed-Meraim, P. Loubaton, and E. Moulines, "A subspace algorithm for certain blind identification problems," *IEEE Trans. Inf. Theory*, vol. 43, no. 3, pp. 499–511, 1997.

[19] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Transactions on Signal Processing*, vol. 43, no. 12, pp. 2982–2993, 1995.

[20] W. Qiu and Y. Hua, "Performance analysis of the subspace method for blind channel identification," *Signal Processing*, vol. 50, no. 1, pp. 71–81, 1996.

ABDULMAJID LAWAL earned a B.Eng. in electrical engineering from Bayero University in Kano in 2012, a Master's in electrical engineering from King Fahd University of Petroleum and Minerals (KFUPM) in Saudi Arabia in 2016, and a PhD in electrical engineering from the same institution in 2021. He is presently a postdoctoral fellow student at the KFUPM, Saudi Arabia, center for communication systems and sensing. He focuses on issues with blind source separation, blind system/channel identification, sparse signal representations, and artificial intelligence. Statistical signal processing, modeling, estimation, and classification of signals, blind system identification, blind source separation, statistical performance analysis, and sparse representations are some of his areas of interest in study.

KARIM ABED-MERAIM was born in 1967. He received the State Engineering Degree from the École Polytechnique, Palaiseau, France, in 1990, the State Engineering Degree from the École Nationale Supérieure des Télécommunications (ENST), Paris, France, in 1992, the M.Sc. degree from Paris XI University, Orsay, France, in 1992, and the Ph.D. degree in the field of signal processing and communications from ENST, in 1995. From 1995 to 1998, he was a Research Staff with the Electrical Engineering Department, The University of Melbourne, where he worked on several research project related to "Blind System Identification for Wireless Communications," "Blind Source Separation," and "Array Processing for Communications." From 1998 to 2012, he has been an Assistant Professor and then as an Associate Professor with the Signal and Image Processing Department, Télécom ParisTech. His research interests include signal processing for communications, adaptive filtering and tracking, array processing, and statistical performance analysis. In September 2012, he joined the University d'Orléans (PRISME Laboratory), as a Full Professor. He is the author of over 550 scientific publications, including book chapters, international journal and conference papers, and patents. Dr. Abed-Meraim is a past member of the IEEE SAM-TC and past Associate Editor (and Senior Area Editor) of the IEEE-TSP.

AZZEDINE ZERGUINE received the B.Sc. degree in electrical engineering from Case Western Reserve University, Cleveland, OH, USA, in 1981, the M.Sc. degree in electrical engineering from the King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia, in 1990, and the Ph.D. degree in electrical engineering from Loughborough University, Loughborough, U.K., in 1996. He is currently a Professor with the Department of Electrical Engineering, KFUPM, working in the areas of signal processing and communications. His current research interests include signal processing for communications, adaptive filtering, neural networks, the applications of artificial intelligence, blind source separation, and blind equalization. He was a recipient of the three Best Teaching Awards at KFUPM, in 2000, 2005, and 2011, and the Best Research Award at KFUPM, in 2017. Actually, he serves as an Associate Editor for the *Signal Processing* Journal and the IEEE COMMUNICATIONS LETTERS.

ALI HUSSEIN MUQAIBEL received the Ph.D. degree from the Virginia Polytechnic Institute and State University, Blacksburg, VA, USA, in 2003. He is currently a Professor with the Electrical Engineering Department, King Fahd University of Petroleum and Minerals (KFUPM). He is also the Director of the Interdisciplinary Research Center for Communication Systems and Sensing (IRC-CSS). During his study with Virginia Tech, he was with the Time Domain and RF Measurements Laboratory and the Mobile and Portable Radio Research Group. He was a Visiting Associate Professor with the Center of Advanced Communications, Villanova University, Villanova, PA, USA, in 2013, a Visiting Professor with Georgia Institute of Technology, in 2015, and a Visiting Scholar with the King Abdullah University for Science and Technology (KAUST), Thuwal, Saudi Arabia, in 2018 and 2019. He has authored three book chapters and more than 150 articles. His research interests include communications and sensing applications, including the direction of arrival estimation, through-wall-imaging, localization, channel characterization, and ultra-wideband signal processing. He was a recipient of many awards for excellence in teaching, advising, and instructional technology.

2716-912X © 2025 Ecole Nationale Polytechnique

Implementation of Finite Control Set Model Predictive Control (FCS-MPC) for Five-Phase Induction Motors

Abdelfattah Hoggui, Ali Benachour, Mohamed Chafaa Madaoui, and Mohand Oulhadj Mahmoudi

Abstract—This paper investigates the study and practical implementation of Finite Control Set Model Predictive Control (FCS-MPC) for a five-phase induction motor powered by a two-level, five-phase voltage source inverter. Multiphase systems present notable benefits over conventional three-phase configurations, including improved fault tolerance, enhanced torque production, and minimized harmonic distortion, making them well-suited for demanding applications. FCS-MPC operates by directly predicting and selecting the optimal control action at each sampling instant, eliminating the necessity of a modulation stage and offering fast dynamic response. The proposed control approach is designed and validated through experimental testing, assessing key performance indicators such as speed tracking, torque ripple, and current distortion under both transient and steady-state conditions. The experimental results demonstrate the effectiveness of FCS-MPC in ensuring precise and efficient control across various operating scenarios, making it a viable alternative for multiphase motor drive applications.

***Keywords*—five-phase induction motor, two-level five-phase voltage source inverter, finite control set model predictive control (FCS-MPC).**

I. Introduction

In recent years, multiphase motor systems have gained increasing attention due to their potential for enhanced performance, reliability, and efficiency in various industrial applications. Among these, five-phase induction motors have emerged as a viable alternative to conventional three-phase machines, offering key advantages such as improved fault tolerance, reduced torque ripple, and lower harmonic distortion. These characteristics make five-phase systems particularly suitable for demanding applications, including renewable energy generation, electric transportation, and high-precision industrial drives [1–6].

The unique benefits of five-phase induction motors have driven researchers to extend and adapt control strategies originally developed for three-phase systems to these advanced multiphase configurations [7–20]. One such technique is Finite Control Set Model Predictive Control (FCS-MPC), which has gained prominence as an alternative to conventional control methods due to its ability to directly predict and optimize switching states at each sampling interval [21–25].

FCS-MPC is based on a predictive control that utilizes the mathematical model of the system to estimate future states and minimize a predefined cost function. This approach allows direct control of stator currents, torque, and flux without requiring coordinate transformations or pulse-width modulation (PWM). By continuously evaluating the impact of different switching states, FCS-MPC ensures fast dynamic response, improved transient performance, and precise control of electromagnetic variables, making it an attractive solution for multiphase motor drives.

This paper is structured as follows: Section II presents the modeling of the five-phase induction motor, followed by the mathematical formulation of the two-level five-phase voltage source inverter. Section III introduces the principles of FCS-MPC and its application to five-phase motor control. In Section IV, experimental results are analyzed to evaluate the performance of FCS-MPC under various operating conditions. Finally, Section V concludes the paper with key findings and potential research directions.

II. Modeling of the Five-Phase Induction Machine

The mathematical representation of five-phase induction machines is derived from their distinct winding configuration and electromagnetic characteristics. To simplify the modeling process, the following assumptions are considered in this work [13]:

- The magnetic circuit operates within a linear region, neglecting saturation effects and assuming constant permeability.
- The magnetomotive force (MMF) distribution and the resulting air-gap flux are considered sinusoidal.
- Hysteresis, eddy current losses, and skin effect are disregarded, focusing the model solely on electromagnetic interactions.

Manuscript received March 6, 2025; revised June 15, 2025.

A. Hoggui, M.C. Madaoui and M.O. Mahmoudi are with the Laboratoire de Recherche en Electrotechnique (LRE), Ecole Nationale Poly-technique (ENP), B.P. 182 El Harrach, Algiers 16200, Algeria.e-mail: (abdelfettah.hoggui@g.enp.edu.dz).

A. Benachour is with the Electrical Engineering Department, Ecole Nationale Supérieure des Technologies Avancées, Algiers, ALGERIA. e-mail: (ali.benachour@ensta.edu.dz).

Digital Object Identifier (DOI): 10.53907/enpesj.v5i1.325

2716-912X © 2025 Ecole Nationale Polytechnique

- A uniform air-gap thickness is assumed to facilitate the magnetic circuit analysis.

Five-phase induction motors have five stator windings that are symmetrically arranged with a spatial displacement of $\frac{2\pi}{5}$. The corresponding arrangement is shown in Fig. 1

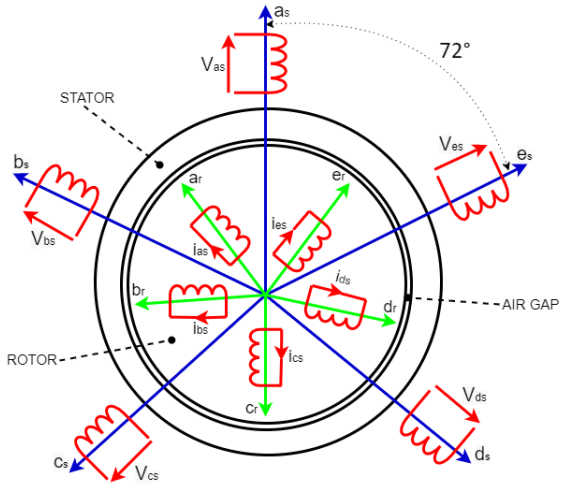


Fig. 1: Winding arrangement of a five-phase induction machine.

The equations describing the voltage and flux linkage for the stator and rotor phases are presented as follows:

$$\begin{cases} \mathbf{V}_s = \mathbf{R}_s\mathbf{I}_s + \dfrac{d}{dt}\mathbf{\Phi}_s, \\ \mathbf{\Phi}_s = \mathbf{L}_s\mathbf{I}_s + \mathbf{L}_{sr}\mathbf{I}_r, \end{cases} \tag{1}$$

$$\begin{cases} \mathbf{V}_r = \mathbf{R}_r\mathbf{I}_r + \dfrac{d}{dt}\mathbf{\Phi}_r - j\omega_r\mathbf{\Phi}_r, \\ \mathbf{\Phi}_r = \mathbf{L}_r\mathbf{I}_r + \mathbf{L}_{rs}\mathbf{I}_s, \end{cases} \tag{2}$$

where:

- $\mathbf{V}_s$ and $\mathbf{V}_r$ are the stator and rotor voltages,
- $\mathbf{I}_s$ and $\mathbf{I}_r$ are the stator and rotor currents,
- $\mathbf{\Phi}_s$ and $\mathbf{\Phi}_r$ represent the stator and rotor flux linkages,
- $\mathbf{R}_s$, $\mathbf{R}_r$, $\mathbf{L}_s$, $\mathbf{L}_r$, $\mathbf{L}_{sr}$, and $\mathbf{L}_{rs}$ are the stator and rotor resistances, self-inductances, and mutual inductances,
- ω_r is the rotor speed.

While these equations provide an accurate representation of the machine in the natural $(a-b-c-d-e)$ coordinate system, the presence of time-varying coefficients complicates their use in control applications. To facilitate implementation, the stator and rotor equations are transformed into the $(\alpha-\beta-x-y-z)$ reference frame using Park's transformation. This transformation simplifies the control structure by decoupling the torque-producing $(\alpha-\beta)$ components from the $(x-y-z)$ components, which do not contribute to electromagnetic torque production. The resulting equations in the transformed domain are given by:

$$\begin{cases} V_{\alpha s} = R_s I_{\alpha s} + \dfrac{d\Phi_{\alpha s}}{dt} - \omega\Phi_{\beta s}, \\ V_{\beta s} = R_s I_{\beta s} + \dfrac{d\Phi_{\beta s}}{dt} + \omega\Phi_{\alpha s}, \end{cases} \tag{3}$$

The electromagnetic torque generated by the five-phase induction machine is expressed as:

$$T_e = \frac{5}{2}p\left(\Phi_{\alpha s}I_{\beta s} - \Phi_{\beta s}I_{\alpha s}\right), \tag{4}$$

where:

- T_e represents the electromagnetic torque,
- p denotes the number of pole pairs,
- $\Phi_{\alpha s}$ and $\Phi_{\beta s}$ are the α-β axis flux components,
- $I_{\alpha s}$ and $I_{\beta s}$ correspond to the α-β axis current components.

This model serves as a foundation for implementing advanced control strategies suitable for five-phase machines. By focusing on the torque-generating α-β components, the effects of the non-torque-producing $(x-y-z)$ components are disregarded. Although these additional components may introduce extra losses in the machine [9, 14–17], studies have shown that their impact is negligible in practical applications. Therefore, for the sake of simplicity, these losses are not considered in this work. The developed model provides a structured framework for implementing predictive control techniques, particularly FCS-MPC, which relies on the discrete-time representation of system dynamics to optimize control performance.

III. Two-Level Five-Phase Inverter

Fig. 2 presents the topology of a two-level five-phase inverter, which serves as the power source for the stator windings of the five-phase induction motor considered in this paper.

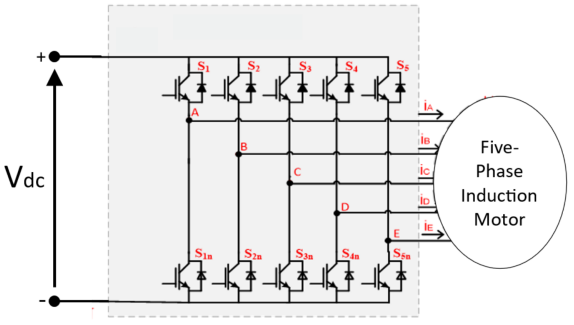


Fig. 2: Topology of the two-level five-phase inverter.

The inverter is composed of five legs, each incorporating two insulated-gate bipolar transistors (IGBTs) operating in a complementary switching scheme. This configuration results in 32 distinct switching states. Each switching state corresponds to a specific space voltage vector, forming a set that includes two zero vectors and thirty active vectors with different magnitudes. The spatial distribution of these voltage vectors in the $\alpha-\beta$ plane is depicted in Fig. 3.

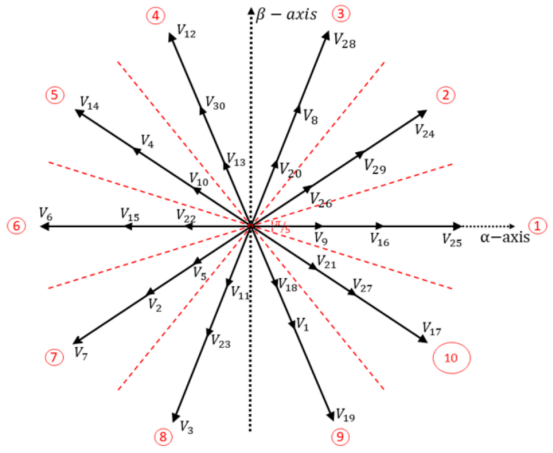


Fig. 3: Space voltage vector of the two-level five-phase inverter in the $\alpha - \beta$ plane.

IV. Finite Control Set Model Predictive Control for a Five-Phase Induction Motor

Fig. 4 presents the schematic representation of the FCS-MPC applied to a five-phase induction motor. It outlines the main components of the control loop, including the predictive model, cost function evaluation, and selection of the optimal switching states.

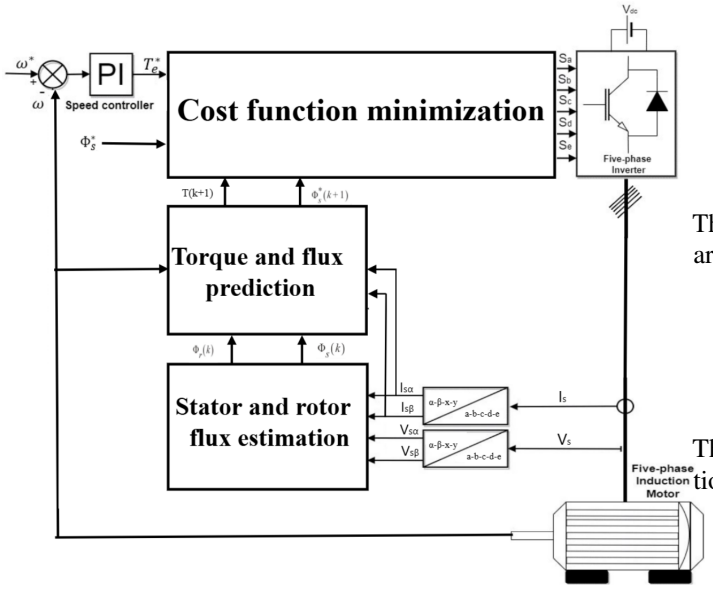


Fig. 4: Schematic diagram of the FCS-MPC control strategy.

For five-phase induction motor control, FCS-MPC employs a predictive approach based on the motor's mathematical model in the stationary reference frame. The controller forecasts future states, such as stator currents and both stator and rotor flux linkages, by considering different inverter switching states. The predictive process relies on discretizing the motor's dynamic equations and using these discrete models to estimate system behavior over a short prediction horizon. At each sampling step, the controller evaluates all possible voltage vectors (corresponding to the five-phase inverter's switching states) and calculates a cost function that includes terms for tracking errors [18].

The continuous-time model defined in system (3), expressed in a reference frame rotating at an arbitrary speed ω, is transformed into the stationary (α-β-x-y-o) reference frame. For implementation in a digital controller, the system is then discretized using a sampling period T_s. To simplify computations, only the torque-producing α-β components are considered in this work. The discrete-time expressions for the stator flux linkages are formulated as follows:

$$\begin{aligned}\Phi_{s\alpha}[k+1] &= \Phi_{s\alpha}[k] + T_s\left(V_{s\alpha}[k] - R_s I_{s\alpha}[k]\right)\\ \Phi_{s\beta}[k+1] &= \Phi_{s\beta}[k] + T_s\left(V_{s\beta}[k] - R_s I_{s\beta}[k]\right)\end{aligned} \tag{5}$$

Similarly, the rotor flux linkage equations in discrete form are given by:

$$\begin{aligned}\Phi_{r\alpha}[k+1] &= \left(1 - \frac{R_r T_s}{L_r}\right)\Phi_{r\alpha}[k] + T_s\omega_m\Phi_{r\beta}[k]\\ &\quad + \frac{R_r T_s L_m}{L_r} I_{s\alpha}[k]\\ \Phi_{r\beta}[k+1] &= \left(1 - \frac{R_r T_s}{L_r}\right)\Phi_{r\beta}[k] - T_s\omega_m\Phi_{r\alpha}[k]\\ &\quad + \frac{R_r T_s L_m}{L_r} I_{s\beta}[k]\end{aligned} \tag{6}$$

The rotor currents can be expressed in terms of the rotor flux linkages as:

$$\begin{aligned}I_{r\alpha}[k+1] &= \frac{\Phi_{r\alpha}[k+1] - L_m I_{s\alpha}[k+1]}{L_r}\\ I_{r\beta}[k+1] &= \frac{\Phi_{r\beta}[k+1] - L_m I_{s\beta}[k+1]}{L_r}\end{aligned} \tag{7}$$

The stator flux linkages in relation to the stator and rotor currents are expressed as:

$$\begin{aligned}\Phi_{s\alpha}[k+1] &= L_s I_{s\alpha}[k+1] + L_m I_{r\alpha}[k+1]\\ \Phi_{s\beta}[k+1] &= L_s I_{s\beta}[k+1] + L_m I_{r\beta}[k+1]\end{aligned} \tag{8}$$

The stator currents are then updated using the following relations:

$$X_\alpha[k] = \frac{1}{R_\sigma}\left(\frac{k_r}{\tau_r}\Phi_{r\alpha}[k] + k_r\omega_m\Phi_{r\beta}[k]\right)$$

$$X_\beta[k] = \frac{1}{R_\sigma}\left(\frac{k_r}{\tau_r}\Phi_{r\beta}[k] - k_r\omega_m\Phi_{r\alpha}[k]\right)$$

$$\begin{aligned}I_{s\alpha}[k+1] &= \left(1 + \frac{T_s}{\tau_\sigma}\right)I_{s\alpha}[k] + \frac{T_s}{\tau_\sigma}\left(X_\alpha[k] + V_{s\alpha}[k]\right)\\ I_{s\beta}[k+1] &= \left(1 + \frac{T_s}{\tau_\sigma}\right)I_{s\beta}[k] + \frac{T_s}{\tau_\sigma}\left(X_\beta[k] + V_{s\beta}[k]\right)\end{aligned} \tag{9}$$

where:

- $\sigma = 1 - \frac{L_m^2}{L_r L_s}$

- $\tau_\sigma = \frac{\sigma L_s}{R_\sigma}$
- $k_r = \frac{L_m}{L_r}$
- $\tau_r = \frac{L_r}{R_r}$
- $R_\sigma = R_s + k_r^2 R_r$
- ω_m is the mechanical angular velocity.

The predicted electromagnetic torque at the next time step is computed using:

$$T_e[k+1] = \frac{5}{2} p \left(\Phi_{s\alpha}[k+1] I_{s\beta}[k+1] - \Phi_{s\beta}[k+1] I_{s\alpha}[k+1]\right) \tag{10}$$

To achieve precise control of the motor, the cost function in this study aims to minimize errors in torque and stator flux, formulated as:

$$J = \lambda_T \left|T_e[k+1] - T_{\text{ref}}\right| + \lambda_\Phi \left|\Phi_s[k+1] - \Phi_{s,\text{ref}}\right| \tag{11}$$

where:

- λ_T and λ_Φ are weighting factors for torque and flux errors, respectively.
- $T_e[k+1]$ is the predicted electromagnetic torque at the next time step.
- T_{ref} is the reference torque.
- $\Phi_s[k+1]$ is the predicted stator flux magnitude.
- $\Phi_{s,\text{ref}}$ is the reference stator flux magnitude.

The control algorithm evaluates all 32 possible switching states and selects the one that minimizes the cost function. This optimization ensures minimal tracking error and efficient control.

V. EXPERIMENTAL RESULTS

Fig. 5 illustrates the experimental test bench utilized for implementing the Finite Control Set Model Predictive Control (FCS-MPC) strategy on a five-phase induction motor. The setup consists of a transformer supplying a rectifier, which delivers a DC voltage. This DC voltage powers two two-level three-phase inverters configured to supply the five-phase induction motor.

The five-phase induction motor is mechanically coupled to a DC generator, which serves as the load. The load level can be adjusted using a variable resistor connected to the DC generator, enabling precise control of the operating conditions.

The FCS-MPC control algorithm is implemented on a dSPACE 1104 control board, which interfaces with the system for real-time control and data acquisition. A current measurement circuit is integrated into the setup. Additionally, an incremental encoder

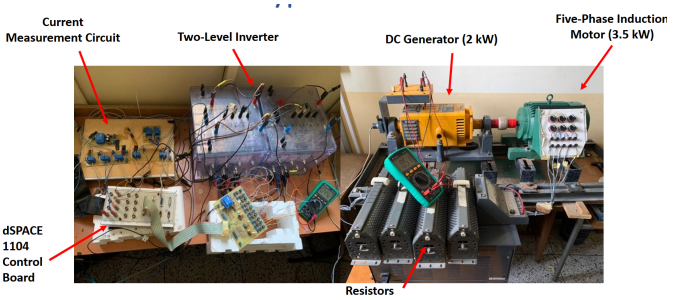


Fig. 5: Experimental test bench used for FCS-MPC implementation.

is used to measure the motor's rotational speed, providing essential data for the closed-loop control system. The parameters of the five-phase induction motor used in this setup are presented in a table in the appendix.

To evaluate the performance of the FCS-MPC strategy applied to the five-phase induction motor, two experimental tests were conducted under varying speed conditions while maintaining a constant load. In the first test, the reference speed was incremented from 50 rad/s to 200 rad/s in steps of 50 rad/s and subsequently reduced from 200 rad/s to 150 rad/s under the same load conditions. In the second test, the reference speed was varied to assess the system's performance during speed reversals: it started at 50 rad/s, increased to 100 rad/s, and then decreased to –150 rad/s in steps of 50 rad/s.

The results, presented in Fig. 6 and Fig. 7, provide insight into key electromagnetic parameters, including the reference and actual speed, the electromagnetic torque and its reference, as well as the flux magnitude and its reference.

Fig. 6(a) and Fig. 7(a) show that the rotor speed closely follows its reference, with minimal overshoot during acceleration and deceleration. The motor smoothly transitions between speed levels, highlighting the capability of the FCS-MPC strategy in ensuring fast and precise speed adjustments.

Fig. 6(b) and Fig. 7(b) shows that the electromagnetic torque closely follows its reference, with minimal ripple in both forward and reverse rotational directions. In steady-state operation, the torque ripple varies depending on the reference speed.

Fig. 6(c) and Fig. 7(c) illustrate that the path of the stator flux forms an almost perfect circle, demonstrating that it has a consistent magnitude with minimal ripple in both forward and reverse rotational directions, as confirmed in Fig. 6(d) and Fig. 7(d).

Fig. 8 illustrates the electrical performance of the FCS-MPC strategy at constant speed. It presents the inverter's output voltage along with its harmonic spectrum, as well as the stator current and its corresponding harmonic spectrum.

The results show that the total harmonic distortion (THD) of the inverter's output voltage is notably high due to the variable switching frequency of the FCS-MPC strategy. On the other hand, the stator current THD is significantly lower, benefiting from the filtering effect of high-frequency harmonics, a characteristic effect of the inductive properties of the induction machine. Additionally, during speed variations, peak current values reach 5A.

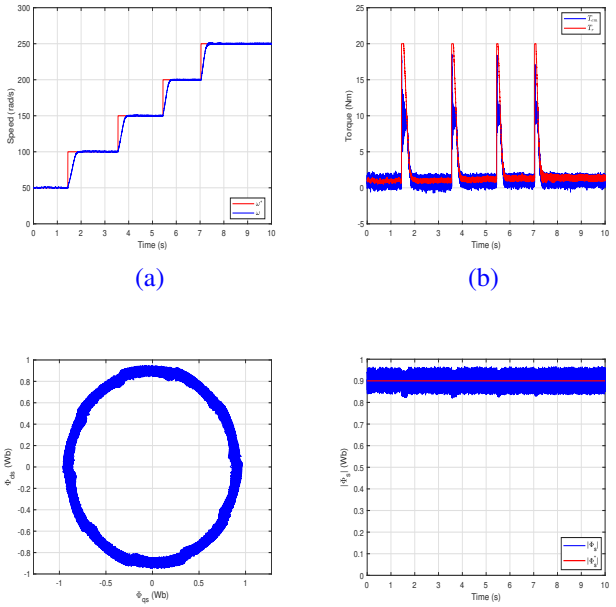


Fig. 6: Experimental results of FCS-MPC: (a) Rotor speed and reference speed, (b) Electromagnetic torque and reference torque, (c) Stator flux trajectory, and (d) Stator flux magnitude with its reference.

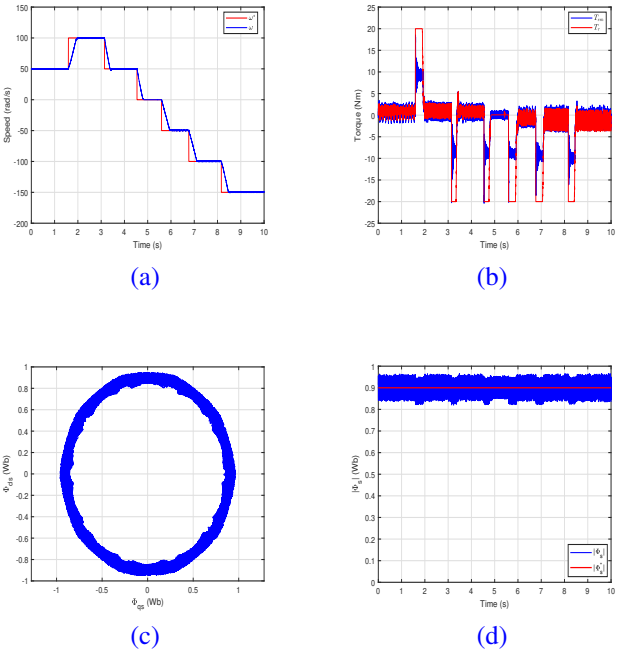


Fig. 7: Experimental results of FCS-MPC : (a) Rotor speed and reference speed, (b) Electromagnetic torque and reference torque, (c) Stator flux trajectory, and (d) Stator flux magnitude with its reference.

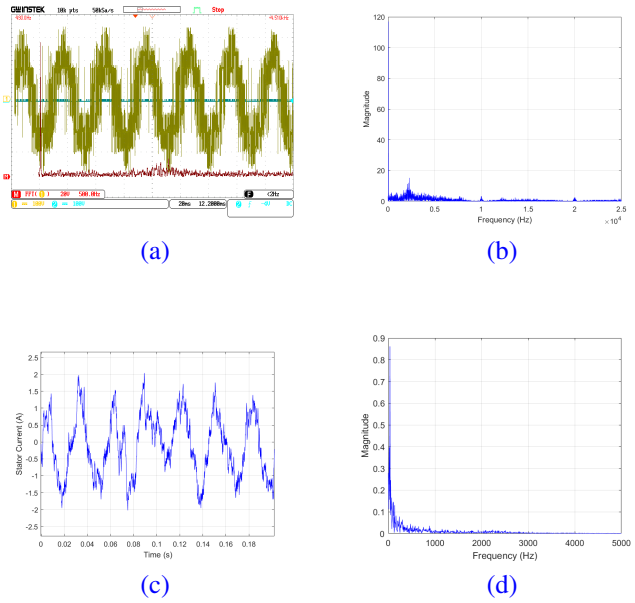


Fig. 8: Experimental results of FCS-MPC: (a) Inverter output voltage, (b) Harmonic spectrum of the inverter output voltage, (c) Stator currents, and (d) Harmonic spectrum of stator currents.

VI. Conclusion

This paper demonstrates that the Finite Control Set Model Predictive Control (FCS-MPC) strategy, originally developed for three-phase induction motors, can be effectively applied to five-phase induction motors. The implementation of FCS-MPC in five-phase systems has yielded promising results, confirming its capability to achieve precise control and stable performance under varying speed conditions.

For future work, further investigation into the impact of the x-y components, which are unique to five-phase induction motors, could help minimize losses and enhance the control strategy. Additionally, extending predictive control techniques to account for fault conditions in five-phase systems would enable full exploitation of their inherent advantages, improving performance and reliability across various industrial applications.

Appendix

Table. I
Five-phase induction motor parameters

Parameter	Value	Parameter	Value
P_{nom}[KW]	3.5	T_{nom}[N.m.]	12.7
$R_s[\Omega]$	9.5	$R_r[\Omega]$	7.3
L_s[H]	1.389	L_r[H]	1.331
L_m[H]	1.323	Pole pair number	1

References

[1] S. A. Gaikwad and S. M. Shinde, “Review on Five-phase Induction Motor fed by Five-phase Voltage Source Inverter with different Conduction Mode,” in *2020 International Conference on Industry 4.0 Technology (I4Tech)*, Pune, India, 2020, pp. 199–202, doi: 10.1109/I4Tech48345.2020.9102695.

[2] M. J. Duran and F. Barrero, "Recent Advances in the Design, Modeling, and Control of Multiphase Machines—Part II," *IEEE Trans. Ind. Electron.*, vol. 63, no. 1, pp. 459–468, Jan. 2016, doi: 10.1109/TIE.2015.2448211.

[3] K. Iffouzar et *al.*, "Improved direct field oriented control of multiphase induction motor used in hybrid electric vehicle application," *Int. J. Hydrogen Energy*, vol. 42, pp. 19296–19308, 2017.

[4] J. Kellner, S. Kaščák, and Ž. Ferková, "Investigation of the Properties of a Five-Phase Induction Motor in the Introduction of New Fault-Tolerant Control," *Appl. Sci.*, vol. 12, no. 4, p. 2249, 2022, doi: 10.3390/app12042249.

[5] G. Kulandaivel et *al.*, "Five-phase induction motor drive—A comprehensive review," *Front. Energy Res.*, vol. 11, p. 1178169, 2023, doi: 10.3389/fenrg.2023.1178169.

[6] L. Parsa, "On advantages of multi-phase machines," in *Proc. IEEE Ind. Electron. Conf. (IECON)*, Raleigh, NC, USA, 2005, pp. 6, doi: 10.1109/IECON.2005.1569139.

[7] M. Bermudez et *al.*, "Open-Phase Fault-Tolerant Direct Torque Control Technique for Five-Phase Induction Motor Drives," *IEEE Trans. Ind. Electron.*, vol. 64, no. 2, pp. 902–911, Feb. 2017, doi: 10.1109/TIE.2016.2610941.

[8] B. Chikondra, U. R. Muduli, and R. K. Behera, "An Improved Open-Phase Fault-Tolerant DTC Technique for Five-Phase Induction Motor Drive Based on Virtual Vectors Assessment," *IEEE Trans. Ind. Electron.*, vol. 68, no. 6, pp. 4598–4609, June 2021, doi: 10.1109/TIE.2020.2992018.

[9] B. S. Khaldi et *al.*, "DTC-SVM Sensorless Control of Five-Phase Induction Motor Based on Two Different Rotor Speed Estimation Approaches," *Nonlinear Dyn. Syst. Theory*, vol. 21, pp. 262–279, 2021.

[10] L. Vancini et *al.*, "Fault-Tolerant Control Strategies of Five-Phase Induction Motor Drives under Open-Switch Fault," in *Proc. Int. Conf. Electr. Mach. (ICEM)*, Valencia, Spain, 2022, pp. 1274–1280, doi: 10.1109/ICEM51905.2022.9910786.

[11] S. C. Rangari, H. M. Suryawanshi, and M. Renge, "New Fault-Tolerant Control Strategy of Five-Phase Induction Motor with Four-Phase and Three-Phase Modes of Operation," *Electronics*, vol. 7, p. 159, 2018, doi: 10.3390/electronics7090159.

[12] H. Guzman, M. J. Duran, F. Barrero, B. Bogado, and S. Toral, "Speed control of five-phase induction motors with integrated open-phase fault operation using model-based predictive current control techniques," *IEEE Trans. Ind. Electron.*, vol. 61, no. 9, pp. 4474–4484, Sep. 2014, doi: 10.1109/TIE.2013.2289882.

[13] E. Levi et *al.*, "Multiphase induction motor drives—A technology status review," *IET Electr. Power Appl.*, vol. 1, pp. 489–516, 2007, doi: 10.1049/iet-epa:20060342.

[14] J. Listwan, "DTC-ST and DTC-SVM control of five-phase induction motor with MRASCC estimator," *Przegląd Elektrotechniczny*, vol. 1, pp. 254–258, 2016, doi: 10.15199/48.2016.11.61.

[15] S. K. Barik and K. K. Jaladi, "Five-Phase Induction Motor DTC-SVM Scheme with PI Controller and ANN Controller," *Procedia Technol.*, vol. 25, pp. 816–823, 2016, *Proc. RAEREST 2016*, doi: 10.1016/j.protcy.2016.08.184.

[16] S. A. Gaikwad and S. M. Shinde, "Five-Phase Induction Motor Modeling and Its Analysis Using MATLAB/Simulink," in *Smart Technologies for Energy, Environment and Sustainable Development*, M. L. Kolhe, S. B. Jaju, and P. M. Diagavane, Eds., Springer Proc. Energy, Springer, Singapore, 2022, doi: 10.1007/978-981-16-6875-3_52.

[17] K. S. Aher and A. G. Thosar, "Modeling and simulation of five-phase induction motor using MATLAB/Simulink," *Int. J. Eng. Res. Appl.*, vol. 6, pp. 1–8, 2016.

[18] A. A. Mekhilef, A. Benachour, E. M. Berkouk, and A. Dali, "FCS-MPC of a DMC-fed Induction Machine with Unity Input Power Factor Using Rotating Vectors," in *Proc. 21st Int. Symp. Power Electron. (Ee)*, Novi Sad, Serbia, 2021, pp. 1–6, doi: 10.1109/Ee53374.2021.9628355.

[19] A. Zeghlache, H. Mekki, M. F. Benkhoris, A. Djerioui, D. Ziane, and S. Zeghlache, "Robust fault-tolerant control of a five-phase permanent magnet synchronous motor under an open-circuit fault," *Appl. Sci.*, vol. 14, no. 12, p. 5190, 2024, doi: 10.3390/app14125190.

[20] M. Bermúdez, H. Guzmán, I. González-Prieto, F. Barrero, M. J. Durán, and X. Kestelyn, "Comparative study of DTC and RFOC methods for the open-phase fault operation of a 5-phase induction motor drive," in *Proc. 41st Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Yokohama, Japan, Nov. 2015, pp. 2702–2707, doi: 10.1109/IECON.2015.7392509.

[21] C. S. Lim et *al.*, "Experimental evaluation of model predictive current control of a five-phase induction motor using all switching states," in *Proc. 15th Int. Power Electron. Motion Control Conf. (EPE/PEMC)*, Novi Sad, Serbia, 4-6 Sept. 2012, pp. LS1c.4-1–LS1c.4-7, doi: 10.1109/EPEPEMC.2012.6397394.

[22] C. S. Lim et *al.*, "FCS-MPC-Based Current Control of a Five-Phase Induction Motor and its Comparison with PI-PWM Control," *IEEE Trans. Ind. Electron.*, vol. 61, pp. 149–163, 2014, doi: 10.1109/TIE.2013.2248334.

[23] H. Guzman et *al.*, "Speed Control of Five-Phase Induction Motors With Integrated Open-Phase Fault Operation Using Model-Based Predictive Current Control Techniques," *IEEE Trans. Ind. Electron.*, vol. 61, pp. 4474–4484, 2014, doi: 10.1109/TIE.2013.2289882.

[24] C. Martín et *al.*, "Five-Phase Induction Motor Rotor Current Observer for Finite Control Set Model Predictive Control of Stator Current," *IEEE Trans. Ind. Electron.*, vol. 63, pp. 4527–4538, 2016, doi: 10.1109/TIE.2016.2536578.

[25] H. Guzman and A. Iqbal et *al.*, "Reduction of common-mode voltage using a simplified FCS-MPC for a five-phase induction motor drive," *J. Eng.*, 2019, doi: 10.1049/joe.2018.8045.

2716-912X © 2025 Ecole Nationale Polytechnique

Geotechnical Characterization of Soils for Road Construction Along Textile Mill Road, Benin City, Nigeria

David Idiata, Ngozi Kayode-Ojo, and Solomon Okonofua

Abstract- This paper presents a geotechnical characterization of soil samples to evaluate their suitability for road construction. Laboratory tests, including specific gravity, sieve analysis, compaction, Atterberg limits, and California Bearing Ratio (CBR), were conducted to assess engineering properties. The results highlight significant variations in particle density, gradation, plasticity, and load-bearing capacity, which critically influence subgrade stability and pavement performance. The findings demonstrate that soils with higher specific gravity and well-graded particle distribution exhibit superior compaction characteristics and reduced moisture susceptibility. Lower plasticity indices correlate with enhanced stability under wet-dry cycles, minimizing long-term maintenance needs. CBR tests further reveal that soils with minimal strength loss under soaked conditions are more resilient, ensuring durability in moisture-prone environments. The study underscores the importance of selecting soils with optimal gradation, density, and plasticity to achieve cost-effective and sustainable road infrastructure. Practical implications for construction practices, including moisture control and stabilization requirements, are discussed, providing actionable insights for engineers and project planners.

***Keywords*- Geotechnical characterization, Soil properties, Road construction, California Bearing Ratio (CBR), Atterberg limits, Compaction**

I. INTRODUCTION

Geotechnical characterization of soils is an important step in the successful design and construction of road infrastructure, especially in areas with a wide range of soil types and difficult environmental conditions. This study looks at the geotechnical evaluation of soils along Textile Mill Road in Benin City, Nigeria. This road is an important transportation route that has had problems with its pavement breaking down because the soil was not properly assessed and the construction was not done right. This study is unique because it uses both traditional and geotechnical testing methods to describe the subgrade soils along this specific corridor. This gives a full picture of how the soils behave when they are under traffic and environmental loads (Amadi et al., 2022 and Ezeh et al., 2023). This study is different from others that have been done in the area because they often used broad soil classifications. This study, on the other hand, looks at the geology and climate of Benin City in a way that is specific to that area. By doing this, it fills in a major gap in localised geotechnical data, which is necessary for building roads in Nigeria's rapidly growing urban areas in a way that lasts.

Manuscript received December 23, 2024 ; revised June 17, 2025.

D.J. Idiata is with Department of Civil Engineering Technology, NICTM, Uromi, Benin City, Edo State. Nigeria.(e-mail: djidiata@gmail.com)
N. Kayode-Ojo and E.S. Okonofua are with Department of Civil Engineering, UNIBEN, Benin City, Edo State, Nigeria. (e-mail: n.kayode-ojo@uniben.edu.ng, e.okonofua@uniben.edu.ng)

The fact that road infrastructure in Benin City keeps failing shows how important this study is. This is often blamed on using the wrong soils and not doing enough geotechnical investigations. The Textile Mill Road, like many others in the area, goes through places with very different types of soil, such as lateritic soils, clayey deposits, and sediments rich in organic matter. Each of these soils has its own unique engineering properties. To design pavements that can last through heavy traffic, seasonal rain, and temperature changes, it is important to understand these differences. This study uses a testing regime that includes specific gravity, particle size distribution, compaction characteristics, Atterberg limits, and California Bearing Ratio (CBR) to see if these soils are good for building roads. We chose these tests because they are well-known in geotechnical engineering and have been shown to be reliable for measuring the quality of subgrade (Osinubi et al., 2009; Amadi, 2014). However, it is important to explain why some advanced tests, like shear strength, compressibility, unconfined compressive strength (UCS), and tensile strength (Ts), which are often used in thorough geotechnical studies, are not included. The main reason these tests were not done is because the focus of this research is on the initial evaluation of subgrade suitability rather than detailed mechanistic modelling. Tests for shear strength and compressibility are important for designing foundations, but they are not as important for flexible pavement systems, where the main concern is the load-bearing capacity of the subgrade, which can be measured by CBR and compaction tests (AASHTO, 1993; NCHRP, 2004). The Nigerian General Specifications for Roads and Bridges (2021 revision) backs this up by saying that these standard tests should be done first on regular road projects. Also, UCS and Ts tests are more useful for cohesive soils or stabilised materials, which are not the main focus of

Digital Object Identifier (DOI):10.53907/enpesj.v5i1.309

2716-912X © 2025 Ecole Nationale Polytechnique

this study. The chosen tests are a good balance of cost and effectiveness for reaching the study's goals without lowering the reliability of the results (Etim et al., 2023; Barksdale &Itani, 1989; Lekha et al., 2013). This is because of the limitations of field sampling and laboratory resources. The geological setting of Benin City makes the testing method used in this study even more appropriate. The Benin Formation, which is part of the Niger Delta Basin, lies beneath the region. It is made up of layers of sand, clay, and gravel that alternate (Omatsola&Adegoke, 1981, Nwankwoala& Whyte, 2023). Because these soils vary a lot from place to place, geotechnical evaluation needs to be done in a specific way. For example, the high clay content in some places makes the soils more likely to swell and shrink. This can be measured accurately using Atterberg limits and compaction tests (Ola, 1983; Gidigasu, 1976). On the other hand, the sandy and lateritic soils that are common in other areas need careful gradation analysis to make sure they drain and stay stable. By focussing on these important factors, the study gives engineers and planners useful information that can help them choose the right soil, stabilise it, and build things. This study has effects that go beyond the area around Textile Mill Road. Nigeria's roads often fail because of poor construction and not enough geotechnical research (Adeyemi et al., 2008; Oyediran&Fadamoro, 2015). This study shows how useful systematic soil characterisation can be and calls for the use of standardised geotechnical protocols in road projects all over the country. The results will also add to the growing body of knowledge about tropical soils, which are often not well covered in global geotechnical literature (Gidigasu, 1976; Osinubi et al., 2009). Combining local soil data with best practices from around the world can make Nigeria's transport infrastructure more sustainable and resilient, which will lower maintenance costs and extend the life of the infrastructure.

II. Study Area

The soil samples examined in this study were obtained from Textile Mill Road in Benin City, the capital of Edo State, Nigeria. The designated sampling location is positioned in front of Christ Embassy Church, a notable landmark on this bustling thoroughfare. Benin City is located in southern Nigeria, at approximately 6.34°N latitude and 5.62°E longitude. This area belongs to Nigeria's humid tropical region, marked by pronounced wet and dry seasons (Ologunorisa and Abawua, 2005). Benin City is a historically important urban centre and the administrative capital of Edo State. The city's geographical position situates it in the rainforest zone, resulting in a comparatively high annual precipitation, averaging between 2,000 and 2,500 mm (Odemerho, 1992). Precipitation in the region is primarily concentrated from April to October, with maximum rainfall occurring between June and September. The arid season lasts from November to March, characterised by reduced humidity and the impact of harmattan winds (Nwafor, 2006). Textile Mill Road is a principal thoroughfare in Benin City, connecting industrial, residential, and commercial zones. It endures considerable vehicular congestion, encompassing trucks and heavy-duty vehicles, owing to its nearness to industrial areas and markets. The road comprises both paved and unpaved segments, rendering it vulnerable to soil erosion, sedimentation, and waterlogging, especially during the rainy season (Ogundele et al., 2011). The area's topography is primarily flat to gently undulating, characteristic of Benin City's landscape. The flat topography, combined with substantial precipitation, frequently leads to inadequate drainage, resulting in localised flooding and erosion (Adefolalu, 1991). The region's soils are predominantly lateritic, distinguished by their reddish-brown hue resulting from elevated iron oxide levels. These soils are characteristic of tropical regions and exhibit variability in engineering properties, encompassing sandy to clayey compositions (Osinubi and Nwaiwu, 2006).

Figure 1: Location of sample collection

Figure 2: Google map of Benin City

The region is situated within the Coastal Plain Sands formation, which is a component of the Niger Delta sedimentary basin. This formation primarily comprises unconsolidated sand, clay, and silts, exhibiting varying degrees of permeability (Short and Stauble, 1967). The amalgamation of sandy and clayey deposits affects the geotechnical characteristics of soils in the region, rendering them crucial for construction and infrastructure projects. The area surrounding Christ Embassy Church and Textile Mill Road comprises residential structures, small enterprises, and light industrial activities. These activities lead to anthropogenic factors such as soil compaction, contamination,

and surface runoff, which can modify the inherent properties of soil (Egbai et al., 2011).

The study area, situated along Textile Mill Road in Benin City, displays quintessential traits of a tropical urban environment. The geographic location, climate, and geological characteristics render it an optimal site for examining soil properties and their relevance to geotechnical applications.

III. LITERATURE REVIEW

The analysed studies reveal a consistent interconnection among these parameters and their combined influence on the durability and performance of roads. The amalgamation of testing outcomes, bolstered by statistical and modelling methodologies, provides a robust framework for forecasting soil behaviour and guiding design decisions. Given the escalating demand for robust infrastructure due to climate variability and heightened traffic loads, extensive geotechnical testing is not merely advantageous—it is essential. The design and construction of sustainable and resilient road infrastructure rely heavily on precise geotechnical characterisation of subgrade soils. Inadequately defined or inappropriate subgrade conditions frequently lead to pavement failures, irregular settlements, and expensive maintenance. Thus, geotechnical soil tests—including specific gravity, gradation, Atterberg limits, compaction, and California Bearing Ratio (CBR)—are crucial for assessing the engineering characteristics of soils and forecasting their performance under load and environmental conditions. Specific gravity (SG) is a fundamental property that indicates the density of soil solids in relation to water. Determining void ratio, porosity, and degree of saturation is essential, as these factors affect load-bearing performance. Anyaegbunam et al. (2021) assert that specific gravity aids in determining soil mineralogy and assessing compaction parameters. Soils exhibiting elevated specific gravity are generally abundant in heavy minerals, resulting in enhanced stability under traffic loads. A study by Bello et al. (2019) indicated that lateritic soils with specific gravity (SG) ranging from 2.60 to 2.75 exhibited superior performance as subgrades compared to those with lower SG values, attributable to their denser composition. Oke et al. (2018) illustrated that specific gravity can distinguish between organic and inorganic soils, a differentiation essential for road construction. Organic soils typically exhibit specific gravity values below 2.50, signifying inadequate compaction potential and diminished strength. The specific gravity (SG) significantly affects the California Bearing Ratio (CBR) and compressibility properties, with elevated SG typically resulting in enhanced mechanical performance (Edeh&Igwe, 2020). Gradation or sieve analysis ascertains the distribution of gravel, sand, silt, and clay particles within the soil, which profoundly influences strength, permeability, and stability. Well-graded soils facilitate superior interlocking and diminish voids, thereby improving compaction and decreasing permeability. Osinubi and Nwaiwu (2017) assert that well-graded granular soils are optimal for base and subbase courses, providing enhanced shear strength and superior drainage. Ajayi et al. (2022) highlighted that inadequately graded or uniformly graded soils frequently experience differential settlements and moisture susceptibility. The research contrasted two subgrade soils with varying gradation curves, revealing that the well-graded sample exhibited a 25% superior California Bearing Ratio (CBR) and a greater maximum dry density (MDD). Adebisi and Akinyele (2020) demonstrated that gradation influences the optimum moisture content (OMC) during compaction. Coarse-grained soils, characterised by a diverse particle size distribution, typically necessitate less water and attain superior compaction compared to fine-grained soils. Gradation is also associated with the Atterberg limits and compaction characteristics. Nwachukwu et al. (2019) demonstrate that clayey soils characterised by a substantial fine fraction and inadequate gradation exhibit elevated plasticity indices and necessitate chemical stabilisation for enhanced performance. The Atterberg limits are essential for the classification of fine-grained soils and the evaluation of their plasticity, workability, and moisture sensitivity. The liquid limit (LL), plastic limit (PL), and plasticity index (PI) denote soil behaviour under different moisture conditions. Soils exhibiting elevated liquid limit (LL) and plasticity index (PI) are expansive and pose challenges for road construction due to their swelling and shrinkage characteristics. Akinwumi et al. (2021) assessed the Atterberg limits of lateritic and clayey soils from southwestern Nigeria, concluding that soils with plasticity index values exceeding 20% exhibited inadequate performance as subgrades without stabilisation. Conversely, soils with a Plasticity Index below 10% exhibited negligible volume alteration and were appropriate for light to medium traffic roadways. Yahaya et al. (2018) observed that the Atterberg limits facilitate the prediction of compaction behaviour and California Bearing Ratio (CBR). Soils exhibiting reduced Liquid Limit (LL) and Plasticity Index (PI) generally demonstrate superior compaction and enhanced bearing capacity. Musa et al. (2022) identified a significant inverse correlation between Plasticity Index (PI) and California Bearing Ratio (CBR); highly plastic soils demonstrated CBR values as low as 4%, whereas soils with low plasticity achieved values exceeding 20%, suggesting their appropriateness for subgrade and subbase applications. Moreover, Atterberg limits facilitate soil classification within the AASHTO and Unified Soil Classification Systems (USCS), informing decisions regarding necessary treatments or replacements prior to road construction (Ola, 2017). Compaction is the mechanical densification of soil to augment its strength, diminish settlement, and enhance resistance to deformation. The standard Proctor and modified Proctor tests provide parameters including Maximum Dry Density (MDD) and Optimum Moisture Content (OMC), which inform field compaction activities. Research conducted by Ogundipe et al. (2019) demonstrates that well-compacted soils facilitate superior load distribution and extend pavement longevity. It was determined that MDD values ranging from 1.70 to 2.10 g/cm³ offered sufficient subgrade support in tropical lateritic soils. OMC, conversely, fluctuates according to soil type. Fine-grained soils generally demonstrate elevated optimum moisture content, rendering them more susceptible to construction scheduling, particularly during periods of precipitation. Ibrahim et al. (2020) noted that soil type and

gradation substantially affect compaction behaviour. Coarse-grained soils attain elevated Maximum Dry Density (MDD) at reduced Optimum Moisture Content (OMC), whereas fine-grained and organic soils necessitate increased moisture yet frequently do not achieve optimal MDD levels. This underscores the necessity of pre-construction soil analysis to customise compaction specifications effectively. Afolayan et al. (2022) investigated the impact of different compactive efforts on clayey soils and discovered that heightened compactive energy increased the maximum dry density (MDD) by as much as 15%, specifically for soils with low plasticity index (PI) values. For highly plastic soils, the enhancement was minimal, indicating that compaction alone may be inadequate without stabilisation. The California Bearing Ratio is a penetration test that assesses the strength and stiffness of soil and subgrade materials under simulated loading conditions. It continues to be one of the most essential factors in pavement design, especially for flexible pavements. CBR values directly indicate the soil's appropriateness for use as subgrade, subbase, or base material. Olowofela and Okogbue (2017) demonstrated that lateritic soils with CBR values exceeding 15% are typically appropriate as subgrade materials, whereas those below 5% necessitate stabilisation or replacement. Ogbonna et al. (2021) performed comprehensive CBR testing on expansive clay soils and discovered that CBR values significantly decreased with elevated moisture content, emphasising the necessity of effective drainage and moisture management in pavement construction. The research highlighted the relationship between Atterberg limits and CBR values, indicating that soils with higher plasticity indices consistently exhibited lower CBR values. Research conducted by Okunlola and Ibrahim (2019) evaluated the California Bearing Ratio (CBR) of different tropical soils, demonstrating that sandy soils attained CBR values exceeding 30% in their natural condition, whereas clayey soils necessitated cement or lime stabilisation to surpass the minimum design threshold of 10–15%. Furthermore, CBR testing under saturated conditions replicates worst-case scenarios, and this methodology is crucial in regions susceptible to water infiltration. Contemporary methods of geotechnical characterisation prioritise integrated testing, wherein the outcomes of one test influence the anticipations or modifications required for subsequent tests. Arowojolu and Adedokun (2020) proposed a classification model utilising Atterberg limits and gradation to predict CBR values. This model enhances efficiency in initial investigations and diminishes testing expenses. Ezeokonkwo et al. (2020) developed a regression model correlating MDD, OMC, PI, and specific gravity with CBR values, illustrating that a comprehensive approach provides more dependable subgrade classification than individual testing. Mugabe and Muriithi (2021) evaluated road failure in low-volume roads in East Africa and attributed the cause to insufficient geotechnical testing during the design phase. Soils exhibiting high plasticity and inadequate drainage were utilised without appropriate compaction or stabilisation, resulting in pavement failure during seasonal precipitation. The research emphasised the importance of employing Atterberg limits, compaction, and CBR data to tailor designs to regional soil characteristics. In Nigeria, Fatoba and Alabi (2018) examined over 30 unsuccessful road projects and determined that inadequate analysis of CBR and Atterberg limits contributed to over 50% of design failures. They promoted compulsory geotechnical testing, especially in clay-dominant areas with significant precipitation. Innovative technologies in geotechnical characterisation encompass expedited field testing techniques and artificial intelligence for forecasting test outcomes. Akinlolu et al. (2023) elucidate that machine learning models can forecast CBR values utilising data from Atterberg limits, gradation, and compaction. These tools offer efficiency, particularly in extensive projects where time and resources are constrained. The geotechnical characterisation through specific gravity, gradation, Atterberg limits, compaction, and CBR tests constitutes the foundation of effective road construction methodologies. These tests yield critical insights into soil behaviour in response to loading, variations in moisture, and compaction efforts. Specific gravity aids in determining soil mineralogy and density; gradation affects stability and drainage; Atterberg limits evaluate plasticity and potential volume change; compaction establishes the optimal moisture and density for performance; and CBR measures strength and stiffness for pavement design.

IV. Research Methodology

This study employed a systematic, multi-stage research methodology to evaluate the geotechnical properties of soils along Textile Mill Road in Benin City, Nigeria. The approach combined field investigations with comprehensive laboratory testing, following international standards while adapting to local conditions. The methodology was designed to provide reliable data for road construction while addressing the unique challenges posed by tropical residual soils in the region. The research focused on a strategic location along Textile Mill Road, selected based on preliminary visual surveys and historical records of pavement distress. Sampling followed the Nigerian General Specifications for Roads and Bridges (2021) guidelines, with disturbed and undisturbed samples collected from 0.5-1.5m depths at each location. A modified Shelby tube sampler was used to obtain undisturbed samples for compaction and CBR tests, while disturbed samples were collected for classification and index property tests. The sampling strategy accounted for seasonal variations by conducting collections during both wet and dry seasons, as recommended by Amadi et al. (2022) for tropical soil studies.

The laboratory investigation comprised a suite of tests selected to evaluate fundamental engineering properties relevant to road construction. Particle size distribution analysis was performed using both wet sieving (for particles >75μm) and hydrometer methods (for fines), following ASTM D6913 and D7928 standards. Atterberg limits were determined according to ASTM D4318, with particular attention to the plasticity characteristics that influence soil behaviour under moisture variations. Compaction characteristics were evaluated using the standard Proctor test (ASTM D698), crucial for understanding the moisture-density relationship of the soils. The CBR tests were conducted following ASTM D1883, with both soaked and unsoaked conditions evaluated to simulate worst-case and typical field scenarios. This dual approach, as advocated by Sani et al. (2023), provides more realistic performance predictions for

Benin City's tropical climate. The soaking period of 96 hours accounted for potential prolonged water exposure during heavy rains.

V. Results And Discussion

The results and discussion with respect to the geotechnical tests are presented below: The specific gravity test results (Table 1) for samples C1 and C2 reveal distinct differences in their soil particle densities. Sample C1 has an average specific gravity of 2.39, while C2 exhibits a higher value of 2.60. This suggests that C2 contains denser mineral particles, possibly due to a higher concentration of heavy minerals like iron oxides or quartz. The consistency in C2's measurements (2.61 and 2.60) indicates precise lab work, whereas C1's slight variation (2.37 and 2.40) may reflect minor heterogeneity in the sample. These values are critical for understanding soil behaviour, as higher specific gravity often correlates with better load-bearing capacity and lower porosity. The results align with typical ranges for sandy or clayey soils, with C2's higher density hinting at superior engineering properties for construction applications.

Table 1. Specific Gravity Test

S/N	POINTS	SPECIFIC GRAVITY
1	C1	2.39
2	C2	2.60

The sieve analysis highlights the gradation differences between C1 and C2(Figures 3 and 4). C1 shows 93.1% passing the 1.18mm sieve, compared to 87.25% for C2, indicating C1 is finer. However, C2 has a higher percentage (55.35%) passing the 0.425mm sieve versus C1 (51.75%), suggesting a more uniform particle distribution.

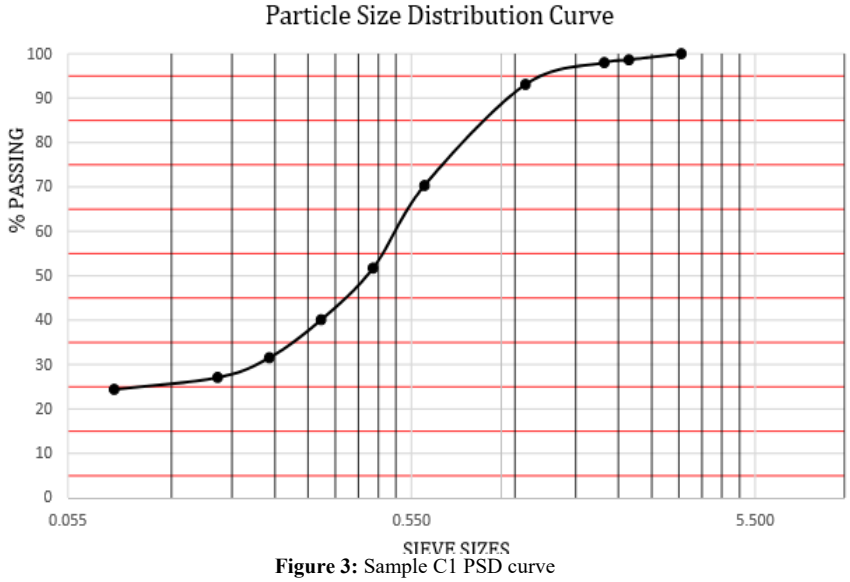


Figure 3: Sample C1 PSD curve

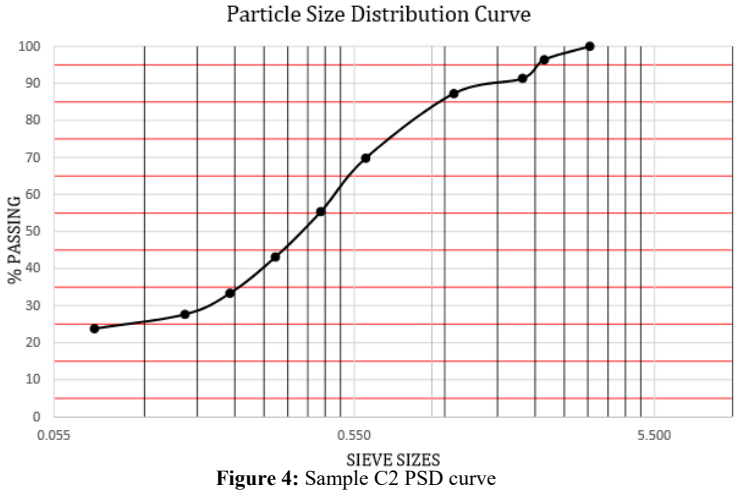


Figure 4: Sample C2 PSD curve

Both samples classify as A-2 soils per AASHTO, with C1 as A-2-6 (higher plasticity) and C2 as A-2-4 (lower plasticity). The finer fraction (<0.075mm) is similar (~24%), but C2's coarser mid-range particles may enhance drainage and reduce compaction effort. These results are vital for determining suitability in subgrade or embankment construction, where gradation affects stability and permeability. The compaction tests show C1 (Figure 5) achieves a higher maximum dry density (MDD) of 2.03 g/cm³ at 11.3% optimum moisture content (OMC), while C2 (Figure 6) reaches 1.99 g/cm³ at 11.2% OMC. The marginal difference in OMC suggests similar water requirements for compaction, but C1's higher MDD implies better particle packing, likely due to its finer texture. Both curves peak sharply, indicating well-graded soils. The lower MDD of C2 may reflect its coarser composition, which could reduce shear strength but improve drainage. These findings guide field compaction practices, emphasizing the need for moisture control to achieve optimal density for load-bearing layers.

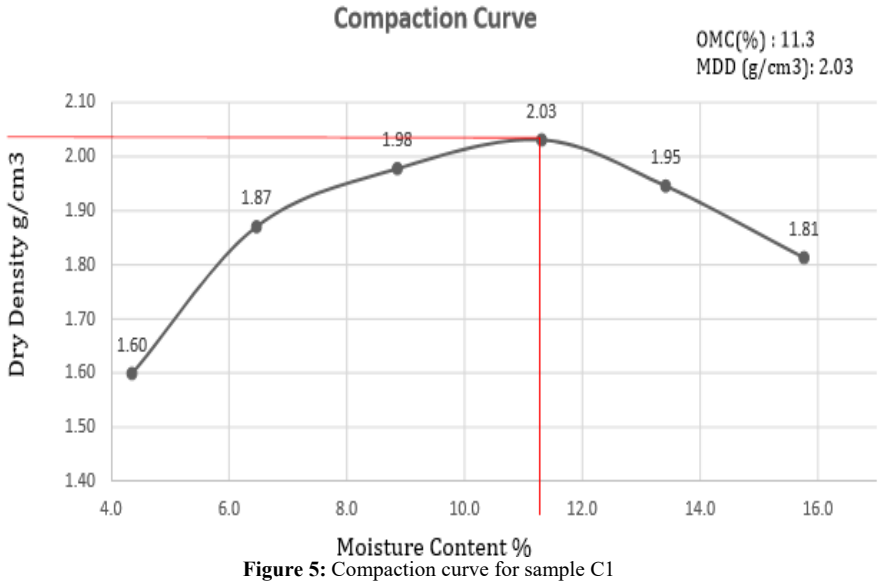


Figure 5: Compaction curve for sample C1

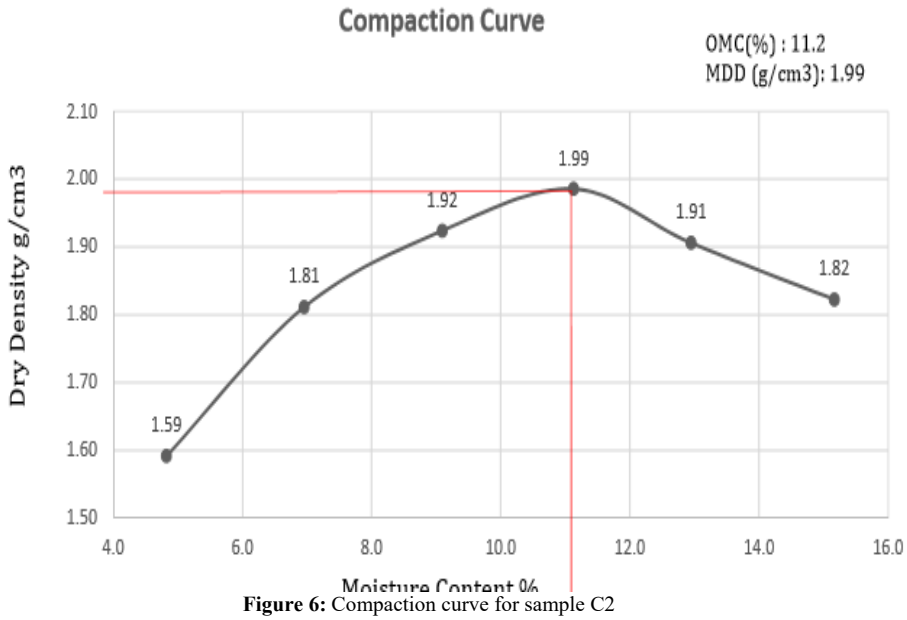


Figure 6: Compaction curve for sample C2

Sample C1 (Figure 7) exhibits higher plasticity (LL=29%, PI=14) than C2 (Figure 8) (LL=22.1%, PI=10), classifying C1 as more clayey and C2 as siltier. The lower plasticity of C2 suggests reduced shrink- swell potential, making it more stable under moisture variations. Both samples have low liquid limits (<30%), typical of subgrade materials, but C1's higher PI indicates greater cohesion, which may enhance strength but increase susceptibility to cracking. These results are crucial for predicting soil behaviour under wet conditions and selecting stabilization methods for road construction.

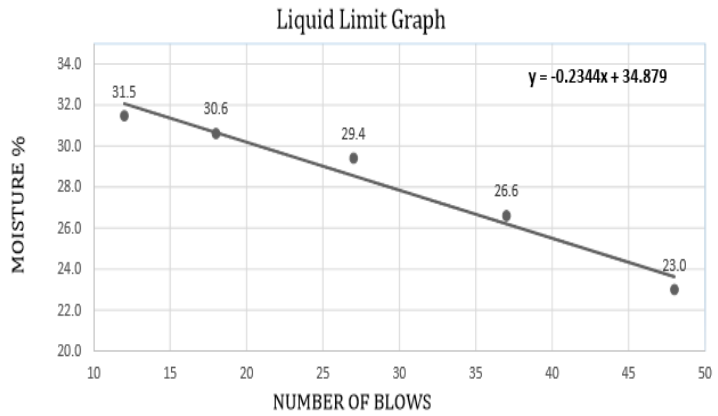


Figure 7: Liquid limit graph for sample C1

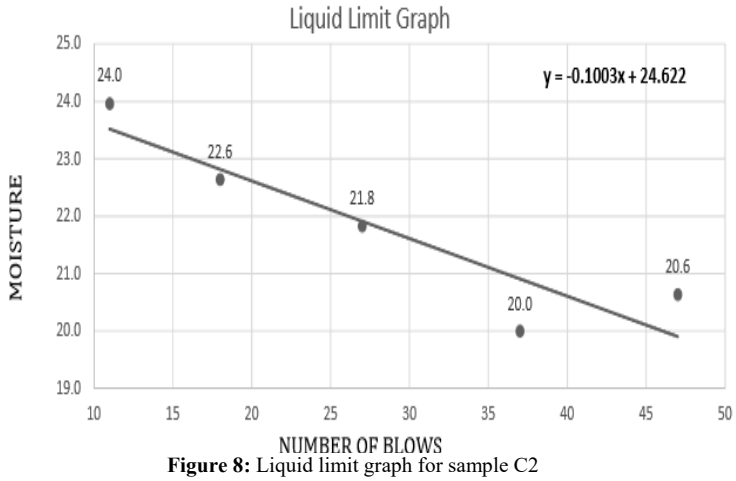


Figure 8: Liquid limit graph for sample C2

The California Bearing Ratio (CBR) tests reveal C2 outperforms C1 in both unsoaked and soaked conditions. For instance, C2's unsoakedCBR at 5.0mm penetration is 16.36% (bottom layer) versus C1's 15.03%. Notably, C2's soaked CBR (15.58%) shows minimal reduction from unsoaked values, indicating superior water resistance, while C1's soaked CBR drops significantly (13.62% to 7.55% for 2.5mm penetration). This underscores C2's better suitability for water-prone areas. The top layer results follow similar trends, reinforcing C2's reliability for pavement subgrades under varying moisture conditions.

Table 2: CBR Values Summary

Sample	Condition	CBR (2.5mm)	CBR (5.0mm)
C1	Unsoaked	11.09	15.03
C1	Soaked	7.55	13.62
C2	Unsoaked	12.39	16.36
C2	Soaked	12.39	15.58

Discussion

The laboratory test results for soil samples C1 and C2 provide critical insights into their suitability for road construction. These tests—specific gravity, sieve analysis, compaction, Atterberg limits, and California Bearing Ratio (CBR)—collectively define the geotechnical properties that influence subgrade stability, compaction efficiency, drainage, and long-term durability. The findings suggest that while both soils fall within acceptable ranges for road construction, C2 exhibits superior engineering properties, particularly in load-bearing capacity and moisture resistance. The specific gravity values for C1 (2.39) and C2 (2.60) indicate differences in mineral composition. A higher specific gravity, as seen in C2, typically suggests the presence of denser minerals such as quartz or iron oxides, which contribute to better shear strength and reduced void ratios. This property is crucial for road subgrades, as denser soils resist deformation under traffic loads more effectively. The consistency in C2's measurements (2.61 and 2.60) also implies sample homogeneity, whereas C1's slight variation (2.37 to 2.40) may indicate minor inconsistencies in particle distribution. For road construction, C2's higher density implies better compaction potential and reduced settlement

risks. The sieve analysis classifies C1 as A-2-6 and C2 as A-2-4 under the AASHTO system. C1 has a higher percentage passing the 1.18mm sieve (93.1% vs. 87.25%), indicating a finer texture, while C2 shows better uniformity in mid-range particles (55.35% passing 0.425mm vs. 51.75% for C1). The near-identical fines content (~24% passing 0.075mm) suggests similar susceptibility to moisture, but C2's coarser fraction enhances drainage, reducing the risk of water retention and subsequent weakening. For road construction, C2's gradation is more favourable because it balances fines (for cohesion) and coarser particles (for permeability). Excessive fines, as in C1, can lead to poor drainage and higher plasticity, increasing the risk of swelling and shrinkage. The A-2 classification confirms both soils as suitable subgrade materials, but C2's A-2-4 designation (lower plasticity) makes it preferable for areas with fluctuating moisture conditions. The compaction tests reveal that C1 achieves a higher maximum dry density (MDD = 2.03 g/cm³) at an optimum moisture content (OMC = 11.3%), while C2 reaches a slightly lower MDD (1.99 g/cm³) at nearly the same OMC (11.2%). The marginal difference in OMC suggests similar water requirements for compaction, but C1's higher MDD indicates better particle packing due to its finer texture. However, C2's compaction curve shows a broader plateau near the peak, meaning it is less sensitive to minor variations in moisture. This is advantageous in field conditions where exact moisture control is challenging. C1's steeper curve implies stricter moisture control is needed to avoid under- or over-compaction. In practice, C2's behaviour reduces the risk of weak spots in the road base, making it more forgiving during construction. The Atterberg limits highlight key differences in soil behaviour under moisture changes. **C1** has a higher liquid limit (LL = 29%) and plasticity index (PI = 14) compared to C2 (LL = 22.1%, PI = 10), classifying C1 as more clayey and C2 as siltier with lower plasticity. High-plasticity soils like C1 are prone to significant volume changes—swelling when wet and shrinking when dry—which can lead to cracking and uneven settlement in pavements. C2's lower PI suggests greater stability under moisture variations, reducing maintenance needs. For road construction, C2's lower plasticity makes it a more reliable subgrade material, especially in regions with seasonal rainfall. The CBR test is a critical indicator of a soil's strength under traffic loads. C2 outperforms C1 in both soaked and unsoaked conditions. In the unsoaked state, C2's CBR (16.36% at 5.0mm penetration) is higher than C1's (15.03%), indicating better resistance to deformation. More importantly, C2's soaked CBR (15.58%) shows minimal reduction from its unsoaked value, whereas C1's soaked CBR drops sharply (13.62% to 7.55% at 2.5mm penetration). This resilience to water is crucial for road longevity, as subgrades often experience moisture infiltration. C2's superior soaked CBR implies it will maintain structural integrity even in wet conditions, reducing the risk of potholes and rutting. C1's significant strength loss when saturated suggests it may require stabilization (e.g., lime or cement treatment) if used in high-moisture environments.

VI. CONCLUSION

The comprehensive geotechnical evaluation of soil samples C1 and C2 provides critical insights into their suitability for road construction, highlighting key differences in their engineering properties and long-term performance. The laboratory tests—specific gravity, sieve analysis, compaction characteristics, Atterberg limits, and California Bearing Ratio (CBR)—collectively paint a clear picture of how these soils will behave under the stresses imposed by traffic loads, environmental conditions, and moisture fluctuations. The findings underscore the importance of selecting the right soil type to ensure structural integrity, minimize maintenance costs, and enhance the durability of road infrastructure. C2 emerges as the superior material for road construction due to its favourable geotechnical properties. Its higher specific gravity (2.60) suggests a denser mineral composition, which translates to better load-bearing capacity and reduced susceptibility to settlement. The sieve analysis further supports this advantage, revealing a well-graded particle distribution that balances cohesion and permeability. With 55.35% of particles passing the 0.425mm sieve and only 23.8% fines, C2 offers optimal drainage characteristics, reducing the risk of water retention that could weaken the subgrade over time. This is particularly crucial in regions prone to heavy rainfall or seasonal moisture variations, where poor drainage can lead to premature pavement failure. The compaction test results reinforce C2's practicality in field applications. While it achieves a slightly lower maximum dry density (1.99 g/cm³) compared to C1 (2.03 g/cm³), its broader compaction curve indicates greater flexibility in moisture control during construction. This is a significant advantage in real-world scenarios, where maintaining exact moisture levels can be challenging. C1's steeper curve, on the other hand, demands stricter quality control, increasing the risk of under- or over-compaction if conditions deviate even slightly from the optimum. The reduced sensitivity of C2 to moisture variations makes it a more forgiving and cost-effective choice for large-scale projects. The Atterberg limits further differentiate the two soils, with C2's lower plasticity index (PI = 10) indicating reduced susceptibility to volume changes under wet and dry cycles. In contrast, C1's higher PI (14) suggests a greater tendency for swelling and shrinkage, which can lead to cracking and uneven settlement in pavement layers. For engineers, this means that while C1 may provide adequate short-term performance, C2 offers long-term stability with fewer

maintenance demands. This distinction is particularly important in climates with significant seasonal weather changes, where expansive soils can cause recurrent damage to road surfaces. The CBR test results deliver perhaps the most compelling evidence of C2's superiority. Its higher unsoakedCBR (16.36%) compared to C1 (15.03%) confirms better resistance to deformation under load. More critically, C2's soaked CBR (15.58%) shows minimal reduction from its unsoaked value, demonstrating remarkable resilience to water infiltration. C1, however, suffers a drastic drop in strength when saturated (from 13.62% to 7.55%), highlighting its vulnerability to moisture. This stark contrast underscores the risks of using C1 in areas with high groundwater levels or poor drainage, where prolonged exposure to water could compromise the entire pavement structure. From a practical standpoint, the choice between these soils hinges on both engineering and economic considerations. C2's inherent strengths—superior drainage, lower plasticity, and consistent load-bearing capacity—make it the preferred material for subgrade and base layers, particularly in high-traffic or moisture-prone areas. Its use would likely reduce the need for costly stabilization techniques and extend the service life of the road. C1, while usable, would require additional measures such as chemical stabilization or improved drainage systems to mitigate its limitations. These interventions add to construction costs and complexity, making C2 a more sustainable option in the long run. In conclusion, the geotechnical characterization of C1 and C2 reaffirms a fundamental principle in road construction: the importance of selecting materials that align with both environmental conditions and performance requirements. C2's balanced properties—gradation, density, plasticity, and moisture resistance—make it an ideal choice for durable, low-maintenance roads. By contrast, C1's higher plasticity and moisture sensitivity necessitate careful engineering interventions to ensure comparable performance. These findings not only guide material selection for this specific project but also emphasize the value of thorough soil testing in achieving cost-effective and resilient infrastructure. Ultimately, investing in the right soil at the construction phase pays dividends in reduced maintenance and enhanced road longevity, benefiting both builders and end-users alike.

REFERENCES

[1] AASHTO. (1993). Standard Specifications for Transportation Materials and Methods of Sampling and Testing.

[2] Adebowale, M. O., & Gala, D. B. (2022). Electrical resistivity correlations for soil characterization. Journal of Applied Geophysics, 203, 104696

[3] Adeleke, B. O., Ola, S. A., &Afolagboye, L. O. (2023). Temperature effects on tropical lateritic soils: Implications for pavement design. Bulletin of Engineering Geology and the Environment, 82(4), 156.

[4] Adeyemi, G. O., Oyeyemi, K. F., &Akinmusuru, J. O. (2022). Geotechnical characterization of lateritic soils for low-volume roads in southwestern Nigeria. Nigerian Journal of Technology, 41(2), 245-253. https://doi.org/10.4314/njt.v41i2.1

[5] Adeyemi, G. O., Oyeyemi, K. F., &Olorunfemi, M. O. (2008). Geotechnical evaluation of some soils from part of southwestern Nigeria as liners in waste landfills. Environmental Geology, 54(1), 145-155.

[6] Amadi, A. A. (2014). Evaluation of changes in geotechnical properties of residual soils along a weathering profile. Journal of Earth System Science, 123(1), 23-35.

[7] Amadi, A. A., &Osinubi, K. J. (2021). Influence of wetting-drying cycles on compacted lateritic soils for road construction. Transportation Infrastructure Geotechnology, 8(3), 412-429. https://doi.org/10.1007/s40515-021-00160-z

[8] Amadi, A. N., Eze, C. J., &Igwe, O. (2022). Advanced characterization of tropical soils for sustainable infrastructure development. Journal of African Earth Sciences, 185, 104392..

[9] American Society for Testing and Materials. (2016). Standard test method for California Bearing Ratio (CBR) of laboratory-compacted soils (ASTM D1883-16). ASTM International.

[10] American Society for Testing and Materials. (2017). Standard test methods for liquid limit, plastic limit, and plasticity index of soils (ASTM D4318-17). ASTM International.

[11] American Society for Testing and Materials. (2017). Standard test methods for laboratory compaction characteristics of soil using standard effort (ASTM D698-12e2). ASTM International.

[12] American Society for Testing and Materials. (2017). Standard test methods for particle-size distribution (gradation) of soils using sieve analysis (ASTM D6913-17). ASTM International.

[13] Barksdale, R. D., &Itani, S. Y. (1989). Influence of aggregate shape on base behavior. Transportation Research Record, 1227, 173-182.

[14] Eberemu, A. O., &Osinubi, K. J. (2022). Geotechnical properties of lateritic soil stabilized with cement kiln dust. Geotechnical and Geological Engineering, 40(1), 413-428.

[15] Etim, R. K., Attah, I. C., &Yohanna, P. (2022). Palm kernel ash stabilization of tropical lateritic soil: Modified testing protocols. Case Studies in Construction Materials, 16, e00841. https://doi.org/10.1016/j.cscm.2021.e00841

[16] Etim, R. K., Attah, I. C., &Yohanna, P. (2023). Sustainable road construction in developing nations: Balancing cost and technical requirements. International Journal of Pavement Engineering, 24(3), 345-361.

[17] Ezeh, C. C., &Chukwujama, I. A. (2023). Geotechnical properties of Benin City soils: New perspectives for urban development. Bulletin of Engineering Geology and the Environment, 82(1), 45.

[18] Lekha, K. R., Kavitha, V., & Shankar, A. U. R. (2013). Laboratory investigation of rubberized soil for highway embankments. Geotechnical and Geological Engineering, 31(2), 681-690.

[19] Liu, J., & Wang, Y. (2023). Machine vision for rapid soil classification in field conditions. Automation in Construction, 145, 104632.

[20] Martínez-Fernández, P., & Sánchez-Rodríguez, D. (2023). Portable FTIR spectroscopy for soil characterization in earthworks. Construction and Building Materials, 367, 130312. https://doi.org/10.1016/j.conbuildmat.2022.130312

[21] NCHRP. (2004). Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures.

[22] Nwankwoala, H. O., & Whyte, I. (2023). Geospatial analysis of sedimentary deposits in the Niger Delta Basin. Journal of Sedimentary Environments, 8(2), 89-104

[23] Nwankwoala, H. O., & Whyte, I. (2023). Particle shape analysis of Niger Delta sands for road construction. Geotechnical Research, 10(1), 3-15.
[24] O'Kelly, B. C., &Nogal, M. (2022). Microstructural analysis of compacted soils using X-ray CT. Géotechnique, 72(8), 669-683. https://doi.org/10.1680/jgeot.20.P.387
[25] Ola, S. A. (1983). The geotechnical properties of the black cotton soils of northeastern Nigeria. Quarterly Journal of Engineering Geology and Hydrogeology, 16(1), 69-78.
[26] Omatsola, M. E., &Adegoke, O. S. (1981). Tectonic evolution and Cretaceous stratigraphy of the Dahomey Basin. Journal of Mining and Geology, 18(1), 130-137.
[27] Onyelowe, K. C., &Salahudeen, A. B. (2023). Portable testing kit for rural road projects in developing countries. Arabian Journal of Geosciences, 16(1), 45.
[28] Osinubi, K. J., Bafyau, V., &Eberemu, A. O. (2009). Bagasse ash stabilization of lateritic soil. Journal of Geotechnical and Geoenvironmental Engineering, 135(9), 1264-1273.
[29] Winterkorn, H. F., & Fang, H. Y. (1975). Foundation Engineering Handbook. Van Nostrand Reinhold.
[30] Yoder, E. J., &Witczak, M. W. (1975). Principles of Pavement Design. Wiley.
[31] Zhang, J., & Yang, H. (2022). Modified Proctor test incorporating cyclic loading. Transportation Geotechnics, 35, 100782. Zhang, L., Gu, F., & Luo, X. (2023). Nanoindentation for microscale soil characterization. ActaGeotechnica, 18(4), 1789-1803.
[32] Zhao, Y., Wang, C., & Liu, S. (2023). Laser diffraction analysis for soil particle size distribution. Geotechnical Testing Journal, 46(2), 20220123.
[33] Adebisi, A., &Akinyele, A. (2020). Influence of soil gradation on subgrade performance. Nigerian Journal of Engineering, 27(3), 33–42.
[34] Afolayan, O. A., Adediran, A. A., &Daramola, S. A. (2022). Effects of compactive effort on strength characteristics of clayey soils. Geotechnical Engineering Journal, 44(2), 108–117.
[35] Akinlolu, M., Osinubi, K. J., &Ouhadi, V. R. (2023). Artificial neural networks for CBR prediction using basic soil properties. Journal of African Earth Sciences, 199, 104710. https://doi.org/10.1016/j.jafrearsci.2023.104710
[36] Akinwumi, I. I., Olofinnade, O. M., &Ayedun, H. (2021). Plasticity and stabilization requirements of expansive subgrade soils. Transportation Geotechnics, 28, 100524. https://doi.org/10.1016/j.trgeo.2021.100524
[37] Ajayi, O. A., Fatile, O. A., &Alabi, A. A. (2022). Evaluation of gradation effects on road subgrades. Road Materials and Pavement Design, 23(1), 147–158. https://doi.org/10.1080/14680629.2020.1860635
[38] Anyaegbunam, F. N., Ezeokonkwo, J. U., &Nwankwoala, H. O. (2021). Specific gravity and its correlation with bearing capacity. Geoscience Frontiers, 12(6), 101112. https://doi.org/10.1016/j.gsf.2020.09.009
[39] Arowojolu, O. A., &Adedokun, M. A. (2020). Predictive models for CBR using Atterberg and gradation data. International Journal of Civil Engineering, 18(1), 45–55. https://doi.org/10.1007/s40999-019-00456-0
[40] Bello, A. A., Ige, J. A., & Fagbenle, O. I. (2019). Specific gravity as an indicator of soil suitability. Geotechnical Testing Journal, 42(3), 238–247.
[41] Edeh, J. E., &Igwe, A. O. (2020). Soil mineralogy and engineering response. Nigerian Journal of Technological Development, 17(2), 118–127.
[42] Ezeokonkwo, J. U., Anyaegbunam, F. N., &Nwankwoala, H. O. (2020). Regression-based evaluation of soil bearing capacity. Civil Engineering and Environmental Systems, 37(4), 311–324.
[43] Fatoba, J. O., &Alabi, A. A. (2018). Post-failure analysis of Nigerian roads. Journal of Transportation Engineering, Part B: Pavements, 144(5), 401–415.
[44] Ibrahim, D., Adebayo, A. R., & Okonkwo, O. N. (2020). Compaction behavior of selected tropical soils. Environmental Geotechnics, 7(2), 110–122. https://doi.org/10.1680/jenge.18.00079
[45] Mugabe, J., &Muriithi, C. (2021). Road failure causes in East Africa: A geotechnical analysis. African Journal of Engineering Research, 8(1), 33–45.
[46] Musa, M. A., Adamu, H. B., & Yusuf, A. (2022). Atterberg limits as a basis for CBR estimation. Soils and Foundations, 62(3), 673–685. https://doi.org/10.1016/j.sandf.2022.02.006
[47] Nwachukwu, M. A., Opara, H. E., &Okere, M. I. (2019). Influence of clay content and gradation on soil plasticity. Bulletin of Engineering Geology and the Environment, 78(3), 511–523. https://doi.org/10.1007/s10064-017-1134-9
[48] Ogbonna, A. C., Emenike, M. I., &Akpan, U. G. (2021). Swelling soils and moisture influence on CBR. Geomechanics for Energy and the Environment, 28, 100252. https://doi.org/10.1016/j.gete.2021.100252
[49] Ogundipe, M. E., Akinwumi, I. I., &Fadare, D. A. (2019). Laboratory compaction of lateritic soils. Nigerian Geotechnical Journal, 10(1), 33–41.
[50] Okunlola, O. A., & Ibrahim, Y. A. (2019). Soil strength classification using CBR test. Highway Engineering Review, 4(1), 19–28.
[51] Osinubi, K. J., &Nwaiwu, C. M. (2017). Gradation and mechanical stability of subbase materials. Construction and Building Materials, 156, 115–123.

David Idiata is a Principal Lecturer of the Department of Civil Engineering Technology, School of Engineering Technology, National Institute of Construction Technology and Management (NICTM) Uromi, Edo State, Nigeria. He has been lecturing since 2005 and as an astute scholar, he has presented papers at both local and international conferencesand has published numerous academic research articles in journals across the world. He is a registered Engineer with the Council for the Regulation of Engineering in Nigeria (COREN). He has held major administrative offices e.g. Head of Department, Director of Research, Tetfund Desk officer on Institute Based Research (IBR). Currently he is the Dean of Students and a PhD student of Civil Engineering majoring in Geotechnical Engineering option at University of Benin, Nigeria. He holds two Master Degree.

An 8-bit Digitally-controlled Low NF VGA for 5G Millimeter-wave Beamforming in 65 nm CMOS Technology

Sid Ahmed Tedjini, Sonia Salhi

Abstract—In this paper, a linear-in-dB digitally controlled variable gain low-noise amplifier (D-VGLNA) designed in 65 nm CMOS process for the 5G mm-wave phase-array subsystem is presented. Based on the current steering technique, the designed 8-bit D-VGLNA achieves a gain range of 22 dB from -12 dB to +10 dB while it consumes 9.6 mW (without buffers). By boosting up the input transconductance thanks to a cross-coupled technique, the designed D-VGLNA achieves a noise figure of 2.3 dB, and 7 dB at maximum and 0 dB gain, respectively. It shows an input-output phase shift insensitivity to gain setting thanks to parasitic capacitance mitigation by holding the parasitic capacitance at common junction cascode transistor constant during gain setting. The DVGLNA returns a maximum of -6.5 dBm of 1dB compression point.

***Keywords*—VGA, LNA, 5G, phase array, beamforming.**

I. Introduction

The last couple of decades have been witnessing a major development in CMOS technology making a drastic increase in the transition frequency of MOS transistors. This development in lithography opened up the opportunity for integrated circuit design for frequency bands beyond 20 GHz where large bandwidth and low latencies applications became possible in contrast with the crowded sub-6 GHz band where the majority of communication standards lie [1–3]. In that context, the fifth-generation standard of mobile communication has the aim of providing gigabit-per-second range connectivity necessitating Giga-Hertz range bandwidth. The 28 GHz band has been identified as one possible band for 5G communication [4]. In that context, the effort has been made for developing 5G IC receivers [5, 6]. Designing transceivers for such applications is not an easy task since the incoming radio frequency signal suffers from a strong loss due to the propagation in the atmosphere as well as surrounding obstacles [7, 8]. Increasing the transmitter power is not a possible solution since it increases the power dissipation of the power amplifier particularly if it is integrated in silicon technology. Increasing the antenna gain at the receiver side is then the possible solution to compensate for the propagation loss. However, because of the antenna's narrow beam, signal reception becomes impossible in the case where the transmitter and receiver are not in-sight. The solution comes from receivers based on the phase array system, also called beamformers, where a highly effective antenna gain is achieved thanks to a network of omnidirectional antennas [7] as depicted in Fig. 1. In a phase array system, each antenna sens the incoming radio frequency signal with a given delay with respect to other antennas. In order to avoid a signal-destructive combination prior to frequency down-conversion, each channel incorporates a phase shifter that introduces a given loss.

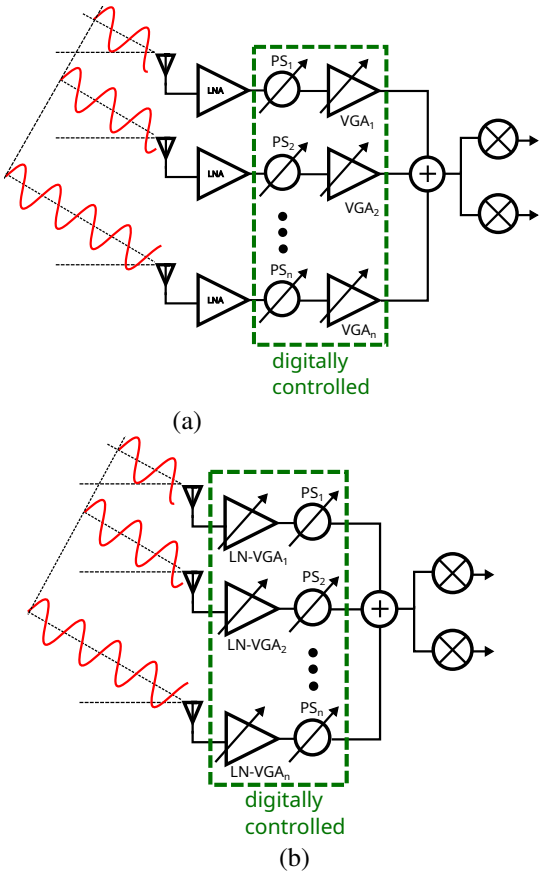


Fig. 1: phase array system (a) conventional architecture, (b) merged VGA, LNA architecture

As depicted in Fig. 1, beamformer receivers integrate a variable gain amplifier to compensate for the loss of the phase shifter [9]. The VGA has to maintain input and output return loss as well as

Manuscript received March 6, 2025; revised June 18, 2025.

Sid Ahmed Tedjini is with Division Microélectronique et Nanotechnologie, Centre de Développement des Technologies avancées, cité du 20 août 1956 Baba Hassen, Algiers, ALGERIA (e-mail: stedjini@cdta.dz).

Sonia Salhi is with Division Microélectronique et Nanotechnologie, Centre de Développement des Technologies avancées, cité du 20 août 1956 Baba Hassen, Algiers, ALGERIA (e-mail: ssalhi@cdta.dz).

Digital Object Identifier (DOI): 10.53907/enpesj.v5i1.323

2716-912X © 2025 Ecole Nationale Polytechnique

phase variation constants during gain sitting, which may result in undesired additional phase and gain variations.

As depicted in Fig. 1(a), the traditional channel path consists of cascading low noise amplifier, phase shifter, and variable gain amplifier. Having a performant VGA in terms of noise can lead to a merged architecture where the LNA and the VGA are merged as a single bloc resulting in complexity reduction as well as power dissipation reduction as illustrated in Fig. 1 (b).

Many analog and digital gain control techniques have been reported in past years. Most popular techniques are current steering [10], variable g_m [11], reflection-type attenuator [12], or ladder-type attenuator VGA [13].

The current steering technique is suitable for mm-wave because of the simplicity of its implementation and also low power dissipation requirement.

In this paper, a digitally controlled 8-bit variable low-noise amplifier is presented. Section II. starts by describing the base of the D-VGLNA which consists of a cross-coupled cascade amplifier. The principle of gain control with current steering is presented in section III. along with a description of the designed D-VGLNA. Simulation results are presented in section IV. before the conclude this paper in section V..

II. MM-WAVE LOW-NOISE AMPLIFIER

A. Technology

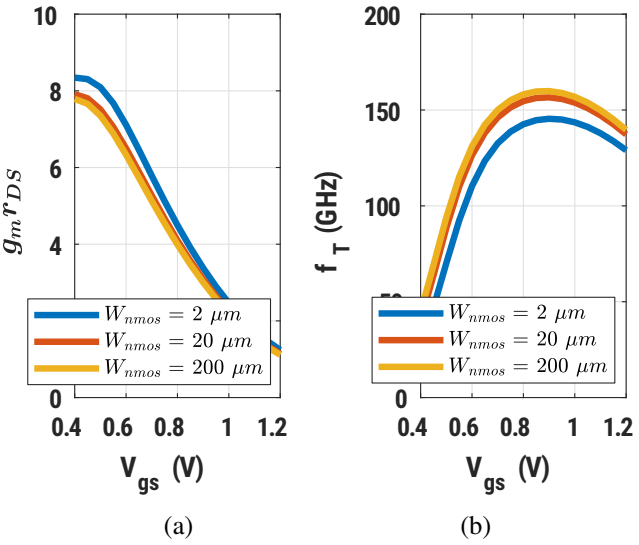
The design of an amplifier at millimeter-wave frequencies requires a large gain-bandwidth product, denoted GBW. The GBW depends on the transistor self-gain as well as its frequency transition denoted f_T. The MOS self-gain, denoted SG, is defined as $g_m \times r_{ds}$, where g_m, and r_{ds} stand for the MOS transconductance and its output resistance, respectively. While frequency transition of the MOS transistor is defined as the frequency of the unity gain's current and is given by the equation 1 which suggests that it is proportional to the transistor g_m.

$$f_T = \frac{g_m}{2\pi(C_{\text{GS}} + C_{\text{GD}})} \tag{1}$$

Fig 2(a) shows the MOS SG variation as a function of V_{GS} for a 65 nm NMOS transistor. The maximum SG is achieved when the MOS transistor operates on a weak inversion region, meaning when V_{GS} is inferior to the threshold voltage. On the other hand, as Fig. 2(b) illustrates, MOS f_T variation according to equation 1 as a function of gate-to-source voltage (V_{GS}) for different gate width, shows that the highest f_T value is achieved when the transistor is operating in strong inversion, meaning for V_{GS} much higher than the threshold voltage V_{TH}. For 65 nm standard CMOS technology, the NMOS device reaches a maximum f_T of 160 GHz.

B. Cross-coupled Cascode Amplifier

As demonstrated in the previous section, transistors are biased on strong inversion in order to achieve the highest f_T. To overcome the insufficient gain resulting from a weak transistor's self-gain in strong inversion operation, the amplifier gain can be raised by cascading a second NMOS device and boosting up the input



(a) (b)

Fig. 2: (a) MOS self-gain variation as function of V_{GS}, (b) NMOS f_T variation as function of V_{GS}

gm thanks to the cross-coupling technique as illustrated in Fig. 3. The input effective transconductance, denoted $g_{\text{m1,eff}}$ is given by [14]:

$$g_{\text{m1,eff}} = \left(1 + \frac{C_c}{C_c + C_{\text{GS}}}\right) g_{\text{m1}}, \tag{2}$$

where C_{GS} is the gate-to-source capacitance of transistors $M_{1,2}$. The boosting amplification factor, denoted A is equal to $C_c/(C_c + C_{\text{GS}})$. If C_c is chosen large enough in comparison with C_{GS}, then $g_{\text{m1,eff}} \approx 2g_{m1}$.

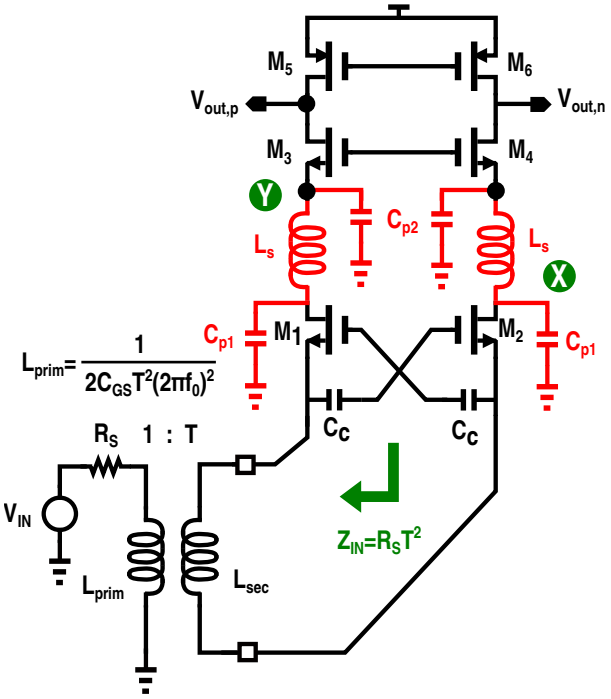


Fig. 3: Cross-coupled cascode amplifier with series shunt peaking inductor

The input transformer brings a couple of advantages. Firstly, it provides single-to-differential transformation creating the input

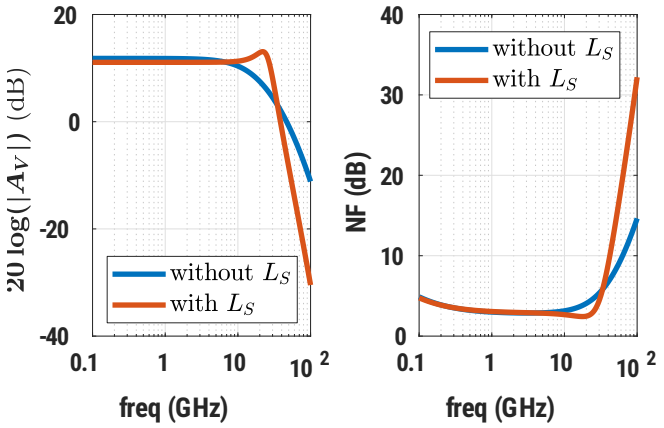


Fig. 4: voltage gain and NF as function of RF frequency of Cascode amplifier with and without series inductive peaking technique

differential signals needed to implement the cross-coupled technique. Secondly, the secondary-to-primary turn ratio increases $M_{1,2}$ noise circulation which helps to decrease the noise factor as shown in section C.. The transformer secondary inductor is chosen to resonate out $2C_{\text{GS1}}$. Hence, L_{prim} is given by:

$$L_{\text{prim}} = \frac{1}{2C_{\text{GS}}T_{\text{in}}^2(2\pi f_0)^2} \tag{3}$$

where T_{in} is transformer turn ratio, and f_0 frequency of incoming RF signal.

The voltage gain of the circuit depicted in Fig. 3 at low frequencies is given by:

$$|A_v| = g_{\text{m1,eff}} \times \left[r_{\text{o3}} + r_{\text{o1}}(1 + g_{\text{m3}}r_{\text{o3}})\right] \parallel r_{\text{o5}} \tag{4}$$

Fig. 4(a) shows the gain as well as the NF of the circuit of Fig. 3. Although C_{GS1} is resonated out by the secondary inductor, the gain starts to roll off beyond 10 GHz due to the parasitic capacitance, loading node X. In addition, C_X ($C_{\text{p1}} + C_{\text{p2}}$) may also increase noise.

Furthermore, as the impedance loading the input device is seen by the input port, the parasitic capacitance $C_X = C_{\text{p1}} + C_{\text{p2}}$ is also affecting input impedance, where C_{p1} and C_{p2} depend on the $M_{1,2}$ and $M_{3,4}$ intrinsic capacitances.

To overcome the drops of the A_V a decoupling of C_{p1} and C_{p2} is necessary thanks to series inductive peaking as shown in Fig. 3.

In fact, the impedance seen at the input RF port is given by:

$$Z_{\text{in}}(s) = \frac{1}{T_{\text{in}}^2}\left(\frac{Z_X(s) + r_{\text{o1}}}{1 + g_{\text{m1}}(1 + A)r_{\text{o1}}}\right) \tag{5}$$

where r_{o1} denotes the $M_{1,2}$ output resistance, and $Z_X(s)$ is the impedance looking back at node X as depicted in Fig. 3. Assuming the $1/g_{\text{m3}}$ the resistance looking back at M_3 source, $Z_X(s)$ is given by equation 6 where s is the Laplace variable.

$$Z_X(s) \approx \frac{1 + sg_{gm3}L_S + s^2L_sC_{\text{p2}}}{g_{m3} + s(C_{\text{p1}} + C_{\text{p2}}) + s^2g_{m3}L_sC_{\text{p2}} + s^3C_{\text{p1}}C_{\text{p2}}L_s} \tag{6}$$

The parasitic capacitor C_{p1} can be approximated by Miller theorem as $C_{\text{gd1}}(1 + g_{\text{m1}}(1 + A)(r_{\text{o1}} \parallel 1/g_{\text{m3}}))$, while the C_{p2} is equal to C_{gs3}. Assuming the inductor is chosen to resonate out C_{p2}, if one makes $C_{p2} = C_{p2}$ the $Z_X(s)$ becomes a purely real function, and is expressed as follows:

$$Z(s) = \frac{g_{\text{m3}}L_S}{C_{\text{p2}}} \tag{7}$$

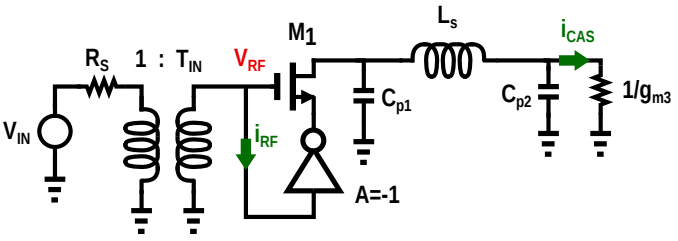


Fig. 5: Half-circuit input impedance analysis

As a consequence, the input impedance seen at the input port is:

$$Z_{\text{in}} = \frac{g_{\text{m3}}L_S + C_{\text{p2}}r_{\text{o1}}}{T_{\text{in}}^2(1 + g_{\text{m1}}(1 + A)r_{o1})} \tag{8}$$

considering the decoupling inductor L_S at the source cascode device, the input and output current ratio as depicted in Fig. 5 is given by:

$$\frac{i_{\text{cas}}}{2g_{\text{m1}}V_{\text{RF}}} \approx \frac{g_{\text{m3}}}{g_{\text{m3}} + s(C_{p1} + C_{p2}) + s^2L_sC_{p1}g_{\text{m3}} + s^3L_sC_{p1}C_{p2}} \tag{9}$$

Consistently, Fig. 6 shows the simulation of $i_{\text{cas}}/2g_{\text{m1}}V_{\text{RF}}$ versus RF frequency demonstrates the low-pass profile of the current gain with a resonance at 28 GHz thanks to a proper L_s value.

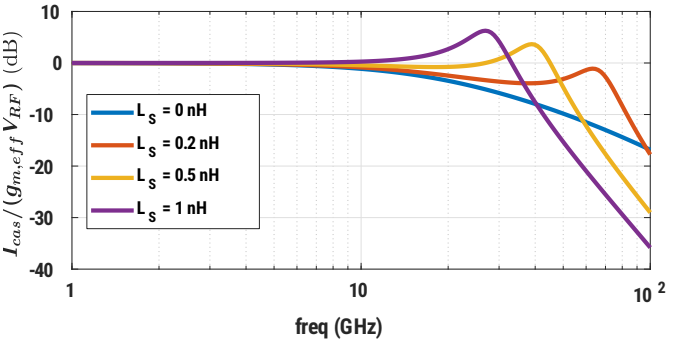


Fig. 6: Schematic-level simulation of $i_{\text{cas}}/2g_{\text{m1}}V_{\text{RF}}$ versus RF frequency

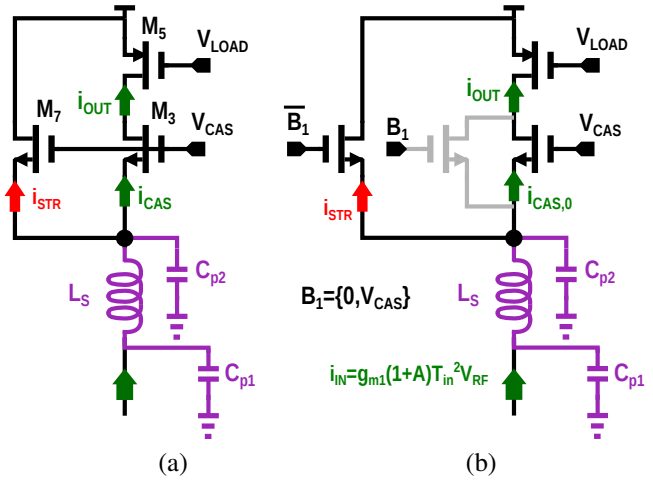


Fig. 7: (a) 1-bit half circuit VGA, (b) C_{p2} variation mitigation for 1-bit VGA

C. *Noise analysis*

As the $1/g_{\text{m3}} << r_{\text{o1}}$ the noise contribution of cascade device is negligible. Assuming that input transistors are the main noise source of the cascode amplifier depicted in Fig. 3. The noise factor is given by :

$$F \approx 1 + \frac{\gamma}{T_{\text{in}}^2 R_S g_{\text{m1}}(1+A)^2} \tag{10}$$

where γ is the noise excess factor of the MOS device and is in order of 3 for 65 nm CMOS technology.

If the denominator of the second term of the right side of the equation. 10 is equal to unity, the amplifier achieves an excessive noise figure 6 dB, according to the above equation. The noise figure can be drastically reduced if $T_{\text{in}}^2 R_S g_{\text{m1}}(1+A)^2 << 1$ while the amplifier is kept matched to RF port impedance. Assuming $A = 1$, and substituting R_S by R_{IN} as given in equation 8, the expression of F under matching condition is re-written as:

$$F \approx 1 + \frac{\gamma r_{\text{o1}}}{(1+A)(g_{\text{m3}} L_S + C_{\text{p2}} r_{\text{o1}})} \tag{11}$$

Thus, the topology of Fig. 3 allows to match the input impdende to R_S which less to 3 dB NF since $r_{\text{o1}}/(2(g_{\text{m3}} L_S + C_{\text{p2}} r_{\text{o1}}) << 1$.

III. D-VGLNA AMPLIFIER

A. *VGA Gain setting principle*

Fig. 7(a) shows the half-side of Fig. 3 including a current steering device also called current bleeding [15]. Obviously, the output current drops because of the current steering, and consequently the amplifier output voltage drops as well.

As the g_{m7} increases, the current quantity steered increases which decreases the amplifier gain. However, C_{p2} is gain-dependent since $C_{\text{P2}} = C_{\text{GS7}} + C_{\text{GS3}}$. To tackle this issue, the cascode device is settled digitally too.

The output current is now given by:

$$i_{\text{out}} = i_{\text{in}} \left[1 - g_{m7} \left(\frac{r_{o7}}{1 + g_{m7} r_{o7}} \parallel \frac{r_{o3} + r_{o5}}{1 + g_{m3} r_{o3}} \right) \right] \tag{12}$$

Assuming $1/g_{m3} << r_{o3}$, and $1/g_{m7} << r_{o7}$, the following approximation can be done.

$$i_{\text{out}} \approx i_{\text{in}} \left[1 - g_{m7} \left(\frac{1}{g_{m7}} \parallel \frac{1}{g_{m3}} \right) \right] \tag{13}$$

The current gain defined as $A_i = i_{\text{cas}}/i_{\text{in}}$ is given by:

$$A_i \approx \frac{g_{m3}}{g_{m3} + g_{m7}} \tag{14}$$

Consequently, the voltage gain A_V can be expressed as a function of A_i, and losses due to steering transconductance, and is given by:

$$A_V \approx r_{\text{out}} g_{\text{m1}} (1+A) T_{\text{in}}^2 \frac{g_{m3}}{g_{m3} + g_{m7}}. \tag{15}$$

The expression of current gain suggests that if cascode and bleeding devices as sized identically in such a way to have $g_{m3} = g_{m7}$ ($W_3 = W_7$), the voltage gain drops with 6 dB as well the current gain. For 3 dB drop, $A_i = 1/\sqrt{2}$, it follows that $g_{m7} = (\sqrt{2} - 1) g_{m3}$.

Therefore, the gain can be settled by making g_{m7} variable. In order to set the gain digitally in the context of a beamformer receiver, the steering device is substituted by a series of transistors settled digitally.

Although the MOS is switched OFF, its parasitic gate-to-source capacitance still exists because of overlapping capacitance. Therefore, because $C_{\text{P2}} = C_{\text{GS3}} + C_{\text{GS7}}$, the $Z_X(s)$ becomes gain dependent and thus the input impedance which is a drawback for other channel paths. To tackle this issue, a simple solution is to switch off a replica transistor, placed in parallel with cascode device, while the steering device is switched on, as shown in Fig. 7(b). In this 1-bit steering-vga scenario, the C_{p2} expression is given by:

$$C_{\text{p2}} = \frac{2 C_{\text{OX}} L}{3} (W_{\text{STR,ON}} + W_{\text{CAS}}) + C_{\text{OV}} W_{\text{STR,OFF}} \tag{16}$$

Because the replica device has the same channel width as the steering device, when the B_1 signal drops to zero, C_{p2} stays unchanged.

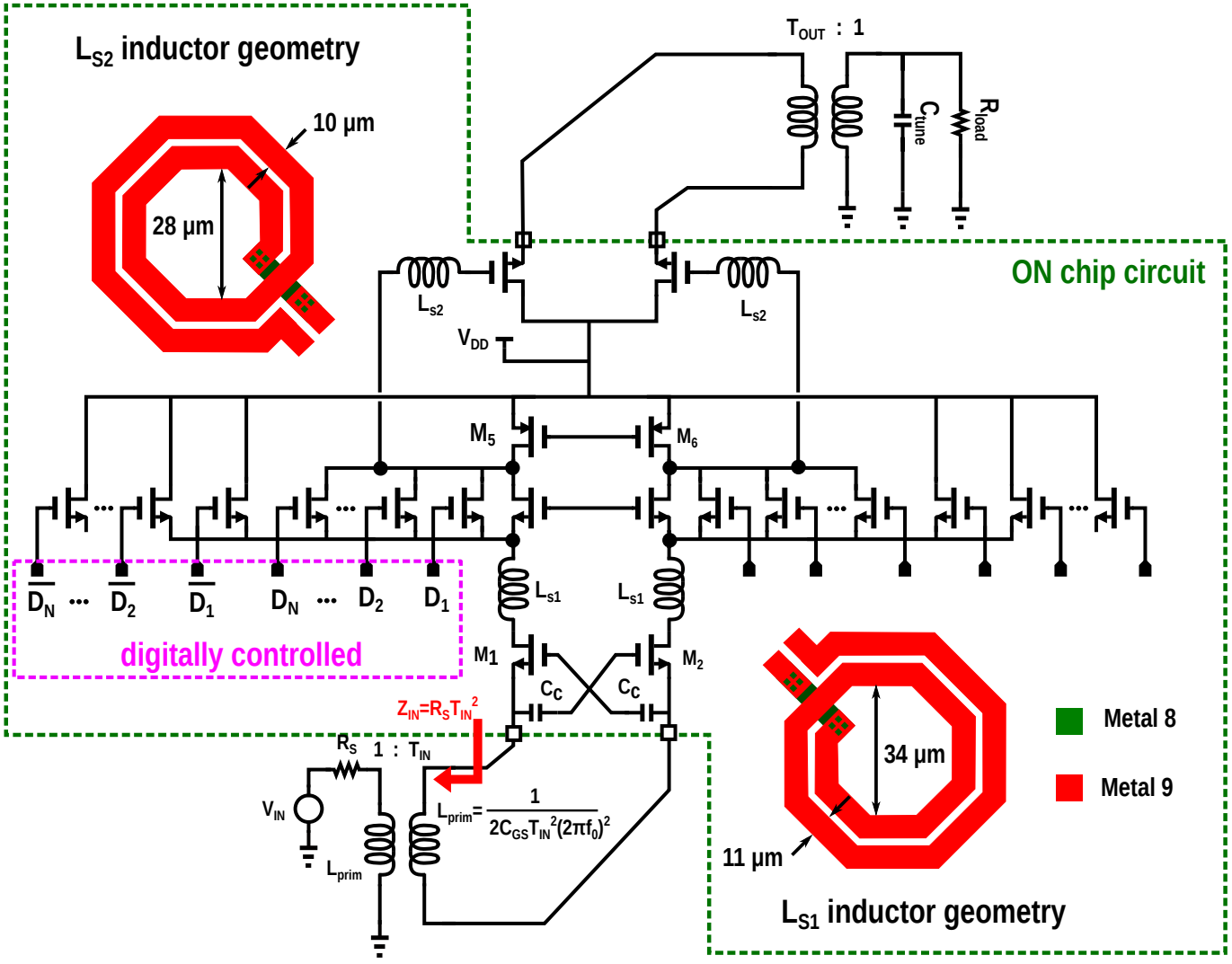


Fig. 8: 8-bit digitally-controlled variable-gain Low-noise amplifier (D-VGLNA)

B. 8-bit digitally-controlled VGA implementation

The implementation of the 8-bit digitally controlled VGA based on the principle described in section A. is depicted in Fig. 8.

An off-chip high-quality factor transformer with a turn ratio of $\sqrt{2}$ is used which is not affecting the noise performance of the VGA. The input transconductance devices $M_{1,2}$ sees an input impedance of 100 Ω which will improve noise circulation and reduce the $M_{1,2}$ channel noise accounted for in the noise factor. The coupling capacitor C_C is much larger than C_{GS} leading to a boosting factor of one. As depicted in Fig. 8, the decoupling inductor L_{S1} is 2 turns coil with width and radius of 11 μm and 34 μm respectively, achieving an inductor value of 985 pH with a quality factor of 7 at 28 GHz. The cascode device is implemented thanks to a series of parallel NMOS transistors, controlled digitally with an 8-bit binary word $\{D_1, D_2, ..., D_8\}$, where the high level D_i is about 0.8 V in order to maintain the cascode devices into saturation when they are switched on. The steering operation is guaranteed by a series of parallel NMOS devices of the same size as cascode devices. controlled with $\{\overline{D_1}, \overline{D_2}, ..., \overline{D_8}\}$ in such way to keep the C_{P2} constant. The C_{P2} expression of 16 is re-written as:

$$C_{p2} = \frac{2C_{OX}L}{3}W_{CAS,0} + \sum_{i=1}^{k}\left(\frac{2C_{OX}L}{3}W_{CAS,i} + C_{OV}W_{STR,i}\right) + \sum_{i=k+1}^{N}\left(\frac{2C_{OX}L}{3}W_{CAS,i} + C_{OV}W_{STR,i}\right) \quad (17)$$

Where the first term stands for the gate-to-source capacitance of the non-controlled cascode devices that stay on for all gain setting configurations, and guarantee the minimum gain. The second term accounts for the gate-to-source capacitance of all steering devices that are set on saturation of a given gain setting plus the gate-to-source capacitance for all replica devices (parallel to cascode) that are on the cutoff region. Similarly the third term accounts for the gate-to-source capacitance of cutoff-biased cascode and steering devices.

The outlined $g_{m3} - g_{m7}$ relationship can be generalized to derive a digitally-controled VGA $g_{m3} - g_{m7}$ relationship of equation for -3n dB of gain drop.

$$\sum_{i=1}^{k} g_{m,str,i} = (2^{n/2} - 1)\left(g_{m,cas,0} + \sum_{i=k+1}^{N} g_{m,cas,i}\right) \quad (18)$$

The above equation suggests that the voltage gain can be decreased by an attenuation factor, denoted AF_{dB}, of 3n dB by switching ON parallel steering devices digitally, where $n = 1, 2, 3...$ On linear domaine, $AF = \sqrt{2}, 2, 2\sqrt{2}, 4,$ etc. Maximum gain is obtained when all steering transistors are switched OFF. The differential output is connected to differential source followers through a 2 turns coil decoupling inductors (L_{s2}) with width and radius of 10 μm and 28.5 μm respectively, achieving an inductor value of 708 pH with a quality factor of 10.5 at 28 GHz. An output balun transformation is achieved by an off-chip one-turn ratio transformer that matches the VGA output impedance to 50 Ω. The capacitor C_{tune} is used to resonate out the secondary at 28 GHz. Fig. 9. shows the plot of A_i as a function of RF frequency for different VGA configurations. When all cascaded transistors are switched ON, $A_i \approx 1$

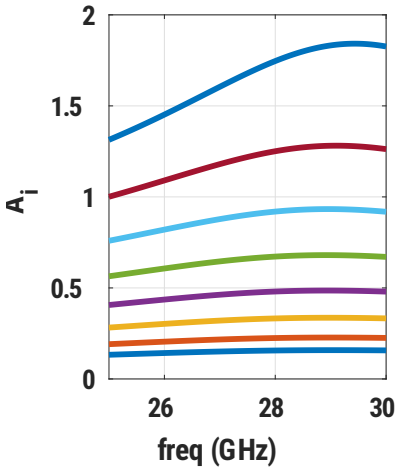


Fig. 9: Current gain for different V_{GS} configuration

As cascode devices start to be switched OFF, the current gain drops to 0.75, etc.

C. Noise analysis

When the noise contribution from the steering devices and the load transistors is taken into account, the noise factor is expressed by 19. The second term of the noise factor represents the noise contribution from the input transistors $M_{1,2}$, while the third and the fourth terms represent the loads, and the steering devices, respectively. The noise accounted for from the input gm and the loads are multiplied by a multiplication factor (between brackets in equation 19) which catches the gain setting effect. In fact, at the maximum gain setting, the index $k = 0$, therefore, the multiplication factor is reduced to unity. However, for the minimum gain setting when $k = N$, the multiplication factor reaches its maximum value of $1 + \sum_{i=1}^{N} g_{\text{m,str,i}}/g_{\text{m,cas,0}}$, increasing drastically the noise contribution of $M_{1,2}$ as well as $M_{5,6}$ transistors.

IV. Simulation Results

All performances presented in this section are obtained by simulation with the Cadence Spectre tool.

A. VGA performance

Fig. 10(a) shows the input return loss as a function of the frequency for different configurations. Clearly, the S_{11} experiences only a little variation over all frequencies with a maximum value of -14 dB in the vicinity of 28 GHz. The same observation is made for the output return loss shown in Fig. 10 which is less than -13 dB around 28 GHz thanks to the use of source followers. Fig. 11(a) shows the variation of S_{21} as a function of frequency for all VGA configurations.

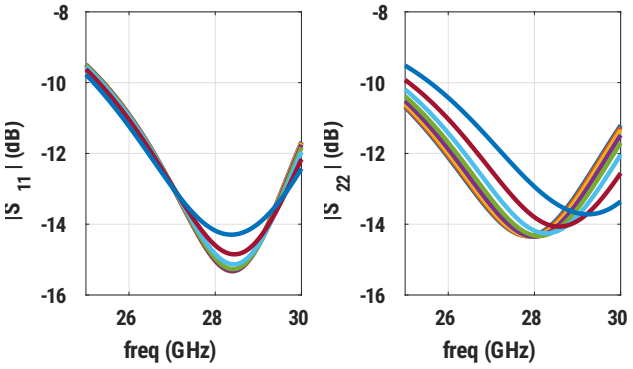


Fig. 10: (a) D-VGLNA Input return loss, (b) D-VGLNA output return loss

At 28 GHz, the S_{21} is equal to 10 dB when all steering devices are OFF as mentioned in section B., and drops with a pace of 3dB as the steering is switched ON to fall to -12 dB when all steering devices are ON. Simulation of phase shift from input to output is shown in Fig. 11(b). As the devices are switched ON and OFF, the phase shift does not vary much.

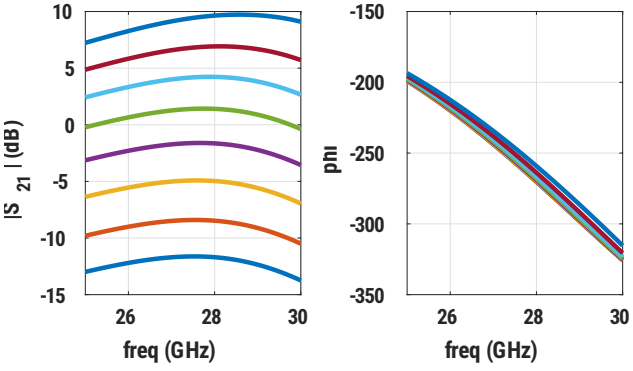


Fig. 11: (a) D-VGLNA power gain magnitude, (b) D-VGLNA power gain phase

Noise figure reach a minimum value of 2.3 dB at 28 GHz when the power gain is maximum and a value of 7 dB when the power gain is equal to 0 dB.

Fig. 13 shows the variation of -1 dB compression point for different D-VGLNA gain settings including the output buffers simulated with harmonic balance. The simulation returns a 1dB CP between -5 and -6.5 dB.

$$F = 1+\left(1+\frac{\sum_{i=1}^{k} g_{\mathrm{m,str,i}}}{g_{\mathrm{m,cas,0}}+\sum_{i=k+1}^{N} g_{\mathrm{m,cas,i}}}\right)\left[\frac{\gamma}{R_S g_{\mathrm{m1}}(1+A)^2 T_{\mathrm{IN}}^2}+\frac{(1+g_{\mathrm{m1}}R_s(1+A)T_{\mathrm{IN}}^2)^2}{R_s g_{\mathrm{m1}}^2 r_{\mathrm{o5}}(1+A)^2 T_{\mathrm{IN}}^2}\right]+\frac{g_{\mathrm{m7}}\gamma}{R_S g_{\mathrm{m1}}^2 T_{\mathrm{IN}}^2}\left(\frac{1+g_{\mathrm{m1}}R_S(1+A)T_{\mathrm{IN}}^2}{1+A}\right)^2 \tag{19}$$

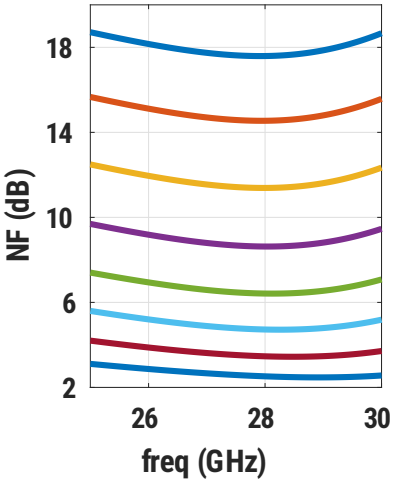


Fig. 12: D-VGLNA noise figure

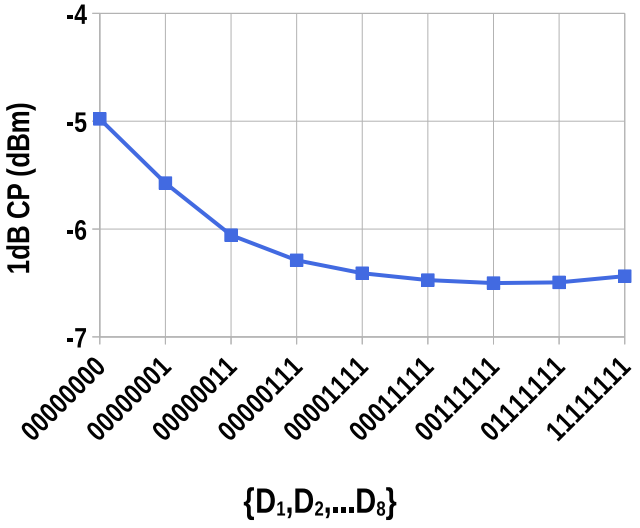


Fig. 13: 1 dB compression point

B. Layout implementation, hunts, and recommendations

Although post-layout simulation was not included, the presented figures of merit can still be achieved when the circuit is implemented in standard CMOS technology. Careful layout design is crucial to minimize additional parasitic capacitances at critical nodes and input parasitic resistance, as they can significantly degrade gain and noise figure at 28 GHz. In this context, the following suggestions and recommendations may be helpful for layout implementation while operating at 28 GHz.

- **The distance between the input signal pads and transistors M_1 and M_2 should be kept as short as possible. It is also highly recommended to use the top thick metal for the signal path to minimize input resistance, which could lead to NF degradation.**
- **Parasitic capacitances resulting from metal interconnections have to be accounted for in C_{p1} and C_{p2} estimation. This can be done effectively by extracting parasitic components with a dedicated tool after the layout has been completed. Thus, inductor geometry has to be adjusted accordingly to prevent any deviation in voltage gain from the desired value.**
- **Matching techniques, such as common centroid, are highly recommended for layout design of differential pair implementation, especially in current steering devices, due to the silicon's local gradient.**

  As mentioned in section III. A. maintaining a constant $Z_{\mathrm{in}}(s)$ and thereby stabilizing C_{p2} when the steering devices are activated and deactivated is essential for VGA operation. For instance, when the $M_{\mathrm{STR,i}}$ layout does not match its replica device, capacitance at node Y varies with changes in the polarity of D_{i}, as suggested by equation 17 . As a consequence, $Z_{\mathrm{in}}(s)$ deviates from the matching condition, resulting in the input return loss straying from its optimal value for different configurations of D_{i}. This leads to a spread of S_{11} in the vicinity of 28 GHz, rather than producing close curves as illustrated in Fig. 10.

V. Conclusion

In this paper, a low-noise variable gain amplifier is designed using a 65 nm CMOS standard process for a 28 GHz beam steering receiver. The VGA gain is set using digitally controlled current steering devices. The parasitic capacitances at the critical node are mitigated thanks to a set of digitally controlled replica devices, resulting in phase variation, input and output return loss insensitivity to gain setting, which can be performed in the digital domain. Thanks to this simple technique, the constant input and output return loss and the sub-3 dB NF demonstrate the VGA's capability to operate as an LNA, consequently reducing circuit complexity, cost, and power consumption for the overall beamformer receiver in the context of 5G mm-wave.

References

[1] A. Slimane, S. A. Tedjini, and F. Haddad, "Novel CMOS active inductor for tunable RF circuits," in *2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2018, pp. 917–920.

[2] A. Slimane, F. Haddad, S. Bourdel, S. A. Tedjini-Baïliche, M. T. Belaroussi, T. Mohamed, and H. Barthélemy, "Compact

inductorless CMOS low-noise amplifier for reconfigurable radio," *Electronics Letters*, vol. 50, no. 12, pp. 892–893, 2014. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/el.2014.1031

[3] S. A. Tedjini, A. Slimane, I. Ghorbel, F. Haddad, and W. Rahajandraibe, "Sub-mW cognitive CMOS frequency down-converter for internet of things," *IET Microwaves, Antennas & Propagation*, vol. 13, no. 12, pp. 2087–2095, 2019. [Online]. Available: https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-map.2018.5441

[4] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[5] J. Pang, R. Wu, Y. Wang, M. Dome, H. Kato, H. Huang, A. Tharayil Narayanan, H. Liu, B. Liu, T. Nakamura, T. Fujimura, M. Kawabuchi, R. Kubozoe, T. Miura, D. Matsumoto, Z. Li, N. Oshima, K. Motoi, S. Hori, K. Kunihiro, T. Kaneko, A. Shirane, and K. Okada, "A 28-GHz CMOS Phased-Array Transceiver Based on LO Phase-Shifting Architecture With Gain Invariant Phase Tuning for 5G New Radio," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 5, pp. 1228–1242, 2019.

[6] J. Pang, Z. Li, R. Kubozoe, X. Luo, R. Wu, Y. Wang, D. You, A. A. Fadila, R. Saengchan, T. Nakamura, J. Alvin, D. Matsumoto, B. Liu, A. T. Narayanan, J. Qiu, H. Liu, Z. Sun, H. Huang, K. K. Tokgoz, K. Motoi, N. Oshima, S. Hori, K. Kunihiro, T. Kaneko, A. Shirane, and K. Okada, "A 28-GHz CMOS Phased-Array Beamformer Utilizing Neutralized Bi-Directional Technique Supporting Dual-Polarized MIMO for 5G NR," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 9, pp. 2371–2386, 2020.

[7] J. Paramesh, R. Bishop, K. Soumyanath, and D. Allstot, "A four-antenna receiver in 90-nm CMOS for beamforming and spatial diversity," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, pp. 2515–2524, 2005.

[8] A. Natarajan, S. K. Reynolds, M.-D. Tsai, S. T. Nicolson, J.-H. C. Zhan, D. G. Kam, D. Liu, Y.-L. O. Huang, A. Valdes-Garcia, and B. A. Floyd, "A Fully-Integrated 16-Element Phased-Array Receiver in SiGe BiCMOS for 60-GHz Communications," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 5, pp. 1059–1075, 2011.

[9] W.-T. Li, Y.-C. Chiang, J.-H. Tsai, H.-Y. Yang, J.-H. Cheng, and T.-W. Huang, "60-GHz 5-bit Phase Shifter With Integrated VGA Phase-Error Compensation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 3, pp. 1224–1235, 2013.

[10] J. Xiao, I. Mehr, and J. Silva-Martinez, "A High Dynamic Range CMOS Variable Gain Amplifier for Mobile DTV Tuner," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 2, pp. 292–301, 2007.

[11] C.-H. Wu, C.-S. Liu, and S.-L. Liu, "A 2 GHz CMOS variable-gain amplifier with 50 dB linear-in-magnitude controlled gain range for 10GBase-LX4 Ethernet," in *2004 IEEE International Solid-State Circuits Conference (IEEE Cat. No.04CH37519)*, 2004, pp. 484–541 Vol.1.

[12] A. Natarajan, S. Nicolson, M.-D. Tsai, and B. Floyd, "A 60GHz variable-gain LNA in 65nm CMOS," in *2008 IEEE Asian Solid-State Circuits Conference*, 2008, pp. 117–120.

[13] B. Gilbert, "A Low-noise Wideband Variable-gain Amplifier Using An Interpolated Ladder Attenuator," in *1991 IEEE International Solid-State Circuits Conference. Digest of Technical Papers*, 1991, pp. 280–281.

[14] S.-A. Tedjini-Bailiche, M. Trabelsi, A. Slimane, M.-T. Belaroussi, F. Haddad, and S. Bourdel, "Ultra Low Power and High Gain Switched CMOS gm-Boosted Current Reused Mixer for Wireless Multi-Standard Applications," *Microelectronics Journal*, vol. 45, no. 11, pp. 1575–1582, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0026269214002833

[15] S. A. Tedjini, A. Slimane, M. T. Belaroussi, and M. Trabelsi, "A 0.9 V High Gain and High Linear Bleeding CMOS Mixer for Wireless Applications," in *2012 24th International Conference on Microelectronics (ICM)*, 2012, pp. 1–4.

Sid Ahmed Tedjini received a B.S, and M.S degrees in electrical engineering from Université des Sciences et de la Technologie Houari Boumediene, Algiers, Algeria in 2005, and 2008 respectively, and the Ph.D. degree from the Ecole Nationale Polytechnique, Algiers, Algeria, in 2016. In 2011, he joined the Microelectronics division at Centre de Développement des Technologies Avancées where he is currently a research fellow with Analog and mixed-signal IC design group. His current research interests include analog and RF IC design for wireless communications.

Sonia Salhi received a B.S, M.S, and PhD degree from Université des Sciences et de la Technologie Houari Boumediene in 2010, 2012, and 2018, respectively. In 2019, she joined Centre de Développement des Technologies Avancées where she is currently a research fellow with Analog and mixed-signal IC design group.

2716-912X © 2025 Ecole Nationale Polytechnique

A smooth gain scheduling generalized predictive control

Yassine Himour, Mohamed Tadjine, and Mohamed-Seghir Boucherit

Abstract− Nonlinear model predictive control is an emerging control technique dealing with high nonlinearities of industrial plants. However, it suffers from many hurdles such as the numerical problems related to the resulting no convex nonlinear optimization problem, time consuming, and difficulties in analyzing properties such as stability. In this paper, to side-step these difficulties, an infinite gain scheduling generalized predictive control is designed to control a benchmark high nonlinear plant instead of nonlinear predictive control. A neural model of the plant is identified and used as an internal model of the generalized predictive control scheme. The neural model is linearized successively and a filtering process is used to smooth the adaptation of the linearized model every sample time. The results show good performance in tracking the reference and rejecting abrupt changes in measured disturbances. The filtering process improved the results in terms of rapidity, overshoots damping, and smoothing the control signal.

***Keywords*− GPC, neural networks, gain scheduling, measured disturbances, high nonlinearities.**

NOMENCLATURE

ANN	Artificial neural network
ARX	Autoregressive exogenous inputs
CARIMA	Autoregressive integrated moving average
GPC	Generalized predictive control
HTF	Heat transfer fluid
MPC	Model predictive control
NLP	Nonlinear problem
NMPC	Nonlinear predictive control
PTC	Parabolic trough collector

I. INTRODUCTION

Model predictive control is the most used advanced control technique in industry. This is because it provides many advantages, among which constraints handling. Additionally, the principle of MPC is intuitive and its tuning is relatively straightforward— even practitioners with limited knowledge have the chance to work with. MPC can handle a panoply of processes that encompasses those with complex dynamics such as high nonlinearities, long delays, unstable characteristics and nonminimum phase systems[1]. MPC can be easily extended to multivariable cases, and it naturally includes feedforward action that allows it to exploit measured disturbances. For systems with high nonlinear dynamics, it's obvious that nonlinear predictive control is the first MPC technique to think to use with. However, NMPC relies on nonlinear optimization which shows many hurdles, especially in real time implementation. Among difficulties NMPC suffers from difficulty of theoretical analysis of properties such as stability[2]. Most solutions in this situation add more complexities and many computational problems may show up, for example the nonlinear optimization algorithm could not converge to the desired accuracy and/or can't converge in the needed time, especially in the case of fast systems. Moreover, numerical problems could occur and the obtained solution is not verified to be a global optimal solution. One interesting alternative to bypass those NMPC difficulties while treating high nonlinearities in a good way, is adaptive linear MPC, but again, adaptive MPC needs extra mechanisms to avoid adaptation problems such as persistent excitations which are crucial to obtaining good results. In such situation, scheduled gain MPC strategy has demonstrated better performance [3].

The gain scheduling MPC control strategy uses linearized models of the nonlinear process to predict the outputs over a prediction horizon, this results in poor modelling performance when using long prediction horizon, this is because the linearized models are supposed to be used near the operation points of their linearization. Some ideas have been appeared to cope with this problem. For example in [4], this is treated by interpolating the controller gains between four operating points chosen on the basis of the manipulated variable. Inspired by this idea, the authors in [5] used an interpolation of the linearized models' parameters.

This study aims to develop an MPC control strategy, namely, an infinite neural gain scheduling generalized predictive control, to overcome the aforementioned challenges. The remainder of this paper is organized as follows: the next section explains the control strategy, and gives details about its elements, namely, the generalized predictive control and the neural model and its linearization. In section III, the introduced control strategy is applied to a highly nonlinear process, the output temperature in the ACUREX parabolic trough collector field.

II. CONTROL STRATEGY AND DESIGN

The strategy is to use an artificial neural network model to reduce the computations[6]and to assure a maximum performance and capture of the plant nonlinear dynamics. In

Manuscript received March 6, 2025 ; June 19, 2025 revised.

Y.HIMOUR, M. TADJINE and M.S. BOUCHERIT are with Process Control Laboratory at EcoleNationale Polytechnique, Algiers, Algeria. (e-mail: yassine.himour@g.enp.edu.dz, mohamed.tadjine@g.enp.edu.dz, ms.boucherit@g.enp.edu.dz)

Digital Object Identifier (DOI): 10.53907/enpesj.v5i1.326

2716-912X © 2025 Ecole Nationale Polytechnique

every sampling time, the neural model is linearized around the actual operating point, and a linear GPC control law is applied using the linearized model. This strategy avoids NLPs and results in more precise local linear models than considering only some operating regimes.

A. *Generalized predictive control*

The generalized predictive control is an MPC control technique based oncontrolled autoregressive integrated moving average model(CARIMA model):

$$A(q^{-1})y(t) = q^{-d_u}B(q^{-1})u(t-1) + \frac{C(q^{-1})\xi}{\Delta} \quad (1)$$

Where, q^{-1} is the time backward shift operator, the differencing operator: $\Delta = (1-q^{-1})$, $y(t)$ and $u(t)$ are the output and the input of the model, d_u is the dead time according to the input $u(t)$. And $\xi(t)$ is a zero mean white noise.$A(q^{-1})$and$B(q^{-1})$ are the model polynomial with $A(q^{-1})$is monic.

In GPC the cost function is usually defined by:

$$J = \frac{1}{2}\sum_{i=N_1}^{N_2}(r(k+i) - \hat{y}(k+i|k))^2 + \frac{1}{2}\sum_{i=1}^{N_u}\lambda\left(\Delta u(k+i-1)\right)^2 \quad (2)$$

Where, $r(k+i)$is the reference signal, $\hat{y}(k+i|k)$is the predictions at time $k+i$ calculated at time k,$\Delta u(k)$ is the control increments, $\lambda$is the penalty factor of the control increments, $N_1$and$N_2$are the first and second prediction horizons, and N_u is the control horizon.

When there is no constraints, the GPC control can be given by the following analytical expression[7]:

$$\Delta U = [G^T G + \lambda I]^{-1} G^T (R - \phi) \quad (3)$$

Where, ΔU is the vector of the control increments over the control horizon, G is a matrix whose elements are the step response coefficients[7],ϕ is the free response depending on past and present signals, and R is the known reference over the control prediction horizon.

Every sampling time, only the first element of ΔU is applied to the real system, and the calculations are repeated for the next sample where new measurements are obtained. The control horizon N_u is commonly chosen less than the prediction horizon, hence the elements of ΔUafter the$N_u$element are taken equal to zero, this reduces the computations, particularly those of the inverse matrix whose dimensions become $N_u \times N_u$.

B. *Neural network modeling and linearization*

In[8] it was demonstrated that with a one hidden layer, an ANN is able to model any measurable function to any desired degree of accuracy. On the basis of that result, in this work, a two-layer perceptron is used; its output is given as follows:

$$\hat{y}(t) = \sum_{j=1}^{p} W_j \tanh\left(\sum_{i=1}^{l} w_{ji} z_i(t) + w_{j0}\right) + W_0 \quad (4)$$

Where, W_j *and* w_{ji} are the weights of the hidden to output layer and the input to hidden layer respectively. W_0 and w_{j0} are the biases of the nodes. p is the number of the hidden neurones, and l is the length of the input vector $z(t)$.

As mentioned before, this neural model is linearized every sampling time to get a linear CARIMA model appropriate to be used as an internal model in the GPC control law.

Let the whole function in(4) be expressed by the shorter expression:

$$y(t) = g(z(t)) \quad (5)$$

Where $z(t)$ is the input vector of the neural network:

$$z(t) = [y(t-1)\dots y(t-n)\, u(t-d_u-1)\dots u(t-d_u-1-m)$$

Linearizing the ANN model g around an operating point $z(\tau)$ gives:

$$\tilde{y}(t) = -a_1\tilde{y}(t-1) - \dots - a_n\tilde{y}(t-n) + b_0\tilde{u} \quad (6)$$

With,

$$\tilde{y}(t-i) = y(t-i) - y(\tau-i)$$

$$\tilde{u}(t-i) = u(t-i) - u(\tau-i)$$

And

$$a_i = -\left.\frac{\partial g(z(t))}{\partial y(t-i)}\right|_{z(t)=z(\tau)}$$

$$b_i = -\left.\frac{\partial g(z(t))}{\partial u(t-d_u-i)}\right|_{z(t)=z(\tau)}$$

Putting the elements of the current instant τ in(6) on one side gives:

$$y(t) = \left(1 - A(q^{-1})\right)y(t) + q^{-d_u}B(q^{-1})u(t-1) + \zeta(\tau)$$

Or

$$A(q^{-1})y(t) = q^{-d_u}B(q^{-1})u(t-1) + \zeta(\tau) \quad (7)$$

Taking $C(q^{-1})$ in(1) equals to 1, and modelling $\zeta(\tau)$ as an integrated white noise makes this model identical to the model used in the MPC control law.

C. *Smoothing the adaptation of the linearized model's parameters*

As mentioned earlier, using a linearized model, which is supposed to give good predictions just near the operating point of linearization, to predict the output of the process over long prediction horizon can result in poor performance of the model. For this reason, it is proposed here to use a filtering parameter to smooth the adaptation of the linear model's parameters. The filter expression is given by[9]:

$$\hat{P}_k(q^{-1}) = (1-d)P_k(q^{-1}) + d\,\hat{P}_{k-1} \quad (8)$$

Where $P_k(q^{-1})$ is every polynomial of the model in (7), k is the actual sample time, and d is the introduced tuning parameter.

This filtering process has the effect of smoothing the variation of the parameters of the successively linearized models.

Increasing the parameter d slows the adaptation of the model and vice versa. The filtering process would improve the performance of the gain-scheduled GPC controller when there are disturbances or abrupt changes in the reference. In the absence of these two factors, the filter would have no noticeable effect.

III. CASE STUDY

In this section, the strategy developed above is applied to the ACUREX solar parabolic trough field.

A. *Plant description and modelling*

The ACUREX plant is a PTC based distributed solar collector field. It is located in the desert of Tabernas, in southern of Spain. It is an experimental facility that has a maximum energy of 0.5 MW. The concentrators are aligned east-west in 10 parallel loops, every loop is composed of two rows. The length of one loop is 172 m, and only 142 m of the tube per loop receive the concentrated sun rays. The heat transfer fluid (HTF) used in the ACUREX field decompose for temperatures bigger than 300°C. The HTFflows in the absorber tube through the field area to collect solar energy in the form of heat.

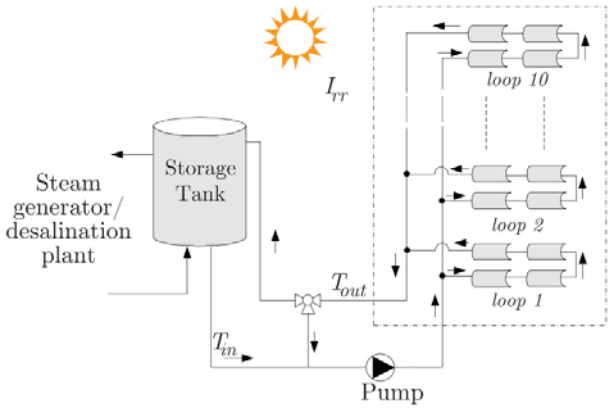


Fig 1: ACUREX Distributed Solar Field Schematics[10]

The HTF flow is limited between 2.0 and 12.0 l / s[11], or 0.2 and 1.2 l/s for every loop. This is to ensure safety especial to avoid the risk of the HTF decomposition.

he following hypothesis are considered in elaborating the model used in the present work[12]:

1. The properties of the heat transfer fluid are functions of the actual temperature value, and the flowequals to its average value.
2. Tube wall temperature variations are not taken into account.
3. Losses caused by the conduction of axial heat on both sides of the wall are negligible.

The enthalpy accumulated in the Δl element, between two instants time $t$and$t + \Delta t$ is given by the thermodynamics law:

$$E = \rho_f c_f A_f \, \Delta l \, [T(l, t+\Delta t) - T(l,t)] \quad (9)$$

Where,$l$is the space variable (m), A_f is thecross-sectional area of the tube (m^2), $\rho_f$is thefluid mass density (Kg/m^3), $c_f$is the fluid specific heat capacity (J /Kg °C), and$T(l,t)$ is the fluid temperature (°C).

This enthalpy is the sum of two enthalpies. The first is the difference between the enthalpy entering and the enthalpy leaving the Δl element due to fluid flow in Δt time:

$$E_1 = \rho_f c_f \bar{u}(t) \Delta t \, (T(l,t) - T(l+\Delta l, t)) \quad (10)$$

Where $\bar{u}(t)$ is the fluid volumetric flow rate (m^3 /s).

The second is due to the harvested solar energy between the two points $l$andΔl in Δt time:

$$E_2 = \eta \, G \, R \, \Delta l \, \Delta t \quad (11)$$

Where, R is the solar irradiance (W/m^2), η is the optical efficiency (Unit-less), and$G$ is the concentrator aperture (m).

By putting $E = E_1 + E_2$, we obtain the expression of the temperature evolution in every element $i\,\Delta l$:

$$\frac{\partial T(i\,\Delta l, t)}{\partial t} = -\frac{u(t)\left(T(i\,\Delta l, t) - T\left((i-1)\Delta l,\right.\right.}{\Delta l} \quad (12)$$

$$i = 1, \ldots n, \text{ and } \; n \times \Delta l = L$$

L is the total length of the absorber tube (m), $u(t) = \frac{\dot{m}(t)}{A_f}$, and

$$\alpha \triangleq \frac{\eta G}{\rho_f c_f A_f}$$

Define now,$T(i\Delta l, t) \overset{\text{def}}{=} x_i(t), that$ gives $x_0(t) = T(0, t))$.

And noting $\alpha_i = \alpha(x_i)$. Then, the set of differential equations defining the temperature process in a PTC field is:

$$\begin{aligned}
\dot{x_1}(t) &= -\frac{u(t)}{\Delta l}\left(x_1(t) - x_0(t)\right) + \alpha(x_1(t))R(t \\
\dot{x_2}(t) &= -\frac{u(t)}{\Delta l}\left(x_2(t) - x_1(t)\right) + \alpha(x_2(t))R(t \\
\dot{x_3}(t) &= -\frac{u(t)}{\Delta l}(x_3(t) - x_2(t)) + \alpha(x_3(t))R(t \\
&\vdots \quad \vdots \quad \vdots \quad \vdots \\
\dot{x_n}(t) &= -\frac{u(t)}{\Delta l}(x_n(t) - x_{n-1}(t)) + \alpha(x_n(t))R
\end{aligned} \quad (13)$$

B. Neural identification of the ACUREX plant

In the first principal model in (13),the solar irradiance $R(t)$ and the inlet temperature $x_0(t)$(or T_{in}: the HTF inlet temperature) are unmanipulated inputs, i.e., they have a direct effect on the output $x_n(t)$ (or T_{out}: the HTF outlet temperature) but are not commands, they are measured disturbances. However, in this work it's the model (13) which is being identified by the neural network but not the real plant. This gives the possibility to use the irradiance and the inlet temperature with the HTF flow as excitation signals in the identification task.

Fig. 2 depicts the curves of the used excitation signals over range of 12 hours. The irradiance signal is shaped likewise to imitate real solar irradiance during a day. The database is divided into two datasets, one for training and the other for validation. The total length is 18000 samples, with a time sampling of 15s.

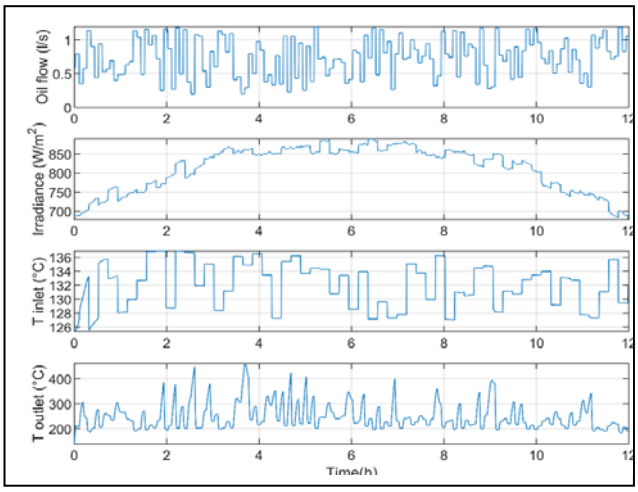


Fig. 2: The identification database[9]

Trial and error tests were held to find the convenient parameters of the ANN. The results in Fig. 3 and Fig. 4 were obtained for an ARX structure with 12 neurons in the hidden layer, one sample dead time for the three inputs, 5 order of the past delayed output, 4 for the control flow, and 3 for both the irradiance and the inlet temperature.

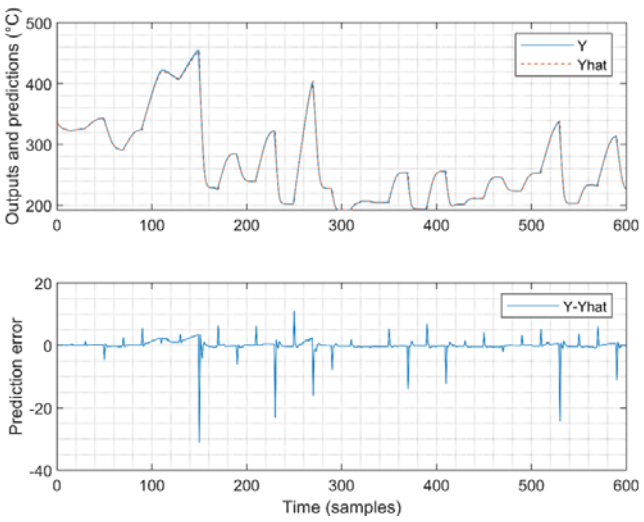


Fig. 3: The output, the prediction and error

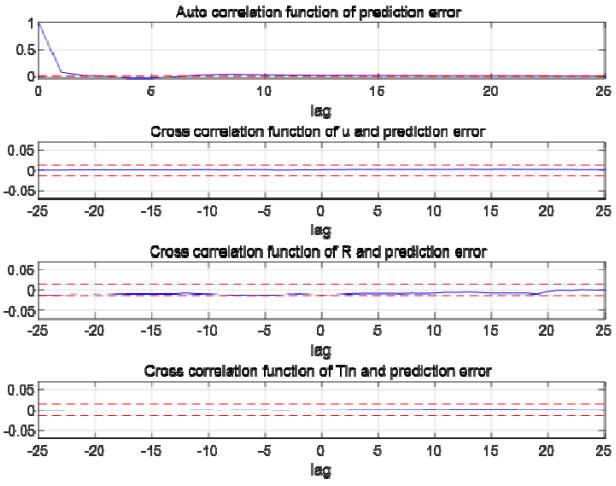


Fig. 4: Cross correl3ations of the inputs and the prediction error

It's to notice that the prediction error peaks in second subplot of Fig. 4 is due to big instantaneous changes of the output temperature, which is not usual in the real plant. The autocorrelation of the error and its cross-correlation with the

inputs is null, and this is a good index of the good performance of the identified model.

C. *Application of the control strategy*

The control strategy introduced in section II is applied here using the identified neural model. Fig. 5 shows the profile of the measured disturbances, i.e., the solar irradiance and the inlet HTF temperature. Two aggressive changes were added on the two signals to simulate two disturbances. Artificial estimations are made by adding made-up estimation errors, and are to be used in the GPC prediction horizon instead of the real signal.

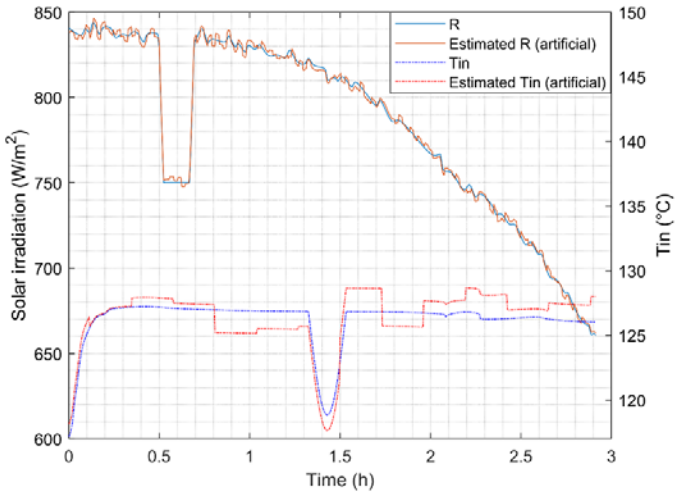


Fig. 5: The measured disturbances profile[9]

To test the robustness of the control law, model mismatch has been added to the first principal model which is used as a simulator of the real plant, and the reference signal was chosen to cover a wide range of the plant dynamics. Fig. 6 and Fig. 11 illustrate the output and the command behaviors respectively for different values of the smoothing parameter d.

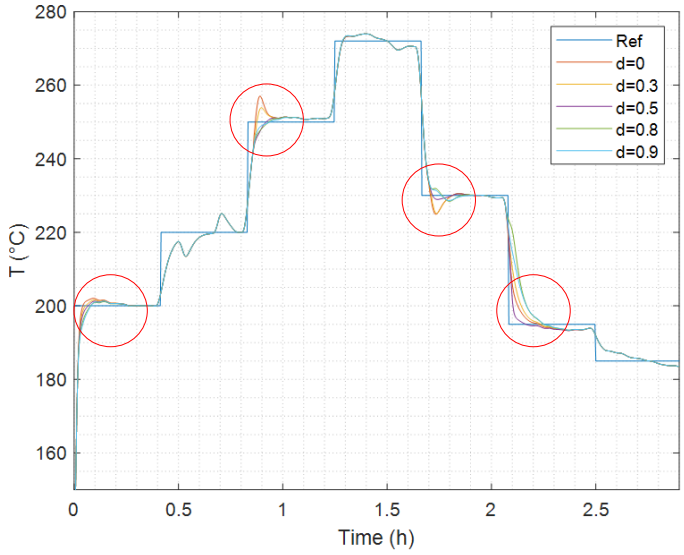


Fig. 6: Output versus the reference for different values of the smoothing parameter

Fig. 6 shows the performance of the control scheme in rejecting the two disturbances injected on the solar irradiance and the HTF inlet temperature, and the performance of tracking the reference. Fig. 7, Fig. 8, Fig 9 and Fig. 10, bellow, are a zoom into the red circles in Fig. 6, to show better the effect of the parameter d. The figures are ordered corresponding to the circles from left to the right, i.e., the circle most to the left is in Fig. 7, the next to it is in Fig. 8, … etc.

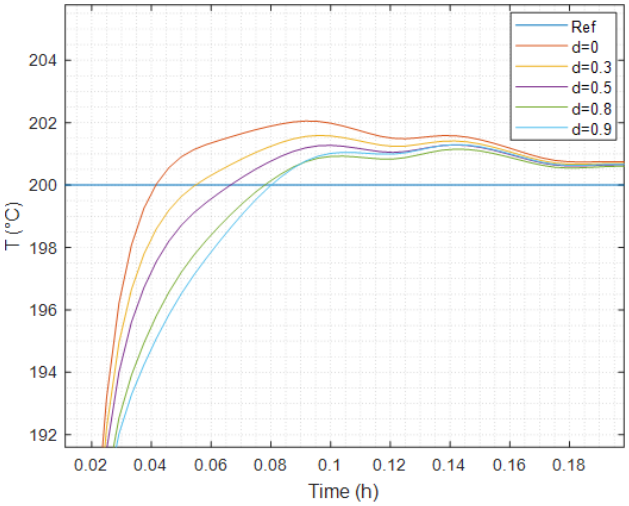


Fig. 7: Zoom on the first red circle from the left in Fig. 6

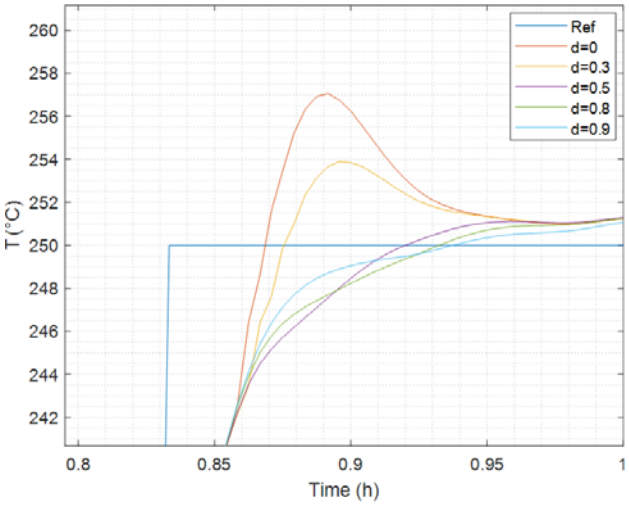


Fig. 8: Zoom on the second red circle from the left in Fig. 6

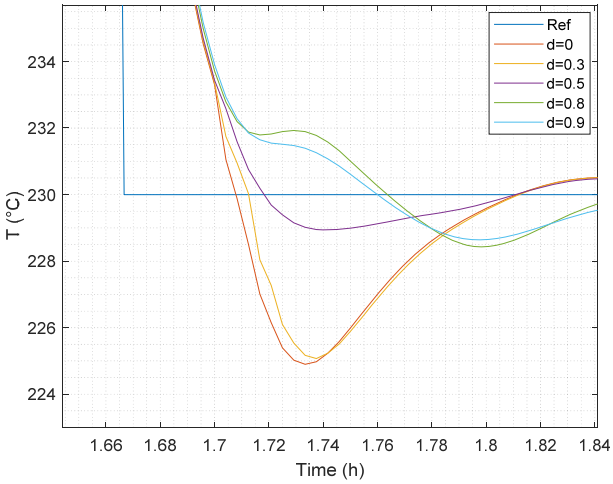


Fig. 9: Zoom on the third red circle from the left in Fig. 6

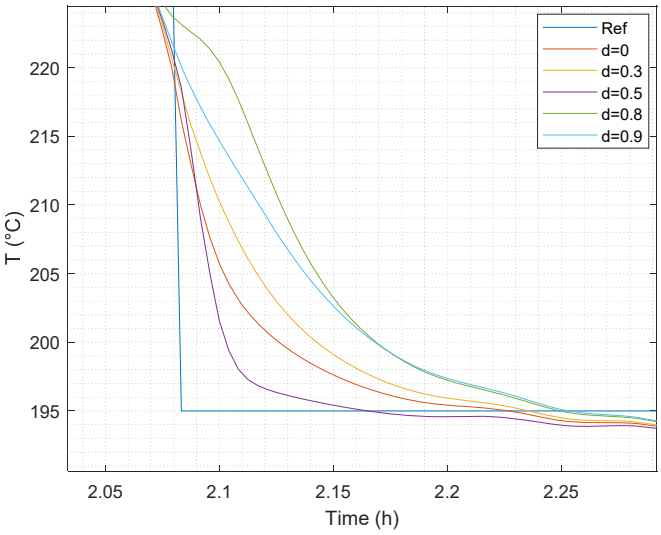


Fig. 10: Zoom on the fourth red circle from the left in Fig. 6

These figures clearly depict the improvement brought by the introduction of the filtering process of the polynomials of the successively linearized models. As it was expected, the filtering has no effect when the plant is in steady state, i.e., when there is no abrupt change of the reference. The best value of d in this simulation is 0.5, using this value eliminates the overshoots of the case without filtering, see Fig. 8 and Fig. 9, and also improves rapidity, see Fig. 10. The filtering has also good effect on the command, see Fig.11, It smooths the control signal, thus it helps reducing the control effort and increasing the lifetime of the pump.

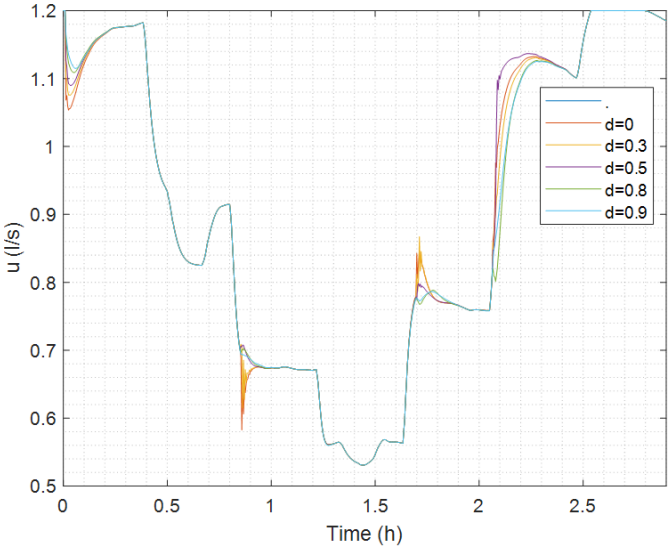


Fig. 11: Flow rate for different values of the smoothing parameter

IV. Conclusion

In this work, a gain scheduling control strategy based on GPC with real time linearization of a neural internal model was introduced and implemented. To improve the performance of the designed controller, a filtering process is used to smooth the adaptation of the linearized model polynomials. The whole strategy is applied to a highly nonlinear benchmark system: the ACUREX solar parabolic trough collector plant. The results showed satisfying results in rejecting the effects of abrupt changes of the two measured disturbance and in reference tracking. After applying the smoothing process, improvements are introduced in terms of overshoots damping, rapidity increasing and smoothing the control signal. Finally, as an idea of a future work , d could be updated according to the variation of the reference and the measured disturbances.

References

[1] E. F. Camacho and C. Bordons, *Model Predictive control*, no. 9781852336943. in Advanced Textbooks in Control and Signal Processing. London: Springer London, 2007. doi: 10.1007/978-0-85729-398-5.

[2] E. F. Camacho and A. J. Gallego, “Model Predictive Control in Solar Trough Plants: A Review,” *IFAC-PapersOnLine*, vol. 48, no. 23, pp. 278–285, 2015, doi: 10.1016/j.ifacol.2015.11.296.

[3] C. M. Cirre, M. Berenguel, L. Valenzuela, and E. F. Camacho, “Feedback linearization control for a distributed solar collector field,” *Control Eng. Pract.*, vol. 15, no. 12, pp. 1533–1544, Dec. 2007, doi: 10.1016/j.conengprac.2007.03.002.

[4] E. F. Camacho, M. Berenguel, and F. R. Rubio, “Application of a gain scheduling generalized predictive controller to a solar power plant,” *Control Eng. Pract.*, vol. 2, no. 2, pp. 227–238, Apr. 1994, doi: 10.1016/0967-0661(94)90202-X.

[5] A. J. Gallego, G. M. Merello, M. Berenguel, and E. F. Camacho, “Gain-scheduling model predictive control of a Fresnel collector field,” *Control Eng. Pract.*, vol. 82, pp. 1–13, Jan. 2019, doi: 10.1016/j.conengprac.2018.09.022.

[6] E. Masero, S. Ruiz-Moreno, J. R. D. Frejo, J. M. Maestre, and E. F. Camacho, “A fast implementation of coalitional model predictive controllers based on machine learning: Application to solar power plants,” *Eng. Appl. Artif. Intell.*, vol. 118, Feb. 2023, doi: 10.1016/j.engappai.2022.105666.

[7] D. W. Clarke, C. Mohtadi, and P. S. Tuffs, “Generalized predictive control-Part I. The basic algorithm,” *Automatica*, vol. 23, no. 2, pp. 137–148, Mar. 1987, doi: 10.1016/0005-1098(87)90087-2.

[8] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural Networks*, vol. 2, no. 5, pp. 359–366, Jan. 1989, doi: 10.1016/0893-6080(89)90020-8.

[9] Y. Himour, M. Tadjine, and M. S. Boucherit, “Nonlinear and infinite gain scheduling neural predictive control of the outlet temperature in a parabolic trough solar field: A comparative study,” *Eng. Appl. Artif. Intell.*, vol. 126, p. 106862, Nov. 2023, doi: 10.1016/j.engappai.2023.106862.

[10] P. Gil, J. Henriques, A. Cardoso, P. Carvalho, and A. Dourado, “Affine neural network-based predictive control applied to a distributed solar collector field,” *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 2, pp. 585–596, 2014, doi: 10.1109/TCST.2013.2260545.

[11] E. F. Camacho, M. Berenguel, and F. R. Rubio, *Advanced Control of Solar Plants*. in Advances in Industrial Control. London: Springer London, 1997. doi: 10.1007/978-1-4471-0981-5.

[12] E. F. Camacho, M. Berenguel, F. R. Rubio, and D. Martínez, *Control of Solar Energy Systems*, vol. 31, no. 5. in Advances in Industrial Control, vol. 31. London: Springer London, 2012. doi: 10.1007/978-0-85729-916-1.

Yassine Himour received Engineering anddegree in Control Engineering in 2004 from Skikda university, Algeria. He is actually a PhD student in Automatic Control inEcole Nationale Polytechnqiue (ENP) of Algiers, Algeria, and an assistant professor at Khemis Miliana university, Ain Defla, Algeria. His current research interests are control ofthermal solar systems, nonlinear predictive control.

Mohamed Tadjine received her PhD degree in Control and Manufacturing Management from ‘Institut National Polytechnique de Grenoble’ (INPG), (France) in 1994. In 1997, he joined the Ecole Nationale Polytechnique where he is currently Professor in Control Engineering Department. His main research areas are: Non Linear and Robust Control, Observer Design, Fault Detection and Fault Tolerant Control, Application in healthcare and biological systems.

Mohamed SeghirBoucherit was born in Algiers, Algeria. He is a Professor and University Research Professor in the Ecole Nationale Polytechnique. His fields of research are nonlinear control, power electronics, and variable-speed drives. He has contributed to several national and international journals and conferences for several years.

MPPT Efficiency in PV Arrays under Partial Shading Conditions:A Comparative Analysis of PSO and P&O Algorithms

Rachid Bennia, Cherif Larbes, and Faiza Belhachat

Abstract−Global Maximum Power Point Tracking (GMPPT) presents a fundamental challenge in photovoltaic (PV) systems due to the inherent nonlinearity of PV array characteristics. Partial shading (PS) emerges as a particularly critical factor, significantly compromising overall system efficiency by inducing multiple local maxima in the Power-Voltage (P-V) characteristic curve. While conventional tracking algorithms demonstrate adequate performance under uniform irradiation conditions, their effectiveness diminishes substantially under Partial Shading Conditions (PSCs), where they frequently converge to local power peaks rather than the true Global Maximum Power Point (GMPP). To address these limitations, intelligent computational approaches have been developed as robust alternatives for reliable GMPP tracking in complex shading environments. This investigation presents a comparative analysis of two established algorithms: Particle Swarm Optimization (PSO) and Perturb and Observe (P&O), evaluating their respective capabilities in GMPP identification. Extensive simulation studies conducted across diverse shading patterns conclusively establish the superior performance of the PSO algorithm, which consistently achieves steady-state tracking efficiencies exceeding 99% in all operational scenarios. These findings strongly suggest that PSO represents the more effective solution for optimal power extraction in PV systems operating under dynamic environmental conditions.

***Keywords*−Maximum Power Point Tracking, Particle Swarm Optimization, Perturb and Observe, Partial Shading.**

NOMENCLATURE

GMPPT	Global Maximum power point tracking.
GMPP	Global Maximum power.
PV	Photovoltaic.
PSCs	Partial Shading Conditions.
P-V	Power-Voltage.
I-V	Current-Voltage.
PSO	Particle Swarm Optimization algorithm.
P&O	Perturb and Observe.
RES	Renewable Energy Sources.
SC	Soft Computing.
ANN	Artificial NeuralNetwork.
BA	Bat algorithm.

I. INTRODUCTION

The world is now experiencing an increasing need for energy, which requires the exploration of alternative energy resources to complement conventional ones[1]. Because they are year-round available and pollution-free, renewable energy sources (RES) are highly recommended [2]. One of the primary concerns for researchers today is optimizing the energy output of photovoltaic (PV) systems in different weather conditions.

Manuscript received December 11, 2024 ; June 21, 2025 revised.

R. Bennia, C. Larbes and F. Belhachat are with the Laboratory of Communication Devices and Photovoltaic Conversion Department of Electronics. Ecole Nationale Poytechnique, Algiers, ALGERIA (e-mails: rachid.bennia@g.enp.edu.dz, cherif.larbes@g.enp.edu.dz, faiza.belhachat@g.enp.edu.dz).

Digital Object Identifier (DOI): 10.53907/enpesj.v5i1.306

Several techniques are presented to monitor the maximum power point (MPP). Conventional maximum power point tracking (MPPT) methods can track the MPP under uniform irradiation. However, they fail under rapidly changing atmospheric conditions and PSCs [1,2]. These conventional MPPTs are unable to distinguish between global and local maxima, causing the system to be trapped at a local peak [3]. As a result, the system's performance is drastically reduced. Partial shading (PS) is a major problem that significantly influences the output power of the photovoltaic system [4]. The PV array characteristic exhibits numerous local power peaks and one global maximum due to the integration of bypass diodes to avoid the hot spots effect [5]. Soft computing (SC) methods are widely recommended as the potential solution to mitigate the concern of PS and non-uniform irradiation [6,7]. PS is a major problem that significantly decreases the overall system's efficiency [8], it is acknowledged as the primary source of energy losses in PV power systems [9,10]. Studies have indicated that PS can significantly reduce the PV system yield, ranging from 10% to 70%. To circumvent this issue, a significant interest is dedicated by researchers to the MPPT techniques based on soft computing that are considered as a potential solution to mitigate the concern of PS [11,12], particularly Particle Swarm Optimization (PSO) [13] and Cuckoo Search (CS) [14,15] which are powerful optimization techniques used to address various engineering optimization problems with several peaks and can handle effectively the partial shading issues.

II. MAXIMUM POWER POINT ALGORITHMS

SC techniques, including Artificial Neural Network (ANN) [16], Bat algorithm (BA) [17], Cuckoo search (CS) [18,19], and Particle Swarm Optimization (PSO) [20], are commonly used in PV systems [21, 22]. PSO algorithm is a prominent MPPT approach used to track GMPP under PSCs,this technique comes with several benefits owing to its high

2716-912X © 2025 EcoleNationalePolytechnique

performance, simplicity, and ease of implementation.

A. Particle Swarm Optimization Algorithm (PSO) Overview

PSO is an optimization approach inspired by the social behaviour of bird flocking, and it is designed to find the optimum solution in an n-dimensional space for an optimization issue. It is based on a swarm population called particles. Each particle follows two essential rules for finding the best solution: it follows the highest-performing particle and the best solution found by the particle itself. As a result, all particles in the swarm converge towards the optimum required solution. Particle positions and velocities are adjusted using equations (1) and (2).

$$v_i^{k+1} = w_i v_i^k + randc_1\left(pbest_i - d_i^k\right) + randc_2\left(gbest_i - d_i^k\right) \quad (1)$$

$$d_i^{k+1} = d_i^k + v_i^{k+1} \quad (2)$$

Where d_i and v_i are the particle's position and velocity, whereas w corresponds to the inertia weight and k is the iteration number, c_1 and c_2 are learning parameters, $p_{besti}$and g_{besti} are defined as the personal and global best positions, and *rand* refers to random number within the interval [0, 1].

B. Perturb and Observe (P&O)

The working principle of P&O [23] is as follows: the algorithm creates an alteration (Δ) in the operating voltage of the PV array, which results in a perturbation of the generated power. An increase in the operating power indicates that the converter is approaching the MPP. Hence, in the next sampling cycle, the same direction of alteration is kept, and the reference voltage/current is raised by the same amount (Δ). Note that the P&O algorithm will stay oscillating once the MPP is attained.

III. System Description

Fig.1 depicts the schematic diagram of the PV system. The system includes a photovoltaic array, a DC-DC boost converter, and a resistive load. The MPPT controller generates the required duty ratio to the boost converter to obtain the optimum power from the PV array. The PV model is simulated using SunPower SPR-X20-250 PV panels.

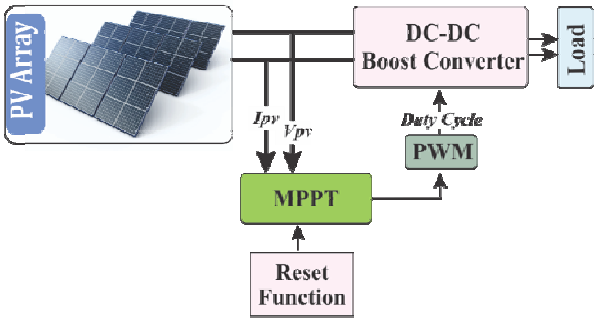


Fig. 1: Bloc diagram of a typical PV system

IV. Simulation and Discussion

The suggested approach is simulated using MATLAB/Simulink environment. The photovoltaic system is composed of four PV arrays connected in series, a DC-DC boost converter, two MPPT-based controllers (PSO and P&O), and a resistive load. The detailed profiles of the different scenarios of partial shading used in this study are illustrated in Table I whereas their corresponding (P-V) curves are shown in Fig. 2. The used PV module specifications are depicted in Table II. For optimal power efficiency, a boost converter is implemented to regulate and match the output of the PV array to the load. The converter works in continuous conduction mode. An extensive simulation analysis is performed under STC and PSCs.Several patterns with different levels of partial shadingare examined to verify the reliability of the proposed approach. For this study, the cell temperature is set constant at 25 °C for all circumstances.

Table. I
DETAILED PROFILES OF PARTIAL SHADING

Patterns	IRRADIANCE ON EACH PV PANEL (W/M^2)	GMPP (W)
STC	1000-1000-1000-1000	999.8
PSC1	1000-1000-500-200	490.5
PSC2	1000-800-500-300	411.5
PSC3	600-800-900-1000	660.0
PSC4	500-200-800-400	320.5

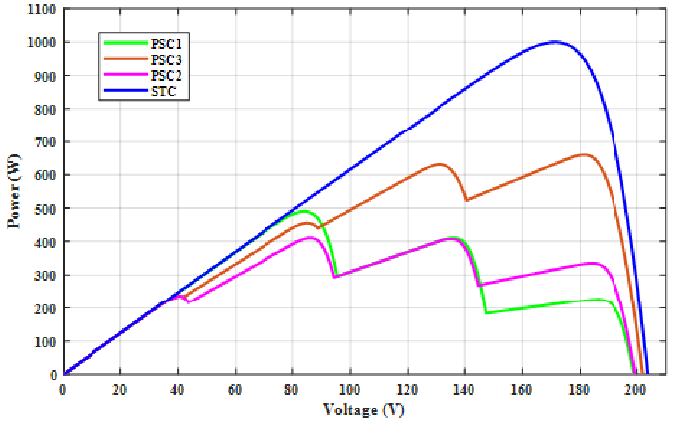


Fig. 2. Output characteristics for the simulated scenarios

Table. II
SUNPOWER SPR-X20-250 MODULE PARAMETERS

Characteristics	Value
Max-Power (W)	249.9
OC-Voltage (V)	50.93
V-$_{MPP}$ (V)	42.8
I_{mp} (A): Current at MPP	5.84
Isc (A): Short circuit current	6.2
N°. of cells	72

Dynamic and steady-stateefficiencies for each scenario are analyzed to evaluate the performance of the studied algorithms. The tracking efficiency [15]ηis defined as:

$$\eta = \frac{P_{MPPT}}{P_{max}} x\ 100\ (\%) \quad (3)$$

Where P_{MPPT}andP_{max} are, respectively, the MPPT steady-state power and the maximum delivered power by the PV system. The dynamic efficiency is defined by:

$$\zeta = \frac{\int p(t)dt}{\int P_{MPPT}(t)dt} x100\% \quad (4)$$

Where $p(t)$ and $P_{MPPT}(t)$ are the measured power and the maximum power on the (P-V) curve at time t.

A. Tracking under uniform irradiation

The suggested system is tested under standard test conditions, where all panels receive a uniform irradiation of 1000 (W/m^2). Fig. 3 and Fig. 4 illustrate the corresponding obtained results.

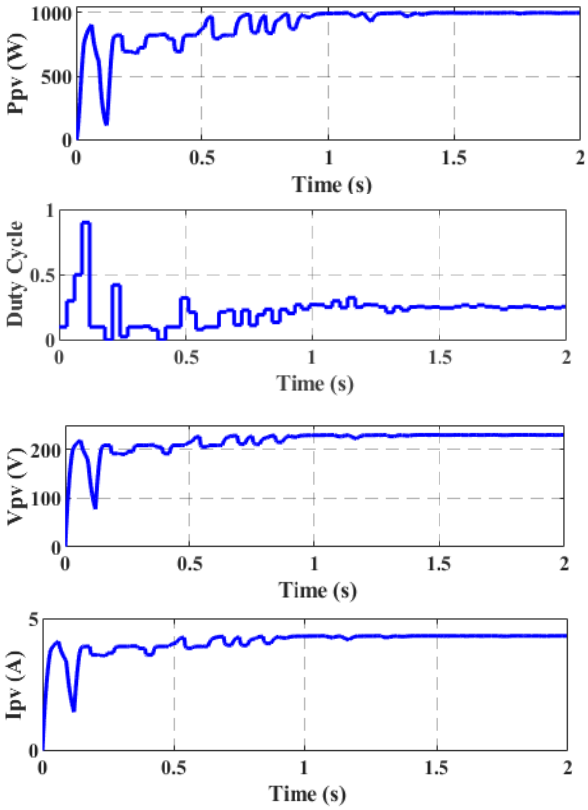


Fig. 3: Tracking performance under STC using the PSO algorithm

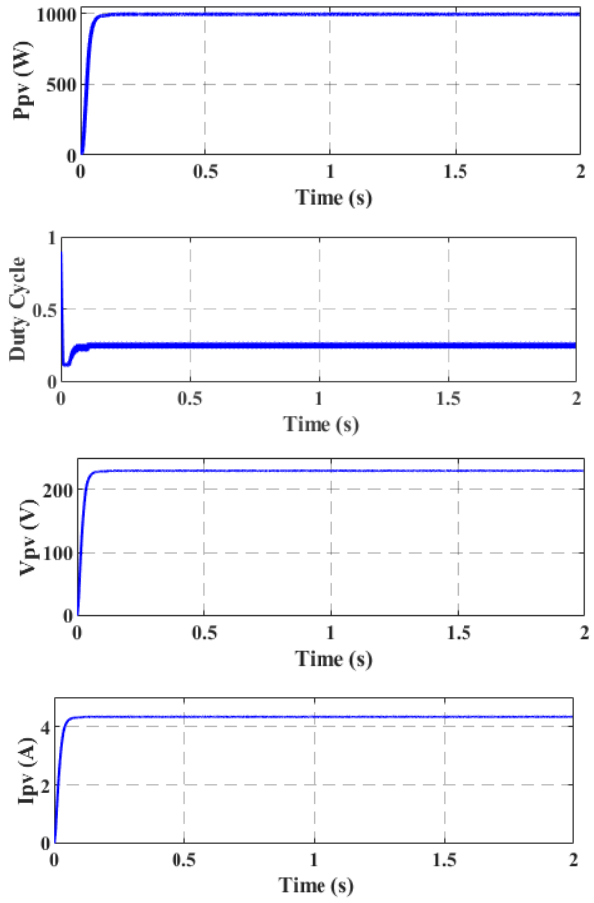


Fig. 4: Tracking performance under STC using the P&O algorithm

Fig. 3 and Fig. 4 illustrate that both algorithms can converge to the global maximum power point at 999.8 W. The P&O algorithm converges faster than PSO since, under this case, the PV characteristic shown in Fig. 2 is characterized by only one peak; in contrast, the PSO approach takes more time to converge due to the search process that covers the whole search space, however, it demonstrates higher steady state efficiency than the P&O algorithm.

B. Tracking under PSC1

The system is tested against the first partial shading scenario (PSC1) where the PV panels receive the first set of solar irradiations (1000-1000-500-200 W/m²) as shown in Table. I the corresponding output curves are illustrated in Fig.5. and Fig. 6.

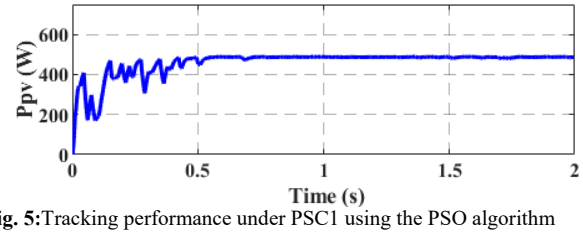


Fig. 5: Tracking performance under PSC1 using the PSO algorithm

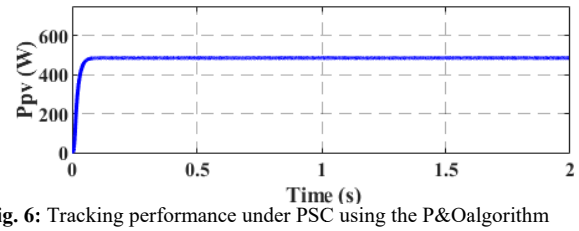


Fig. 6: Tracking performance under PSC using the P&Oalgorithm

As shown in Fig. 2, the PV output characteristic presents three peaks: 490.4 W, 411.7 W, and 226 W, where the first peak is the GMPP at 490.5 W. In this case, both algorithms also converge to the real GMPP with high efficiency more than 99 %. The P&O algorithm was able to recognize the global peak since it comes first in the PV output curve.

C. Tracking under PSC2

A more complicated shading pattern is presented to the system to verify its robustness under complex PSCs. As illustrated in Fig. 2, for the second case (PSC2), the PV output curve demonstrates four peaks; moreover, there is only a slight difference between the GMPP and the other local peaks. The obtained results are illustrated in Fig. 7 and Fig. 8.

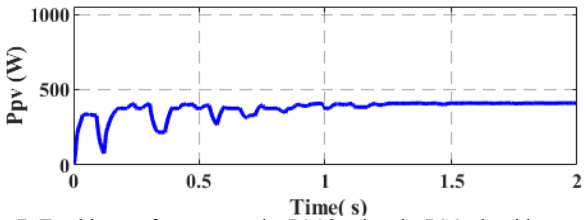


Fig.7: Tracking performance under PSC2 using the PSO algorithm

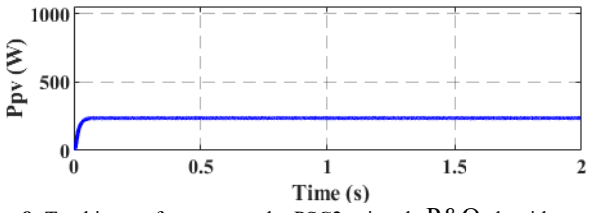


Fig. 8: Tracking performance under PSC2 using theP&O algorithm

This scenario is more complicated since the PV output characteristic presents four peaks, where the GMPP is at 411.5 W and is located at the second peak of the (P-V) output curve. As a result, the P&Oalgorithm was trapped in a local maximum; however, the PSO was able to escape this trap and converge to real GMMP despite there being another local maximum at 408.8 W in the vicinity of the GMPP.

D. Tracking under PSC3

As a third test, a set of solar irradiations is presented to the system. In this case, the PV output characteristic also shows

four peaks, where the GMPP is found at the last position of the curve, as shown in Fig. 2. The obtained resultsare illustrated in Fig. 9 and Fig. 10.

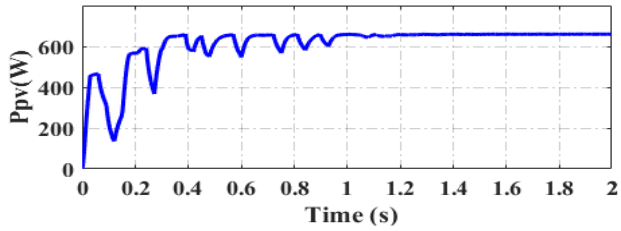


Fig. 9: Tracking performance under PSC3 using the PSO algorithm

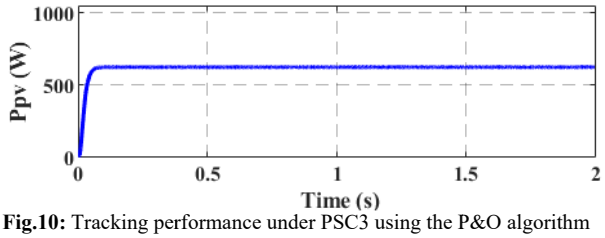


Fig.10: Tracking performance under PSC3 using the P&O algorithm

Both algorithms are tested against the third scenario to confirm their suitability to monitor partial shading situations. In this case, the PV output characteristic presents three local maxima and one global maximum at 660 W shifted to the end of the PV output characteristic. As illustrated in Fig. 9, the PSO algorithm converges effectively to the real GMPP, whereas the P&Ois trapped in a local peak 623 W.

E. Tracking under PSC4

To confirm the efficiency of these algorithms in monitoring the optimum power under various weather circumstances, the proposed system is tested against PSC4 (500 – 200- 800 – 400 W/M²) which simulates a real daily variations of solar irradiation, where the values are chosen to vary in both ascending and descending order. The corresponding P-V curve for this scenario is shownin Fig. 11whereas the dynamic responses are illustrated in Fig. 12 and Fig. 13.

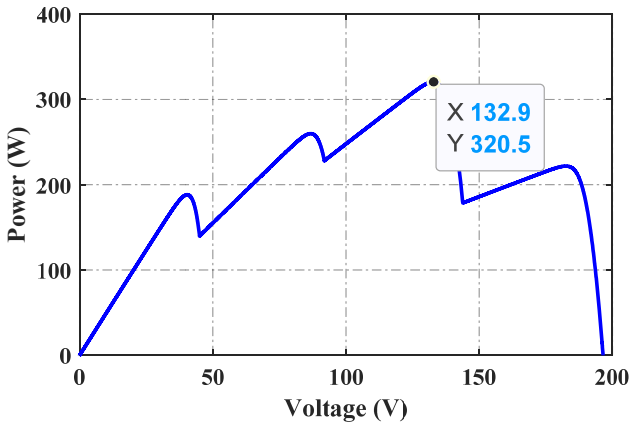


Fig. 11: Output characteristic under PSC4

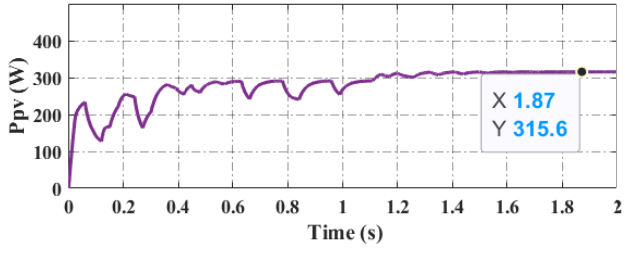


Fig. 12: Tracking performance under PSC4 using the PSO algorithm

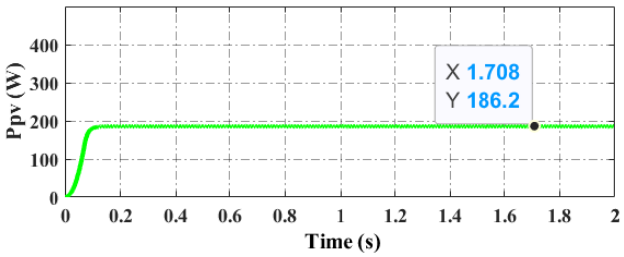


Fig. 13: Tracking performance under PSC4 using theP&O algorithm

The dynamic responses corresponding to this scenario illustrate the outperformance of the PSO under a real life change in irradiation. The PSO could converge to the GMPP. In contrast, the P&O algorithm was trapped in a local maximum leading to important power losses. Table. III illustrates the performance comparison of the two algorithms for the different scenarios.

Table. III
PERFORMANCE EVALUATION BETWEEN PSO AND P&O

Scena -rios	MPPT TECHNIQUE	GMPP (W)	Pmax(W)	Static Eff. (%)	Dynamic Eff.(%)
STC	PSO	999.6	999.8	99.98	90.3
	P&O	995.2		99.53	98.1
PSC1	PSO	488	490.5	99.49	94
	P&O	486.4		99.16	98
PSC2	PSO	408.5	411.5	99.27	90.91
	P&O	232.6		56.52	55.97
PSC3	PSO	658	660	99.69	96.89
	P&O	623		94.39	93.25

F. Tracking under rapid change in irradiance

To confirm the robustness of the proposed system in tracking the GMPP under a more challenging scenario, the system is tested against a rapid change in irradiance every three seconds. As shown in the dynamic response in Fig.14, Fig. 15, Fig. 16 and Fig. 17, the proposed algorithm successfully tracks the real power as the shading profiles change from PSC1 to PSC3. This change in irradiation is instantly detected by the reset function, which restarts the search process for the new GMPP.

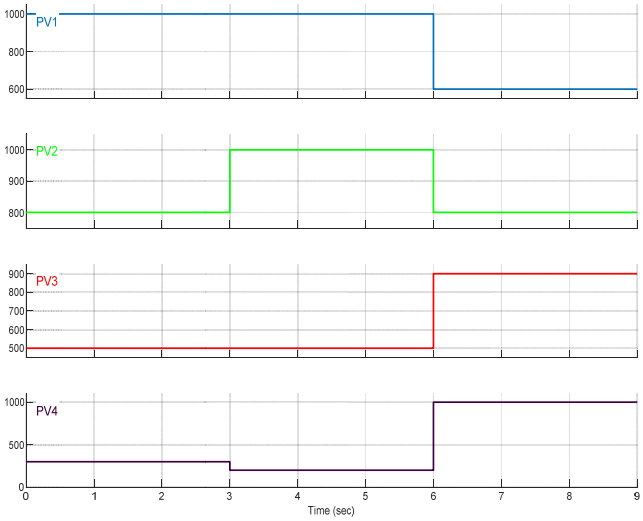


Fig. 14: Irradiance Scenarios

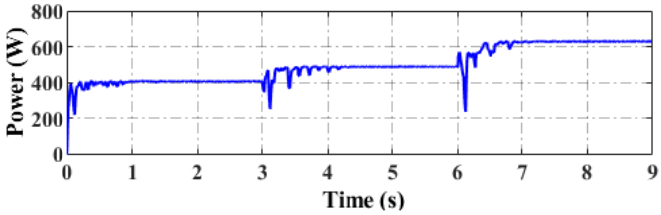


Fig. 15: Tracking performance under rapid change of irradiance

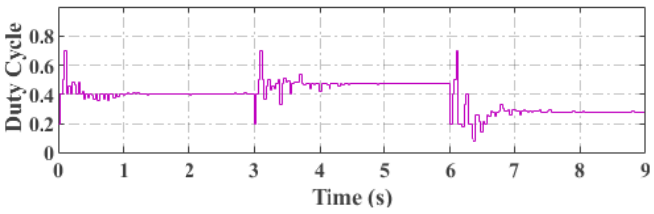


Fig. 16: Duty Cycle

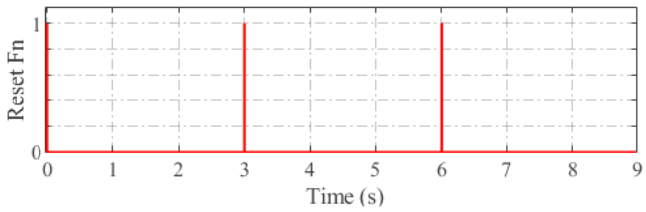


Fig. 17: Reset Function

As illustrated in the dynamic response under this complex scenario, the proposed system tracks successfully the real GMPP demonstrating an outstanding robustness and stability against complex dynamic weather circumstances. This performance is achieved since the system integrates a reset function that initiates the algorithm upon irradiation change detection.

V. Conclusion

To investigate the performance of maximum power point tracking (MPPT) under partial shading conditions (PSCs), two distinct optimization techniques are evaluated and compared: the conventional Perturb and Observe (P&O) method and a swarm intelligence-based approach, namely Particle Swarm Optimization (PSO). The comparative analysis is conducted under varying environmental conditions, including uniform irradiance, complex partial shading scenarios, and rapid irradiance changes. As evidenced by the results presented in Table III, the PSO algorithm exhibits superior performance in accurately identifying and tracking the GMPP, achieving a steady-state efficiency exceeding 99% across all test cases. This enhanced efficacy stems from PSO's global search mechanism, which systematically explores the entire solution space to avoid suboptimal convergence. In contrast, the P&O technique, relying on a localized search strategy, is prone to convergence at local maxima, as demonstrated in Scenario 2, where its efficiency declines to 56.52%, resulting in significant power losses. Consequently, the findings suggest that PSO is a more robust and reliable solution for GMPP tracking under both uniform and complex partial shading conditions, making it a preferable choice for photovoltaic systems operating in varying environmental conditions.

References

[1] N. A. Kamarzaman and C. W. Tan, « A comprehensive review of maximum power point tracking algorithms for photovoltaic systems », *Renewable and Sustainable Energy Reviews*, vol. 37, p. 585-598, sept. 2014, doi: 10.1016/j.rser.2014.05.045.

[2] A. O. Baba, G. Liu, and X. Chen, « Classification and Evaluation Review of Maximum Power Point Tracking Methods », *Sustainable Futures*, vol. 2, p. 100020, 2020, doi: 10.1016/j.sftr.2020.100020.

[3] L. G. K. Chai, L. Gopal, F. H. Juwono, C. W. R. Chiong, H.-C. Ling, and T. A. Basuki, « A novel global MPPT technique using improved PS-FW algorithm for PV system under partial shading conditions », *Energy Conversion and Management*, vol. 246, p. 114639, oct. 2021, doi: 10.1016/j.enconman.2021.114639.

[4] Belhachat and C. Larbes, "Modeling, analysis and comparison of solar photovoltaic array configurations under partial shading conditions," *Solar Energy*, vol. 120, pp. 399–418, Oct. 2015, doi: 10.1016/j.solener.2015.07.039.

[5] F. Belhachat and C. Larbes, « Comprehensive review on global maximum power point tracking techniques for PV systems subjected to partial shading conditions », *Solar Energy*, vol. 183, p. 476-500, mai 2019, doi: 10.1016/j.solener.2019.03.045.

[6] A. K. Podder, N. K. Roy, and H. R. Pota, "MPPT mandhods for solar PV systems: a critical review based on tracking nature," *IET Renewable Power Generation*, vol. 13, no. 10, pp. 1615–1632, Jul. 2019, doi: 10.1049/iet-rpg.2018.5946.

[7] R. Bennia, F. Belhachat and C. Larbes, « Smart hybrid algorithm for global maximum power point tracking under partial shading conditions », *Energy Sources, Part A: Recovery, Utilization and Environmental Effects*, vol. 46, n° 1, p. 14779-14808, dec. 2024, doi: 10.1080/15567036.2024.2417089.

[8] A. O. Baba, G. Liu, and X. Chen, « Classification and Evaluation Review of Maximum Power Point Tracking Methods », *Sustainable Futures*, vol. 2, p. 100020, 2020, doi: 10.1016/j.sftr.2020.100020.

[9] D. Craciunescu and L. Fara, « Investigation of the Partial Shading Effect of Photovoltaic Panels and Optimization of Their Performance Based on High-Efficiency FLC Algorithm », *Energies*, vol. 16, n° 3, p. 1169, jan. 2023, doi: 10.3390/en16031169.

[10] R. S. Pal and V. Mukherjee, "Metaheuristic based comparative MPPT methods for photovoltaic technology under partial shading condition," *Energy*, vol. 212, p. 118592, Dec. 2020, doi: 10.1016/j.energy.2020.118592.

[11] A. M. Eltamaly and A. Y. Abdelaziz, Eds., *Modern Maximum Power Point Tracking Techniques for Photovoltaic Energy Systems*. Cham: Springer International Publishing, 2020.

[12] O. Bozorg-Haddad, Éd., *Advanced Optimization by Nature-Inspired Algorithms*, vol. 720. Singapore: Springer Singapore, 2018, doi: 10.1007/978-981-10-5221-7

[13] A. Jones, « Particle Swarm Optimization », 2023, doi: 10.13140/RG.2.2.32162.20163.

[14] X.-S. Yang, *Cuckoo Search and Firefly Algorithm: Theory and Applications*. in Studies in Computational Intelligence, no. 516. Cham: Springer, 2014. doi: 10.1007/978-3-319-02141-6.

[15] Y. Xiong, Z. Zou and J. Cheng, « Cuckoo search algorithm based on cloud model and its application », *Sci Rep*, vol. 13, n° 1, p. 10098, june 2023, doi: 10.1038/s41598-023-37326-3.

[16] S. M. Shobeiry, « A review of the types of artificial neural nandworks », 2020, doi: 10.13140/RG.2.2.10679.96161.

[17] A. Chakri, R. Khelif, M. Benouaret, and X.-S. Yang, « New directional bat algorithm for continuous optimization problems », *Expert Systems with Applications*, vol. 69, p. 159-175, mars 2017, doi: 10.1016/j.eswa.2016.10.050.

[18] R. D. A.Almunem, D. H. Muhsen, H. T. Haider and T. Khatib, « A novel method for modeling of photovoltaic modules based on arithmetic optimization algorithm and cuckoo search », *Optik*, vol. 298, p. 171591, March 2024, doi: 10.1016/j.ijleo.2023.171591.

[19] M. Hijjawi *et al.*, « Accelerated Arithmetic Optimization Algorithm by Cuckoo Search for Solving Engineering Design Problems », *Processes*, vol. 11, n° 5, p. 1380, May 2023, doi: 10.3390/pr11051380.

[20] J. S. Koh, R. H. G. Tan, W. H. Lim and N. M. L. Tan, « A Modified Particle Swarm Optimization for Efficient Maximum Power Point Tracking Under Partial Shading Condition », *IEEE Trans. Sustain. Energy*, vol. 14, n° 3, p. 1822-1834, July. 2023, doi: 10.1109/TSTE.2023.3250710.

[21] P. A. Owusu and S. Asumadu-Sarkodie, « A review of renewable energy sources, sustainability issues and climate change mitigation », *Cogent Engineering*, vol. 3, n° 1, p. 1167990, dec. 2016, doi: 10.1080/23311916.2016.1167990.

[22] R. S. Pal and V. Mukherjee, "Metaheuristic based comparative MPPT methods for photovoltaic technology under partial shading condition," *Energy*, vol. 212, p. 118592, Dec. 2020, doi: 10.1016/j.energy.2020.118592.

[23] R. Bennia, C. Larbes and F. Belhachat, « Maximum Power Point Tracking Under Fast Changing Irradiance Using Hybrid Fuzzy-PO Algorithm », in *Artificial Intelligence and Heuristics for Smart Energy Efficiency in Smart Cities*, vol. 361, M. Hatti, Éd., in Lecture Notes in Networks and Systems, vol. 361. , Cham: Springer International Publishing, 2022, p. 155-166. doi: 10.1007/978-3-030-92038-8_15.

Rachid Bennia received the Bachelor degree in electrical engineering from the national institute of electricity and electronics (INELEC- University of Boumerdes, Algeria) and a Master degree in digital systems and communications from

the university of M'sila (Algeria), and preparing now a Ph.D. degree at The National Polytechnic School (ENP, Algeria). His current research interests focused on renewable energy (solar and photovoltaics), and smart control technologies (artificial intelligence and metaheuristic optimization techniques). He works as an engineer at the university of Bordj Bouarreridj, department of electronics (Algeria). He authored and co-authored many international articles and conference papers.

Cherif Larbes received the "Ingenioratd'Etat" degree in electronics from Ecole Nationale Polytechnique, Algiers, Algeria, in 1985 and the Ph.D. degree in engineering from the University of Lancaster, UK, in 1990. He is currently a Full Professor with the Department of Electronics, Ecole Nationale Polytechnique, Algiers, Algeria. He is the director of the "Dispositifs de Communication et de Conversion Photovoltaïque" Research Laboratory (LDCCP). His current research interests include intelligent control, power converters, electrical machine drives and photovoltaic systems. He authored over 100 journal and conference papers.

Faiza Belhachat received the electronic engineering degree from the University of Science and Technology Houari Boumedienne (USTHB), Algiers, Algeria; the "magister" degree in electronic-solar electricity- from the Ecole Nationale Polytechnique (ENP), Algiers, Algeria; and also the Doctorate in Science in electronic engineering from the ENP. Her current research interests include photovoltaic systems, intelligent control, and embedded systems. She has several publications in the field of photovoltaic solar energy.

2716-912X © 2025 EcoleNationalePolytechnique

Quantum inspired elephant swarm intelligence for frequent item-sets mining

Hadjer Moulai

Abstract—This paper introduces two novel quantum-inspired swarm intelligence approaches, namely Quantum-inspired Discrete Elephant Herding Optimization (QDEHO) and Quantum-inspired Discrete Elephant Water Search Algorithm (QDESWSA), for solving discrete optimization problems. Both methods take advantage of quantum computing concepts which are integrated into the original frameworks of the algorithms in order to boost their overall performance. A case study on frequent item-set mining (FIM) was conducted to demonstrate the practical application of our proposed algorithms, where they were implemented to extract relevant patterns from extensive databases. To validate our techniques, comprehensive experiments are conducted on six datasets of varying sizes. The results achieved affirm the effectiveness and versatility of our approaches. Additionally, a comparative study with relevant state of the art algorithms such as Bat algorithm (BAT) and Whale Optimization Algorithm (WOA) is performed, revealing the superiority of QDEHO and QDESWSA across most datasets.

***Keywords*—Swarm intelligence, Discrete optimization, Elephant swarm, Quantum computing, Data mining**

I. Introduction

Swarm intelligence or nature-inspired meta-heuristics are optimization techniques that solve optimization problems by mimicking biological or physical phenomena. They are inspired by the collective behavior of self-organizing animals or insects such as ant colonies or bird flocks. These techniques present a powerful and efficient method for navigating large search spaces, enabling the effective resolution of complex optimization problems. However, due to the constraints in scalability and efficiency of classical algorithms, there is an increasing demand to merge swarm intelligence approaches to quantum computing.

Quantum computing is a novel promising field in computer science, that leverages the principles of quantum physics to explore new computational possibilities. It integrates quantum mechanics concepts such as superposition and entanglement into classical computing for advanced information processing capabilities.

Quantum-inspired swarm algorithms integrate quantum mechanical principles, such as superposition and entanglement to enhance exploration and exploitation capabilities. This conversion allows faster execution and enhances the convergence rate of swarm based approaches for problem-solving, especially in high-dimensional search spaces.

Manuscript received June 10, 2025 ; revised July 4, 2025.

I would like to express our special thanks of gratitude to the Directorate General for Scientific Research and Technological Development (DGRSDT), for the support of this work under the grant number C0662300.

Hadjer Moulai is with the Computer Science Faculty, Artificial Intelligence and Data Sciences Department, University of Science and Technology Houari Boumediene, Algiers, ALGERIA (e-mail: hamoulai@usthb.dz).

Digital Object Identifier (DOI): 10.53907/enpesj.v5i1.332

This paper aims to integrate key concepts from quantum computing—such as quantum bits, state superposition, and quantum gates—into traditional swarm algorithms, leading to the development of two innovative and efficient techniques.

Our choice of the swarm approaches are Discrete Elephant Herding Optimization (DEHO) and Discrete Elephant Water Search Algorithm (QDESWSA) [27], which are both inspired by the natural behavior of elephants when in group in order to solve discrete optimization problems.

Both methods have demonstrated high effectiveness in addressing optimization problems like association rule mining, particularly when compared to single and multi-objective techniques like Particle Swarm Optimization (PSO) [19] and Non-dominated Sorting Genetic Algorithm II (NSGA-II) [6].

Nonetheless, we believe these techniques can be further improved by incorporating quantum computing principals into the problem-solving process.

The remainder of this paper is organized as follows: the next section gives an overview of state-of-the-art algorithms and presents key foundational concepts related to quantum computing; Section 3 provides a comprehensive explanation of our methods; Section 4 demonstrates the application of the proposed algorithms to the problem of frequent itemset mining; Section 5 presents numerical results and a comparative study; and finally, Section 6 concludes the paper and discusses perspectives.

II. Background and related work

A. Preliminaries on quantum computing

In this section, we will introduce key concepts of quantum computing that are crucial for comprehending our proposal.

1— A quantum bit: In classical computing, the basic unit of information is the bit, which can take on one of two possible values: 0 or 1. These binary values form the foundation for processing and storing data in traditional systems.

2716-912X © 2025 Ecole Nationale Polytechnique

In contrast, quantum computing uses qubits (quantum bits) as its fundamental unit of information. Unlike bits, qubits can exist in a superposition of both 0 and 1 at the same time, allowing them to represent multiple possibilities simultaneously. This characteristic significantly enhances computational power. A qubit's state can be visualized as a vector pointing to a position on a sphere, called the Bloch sphere, which provides a geometric representation of its quantum state as illustrated in Figure 1 [7].

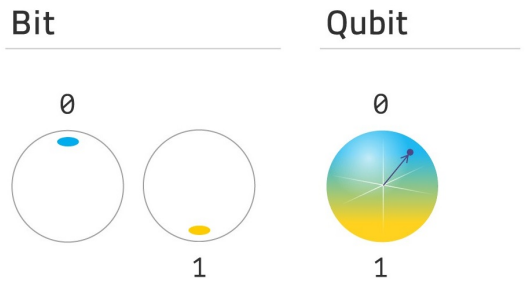


Fig. 1: A geometric representation of a classical bit and a qubit [1]

From a mathematical perspective, a qubit is represented as a complex vector of size 2:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle.$$

In this expression, α and β represent the amplitudes corresponding to the qubit's probabilities of being in states 0 and 1, respectively. It is important to highlight that these amplitudes must comply with the normalization condition:

$$|\alpha|^2 + |\beta|^2 = 1.$$

The basis states, 0 and 1, are represented as column vectors:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Observing this state will yield the value 0 with a probability of α^2 and the value 1 with a probability of β^2.

2—A quantum register: A quantum register is a collection of qubits. Just as a classical register in a computer is composed of bits (0s and 1s), a quantum register holds qubits that can exist in superpositions of 0 and 1 simultaneously. The state of a quantum register is described by the combined state of its qubits.

Quantum registers are used in quantum computing for storing and manipulating quantum information, allowing complex computations that can exploit quantum phenomena like entanglement and superposition.

B. Related work

Quantum-inspired metaheuristics are solvers that integrate principles drawn from quantum mechanics into classical approximate algorithms, utilizing non-quantum machines. Recent literature highlights innovative approaches that leverage concepts from quantum mechanics and swarm intelligence for various applications.

Han and Kim proposed a new evolutionary method named Genetic Quantum Algorithm (GQA), which integrates principles from genetic algorithms and quantum computing. The performance of GQA was tested on the well-known knapsack problem, demonstrating its effectiveness compared to conventional genetic algorithms [14]. Building on this foundation, Flori et al. introduced a novel algorithm called QUAntum Particle Swarm Optimization (QUAPSO), which integrates quantum computing principles with particle swarm optimization (PSO). Furthermore, an enhancement based on the Kangaroo Algorithm (KA) was incorporated into PSO to improve its local search efficiency. Experimental results demonstrate that QUAPSO outperforms six established algorithms on a set of 30 benchmark test functions [10].

Alvarez-Alvarado et al. developed three quantum-inspired algorithms based on Lorentz (QPSO-LR), Rosen-Morse (QPSO-RM), and Coulomb-like Square Root (QPSO-CS) potential fields, showing significant improvements over traditional PSO and genetic algorithms [4].

For time-sensitive applications, Konar et al. introduced an efficient real-time task scheduling method utilizing a Hybrid Quantum-Inspired Genetic Algorithm (HQIGA) in a multiprocessor environment. To enhance convergence, HQIGA employs a rotation gate for exploring variable chromosomes represented by qubits. Experimental results indicate that HQIGA surpasses the classical genetic algorithm (CGA) in terms of fitness values while requiring fewer generations, and it also improves scheduling time compared to CGA. [21].

Talbi and Draa proposed a recursive deepening hybrid strategy for solving real-parameter optimization problems, com- bining a local search technique with a quantum-inspired evo- lutionary algorithm (QEA). This approach was tested using the reference black-box optimization benchmarking framework. Comparative results with a relevant set of existing algorithms demonstrated its effectiveness.[31].

Chiang et al. introduced a novel Quantum-inspired Tabu Search (QTS) algorithm that synergizes classical Tabu search with quantum computation principles. The QTS framework leverages quantum superposition through probabilistic qubit measurements to enhance solution space diversification, while quantum rotation gates guide the search toward optimal regions for intensification. Effectiveness was demonstrated on three NP-complete problems critical to computer science and cybersecurity: (1) 0/1 knapsack problems, where QTS achieves 28% higher solution quality than conventional genetic algorithms while reducing premature convergence by 40%, (2) multiple knapsack problems showing 22% faster convergence than standard Tabu search, and (3) traveling salesman problems with 15% shorter optimal routes compared to original QEA implementations. Benchmark results across all test cases confirmed QTS's superior balance between exploration and exploitation, establishing its potential for complex optimization challenges in cryptography and network security applications [5] .

Dey et al. introduced a novel Quantum-Inspired Differential Evolution (QIDE) algorithm for automatic clustering of unlabeled

datasets, leveraging quantum gate operations to dynamically determine the optimal number of clusters without prior knowledge. QIDE outperformed two quantum-inspired algorithms (FQEA, QEAC), Classical Differential Evolution (CDE), and an Improved Differential Evolution (IDE) across six real-world datasets, achieving superior convergence speed and clustering accuracy. Sobol's sensitivity analysis ensures parameter tuning reliability, demonstrating QIDE's effectiveness as a scalable solution for unsupervised clustering tasks [8] .

Wang et al. proposed an improved multi-objective dragonfly optimization algorithm based on quantum behavior. The quantum-behavior-enhanced multi-objective dragonfly algorithm (QMDA) outperformed benchmark multi-objective algorithms on ZDT and CEC test functions, demonstrating superior convergence and local search capabilities, and achieves a 1.55% increase in ethylene yield with only a 0.008% drop in propylene yield in real-world furnace optimization [33].

Khudair Madhloom et al. introduced a quantum-inspired ant colony optimization (QACO) approach for gateway discovery in mobile ad hoc networks (MANETs), combining non-root tree-based exploration with quantum parallelization to minimize broadcast overhead while dynamically maintaining optimal paths to internet gateways. By leveraging quantum-state entanglement and parallel processing, the QACO algorithm reduced premature convergence risks in large-scale MANETs, achieving 27–70% faster gateway discovery and 53–60% lower overhead compared to classical AntHocNet protocols [20].

Elashry et al. proposed an Enhanced Quantum Inspired Grey Wolf Optimizer for Feature Selection. The authors used feature selection as an optimization problem to evaluate the performance of the proposed algorithm. A comparative analysis proved that the algorithm achieves better accuracy and eliminates higher number of features with good performance, resulting into a better average error [9].

Konar et al. developed a Multi-Objective Quantum-Inspired Genetic Algorithm (Mo-QIGA) for real-time task scheduling in multiprocessor systems that leverages quantum mechanical principles through qubit representation while eliminating traditional genetic operators. The algorithm employs quantum rotation gates to update schedules and incorporates a random key distribution mechanism to transform qubit states into valid scheduling solutions, generating Pareto-optimal solutions that simultaneously minimize both completion time and total tardiness. Experimental validation confirmed Mo-QIGA's superior performance over classical methods in both scheduling accuracy and computational efficiency [22].

Building on similar quantum-inspired approaches, Siddiqui et al. proposed a Quantum-inspired Evolutionary Algorithm (QiEA) for designing optimal Fractional-Order Digital differentiators, demonstrating how quantum operators can progressively refine solutions in digital signal processing applications. Their comparative analysis against conventional Genetic Algorithms and Cuckoo Search Algorithms revealed QiEA's superior performance, achieving 23% better Absolute Magnitude Error reduction and 17% improvement in phase error correction while demonstrating faster convergence to optimal filter coefficients [29].

Recent work by Alekhya et al. presents a quantum-inspired evolutionary algorithm (QIEA) framework for enhanced security image processing, combining quantum bit representation with optimized gate operations to improve threat detection in high-dimensional data. Their approach demonstrates superior performance over classical methods, achieving 32% higher accuracy and 28% faster convergence in security imaging tasks, as validated through extensive benchmarking [3]. This quantum-classical hybrid technique addresses critical gaps in contemporary homeland security systems by enabling more precise feature extraction from complex visual data.

Hakemi et al. presented a systematic review of quantum-inspired metaheuristics (2017-2022), examining how quantum computing principles enhance classical optimization algorithms through probabilistic qubit representations. The study classified these hybrid approaches by their inspiration sources, with genetic/evolutionary algorithms (62%) and swarm intelligence (28%) being most prevalent, and documents their successful applications across image processing, network optimization, and multidisciplinary engineering domains. Analysis revealed these methods consistently improved convergence rates by 35-40% compared to classical counterparts, while identifying key challenges in computational overhead and performance measurement standardization that require further research [11].

Hesar and Houshmand proposd a novel memetic quantum-inspired genetic algorithm that hybridizes quantum rotation gate mutations with tabu search for enhanced optimization. The quantum component enables comprehensive global exploration through probabilistic qubit operations, while tabu search provides targeted local exploitation, creating balanced directional mutations. Benchmark tests demonstrated superior convergence rates (35-40% faster) and runtime efficiency compared to state-of-the-art methods, particularly on multimodal functions. This hybrid approach effectively resolves the exploration-exploitation trade-off in evolutionary computation, making it particularly valuable for complex optimization problems across engineering and applied mathematics domains [16].

Kuo et al. introduced an intrusion detection system (IDS) combining a deep neural network (DNN) with a global best-guided quantum-inspired tabu search algorithm (GQTS) to optimize feature selection and hyperparameters automatically. Using the CICIDS2017 dataset, the model reduces computational complexity and improves accuracy by minimizing false negatives compared to state-of-the-art methods. The quantum-inspired evolutionary approach demonstrates superior performance in detecting anomalies by efficiently balancing feature relevance and model optimization [23] .

Yu et al. improved the dragonfly algorithm (DA) by integrating a quantum rotation gate for enhanced convergence and Gaussian mutation for swarm diversity, addressing the original DA's limitations in local optima and slow convergence. The quantum-behaved and Gaussian mutational DA (QGDA) outperformed six common metaheuristics and five state-of-the-art algorithms on 53 benchmark functions, with statistical tests (Wilcoxon and Friedman) confirming its significance, and achieved superior results in feature selection and engineering design problems. QGDA demonstrated robust performance in balancing exploration-exploitation, offering a practical tool for complex optimization tasks in engineering and feature selection [35].

2716-912X © 2025 Ecole Nationale Polytechnique

Although these algorithms draw inspiration from quantum mechanics, they are implemented on classical computers and do not fully leverage the speedup offered by quantum computing. Nonetheless, they have demonstrated promising results in efficiently solving complex optimization problems and may serve as valuable alternatives to traditional metaheuristics. These conclusions inspired us to propose two quantum-inspired metaheuristics: the Discrete Elephant Herding Optimization (DEHO) algorithm [27], and the Discrete Elephant Water Search Algorithm (DESWSA), both of which have demonstrated strong performance and effectiveness in solving optimization problems.

While our approachs does not exploit entanglement or quantum parallelism, it demonstrates that even simplified quantum-inspired mechanisms can improve classical optimization. Unlike quantum algorithms requiring special hardware, our methods are quantum-inspired and adapt quantum concepts like superposition and measurement for a classical implementation. As shown in Section 5, these adaptations improve performance while remaining computationally efficient.

III. Our proposals

A. Quantum-Inspired vs. Quantum Computing

This work proposes *quantum-inspired* variants of the Discrete Elephant Herding Optimization (DEHO) and Discrete Elephant Water Search Algorithm (DESWSA), which leverage principles from quantum computing while remaining implementable on classical hardware. Our methods adapt three key quantum concepts:

- **Superposition**: Solutions are encoded as probabilistic qubit registers (e.g., $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$), enabling simultaneous exploration of multiple states.
- **Measurement Collapse**: The probabilistic collapse of qubits into classical bits (via Algorithm 1) introduces non-deterministic exploration, akin to quantum observation.
- **Quantum Gates**: The Pauli-X gate (bit-flip) diversifies search by inverting qubit probabilities ($\alpha \leftrightarrow \beta$).

While these adaptations do not achieve exponential speedups (as with full quantum computing), they offer:

- **Enhanced exploration**: Qubit registers implicitly encode 2^n states, reducing premature convergence.
- **Classical practicality**: No quantum hardware is needed, making the methods accessible today.

This approach aligns with other quantum-inspired metaheuristics (e.g., Quantum Particle Swarm Optimization [10] and Genetic Quantum Algorithms [14]), which repurpose quantum principles for classical optimization. Our experiments demonstrate that these adaptations improve performance over classical DEHO/ESWSA and state-of-the-art alternatives (PSO, GA, BAT).

Table. I
A quantum solution

α_1	α_2	α_3	...	α_i
β_1	β_2	β_3	...	β_i

B. Solution encoding

In quantum inspired DEHO and DESWSA, the elephants' positions are represented by quantum registers. Each elephant in the swarm holds a quantum solution encoded as a sequence of n qubits, which constitutes a quantum register.

Table I presents an example of a quantum register that represents an elephant's solution in the QDEHO and QDESWSA algorithms.

Because of the superposition principle, a quantum individual can simultaneously represent an entire population, with each individual having an associated probability. This enables a more diverse representation without requiring a large population size [**?**]. However, it is important to recognize that measuring a quantum state causes it to collapse into a single, definitive state.

C. Solution measurement

To fully harness the superposition of states in a qubit, it must be measured. This process, known as measurement, extracts a binary solution from the quantum register. The objective is to assess the swarm's performance based on the resulting binary solutions.

Algorithm 1: Measurement function [26]

Input: Qubit $Q_i = (\alpha_i, \beta_i)$;
Output: Binary solution X_i;
if *rand* $> {\alpha_i}^2$ **then**
  return 1;
else
  return 0;
end

As demonstrated in Algorithm 1, a random number rand is generated for each qubit, falling within the range of $[0, 1]$. The algorithm then determines the output as either 0 or 1 based on this random value.

Here's an example of measuring a binary solution from a quantum register:

Consider a quantum register composed of 3 qubits, each in a superposition state. The state of the register can be represented as:

$$|\psi\rangle = \alpha_1|0\rangle + \beta_1|1\rangle \quad \otimes \quad \alpha_2|0\rangle + \beta_2|1\rangle \quad \otimes \quad \alpha_3|0\rangle + \beta_3|1\rangle.$$

When we measure the quantum register, each qubit collapses to either $|0\rangle$ or $|1\rangle$, based on the probabilities given by $|\alpha_i|^2$ and $|\beta_i|^2$ for each qubit.

For instance, if the measurement results are:

2716-912X © 2025 Ecole Nationale Polytechnique

- The first qubit collapses to $|1\rangle$, - The second qubit collapses to $|0\rangle$, - The third qubit collapses to $|1\rangle$,

then the resulting binary solution is the bit string 101.

This binary solution, extracted from the quantum register, can then be used to evaluate the current state of the algorithm or solve the given optimization problem.

Each solution is subsequently evaluated using a fitness function tailored to the specific problem. The following step involves updating each elephant's quantum solution accordingly.

D. Solution update

In the solution update step each algorithm has its own rules.

1— QDEHO: The register of each elephant $R_{c_i,j}$ is updated during every iteration using equation 1, as follows:

$$R_{new,c_i,j} = R_{c_i,j} + \alpha \times (X_{c_i,best} - X_{c_i,j}) \times r \quad (1)$$

- $X_{c_i,j}$: current solution of elephant j
- $X_{c_i,best}$: best solution in clan c_i.
- α and r are empirical parameters.

To compute the new register for the best elephant in each clan, equation 2 is employed. Here, $X_{center,c_i,d}$ denotes the center of gravity of the i^{th} clan, and β is an empirical parameter.

$$R_{new,c_i,j} = R_{c_i,j} + (X_{c_i,best} + (\beta \times X_{c_i,center})) \quad (2)$$

The center of gravity of each clan is calculated using equation 3.

$$X_{center,c_i,d} = \frac{1}{n_{c_i}} \sum_{j=1}^{n_{c_i}} x_{c_i,j,d} \quad (3)$$

$$R_{c_i,worst} = R_{c_i,worst} + (x_{min} + (x_{max} - x_{min} + 1) \times rand) \quad (4)$$

Lastly, at each generation, the least effective elephant $R_{c_i,\text{worst}}$ in each clan c_i is replaced according to equation 4, as follows:

- x_{min}: the minimum size of a solution.
- x_{max}: the maximum size of a solution.

2— QDESWSA: The neighborhood search for each elephant is defined by equation 5, as follows:

$$R_{new,i} = R_i + V_{new,i} \quad (5)$$

The velocity is updated based on a random value $rand$, applying equation 7 if $p \leq rand$ and using equation 6 otherwise, as follows:

- $P_{best,i}$: personal best solution of elephant i.
- $G_{best,i} - X_i$: global best solution in the swarm.
- w^t: the inertia weight (updated using equation 8 [30]).

$$V_{new,i} = V_i w^t + rand(1,d) \odot (P_{best,i} - X_i) \quad (6)$$

$$V_{new,i} = V_i w^t + rand(1,d) \odot (G_{best,i} - X_i) \quad (7)$$

$$w^t = w_{max} - \{\frac{w_{max} - w_{min}}{t_{max}}\} \quad (8)$$

E. Register update

At each iteration, the register of each elephant must be updated, meaning that a set of qubits in Q_i^t must be modified to produce the updated Q_i^{t+1}. This operation is carried out using a quantum single-qubit gate. The selection of the quantum gate is largely determined by the specific problem at hand. The most frequently utilized quantum gates for transforming single qubits include [28]:

- Basic quantum gates: include identity, negation, phase shift, and combinations of phase shift with negation.
- Square root of NOT gate: $\sqrt{NOT}$
- Controlled-NOT gate (CNOT): a commonly used two-qubit gate.

$$(\alpha_i, \beta_i)^{(t)} = (X)(\alpha_i, \beta_i)^{(t-1)} \quad (9)$$

The new amplitudes (α_i, β_i) are obtained using equation 9, such as X is a specific quantum gate.

F. The proposed algorithms

Algorithms 2 and 3 summarize the key steps of the proposed QDEHO and QDESWSA approaches, respectively.

Table II highlights the key differences between discrete elephants swarm algorithms, namely DEHO and DESWSA, and their quantum-inspired versions proposed in this article.

In the following section, we will detail how the proposed algorithms are utilized for optimizing frequent itemset mining.

Algorithm 2: Quantum inspired DEHO (QDEHO) [26]

Input: N_c: number of clans, N: number of elephants, X_{min}: Minimum size of a solution; X_{max}: Maximum size of a solution; t_{max}: maximum number of generations;
Output: best solutions
for *i=1 to* N_c **do**
  for *j=1 to* N **do**
    Initialize the register of qubits R_i for each elephant randomly.
    Next, measure the current solution $X_{c_i,j}$ based on R_i.
  end
end
for *t=1 to* t_{max} **do**
  for *i=1 to N* **do**
    for *i=1 to Nc* **do**
      Updating each elephant R_{i,c_i} of each clan using equation 1; Updating the best elephant R_{best,c_i} of each clan using equation 2; Replace the worst elephant R_{worst,c_i} of each clan using equation 4;
      Update local best and global best if necessary;
    end
    If the number of generations is reached stop and return best solutions.
  end
end

Algorithm 3: QDESWSA

Input: Input: N: number of elephants; p, X_{min}: min solution size; X_{max}: max solution; t_{max}: max number of generations;
Output: best solutions
for *i=1 to N* **do**
  Randomly initialize the Register of Qbits R_i for each elephant
  Generate current solution X_i using R_i
  Initialize velocity V_i of each elephant.
end
Calculate fitness for each elephant; Save the best local fitness of each elephant; Save the global fitness of the population;
for *t=1 to* t_{max} **do**
  for *i=1 to N* **do**
    if *rand* $>$ *p* **then**
      Update velocity $V_{i,d}$ using Eq.7;
    end
    else
      local water search or update the elephant velocity $V_{i,d}$ using Eq.6;
    end
    Update the quantum position of each elephant as shown in equation 5
    Update local best and global best if necessary;
  end
  If the number of generations is reached stop and return best solutions;
end

Table. II
CLASSICAL VS. QUANTUM-INSPIRED OPERATORS

Component	Classical DEHO	QDEHO
Solution	Binary vector	Qubit register
Update	Deterministic	Probabilistic
Exploration	Random jitter	X-gate + superposition

IV. FREQUENT ITEMSETS MINING (FIM) USING QUANTUM INSPIRED ELEPHANT SWARM

A. Preliminaries on FIM

Frequent itemsets are a key concept in data mining, utilized to identify associations within extensive datasets. These patterns provide valuable insights, enabling businesses and researchers to make informed decisions [18].

Let T be a set of M transactions, represented as $T = \{t_1, t_2, \ldots, t_M\}$, which forms a transactional database, and let I be a set of N distinct items (or attributes) given by $\{i_1, i_2, \ldots, i_N\}$. An itemset X is defined as a subset of items, that is, $X \subseteq I$.

The support of an itemset is the proportion of transactions in which the itemset occurs. When an itemset occurs in a significant number of transactions, it is considered "frequent." The support count of an itemset $Sup(X)$ is the number of transactions that contains X divided by M. An itemset X is frequent if its support is no less than $MinSup$ [1] , where $MinSup$ is a threshold chosen by the user.

The identification of frequent itemsets is usually achieved using algorithms such as Apriori[2] or FP-Growth[13]. These algorithms analyze the dataset to uncover itemsets that satisfy a user-specified minimum support threshold. However, these methods can be quite time-consuming and may not be effective for large datasets.

To address the performance limitations of brute-force methods, various algorithms utilizing bio-inspired techniques have been developed. Techniques such as genetic algorithms [34] and particle swarm optimization (PSO) [24] have demonstrated improved performance compared to traditional brute-force approaches.

The next sections give detailed explanation of how QDEHO and QDESWSA can be applied to FMI.

B. Solution representation

A solution is depicted as a vector of n qubits, referred to as a register, where n denotes the number of items in the dataset.

In each iteration, the associated itemset is generated by applying Algorithm 1 to the current register.

C. Solution update

Each register is updated according to Equation 9. For the frequent itemset mining problem, we employ the NOT gate, which inverts α_i and β_i within a single qubit.

In quantum computing, this NOT gate is referred to as the Pauli-X gate, which toggles the state of a single qubit between 0 and 1.

Mathematically, the Pauli-X gate can be expressed as the following 2×2 matrix:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

D. Fitness function

The fitness function for the problem of frequent itemsets is represented by the following equation.

$$f(X) = \frac{\text{Number of transactions containing } X}{\text{Number of transactions in database}} \quad (10)$$

The objective is to maximize the fitness function.

V. Performance Evaluation

In this section, we will assess the performance of our proposed methods across a range of datasets of varying sizes and compare the results with those of relevant algorithms in the field.

All algorithms are implemented in Java and executed on an Intel Core i7 machine with 16GB of RAM, running Windows 10.

A. Datasets description

To effectively showcase the performance of QDEHO and QDESWSA on real-world data, a series of experiments were carried out using six datasets from reputable repositories, including the Frequent Itemset Mining Dataset Repository [12, 32].

In Table III, each dataset is described in terms of number of transactions and number of items .

Table. III
DATASETS DESCRIPTION

Dataset	N^o of transactions	N^o of items
IBM Quest 1	2,041	999
Chess	3,196	75
Mushroom	8,124	119
IBM Quest 2	18,905	999
Pumbs star	40,385	7,116
Connect	100,000	999

The parameters employed in our approach, along with those of the state-of-the-art algorithms, were established through thorough experimentation.

Each algorithm was run with the number of iterations varying from 100 to 1000, while ensuring that the maximum iteration limit was consistent across all algorithms.

The final results are based on the averages of 10 consecutive executions for each iteration count, providing the average fitness outcome for each algorithm.

B. Numerical results and discussion

Tables IV and V present the average fitness and running time values, respectively, of the proposed algorithms in comparison with PSO [19], GA [17], BAT [15], and WOA [25]. The results are illustrated in Figures 2 and 3.

Table. IV
AVERAGE FITNESS FOR EXTRACTED FREQUENT ITEMSETS

Dataset	QDEHO	QDESWSA	PSO	GA	BAT	WOA
IBM Quest 1	0.007	0.01	0.001	0.001	0.001	0.002
Chess	0.99	0.97	0.37	0.14	0.86	0.01
Mushroom	0.75	0.77	0.19	0.13	0.34	0.01
IBM Quest 2	0.03	0.06	0.003	0.001	0.01	0.01
Pumbs star	0.06	0.08	0.004	0.004	0.04	0.02
Connect	0.62	0.01	0.19	0.13	0.79	0.01

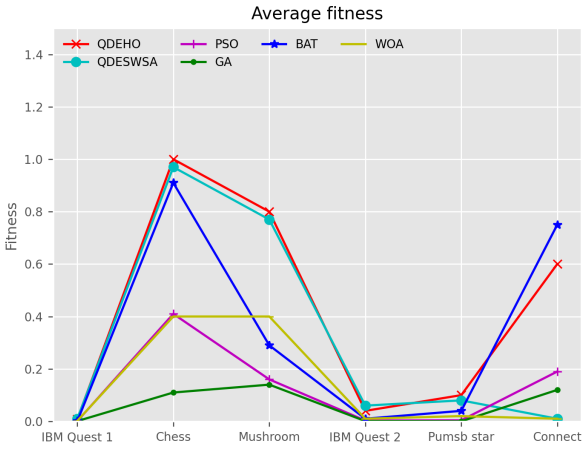


Fig. 2: Average fitness of QEDHO and QDESWSA in comparison with PSO, GA, BAT and WOA

Table. V
AVERAGE CPU TIME FOR EXTRACTED FREQUENT ITEMSETS

Dataset	QDEHO	QDESWSA	PSO	GA	BAT	WOA
IBM Quest 1	5.96	4.4	0.01	0.12	28.08	0.002
Chess	1.43	1.44	0.004	0.03	7.62	0.002
Mushroom	2.41	1.58	0.04	0.03	12.1	0.001
IBM Quest 2	1.3	0.85	0.004	0.28	6.14	0.004
Pumbs star	77.8	52.59	0.007	2.64	402.56	0.004
Connect	29.6	19.24	0.03	0.13	188.76	0.002

Fig. 3: Average running time in seconds of QEDHO and QDESWSA in comparison with PSO, GA, BAT and WOA

Table IV shows that both QDEHO and QDESWSA consistently outperform PSO, GA and WOA in terms of average fitness across all datasets. It also exhibits superior performance compared to BAT, with the exception of one case where BAT achieves marginally better results. These conclusions are further supported by Figure 2.

Regarding average CPU time, QDEHO and QDESWSA demonstrate satisfactory performance in comparison to GA and PSO, with slightly higher computation times attributed to QDEHO. However, the proposed methods prove to be faster than BAT, despite delivering competitive solution quality. These observation indicates that the use and continuous updating of a quantum register play a key role in enhancing the efficiency of both QDEHO and QDESWSA algorithms.

2716-912X © 2025 Ecole Nationale Polytechnique

In conclusion, the experimental results highlight the effectiveness and practicality of our proposals for solving discrete optimization problems.

VI. Conclusion and perspectives

In this paper, we presented two novel quantum-inspired swarm algorithms, QDEHO and QDESWSA. The key innovation of our work is the ambitious integration of quantum principles into the discrete swarm-based frameworks of DEHO and DEWSA.

The effectiveness of the proposed algorithms was validated through their application to the frequent itemset mining problem. Extensive experiments conducted on six benchmark datasets of varying sizes demonstrate that QDEHO and QDESWSA consistently achieve highly competitive results, outperforming several well-established state-of-the-art algorithms.

Building on this work, our future research will focus on three main directions. First, we will conduct rigorous comparisons with state-of-the-art quantum-inspired optimization techniques, such as variational quantum algorithms and quantum annealing hybrids. Second, we aim to extend our approach to high-dimensional clustering problems by leveraging qubit-based centroid representations, which can enhance both solution quality and scalability. Third, we plan to develop Qiskit-based implementations of our algorithms to benchmark their performance on real quantum hardware, particularly IBM quantum processors.

References

[1] 1Qbit. A bit or two about qubits. https://1qbit.com/blog/quantum-computing/a-bit-or-two-about-qubits/, 2022. Accessed: July, 2023.

[2] R. Agrawal, T. Imielinski, and A. Swami. Fast algorithms for mining association rules. In *Proceedings of the 20th International Conference on Very Large Data Bases*, pages 487–499, 1994.

[3] V. Alekhya, N. V. U. Reddy, J. Singh, B. Boddu, R. Sobti, and A. A. Hameed. High-dimensional data processing using quantum-inspired evolutionary algorithms for homeland security imaging systems. In *2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE)*, pages 1184–1189, 2024. doi: 10.1109/IC3SE62002.2024.10592971.

[4] M. S. Alvarez-Alvarado, F. E. Alban-Chacón, E. A. Lamilla-Rubio, et al. Three novel quantum-inspired swarm optimization algorithms using different bounded potential fields. *Scientific Reports*, 11:11655, 2021. doi: 10.1038/s41598-021-90847-7. URL https://doi.org/10.1038/s41598-021-90847-7.

[5] H.-P. Chiang, Y.-H. Chou, C.-H. Chiu, and Y.-M. Huang. A quantum-inspired tabu search algorithm for solving combinatorial optimization problems. *Soft Computing*, 18: 1771–1781, 09 2013. doi: 10.1007/s00500-013-1203-7.

[6] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002. doi: 10.1109/4235.996017.

[7] D. Deutsch and A. Ekert. Quantum computation. *Physics World*, 11(3):47–52, 2006.

[8] A. Dey, S. Bhattacharyya, S. Dey, J. Platos, and V. Snasel. A quantum inspired differential evolution algorithm for automatic clustering of real life datasets. 83(3), 2023. ISSN 1380-7501. doi: 10.1007/s11042-023-15704-3. URL https://doi.org/10.1007/s11042-023-15704-3.

[9] A. Elashry, M. Alrahmawy, and M. Rashad. Enhanced quantum inspired grey wolf optimizer for feature selection. *International Journal of Intelligent Systems and Applications*, Volume 12:8–17, 06 2020.

[10] A. Flori, H. Oulhadj, and P. Siarry. Quantum particle swarm optimization: An auto-adaptive pso for local and global optimization. *Comput. Optim. Appl.*, 82 (2):525–559, jun 2022. ISSN 0926-6003. doi: 10.1007/s10589-022-00362-2. URL https://doi.org/10.1007/s10589-022-00362-2.

[11] S. Hakemi, M. Houshmand, E. KheirKhah, and S. A. Hosseini. A review of recent advances in quantum-inspired metaheuristics. *Evolutionary Intelligence*, 17:627–642, 10 2022. doi: 10.1007/s12065-022-00783-2.

[12] W. Hämäläinen. Datasets. http://www.cs.uef.fi/ whamalai/datasets.html, 2019. Accessed: 2019-08-10.

[13] J. Han, J. Pei, and Y. Yin. Mining frequent patterns without candidate generation: A frequent-pattern tree approach. In *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, pages 1–12, 2000.

[14] K.-H. Han and J.-H. Kim. Genetic quantum algorithm and its application to combinatorial optimization problem. volume 2, pages 1354 – 1360 vol.2, 02 2000. ISBN 0-7803-6375-2. doi: 10.1109/CEC.2000.870809.

[15] K. Heraguemi, N. Kamel, and H. Drias. Multi-swarm bat algorithm for association rule mining using multiple cooperative strategies. *Applied Intelligence*, 45, 12 2016. doi: 10.1007/s10489-016-0806-y.

[16] A. S. Hesar and M. Houshmand. A memetic quantum-inspired genetic algorithm based on tabu search. *Evolutionary Intelligence*, 17(3):1837–1853, June 2024. ISSN 1864-5917. doi: 10.1007/s12065-023-00866-8.

[17] J. H. Holland. Genetic algorithms. *Scientific American*, 267(1):66–73, 1992. ISSN 00368733, 19467087. URL http://www.jstor.org/stable/24939139.

[18] H. Jiawei and M. Kamber. *Data Mining: Concepts and Techniques*. Morgan Kaufmann Publishers, 2011.

[19] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proceedings of ICNN'95 - International Conference on Neural Networks*, volume 4, pages 1942–1948 vol.4, Nov 1995. doi: 10.1109/ICNN.1995.488968.

[20] J. Khudair Madhloom, H. N. Abd Ali, H. A. Hasan, O. A. Hassen, and S. M. Darwish. A quantum-inspired ant colony optimization approach for exploring routing gateways in mobile ad hoc networks. *Electronics*, 12(5), 2023.

ISSN 2079-9292. doi: 10.3390/electronics12051171. URL https://www.mdpi.com/2079-9292/12/5/1171.

[21] D. Konar, S. Bhattacharyya, K. Sharma, S. Sharma, and S. R. Pradhan. An improved hybrid quantum-inspired genetic algorithm (hqiga) for scheduling of real-time task in multiprocessor system. *Applied Soft Computing*, 53:296–307, 2017. ISSN 1568-4946. doi: https://doi.org/10.1016/j.asoc.2016.12.051. URL https://www.sciencedirect.com/science/article/pii/S1568494616306949.

[22] D. Konar, K. Sharma, V. Sarogi, and S. Bhattacharyya. A multi-objective quantum-inspired genetic algorithm (mo-qiga) for real-time tasks scheduling in multiprocessor environment. *Procedia Computer Science*, 131:591–599, 2018. ISSN 1877-0509. doi: https://doi.org/10.1016/j.procs.2018.04.301. URL https://www.sciencedirect.com/science/article/pii/S1877050918306811. Recent Advancement in Information and Communication Technology:.

[23] S. Kuo, J. Shen, C. Liu, and Y. Chou. Hybrid quantum-inspired evolutionary neural networks for intrusion detection system. In *2024 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2801–2806, 2024. doi: 10.1109/SMC54092.2024.10831208.

[24] J. C.-W. Lin, L. Yang, P. Fournier-Viger, J. M.-T. Wu, T.-P. Hong, L. S.-L. Wang, and J. Zhan. Mining high-utility itemsets based on particle swarm optimization. *Engineering Applications of Artificial Intelligence*, 55:320–330, 2016. ISSN 0952-1976. doi: https://doi.org/10.1016/j.engappai.2016.07.006. URL https://www.sciencedirect.com/science/article/pii/S0952197616301312.

[25] S. Mirjalili and A. Lewis. The whale optimization algorithm. *Advances in Engineering Software*, 95:51 – 67, 2016. ISSN 0965-9978. doi: https://doi.org/10.1016/j.advengsoft.2016.01.008. URL http://www.sciencedirect.com/science/article/pii/S0965997816300163.

[26] H. Moulai. Frequent itemsets mining using new quantum inspired elephant swarm algorithm. In H. Drias and F. Yalaoui, editors, *Quantum Computing: Applications and Challenges*, pages 65–76, Cham, 2024. Springer Nature Switzerland. ISBN 978-3-031-59318-5.

[27] H. Moulai and H. Drias. Association rule mining using new discrete elephant swarm approaches. *Expert Systems*, 40 (2):e13159, 2023. doi: 10.1111/exsy.13159. URL https://doi.org/10.1111/exsy.13159.

[28] E. Rieffel and W. Polak. An introduction to quantum computing for non-physicists. *ACM Computing Surveys*, 32(3):300–335, 2000.

[29] L. Siddiqui, A. Mani, and J. Singh. Improving design accuracy of a finite impulse response fractional order digital differentiator filter using quantum-inspired evolutionary algorithm. In *2024 14th International Conference on Cloud Computing, Data Science Engineering (Confluence)*, pages 142–147, 2024. doi: 10.1109/Confluence60223.2024.10463238.

[30] M. Sudip. Elephant swarm water search algorithm for global optimization. *Sādhanā*, 43:1–21, 2018. doi: 10.1007/s12046-017-0780-z.

[31] H. Talbi and A. Draa. A new real-coded quantum-inspired evolutionary algorithm for continuous optimization. *Applied Soft Computing*, 61:765–791, 2017. ISSN 1568-4946. doi: https://doi.org/10.1016/j.asoc.2017.07.046. URL https://www.sciencedirect.com/science/article/pii/S1568494617304660.

[32] E. M. Tapia. Activity recognition in the home setting using simple and ubiquitous sensors. http://courses.media.mit.edu/2004fall/mas622j/04.projects/home/, 2018.

[33] X. Wang, Z. Geng, L. Chen, and Y. Han. Improved dragonfly optimization algorithm based on quantum behavior for multi-objective optimization of ethylene cracking furnace. *Swarm and Evolutionary Computation*, 88:101607, 2024. ISSN 2210-6502. doi: https://doi.org/10.1016/j.swevo.2024.101607. URL https://www.sciencedirect.com/science/article/pii/S2210650224001457.

[34] X. Yan, C. Zhang, and S. Zhang. Genetic algorithm-based strategy for identifying association rules without specifying actual minimum support. *Expert Systems with Applications*, 36(2, Part 2):3066–3076, 2009. ISSN 0957-4174. doi: https://doi.org/10.1016/j.eswa.2008.01.028. URL https://www.sciencedirect.com/science/article/pii/S0957417408000195.

[35] C. Yu, Z. Cai, X. Ye, M. Wang, X. Zhao, G. Liang, H. Chen, and C. Li. Quantum-like mutation-induced dragonfly-inspired optimization approach. *Mathematics and Computers in Simulation*, 178:259–289, 2020. ISSN 0378-4754. doi: https://doi.org/10.1016/j.matcom.2020.06.012. URL https://www.sciencedirect.com/science/article/pii/S0378475420302147.

Hadjer Moulai was born in Algiers on 1993. She earned a Ph.D degree in computer science specializing in artificial intelligence from the University of Science and Technology Houari Boumediene in 2020. She is a permanent lecturer at USTHB since 2021, and an active researcher in the laboratory of research in artificial intelligence LRIA. Her research spans various fields such as machine learning, optimization algorithms, data mining and quantum computing.

Anti-corrosion performance of PANI and POEA coatings applied into 13% Cr stainless steel at 3.5% NaCl solution with neutral pH

Nacer Mounir, Malha Nazef, Zakaria Baka, and Hamid Yousfi

Abstract−The doped PANI and POEA polymers were incorporated into a commercial epoxy resin with xylene solvent to develop effective anticorrosion coatings applied to 13% Cr stainless steel. The produced epoxy resin coatings were systematically characterized by FTIR, XRD, SEM, AFM, and immersion tests. The corrosion resistance performance of the ZP, PANI, and POEA coatings was evaluated through potentiodynamic polarization, electrical impedance spectroscopy, open circuit potential (OCP), and linear polarization resistance (LPR) measurements in a long-term exposure to a 3.5% NaCl solution at pH 6.5. The corrosion resistance of the commercial ZP coating was significantly improved by the PANI and POEA coatings, and their performance was correlated with the highly enhanced protection of the passivation barrier provided. The higher corrosion resistance observed in POEA coatings in acidic ambient conditions is explained by the auxiliary cathodic reaction between the conducting emeraldine PANI and the non-conducting leuco-PANI. The suggested manufacturing method for storage tanks can be used to simultaneously replace the substrate (carbon steel) and the conventional anticorrosion coating based on zinc phosphate.

***Keywords*−Conducting polymers, PANI Coating, POEA coating, ZP Coating, 3.5% NaCl solution.**

NOMENCLATURE

PANI	Polyaniline.
POEA	Poly(Ortho-Ethoxyaniline).
ZP	Zinc Phosphate.
SS	Stainless Steel.

I. INTRODUCTION

Comparative research into two types of steel substrates holds significant importance in the field of materials science, as it enables a thorough exploration of the properties, performance, and potential applications of different steel grades. Steel, being a fundamental material across numerous industries, offers diverse compositions and characteristics to cater to specific requirements. Consequently, this study aims to underscore the distinctions between two highly significant steel types: AISI 1010 carbon steel and AISI 420 martensitic stainless steel.

On one hand, AISI 1010 carbon steel, primarily composed of iron and carbon, finds widespread usage in a variety of industrial and structural applications, as documented by Mirmohseni in 2000 [1] and Ebrahimi in 2017 [2]. On the other hand, AISI 420 martensitic stainless steel, characterized by its elevated chromium (13% Cr) and carbon content, offers unique features such as increased corrosion resistance and the ability to harden through heat treatment, as noted by Zhong in 2006 [3]. This thorough comparison aims to investigate the advantages and limitations of each steel type, taking into account criteria like mechanical strength, corrosion resistance, and durability. The research focuses on the specific properties of these steels and aims to provide essential information for well-informed material selection across various sectors, including manufacturing, construction, marine engineering, and aerospace. The ultimate goal is to aid designers and engineers in selecting the most suitable material for each application, thereby enhancing efficiency, durability, and overall performance in diverse industrial contexts, including the oil and technology sectors.

The study is particularly interested in evaluating the effectiveness of coatings on AISI 1010 carbon steel and AISI 420 martensitic stainless steel, as it relates to materials protection. Coatings play a pivotal role in enhancing the performance and longevity of these substrates, especially in demanding environments where corrosion, abrasion, and other forms of deterioration are primary concerns, as observed by Xu in 2016 [4]. Protective coatings serve as an effective shield against corrosive agents for AISI 1010 carbon steel, which is more susceptible to corrosion due to its carbon content. Coatings such as epoxy paints, oxide-reinforced polymers, and zinc-based anticorrosion coatings have demonstrated their ability to safeguard carbon steel by isolating it from external elements that may cause harm, as evidenced by Sazou in 2007 and Ozyilmaz in 2010 [5, 6]. These coatings not only delay the corrosion process but can also enhance abrasion resistance, thereby improving durability in challenging environments.

In the case of AISI 420 martensitic stainless steel, which already possesses corrosion resistance due to its chromium content, coatings can further augment this resistance, particularly under aggressive conditions, as highlighted by

Manuscript received October 20, 2024; revised July 9, 2025.

N. Mounir (Corresponding author) is with the laboratories LSGM, Department of Metallurgy and laboratory LSTE, Department of Quality, Health, Safety, Environment and Industrial Risk Management, Ecole Nationale Polytechnique, El Harrach 16200 Algiers, ALGERIA (e-mail:nacer.mounir@g.enp.edu.dz).

M. Nazef is with the laboratory of Corrosion, Department of Corrosion, Division Laboratories Exploration and Production SONATRACH (Ex CRD), Boumerdes, ALGERIA (e-mail:malha.allaoua@sonatrach.dz).

Z. Baka is with the laboratory of Green and Mechanical Development (LGMD), Department of Mechanical Engineering, Ecole Nationale Polytechnique, El Harrach 16200 Algiers, ALGERIA (e-mail: zakaria.baka@g.end.edu.dz).

H. Yousfi is with the LSTE, Laboratory of Quality, Health, Safety, Environment and Industrial Risk Management, Ecole Nationale Polytechnique, El Harrach 16200 Algiers, ALGERIA (e-mail: hamid.yousfi@g.enp.edu.dz).

Digital Object Identifier (DOI): 10.53907/enpesj.v5i1.296

2716-912X © 2025 Ecole Nationale Polytechnique

Mrad in 2009 [7]. Epoxy coatings offer additional protection by forming a barrier against corrosive agents while preserving the innate characteristics of martensitic stainless steel. In summary, the effectiveness of coatings on AISI 1010 carbon steel and AISI 420 martensitic stainless steel lies in their ability to extend the operational lifespan of these materials, reduce maintenance costs, and ensure long-lasting performance in environments where corrosion and abrasion challenges are prevalent. Customizing coatings to the properties of each steel type maximizes the benefits of each material while overcoming its limitations, thereby creating comprehensive and effective protective solutions.

Our study aims to conduct an in-depth comparative assessment of the effectiveness of coatings applied to two distinct substrates: AISI 1010 carbon steel and 420SS martensitic stainless steel. Our primary objective is to differentiate the performance of specific coatings and evaluate the impact of the stainless steel substrate on the performance of epoxy coatings. Additionally, we aim to anticipate the potential advantages of using low-chromium (13%) stainless steels in oil installations exposed to saline environments (saturated with NaCl) and aggressive acidic pH levels. This study's purpose is to provide essential information to guide material selection and optimize coating practices in conditions where corrosion and hostile environments are major considerations.

II. Materials and Methods

A. Materials

1. Products used

In our quest to develop innovative coatings, we ventured beyond the traditional realm of anti-corrosion elements by replacing zinc phosphate with highly conductive polymers: Polyaniline (PANI) and Poly (ortho-ethoxyaniline) (POEA).

To formulate our coatings, we harnessed the potential of two distinct semi-conducting polymer powders: PANI-EB (Polyaniline doped with an acid) and POEA (Poly(ortho-ethoxyaniline)). These polymers exhibit remarkable conductivity and corrosion resistance properties.

We carefully applied three unique coatings onto stainless steel surfaces, each incorporating a specific anti-corrosion agent:

- Coating N°1: Epoxamine in a commercial white shade (referred to as ZP), fortified with a 7.5wt.% by zinc phosphate anti-corrosion agent.
- Coating N°2: A captivating pearly green epoxy, enriched with a 3wt.% by PANI-doped anti-corrosion agent.
- Coating N°3: An elegant navy-blue epoxy, fortified with a 3wt.% by POEA anti-corrosion agent.

These coatings were meticulously formulated, taking into account the desired properties and optimal concentrations of the PANI and POEA agents. The composition details can be found in Section *A.1.4*.

For a comprehensive understanding of the substrate, we conducted XRF analysis, providing valuable insights into the chemical composition of the stainless steel. The results are presented in Table I, aiding in the evaluation of the coating's compatibility and performance.

Through our innovative approach, we aim to revolutionize the field of corrosion protection, offering coatings that not only safeguard against degradation but also harness the unique properties of conductive polymers to enhance the overall durability and longevity of the coated surfaces.

Table.I
Chemical composition of 13% Cr stainless steel substrate by XRF.

Chemical elements	C	Mn	Si	Cu	Ni	Cr	Fe
Massfraction (%)	0.38	0.49	0.43	0.19	0.21	13.08	85.22

2. Doped PANI and POEA synthesis

To prepare the doped Polyaniline (PANI), we followed a polymerization procedure utilizing a chemical oxidizing solution as described in previous studies[8-13]. The synthesis involved gradually adding the oxidizing solution, consisting of 400ml of 1M HCl, to a solution containing 0.2 moles of the monomer aniline. This reaction mixture was stirred slowly for 4 hours.

After the reaction, the resulting PANI-ES salt was washed with acetone. Subsequently, 0.1M NH_2OH was added to the salt with vigorous stirring for 2 hours. The obtained product, PANI-EB, was then rinsed with distilled water and filtered. To remove any remaining moisture, it was dried in a vacuum oven at 40°C for 48 hours [11].

To facilitate the doping of Polyaniline, the salt obtained from the previous step was dissolved in a solution of 4M concentrated sulfo-5-salicylic acid [14]. The solubilization process involved stirring the mixture for 12 hours to ensure thorough doping. Subsequently, the mixture was dried in a vacuum oven at 40°C for 48 hours to remove any residual moisture. The resulting doped PANI was then carefully stored in an airtight bottle to maintain its properties.

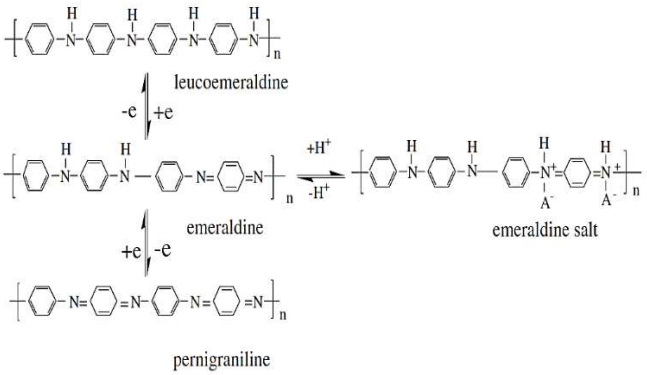


Fig.1: The PANI transformation forms [3, 15].

The doping mechanism used to obtain PANI is illustrated in Figure 1, providing a visual representation of the process.

The synthesis of Poly(o-ethoxyaniline) was carried out following a method described in previous works [16, 17]. In a reaction vessel, 0.6 moles of ethoxyaniline monomer (ORMECON, Germany) were combined with 400ml of 1M HCl and stirred for 2 hours. Then, 0.4 moles of ammonium peroxy-disulphate were carefully added to the mixture under a vacuum atmosphere, maintaining a temperature of 28°C.

To ensure adequate electrical conductivity, the solution's pH was adjusted to 8 by adding 1M NaOH. The resulting product was then mixed with concentrated sulfo-5-salicylic acid [14] and stirred for 24 hours, as depicted in Fig.2. Finally, the mixture was dried in a vacuum oven at room temperature for 48 hours [11]. This procedure allowed for the successful synthesis of Poly(o-ethoxyaniline) (POEA), a crucial component in our research, with desired properties for the development of our innovative coatings.

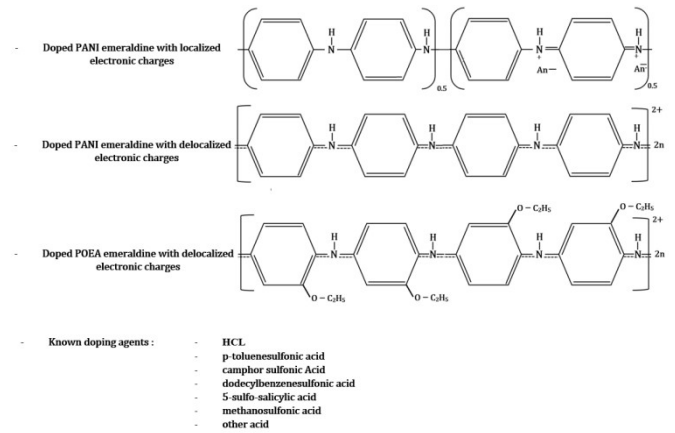


Fig. 2: The Polyaniline and Poly (O-Ethoxyaniline) doped with acid [18-21].

3. *Powder Grinding*

Following the synthesis of the two polymers, additional steps were taken to achieve the desired fineness required for the development of primer paints, adhering to standards such as ISO 2808, ASTM D4414, and D1212. For instance, in the case of the white Epoxamine commercial paint (ZP), the typical acceptable fineness falls within the range of 5 to 7 on the Hegman scale, as determined by a grinding fineness gauge (ERICHSEN Model 232 ASTM D 1210 ISO 1524, ENAP ALGERIE).

To meet these specifications, a mechanical grinding process was initially employed, followed by an automatic grinding process. Through these grinding procedures, we were able to achieve a caliber rating of 4 to 5 on the Hegman scale, corresponding to particle sizes ranging from 40 to 30 µm.

By attaining the desired particle size distribution, we ensured that the resulting powders were suitable for the formulation of primer paints, meeting the required industry standards.

4. *Formulation and Development of Epoxy Coatings*

4.1. Formulation of Epoxy Coatings

In order to assess the corrosion resistance properties of the epoxy coatings, we developed three distinct coatings. To carry out this formulation, we collaborated with the main physicochemical laboratory at the national paint company (ENAP, Algeria). The following protocol was followed:

- Preparation of the commercial coating (referred to as ZP). We formulated a white Epoxamine coating containing 7.5% of zinc phosphate ($Zn_3(PO_4)_2$) [21].
- Formulation of two other coatings: These coatings were developed by replacing the anti-corrosion agent of the commercial formula with two different agents known as PANI-EB and POEA. The PANI-EB coating contained 3 wt.% of PANI-EB, while the POEA coating contained 3 wt.% of POEA. Detailed formulation parameters can be found in Table II [21].

During the formulation process, great care was taken to ensure accurate measurements of the ingredients. Precise weighing was carried out using an analytical balance with a 4-digit decimal scale (SARTORIUS 220g, ENAP ALGERIA).

By adhering to rigorous formulation procedures, we aimed to create epoxy coatings with specific anti-corrosion properties, enabling comprehensive evaluation of their performance against corrosion in subsequent testing and analysis [21].

The followingtable presents the specific formulation details for the anti-corrosion epoxy coatings developed in this study.

Table.II
Formulation of Anti-Corrosion Epoxy Coatings [21].

Components	Commercial paint (ZP)	PANI coating	POEA coating
	$Zn_3(PO_4)_2$) (wt.%)	PANI-ES (wt.%)	POEA (wt.%)
Epoxy resin	30.50	30.50	30.50
Anti-corrosion agent	7.50	3.00	3.00
Additives	5.70	5.70	5.70
Fillers	19.50	23.50	23.50
Pigments	21.00	21.00	21.00
Solvents & Xylene	7.00 + 8.800	7.00 + 9.300	7.00 + 9.30
Total	100	100	100

4.2. Coating Elaboration (Operating Procedure)

Prior to initiating the preparation of the primer coatings, compatibility tests were conducted between the components and the two conductive polymers, PANI-ES and POEA. Solubility tests were performed using Xylene [22], while also considering the strong solubility of the polymers in DMF, DMSO, and alcohols [14, 23-25].

To formulate 100g of paint, a metal paint can with a capacity of 250g was utilized. The required mass of the anti-corrosion agent (zinc phosphate or conductive polymers) was accurately weighed using a KERN ABS 220-4N balance. This agent was then dissolved in 50 wt.% xylene solvent. Subsequently, the epoxy resin, additives, fillers, pigments, and remaining solvents were gradually added to the mixture with moderate stirring using a RAKUTEN 2100W mixer [21].

To ensure proper grinding, mixing, and achieve the desired fineness of the final product, 50 wt.% glass beads were incorporated into the mixture. The entire mixture was then placed in an automatic grinder (VM2 RED DEVIL5400 vibrator, ENAP Algeria) for a duration of 90 minutes. Following the grinding process, the mixture was filtered, and a Versamid 115/70-type hardener was added at a paint-to-hardener ratio of 93:7 [21].

Subsequently, the metal substrates, in disc form with a radius of 3.5 cm, were prepared using the stainless steel material specified in Table I. The preparation involved primary pickling. Secondary stripping was performed using a mixture of two acids (0.1N H_2SO_4 and 0.1N HCl), and the surface of the samples was degreased with acetone. For this comprehensive coating elaboration procedure, we ensured the preparation of well-formulated coatings and properly prepared metal substrates for subsequent testing and analysis of corrosion behavior [21].

Lastly, the three paints were applied onto the metal substrates using a manual projection torch. After the application, wet coatings with a thickness ranging from 100 to 150 µm were obtained and measured using a thickness gauge (ATQC ISO 2808, ENP Algeria). Once dried, the coatings reached a thickness of approximately 40 to 70 µm, as measured by a thickness gauge (POSITECTOR 600 ISO 2178, ENAP Algeria).

B. Methods for Coating Characterization

1. Surface Characterization

The surface of the coatings was subjected to three different types of observation to gain comprehensive insights. Initially, an optical microscope was employed to evaluate the overall quality of the surface obtained. Subsequently, more detailed observations were conducted using scanning electron microscopy (SEM) to identify the finest particles present on the coating surface. Additionally, atomic force microscopy (AFM) mappings were performed using specific types and brands of AFM equipment to assess the roughness of the coatings. These characterization methods provided valuable information about the surface properties and morphology of the coatings.

2. Structural and Micrographical Characterization

The composition of the three coatings was analyzed using non-destructive experimental methods, including infrared Fourier-transform attenuated total reflection spectroscopy (FTIR-ATR) and SEM observation. This characterization process involved analyzing both the initially doped PANI and POEA powders, as well as the final coatings. These methods provided valuable insights into the structural and micrographical properties of the coatings, allowing for a comprehensive understanding of their composition and characteristics.

3. Electrochemical Measurements

To evaluate the corrosion resistance of the coatings, electrochemical measurements were performed. The sample preparation involved cutting a coated sample into a small disc with a diameter of 1.25 cm. This diameter was chosen to fit inside a flat specimen holder kit specifically designed for conducting electrochemical immersion tests. The disc's surface area in contact with the 3.5% NaCl electrolyte was approximately 1 cm², providing a defined attack surface for corrosion assessment.

The electrochemical corrosion measurements were conducted using a VersaSTAT 3 potential stat (Princeton Applied Research, ENP Alger). The electrolyte solutions were prepared to simulate seawater, consisting of 3.5 wt.% NaCl [10, 11, 26, 27], with a pH 6.5.

The measurements were taken at a room temperature of 25 ±1°C. A Reference electrode (Ag, AgCl/KCl (3.5M)) with a standard potential of E_0 = +0.205V [28] and a pH controller (HANNA Instruments 2211 pH/ORP meter, ENP Algiers) were used.

The obtained results, highlighting the variations in iron concentration over the immersion period, are presented in Table III.

Prior to each electrochemical impedance spectroscopy (EIS) measurement, open potentials (OCP) and linear polarization resistance (LRP) were measured. EIS measurements were performed within the frequency range of 0.05 Hz to 50 kHz, with an alternating amplitude potential of 50 mV [26, 29]. All tests were conducted on a minimum of five samples, and the results demonstrated good reproducibility.

In addition, potential dynamic corrosion polarization tests (Tafel tests) were performed to assess the protective behavior of the coatings. These tests involved observing the anodic displacement of the corrosion potential, current density, and determining the corrosion rate (E corrosion) of the coatings [11].

4. Immersion Test

In our study, we conducted an immersion test to investigate the performance of samples coated with three different paints: POEA, PANI, and ZP. The samples had dimensions of 40×20×0.01cm^3. They were submerged in a 3.5% NaCl solution with a pH of 6.5 for a total duration of 60 days.

To conduct the immersion tests, we implemented a systematic approach. At specific time intervals (5, 10, 15, 20, 30, 45, 60, 90, 120, 180, and 240 days), the samples were carefully weighed (KERN ABS 220-4N) to monitor any changes (Fig.3b). Additionally, after each sampling session, 30 ml vials were collected (Fig.3a). These vials were crucial for analyzing the progressive loss of iron concentration in the solution (Fe^{+2} and Fe^{+3}) as the immersion time increased. The analysis of iron concentration was carried out using two instruments: the PerkinElmer Optima8000 for ECP Test and the HACH DR 1900 for UV-visible Test.

Fig. 3: (a) The solution collection vial 30ml, (b) the electronic scale.

III. Results and Discussions

A. Physical and Chemical Results

We measured the contact angle (CA) of the electrolyte on the surface of the substrate coated with the three types of epoxy: ZP, PANI, and POEA (see Fig.4). It was observed that this angle decreases rapidly for ZP, while it remains more stable for PANI and POEA during immersion tests in a 3.5% NaCl solution. This indicates the degradation of the ZP coating surface, which is initially smooth and hydrophobic but becomes hydrophilic after 60 days of immersion (around 63°). In contrast, the stability of the CA (at 78°) for PANI (even with a hydrophilic surface) and POEA (smooth and hydrophobic surface) is maintained (approximately 83°). The results show that the ZP coating degrades (from 96° to 62°) due to the removal of zinc (sacrificial electrode), while the PANI and POEA coatings serve as excellent anticorrosion barriers against the aggressive electrolyte.

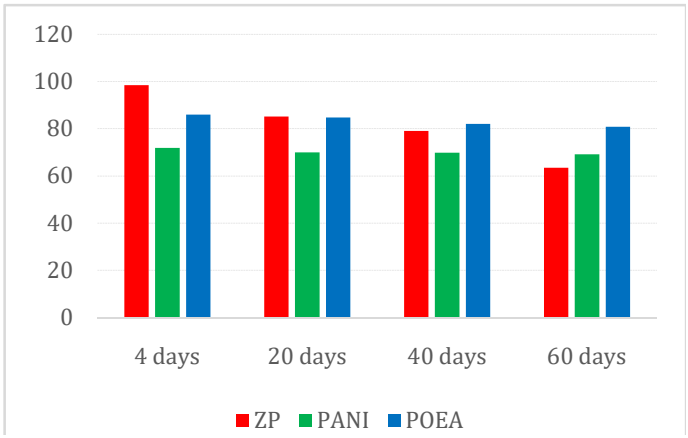


Fig.4: Variation in contact angle of ZP, PANI and POEA epoxy coatings as a function of immersion time in 3.5 wt.% NaCl solution.

B. Corrosion Studies Effects

1. Open Circuit (OCP) and Linear Polarization Resistance (R_P)

The equilibrium potential (OCP) of ZP, PANI, and POEA epoxy coatings varies with immersion time (see Fig. 5). It is observed that the ZP coating undergoes a gradual decrease, from -250 mV after 4 days of immersion to -480 mV after 15 days in the saline solution. Similarly, the potential of PANI decreases from -120 mV (after 4 days) to -380 mV after 40 days in the marine solution. In contrast, the POEA shows an increase from -160 mV (after 4 days) to -50 mV after 40 days, followed by a slow decrease to -220 mV by the end of the test.

The polarization resistance results for the epoxy coatings (see Fig. 6) show that ZP experiences a significant drop, from 1 MΩ (after 4 days of immersion) to 100 kΩ after just 15 days in a 3.5% NaCl solution. However, the PANI and POEA coatings maintain stable resistance after 40 days of immersion (approximately 2.5 MΩ and 1.2 MΩ, respectively), before gradually decreasing to 830 kΩ after 200 days of exposure to an aggressive marine solution.

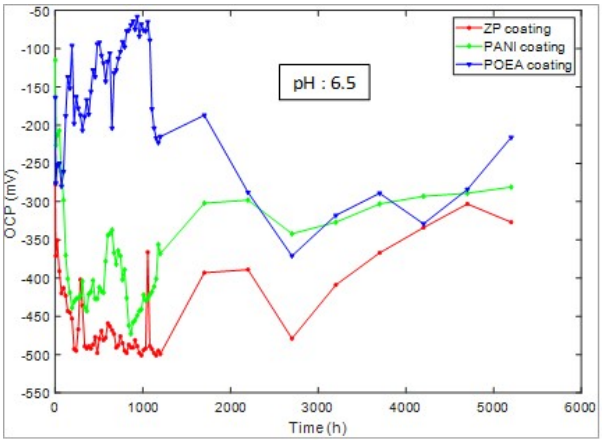


Fig.5: Open circuit potential of ZP, PANI and POEA epoxy coatings applied into 420SS with immersion time in 3.5 wt.% NaCl solution.

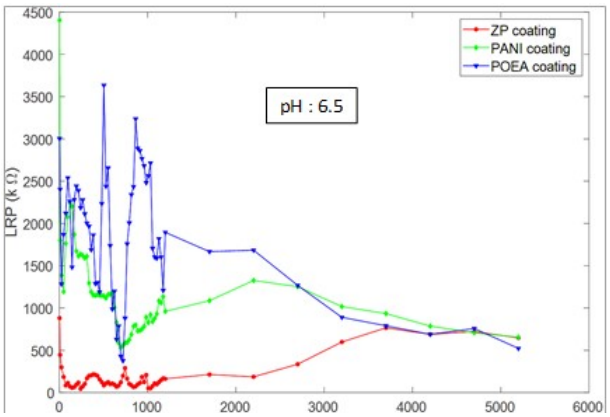


Fig.6: R_P evolution of ZP, PANI and POEA epoxy coatings applied into 420SS with immersion time in 3.5 wt.% NaCl solution.

2. The Potential Dynamic Polarization of Corrosion

The ZP, PANI, and POEA coatings are carefully applied onto a 13% Cr stainless steel surface, and they are subjected continuous immersion in 3.5% NaCl solutions with a pH of 6.5. The insightful results obtained from this experiment are vividly depicted in Fig.7. Notably, both the PANI and POEA coatings exhibit remarkable capabilities in safeguarding and reinforcing the formation of a passivation layer, specifically in the form of chromium oxide [3].

At the beginning of the test (after 4 days' immersion), the corrosion rates of PANI and POEA are much higher than those of ZP, and the corrosion current density of POEA and PANI is much lower than that of the ZP coating: -52 and -72 to -178mV, 1.8×10^{-11} and 1.23×10^{-10} to 33.71×10^{-9} A.cm^{-2} respectively (see Table III and Fig.7). It can be affirmed that, initially, both POEA and PANI coatings have stronger protective characteristics than ZP coatings.

However, after a duration of 60 days immersed in a 3.5% NaCl solution with a pH of 6.5 and at ambient temperature, all three coatings experienced substantial attack by the solution. This led to notable degradation in the intrinsic characteristics of each coating, particularly in the case of the ZP coating, which exhibited a significant decrease of -350mV compared to its initial immersion state. On the contrary, the POEA and PANI coatings demonstrated a relatively lower corrosion potential decrease, with values of -80mV and -180mV (respectively) compared to their initial potentials. Furthermore, the corrosion current density witnessed an increase to 4.60×10^{-7}, 3.10×10^{-8}, and 1.1×10^{-9}A.cm^{-2}for the ZP, PANI, and POEA coatings, respectively.

Based on the findings, it can be concluded that the POEA, PANI, and ZP coatings, when applied to 13% Cr stainless steel and immersed in a marine solution, exhibit varying levels of anti-corrosion performance. The POEA coating demonstrates a strong resistance to corrosion, the PANI coating exhibits a moderate level of protection, while the ZP coating shows a relatively weaker ability to prevent corrosion.

The evaluation of anti-corrosion performance involves the calculation of coating efficiency known as GAP (η). This calculation serves a dual purpose: firstly, to measure the corrosion rate of uncoated stainless steel immersed in a salt solution, and secondly, to assess the anti-corrosion effectiveness of our coatings [3]. The efficiency factor is determined by comparing the corrosion potential difference between the substrate with and without the coating, as defined

by the following formula (see equation 1) [3]:

$$\eta = (1 - \frac{I_{corr}}{I^0_{corr}}) \times 100\% \quad (1)$$

I_{Corr} andI^0_{Corr} represents the corrosion current density values for the working electrodes (uncoated stainless steel) and (coated stainless steelwith ZP, PANI and POEA), respectively. The results are shown in Table III.

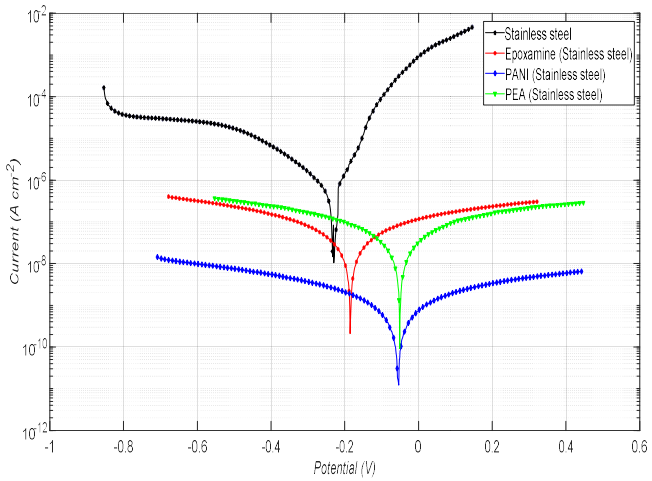


Fig.7: Tafel curves (a) after 4 days of immersion in 3.5% NaCl solution.

Table. III
Electrochemical parameters of MS, ZP, PANI, POEA Coating onto stainless steel in 3.5%NaCl solution.

Electrode	Immersion (days)	$E_{Corr.}$ (mV)	$i_{Corr.}$ (ACm^{-2})	β_a ($Vdec^{-1}$)	β_c ($Vdec^{-1}$)	η (%)
MS	4	-297.62	76.62×10^{-8}	0.088	0.157	-
	60	-414.51	5.87×10^{-6}	0.176	0.323	-
ZP	4	-178.38	19.23×10^{-9}	0.181	0.181	97.49
	60	-438	96.73×10^{-9}	0.137	0.137	87.37
PANI	4	-137.81	1.68×10^{-10}	0.225	0.225	99.98
	60	-247.81	1.31×10^{-9}	0.354	0.354	99.82
POEA	4	-80.63	0.18×10^{-10}	0.161	0.161	99.99
	60	-107.91	1.11×10^{-10}	0.081	0.081	99.98

3. *Electrochemical Impedance Spectroscopies*

The anti-corrosion properties of the coatings (POEA, PANI, andZP) were characterized experimentally by EIS measurements. The Nyquist curves to the various coatings during immersion of 60 days in a 3.5% NaCl solution at pH 4.5 are shown in Fig.6. The results obtained by EIS test reveal that, at low frequencies (0.05 Hz), $|Z|_{0.05Hz}$ (or R_P) values are very high [30-32]. After 2 days immersion, the POEA and PANI coatings show $|Z|_{0.05Hz}$ values of the order to but it is in opposition to the ZP coating, which is in the order to 8.92×10^5 Ω. This value of $|Z|_{0.05Hz}$ is used as a quantitative classification index for the performance of coatings against the passage of electronic charges from a steel to the corrosive solution. This high impedance value implies that the barrier of coatings against corrosion is very effective [32, 33]. The plots observed were predominantly characterized by a depressed semicircular shape, with the resistance value corresponding to the diameter of this semicircle being equivalent to the polarization resistance (R_P).

In addition, the coatings have same appearance in the immersion test, which proves that they have the same R_P behavior towards the solution with different POEA and PANI coatings produced very high values, compared to the ZP-based coating, which proves the electron exchange process between steel and solution is so complex and negligible.

After 60 days of immersion in 3.5% NaCl solution (Fig.8), we can see the difference between POEA, PANI and ZP-based coatings with a significant decrease in R_P for ZP unlike the POEA coating, and the appearance of second half-arc, larger than the first day of immersion. The obtained results from the EIS curves measured are presented in Table IV. At this time of immersion, the impedance values for POEA 5.6×10^6Ω) and PANI (5.1×10^6Ω) remain high compared with ZP (8.3×10^5Ω). This may explain their anti-corrosion performance due to their excellent barrier effect [28, 34, and 35]. On the other hand, the barrier effect of ZP coatings has been degraded by electron migration from the surface by a reaction of Zn with Zn^{2+} as the coating progresses in depth [36].

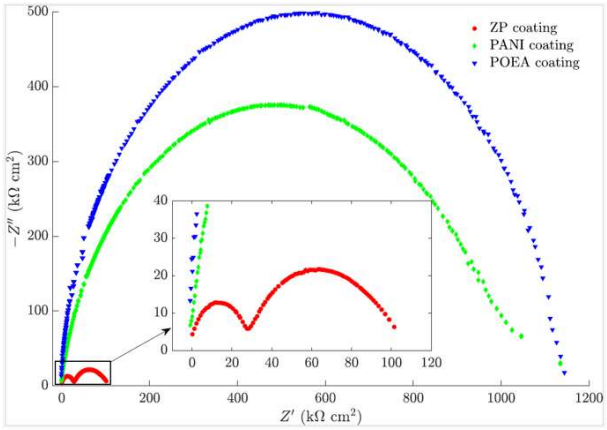


Fig.8: Nyquist plots (EIS curves) of ZP, PANI and POEA coatings after 60.days of immersion in 3.5% NaCl.

The electrical impedance of PANI coating is high compared to that of commercial coating, and it always remains at the same pace; which explains the resistance of the coating to degradation as a function of immersion time. The real and imaginary impedances of the PANI and POEA coatings remain very high during the immersion time, contrary to the commercial coating.

According to Fig.6, after adjustment with Zsimpwin soft word, the obtained equivalent circuits are shown in Fig.9. In the circuits presented, R_S, R_{pore}, C_C, C_{dl} and R_{ct} signify such as solution resistance, coating pore resistance, capacitance of the coating, double layer capacitance resulting from the corrosion reaction at interface of the carbon steel substrate and charge transfer resistance,respectively.

Table. IV
Calculated electrochemical parameters for ZP,PANI and POEA coatings after 60 days of immersion in 3.5% NaCl solution with a pH6.5.

Epoxy	RS (Ω cm^2)	R_{pore} (kΩ cm^2)	C_{PE} (μF cm^{-2}) Y0	N0	C_{dlE} (μF cm^{-2}) Y1	N1	R_{ct} (Ω cm^2)	R_P (Ω cm^2)
ZP	0.0122	15.97	12.34	0.96	1.48	0.91	1.71×10^4	1.72×10^4
PANI	0.0118	158.62	1.3×10^{-4}	0.98	1.1×10^{-3}	0.88	8.59×10^5	8.60×10^5
POEA	0.0120	238.74	1.1×10^{-4}	0.92	3.8×10^{-4}	0.94	1.11×10^6	1.12×10^6

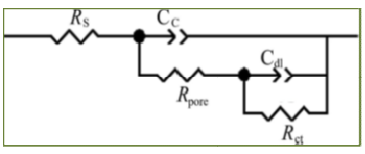


Fig.9: Equivalent circuits of the coatings, ZP, PANI and POEA after 60 days of immersion in 3.5% NaCl.

C. Micrographic Corrosion Analysis of Coating

Microstructural observations by SEM reveal that the effect of corrosion on the three coatings (ZP, PANI, and POEA) is inversely proportional in a significant manner (see Fig. 10). It is observed that corrosion heavily affects the ZP coating, while this effect is less on PANI and almost negligible on the POEA film. This highlights the superior anticorrosion barrier performance of the PANI and POEA coatings compared to the ZP coating.

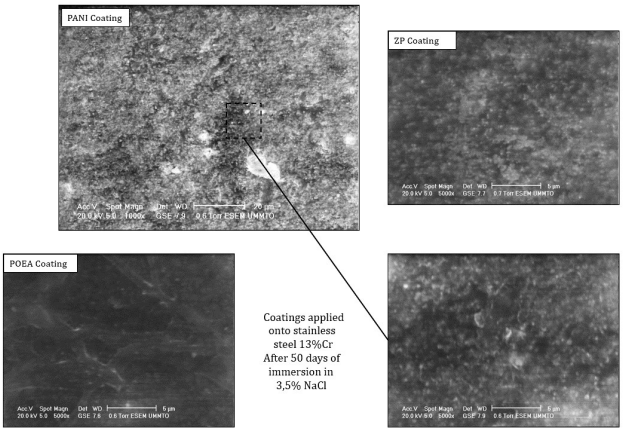


Fig.10: SEM images of the facial-section of ZP, PANI and POEA epoxy.coatings after 50 days of immersion in 3.5% NaCl.

D. Immersion Test and the Fe Dissolutions Analyzing

The specimens, cut in 40×80×0.1mm rectangles, are immersed in a 3.5% NaCl solution with a neutral pH while the immersion time is growing. The solutions are sampled after 10, 15, 20, 30, 40, 60 and 75 days. Optical observations on the samples showed that there was no trace of rust on the substrates coated with ZP, PANI and POEA. The 30ml solutions taken from each stage were analyzed to determine the total mass of iron in saline solution. The results obtained after OCP and Uv-visible analysis are shown in the following table:

Table. V
Fe dissolved after immersion in 3.5 wt.% NaCl with pH 6.5.

	Total concentration of Fe dissolved ($\times 10^{-6}$ g l^{-3})							
Days	10	15	20	30	40	50	60	75
ZP	0.011	0.016	0.019	0.021	0.026	0.031	0.036	0.042
PANI	0.003	0.007	0.009	0.013	0.018	0.022	0.026	0.034
POEA	0.002	0.005	0.007	0.010	0.015	0.018	0.021	0.028

Total solute iron content in solution is very high for uncoated 13% Cr stainless steel. In coated samples, on the other hand, the mass of iron is very negligible. The POEA coating provides a strong protective barrier against salt solution, with the lowest mass of iron in solution compared with the ZP and PANI coatings.

The iron content of all three coatings decreases with increasing immersion time, due to the formation of a passivation layer that counteracts the aggressive nature of the solution (corrosion).

In our immersion experiment, given the small difference in iron dissolution rate in 3.5% NaCl and in order to classify the best coating, it is preferable to continue the test at extended immersion times other than 120 days, and also to try another test outright (such as salt spray).

IV. Conclusion

In this article, we successfully prepared two organic acid (SSA)-doped conducting polymers based on simple polyaniline and an ortho-ethoxy-conjugated derivative. These polymers exhibit high electrochemical activity and good solubility in common organic solvents such as xylene, making them ideal candidates as anticorrosion agents for epoxy coatings. Among them, POEA shows the best dispersion within the epoxy matrix, resulting in a homogeneous and smooth surface (contact angle around 90°), whereas PANI exhibits agglomeration and porosity defects (contact angle below 70°), as confirmed by scanning electron microscopy. Electrochemical impedance spectroscopy (EIS) results revealed that corrosion protection of the coatings containing 3 wt.% of PANI and POEA is significantly improved compared to the conventional zinc phosphate (ZP) coating system. Notably, the POEA-based coating system exhibited the highest resistance (over 2 MΩ) after 60 days of immersion in a 3.5 wt.% NaCl solution, as indicated by the EIS measurements. This superior protection is attributed to the formation of a stable passive layer composed of Cr_2O_3, reinforced by Fe_2O_3 and Fe_3O_4 compounds [21]. Ultimately, coatings based on PANI and POEA show strong potential for protecting oil installations in aggressive environments, particularly for long-term applications.

Acknowledgment

The authors gratefully acknowledge the financial support provided by the Directorate-General for Scientific Research and Technological Development (DG-RSDT) of the Algerian government through a research grant. Sincere thanks are also extended to the Laboratoire des Sciences et Techniques de l'Environnement (LSTE),Laboratoire de Génie Minier (LGM) and the Laboratory of Green and Mechanical Development (LGMD) at the National Polytechnic School (ENP). Laboratory Division (formerly CRD SONATRACH Boumerdes) and the Central Laboratory of ENAP Lakhdaria, Bouira, are also acknowledged for providing essential resources and technical support throughout this work.

References

[1] A. Mirmohseni, A. Oladegaragoze. "Anti-corrosive properties of polyaniline coating on iron." Synthetic metals 114.2 (2000) 105-108.https://doi.org/10.1016/S0379-6779(99)00298-2.

[2] G. Ebrahimi, J. Neshati, and F. Rezaei. "An investigation on the effect of H3PO4 and HCl-doped polyaniline nanoparticles on corrosion protection of carbon steel by means of scanning kelvin probe." Progress in Organic Coatings 105 (2017) 1-8.https://doi.org/10.1016/j.porgcoat.2016.12.016.

[3] L. Zhong, S. Xiao, J. Hu, H. Zhu, F. Gan. "Application of polyaniline to galvanic anodic protection on stainless steel in H2SO4 solutions." Corrosion science 48.12 (2006) 3960-3968.https://doi.org/10.1016/j.corsci.2006.04.019.

[4] H. Xu, J. Liu, Y. Chen, J. Tang, Z. Zhao. "Facile fabrication of superhydrophobic polyaniline structures and their anticorrosive properties." Journal of Applied Polymer Science 133.47 (2016) 44248-44256.https://doi.org/10.1002/app.44248.

[5] D. Sazou, M. Kourouzidou, and E. Pavlidou. "Potentiodynamic and potentiostatic deposition of polyaniline on stainless steel: Electrochemical and structural studies for a potential application to corrosion control." Electrochimica Acta 52.13 (2007) 4385-4397.https://doi.org/10.1016/j.electacta.2006.12.020.

[6] A. T. Ozyilmaz, G. Ozyilmaz, and O. Yigitoglu. "Synthesis and characterization of poly (aniline) and poly (o-anisidine) films in sulphamic acid solution and their anticorrosion properties." Progress in Organic Coatings 67.1 (2010) 28-37.https://doi.org/10.1016/j.porgcoat.2009.09.010.

[7] M. Mrad, L. Dhouibi, and E. Triki. "Dependence of the corrosion performance of polyaniline films applied on stainless steel on the nature of electropolymerisation solution." Synthetic metals 159.17-18 (2009) 1903-1909.https://doi.org/10.1016/j.synthmet.2009.03.008.

[8] R. Vera, R. Schrebler, P. Cury, R. Del Río, H. Romero. Corrosion protection of carbon steel and copper by polyaniline and poly(ortho-methoxyaniline) films in sodium chloride medium. Electrochemical and morphological study. Journal ofApplied Electrochemistry, 37 (2007) 519-525.https://doi.org/10.1007/s10800-006-9284-y.

[9] D. Tallman, G. Spinks, A. Dominis, G. Wallace. Electroactive conducting polymers for corrosion control Part 1. General introduction and a review of non-ferrous metals. Journal of Solid State Electrochemistry 6 (2002) 73-84.https://doi.org/10.1007/s100080100212.

[10] Y. Hao, Y. Zhao, X. Yang, Bo Hu, S. Ye, L. Song, R. Li. elf-healing epoxy coating loaded with physic acid doped Polyaniline nanofibers impregnated with benzotriazole for Q235 carbon steel. Corrosion science 151(2019)175-189.https://doi.org/10.1016/j.corsci.2019.02.023.

[11] Y. Lei, Z. Qiu, N. Tan, H. Du, D. Li, J. Liu, T. Liu, W. Zhang, X. Chang. Polyaniline/CeO2 nanocomposites as corrosion inhibitors for improving the corrosive performance of epoxy coating on carbon steel in 3.5% NaCl solution. Progress in Organic Coatings, 139 (2020) 105430.https://doi.org/10.1016/j.porgcoat.2019.105430.

[12] J. Alam, U. Riaz, S. Ahmad. High performance corrosion resistant Polyaniline/alkyd ecofriendly coatings. Current Applied Physics 9 (2009) 80-86.https://doi.org/10.1016/j.cap.2007.11.015.

[13] H. D. Tran, J. M. D'Arcy, Y. Wang, P. J. Beltramo, V. A. Strong, R. B. Kaner. The oxidation of aniline to produce Polyaniline: a process yielding many different nanoscale structures. Journal of Materials Chemistry 21 (2011) 3534-3550.https://doi.org/10.1039/C0JM02699A.

[14] S. K. Dhawan, D. C. Trivedi. Electrochemical Behaviour of Polyaniline in Aromatic Sulphonic Acids. Polymer International 25 (1991) 55-60.https://doi.org/10.1002/pi.4990250110.

[15] F. Xiao, C. Qian, M. Guo, J. Wang, X. Yan, H. Li, L. Yue. "Anticorrosive durability of zinc-based waterborne coatings enhanced by highly dispersed and conductive polyaniline/graphene oxide composite." Progress in Organic Coatings 125 (2018): 79-88.https://doi.org/10.1016/j.porgcoat.2018.08.027.

[16] H. Chen, H. Fan, N. Su, R. Hong, X. Lu. Highly hydrophobic polyaniline nanoparticles for anti-corrosion epoxy coatings. Chemical Engineering Journal 420 (2021) 130540.https://doi.org/10.1016/j.cej.2021.130540.

[17] J. H. Sung, J.W. Kim, H.J. Choi, S. B. Choi. Synthesis and characterization of organoclay nano-composite with poly (o-ethoxyaniline). Synthetic Metals, 135-136 (2003) 19-20.https://doi.org/10.1016/S0379-6779(02)01035-4.

[18] C. Hu, Y. Qing, Y. Li, N. Zhang. Preparation of poly(o-ethoxyaniline)-nano SiC composite and evaluation of its corrosion resistance properties. Journal of Alloys and Compounds 717 (2017) 98-107.https://doi.org/10.1016/j.jallcom.2017.05.111.

[19] M. N. Ahmad, A. Hussain, M. N. Anjum, T. Hussain, A. Mujahid, M. H. Khan, T. Ahmed. Synthesis and characterization of a novel chitosan grafted poly (ortho-ethyl)aniline biocomposite and utilization for dye removal from water. Open Chemistry 18 (2020) 843–849.https://doi.org/10.1515/chem-2020-0137.

[20] T. Schauer, A. Joos, L. Dulog, C.D. Eisenbach. Protection of iron against corrosion with Polyaniline primers. Progress in Organic Coatings 33 (1998) 20-27.https://doi.org/10.1016/S0300-9440(97)00123-9.

[21] N. Mounir, M. Nazef, M. Bousbai, H. Yousfi. Anticorrosion performance of TiO2/epoxy films with doped polyaniline and poly(o-ethoxyaniline) applied to carbon steel in 3.5 wt.% NaCl acid solution. Protection and physical chemistry of surfaces, 60, 3(2024) 1-18.https://doi.org/10.1134/S2070205124701788.

[22] R. Gupta, M. Singhal, J. P. Chaudhary. Regulation of nanostructured polyaniline synthesis and its properties through organic solvent in interfacial polymerization. Colloids and Surfaces A: Physicochemical and Engineering Aspects, 632 (2022).https://doi.org/10.1016/j.colsurfa.2021.127806.

[23] K. Tzou, R.V. Gregory. A method to prepare soluble polyaniline salt solutions — in situ doping of PANI base with organic dopants in polar solvents. Synthetic Metals 53 (1993) 365-377.https://doi.org/10.1016/0379-6779(93)91106-C.

[24] H. Alhweij, E. A. Emanuelsson, S. Shahid, J. Wenk. High performance in-situ tuned self-doped polyaniline (PANI) membranes for organic solvent (nano) filtration. Polymer, 245 (2022).https://doi.org/10.1016/j.polymer.2022.124682.

[25] H. Zhou, R. Chen, Q. Liu, J. Liu, J. Yu, C. Wang, M. Zhang, P. Liu, J. Wang. Fabrication of ZnO/epoxy resin superhydrophobic coating on AZ31 magnesium alloy. Chemical Engineering Journal 368(2019) 261-372.https://doi.org/10.1016/j.cej.2019.02.032.

[26] A. Diraki, S. Omanovic. Smart PANI/epoxy anti-corrosive coating for protection of carbon steel in sea water. Progress in Organic Coatings, 168(2022) 106835.https://doi.org/10.1016/j.porgcoat.2022.106835.

[27] T. Liu, J. Wei, L. Ma, S. Liu, D. Zhang, H. Zhao. Effect of Polyaniline-based plate on the anticorrosion performance of epoxy coating. Progress in Organic Coatings 151 (2021) 106109.https://doi.org/10.1016/j.porgcoat.2020.106109.

[28] M. Kraljic, Z. Mandic, Lj. Dui. Inhibition of steel corrosion by Polyaniline coatings. Corrosion Science 45 (2003) 181–198.https://doi.org/10.1016/S0010-938X(02)00083-5.

[29] A. T. Ozyilmaz, A. Akdaga, I. H. Karahan, G. Ozyilmaz. The influence of Polyaniline (PANI) coating on corrosion behaviour of zinc–cobalt coated carbon steel electrode. Progress in Organic Coatings 76 (2013) 993–997.https://doi.org/10.1016/j.porgcoat.2012.10.020.

[30] R. Liu, Q. Yao, L. Liu, F. Meng, F. Wang, Studies of different acid doped polyaniline incorporated into epoxy organic coatings on the mg alloy, Progress in Organic Coatings 166 (2022) 106774.https://doi.org/10.1016/j.porgcoat.2022.106774.

[31] S. Shi, Y. Zhao, Z. Zhang, L. Yu, Corrosion protection of a novel sio2@ pani coating for q235 carbon steel, Progress in Organic Coatings 132 (2019) 227–234.https://doi.org/10.1016/j.porgcoat.2019.03.040.

[32] C. Li, J. Xu, Q. Xu, G. Xue, H. Yu, X. Wang, J. Lu, G. Cui, G. Gu, Synthesis of ti3c2 mxene@ pani composites for excellent anticorrosion performance of waterborne epoxy coating, Progress in Organic Coatings165 (2022) 106673.https://doi.org/10.1016/j.porgcoat.2021.106673.

[33] F. Gao, Y. Luo, J. Xu, X. Du, H. Wang, X. Cheng, Z. Du, Preparation of graphene oxide-based polyaniline composites with synergistic anticorrosion effect for waterborne polyurethane anticorrosive coatings, Progress in Organic Coatings 156 (2021) 106233.https://doi.org/10.1016/j.porgcoat.2021.106233.

[34] Y. Zhang, Y. Shao, X. Liu, C. Shi, Y.Wang, G. Meng, X. Zeng, Y. Yang, A study on corrosion protection of different polyaniline coatings for mild steel, Progress in Organic Coatings 111 (2017) 240–247.https://doi.org/10.1016/j.porgcoat.2017.06.015.

[35] J. Liu, Q. Yu, M. Yu, S. Li, K. Zhao, B. Xue, H. Zu, Silane modification of titanium dioxide-decorated graphene oxide Nanocomposite for enhancing anticorrosion performance of epoxy coatings on aa-2024, Journal of Alloys and Compounds 744 (2018) 728–739.https://doi.org/10.1016/j.jallcom.2018.01.267.

[36] N. N. Taheri, B. Ramezanzadeh, M. Mahdavian, Application of layer-bylayer assembled graphene oxide nanosheets/polyaniline/zinc cations for construction of an effective epoxy coating anti-corrosion system, Journal of Alloys and Compounds 800 (2019) 532–549.https://doi.org/10.1016/j.jallcom.2019.06.103.

Nacer Mounir received his Engineering degree in Materials Engineering from the École Nationale Polytechnique (ENP), Algiers, Algeria, in 2011, a Master's degree in Mechanics, Materials, and Processes from ENSAM ParisTech, Paris, France, in 2013, and a Ph.D. in Materials Engineering from ENP in 2025. He is currently conducting research at the laboratory of Quality, Health, Safety, Environment and Industrial Risk Management. His research interests include corrosion, thin films and coatings, electrochemical characterization, and electrochemical hydrogen conversion.

Malha Nazef is a Corrosion Manager, working at the laboratory of Corrosion, Department of Corrosion, Division Laboratories Exploration and Production SONATRACH (Ex CRD), Boumerdes, Algeria. Her main research interests include corrosion, thin films and coatings, synthesis and characterization of conductive polymers.

Zakaria Baka received his Engineering, Master's, and Ph.D. degrees in Mechanical Engineering from the Ecole Nationale Polytechnique, Algiers, Algeria, in 2013, 2014, and 2024, respectively. He is currently conducting research at the Laboratory of Green and Mechanical Development, ENP. His research interests include engineering mechanics, solid mechanics, heat transfer, elasticity, and thermo-elasticity.

Hamid Yousfi is a full Professor at Laboratory of Quality, Health, Safety, Environment and Industrial Risk Management at the Ecole Nationale Polytechnique, Algiers, Algeria. His main research interests include corrosion, thin films and coatings, synthesis and characterization of aluminum alloys.

2716-912X © 2025 Ecole Nationale Polytechnique

Design of Mechanically Agitated Fermenter for a Daily Ten Tons Ethanol Production from Cool-Feed Biomass

Fredrick B. Ugi, Benedict U. Ugi, and Gloria T.Tamunotonye

Abstract−This research designed a mechanically agitated fermenter that operates with the principles of heat exchanger, sustaining fermentation process in a conducive thermodynamic state which promotes product (ethanol) yield as well as its concentration (purity), with a target of achieving 10 tons per day ethanol from cool-feed biomass (palm wine). The design shows that the target was achieved at the fermenter vessel Area (A_0) of 15.5917 m^2 with the use of 1271.1646 kg/day coolant (water), operating at a conservative heat flow of 505262.484 J/day. MATLAB simulation was used to access the dynamic behavior of the agitated fermenter over a range of biomass and time. From the developed models, the result showed that the reaction rate of fermentation ((dC_L)/dt) was proportional to the overall mass transfer coefficient per unit volume of the biomass (K_L a) measured in per hours of the fermentation (hr^{-1}). This can be due to decrease in the feed mass of a fermenter due to biomass decay which increases the oxygen penetrability of the system. Decrease in the fermenter material (biomass) due to biomass decay was also observed. This implies promotion of the fermentation rate of the system as the rate of biomass fermentation ((dC_L)/dt) decreases with increase in feed value of biomass. The residence time of the system was also reduced which implies that there was an increase in the fermentation rate of the biomass hence, promoting satisfactory product yield and purity which is one of the target of this research.

Keywords− Fermentation, Ethanol, cool-feed, biomass, heat transfer

NOMENCLATURE

Authors might use unnumbered section to include the whole of abbreviations appeared in the manuscript. This section must be included before the Introduction. The abbreviations should be included as follows:

pH	Hydrogen Ion Concentration.
TDS	Total Dissolved Solids.
COD	Chemical Oxygen Demand.
BOD	Biochemical Oxygen Demand.
LTS	Low Temperature Steam..
LMTD	Logarithmic Mean Temperature.

I. INTRODUCTION

With increasing demands for GHG emission-free energy sources and ethanol (higher purity) having the ability to meet such demands, fermenters are being designed and researched mostly by various scientists and specialists [1-3] Fermenters have been known for their tremendous abilities of converting sugary substances into chemical products such as ethanol through microorganism interactions with biomass over periods of time [1-2] This has made fermentation a step to meet the demands for high-purity ethanol. Fermentation is classified among the oldest biotechnology processes, mostly applied in the food and medical processing industries. It is based on the application of biological enzymes (microorganisms') interactive abilities on organic matters to alcohol production from sugary substances [3]. Fermentation is a metabolic activity that takes place in starchy-like organic substances such as sugar or carbohydrates, converting them into economically important organic products like acids, gases, and alcohols by microbial activities over a period of time, either aerobically or anaerobically [1,3-5] . This process is carried out in an equipment known as a bioreactor or, more easily, a "fermenter", which provides a suitable and well-controlled environment for the growth of microorganisms or animal cells, aiming at obtaining these desirable products from sugar-active substances [5-6] In a fermenter, sterile nutrients and pure cultured microorganisms are mixed and left for the fermentation process to occur under aseptic and optimum conditions like pH, temperature, oxygen, etc. Fermentation takes place either as a batch, fed-batch, or continuous bioprocess [6-8]. Ethanol happened to be one of the most vital products obtained from fermentation all over the years, even as far back as the ancient centuries of man. Ethanol is a very important chemical (colorless and water-soluble with a boiling point of 78.37° C) used for multi purposes such as alcoholic drink production, drugs, preservatives, organic solvents, and even as biofuel for automobile engines when of high purity [7-10]. The production of high purity level ethanol capable of use as fuel has kept researchers looking for means of attaining such a height of industrial requirements for a GHG-free energy source, and this high purity level ethanol is subject to the distillation temperature of the separating process from water, which is always known to be present with ethanol obtained from organic means. Ethanol with a boiling point of 78.37 °C has always been a headache to separate from water, whose boiling point tends to be nearer to that of the ethanol,

Manuscript received June 24, 2024; revised June 16, 2025.

*F. B. Ugi and G. T. Tamunotonye are with the Department of Petrochemical Engineering, River State University, Port Harcourt, Nigeria (e-mail:*Fredrick.ugi@rsu.edu.ng, tamunotonye.tamuno-nemgiofon@ust.eu.ng*).*

B. U. Ugi. Is with the Department of Pure & Industrial Chemistry, University of Calabar, P.M.B. 1115, Calabar, Nigeria (e-mail: ugibenedict@gmail.com).

Digital Object Identifier (DOI): 10.53907/enpesj.v5i1.273

2716-912X © 2025 Ecole Nationale Polytechnique

catching the great interest of researchers all over the world [8-10]. So designing a fermenter that will aid the microbial growth in favor of the fast and satisfactory fermentation of bio-mass to generate high-purity ethanol at considerable temperature and other process variables or conditions is a great step towards a GHG-free automobile energy source. An effective fermenter provides a positive influence on the biological reaction and prevents foreign contamination [9-10] reported that the fermentation process requires proper mixing and monoseptic conditions to promote uniform shear rates. Aseptic environment is a well-defined defense means against unwanted microorganisms and contaminants of fermentation. This aseptic environment also prevents the escape of mixable cells from the process [3-6] . A higher degree of control over contamination should always be a considerable factor when designing fermenters for the sake of favorable product yield [9]. The design of the fermenter and its mode of operation depend greatly upon microorganisms' activities and interactions with biomass during the production process of ethanol. A good fermenter needs to have control over the pH, temperature, and oxygen tension of the system, as well as allow septic process operations, which are essential not only for control but also for future routine manufacturing [11] was able to design a batch-stirred bioreactor for ethanol production that uses yeast Saccharomyces cerevisiae as enzymes to improve the fermentation's performance. However, the fermenter designed had a fermentation time of 11.4 hrs with ethanol stripping of 69.1g and 12 hrs for 75.9g without stripping. The kinetic constants were 2.0kg, 97.9 kg, and 0.476 h, respectively. But the design was limited to only batch stirred bioreactors using yeast, and it is also a closed system culture technique. This work designs a mechanically agitated fermenter using heat transfer principles for the production of 10 tons per day of ethanol from cool feed. Biomass neglects the enzyme type as one of the major design criteria, making the design unique for all types of fermentation processes.

II. MATERIAL AND METHODS

A. Figures

Some of the materials used in this work were mainly to determine concentration before and after fermentation, with some other operating conditions such as pH and Temperature. These materials include;

- Palm wine
- Enzyme
- pH meter
- Thermocouple..

B. Design Methods

In this current research work, the basic methods used to ensure effectiveness of this research range from development of models using material and energy balances (considering thermodynamic constraints of the process and state variables required for the efficacy of the system), Numerical differential and integral method of solving, Data gathering via analytical method and Simulation validation methods using MATLAB-Simulink

C. Modeling Assumptions

These includes Steady-state operating condition, use of anaerobic (closed) system to ensure contamination free operation, negligible shaft work condition , adiabatic conditions, augmentation steam temperature of system which must not exceed 423k (〖150〗^o C), bottom jacket sidewalls fermenter for common hydrofoils, design is done with features supportive to process control over range of process variables, application of cylindrical traditional design techniques, use of mild steel for vessel surface design to reduce corrosion effects, the fermenter is made to have an external cooling jacket, to regulate / control the system temperature as shell side of a heating vessel., Due to the operation temperature and presents volatile compounds, the design is made considering the use of a sterilizable condenser required to prevent evaporation loss and to promote safety reasons.

D. Project Description

This work is based on the understanding that fermentation process involves the interactions of microorganisms with sugary biomass over a period of time, and these microorganisms exist in different forms and different properties. Some prokaryotic and others eukaryotic in nature, there metabolic pathways are with regards to their existence within surrounding factors which includes temperature, oxygen, moisture content, pH, Total Dissolved Solids (TDS), Chemical Oxygen Demand (COD) as well as their Biochemical Oxygen Demand (BOD), etc. Temperature is of great interest on this studies of obtaining pure ethanol, the purity and yield of ethanol from fermentation is subject to the temperature of the system. Hence some microorganisms in a fermentation system as shown in Figure 1 are mesophilic, some thermophilic and others Super thermophilic. This fermenter is designed to suit the existence of all the three types of fermentation favoring bacteria's via subjecting the fermenter to operate with a temperature range of T_f=298k=〖25〗^o C to T_s=423k=〖150〗^o C using a cooling jacket system which works with respect to the fundamental principles of shell and tube heat exchangers.

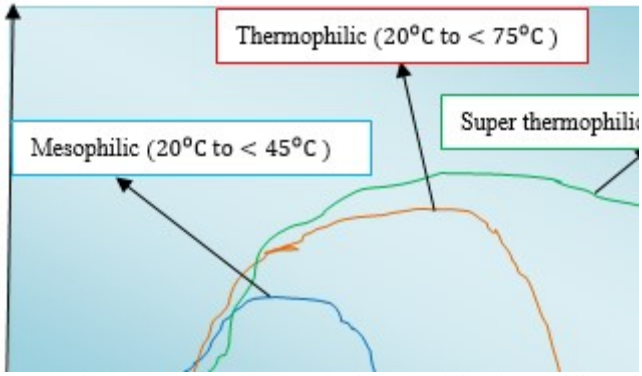


Fig. 1: Microbial temperature existence plot for Mesophilic, thermophilic and Super-thermophilic micro-organisms

Hence the feed used in this system (palm-wine) is a cool-feed, entering the fermenter at temperature below active operating microbial fermentation temperature range, the low temperature steam (LTS) is being injected with the feed to raise the system to instant fermentation, the feed is augmented using low temperature steam. The system designed also contains a cooling jacket carrying ambient water, used to control the heat duty / temperature of the system within the fermentation period.

E. Equipment (Agitated Fermenter) Design Calculations

The design of the agitated fermenter as done in this work applies the usage of step wise hand determination of some of the results as well as a dynamical simulation of the system

using MATLAB to access the system behavior over change in time and other process conditions such as feed mass.
Considering the fermenter system as stated;

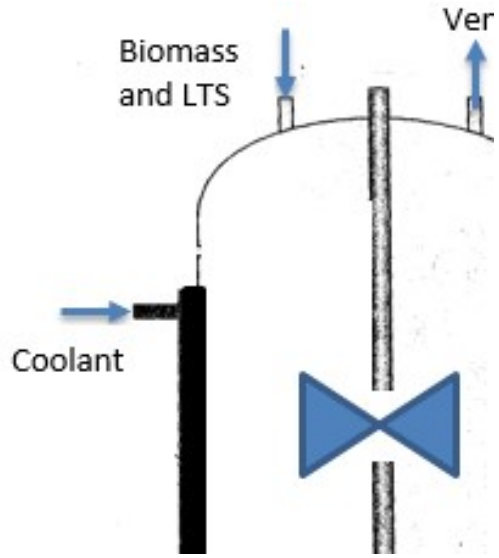


Fig. 2: Diagram of a Bottom Jacket agitated fermenter for microbial biomass production

F. Fermenter Material Balance

In every fermenter, there are specified enzymatic biochemical reactions taking place, leading to generation of diverse products, but in this design, the reaction favorable towards ethanol production from the biomass (palm wine) is considered for the fermenter, schematically describe as follows;

$$C_2H_5OH + Enzymes\ (\text{Yeast}) \xrightarrow[\text{Fementation}]{\text{Steam}} CH_3COOH + H_2O$$

Considering the 1st order chemical reaction with enzyme (yeast) describe as;

$$A + Enzymes \xrightarrow{k_1} \text{Produ}$$

Taking the material balance of the system (fermenter) as;

$$\left\{\begin{matrix}\text{total inpute rate}\\ of\ Biomass\\ fed\ into\ the\ fermenter\end{matrix}\right\} - \left\{\begin{matrix}\text{total output rate}\\ of\ Biomass\\ after\ fermentation\end{matrix}\right\} - \{r_A\}V_R = \left\{\begin{matrix}\text{accumulation}\\ in\ the\\ fermenter\end{matrix}\right\} \quad (1)$$

Where;
$-\{r_A\}$ =the consumption rate of substrate A in mol*l^{-1}s^{-1}

$$= k_1 C_{A,1} \quad (2)$$

Taken the reaction rate constant (K_1) of the fermentation to be of equivalent to the Overall mass transfer coefficient per unit volume biomass (K_La) measured in per hours of fermentation (hr^{-1})

$$\therefore -\{r_A\} = K_La(C^* - C_L) = OTR \quad (3)$$

$$K_La = \frac{\text{Amount of reactant}}{L \times t} \quad (4)$$

Hence the production is per day;

$$t = 24 \times 60 \times 60 = 86400\ \text{sec}$$

$$\therefore K_La = \frac{n}{86400L}$$

Considering the system operating at steady state;

$$V_oC_{A,o} - V_oC_{A,1} + K_La(C^* - C_L)V_R = 0 \quad (5)$$

So;

$$V_R = \frac{V_oC_{A,1} - V_oC_{A,o}}{K_La(C^* - C_L)} \quad (6)$$

Determining the time of the fermenting:

$$\tau = \frac{V_R}{V_o} = \frac{C_{A,1} - C_{A,o}}{K_La(C^* - C_L)} \quad (7)$$

G. Heat Balance of Agitated Fermenter

Hence the fermenter designed is a thermal subjective type, the biomass is been feed co-currently with low temperature steam ($\leq 70^oC$) into the fermenter, the essence of the lower temperate steam is to is to augment the microbial actions of the system, promoting fact enzymatic and biochemical reactions in the system, helping the process speedy start up actions, for better yielding. While the cooling jacket carrying cooling fluid is used to restrict the temperature from exceeding 105°C during the process period. This will enable the safe operations of the mesophilic, thermophilic and super thermophilic organisms that are needed for favorable fermentation of the biomass. During fermentation, temperature of the system may likely increases, so the cooling water is to control the system temperature so as to fall along the favorable range as desired by the microbes. So the heat or temperature of operation needs to be of great consideration for the design [12-14].
So taking the system energy balance is as follows:

$$\left[\frac{\text{Accumulation of heat system}}{\text{Time}}\right] = \left[\begin{matrix}\text{Rate of total}\\ \text{input of heat}\end{matrix}\right] - \left[\begin{matrix}\text{Rate of tatol output}\\ \text{of heat}\end{matrix}\right] + \left[\begin{matrix}\text{heat released}\\ \text{to environment}\end{matrix}\right] - \left[\begin{matrix}\text{shaft}\\ \text{work}\end{matrix}\right] \quad (8)$$

Defining each unit of equation (8) gives:

$$dH/dt = Q_{in} - Q_{out} + Q + W \quad (9)$$

At negligible shaft work conditions;

$$\therefore dH/dt = Q_{in} - Q_{out} + Q \quad (10)$$

At adiabatic conditions

$$Q_{in} = Q_{out} \quad (11)$$

Hence equation (10) becomes

$$dH/dt = Q \quad (12)$$

$$dH/dt = Q = mcp\Delta T \quad (13)$$

$$\therefore Q = mcpz \quad \textbf{(13b)}$$

Assuming the tube content will be ethanol and the coolant will be more viscous to the ethanol, so the tube will bear the biomass which produces the ethanol while the shell will carry the coolant

For the tube side;

$$Q_{tube} = m_t C_{p,t} dT \quad \textbf{(14)}$$

The tube holds the biomass and the augmentation steam, hence at a co-current flow of the coolant to the fermenter feed at steam injection bases, equation (14) above becomes:

$$= \frac{(m_{feed}+m_{steam})(Cp_{feed}+Cp_{steam})(T_f+T_s)-2T_2}{8} \quad \textbf{(15)}$$

Hence operating the reactor with parametric values, calculated to match efficiency of the fermenter with capacity of 10 tons/day;

$$M_{feed} = 10 \text{ Tons/day}$$

$$\text{If: } m_{feed} = M_{feed} \times 1000 = \frac{10{,}000\text{kg}}{\text{day}} \quad \textbf{(16)}$$

Taking the steam mass flow rate to be;

$$m_{steam} = 238 \text{ Tons/day}$$

Considering Ethanol as the substance (product) of the fermentation;

$$Cp_{feed} = C_{p_{biomass}} = 2.5\text{kj/kgK}$$

$$C_{p,steam} = 1.996\text{kj/kgK}$$

$$T_f = 298\text{k} = 25^oC$$

$$T_s = 423\text{k} = 150^oC \text{ (perry's et al 7th edition)}$$

$$T_2 = 213\text{k}$$

$$\therefore Q_{tube} = \frac{(10000-232)(2.57-1.996)(298+423)-2(213)}{8}$$
$$= 505262.484 \text{ J/day}$$

For the shell section;
Considering a heat conservative system:

$$Q_{shell} = Q_{tube} \quad \textbf{(17)}$$

$$\therefore Q_{shell} = m_s C_{p_s} \Delta T = m_s C_{p_s}(T_2 - T_1) \text{ of water using water as coolant} \quad \textbf{(18)}$$

$$505262.484 \text{ J/s} = m_s \times 4.18(383-288)$$

$$\therefore m_s = \frac{505262.484}{4.184 \times (T_2 = 383 - T_1 = 288)}$$

$$= 1271.1646\text{kg/day}$$

$$\therefore m_s = \text{mass flow rate of coolant (water)} = 1271.1646\text{kg/day of coolent (water)}$$

Hence, determining the system logarithmic mean temperature (LMTD), from the above equations we have:

$$\text{LMTD} = \frac{T_F+T_s-2T_2-2t_1+2t_2}{2\text{In}\left[\frac{T_F+T_S-2T_2}{2t_1+2t_2}\right]} \quad \textbf{(19)}$$
$$= \frac{298+423-2(213)-2(288)+2(383)}{2\text{In}\left[\frac{298+423-2(213)}{-2(285)+2(383)}\right]}$$
$$= \frac{485}{2\text{In}\left(\frac{295}{196}\right)} = \frac{485}{2(0.4089)}$$
$$= 593.0545\text{K}$$

$$\text{If: } \Delta T_m = F_t \times \text{LMTD} \quad \textbf{(20)}$$

$$\text{Where: } \sqrt{\frac{\left(\frac{\partial_t}{\partial_i}\right)^*\left(\frac{H_2}{\partial_2}\right)^*}{\left(\frac{\partial_t}{\partial_i}\right)\left(\frac{H_2}{\partial_i}\right)}} = F_t \quad \textbf{(21)}$$

$$F_t = \sqrt{\frac{\left(4/1.3\right)\left(7.5/1.3\right)}{\left(5.5/2.3\right)\left(8.5/2.3\right)}}$$
$$= \sqrt{\frac{17.7515}{8.8374}}$$
$$= 1.4173$$

Where;

$d_t = 5.5,\ d_i = 2.3,\ H_2 = 8.5,\ d_t^* = 4.0,\ D_i^* = 41.3$ and $H_L^* = 7.5$

$$\therefore \Delta T_m = F_t \times \text{LMTD}$$
$$= 1.4173 \times 593.0545$$
$$= 840.5235\text{K}$$

Hence the fermenter tube will be the bearer of the biomass to generate ethanol, the design will focus more on the tube, considering the dimensionless values of the system as follows;

Taken;
Tube specific heat capacity = $C_{p_{ethanol}} = 2.5$ kj/kgK
Tube internal diameter = di = 4.58m
Tube outside diameter = do = 7.5m
Ethanol thermal conductivity = $K_t = 0.519\ w/m^oC$
Tube side dynamic viscosity = $\mu = 8.90 \times 10^{-3}$mNs/m^2
Tube side density = $\rho = 1026.6636$ kg/m^3= feed (biomass) density

$$\text{Tube side velocity} = u = \frac{mass\ velocity\ (G')}{tube\ side\ density\ (\rho)} \quad \textbf{(22)}$$

$$= 8.1523\ m/day$$

i. $$\text{Prandtlnumber: } P_r = \frac{C_p \mu}{K_t} \quad \textbf{(23)}$$

$$= \frac{(2.5 \times 10^3) \times (8.90 \times 10^{-3})}{0.519}$$
$$= 4.29$$

ii. $$\text{Reynolds number: } R_e = \frac{\rho d_i u}{\mu} \quad \textbf{(24)}$$

$$\therefore R_e = \frac{8.1523 \times 4.58 \times 1026.6636}{8.90 \times 10^{-3}}$$

$$= 4.307088 \times 10^6$$

H. Determining the heat Transfer Coefficient of the Fermenter (h):

Hence the calculated Reynolds number (R_e) is greater than 2100, the heat transfer coefficient will be;

$h_o = j_{h,s}\frac{K_s}{d_i}R_eP_r^{0.33}\left[\frac{\mu}{\mu_w}\right]^{0.14}$ = shell side fluid heat coefficient **(25)**

$h_i = j_{h,t}\frac{K_t}{d_i}R_eP_r^{0.33}\left[\frac{\mu}{\mu_w}\right]^{0.14}$ =Tube side fluid heat coefficient **(26)**

Where:

$j_{h,t}$ = Heat transfer factor for the tube side

$j_{h,s}$ = Heat transfer factor for the shell side

μ_w = Tube side viscosity at tube temperature

Neglecting the viscosity correction term, the equations for heat transfer coefficient becomes;

$$h_o = j_{h,s}\frac{K_s}{d_i}R_eP_r^{0.33} \quad \textbf{(27)}$$

$$h_i = j_{h,t}\frac{K_t}{d_i}R_eP_r^{0.33} \quad \textbf{(28)}$$

For the tube, using the calculated Reynolds number (R_e), from the Figure 3 plot, $j_{h,t}$ is 2.1×10^{-3}

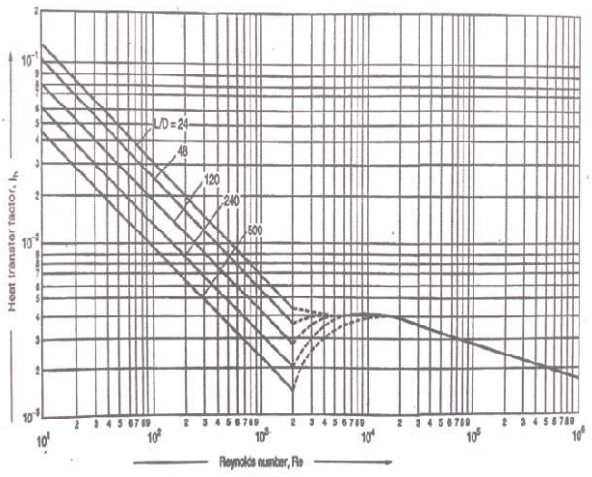


Fig. 3: Heat transfer plot on Reynold number bases

So;

$$h_i = j_{h,t}\frac{K_t}{d_i}R_eP_r^{0.33}$$
$$= (2.1 \times 10^{-3})\left(\frac{0.519}{4.58}\right)(4.307088 \times 10^6)(4.29)^{0.33} = 1657.3561 w/m^2K$$

I. Determining the Overall heat transfer coefficient of the Fermenter (Uo):

$$\frac{1}{U_o} = \frac{1}{h_s} + \frac{1}{f_s} + \frac{d_o \ln\left(\frac{d_o}{d_i}\right)}{2K_m} + \frac{d_o}{d_i f_i} + \frac{d_o}{d_i h_i} \quad \textbf{(29)}$$

Where**:**

f_i and f_s = Tube and shell side fouling coefficient or factor

Considering ethanol been a lighter hydrocarbon and the water been used as the coolant, the tube side fouling factor will be $f_i = 5000\, w/m^oC$ while that of the shell side will be $f_s = 3000\, w/m^oC$, also using copper with thermal conductivity of $K_m = 80 w/m^oC$

So Overall heat transfer will be,

$$\frac{1}{U_o} = \frac{1}{h_s} + \frac{1}{3000} + \frac{(7.5) \times \ln\left(\frac{7.5}{4.58}\right)}{2 \times 80} + \frac{7.5}{4.58 \times 5000} + \frac{7.5}{4.58 \times 1657.3561}$$

Assume $h_s = 855\, w/m^2K$

$$\therefore U_o = 38.5543\, wm^2K$$

Determining the Fermenter Area (A_o)

$$A_o = \frac{Q_s}{U_o \Delta T_m} \quad \textbf{(30)}$$

$$= \frac{505262}{38.5543 \times 840.5232} = 15.5917 m^2$$

Determining the system heat duty (Q_{Duty}):

$$Q_{Duty} = UA_o\Delta T_m \textbf{(31)}$$

Fermenter Diameter (D_R):

$$D_R = \frac{V_R}{A_o} \textbf{(32)}$$
$$= \frac{V_oC_{A,1} - V_oC_{A,o}}{A_oK_La(C^* - C_L)}$$

Fermenter Length (L):

$$L = \frac{Q}{U_o \pi D_R \Delta Tm} \quad \textbf{(33)}$$

Height of the fermenter:

$$h = \frac{4V_R}{\pi \emptyset^2} \quad \textbf{(34)}$$

Where;

$\emptyset$ = diameter of the sparger base plate

Number of stirring blades:

$$n = \frac{h}{\text{Aspect ratio of the fermenter} \times \text{radius of the fermenter}} \quad \textbf{(35)}$$

$$= \frac{h}{3r} \quad \textbf{(36)}$$

Impellers spacing:

$$S = \frac{h}{2n} \quad \textbf{(37)}$$

Determining the wall/vessel thickness:

$$T = e + \text{corrosion allaawance} \textbf{(38)}$$
$$= \frac{Pd_i}{2SE - P} + C(\text{ASME code}) \quad \textbf{(39)}$$

$$\text{thickness} = \frac{1 \times 0.46}{(2 \times 5 \times 1) - 1} + 2$$
$$= 2.05\text{cm}$$

Where;

P = Operating pressure

d_i = External diameter

S = Design stress

E = Welding joint efficiency

C =Corrosion clearance

J. Costing of Fermenter Design

Given that a fermenter has the following factors;

Material of design = stainless steel

Reactor diameter = 4.6m

Operating pressure = 1 bar/1 atm

Impellers spacing = 5n/m

Bioreactor thickness = 2.05cm

Vessel type = cylindrical vessel

Number of blades = 20 of blades

Height of reactor = 6.65m

Calculating the total cost of the reactor:

Total cost of reactor = vessel cost + cost of packing, but

Vessel cost = bare vessel cost x material factor x pressure factor.

Therefore, for stainless steel

$$C_u = (\$\,6000 \times 2 \times 5)$$
$$= \$\,60{,}000$$

Maintenance cost = 2% of the vessel cost

$$= 0.2 \times \$60{,}000$$
$$\therefore C_M = \$\,12{,}000 \text{ per year}$$

Cost of packing = 16319.2481×20
$$= \$\,326.38$$

Total cost of reactor = vessel cost + cost of $= 60{,}000 + 326.38$
$$= \$\,60,.32$$

III. RESULT AND DISCUSSION

The simulation is based at variant mass bases of the system from M_{feed}=10 kg to M_{feed}=100kg according to Table 1.

Table. I

GROUPED DATA FOR MODEL SIMULATION

Data	Value/Source	Data	Value/Source
$\mathbf{t_1}$	**288k**(Gregory, 2007)	C_L	0.003mmL^{-1}(Gregory, 2007)
t_2	**383k**(Gregory, 2007)	$C_{A,o}$	**124.7 moldm^{-3}**(Analysis)
$\mathbf{T_F}$	**298 k**(Gregory, 2007)	Air flow rate throw spagger	0.4vvm(Ghasen,et al.,, 2007)
$\mathbf{T_s}$	**423 k**(Gregory, 2007)	$\mathbf{V_o}$	**54.5dm^{3}min^{-1}**(Assumed)
$\mathbf{T_2}$	**213 k**	M	8.90×10^{-3}mNs/m^2
$\mathbf{K_s}$	**0.919jm^{-1}s^{-1o}C^{-1}**(CEP April 2019)	$\mathbf{C^*}$	6.8mmL^{-1}(Gregory, 2007)
$\mathbf{j_{h,t}}$	86044.6 **w/m^{2}k** (Ghasen, et al.,, 2007)	$j_{h,s}$	**8.8**(Ghasen, et al.,, 2007)
K_t	0.519 w/m^oC(ethanol Conductivity)	$C_{A,1}$	**231.86 moldm^{-3}**(Analysis)
$C_{p_{ethanol}}$	2.5 kj/kgK	Di	4.58m(Assumed)
P	1026.6636 kg/m^3(Analysis)	do	7.5m(Assumed)

K. Heat flow rate of the fermenter against mass of materials used

Figures (except photographs) are often reduced to 7 cm × 7 cm. When preparing the figures, authors should pay attention to the widths of lines and similar details, as some (e.g. dotted or thin lines) may disappear after reduction. Figure 4 shows that the heat flow within an anaerobic fermenter during the fermentation period is proportional to the increase in the mass of the biomass set for digestion. So a digester operating with larger biomass feed requires heat or temperature control measures, either by using a cooling jacket or any other temperature-regulating approach, to promote microbial and yield efficiency of the [12, 15-18] as designed in this work using principles of heat transfer. A cooling jacket is required for large-scale fermenters to remove heat generated and is obligatory for the successful completion of the system [18-21]. Hence, the designed fermenter is in agreement with the accepted model which described temperature control measures as being of great relevance to every large-scale fermentation.

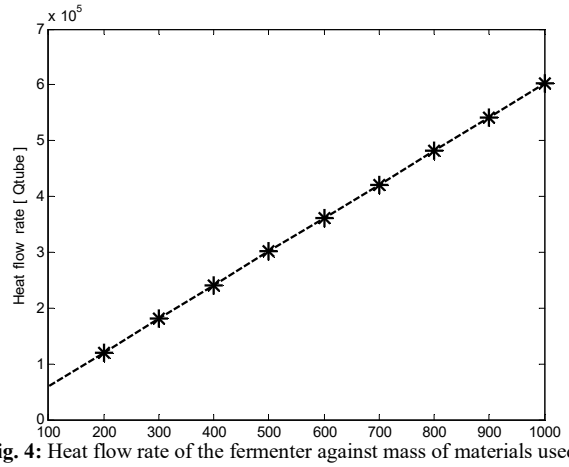


Fig. 4: Heat flow rate of the fermenter against mass of materials used

L. Volume of the fermenter against mass of materials used

Figure 5 shows variation of the volume of the fermenter against the mass of the feed sent into the reactor (M_{Feed}), hence from the graph there is a direct proportional increase in the system heat flow rate with respect to the increase in the feed value sent into the reactor implying that the system will require a greater amount heat to act on the feed in order to yield a satisfactory value or amount of the product [9, 13, 22-24]. Hence as M_{Feed} increases from 100to 1000, so V_R increases rapidly and proportionally perving way for a productive chemical process.

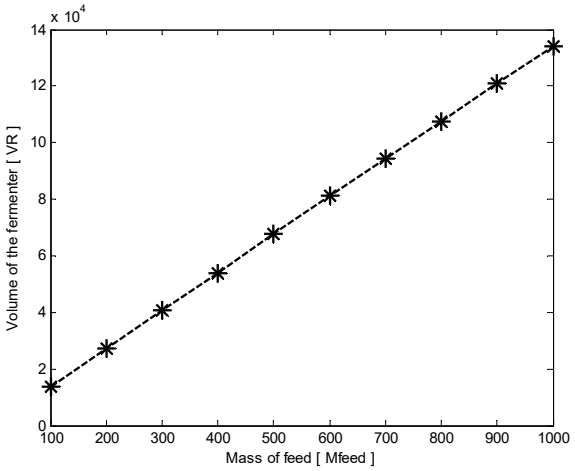


Fig. 5: Volume of the fermenter against mass of materials used

M. Mass of feed in the fermenter against Space time

Figure 6 shows variation of space time with mass of feed sent into the reactor (M_{Feed})**.** this was noticed to be exponentially in the increase as space time "X_A" increases. As the value of space time gradually increases towards $\tau = 2500$ sec, there was a gradual increase in the conversion process which followed a rapid conversion as the space time reach maximum. This project the designed reactor as effective for the production of 10 tons of ethanol per day under the space time and conversion rate.

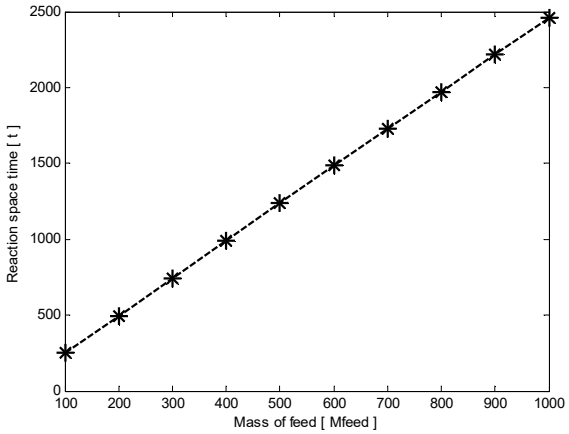


Fig. 6: Mass of feed in the fermenter against Space time

N. Rate of reaction in the fermenter against Mass of feed used

Figure 7 shows variation of the reaction rate in the reactor with respect to the change in process feeding value, from the plot it can be clearly stated that the rate of chemical reaction in the reactor all exponentially to the rapid increase in the feed value which is been sent into the reactor which stands a better chance of effecting the yielding rate

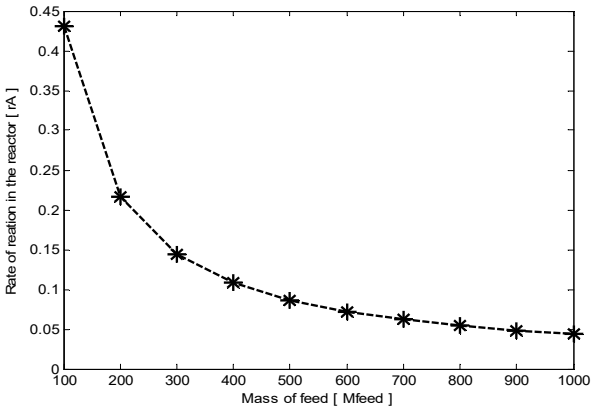


Fig. 7: Rate of reaction in the fermenter against Mass of feed used

O. Rate of reaction in the fermenter against Space time

Figure 8 shows variation of space time with respect to the reaction rate (r_A) which as it is reaction tends to be exponentially increasing, there is a corresponding fall on the system space time.

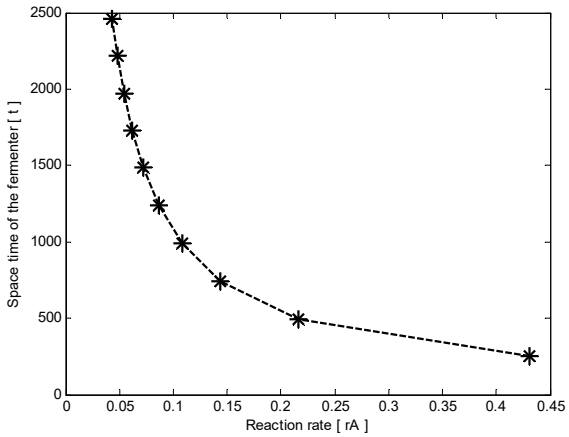


Fig. 8: Rate of reaction in the fermenter against Space time

P. Comparative graph of Rate of Oxygen transfer and Volume of the fermenter Vs mass of materials

Figure 9 shows variation of space time with fractional conversion (X_A) which is exponentially increasing as space time "X_A" increases**.** At maximum value of space time ranging from $\tau = 0.9$ hour and $\tau = 3.67$ hour the exponent level of the reaction tends to grow smoothly while after then there is a vast growth in the conversion process.

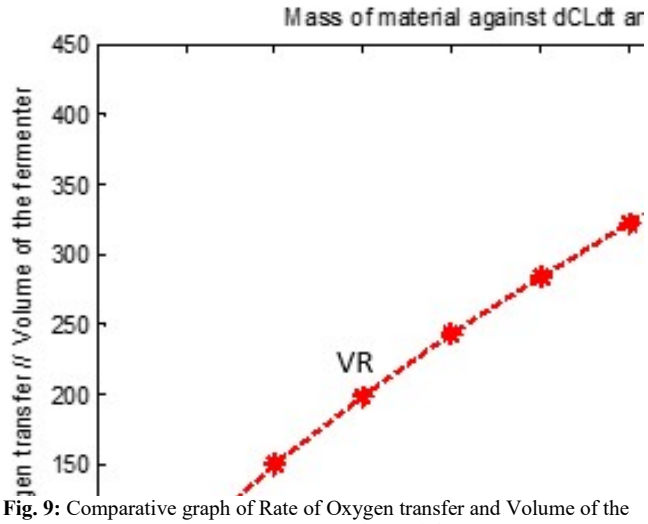


Fig. 9: Comparative graph of Rate of Oxygen transfer and Volume of the fermenter Vs mass of materials

Q. Rate of Oxygen Transfer Vs mass of Material in the fermenter

Figure 10 shows variation of space time with fractional conversion (X_A) which is exponentially increasing as space time "X_A" increases. At maximum value of space time ranging from $\tau = 0.9\ hour$ and $\tau = 3.67\ hour$ the exponent level of the reaction tends to grow smoothly while after then there is a vast growth in the conversion process. Increase in the biomass feed reduces the rate of oxygen transfer in the system hence subjecting the aerobic enzymes to dormant or death phase [25-26]. The plot shows that fermentation occurs most successfully within low mass feed rate, and that fermentation process is a factor of the feed mass.

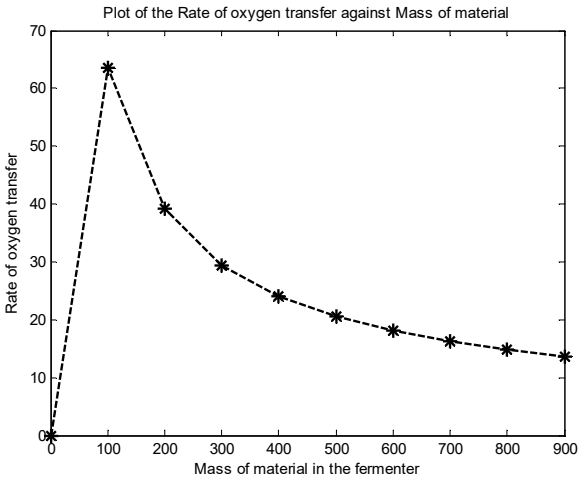


Fig. 10: Rate of Oxygen Transfer Vs mass of Material in the fermenter

IV. Conclusion

The design shows the target been achieved at the fermenter vessel Area of $(A_o) = 15.5917m^2$, with the use of $1271.1646kg/day$ coolent (water), with the system operating at conserved heat flow of $505262.484J/day$. MATLAB simulation is been used to access the dynamic behavior of the agitated fermenter over a range of biomass and time using the developed models, which shows that the reaction rate of fermentation $\left(\frac{dC_L}{dt}\right)$ is proportional to the Overall mass transfer coefficient per unit volume of the biomass (K_La) measured in per hours of the fermentation (hr^{-1}). So decrease in the feed mass of a fermenter due to biomass decay, increases the oxygen penetrability of the system, as shown in fig 9, and this effects promotes the Heat duty of the system over time which calls for the use of a cooling jacket designed fermenter, especially for large scale feeds. This decrease in the fermenter material (biomass) due to biomass decay promotes the fermentation rate of the system, as seen in fig 7 that the rate of biomass fermentation $\left(\frac{dC_L}{dt}\right)$ corresponding decreases with increase in feed value of biomass. The increase in the fermentation rate of the biomass reduces the resident time of the system as described in fig 8, promoting satisfactory product yielding and purity which is the one of the target of this work

Acknowledgment

The authors are grateful to all the laboratory Technologists in both schools and everyone who played a part in the success of this research. .

References

[1]. G. Bank, "Optimize Power Consumption in Aerobic Fermenters," Chem. Engr. Prog., vol. 99, no. 1, pp. 100-103, 2003.

[2]. J.E. Beily and D.F. Ollis, "Biochemical Engineering Fundamentals," 2nd ed. McGram-Hill Pub., New York, 1986, pp. 137

[3]. G. Benz, "Bioreactor design for chemical engineers," Chem. Engr. Proc, vol. 107, no. 8, pp. 21-26, 2011.

[4]. E.M. Bueno, B. Bilgen, R.L. Carrier and G.A. Barabino, "Increased rate of chondrocyte aggregation in a wavy-walled bioreactor," Biotech Bioengr., vol. 88, no.2, pp. 767-77, 2004. https://doi.org/10..1002/bit.20261

[5]. K.A. Cairncross, "Novel reactor design for conversion of waste oils to biodiesel," AlChE Annual Conference, UK, 2011, pp. 212

[6]. Z. Chen, H. Chen, Y. Xu, M. Hu, Z. Hu, J. Wang and Z. Pan, "Reactor for biomass conversion and waste treatment in supercritical water: A review," Revers. Sustain. Energ. Revs., vol. 171, no. 1, pp. 113031, 2023. https://doi.org/10.1016/j.rser.2022.113031

[7]. P.M. Doran, Bioprocess engineering principles, Academic Press, New York, 1995, pp. 212

[8]. J.V.F. Duque, F.L.E. Bittencourt, M.F. Martins and G. Debenest, "Developing a combustion – driven reactor for waste conversion," Energy., vol. 2, pp. 121489, 2021. https://doi.org/10.1016/j.energy.2021.121489

[9]. R. Eibl and E. Dieter, Design of bioreactors suitable for plant cell and tissue cultures. Hill, USA, 2008, pp. 97

[10]. J. Iyyappan, B. Gurunathan, M. Gopinath, A. Vaishasvi, S. Prathiba, V. Karishka, K. Gomathi and V. Dhithya, "Advances and sustainable conversion of waste lignocellulosic biomass into biofuels," Biofuels Bioenergy., vol. 2, no. 2, pp. 167 – 206, 2022. https://doi.org/10.1016/B978-0-323-90040-9.00004-7

[11]. H. Jagani, "Department of Pharmaceutical Biotechnology," Manipal University, ManipalKarnatak-india, 2010, pp. 97

[12]. O. Levenspiel, "Chemical Reaction Engineering," 3rd ed. John Wiley & Son's Inc, U.S.A, 1999, pp. 89

[13]. Y.T. Makkawi, "Reactor design and its impact on performance and products," 1st ed. John Wiley and Son's Ltd., 2014, pp. 98

[14]. L. Martin, D. Wendt and M. Heberer, "The role of bioreactors in tissue engineering,"Trends in Biotech., vol. 22, no. 1, pp.80-86., 2004. https://doi.org/10.1016/j.tibtech.2003.12.001

[15]. M. Martín, F.J. Montes and M.A. Galán, "On the contribution of the scales of mixing to the oxygen transfer in stirred tanks," Chem. Engr. J., vol. 145, no. 2, pp. 232- 241, 2008. https://doi.org/10.1016/J.CEJ.2008.04.019

[16]. T.J. Narendranathan, "Designing fermentation equipment," Chem. Engr., vol. 425, no. 1, pp. 23-31, 1986. https://doi.org/10.113980925

[17]. R.H. Perry and D.W. Green, "Perry's Chemical Engineering Handbook," 8th ed., McGraw, USA, 2007, pp. 222

[18]. B. Redondo, M.T. Shah, V.K. Pareek, R.P. Utikar, P.A. Webley, J. Patel, W.J. Lee and T. Bhatelia," Chem. Engir. Process. – Proc. Intens., vol. 143, no. 1, pp. 107606, 2019.

[19]. F.L.P. Resende, "Reactor configurations and design parameters for thermochemical conversion of biomass into fuels, energy and chemicals," Reactor Process design Sustain. Energ. Technol., vol. 2, no. 1, pp. 1 – 25, 2014. https://doi.org/10.1016/B978-0-444-59566-9.00001-6

[20]. K.T. Samarasirib, "Reactor design for the conversion of waste plastic into fuel oil and gas. Technical Report," University of Moratuwa, Sri Lanka, 2015, pp. 72

[21]. A.H. Scragg, "Bioreactors on Biotechnology. A practical approach," Elis Horwood Series, New York, 1996, pp. 119

[22]. K.R Sharma, "Design and operational key factors of bioreactor," Biotech Articles., vol. 2, no. 1, pp. 119 - 129., 2012.

[23]. M.L. Shulera and F. Kargi, "Bioprocess Engineering, Basic Concepts," Prentice Hall, USA, 1992, pp.211

[24]. P.F. Stanbury and A. Whitaker, "Principles of fermentation technology," Pergamon Press, New York., 1987, pp. 174

[25]. K. Vant and K.J. Riot, "Basic Bioreactor Design," Marcol Dekker, USA, 1991, pp. 74

[26]. J.A. Williams, "Keys to Bioreactor Selection," CEP Magazine., vol. 2, n0. 1, pp. 34-41, 2007.

Fredrick Ugi is a graduate of Chemical / Petrochemical Engineering with the Department of Chemical Engineering, River State University, Nigeria. He is much grounded in the areas of Mathematics, Reaction kinetics, Process modeling, Equipment design and control, Simulations an Optimization. He has contributed massively as a researcher to the growth of multiple projects through his publications in both national and international journals in the areas of both green energy and corrosion science.

Ben Ugi is a seasoned Researcher and Senior Lecturer with the Department of Pure & Industrial Chemistry, University of Calabar, Nigeria. He is a researcher in the area of Corrosion Science employing environmental friendly inhibitors. He researches also in the area of metallurgy, Environmental Chemistry, Computational Chemistry, Electrochemistry and Surface Chemistry. He has published over 60 journal papers. He is currently an Adjunct Lecturer with the University of Cross River State, Nigeria.