

TRƯỜNG ĐẠI HỌC QUY NHƠN  
QUY NHON UNIVERSITY

TẠP CHÍ KHOA HỌC  
TRƯỜNG ĐẠI HỌC QUY NHƠN

QUY NHON UNIVERSITY  
JOURNAL OF SCIENCE

KHOA HỌC TỰ NHIÊN VÀ KỸ THUẬT  
NATURAL SCIENCES AND ENGINEERING

19 (1)

2025

FEBRUARY 2025





## CONTENTS

|     |                                                                                                                                                                                                                                              |     |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 1.  | Research and application of industrial communication networks in the domestic wastewater treatment system for Quy Nhon city<br><b>Do Van Can, Tran Xuan Khoa</b> .....                                                                       | 5   |
| 2.  | New stabilization criteria for time-delayed bimodal piecewise linear systems<br><b>Nguyen Hoang Tuong Vy, Le Quang Thuan</b> .....                                                                                                           | 17  |
| 3.  | Research on data collection and monitoring voltage quality using Ad-hoc LoRa and Siemens PLC through Modbus TCP/IP communication<br><b>Tran Xuan Khoa, Huynh Duc Hoan, Phan Gia Tri</b> .....                                                | 29  |
| 4.  | Weighted composition operators from Bloch-type spaces into growth spaces on the unit ball of a Hilbert space<br><b>Duong Thanh Vy</b> .....                                                                                                  | 49  |
| 5.  | Investigating the application of ginger essential oil in Ramie Leaf Rice Cake preservation<br><b>Vo Thi Trong Hoa, Le Duy Thanh, Ngo Kim Khue, Nguyen Thanh Hang, Le Canh Dinh</b> .....                                                     | 61  |
| 6.  | Effects of water depths on growth and survival of black apple snail ( <i>Pila polita</i> ) in rearing condition<br><b>Ngo Kim Khue, Vo Van Chi</b> .....                                                                                     | 71  |
| 7.  | An alternative equivalent condition for a finite family of normal matrices to be simultaneously diagonalizable via congruence<br><b>Vo Duc Tu Duy, Le Thanh Hieu</b> .....                                                                   | 79  |
| 8.  | Complete spectral analysis of berberine, a key compound in Vietnamese traditional remedy “Đại tràng hoàn Bà Giằng”<br><b>Nguyen Hoang Sa</b> .....                                                                                           | 87  |
| 9.  | The $C_{sp^2}$ -H...O/Se/Te and O-H...Se/Te nonconventional hydrogen bonds in the systems of formic acid with selenoformaldehydes and telluroformaldehydes<br><b>Nguyen Tien Trung, Nguyen Truong An, Nguyen Ngoc Tri, Vu Thi Ngan</b> ..... | 99  |
| 10. | Efficient interpretable prediction of protein-ligand interactions using Gradient Boosting models and explainable AI<br><b>Truong Thi Cam Mai</b> .....                                                                                       | 115 |



# Nghiên cứu và ứng dụng mạng truyền thông công nghiệp trong hệ thống xử lý nước thải sinh hoạt tại thành phố Quy Nhơn

Đỗ Văn Cần\*, Trần Xuân Khoa

*Khoa Kỹ thuật và Công nghệ, Trường Đại học Quy Nhơn, Việt Nam*

*Ngày nhận bài: 29/05/2024; Ngày sửa bài: 04/10/2024;  
Ngày nhận đăng: 16/10/2024; Ngày xuất bản: 28/02/2025*

## TÓM TẮT

Xử lý nước thải là rất quan trọng đối với các khu vực đô thị để đảm bảo bảo vệ môi trường và sức khỏe cộng đồng. Ở thành phố Quy Nhơn, việc quản lý hiệu quả nước thải sinh hoạt liên quan đến việc thu thập nó từ các khu dân cư khác nhau và xử lý tại một cơ sở xử lý tập trung. Việc tích hợp mạng truyền thông công nghiệp và hệ thống giám sát và thu thập dữ liệu (SCADA), vào quá trình này có thể cải thiện đáng kể hiệu quả vận hành và giảm chi phí. Bài báo này nhóm tác giả đã nghiên cứu áp dụng mạng truyền thông công nghiệp trong hệ thống xử lý nước thải của thành phố Quy Nhơn để giảm chi phí vận hành và nâng cao hiệu quả quản lý trong việc xử lý các sự cố vận hành. Kết quả thực nghiệm tại Phòng thí nghiệm Mạng truyền thông công nghiệp và SCADA (314A7), Trường Đại học Quy Nhơn đã cho thấy hiệu quả rất cao.

**Từ khóa:** *Mạng truyền thông công nghiệp, SCADA, trạm bơm, hệ thống xử lý nước thải, thành phố Quy Nhơn.*

*\*Tác giả liên hệ chính.*

*Email: dovancan@qnu.edu.vn*

# Research and application of industrial communication networks in the domestic wastewater treatment system for Quy Nhon city

Do Van Can\*, Tran Xuan Khoa

Faculty of Engineering and Technology, Quy Nhon University, Vietnam

Received: 29/05/2024; Revised: 04/10/2024;

Accepted: 16/10/2024; Published: 28/02/2025

## ABSTRACT

Wastewater treatment is crucial for urban areas to ensure environmental protection and public health. In Quy Nhon city, effective management of domestic wastewater involves collecting it from various residential areas and treating it at a centralized treatment facility. Integrating industrial communication networks and Supervisory Control and Data Acquisition (SCADA) systems into this process can significantly improve operational efficiency and reduce costs. This paper examines the application of industrial communication networks in the wastewater treatment system of Quy Nhon city to reduce operational costs and enhance management efficiency in handling operational incidents. Experimental results from the Industrial Communication Network and SCADA Laboratory (314A7) at Quy Nhon University have shown very high effectiveness.

**Keywords:** Industrial communication networks, SCADA, pumping stations, wastewater treatment system, Quy Nhon city.

## 1. INTRODUCTION

Wastewater management is a fundamental aspect of life and development in modern societies. The collection and treatment of wastewater significantly affect both the environment and the economy, on local and global scales. Thus, it is essential to engage in activities that guarantee proper wastewater management. This is especially true in urban and densely populated

areas, where finding optimal solutions for sewage systems, including wastewater treatment plants, is crucial.<sup>1,2</sup>

The primary objective of a wastewater treatment plant is to treat water from various sources by eliminating its pollutants.<sup>3,4</sup> Typically, this process is carried out in four or five stages, as depicted in Figure 1.

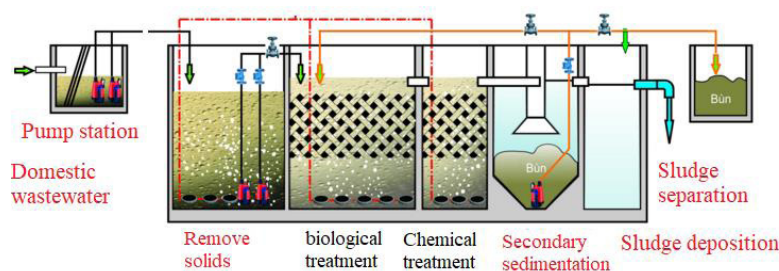


Figure 1. Wastewater treatment process.<sup>3</sup>

\*Corresponding author.

Email: dovanacan@qnu.edu.vn

<https://doi.org/10.52111/qnjs.2025.19101>

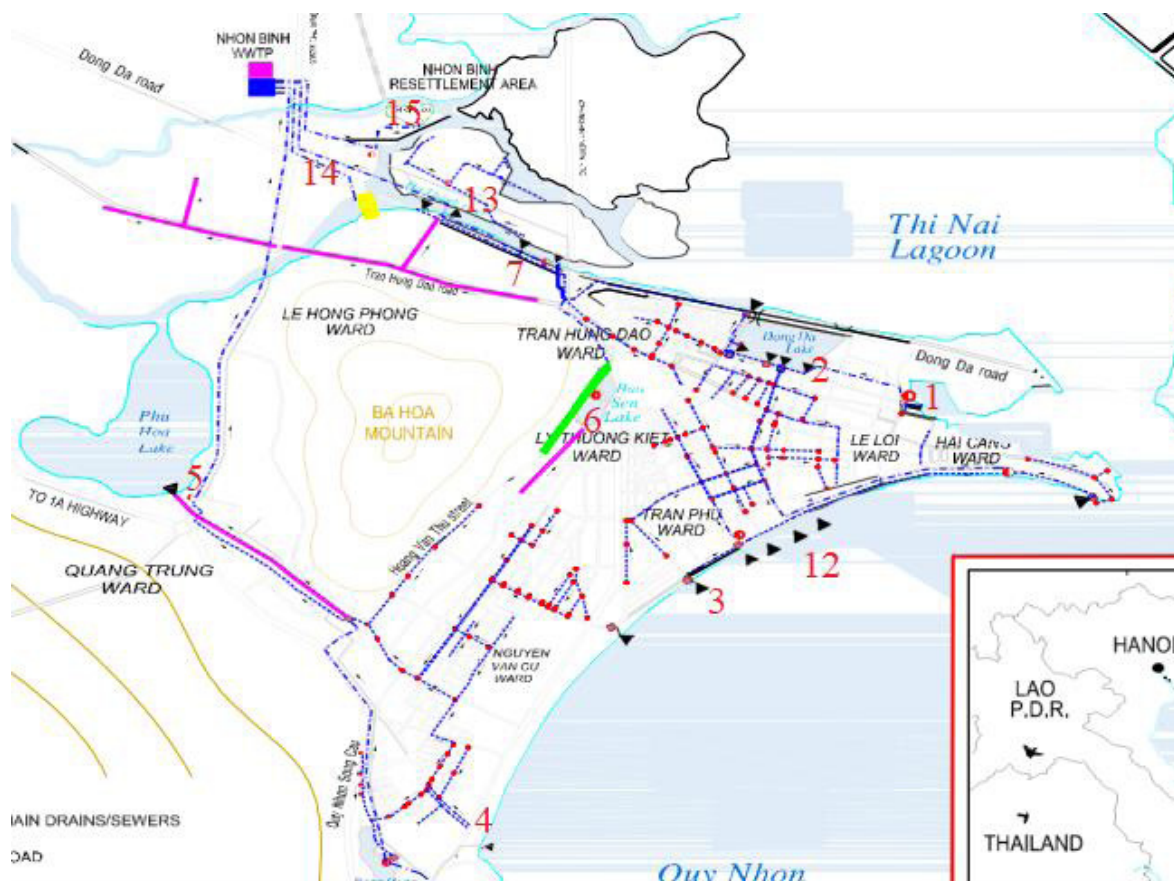
- Pre-treatment, or mechanical treatment, involves removing solids and larger papers from the water.
- Following pre-treatment is biological treatment.
- Chemical treatment occurs after the biological treatment.
- The subsequent step is secondary sedimentation, which entails gathering all the waste and sludge from the treatment process and separating it from the purified water.
- Lastly, the sludge is collected and dried through various methods.

The wastewater treatment system in Quy Nhon city, operational since 2014 following significant investment, has markedly enhanced the city's environmental sanitation. Wastewater is collected and conveyed to the Nhon Binh Treatment Plant, which has a capacity of 14,000 cubic meters per day.

In recent years, a significant number of centralized wastewater treatment plants have been established in both large and medium-sized urban areas, including Hanoi, Bac Ninh, Vinh, and Can Tho, with the support of the Vietnamese Government and various international organizations.

Consequently, the presence of toxins in wastewater poses a significant health hazard to operational personnel. Therefore, the adoption of automated or remote control methodologies emerges as a critical necessity in the management of wastewater pumping and treatment processes.

Presently, the Binh Dinh Civil and Industrial Project Management Board has recommended that the Provincial People's Committee enter into an agreement with URENCO and UDC for the management and operation of the facility and 12 wastewater pumping stations, employing on-site personnel.<sup>5</sup>



**Figure 2.** Diagram of Quy Nhon city wastewater pumping and treatment system.

Wastewater treatment systems globally collect water from various stations and transport it to treatment centers. There is an increasing emphasis on implementing remote management solutions to reduce issues encountered during system operation, particularly concerning pipelines and pumping stations.<sup>6,7</sup> The City of Seattle's wastewater collection system ranks among the largest in Washington State, encompassing sanitary, combined, and partially separated combined sewers. In areas served by sanitary sewers, stormwater runoff is channeled into a storm drainage system, while sewage is transported through city sewers to extensive pipelines and treatment facilities. These larger infrastructures are owned and operated by the King County Department of Natural Resources and Parks (DNRP).<sup>6,8</sup>

The Nhon Binh Wastewater Treatment Plant is a component of the Quy Nhon City Environmental Sanitation Project, financed through non-refundable aid from the global environmental fund. Officially inaugurated in 2017, the plant utilizes advanced negative filtration technology, incorporating air and biological trickling filtration processes.

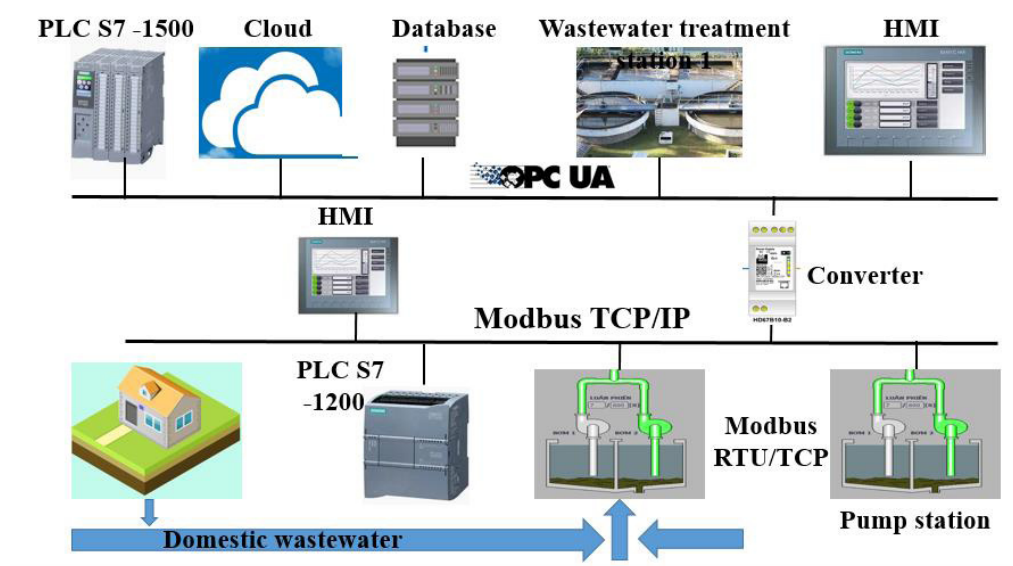
Currently, the new technology application solution (OPC UA) has not been applied to the

water treatment system in general, and Quy Nhon city in particular. The authors' proposal will bring many technical and economic benefits.

The authors propose the adoption of Supervisory Control and Data Acquisition (SCADA) for the wastewater treatment system of Quy Nhon city (Figure 3).

Data is stored in the SCADA database and preprocessed to eliminate outliers and fill in missing values. The process monitoring and control functions are collectively organized, demonstrating their ongoing and collaborative use in the system. This preprocessed data is crucial for detecting and identifying sensor faults, as well as for monitoring the operational processes of the plant.<sup>9,10</sup>

Significant advancements in computer technology have led to the development of both wireless and wired communication protocols, such as MODBUS, OPC DA, and OPC UA. These innovations make it feasible to collect and compile extensive datasets. Consequently, this provides operators, engineers, and maintenance staff with sophisticated analytical tools, facilitating cost reduction, the avoidance of unnecessary repairs, and improved time and plant management. Importantly, these advancements also contribute to safeguarding the health of operators.<sup>11,12</sup>



**Figure 3.** Applying industrial network to manage wastewater treatment system of Quy Nhon city.



- Benefits of wastewater treatment from the SCADA system are:
- Reduce chemical exposure
- Control the amount of wastewater and treat it appropriately
- Economic efficiency
- Integrated control of environmental and weather impacts
- Record the history of the entire process
- Monitor the status of remote pumping systems

## 2. IMPLEMENTATION

### 2.1. Management solution for wastewater collection station in Quy Nhon city

In certain scenarios, wastewater treatment systems are established directly at the discharge location. Within the scope of this study, the oversight of the Wastewater Treatment Plant falls under the jurisdiction of the Department of Power and Utility. The outcomes of the treatment process include water that has been neutralized to a pH level devoid of hazardous and toxic substances, and sludge that is prepared for disposal at a designated site for hazardous and toxic materials.<sup>3</sup> In the city of Paraty, collection systems utilize piping networks. A significant challenge associated with the vacuum system was the anticipated operational and maintenance costs, given the city's limited experience and expertise with such collection systems. It has been determined that for Paraty, adopting a uniform, consistent, and straightforward collection system would be the most suitable approach.

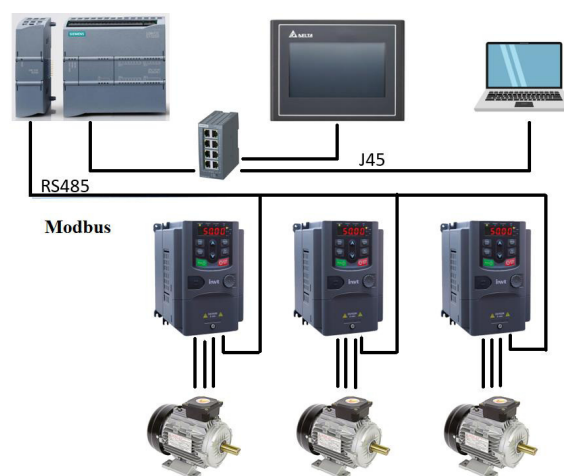
Implementing a comprehensive wastewater management solution for Quy Nhon city entails the integration of monitoring systems, the optimization of collection infrastructure, and the assurance of efficient treatment processes. Collaboration with local authorities, the deployment of sensor networks, and the application of data analytics

are key strategies to improve operational efficiency and the assessment of environmental impacts. To manage the wastewater collection station system effectively, the authors utilize the Modbus communication network (Figure 4).

Modbus is a Client/Server protocol. In simple terms, this means that the server, often a PLC, does not send any data until the client, in this case, the SCADA, requests it. In other words, there is no way for a PLC to send data to the SCADA if a sensor has a change in value. In practice, this means that the SCADA will have to request data from all PLCs to look for a change in the data, even if the data might have been unchanged since the last time. An exception to this is OpenMbus, where the client is able to subscribe to the server which then sends an update every time a value changes.

First, the program from the computer connects via an ethernet network through a switch including PLC, HMI, and computer to upload the program to the PLC. This program controls the motor through the inverter using Modbus RTU communication.

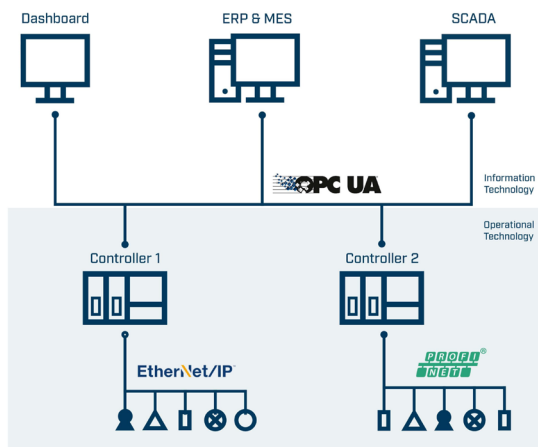
Then, data at the pumping stations will be sent to the wastewater treatment center via the industrial communication network (Ethernet) by fiber optic cable.



**Figure 4.** Wastewater collection station management system.

### 2.3. Communication solutions for the system

To exchange data between pumping stations and processing centers, we create a communication system. They supervisory control and data acquisition. The radius of Quy Nhon city's wastewater pumping and treatment system is about 10 km, suitable for applying Ethernet network based on OPC UA platform.



**Figure 5.** Ethernet industrial communication protocol.

OPC-UA was previously OPC Classic, based solely on technologies developed by Microsoft. Although useful and efficient, the system's reliance on Windows has prevented its development in the modern factory. Released since 2015, OPC UA retains all the functionalities of OPC Classic, except that it is 100% independent of the platform used, simplifying the implementation and maintenance of information systems. Unlike its predecessor, OPC UA can be used on Linux, Mac OS, Microsoft Windows, iOS or Android. Figure 5 is an OPC UA protocol in industrial communication.

Having an increase in determinism within the network will help to perform real-time data transmission. Within the WWTP facilities, measurement devices, control and communication equipment are required to report real-time data to the SCADA system, but other

equipment like VoIP or surveillance cameras will not require this. An industrial managed Ethernet switch with network management software built-in is required in order to increase determinism for all field equipment data transmission.

Set the IP Address with the BOOTP/DHCP Tool. The BOOTP/DHCP tool is a standalone server that you can use to set an IP address. The BOOTP/DHCP tool sets an IP address and other TCP parameters.

The EtherNet/IP Address Commissioning Tool provides additional functionality than the BootP/DHCP tool.

Access the BOOTP/DHCP tool from one of these locations:

- Programs > Rockwell Software® > BOOTP-DHCP Tool > BOOTP-DHCP

- Tools directory on the Studio 5000® environment installation CD.

Add the Device to the Controller Organizer:

After installing a device and setting the IP address, add the device to the Controller Organizer in a programming software project. This addition establishes I/O control.

Before initiating operation, it is essential to download the project to the host controller. Once the controller starts, it establishes a connection with the device, and the device's configuration dictates its behavior. To construct the I/O configuration for a standard I/O network, adhere to the following steps:

1. Add the device.
2. Add the remote device for distributed I/O.
3. Add the I/O modules. This graphic shows the I/O configuration of the consumer controller after distributed I/O modules are added.





1. Make sure that the device is installed, started, and connected to the controller.
2. In the Controller Organizer, right-click the device and choose Properties.

Wastewater treatment encompasses a range of processes and equipment units, including screening, pumping, aeration, sedimentation, filtration, neutralization, chlorination, residual removal, sludge digestion, and effluent discharge. These processes necessitate an automated monitoring system to verify the proper operation of each equipment unit and minimize worker exposure to pollutants, pathogens, or hazardous chemicals. The adoption of Open Platform Communications Unified Architecture (OPC UA Figure 7) can potentially lower the costs associated with engineering and maintaining an IoT network, as well as bolster data security from the field level to the cloud.<sup>13</sup>

An application architecture was designed by letting the application act as both a SCADA and HMI. Depending on what level of the architecture you place it and to where it stores its data, this becomes possible.



Take a pump as an example:

- Power consumption (kW)
- Pump flow (m<sup>3</sup>/h)
- Setpoint for pump flow (m<sup>3</sup>/h)
- Is active (true/false)
- Has any active alarms (true/false)
- Manual mode active (true/false)

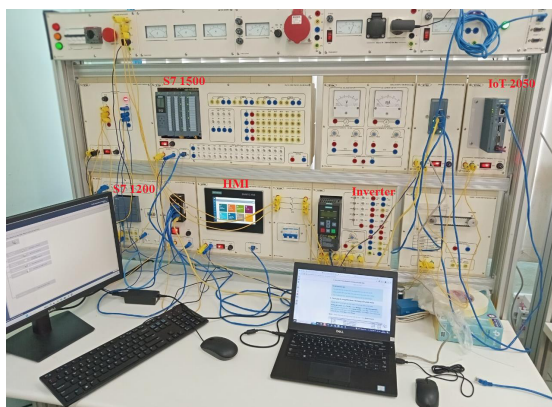
## 2.5. Experimental results and discussion

To evaluate the proposed solution, the authors tested each function of system X in the laboratory. From there, it confirms its feasibility when applied in practice.

The results are divided into three layers where layer one describes the overview of all systems and correlates to level one in the high performance HMI as described section. Layer two correlates to level two in the high performance HMI hierarchy and speaks of sub systems and objects in the workviews. Finally, layer three presents the details of a specific object and correlates to levels three and four of the hierarchy.

### - Experimental equipment

The authors introduce some of the main equipment available in the laboratory to serve the experimental process of this research content. The main control unit for the treatment plant is the S7-1500, IoT 2050 Ethernet communication module, the main control of the pumping station is the S7-1200 and the pump control inverter as Figure 8.



**Figure 8.** Experimental model of wastewater pumping and treatment system.

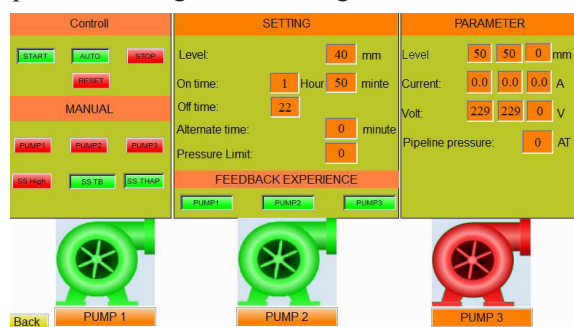
### - Modbus communication for pumping stations

The model is implemented with three inverters, each linked to a specific wastewater pump responsible for transferring water from the storage tank to the pressure pipeline. The control and programming for this pumping station setup are executed using the PLC S7-1200 (Figure 9).



**Figure 9.** Pump station monitoring and control model based on Modbus RTU network.

Parameters are within allowable limits and are updated continuously every 1 s, ensuring real-time calculation for shrimp pond monitoring systems. In addition, the above system can provide a common database for scientists, shrimp hatcheries, fishery extension centers and so on. In addition, the above data is also daily, monthly, and yearly graphed to store historical parameters Figure 9, and Figure 10.



**Figure 10.** Results of running the model in Auto mode.

Start the system running in Auto mode, and set the system's on/off time, alternating time, and pipe pressure limit. When the low level sensor receives, pumps 1 and 2 run alternately according to the set time and frequency. If the level sensor receives an average value, both pumps 1 and 2 will run at the set frequency.

If the level sensor receives a high value, all 3 pumps 1, 2 and 3 will run at the set frequency. If the pipeline pressure exceeds the set pressure, the system will stop the pump and warn.

During operation, the system continuously monitors the pump's operating frequency, voltage, and current.

- SCADA system for wastewater treatment plant

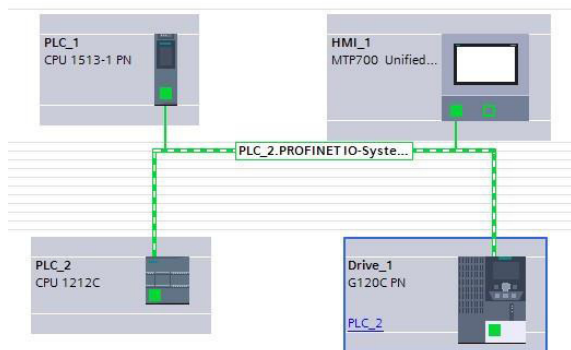


Figure 11. Experimental model of control center.

In wastewater treatment plants, the SCADA system operates in a more complex manner, with the program being implemented on the PLC S7-1500.

Like the pumping station, the treatment plant features a vast array of inputs and outputs, facilitating extensive data exchange. The SCADA system onsite constructs parameters, enabling data communication with 12 pumping stations via the Ethernet network. This setup, grounded in the OPC UA platform, ensures compatibility across all devices and various operating systems. Some parameters were selected to experiment with the research content, to reflect the applicability of the content of this paper (Figure 12).

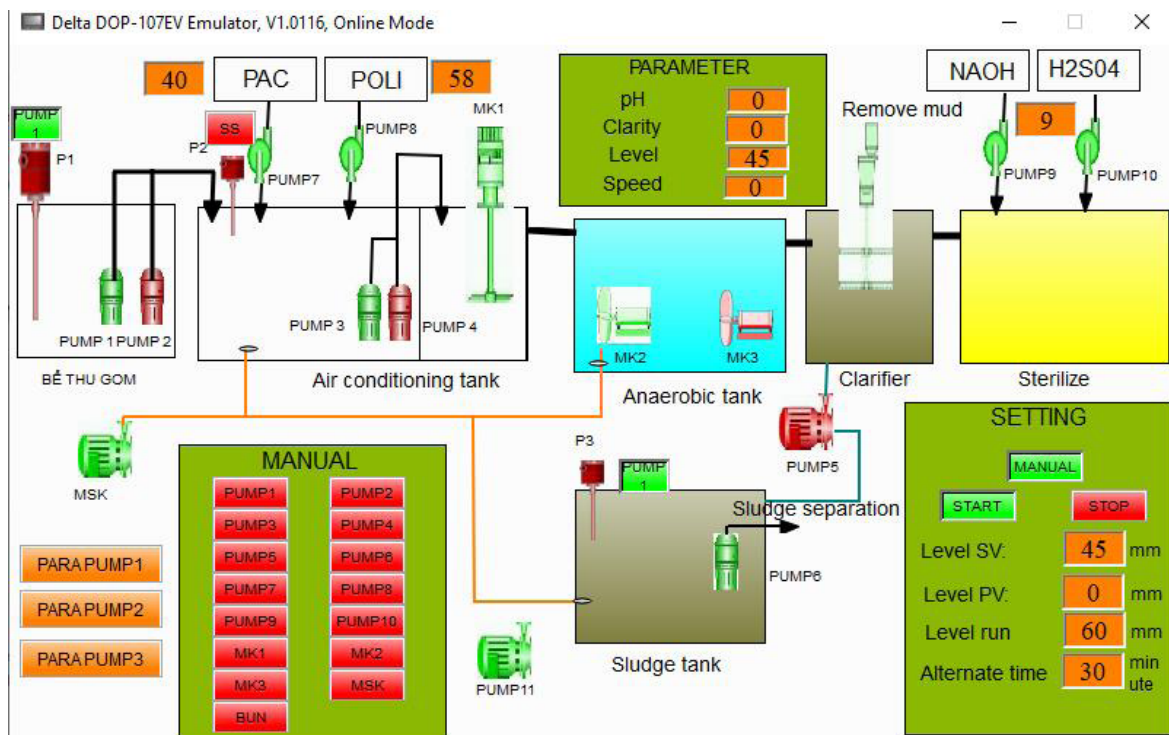


Figure 12. Experimental results of SCADA system at wastewater treatment plant.

If SS P1 is in state 1 ( $P1=1$ ) and SS P2 is in state 0 ( $P2=0$ ), then pump 1 and pump 2 are eligible to operate. These two pumps operate alternately according to the set time, progressing.

Pump water from the collection tank to the equalization tank to regulate the amount of water. Next, pump 3 and pump 4 in the equalization tank will operate alternately to pump water to



the tank where the stirrer is located. If pump 3 or pump 4 operates, MK1 will operate. Pump 7 and pump 8 will operate for x minutes and rest for x minutes (x is the previously set time).

MK2 and MK3 operate alternately. At the sedimentation tank, the sludge scraper will operate for 10 minutes and rest for 5 minutes. The sterilization tank and chemical pumps will operate according to the set time. At the sludge tank, if SS P3 ( $P3=0$ ) has not yet been activated, pump 5 operates to pump the amount of water and sludge settled at the bottom of the tank to the sludge tank. At the sludge tank, pump 6 operates to pump the unsatisfactory amount of water. The treatment request is returned to the collection tank to continue the treatment process. Otherwise, when P3 ( $P3=1$ ) is activated, pump 5 stops operating, and at this time, pump 11 operates to pump the sludge to the sludge drying yard for treatment. mud. SS-pH and SS-sludge sensors are two sensors for measuring water pH and turbidity.

At the disinfection tank, pump 9 and pump 10 will pump CLO and NaOH chemicals to disinfect water to meet standards before being discharged into the environment. The aerator operates for 10 minutes and has a 5-minute break to blow air into the tanks in the system.

In case SS P2 is active ( $P2=1$ ), pump 1 and pump 2 in the collection tank stop operating and stop the process of pumping water to the equalization tank to avoid filling the tank and draining water, when the amount of water is pumped to another tank of the tank. The system reaches the point where sensor P2 is no longer active ( $P2=0$ ), then pumps 1 and 2 will operate normally again.

In case SS P3 operates, pump 5 stops working and pump 11 operates and will pump the sludge from the sludge tank to the sludge drying tank for treatment...

If an unexpected problem occurs during system operation, the system alarms and the operator presses Stop. After fixing the problem,

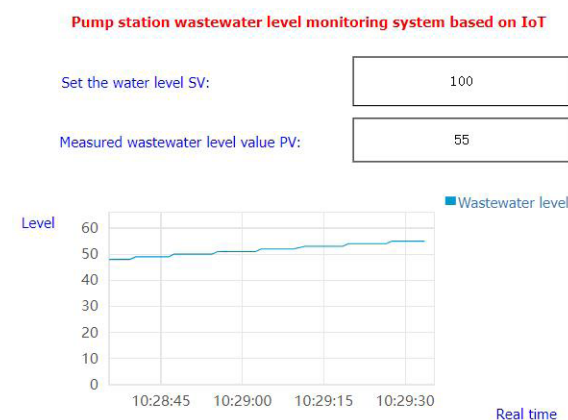
the system continues to operate again when we press Start.

### - IoT communications for wastewater treatment systems

Equipment for selecting experimental configurations in the laboratory is V-BOX H-WF, Server-based <https://web.asean.v-box.net>

**Figure 13.** Experimental configuration of IoT communication system.

The monitored parameter is the wastewater level at 12 internal pumping stations. The optimal control algorithm for the pumping system is mentioned in other content. The experimental value is set at 0 - 100 cm, corresponding to the monitored wastewater level and pile pressure so that the control system at the stations can operate 1 - 3 pumps.



**Figure 14.** Monitor wastewater levels at pumping stations.

The results show that controlling remote values through research content works stably. Parameters can be displayed and set remotely via the S7-1200 PLC control system (Figure 14).

### 3. CONCLUSION

The paper has achieved the following results:

The authors offer a solution to manage the entire wastewater pumping and treatment system for Quy Nhon City.

At pumping stations, the paper proposes to use the Modbus industrial network system to manage pumps and pipeline pressure.

At the central processing plant, the paper proposes to use an industrial Ethernet network system based on the OPC UA platform to manage the processing system and upload data to the Cloud.

It uses industrial IoT communications to link data exchange between the pumping station and the wastewater treatment center.

After the paper's content is deployed, it brings many benefits such as: Increasing system reliability, managing data via Cloud, automatically controlling the amount of wastewater, especially reducing health damage to operators.

Experimental results at the Industrial Network and SCADA Laboratory of Quy Nhon University, the research results of the paper are applicable in practice.

### Acknowledgment

*This research is conducted within the framework of science and technology projects at institutional level of Quy Nhon University under the project code T2024.845.15.*

### REFERENCES

1. A. J. Gusman, P. E. Project manager RFP. *Water distribution and wastewater collection system technical study*, Support of the Program Environmental Impact Report for the Mira Mesa Community Plan Update, 2022.
2. V. C. Georgescu. Operations management in water and wastewater treatment plants, *Applied Mechanics and Materials*, **2013**, 245, 179-184.
3. R. Arbour. *Basic wastewater collection systems*, Training and Information Management Unit, 1995.
4. K. Neelshetty, Avinash, Jaleel, M. Jaleel, M. Mujeeb, Sharanappa. Waste water treatment using PLC and SCADA, *Iconic Research and Engineering Journals*, **2018**, 1(11), 24-27.
5. Project Management Board for investment in construction of civil and industrial works in Binh Dinh province. *Sustainable environment project for coastal cities*, The Report Recommends Granting Environmental Permits Coastal Cities Sustainable Environment Project, 2023.
6. M. Andres, C. Jagroop, C. Maharaj. Wastewater collection system failures in a capital city: analysis and sustainable prevention, *Water Science and Technology*, **2021**, 83(8), 1958-1972.
7. R. Basreen, R. P. Chavan, U. Rani. Dam water level monitoring and alerting system using IoT, *Journal of English Sciences*, **2022**, 13(06), 707-713.
8. S. Verma. *Wastewater collection system*, Minnesota Pollution Control Agency Rick Arbour, 2023.
9. M. J. Wade, A. Sánchez, M. R. Katebi. On real-time control and process monitoring of wastewater treatment plants: real-time process monitoring, *Transactions of the Institute of Measurement and Control*, **2005**, 27(3), 173-193.
10. J. Lloret, L. Garcia, J. M. Jimenez, S. Sendra, P. Lorenz. New protocol and architecture for a wastewater treatment system intended for irrigation, *MDPI Applied Sciences*, **2021**, 9, 142374-142389.
11. N. T. Duy, T. D. H. Duc, B. V. Vu, D.V. Can. Applying of image processing technology in classifying tomatoes after harvest, *Quy Nhon University Journal of Science*, **2020**, 14(5), 59-67.

12. B. V. Vu, N. D. Thinh, P. T. Binh, D. V. Can. Application of raspberry and PLC Mitsubishi to the sign of tomato classified system automatically, *Quy Nhon University Journal of Science*, **2022**, 15(5), 13-20.
13. B. Barig, K. T. H. Balzereit. *Applying OPC-UA for factory-wide industrial assistance systems*, 15<sup>th</sup> IEEE International Workshop on Factory Communication Systems, Sweden, 2019.



© 2025 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

# Một số tiêu chuẩn mới về ổn định hóa đối với hệ tuyến tính từng phần hai mô hình có trễ thời gian

Nguyễn Hoàng Tường Vy<sup>1</sup>, Lê Quang Thuận<sup>2,\*</sup>

<sup>1</sup>Bộ môn Công nghệ thông tin, Trường Đại học FPT, Campus Quy Nhơn, Việt Nam

<sup>2</sup>Khoa Toán và Thống kê, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 24/07/2024; Ngày sửa bài: 31/10/2024;

Ngày nhận đăng: 04/11/2024; Ngày xuất bản: 28/02/2025

## TÓM TẮT

Trong bài báo này, chúng tôi nghiên cứu tính ổn định và ổn định hóa của các hệ tuyến tính từng phần hai mô hình có trễ thời gian bằng cách sử dụng các phiếm hàm Lyapunov-Krasovskii tron. Những đóng góp mới của bài báo bao gồm: (1) thiết lập một tiêu chuẩn ổn định mới dựa trên phiếm hàm Lyapunov-Krasovskii tron để đảm bảo tính ổn định tiệm cận của hệ điều khiển trong trường hợp không có điều khiển đầu vào và (2) đề xuất một điều kiện đủ cho sự tồn tại một điều khiển ngược tuyến tính trạng thái để ổn định tiệm cận hệ thống khi có điều khiển đầu vào. Cuối cùng, một số ví dụ số được chọn lọc để minh họa cho tính hiệu quả của phương pháp đã đề xuất.

**Từ khóa:** Hệ tuyến tính từng phần có trễ thời gian, phiếm hàm Lyapunov-Krasovskii bậc hai từng phần, ổn định hóa hệ hai mô hình.

\*Tác giả liên hệ chính.

Email: lequangthuan@qnu.edu.vn



# New stabilization criteria for time-delayed bimodal piecewise linear systems

Nguyen Hoang Tuong Vy<sup>1</sup>, Le Quang Thuan<sup>2,\*</sup>

<sup>1</sup>Information Technology Division, FPT University, Quy Nhon Campus, Vietnam

<sup>2</sup>Department of Mathematics and Statistics, Quy Nhon University, Vietnam

Received: 24/07/2024; Revised: 31/10/2024;

Accepted: 04/11/2024; Pubished: 28/02/2025;

## ABSTRACT

In this paper, we study the stability and stabilization of time-delayed bimodal piecewise linear systems via smooth Lyapunov-Krasovskii functionals. The main contributions of the paper are twofold: (1) a new stability criterion based on the proposed smooth Lyapunov-Krasovskii functional is derived to guarantee asymptotic stability in the case of zero inputs and (2) an interesting condition is proposed to design linear state feedback controllers to stabilize the system which is less conservative than those previously reported in the literature. Finally, some numerical examples illustrate the effectiveness of proposed methods.

**Keywords:** Time-delayed piecewise linear systems, piecewise quadratic Lyapunov-Krasovskii functionals, bimodal system stabilization.

## 1. INTRODUCTION

In recent decades, piecewise affine (PWA) systems have received much attention in the field of system and control theory. Each PWA system can be seen as a switching one that is characterized by a finite collection of affine time-invariant dynamics together with a state-dependent switching law that is ruled by a polyhedral partition of the state space.<sup>1</sup> PWA systems also form an important subclass of hybrid systems and they can be found in several engineering applications: power converters, robotics, relay control systems, etc. PWA systems are also interesting models to be used for approximating complex nonlinear dynamics. Analysis and design of PWA systems are therefore important as a first step to establish hybrid control theory.

Among the fundamental problems of system theory, the issues concerning stability and stabilization of PWA systems have been intensively studied for both cases: without and with time delays. For the first

case, these problems are well-studied such as for general vector fields in the papers,<sup>2,3</sup> and for continuous ones in the papers.<sup>4-6</sup> With the appearance of time delays, there have been also existed many works developed over the past years, for instance, in the papers.<sup>7,8</sup> In the paper,<sup>7</sup> the authors investigated a class of piecewise time-delayed systems by using piecewise quadratic functions to derive stability criteria in term of LMIs and matrix equations. However, with these employed results, one can not solve the issue of state feedback controllers design to stabilize the systems. The paper<sup>8</sup> has proposed a method to design a piecewise linear state feedback controller to make the closed-loop system asymptotically stable. In this research direction, there are some restrictions. The first one is that from the LMIs combined with matrix equations that guarantees the continuity of Lyapunov functions for stability it is difficult to develop results about feedback controller designs. The second one is that the system under consideration requires non-Zeno behaviors. Note that the non-Zeno property has been established for continuous piecewise affine sys-

---

\*Corresponding author.

Email: lequangthuan@qnu.edu.vn

tems without time-delays in the papers.<sup>9-11</sup> However, one can not obtain similar results for the case of time-delayed PWA systems. Therefore, checking for non-Zenoness of time-delayed PWA systems becomes an impossible task.

Motivated by the above mentioned challenges, we study the stabilization of time-delayed bimodal piecewise linear systems in this paper. Our approach also uses piecewise quadratic Lyapunov-Krasovskii functionals. However, the functionals are developed such that asymptotic stability works for more general solution concepts, *i.e.*, Carathéodory solutions, and the obtained LMIs can be employed to design a linear state feedback controller to stabilize the system. The main contributions are that by employing the special structure of such a functional, stability criteria will be derived for continuous bimodal time-delayed piecewise linear systems. Moreover, the derived LMIs can be employed to design a linear state feedback controller preserving continuity and stabilizing the system. It is worth to mention that there are only a few papers studying the stabilization of PWA systems by linear state feedback controllers taking quadratic Lyapunov function.<sup>4</sup> Finally, our approach is therefore hopefully generalize for more general multi-modal piecewise affine systems.

The rest of this paper is organized as follows. Section 2 introduces notations and preliminaries. In Section 3, we introduce time-delayed bimodal piecewise linear systems and present related preliminaries. This will be followed by stating and proving the main results of stability issue and stabilization of continuous time-delayed PWL systems in Section 4. The proposed theoretical results are validated by numerical examples in Section 5, before concluding the paper in Section 6.

## 2. NOTATIONS AND PRELIMINARIES

Denoted by  $\mathbb{R}$  the set of all real numbers,  $\mathbb{R}_+$  the set of all non-negative real numbers, and  $\mathbb{R}_+^n$  the set of all  $n$ -tuple non-negative real numbers. The notation  $\mathbb{R}^{n \times m}$  denotes the set of all real  $n \times m$  matrices and the transpose of a real matrix  $M \in \mathbb{R}^{n \times m}$  is denoted by  $M^T$ . The notation  $\text{He}(M)$  stands for the matrix  $M + M^T$ . A symmetric matrix  $Q \in \mathbb{R}^{n \times n}$  is said to be positive definite, writing  $Q > 0$ , if  $x^T Q x > 0$  for all non-zero  $x \in \mathbb{R}^n$ . We write  $Q < 0$  if  $-Q > 0$ . For a positive definite matrix  $Q$ , the notation  $\lambda(Q)$  stands for its the maximum eigenvalue. For  $\tau > 0$ ,

$C([-\tau, 0], \mathbb{R}^n)$  denotes the normed space of continuous functions from  $[-\tau, 0]$  to  $\mathbb{R}^n$  endowed with the norm

$$\|\varphi\|_C := \max\{\varphi(s) \mid s \in [-\tau, 0]\}.$$

Also,  $C^1([-\tau, 0], \mathbb{R}^n)$  denotes the space of continuously differentiable functions from  $[-\tau, 0]$  to  $\mathbb{R}^n$ .

Next, we introduce the following auxiliary results.

**Lemma 2.1.** *Let  $P_1, P_2 \in \mathbb{R}^{n \times n}$  be symmetric matrices. The piecewise quadratic function*

$$F(x) = \begin{cases} x^T P_1 x & \text{if } c^T x \leq 0, \\ x^T P_2 x & \text{if } c^T x \geq 0 \end{cases}$$

is

a) *continuous if and only if there exist  $h \in \mathbb{R}^n$  such that*

$$P_2 = P_1 + hc^T + ch^T. \quad (1)$$

b) *continuously differentiable if and only if there exist  $\gamma \in \mathbb{R}$  such that*

$$P_2 = P_1 + \gamma cc^T. \quad (2)$$

*Proof.* a) On the one hand, the function  $F$  is continuous if and only if the following implication holds

$$c^T x = 0 \implies x^T (P_1 - P_2) x = 0.$$

On the other hand, this implication is equivalent to the existence of  $h \in \mathbb{R}^n$  such that

$$\begin{aligned} x^T (P_1 - P_2) x &= 2(c^T x)(h^T x) \\ &= x^T (hc^T + ch^T) x \end{aligned}$$

for all  $x \in \mathbb{R}^n$ , *i.e.*,  $P_1 - P_2 = hc^T + ch^T$ . Therefore,  $F$  is continuous if and only if (1) holds for some  $h \in \mathbb{R}^n$ .

b) It can be seen that  $F$  continuously differentiable if and only if

$$c^T x = 0 \implies P_1 x = P_2 x. \quad (3)$$

Thus, it suffices to show the equivalence between (3) and (2). The implication "(2)  $\Rightarrow$  (3)" is obvious. To prove its converse, the condition (3) implies the existence of  $p \in \mathbb{R}^n$  such that

$$P_1 - P_2 = pc^T.$$

Then, due to the symmetry of  $P_1, P_2$ , we further get  $pc^T = cp^T$  and hence

$$p = \frac{(cp^T)c}{c^T c} = \frac{c(p^T c)}{c^T c} = \frac{p^T c}{c^T c} c = \gamma c$$

where  $\gamma = \frac{p^T c}{c^T c} \in \mathbb{R}$ . □

**Lemma 2.2.** For any vectors  $x, y \in \mathbb{R}^n$  and positive definite matrix  $P \in \mathbb{R}^{n \times n}$ , one has

$$2x^T y \leq x^T P x + y^T P^{-1} y.$$

**Lemma 2.3.** Let  $X, Y \in \mathbb{R}^{m \times n}$ . Let  $F \in \mathbb{R}^{m \times m}$  be a symmetric positive definite matrix. Then, one has

$$X^T F Y + Y^T F X \leq \gamma X^T F X + \gamma^{-1} Y^T F Y$$

for any  $\gamma > 0$ .

*Proof.* It follows from the fact that

$$\left( \sqrt{\gamma} X - \sqrt{\gamma^{-1}} Y \right)^T F \left( \sqrt{\gamma} X - \sqrt{\gamma^{-1}} Y \right) \geq 0,$$

due to the positive definite property of matrix  $F$ .  $\square$

The last lemma is well-known in the field of linear control theory called S-lemma that provides a characterization of symmetric positive definite matrices using Schur complements.

**Lemma 2.4** <sup>(12)</sup>. Let  $Q \in \mathbb{R}^{m \times m}$ ,  $R \in \mathbb{R}^{k \times k}$ ,  $S \in \mathbb{R}^{m \times k}$ . The following statements are equivalent.

- $\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} > 0$ ;
- $R > 0$  and  $Q - S R^{-1} S^T > 0$ ;
- $Q > 0$  and  $R - S^T Q^{-1} S > 0$ .

### 3. TIME-DELAYED BIMODAL PIECEWISE LINEAR SYSTEMS

Consider the time-delayed bimodal piecewise linear systems with inputs

$$\dot{x}(t) = \begin{cases} A_1 x(t) + A_d x(t - \tau) + B u(t) & \text{if } c^T x(t) < 0, \\ A_2 x(t) + A_d x(t - \tau) + B u(t) & \text{if } c^T x(t) \geq 0, \end{cases} \quad (4a)$$

$$x(s) = \varphi(s), s \in [-\tau, 0] \quad (4b)$$

where  $x \in \mathbb{R}^n$  is the state and  $\dot{x}(t)$  denotes its derivative with respect to time  $t$ ,  $u \in \mathbb{R}^m$  is the input, the positive number  $\tau$  is the time delay, the matrices  $A_1, A_2, A_d \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times m}$  and  $c \in \mathbb{R}^n$  are given. The initial function  $\varphi(s)$  is in  $C([-\tau, 0], \mathbb{R}^n)$ . For this work, the right-hand side of (4a) is assumed to be continuous; equivalently, there exists  $e \in \mathbb{R}^n$  such that

$$A_1 - A_2 = e c^T. \quad (5)$$

**Definition 3.1.** Consider the system (4) for a given continuous input  $u \in C(\mathbb{R}_+, \mathbb{R}^m)$ . A continuous function  $x : [-\tau, \infty) \rightarrow \mathbb{R}^n$  is said to be a solution of system (4) for the initial function  $\varphi \in C([-\tau, 0], \mathbb{R}^n)$  if  $x(s) = \varphi(s), \forall s \in [-\tau, 0]$ ,  $x$  is differentiable on  $(0, \infty)$  and satisfies (4a) for all  $t \geq 0$ .

Note that the existence and uniqueness of such a solution follow from the theory of non-homogeneous ordinary differential equations with continuous right-hand-sides. The corresponding solution is denoted by  $x^u(t; \varphi)$ . In the case that  $u(t) \equiv 0$ , it is simply denoted by  $x(t; \varphi)$ .

**Remark 3.1.** For  $\tau = 0$ , the system (4) boils down to bimodal PWA systems that is the main object studied in the paper<sup>4</sup> for stability and stabilization.

**Definition 3.2.** We consider the system (4) without inputs, i.e.,  $u(t) \equiv 0$ . The system (4) is said to be

- stable if for any  $\epsilon > 0$ , there exists  $\delta > 0$  such that  $\|\varphi\|_C < \delta \Rightarrow \|x(t)\| < \epsilon, \forall t \geq 0$ .
- asymptotically stable if it is stable and there is a positive number  $\delta_1$  such that

$$\|\varphi\|_C < \delta_1 \Rightarrow \lim_{t \rightarrow \infty} \|x(t)\| = 0.$$

To study the stability of system (4), we will employ continuous functionals as follows.

**Definition 3.3.** A continuous function  $w : \mathbb{R}^n \rightarrow \mathbb{R}$  is said to be positive definite if  $w(0) = 0$  and  $w(x) > 0$  for all  $x \in \mathbb{R}^n$ .

**Definition 3.4.** We say that a continuous functional  $V : C([-\tau, 0], \mathbb{R}^n) \rightarrow \mathbb{R}$  is positive definite if  $V(0) = 0$  and there exists a positive-definite function  $w : \mathbb{R}^n \rightarrow \mathbb{R}$  such that  $w(\phi(0)) \leq V(\phi)$  for all  $\phi \in C([-\tau, 0], \mathbb{R}^n)$ .

The following proposition yields a sufficient condition to ensure the asymptotic stability of system (4).

**Proposition 3.2** <sup>(13)</sup>. Consider the system (4) without inputs. The system (4) is asymptotically stable if there exist a positive-definite functional  $V(\phi)$  and a positive-definite function  $w(x)$  such that the value of the functional along any selections  $x_t$  of solution  $x(t)$  of the system is differentiable by  $t$ , and its time derivative satisfies the inequality

$$\frac{dV(x_t)}{dt} \leq -w(x(t)) \text{ for all } t \geq 0,$$

where  $x_t(s) := x(t + s), s \in [-\tau, 0]$ .

## 4. MAIN RESULTS

In this section, we will provide a novel method to design state feedback controllers for stabilizing system (4). To do so, a suitable class of smooth Lyapunov-Krasovskii functionals (LKF) is first introduced and discussed. By employing the proposed LKFs, a new criterion on asymptotic stability is derived only in term of linear matrix inequalities (LMIs). Then, these LMIs are used to design a linear state feedback controller to stabilize the system.

### 4.1. Lyapunov-Krasovskii functionals

In the literature of time-delayed piecewise affine systems, piecewise quadratic LKFs have been often used to study the stability of systems.<sup>6-8</sup> Such a functional is basically composed from two parts: a piecewise quadratic Lyapunov function and an integral functional defined on the space  $C^1([-\tau, 0], \mathbb{R}^n)$ , with  $\tau$  is the delay, as

$$V(\varphi) = V_1(\varphi(0)) + V_2(\varphi), \forall \varphi \in C^1([-\tau, 0], \mathbb{R}^n), \quad (6)$$

where the functional  $V_2 : C^1([-\tau, 0], \mathbb{R}^n) \rightarrow \mathbb{R}_+$  is defined as

$$V_2(\varphi) = \int_{-\tau}^0 \varphi^T(s) Q \varphi(s) ds + \int_{-\tau}^0 \int_{\eta}^0 \dot{\varphi}^T(s) R \dot{\varphi}(s) ds d\eta,$$

for some positive definite matrices  $Q, R$  and the quadratic piecewise Lyapunov function  $V_1 : \mathbb{R}^n \rightarrow \mathbb{R}_+$  is defined corresponding with a given polyhedral subdivision  $\{\mathcal{X}_i\}_{i=1}^k$  of  $\mathbb{R}^n$ ; that is,  $V_1(x) = x^T P_i x$  whenever  $x \in \mathcal{X}_i$ . The matrices  $P_i$  are often chosen in such a way that  $V_1$  is positive definite and it is continuous across region boundaries<sup>6-8</sup>. In our point of view, the restriction when one uses this kind of Lyapunov functions is that the stability of solutions only can be applied for the systems whose trajectories do not have Zeno property. To our best of knowledge, checking for non-Zeno property of time-delayed bimodal piecewise linear systems is impossible since there is no available paper about the non-Zenoness of time-delayed piecewise linear systems. Therefore, in this work, we develop Lyapunov-Krasovskii functionals in two aspects: requirements that  $V_1$  is continuously differentiable and relaxation on the integral functional that it has more general form of piecewise quadratic one. It turns out that such requirements impose certain relations on the involved matrices in the literature of bimodal piecewise linear systems.

### 4.2. Stability analysis

By employing the proposed smooth piecewise quadratic LKFs, we now establish a novel stability criterion presented in term of linear matrix inequalities.

**Theorem 4.1.** *For system (4), suppose that there exist the symmetric positive definite matrices  $P, Q, R \in \mathbb{R}^{n \times n}$  and  $h \in \mathbb{R}^n, \gamma \in \mathbb{R}$  such that the following statements hold*

$$P + \gamma c c^T > 0, Q + h c^T + c h^T > 0, \quad (7a)$$

$$\begin{cases} 2\tau A_d^T R A_d^2 - Q < 0, \\ 2\tau A_d^T R A_d^2 - (h c^T + c h^T + Q) < 0, \end{cases} \quad (7b)$$

$$\begin{bmatrix} \Phi_1 & \tau P \\ \tau P & -\tau R \end{bmatrix} < 0, \quad (7c)$$

$$\begin{bmatrix} \Phi_2 & \tau(P + \gamma c c^T) \\ \tau(P + \gamma c c^T) & -\tau R \end{bmatrix} < 0, \quad (7d)$$

where  $\Phi_1 = \text{He}(P(A_1 + A_d)) + Q + 2\tau A_1^T A_d^T R A_d A_1$  and  $\Phi_2 = \text{He}\{(P + \gamma c c^T)(A_2 + A_d)\} + Q + \text{He}(h c^T) + 2\tau A_2^T A_d^T R A_d A_2$ . Then, the system (4) is asymptotically stable.

*Proof.* Let us consider the piecewise quadratic Lyapunov-Krasovskii functional

$$V : C^1([-\tau, 0], \mathbb{R}^n) \rightarrow \mathbb{R}$$

defined by

$$V(\varphi) = V_1(\varphi(0)) + V_2(\varphi),$$

where

$$V_1(z) = \begin{cases} z^T P_1 z & \text{if } c^T z \leq 0, \\ z^T P_2 z & \text{if } c^T z \geq 0, \end{cases}$$

with  $P_1 = P, P_2 = P + \gamma c c^T$  and

$$V_2(\varphi) = \int_{-\tau}^0 F(\varphi(s)) ds + \int_{-\tau}^0 \int_{\eta}^0 \dot{\varphi}^T(s) A_d^T R A_d \dot{\varphi}(s) ds d\eta.$$

where

$$F(z) = \begin{cases} z^T Q_1 z & \text{if } c^T z \leq 0, \\ z^T Q_2 z & \text{if } c^T z \geq 0, \end{cases}$$

with  $Q_1 = Q, Q_2 = Q + h c^T + c h^T$ . Due to (7a) and Lemma 2.1,  $V_2$  is continuous and positive. Thus, one has

$$V(\varphi) = V_1(\varphi(0)) + \int_{-\tau}^0 F(\varphi(s)) ds$$

$$+ \int_{-\tau}^0 \int_{\eta}^0 \dot{\varphi}^T(s) A_d^T R A_d \dot{\varphi}(s) ds d\eta \\ \geq V_1(\varphi(0)), \forall \varphi \in C^1([-\tau, 0], \mathbb{R}^n).$$

This inequality shows that the functional  $V$  is positive definite on  $C^1([-\tau, 0], \mathbb{R}^n)$ .

For any initial function  $\varphi \in C([-\tau, 0], \mathbb{R}^n)$ , let  $x(t; \varphi)$  be the corresponding trajectory of system (4) and define  $x_t(s) = x(t + s; \varphi)$ ,  $s \in [-\tau, 0]$ . Then,  $x_t \in C^1([-\tau, 0], \mathbb{R}^n)$  for each  $t \geq \tau$  and by simple transformations we get

$$V(x_t) = V_1(x(t; \varphi)) + \int_{t-\tau}^t F(x(s; \varphi)) ds \\ + \int_{-\tau}^0 \int_{t+\eta}^t \dot{x}^T(s; \varphi) A_d^T R A_d \dot{x}(s; \varphi) ds d\eta$$

for all  $t \geq \tau$ . Using Newton-Leibnitz formula, one can verify that

$$\dot{x}(t; \varphi) = G(x(t, \varphi)) - A_d \int_{t-\tau}^t \dot{x}(s; \varphi) ds, \quad (8)$$

where

$$G(x) = \begin{cases} (A_1 + A_d)x, c^T x \leq 0 \\ (A_2 + A_d)x, c^T x \geq 0. \end{cases}$$

Observe that by Lemma 2.1 and due to (7a),  $V_1$  is continuously differentiable with respect to  $x$ . On the other hand,  $x(t) := x(t; \varphi)$  is continuously differentiable with respect to  $t$ . Therefore,  $V(x_t)$  is continuously differentiable as a function of variable  $t$  defined on  $[\tau, \infty)$  and its derivative is computed as

$$\frac{d}{dt} V(x_t) = \frac{d}{dt} \left\{ V_1(x(t)) + \int_{t-\tau}^t F(x(s)) ds \right. \\ \left. + \int_{-\tau}^0 \int_{t+\eta}^t \dot{x}^T(s) A_d^T R A_d \dot{x}(s) ds d\eta \right\}. \quad (9)$$

For the first term on the right-hand side of (9), by taking the derivative and employing (8), one gets

$$\frac{d}{dt} V_1(x(t)) = \left\langle \frac{\partial V_1}{\partial x}, \dot{x}(t) \right\rangle \\ = \dot{x}(t)^T P_i x(t) + x^T(t) P_i \dot{x}(t) \\ = 2x^T(t) P_i (A_i + A_d) x(t) - 2x^T(t) P_i A_d \int_{t-\tau}^t \dot{x}(s) ds.$$

On the other hand, due to Lemma 2.2, we can estimate the following term

$$- 2x^T(t) P_i A_d \int_{t-\tau}^t \dot{x}(s) ds \\ = \int_{t-\tau}^t -2x^T(t) P_i A_d \dot{x}(s) ds$$

$$\leq \tau x^T(t) P_i R^{-1} P_i x(t) + \int_{t-\tau}^t \dot{x}^T(s) A_d^T R A_d \dot{x}(s) ds.$$

Thus,

$$\frac{d}{dt} V_1(x(t)) \leq x^T(t) \{ (A_i + A_d)^T P_i + P_i (A_i + A_d) \\ + \tau P_i R^{-1} P_i \} x(t) + \int_{t-\tau}^t \dot{x}^T(s) A_d^T R A_d \dot{x}(s) ds. \quad (10)$$

For the second term on the right-hand side of (9), we have

$$\frac{d}{dt} \int_{t-\tau}^t F(x(s)) ds = F(x(t)) - F(x(t-\tau)) \\ = x^T(t) Q_i x(t) - x^T(t-\tau) Q_j x(t-\tau) \quad (11)$$

for some  $i, j \in \{1, 2\}$ . For the third term on the right-hand side of (9), we note that

$$\frac{d}{dt} \int_{-\tau}^0 \int_{t+\eta}^t \dot{x}^T(s) A_d^T R A_d \dot{x}(s) ds d\eta \\ = \int_{-\tau}^0 \left\{ \frac{d}{dt} \int_{t+\eta}^t \dot{x}^T(s) A_d^T R A_d \dot{x}(s) ds \right\} d\eta \\ = \int_{-\tau}^0 \dot{x}^T(t) A_d^T R A_d \dot{x}(t) d\eta \\ - \int_{-\tau}^0 \dot{x}^T(t+\eta) A_d^T R A_d \dot{x}(t+\eta) d\eta \\ = \tau \dot{x}^T(t) A_d^T R A_d \dot{x}(t) - \int_{t-\tau}^t \dot{x}^T(s) A_d^T R A_d \dot{x}(s) ds$$

and

$$\dot{x}^T(t) A_d^T R A_d \dot{x}(t) \\ = (A_i x(t) + A_d x(t-\tau))^T A_d^T R \\ \times A_d (A_i x(t) + A_d x(t-\tau)) \\ = x^T(t) A_i^T A_d^T R A_d A_i x(t) \\ + x^T(t-\tau) A_d^{2T} R A_d^2 x(t-\tau) \\ + x^T(t) A_i^T A_d^T R A_d A_d x(t-\tau) \\ + x^T(t-\tau) A_d^T A_d^T R A_d A_i x(t) \\ \leq 2x^T(t) A_i^T A_d^T R A_d A_i x(t) + 2x^T(t-\tau) \\ \times A_d^{2T} R A_d^2 x(t-\tau)$$

where the last inequality is due to Lemma 2.2. Therefore, one obtains

$$\frac{d}{dt} \int_{-\tau}^0 \int_{t+\eta}^t \dot{x}^T(s) A_d^T R A_d \dot{x}(s) ds d\eta \\ \leq 2\tau x^T(t) A_i^T A_d^T R A_d A_i x(t) + 2\tau x^T(t-\tau) A_d^{2T} \\ \times R A_d^2 x(t-\tau) - \int_{t-\tau}^t \dot{x}^T(s) A_d^T R A_d \dot{x}(s) ds. \quad (12)$$



A combination of (10), (11) and (12) yields

$$\frac{d}{dt}V(x_t) \leq \begin{bmatrix} x(t) \\ x(t-\tau) \end{bmatrix}^T \begin{bmatrix} \Pi_{11}^i & 0 \\ 0 & \Pi_{22}^j \end{bmatrix} \begin{bmatrix} x(t) \\ x(t-\tau) \end{bmatrix}$$

where  $\Pi_{22}^j = -Q_j + 2\tau A_d^{2T} R A_d^2$  and  $\Pi_{11}^i = \text{He}(P_i(A_i + A_d)) + \tau P_i R^{-1} P_i + Q_i + 2\tau A_i^T A_d^T R A_d A_i$ . By Schur complement (Lemma 2.4) and the assumptions (7b)-(7d), we further obtain

$$\frac{d}{dt}V(x_t) \leq -\omega(x(t)), \forall t \geq \tau$$

where

$$\omega(x) = x^T \max\{-\lambda(\Pi_{11}^1), -\lambda(\Pi_{11}^2)\}x.$$

This fact together with Proposition 3.2 yields the asymptotic stability of system (4).  $\square$

### 4.3. Linear state feedback stabilization

An interesting application of stability conditions derived in Theorem 4.1 is that they can be employed to design a linear state feedback controller

$$u(t) = Kx(t) \quad (13)$$

that makes the following closed-loop system is asymptotically stable

$$\dot{x}(t) = \begin{cases} (A_1 + BK)x(t) + A_d x(t-\tau), & \text{if } c^T x(t) < 0 \\ (A_2 + BK)x(t) + A_d x(t-\tau), & \text{if } c^T x(t) \geq 0 \end{cases} \quad (14a)$$

$$x(s) = \varphi(s), s \in [-\tau, 0]. \quad (14b)$$

**Theorem 4.2.** Consider the system (4). Suppose that there exist symmetric positive definite matrices  $\tilde{P}, \tilde{Q}, \tilde{R} \in \mathbb{R}^{n \times n}$ , a matrix  $U \in \mathbb{R}^{m \times n}$  and scalars  $\gamma > 0, \mu > 0$  such that

$$\begin{bmatrix} -\tilde{\gamma}\tilde{R} & \tilde{R}^2 \\ \tilde{R}^2 & -\tilde{\gamma}\tilde{R} \end{bmatrix} < 0 \quad (15)$$

and the following statements hold

$$\begin{bmatrix} -\tilde{Q} & \tilde{P}c \\ c^T \tilde{P} & -\mu \end{bmatrix} < 0, \quad (16a)$$

$$\begin{bmatrix} -\tilde{Q} & \sqrt{2\tau}\tilde{P}A_d^{2T} & \tilde{P}c \\ \sqrt{2\tau}A_d^{2T}\tilde{P} & -\tau\tilde{R} & 0 \\ c^T \tilde{P} & 0 & -\mu \end{bmatrix} < 0, \quad (16b)$$

$$\begin{bmatrix} \Psi_1 & \sqrt{2\tau}\Gamma_1^T A_d^T \\ \sqrt{2\tau}A_d\Gamma_1 & -\tau\tilde{R} \end{bmatrix} < 0, \quad (16c)$$

$$\begin{bmatrix} \Psi_2 & \sqrt{2\tau}\Gamma_2^T A_d^T & \tau\tilde{P}cc^T & \Delta \\ \sqrt{2\tau}A_d\Gamma_2 & -\tau\tilde{R} & 0 & 0 \\ cc^T \tilde{P} & 0 & -\tau\tilde{R} & 0 \\ \Delta^T & 0 & 0 & \Sigma \end{bmatrix} < 0, \quad (16d)$$

where  $\Psi_i = \text{He}\{(A_i + A_d)\tilde{P} + BU\} + \tilde{Q} + \tau\tilde{R}$ ,  $\Gamma_i = A_i\tilde{P} + BU$ ,  $\Sigma = \text{diag}(-\tilde{\gamma}I, -\tilde{\gamma}I, -\tau\tilde{\gamma}, -\tau\tilde{\gamma})$ , and

$$\Delta = [\tilde{P}(A_2 + A_d)^T + U^T B^T \quad \tilde{P}cc^T \quad \tau\tilde{R}c \quad \tau\tilde{P}c].$$

Then, there exists a linear state feedback controller  $u(t) = Kx(t)$  such that the closed-loop system (14) is asymptotically stable.

*Proof.* Define  $P := \tilde{P}^{-1} > 0$ ,  $Q := \tilde{P}^{-1}\tilde{Q}\tilde{P}^{-1} > 0$ ,  $R := \tilde{R}^{-1} > 0$ ,  $K := U\tilde{P}^{-1}$  and  $h := -c/(2\mu)$ . We prove that the matrices  $P, Q, R, h$  together with the scalar  $\tilde{\gamma}^{-1}$  fulfill the conditions of Theorem 4.1, in the framework of closed-loop system (14).

1) First, it is obvious that  $P + \tilde{\gamma}^{-1}cc^T > 0$  since  $P > 0$  and  $\tilde{\gamma} > 0$ . Next, one has

$$\begin{aligned} Q + hc^T + ch^T &= Q - \mu^{-1}cc^T \\ &= \tilde{P}^{-1}\tilde{Q}\tilde{P}^{-1} - \mu^{-1}cc^T \\ &= \tilde{P}^{-1}(\tilde{Q} - \mu^{-1}\tilde{P}cc^T\tilde{P})\tilde{P}^{-1} > 0 \end{aligned}$$

due to (16a) and Schur complement.

2) To verify the claim (7b), note that it follows from (16b)

$$\begin{aligned} &-\tilde{Q} - [\sqrt{2\tau}\tilde{P}A_d^{2T} \quad \tilde{P}c] \\ &\quad \times \begin{bmatrix} -\tau\tilde{R} & 0 \\ 0 & -\mu \end{bmatrix}^{-1} [\sqrt{2\tau}\tilde{P}A_d^{2T} \quad \tilde{P}c]^T < 0 \end{aligned}$$

or equivalently

$$-\tilde{Q} + 2\tau\tilde{P}A_d^{2T}\tilde{R}^{-1}A_d^2\tilde{P} + \mu^{-1}\tilde{P}cc^T\tilde{P} < 0.$$

By pre-multiplying and post-multiplying by  $\tilde{P}^{-1}$  in the above inequality, one gets

$$\begin{aligned} &2\tau\tilde{P}A_d^{2T}\tilde{R}^{-1}A_d^2\tilde{P} - Q - hc^T - ch^T \\ &= 2\tau\tilde{P}A_d^{2T}\tilde{R}^{-1}A_d^2\tilde{P} - Q + \mu^{-1}cc^T < 0. \end{aligned}$$

Note that  $\mu^{-1}cc^T > 0$ , the above inequality also implies that  $2\tau\tilde{P}A_d^{2T}\tilde{R}^{-1}A_d^2\tilde{P} - Q < 0$ . The claim (7b) is verified.

3) Next, we verify the claim (7c). Due to (16c), we have

$$\begin{aligned} &\text{He}\{(A_1 + A_d)\tilde{P} + BU\} + \tilde{Q} + \tau\tilde{R} \\ &+ 2\tau(A_1\tilde{P} + BU)^T A_d^T \tilde{R}^{-1} A_d (A_1\tilde{P} + BU) < 0 \end{aligned}$$



Substituting  $U = K\tilde{P}$  and then pre-multiplying and post-multiplying by  $\tilde{P}^{-1}$  in the obtained inequality, one gets

$$P(A_1 + BK + A_d) + (A_1 + BK + A_d)^T P + Q + 2\tau(A_1 + BK)^T A_d^T R A_d (A_1 + BK) + \tau P R^{-1} P < 0.$$

This is equivalent to (7c) in the context of closed-loop system. The claim (7c) is verified.

4) Finally, we verify the claim (7d). Note that the (16d) implies

$$\Psi_2 - \begin{bmatrix} \sqrt{2}\tau\Gamma_2^T A_d^T & \tau\tilde{P}cc^T & \Delta \end{bmatrix} \begin{bmatrix} -\tau\tilde{R} & 0 & 0 \\ 0 & -\tau\tilde{R} & 0 \\ 0 & 0 & \Sigma \end{bmatrix}^{-1} \times \begin{bmatrix} \sqrt{2}\tau\Gamma_2^T A_d^T & \tau\tilde{P}cc^T & \Delta \end{bmatrix}^T < 0$$

or equivalently

$$\begin{aligned} & \text{He}\{(A_1 + A_d)\tilde{P} + BU\} + \tilde{Q} + \tau\tilde{R} \\ & + 2\tau\Gamma_2^T A_d^T \tilde{R}^{-1} A_d \Gamma_2 + \tau\tilde{P}cc^T \tilde{R}^{-1} \tilde{P}cc^T \\ & + (\tilde{P}(A_2 + A_d)^T + U^T B^T)\tilde{\gamma}^{-1}((A_2 + A_d)\tilde{P} + BU) \\ & + \tilde{P}cc^T \tilde{\gamma}^{-1}cc^T \tilde{P} + \tau\tilde{R}c\tilde{\gamma}^{-1}c\tilde{R} + \tau\tilde{P}c\tilde{\gamma}^{-1}c\tilde{P} < 0. \end{aligned} \quad (17)$$

Note that in the context of closed-loop system, the LMI

$$\begin{bmatrix} \Phi_2 & \tau(P + \tilde{\gamma}^{-1}cc^T) \\ \tau(P + \tilde{\gamma}^{-1}cc^T) & -\tau R \end{bmatrix} < 0, \quad (18)$$

is equivalent to

$$\begin{aligned} & \text{He}\{(P + \tilde{\gamma}^{-1}cc^T)(A_2 + BK + A_d)\} + Q - \mu^{-1}cc^T \\ & + 2\tau(A_2 + BK)^T A_d^T R A_d (A_2 + BK) \\ & + \tau(P + \tilde{\gamma}^{-1}cc^T)R^{-1}(P + \tilde{\gamma}^{-1}cc^T) < 0. \end{aligned}$$

By pre- and post-multiplying by  $P^{-1}$ , the above inequality is equivalent to

$$\begin{aligned} & (A_2 + BK + A_d)P^{-1} + P^{-1}(A_2 + BK + A_d)^T \\ & + \tilde{\gamma}^{-1}P^{-1}cc^T(A_2 + BK + A_d)P^{-1} \\ & + \tilde{\gamma}^{-1}P^{-1}(A_2 + BK + A_d)^T cc^T P^{-1} + P^{-1}QP^{-1} \\ & + 2\tau P^{-1}(A_2^T + K^T B^T)A_d^T R A_d (A_2 + BK)P^{-1} \\ & - \mu^{-1}P^{-1}cc^T P^{-1} + \tau R^{-1} + \tau\tilde{\gamma}^{-1}P^{-1}cc^T R^{-1} \\ & + \tau\tilde{\gamma}^{-1}R^{-1}cc^T P^{-1} \\ & + \tau P^{-1}cc^T \tilde{\gamma}^{-2}R^{-1}cc^T P^{-1} < 0. \end{aligned} \quad (19)$$

On the other hand, applying Lemma 2.3, we have

$$P^{-1}cc^T(A_2 + BK + A_d)P^{-1}$$

$$\begin{aligned} & + P^{-1}(A_2 + BK + A_d)^T cc^T P^{-1} \\ & \leq P^{-1}cc^T cc^T P^{-1} \\ & + P^{-1}(A_2 + BK + A_d)^T (A_2 + BK + A_d)P^{-1}, \end{aligned}$$

and

$$\begin{aligned} & P^{-1}cc^T R^{-1} + R^{-1}cc^T P^{-1} \\ & \leq P^{-1}cc^T P^{-1} + R^{-1}cc^T R^{-1}. \end{aligned}$$

Moreover, the LMI (15) yields  $\tilde{\gamma}^{-2}R^{-1} \leq R$ . Therefore, the inequality (19) holds if the following one fulfills

$$\begin{aligned} & (A_2 + BK + A_d)P^{-1} + P^{-1}(A_2 + BK + A_d)^T \\ & + \tilde{\gamma}^{-1}P^{-1}cc^T cc^T P^{-1} + P^{-1}QP^{-1} - \mu^{-1}P^{-1}cc^T P^{-1} \\ & + \tilde{\gamma}^{-1}P^{-1}(A_2 + BK + A_d)^T (A_2 + BK + A_d)P^{-1} \\ & + 2\tau P^{-1}(A_2^T + K^T B^T)A_d^T R A_d (A_2 + BK)P^{-1} \\ & + \tau R^{-1} + \tau\tilde{\gamma}^{-1}P^{-1}cc^T P^{-1} + \tau\tilde{\gamma}^{-1}R^{-1}cc^T R^{-1} \\ & + \tau P^{-1}cc^T Rcc^T P^{-1} < 0. \end{aligned}$$

Note that the later inequality is followed from (17) where  $P = \tilde{P}^{-1} > 0$ ,  $R = \tilde{R}^{-1} > 0$ ,  $Q = \tilde{P}^{-1}\tilde{Q}\tilde{P}^{-1} > 0$ . The proof of claim (7d) is done.  $\square$

Without time-delays, we get the corollary.

**Corollary 4.3.** Consider system (4) with  $\tau = 0$ . Suppose that there exist a positive definite symmetric matrix  $\tilde{P} \in \mathbb{R}^{n \times n}$ , a matrix  $U \in \mathbb{R}^{m \times n}$  and scalar  $\tilde{\gamma} > 0$  such that

$$\text{He}((A_1 + A_d)\tilde{P} + BU) < 0,$$

and

$$\begin{bmatrix} \Psi & \Delta & \tilde{P}cc^T \\ \Delta^T & -\tilde{\gamma}I & 0 \\ cc^T \tilde{P} & 0 & -\tilde{\gamma}I \end{bmatrix} < 0,$$

where  $\Psi = \text{He}((A_2 + A_d)\tilde{P} + BU)$ , and  $\Delta = \tilde{P}(A_2 + A_d)^T + U^T B^T$ . Then, there exists a linear state feedback controller  $u(t) = Kx(t)$  such that the closed-loop system (14) is asymptotically stable.

**Remark 4.4.** Our developed results can be applied to discontinuous time-delayed bimodal piecewise linear systems with inputs. In fact, for such systems, we may employ a state feedback controller as

$$u(t) = Kx(t) + \begin{cases} K_1 x(t) & \text{nếu } c^T x(t) \leq 0 \\ K_2 x(t) & \text{nếu } c^T x(t) \geq 0, \end{cases}$$

where the gains  $K_1, K_2$  are first designed in such a way that the closed-loop system is continuous, i.e., satisfying

$$(A_1 - A_2) + B(K_1 - K_2) = hc^T$$

for some  $h \in \mathbb{R}^n$  and the gain  $K$  is designed to stabilize the system

$$\dot{x}(t) = \begin{cases} (\tilde{A}_1 + BK)x(t) + A_d x(t - \tau) & \text{if } c^T x(t) \leq 0 \\ (\tilde{A}_2 + BK)x(t) + A_d x(t - \tau) & \text{if } c^T x(t) \geq 0 \end{cases} \quad (20)$$

with  $\tilde{A}_i := A_i + BK_i$ ,  $i = 1, 2$ .

## 5. NUMERICAL EXAMPLES

In this section, we present two numerical examples to illustrate the effectiveness of the proposed stabilization conditions for both cases: with and without time delays. The first example is considered as a bimodal piecewise linear system with time delays. In the second example, we collect a bimodal piecewise linear system without time delays.

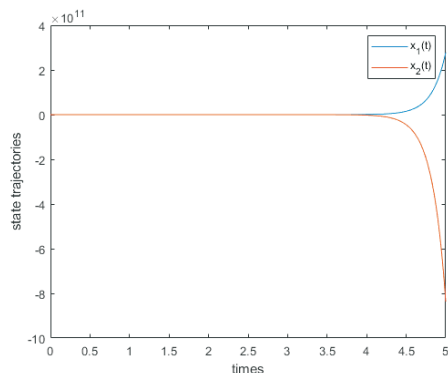
**Example 5.1.** Consider the planar time-delayed bimodal piecewise linear system

$$\dot{x}(t) = \begin{cases} A_1 x(t) + A_d x(t - \tau) + bu(t), & c^T x(t) \leq 0 \\ A_2 x(t) + A_d x(t - \tau) + bu(t), & c^T x(t) \geq 0 \end{cases} \quad (21)$$

where  $c^T = [-1 \ 2]$ ,  $b^T = [0.3 \ 0]$  and

$$A_1 = \begin{bmatrix} -2 & -4 \\ 2 & 3 \end{bmatrix}, A_2 = \begin{bmatrix} -5 & -4 \\ -1 & -2 \end{bmatrix}, A_d = \begin{bmatrix} 1 & 3 \\ -5 & 2 \end{bmatrix}.$$

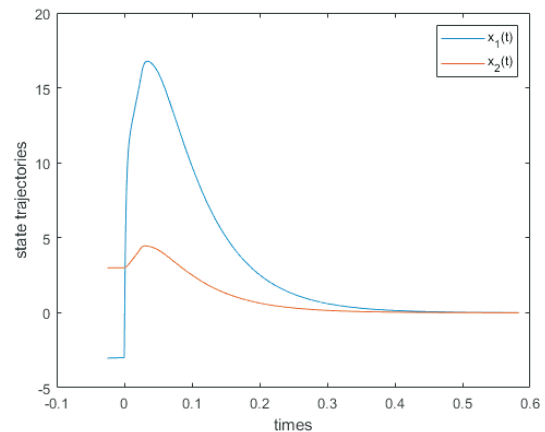
For  $u(t) \equiv 0$  and  $\tau = 0$ , the system (21) is not asymptotically stable as shown in Fig. 1,



**Figure 1.** Trajectories of system (21) for  $\tau = 0$ ,  $u(t) = 0$  and starting at  $x^0 = (-3, 3)^T$ .

For  $\tau = 0.025$ , we would stabilize the system by using linear state feedback controller. To do so, we find

the matrices  $\tilde{P}$ ,  $\tilde{Q}$ ,  $\tilde{R}$ , the numbers  $\tilde{\gamma}$ ,  $\mu$  satisfying (15) and LMIs (16a), (16b), (16c), (16d) of Theorem 4.2. Note that condition (15) is not an LMI. However, we can take  $\tilde{R} = 5I_2$ . Then, it is an LMI in  $\tilde{\gamma}$ . Solving the LMIs, we get  $K = U\tilde{P}^{-1} = \begin{bmatrix} -1513.4 & 5681.6 \end{bmatrix}$ .



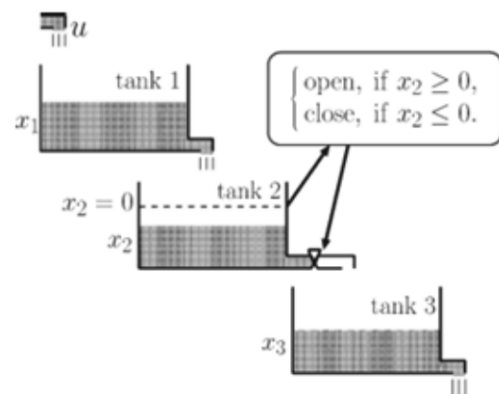
**Figure 2.** Trajectories of the closed-loop system between system (21) and controller (13).

Figure 2 shows the trajectories of the closed-loop system that is composed from the system (21) and the state feedback controller  $u(t) = Kx(t) = \begin{bmatrix} -1513.4 & 5681.6 \end{bmatrix} x(t)$  for initial function

$$\varphi(t) = \begin{bmatrix} -3 + \sin t \\ 2 + \cos t \end{bmatrix}, t \in [-0.025, 0],$$

This trajectory asymptotically converges to the origin.

In the rest of this paper, we validate our method to stabilize a practical bimodal piecewise linear system without time-delays that appeared in the work,<sup>14</sup> Example 21, and compare our achievements with the available methods in<sup>14</sup>.



**Figure 3.** Three water tanks system, source in.<sup>14</sup>

**Example 5.2** <sup>(14)</sup>. Consider a three water tanks system as illustrated in Fig. 3 that is taken from the paper. <sup>14</sup> Let  $x_i$  be the water level of tank  $i$ , ( $i = 1, 2, 3$ ), and  $u$  be the volume of water discharged into tank 1. The valve at tank 2 is open if  $x_2 \geq 0$  and closed if  $x_2 < 0$ . For simplicity, all coefficients are normalized to 1. Then, dynamic equation of the system in the neighborhood of origin is

$$\dot{x}(t) = \begin{cases} A_1 x(t) + bu(t) & \text{if } c^T x(t) \leq 0 \\ A_2 x(t) + bu(t) & \text{if } c^T x(t) \geq 0 \end{cases} \quad (22)$$

where  $c^T = \begin{bmatrix} 0 & -1 & 0 \end{bmatrix}$ ,  $b^T = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$  and

$$A_1 = \begin{bmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}, A_2 = \begin{bmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

Without input, i.e.,  $u(t) = 0$ , the origin is not asymptotically stable. <sup>14</sup> Based on the theory developed in, <sup>14</sup> a state feedback controller has been derived to stabilize the system using piecewise linear functions as follows

$$u(t) = \begin{cases} K_1 x(t) & \text{if } c^T x(t) \leq 0 \\ K_2 x(t) & \text{if } c^T x(t) \geq 0 \end{cases}$$

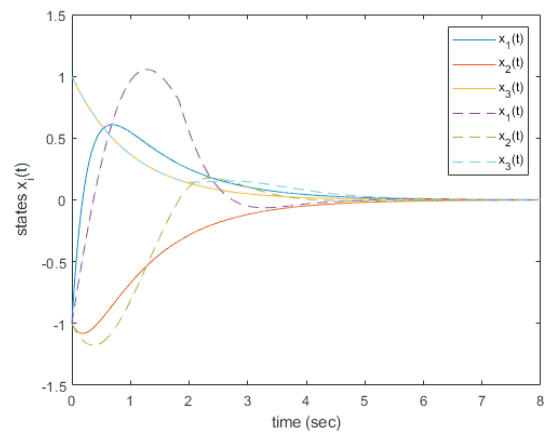
with  $K_1 = \begin{bmatrix} -1 & -2 & 0 \end{bmatrix}$ ,  $K_2 = \begin{bmatrix} 0 & -2 & 0 \end{bmatrix}$ . In fact, such controller transforms a continuous bimodal system into a discontinuous closed-loop one, but it is still well-posed in the sense of Carathéodory solutions. <sup>15</sup> It worth to mention that one does not get such lucky in general.

Taking our proposed approach, we solve the involved LMIs of Corollary 4.3 and get the matrices  $U = \begin{bmatrix} -0.2548 & -0.4668 & -0.0143 \end{bmatrix}$  and

$$\tilde{P} = \begin{bmatrix} 0.2690 & -0.2156 & -0.0517 \\ -0.2156 & 0.3749 & 0.0513 \\ -0.0517 & 0.0513 & 0.6138 \end{bmatrix}.$$

Then, the gain  $K$  of linear state feedback controller is

$$K = U\tilde{P}^{-1} = \begin{bmatrix} -3.6140 & -3.3159 & -0.0504 \end{bmatrix}.$$



**Figure 4.** Trajectories of the closed-loop three tanks system with two kinds of controllers.

Figure 4 shows the trajectories of closed-loop systems for the initial state  $x(0) = \text{col}(-1, -1, 1)$  that is composed from system (22) and state feedback controllers: the linear state feedback controller  $u(t) = Kx(t)$  (solid lines) and piecewise linear state feedback controller (dash lines)

$$u(t) = \begin{cases} K_1 x(t) & \text{if } c^T x(t) \leq 0 \\ K_2 x(t) & \text{if } c^T x(t) \geq 0. \end{cases}$$

Our controller yields a better stabilization.

## 6. CONCLUSIONS

In this paper, we studied the stability and stabilization of time-delayed bimodal piecewise linear systems via smooth Lyapunov-Krasovskii functionals. The main contributions of the paper are including: (1) new stability criteria based on the proposed smooth Lyapunov-Krasovskii functional were derived to guarantee asymptotic stability in the zero inputs and (2) an interesting condition was established to design a linear state feedback controller to stabilize the system which is less conservative than before in the literature. Finally, some numerical examples illustrate the effectiveness of proposed methods.

## Acknowledgment

*This research is conducted within the framework of science and technology projects at institutional level of Quy Nhon University under the project code T2024.831.02.*

## REFERENCES

1. L. Q. Thuan, M. K. Camlibel. Controllability and stabilizability of a class of continuous piecewise affine dynamical systems, *SIAM Journal on Control and Optimization*, **2014**, 52(3), 1914-1934.
2. M. Hovd, S. Olaru. Relaxing PWQ Lyapunov stability criteria for PWA systems, *Automatica*, **2013**, 49(2), 667-670.
3. Z. Sun. Stability of piecewise linear systems revisited, *Annual Reviews in Control*, **2010**, 34(2), 221-231.
4. Y. Eren, J. Shen, M. K. Camlibel. Quadratic stability and stabilization of bimodal piecewise linear systems, *Automatica*, **2014**, 50, 1444-1450.
5. R. Iervolino, F. Vasca, L. Iannelli. Cone-copositive piecewise quadratic Lyapunov functions for conewise linear systems, *IEEE Transactions on Automatic Control*, **2015**, 60(11), 3077-3082.
6. R. Iervolino, D. Tangredi, F. Vasca. Lyapunov stability for piecewise affine systems via cone-copositivity, *Automatica*, **2017**, 81, 22-29.
7. V. Kulkarni, M. Jun, J. Hespanha. *Piecewise quadratic Lyapunov functions for piecewise affine time-delay systems*, in Proceeding of the 2004 American Control Conference, Boston, Massachusetts, 2004.
8. X. Song, S. Zhou, W.X. Zheng. *Delay-dependent stabilization for continuous-time piecewise time-delay systems*, in Proceeding of the 6<sup>th</sup> World Congress on Intelligent Control and Automation, Dalian, China, 2006.
9. J. Shen. Robust non-Zenoness of piecewise affine systems with applications to linear complementarity systems, *SIAM Journal on Optimization*, **2014**, 24(4), 2023-2056.
10. L. Q. Thuan. Non-Zenoness of piecewise affine dynamical systems and affine complementarity systems with inputs, *Control Theory Technology*, **2014**, 12, 35-47.
11. L. Q. Thuan, M. K. Camlibel. Continuous piecewise affine dynamical systems do not exhibit Zeno behavior, *IEEE Transactions on Automatic Control*, **2011**, 56(8), 1932-1936.
12. S. Boyd, L. Vandenberghe. *Convex optimization*, Cambridge University Press, 2004.
13. V. L. Kharitonov. *Time-delay systems: Lyapunov functionals and matrices*, Springer, New York, 2013.
14. Y. Iwatani, S. Hara. Stability tests and stabilization for piecewise linear systems based on poles and zeros of subsystems, *Automatica*, **2006**, 42, 1685-1695.
15. J. Imura, A. V. D. Schaft. Characterization of well-posedness of piecewise-linear systems, *IEEE Transactions on Automatic Control*, **2000**, 45(9), 1600-1619.
16. J. K. Hale. *Theory of functional differential equations*, Springer New York, NY, 2012.



© 2025 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).





# Nghiên cứu thu thập dữ liệu và giám sát chất lượng điện áp sử dụng Ad-hoc LoRa và PLC Siemens thông qua mạng truyền thông Modbus TCP/IP

Trần Xuân Khoa\*, Huỳnh Đức Hoàn, Phan Gia Trí

Khoa Kỹ thuật và Công nghệ, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 09/08/2024; Ngày sửa bài: 12/09/2024;

Ngày nhận đăng: 30/09/2024; Ngày xuất bản: 28/02/2025

## TÓM TẮT

Một trong những vấn đề nghiên cứu quan trọng trong lĩnh vực đo lường thông minh là đảm bảo truyền tải chính xác các số liệu đo lường từ các điểm thu thập dữ liệu đến các thiết bị giám sát. Đảm bảo độ ổn định và tính liên tục của quá trình truyền dẫn luôn là ưu tiên hàng đầu trong các ứng dụng này. Bài báo này nghiên cứu và xây dựng hệ thống giám sát thông số về điện áp của lưới điện của các điểm dân cư, các điểm có hệ thống điện riêng biệt, sử dụng phương thức truyền thông không dây Lora. Mục đích của bài viết này là kiểm tra tính khả thi của hệ thống trong môi trường đô thị khi sử dụng các thiết bị đơn giản, nhỏ gọn và dễ lắp đặt, vận hành và bảo dưỡng đơn giản. Điện áp của điểm cần giám sát được kiểm tra, các giá trị này được gửi từ các Node về trung tâm sử dụng kiến trúc mạng Ad-hoc LoRa network cho khả năng tự thiết lập và duy trì kết nối mạng với nhau mà không phụ thuộc vào cơ sở hạ tầng mạng cố định. Các Node trung gian có thể được thêm vào để đảm bảo tính ổn định và mở rộng cho mạng lưới với chi phí thấp. Dữ liệu sau khi thu thập được hiển thị trên giao diện Wincc Unified của Siemens. Việc sử dụng thiết bị giám sát là PLC cùng giao thức Modbus TCP/IP cho phép tối ưu hóa khả năng tương thích của hệ thống khi triển khai vào thực tế tại các cơ sở vốn luôn có sẵn thiết bị PLC. Hệ thống vận hành trên thực tế với độ chính xác cao, hệ thống hoạt động ổn định, đáp ứng được các yêu cầu về độ tin cậy và thời gian phản hồi của hệ thống giám sát và thu thập dữ liệu. Nghiên cứu này đã chứng minh rằng Ad-hoc LoRa có thể là giải pháp thay thế cho hệ thống truyền thông sử dụng mạng viễn thông hiện nay.

**Từ khóa:** Ad-hoc LoRa network, Modbus TCP/IP, Kalman filter, Wincc Unified, Smart Metering System.

\*Tác giả liên hệ chính.

Email: tranxuankhoa@qnu.edu.vn

# Research on data collection and monitoring voltage quality using Ad-hoc LoRa and Siemens PLC through Modbus TCP/IP communication

Tran Xuan Khoa\*, Huynh Duc Hoan, Phan Gia Tri

*Faculty of Engineering and Technology, Quy Nhon University, Vietnam*

*Received: 09/08/2024; Revised: 12/09/2024;*

*Accepted: 30/09/2024; Published: 28/02/2025*

## ABSTRACT

Ensuring the accurate transmission of measurement data from data collection points to monitoring devices is a crucial research issue in the field of intelligent measurement. Stability and continuity of data transmission processes are of paramount importance in these applications. This paper investigates and develops a voltage monitoring system for the power grid at sites with separate electrical systems, utilizing LoRa wireless communication technology. The objective of this study is to assess the feasibility of the system in urban environments using devices that are characterized by being simple, compact, and easy to install, operate, and maintain. The voltage at each point is measured and then transmitted from the Nodes to the Base using an Ad-hoc LoRa network architecture. This architecture enables self-establishment and maintenance of network connections without relying on fixed network infrastructure. Intermediate Nodes can be added to ensure stability and scalability of the network at a low cost. Collected data is displayed on the WinCC Unified interface of Siemens. Using a PLC as a monitoring device with the Modbus TCP/IP allows optimizing the system's compatibility when deployed in facilities that always have PLC equipment available. The system operates with high accuracy and stability, meeting the requirements of reliability and response time for the monitoring and data collection system. This study has demonstrated that Ad-hoc LoRa can be an alternative solution to the current telecommunication network communication system.

**Từ khóa:** *Ad-hoc LoRa network, Modbus TCP/IP, Kalman filter, Wincc Unified, Smart Metering System.*

## 1. INTRODUCTION

With the rapid development of our nation's economy, residential, commercial, and industrial loads increasingly demand higher standards for the development of the power grid and the quality of electricity supplied.<sup>1,2</sup> Collecting data, monitoring, and analyzing the parameters of the power grid are crucial in the management and operation of the grid. As the demand for electricity and the complexity of the grid continue to grow,

data collection plays a vital role by providing essential and timely information for the analysis, management, and operation of the power grid. The use of advanced monitoring technologies such as smart sensors, automated systems, and artificial intelligence enables operators to gain a comprehensive and detailed view of the grid's performance. Data collected from these devices and systems can be analyzed to detect potential issues early, optimize operational processes, and make strategic decisions regarding the

---

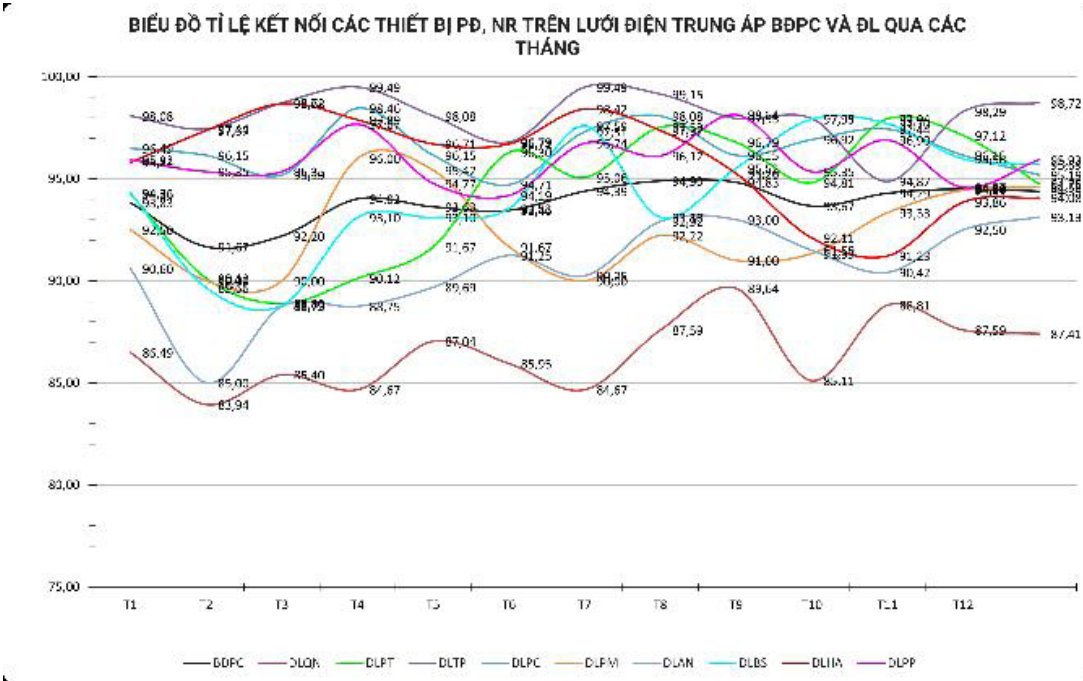
*\*Corresponding author.*

*Email: tranxuankhoa@qnu.edu.vn*

operation, maintenance, upgrading, and expansion of the system.

Nowadays, collecting data from points along transmission lines through hilly and sparsely populated areas using 4G cellular networks provided by telecom operators offers several advantages. This method allows direct data transmission to a central hub, ensuring high-security levels through the use of modems provided by the network operators. However, the requirement to equip 4G modems from the network provider implies initial investment costs and significant operational expenses

(approximately 20-30 million VND per modem and an annual fee of around 1 million VND per Node). In certain cases, power transmission lines traverse complex terrain and encounter limited telecommunications infrastructure. Maintaining stable connectivity under such conditions poses significant challenges. Particularly, the need for continuous 24/24 operation pushes devices into a thermal overload state, resulting in frequent disconnections from the central hub. This not only impedes data collection but also affects the effectiveness of system monitoring and management within the electrical grid.



**Figure 1.** Connection rate of devices on the medium voltage grid of BDPC and DL through the months of 2023.

Here are statistics from Binh Dinh Provincial Power Company regarding the connectivity loss rates of communication devices within the province's power grid in 2023. The chart illustrates the highest connectivity rate at approximately 98%, with most devices maintaining connection levels fluctuating around 95% of total operating time. However, certain segments experience lower connectivity, reaching only 79.4%. Ensuring reliable data collection necessitates maintaining consistent connectivity.

On the other hand, wireless networks such as ZigBee, Wireless Fidelity (Wi-Fi), Narrowband Internet of Things (NB-IoT), SigFox, and Long Range (LoRa) are increasingly developing and have the potential to provide multiple options for wireless communication systems.<sup>3</sup> Among these, low-energy consumption networks with long communication ranges, such as LoRa, help overcome coverage and internet access issues in remote areas where networks like Wi-Fi cannot reach.<sup>4</sup> Wireless technology opens the door for the more efficient and flexible deployment

of monitoring systems, particularly in areas with challenging geographical conditions and limited telecommunications infrastructure. This not only presents opportunities for expanding the range and accuracy of monitoring systems but also helps optimize the cost and reliability of the system.

In this paper, an Ad-hoc LoRa network is employed to develop a network of devices for collecting and transmitting voltage data to a central hub, with a focus on optimizing data transmission and ensuring connection continuity. A complete system has been constructed and deployed in Quy Nhon City to assess the project's feasibility.

## 2. RELATED WORKS

Traditional power grids are continually evolving in the era of smart grids. This evolution necessitates the implementation of precise monitoring strategies to assess system conditions. Eduardo Viciano and colleagues have proposed a multifunctional device model capable of measuring and calculating data related to energy and power quality features in three-phase electrical networks. Designed to perform advanced calculations involving voltage, current, frequency, power, and energy, this device addresses the growing demand for accurate monitoring in our interconnected power systems.<sup>5</sup>

Roach and colleagues deployed smart meters in 129 commercial buildings to investigate the impact of load profiles on building characteristics.<sup>6</sup> Additionally, Czétány and colleagues evaluated the database obtained from using smart meters to determine electricity consumption profiles based on time series data and annual load patterns.<sup>7</sup> Sánchez-Sutil and colleagues developed a device for monitoring and controlling the functions of a power quality analyzer using a LoRa LPWAN (Low Power Wide Area Network).<sup>8</sup>

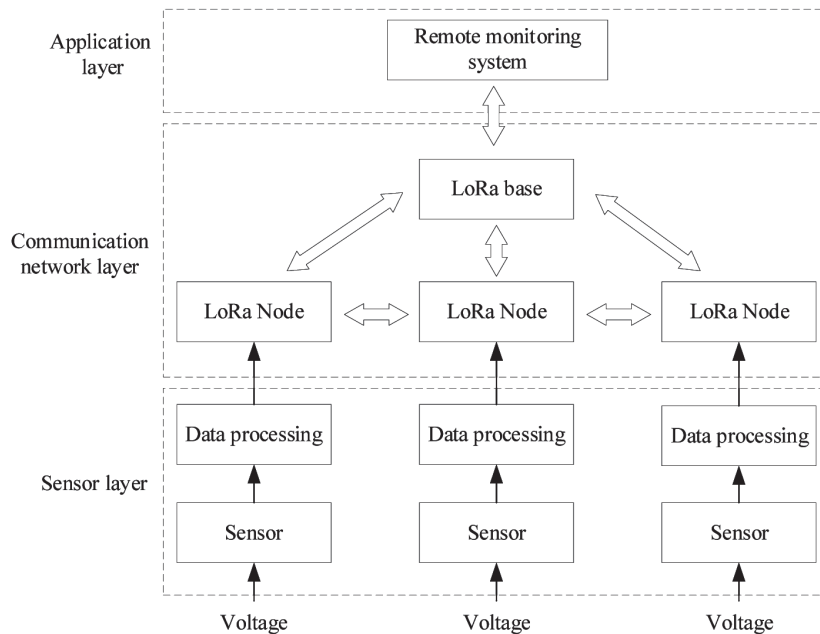
Low Power Wide Area Networks (LPWANs) represent a promising solution

for long-range, low-power communication in Internet of Things (IoT) and Machine-to-Machine (M2M) applications. These networks are resource-constrained and have critical requirements, including long battery life, extended coverage, high scalability, and cost-effective device deployment. LPWANs enable devices to operate for 5-10 years using battery power as their energy source.<sup>9</sup> The advent of these technologies has led to the emergence of new services supporting data collection across diverse organizations in industries, academia, and government. Deployments of such solutions have already begun, utilizing sensor devices to gather various types of data.<sup>10</sup>

Rodrigues-Junior and colleagues utilized smart metering systems to monitor power quality, thereby detecting disturbances in smart grids.<sup>11</sup> Preti Kumari and her colleagues have proposed an intelligent measurement system that saves energy by utilizing Edge computing and Long Range (LoRa) technology.<sup>12</sup>

Due to the Low-Power Wide-Area (LPWA) networks not adhering to the TCP/IP protocol and utilizing their own proprietary protocols, LPWA-equipped sensors face challenges in direct Internet connectivity. Jumpei Sakamoto and colleagues have established a TCP/IP network via LoRa and are exploring ways to enhance TCP communication performance using the Private LoRa interface.<sup>13,14</sup>

This study proposes a solution for data collection and voltage quality monitoring. The system employs voltage sensors connected to an Arduino to read voltage values at the observed nodes. Then, the collected data is transmitted to a central via an Ad-hoc LoRa network. At the central node (Base), the TCP/IP protocol facilitates communication between the Ad-hoc LoRa network and existing industrial devices, such as Programmable Logic Controllers (PLCs). The monitoring interface is built on WinCC Unified, enhancing integration with established systems.



**Figure 2.** Overall structural block diagram of the system.

### 3. Ad-hoc LoRa NETWORK FOR ENERGY GRID

LoRa (Long Range) is a wireless communication technology that enables long-distance data transmission with minimal energy consumption. The Ad-hoc LoRa network for power grids offers a flexible and cost-effective solution for monitoring and controlling electrical systems. This network allows easy deployment without the need for fixed communication infrastructure, reducing costs and deployment time. LoRa devices can connect thousands of network nodes simultaneously, efficiently collecting data from various points within the power system.<sup>15</sup>

By utilizing LoRa networks, grid operators can remotely monitor and manage devices in the power grid. This includes measuring energy consumption, detecting faults, and optimizing grid operations. These capabilities enhance the reliability and efficiency of the electrical system while minimizing downtime and maintenance costs.

#### 3.1. Sensor layer

The data collection layer comprises sensor nodes and measurement devices, serving the primary purpose of gathering and processing

physical data. The collected data may include current intensity, voltage, as well as parameters from smart electricity meters, or variables such as temperature and humidity inside electrical cabinets.

In this study, the voltage sensor was employed to collect voltage data from various locations within the power grid. The three-phase voltage data obtained from these sensors undergo noise filtering and are subsequently forwarded to the next processing layer.

#### 3.2. Communication network layer

The main components of the network include sensors at data collection points, data concentrators, and the data management system. Remote data collection and monitoring of energy consumption are crucial for both end-users and operators (electric utility companies or providers).

In the system, sensor devices (Nodes) transmit collected data to a central hub (referred to as Base) through an Ad-hoc LoRa network. The wireless Ad-hoc LoRa network consists of self-forming nodes that operate independently of existing infrastructure. These nodes manage the network and control tasks using distributed algorithms and multipoint routing, where packets



are relayed by intermediate nodes, enhancing network efficiency.<sup>16</sup>

The key advantages of Ad-hoc LoRa networks include adaptability to specific needs and the ability to form a network with any available nodes. They also offer cost-effective investment, operational efficiency, rapid deployment, simplified reconfiguration due to their distributed architecture, and integrated redundancy features. This flexibility allows for the establishment of a network between nodes without the need for complex infrastructure structures.<sup>17</sup> The data transmission process can be carried out in the following ways:

- Direct transmission from the sending point to the destination: When two points are within the connection range, data will be transmitted directly to the destination.
- Transmission via intermediate nodes: If the destination and sending point are beyond the connection range, intermediate Nodes which located between the sending point and the destination will be used to establish a network. In this configuration, data will be forwarded from point to another until it reaches the requested destination in the packet. This setup allows for expanding the data transmission range and overcoming physical barriers

### 3.3. Application layer

Through LoRa, Nodes transmit data to the Base, including the following information:

- Node Address: Each node device in the system is assigned a unique address for identification and differentiation. The greater the distance from the Node to the Base, the higher the address number (this address is determined during device configuration).
- Data: Nodes collect information as required. in this project, voltage data is collected. After noise filtering, the data is transmitted to the Base for analysis and monitoring of the operational status of the measurement points.
- Received Signal Strength Indicator (RSSI): The LoRa signal strength is a crucial parameter

for assessing the quality of the connection between the node and the Base. It affects the communication range and the system's stability.

- Signal-to-Noise Ratio (SNR): This is the ratio of the received signal power to the noise floor level.
- Timing: The interval between two packets from the Nodes. This parameter is used to evaluate the continuity of the connection between the Nodes.

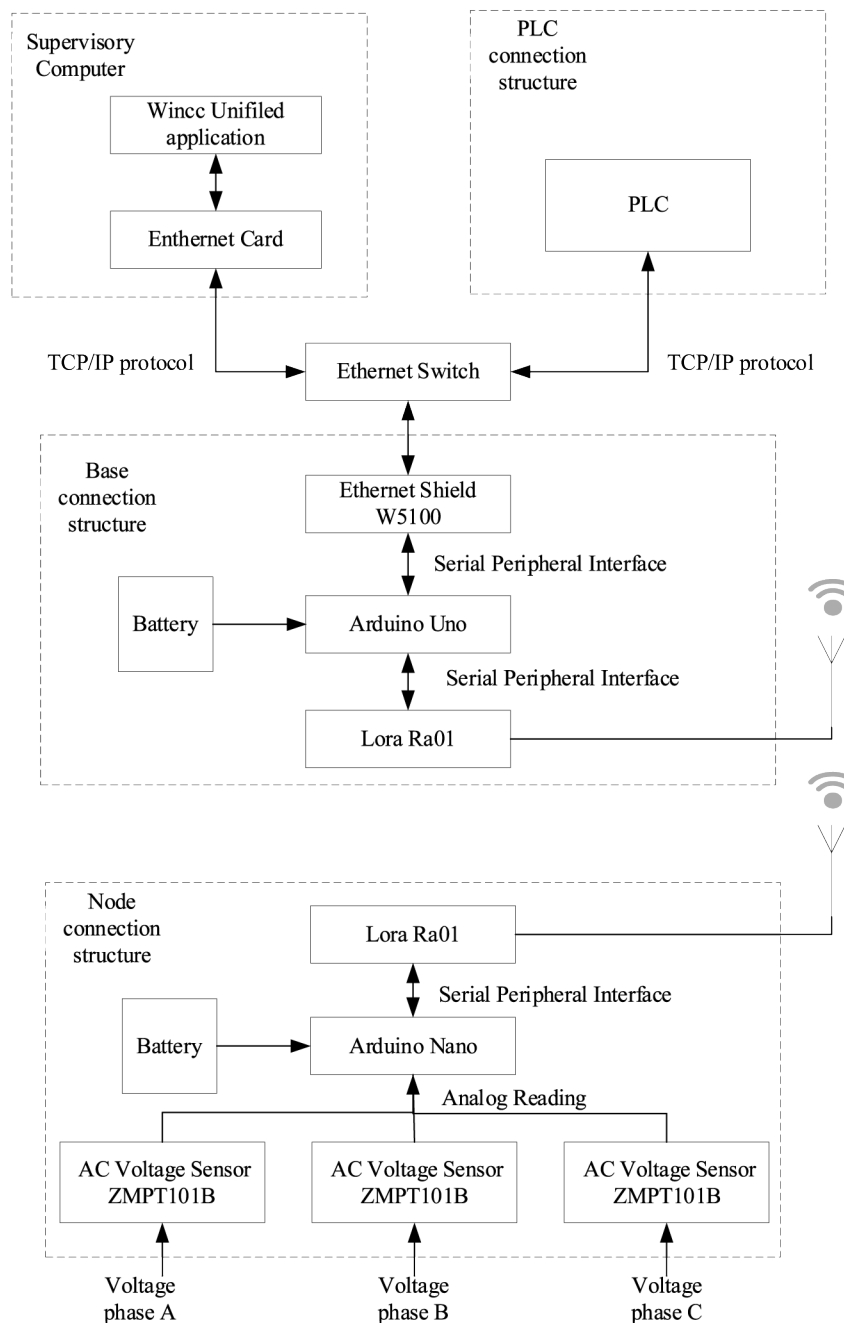
All data transmitted to the Bases will be sent to the SCADA system via the Modbus TCP/IP protocol. The monitoring and operation system of the power sector can easily integrate this data to understand and control the system's performance.

## 4. METHODOLOGY AND DESIGN

In this project, the authors implement an Ad-hoc LoRa network based on LoRa technology to transmit data to the control center, with the primary application being the monitoring of power quality at designated locations. Several parameters, such as voltage stability and phase loss, are collected to preliminarily assess the power quality at the connection points. A practical system has been constructed using voltage sensors, Arduino, LoRa modules, and PLC to evaluate operational capability. In this system, the voltage sensor collects voltage data, transmits it to the Arduino for initial processing, and then sends it to the control center via LoRa signals. At the monitoring center, the WinCC Unified interface on the PLC displays the parameters received from the Arduino through Modbus TCP/IP communication.

- Node connection structure: The node utilizes an Arduino Nano, which offers advantages in terms of size and cost. The entire node is powered by a battery. Three ZMPT101B AC voltage sensors are connected to the Arduino through analog pins. The Lora Ra01 module is interfaced with the Arduino via the SPI communication protocol. The voltage signals, once read, are processed by the Arduino and transmitted via the Lora module.





**Figure 3.** Hardware connection structure.

- Base connection structure: The base connection structure utilizes an Arduino Uno as the microcontroller, with an Ethernet Shield W5100 and a Lora Ra01 module interfaced with the Arduino Uno via the SPI protocol. During operation, the communication alternates between the Lora and Ethernet Shield every 50ms. The Ethernet Shield W5100 enables the Arduino to establish communication with a PLC through the TCP/IP protocol.

- PLC and computer supervisory were connected to a switch where the Base was connected. All connect using the TCP/IP protocol.

- LoRa is used at the Node and Base, which are LoRa Ra01 modules, along with a -12dB antenna and a 1.2m cable. The 433MHz frequency band is suitable for the location in Vietnam. The Bandwidth is 125 kHz, and the power level is 12dBm.

4.1. Data input collection

With the varying requirements of the data acquisition system, different types of sensors can be utilized to collect data. One of the critical functions of the system is the collection and processing of voltage data. The authors employ the AC Voltage Sensor ZMPT101B to measure alternating current (AC) voltage. This sensor is directly connected to an Arduino. After the sensor collects the voltage data, it is converted to a direct current (DC) voltage of 0-5 V and then measured by the Arduino for processing.

To ensure data quality and eliminate unwanted noise, the authors apply the Kalman

filter algorithm. This filter can estimate the true value of the signal from measurements affected by noise. This process is a crucial step to ensure that the collected voltage data is reliable and can be effectively used in management and control.

Comparisons were made using sensors from two nodes simultaneously, where one node transmitted raw, unprocessed data to the control center, and the other node used the Kalman filter to process the signal and remove noise, thereby selecting appropriate coefficients for the filter.

a. Node 200 collected raw data without using the Kalman filter:

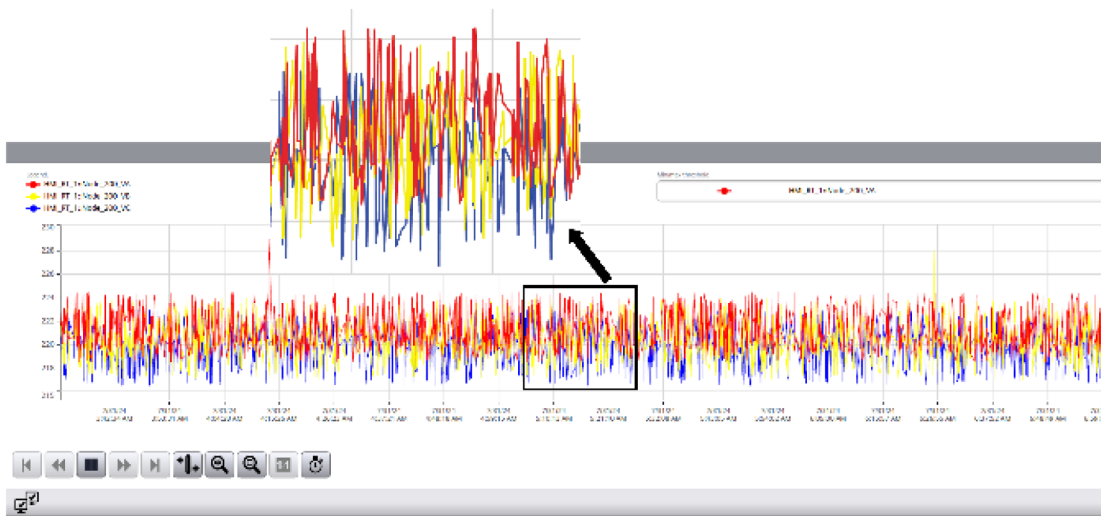


Figure 4. 3 phases voltage data collected from Node 200.

Several criteria were used to evaluate the reliability of the measurement results over 252 minutes:

Table 1. Amplitude of VAC oscillation in Node 200.

|     | Phase A | Phase B | Phase C |
|-----|---------|---------|---------|
| Min | 218.5   | 217     | 216.5   |
| Max | 229     | 228     | 222.99  |

Table 2. % Amplitude of oscillation and standard deviation in Node 200.

|                    | Phase A | Phase B | Phase C |
|--------------------|---------|---------|---------|
| % Min              | 1.27    | 1.56    | 1.56    |
| % Max              | 3.47    | 3.43    | 1.40    |
| Standard Deviation | 1.78    | 1.78    | 1.63    |

The raw data collected from the sensors exhibited significant amplitude fluctuations, ranging from 1.27% to 3.47% compared to the average value. The maximum standard deviation reached 1.78, indicating a considerable dispersion of the measured values from the average. The relative accuracy of the raw data ranged from 0% to 4.5%. These statistical parameters and relative accuracy indicate that the unprocessed raw data is significantly affected by measurement noise and errors, failing to meet the necessary reliability and accuracy requirements for related applications. Therefore, appropriate signal processing and filtering measures must be implemented to improve data quality before use.

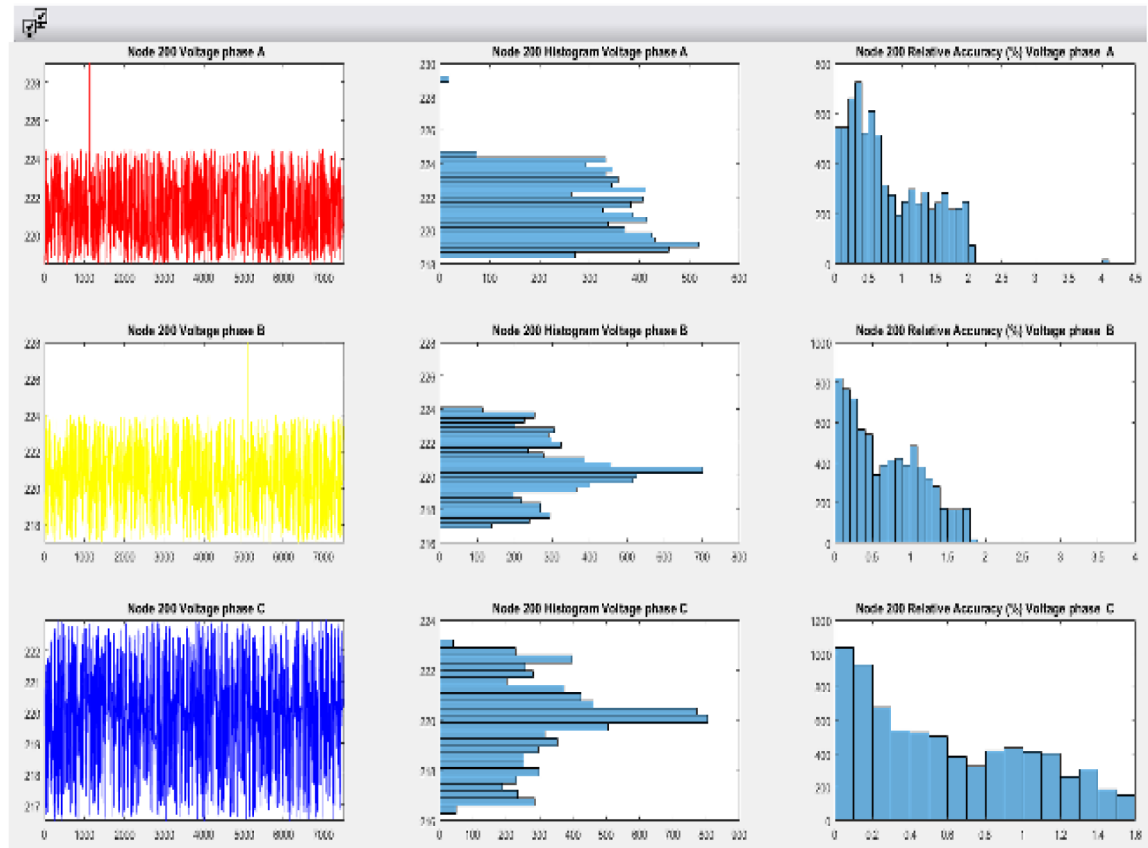


Figure 5. Voltage analysis in Node 200.

**b. Node 201 collects data and processes noise using the Kalman filter**

Several criteria were used to evaluate the reliability of the measurement results over 252 minutes:

**Table 3.** Amplitude of VAC oscillation in Node 201.

|     | Phase A | Phase B | Phase C |
|-----|---------|---------|---------|
| Min | 218.55  | 217.01  | 216.50  |
| Max | 224.49  | 223.95  | 222.98  |

**Table 4.** % Amplitude of oscillation and standard deviation in Node 201.

|                    | Phase A | Phase B | Phase C |
|--------------------|---------|---------|---------|
| % Min              | 0.74    | 1.48    | 1.64    |
| % Max              | 1.96    | 1.67    | 1.30    |
| Standard Deviation | 0.48    | 0.83    | 0.96    |

Based on the provided measurement parameters, an assessment of the accuracy and

reliability of the data after applying the Kalman filter can be performed.

The amplitude fluctuations of the data after applying the Kalman filter ranged only from 0.74% to 1.96% compared to the average value, indicating significant stability and noise reduction. The maximum standard deviation was 0.96, showing a relatively concentrated dispersion of the measured values from the average. The relative accuracy of the filtered data ranged from 0.5% to 1.8%, demonstrating the high suitability and reliability of the measurement results after processing. This ensures that the obtained data can be used with reliable accuracy.

The statistical parameters and relative accuracy indicate that the data after applying the Kalman filter has small amplitude fluctuations, low standard deviation, and high relative accuracy, meeting the necessary reliability and accuracy requirements.

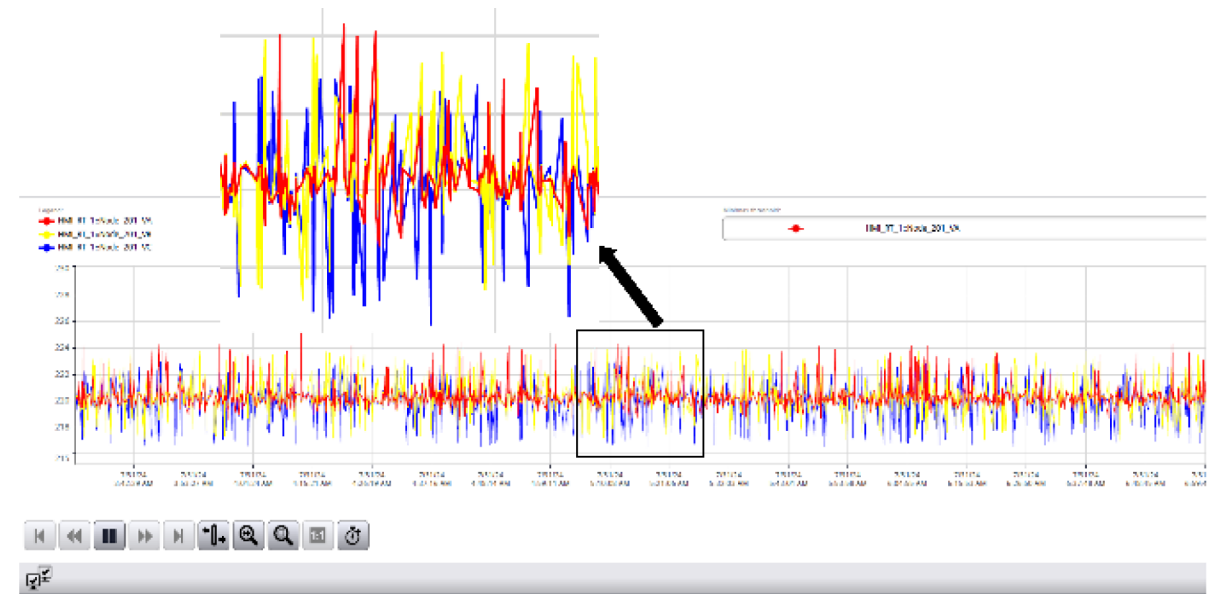


Figure 6. 3 phases voltage data collected from Node 200.

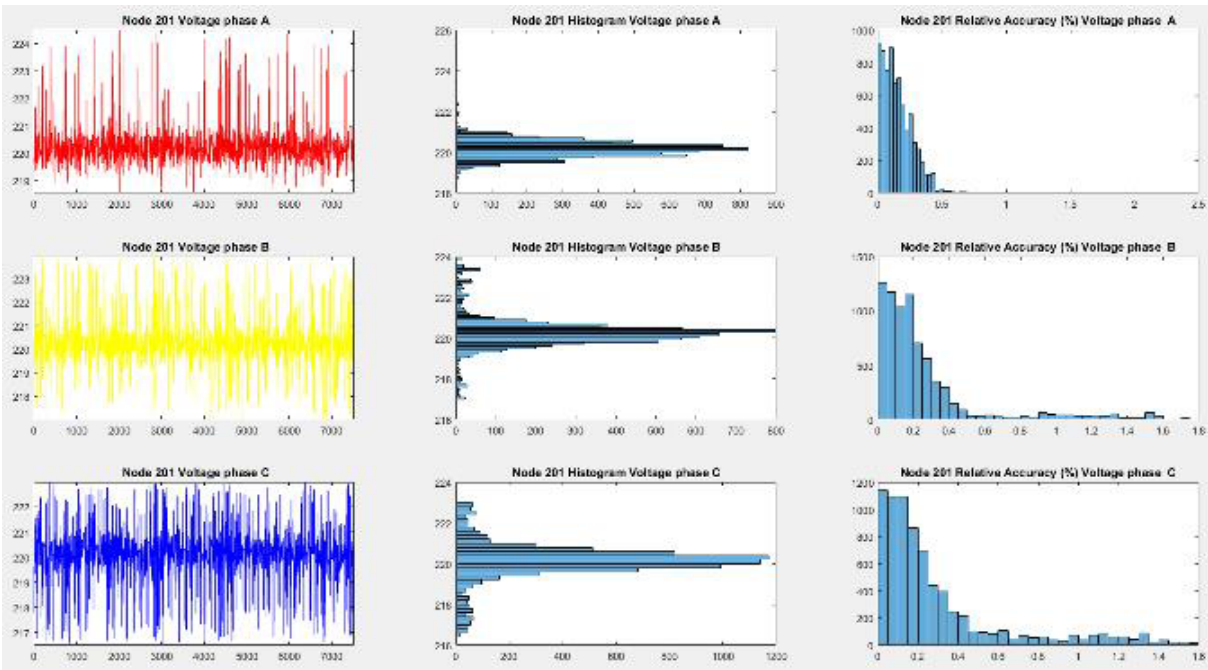


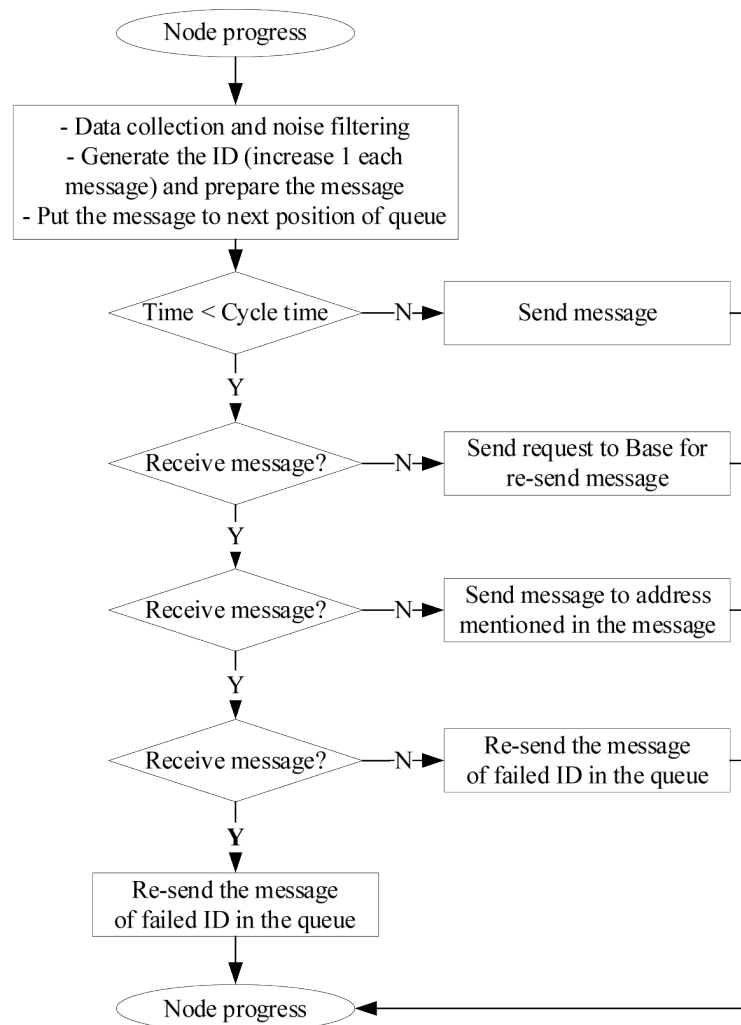
Figure 7. Voltage analysis in Node 201.

4.2. Ad-hoc LoRa network

a. Node

Each node performs functions:

- Collect and process input data.
- Prepare data packets including:
  - + Packet ID
  - + Data to be transmitted
- Transmit signals to the Base. To ensure data integrity, each packet is sent three times.
- After each cycle, if the Node does not receive a response from the Base, it will send a request for the Base to resend the previous packet.
- Receive and process incoming packets:
  - + If the packet is not intended for the node itself, it will forward the packet to the address mentioned in the packet.



**Figure 8.** Node general algorithm flowchart.

+ If the packet is intended for the node, it will check the transmission results:

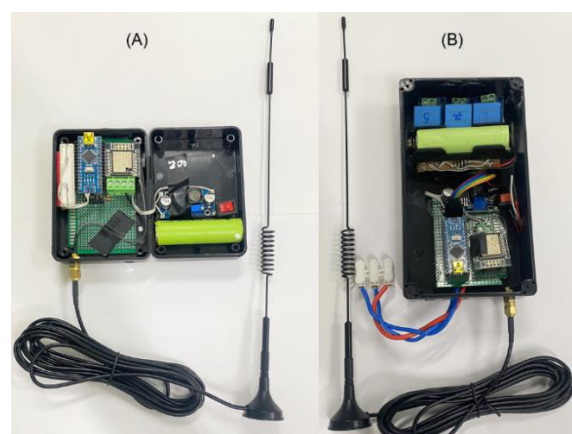
o If the packet was successfully transmitted, it will delete the packet with that ID.

o If the packet requests the retransmission of data for a specific ID, it will resend the corresponding data packet.

- Data collection node: uses Arduino as the MCU to collect and process data from sensors, which is then transmitted via the LoRa Module using a 12db antenna.

- Another configuration that can be deployed to extend the data transmission range is the use of an intermediate Node. This node consists only of an Arduino and LoRa, with the task of forwarding packets to the destination address specified in the packet.

- In this study, the nodes use batteries through a voltage converter to power all their activities. This increases the flexibility of the experiments.



**Figure 9.** Intermediate node (A) Data collection node (B).

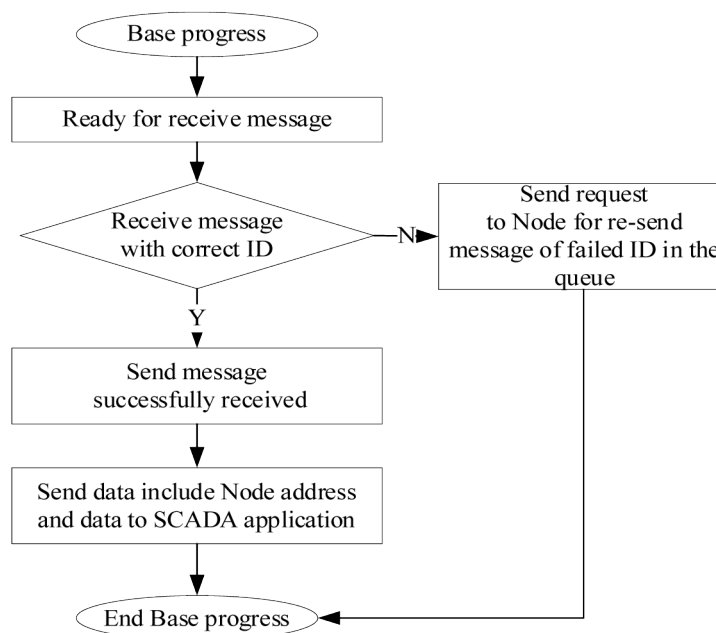
## b. Base

The Base performs functions:

- Receive signals from the Nodes. Check the continuity of the IDs, and if any ID is missing, send a request to the Node to resend that ID.
- Prepare packets for the Nodes, including the following information:
  - + Destination of the packet
  - + Transmission result of the previous packet

+ Resend request with ID (if any)

- Send data to the SCADA software, including the Node address and data.
- The Base uses an Arduino connected to a PLC S7 1200 via the TCP/IP protocol. Data from the Nodes, once received by the Arduino, is transmitted through the PLC to be displayed on the WinCC Unified interface.



**Figure 10.** Base general algorithm flowchart.

## 4.3. Monitoring interface

In monitoring and data acquisition systems, the user interface (UI) serves as a critical component that facilitates user interaction, observation, and manipulation of the displayed data. This interface provides a direct means for users to engage with the system, enabling them to monitor key parameters and important data, as well as perform essential tasks such as adjusting parameters, acknowledging alarms, scheduling, and analyzing historical data.

In this study, a Programmable Logic Controller (PLC) was utilized as the communication device between the LoRa network and the monitoring interface to meet the requirements

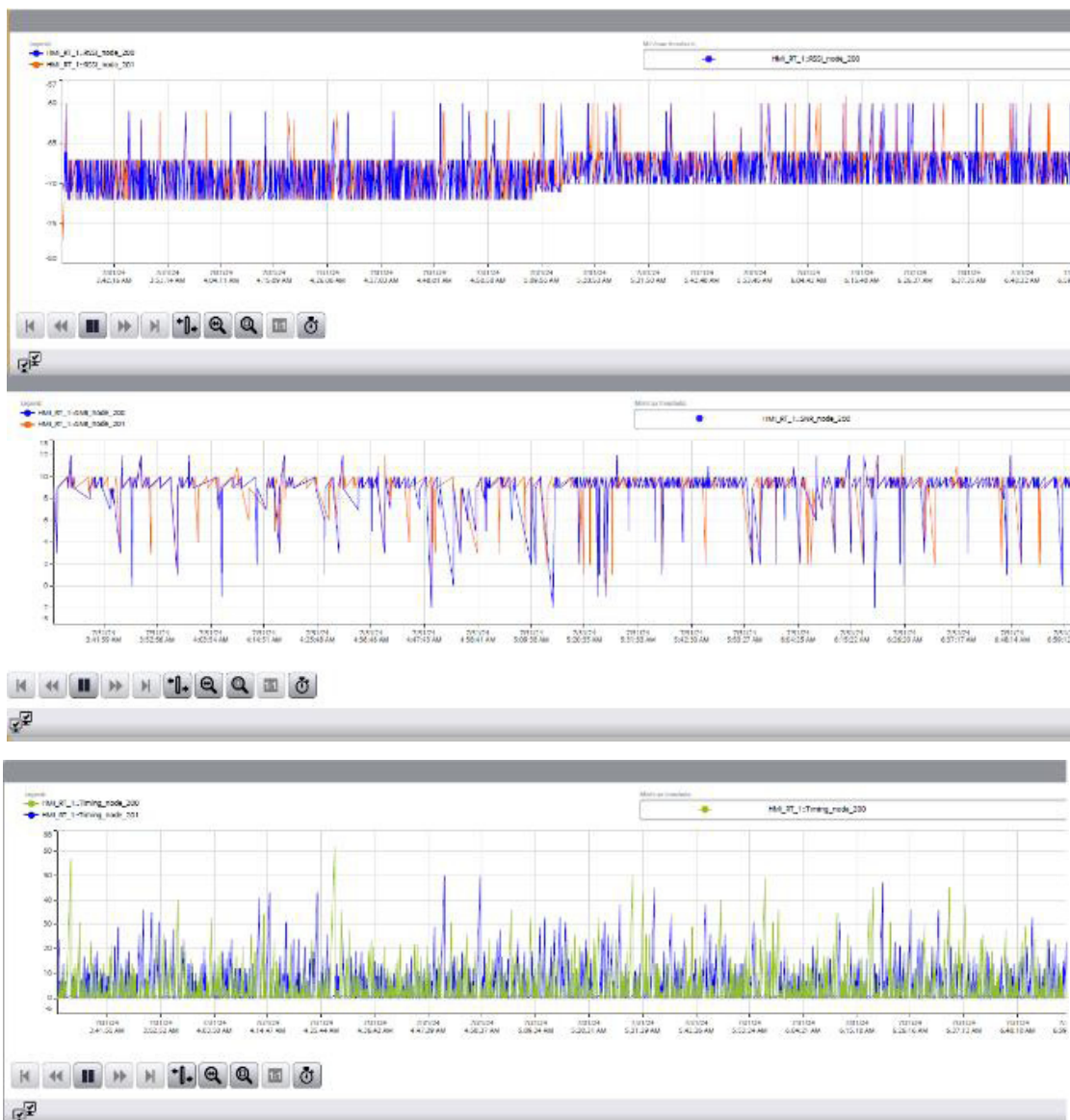
for reliability and compatibility with existing systems. Specifically, a PLC S71200 was employed to connect to the LoRa Base via Modbus TCP/IP communication. The collected data was then transmitted from the Base to the PLC, which subsequently displayed the information on the user interface. To ensure high reliability and compatibility, the WinCC Unified tool was utilized, offering diverse connectivity and seamless integration with PLC systems and other devices. This approach helped optimize the monitoring and control processes of the industrial system.

The user interface will display data including:



- The three-phase voltage at Nodes 200 and 201.
- Received Signal Strength Indicator (RSSI) from the Nodes over time.
- Signal-to-Noise Ratio (SNR), which is the ratio of the received signal power to the noise floor level.

- The interval between two packets from the Nodes to the Base (Timing).
- All data can be exported to an Excel file using the WinCC Unified tool at 2-second intervals throughout the survey period. This data is then imported into Matlab for analysis and graphing.



**Figure 11.** Voltage and signal quality data from Nodes displayed over time on the local web interface using WinCC Unified.

## 5. RESULTS

### A. Case setup

This study aims to evaluate the feasibility and effectiveness of using a LoRa wireless communication network to collect and monitor

voltage data from nodes sent to the control center. Factors such as distance and terrain characteristics are prioritized in the scenarios.

To establish the scenarios, several initial conditions need to be considered when deploying the Ad-hoc LoRa network. Specifically, the

study will examine different cases regarding the placement of the Base and Nodes, as well as the variation in distance between them, including scenarios with obstacles and open spaces.

The scenarios are designed to assess the ability to reliably and efficiently collect and transmit voltage data through the LoRa network. Parameters such as transmission speed, reliability, coverage range, and the ability to cope with terrain factors will be measured and analyzed.



**Figure 12.** System image.

RSSI (Received Signal Strength Indicator) and SNR (Signal-to-Noise Ratio) are two crucial parameters in evaluating signal stability. RSSI is an indicator that measures the signal strength received at the receiver. It provides information about the strength or weakness of the LoRa signal. A higher RSSI generally corresponds to a better connection and higher data transmission capability. SNR is the ratio between the signal strength and the noise. It measures the quality of the LoRa signal. A higher SNR indicates that the LoRa signal is stronger compared to the noise, thus having a better data transmission capability.

In LoRa applications, the continuous monitoring and optimization of RSSI and SNR are imperative to ensure stable connections and optimal data transmission performance. Additionally, the interval between two packets is monitored to assess the continuity of the connection.

The results of this study will provide important information about the capabilities and limitations of using the LoRa network in voltage data monitoring and collection applications. This will enable designers and implementers to make appropriate decisions regarding the application of this technology in real-world systems.

## B. Test results

The authors evaluate real-world scenarios with variations in distance, Node and Base installation locations, and terrain between the Nodes. This assessment aims to evaluate signal stability and the environmental impacts on the experiment.

The energy consumption of each Node is disregarded in these scenarios.

These devices are mounted on poles 1.8 meters above the ground.

*a) Case1: The Base and Node are located in an open area with few obstacles, while Node1 and Node2 are situated along a route flanked by tall buildings on both sides.*

- The Base is located at “PD DH SU PHAM.”
- Node1 is located at “PD EO BIEN,” 476 meters from the Base.
- Node2 is located at “PD CHO KHU 2,” 282 meters from Node1.

The survey was conducted over 252 minutes.

Analysis of the distribution of RSSI (Received Signal Strength Indicator), SNR (Signal-to-Noise Ratio), and Timing values in Case1:



Figure 13. The layout map of the devices in Case 1.

The RSSI values were observed to fluctuate between -60 dBm and -75 dBm, with an

average RSSI of -69.18 dBm for the Base-Node1 connection and -67.67dBm for the Base-Node2 connection. These RSSI levels are considered quite ideal for applications requiring reliable and timely data delivery.<sup>18</sup>

The SNR values also exhibited a similar range of 2 dB to 12 dB across the two connections, with positive SNR indicating that the received signal was operating above the noise floor. The average SNR values of 8.96 dB for Base-Node1 and 8.92 dB for Base-Node2 are within the optimal range for LoRa modulation to perform well.<sup>18</sup>

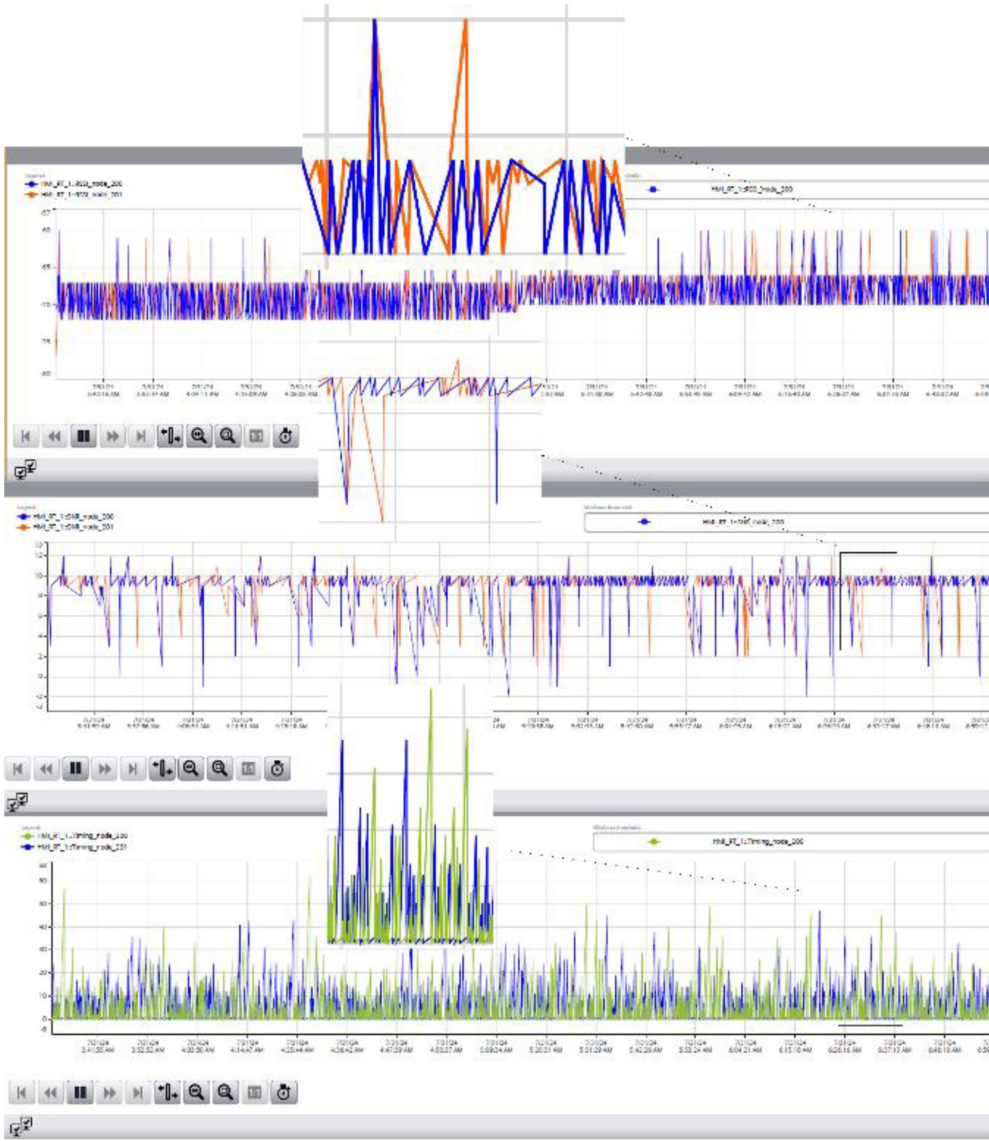


Figure 14. RSSI, SNR, Timing data between packets from Node to Base in Case1.



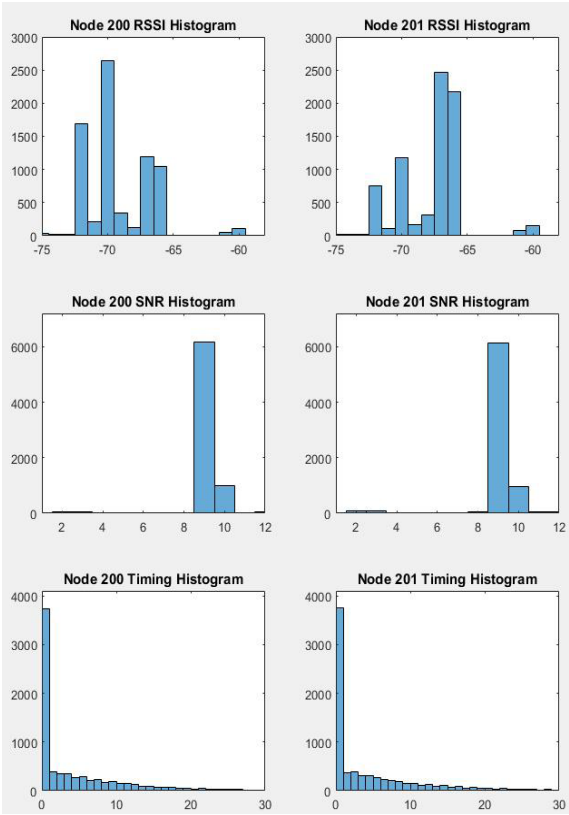


Figure 15. Histogram of RSSI, SNR and Timing values in Case1.

Table 5. Mean prevalent parameter in Case1.

| Mean | Base - Node 1 | Base - Node 2 |
|------|---------------|---------------|
| RSSI | -69,18        | -67,67        |
| SNR  | 8,96          | 8,92          |
| Time | 4,53          | 4,57          |

The maximum time interval observed was 60 seconds for the Base-Node1 link and 50 seconds for the Base-Node2 connection. The average time interval between packets is very similar for both links, at 4.53 seconds for Base-Node1 and 4.57 seconds for Base-Node2. The relatively low average time gaps and the absence of significant deviations suggest that the communication between the nodes and the base station is occurring at a reliable and stable.

The consistent and stable fluctuations in RSSI and SNR, along with the stable data transmission rate, suggest a reliable connection between the Nodes and the Base station. This favorable outcome can be attributed to the relatively unobstructed propagation environment. Overall, the observed RSSI, SNR, and Timing characteristics indicate a robust and reliable communication system, well-suited for the targeted applications.

b) Case 2: Both the Base and Node are located along a road with high-rise buildings on both sides. The Base and Node2 are obstructed by a densely populated urban area with many high-rise buildings.

- The Base is located at "PD EO BIEN"
- Node1 is placed at "PD CHO KHU 2", 282 meters away from the Base
- Node2 is placed at "PD PHAM NGOC THACH", 293 meters away from Node1

The survey was conducted over 300 minutes.

In this case, Node 2 was unable to transmit data directly to Node 1 due to exceeding the transmission range.



Figure 16. The layout map of the devices in Case 2.

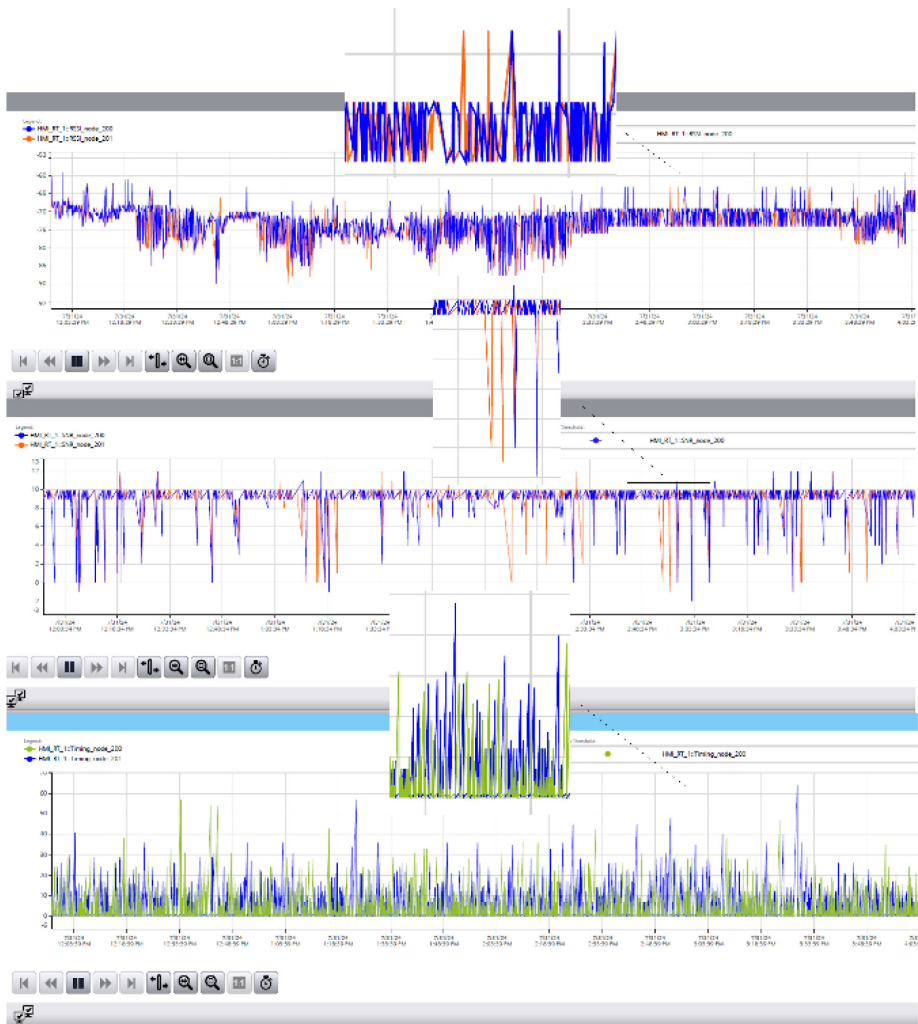


Figure 17. RSSI, SNR, Timing data between packets from Node to Base in Case2.

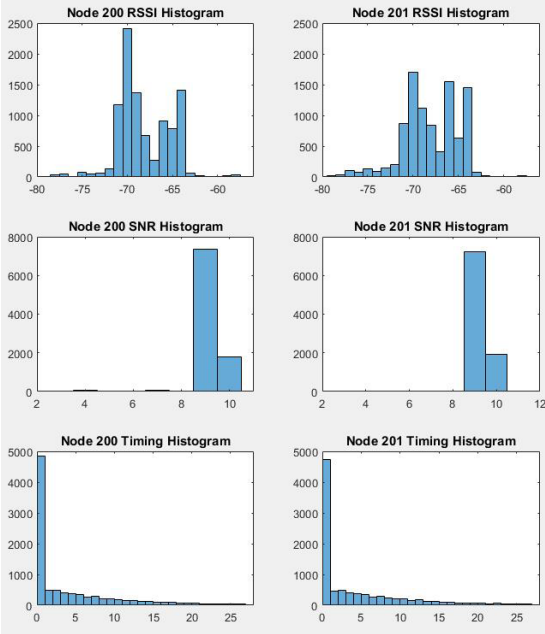


Figure 18. Histogram of RSSI, SNR and Timing values in Case2.

Table 6. Mean relevant parameter in Case2.

| Mean | Base - Node 1 | Base - Node 2 |
|------|---------------|---------------|
| RSSI | -68.16        | -68.05        |
| SNR  | 8.99          | 9.00          |
| Time | 4.35          | 4.68          |

Analysis of the distribution of RSSI (Received Signal Strength Indicator), SNR (Signal-to-Noise Ratio), and Timing values in Case2:

The RSSI parameter of the Nodes fluctuated between -78 dBm and -58 dBm, primarily concentrating in the range of -65 dBm to -70 dBm. The average RSSI for the Node1-Base connection was -68.16 dBm, while for the Node2-Base connection, it was -67.85 dBm. The difference in RSSI values between these two

connections could be attributed to the presence of large trees along the Node2-Base route. Nevertheless, the values indicate stable signal strength, ensuring reliable connectivity.<sup>18</sup>

The SNR parameter for both connections fluctuated similarly, ranging from 4 dB to 12 dB. For the Node2-Base connection, the SNR values were mostly stable around 9-10 dB, unlike the Node1-Base connection, which exhibited more fluctuation. The average SNR was approximately 8.99 dB for the Base-Node1 link and 9 dB for the Base-Node2 link, which is ideal for effective LoRa signal modulation.<sup>18</sup>

Throughout the survey period, the average interval between two packets was 4.35 seconds for the Node1-Base connection and 4.68 seconds for the Node2-Base connection. The stable and high consistency of the RSSI and SNR values indicate the stability of the connection. The data transmission speed remained consistently below 4.7 seconds at most times, demonstrating stable connectivity between the Nodes and the Base station.

In Case2, both the RSSI and Timing indices showed fewer similarities compared to Case1, possibly due to the influence of obstacles such as trees along the route. However, the SNR values remained relatively consistent in both scenarios.

### *c) Summary*

Practical experiments were conducted to analyze the impact of distance, installation location, and terrain conditions on signal strength and data transmission time between the Nodes and the Base. The survey was divided into different scenarios to evaluate the system's performance under conditions with few obstacles and conditions surrounded by tall buildings.

The results indicate that the accuracy of voltage measurement results was significantly improved by the Kalman algorithm. The deployment scenarios for data collection at existing power system locations in Quy Nhon

City demonstrated that the system could operate stably under various terrain and shielding conditions.

The system operated stably in both surveyed scenarios, meeting the reliability and response time requirements of the application. Terrain and shielding factors affected the connection quality, but the system could still function effectively.

In the absence of obstacles (Case1), the connection between the Base and the Nodes was of good quality, with an average RSSI of approximately -64.2 dBm, an SNR of around 8.9 dB, and a response time of less than 4.6 seconds. In the presence of tall buildings and trees (Case2), the connection quality was maintained, with an average RSSI of approximately -68.1 dBm and a response time of around 4.5 seconds.

The practical results show that transmission distance can affect the SNR, while distance, the presence of obstacles, and propagation factors influence the RSSI of the received signal. Along with the stability of the RSSI signal and data transmission time recorded in the practical results, the study indicates that the choice of installation locations for the Node and Base devices is a crucial factor significantly affecting the performance of the wireless communication system.

For applied in practice, when many practical applications use Lora networks up to tens of thousands of nodes with star networks (star nodes) using gateways and data encryption applications according to Lora Alliance standards. The methods to expand the number of system nodes and protect data should be considered. Furthermore, the use of industrial-grade hardware should be regarded as an alternative to the devices proposed in previous studies. Industrial equipment may provide more accurate measurement results, which is crucial for the reliability and performance of LoRa-based applications.



## 6. CONCLUSION

In this paper, the authors propose a measurement system utilizing a Ad-hoc LoRa network to transmit data to a monitoring center connected to commonly available industrial infrastructure. The system is practically deployed in Quy Nhon City, comprising one data collection and monitoring station and two data collection nodes.

Practical experiments were conducted to assess the feasibility of deploying voltage measurement and data collection in an urban environment, as well as to analyze the impact of distance, installation location, and terrain conditions on signal strength and data transmission time between connection points. The experimental results indicate that the system operates stably in an urban environment with appropriate distances, meeting the reliability and response time requirements of the system. However, the LoRa network also includes other parameters, such as the impact of energy consumption of the nodes, to provide a more comprehensive evaluation of the system's performance and applicability. Additionally, the feasibility of deploying the LoRa network on a larger scale with more nodes should be assessed to evaluate its practicality for more realistic scenarios.

## Acknowledgment

*This research is conducted within the framework of science and technology projects at institutional level of Quy Nhon University under the project code T2024.845.15.*

## REFERENCES

1. G. S. Elbasuony, S. H. E. A. Aleem, A. M. Ibrahim, A. M. Sharaf. A unified index for power quality evaluation in distributed generation systems, *Energy*, **2018**, *149*, 607-622.
2. S. Khokhar, A. A. B. M. Zin, A. S. B. Mokhtar, M. Pesaran. A comprehensive overview on signal processing and artificial intelligence techniques applications in classification of power quality disturbances, *Renewable and Sustainable Energy Reviews*, **2015**, *51*, 1650-1663.
3. A. Goldsmith, Andrea. *Wireless networks (8<sup>th</sup> edition)*, Cambridge University Press, New York, 2013.
4. J. P. S. Sundaram, W. Du, Z. Zhao. A survey on LoRa networking: research problems, current solutions and open issues, *IEEE Communications Surveys Tutorials*, **2020**, *22*, 371-388.
5. E. Viciano, F. M. A. Campos, A. Alcayde, R. Baños, F. G. Montoya. All-in-one three-phase smart meter and power quality analyzer with extended IoT capabilities, *Measurement*, **2023**, *206*, 112309.
6. C. Roach. Estimating electricity impact profiles for building characteristics using smart meter data and mixed models, *Energy and Buildings*, **2020**, *211*, 109686.
7. L. Czétány, V. Vámos, M. Horváth, Z. Szalay, Z. D. Béla, T. Csoknyai. Development of electricity consumption profiles of residential buildings based on smart meter data clustering, *Energy and Buildings*, **2021**, *252*, 111376.
8. F. S. Sutil, A. C. Ortega. Design and testing of a power analyzer monitor and programming device in industries with a LoRa LPWAN network, *Electronics*, **2021**, *10*(4), 453.
9. A. Rahman, M. Suryanegara. *The development of IoT LoRa: a performance evaluation on LoS and non-LoS environment at 915 MHz ISM frequency*, The 1<sup>st</sup> International Conference on Signals and Systems (ICSigSys), Bali, Indonesia, 2017.
10. Y. Ishida, D. Nobayashi, T. Ikenaga. *Experimental performance evaluation of the collisions in LoRa communications*, The 6<sup>th</sup> International Conference on Computational Science and Computational Intelligence (CSCI 2018), USA, 2018.
11. W. L. Rodrigues Junior, F. A. S. Borges, A. F. D. S. Veloso, R. D. A. L. Rabêlo, J. J. P. C. Rodrigues. Low voltage smart meter for monitoring of power quality disturbances applied in smart grid, *Measurement*, **2019**, *147*, 106890.

12. P. Kumari, R. Mishra, H. P. Gupta, T. Dutta, S. K. Das. An energy efficient smart metering system using edge computing in LoRa network, *IEEE Transactions on Sustainable Computing*, **2021**, 7, 786-798.
13. J. Sakamoto, D. Nobayashi, K. Tsukamoto, T. Ikenaga, G. Sato, K. Takizawa. *Poster: implementation and performance evaluation of TCP/IP communication over private LoRa*, The 30<sup>th</sup> International Conference on Network Protocols (ICNP), USA, 2022.
14. J. Sakamoto, D. Nobayashi, K. Tsukamoto, T. Ikenaga, G. Sato, K. Takizawa. *Improvement of TCP performance based on characteristics of private LoRa interface*, The 47<sup>th</sup> Annual Computers, Software, and Applications Conference (COMPSAC), Italy, 2023.
15. W. D. Paredes, H. Kaushal, I. Vakilinia, Z. Prodanoff. LoRa technology in flying Ad-hoc networks: a survey of challenges and open issues, *Sensors*, **2023**, 23(5), 2403.
16. J. Wu, J. Zhang, J. Ma, M. Zhao, X. Wang, X. Gao, H. Zhang. *Energy efficient 5G LoRa Ad-hoc network for Smart Grid Communication*, The 11<sup>th</sup> International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, 2021.
17. S. Bertoldo, C. Lucianaz, M. Paredes, M. Allegretti, L. Carosso, P. Savi. *Feasibility study of LoRa Ad-hoc network in an urban noisy environment*, The 18<sup>th</sup> Mediterranean Microwave Symposium (MMS), Istanbul, Turkey, 2018.
18. M. Arshad, M. M. Sandhu, M. Khan, A. Mahmood, T. Muhammad, M. A. Khan. Review of applications and challenges of LoRa technology for the internet of things (IoT), *Journal of Sensor and Actuator Networks*, **2017**, 6(2), 0007.



© 2025 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

# Toán tử hợp có trọng từ không gian kiểu Bloch vào không gian tăng trưởng trên hình cầu đơn vị của không gian Hilbert

Dương Thanh Vỹ\*

Khoa Sư phạm, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 11/08/2024; Ngày sửa bài: 16/10/2024;  
Ngày nhận đăng: 28/10/2024; Ngày xuất bản: 28/02/2025

## TÓM TẮT

Cho  $\nu, \mu$  là các trọng chuẩn tắc trên hình cầu đơn vị  $B_X$  của một không gian Hilbert phức với số chiều tùy ý và  $\psi$  là một hàm chỉnh hình trên  $B_X$ ,  $\varphi$  là một ánh xạ tự chỉnh hình của  $B_X$ . Trong bài báo này, chúng tôi nghiên cứu các đặc trưng cho tính bị chặn và tính compact của toán tử hợp có trọng  $W_{\psi, \varphi}$ ,  $f \mapsto \psi \cdot (f \circ \varphi)$ , từ không gian kiểu Bloch  $\mathcal{B}_\nu(B_X)$  đến không gian tăng trưởng (nhỏ)  $\mathcal{H}_\mu^\infty(B_X)$ ,  $\mathcal{H}_\mu^0(B_X)$  thông qua tính chất của  $\psi$ , các phiếm hàm đánh giá điểm  $\delta_{\varphi(z)}^{\mathcal{B}_\nu(B_X)}$ , và các hạn chế của các đại lượng này lên các không gian con  $m$ -chiều với  $m \geq 2$ . Chúng tôi cũng tính được chính xác công thức của chuẩn toán tử  $W_{\psi, \varphi}$ .

**Từ khóa:** Toán tử hợp có trọng, không gian Bloch, không gian tăng trưởng, tính bị chặn, tính compact.

\*Tác giả liên hệ chính.

Email: duongthanhvy@qnu.edu.vn

# Weighted composition operators from Bloch-type spaces into growth spaces on the unit ball of a Hilbert space

Duong Thanh Vy\*

Faculty of Education, Quy Nhon University, Vietnam

Received: 11/08/2024; Revised: 16/10/2024;

Accepted: 28/10/2024; Published: 28/02/2025

## ABSTRACT

Let  $\nu, \mu$  be normal weights on the unit ball  $B_X$  of an Hilbert space  $X$  with arbitrary dimension and  $\psi$  be a holomorphic function on  $B_X$  and  $\varphi$  a holomorphic self-map of  $B_X$ . In this work, we characterize the boundedness and the compactness of weighted composition operators  $W_{\psi, \varphi}$ ,  $f \mapsto \psi \cdot (f \circ \varphi)$ , from the Bloch-type spaces  $\mathcal{B}_\nu(B_X)$  to the (little) growth spaces  $\mathcal{H}_\mu^\infty(B_X)$ ,  $\mathcal{H}_\mu^0(B_X)$  via function theoretic properties of the symbol  $\psi$  and the point evaluation function  $\delta_{\varphi(z)}^{B_\nu(B_X)}$ , specifically, of the restrictions of functions  $\psi, \varphi$  to the  $m$ -dimensional subspaces for some  $m \geq 2$ . We obtain also the formula of the operator norm of  $W_{\psi, \varphi}$ .

**Keywords:** Bloch spaces, growth spaces, compactness, boundedness, weighted composition operator

## 1. INTRODUCTION

Let  $\mathcal{E}_1, \mathcal{E}_2$  be spaces of holomorphic functions on the unit ball  $B_X$  of a Banach space  $X$ ,  $\psi$  be a holomorphic function on  $B_X$  and  $\varphi$  a holomorphic self-map of  $B_X$ . The weighted composition operator, defined by symbols  $\psi$  and  $\varphi$ , maps from  $\mathcal{E}_1$  to  $\mathcal{E}_2$  and is defined by

$$W_{\psi, \varphi}(f) = M_\psi C_\varphi(f) = \psi \cdot (f \circ \varphi)$$

where  $M_\psi$  represents the multiplication operator with symbol  $\psi$  and  $C_\varphi$  is the composition operator with symbol  $\varphi$ .

In recent years, there has been significant interest in studying weighted composition operators. A famous theorem developed by Banach asserts that for a compact metric space  $K$ , the surjective linear isometries of  $C(K)$

are given by  $Tf = u(f \circ \varphi)$  where  $|u(x)| = 1$  for all  $x \in K$ , and  $\varphi : K \rightarrow K$  is a homeomorphism. Inspired by this theorem, ongoing research on the characterization of isometries in Banach spaces of analytic functions has revealed that weighted composition operators define the isometries of many such spaces, including the Hardy space  $H^p$  (for  $1 \leq p \leq \infty$ ,  $p \neq 2$ ), the weighted Bergman space, and the disk algebra.<sup>1</sup>

For a comprehensive overview of various aspects of the theory of (weighted) composition operators acting on several spaces of holomorphic functions, we refer to a standard reference.<sup>2</sup>

There is extensive literature on weighted composition operators and integral opera-

\*Corresponding author:

Email: duongthanhvy@qnu.edu.vn

tors between specific holomorphic function spaces. To address these spaces in a unified way, certain frameworks for Banach spaces of holomorphic functions on the unit disk have been introduced.<sup>3,4</sup>

For instance, in reference,<sup>3</sup> established certain topological and function-theoretic conditions for the domain space and provided criteria for boundedness and compactness, along with estimates for the operator norm and the essential norm of the weighted composition operators that map to the weighted-type space or the Bloch-type space on the unit disk. In recent years, there has been significant interest in the study of weighted composition operators. More recently, attention has also focused on composition operators and operator-valued multipliers in various vector-valued analytic function spaces, particularly when  $X$  is an infinite-dimensional Hilbert space.<sup>5–10</sup>

In this paper, we aim to consider the compactness and boundedness of  $W_{\psi,\varphi}$  when  $\mathcal{E}_1$  is a general Banach space of holomorphic functions and  $\mathcal{E}_2$  is either growth space  $\mathcal{H}_\mu^\infty(B_X)$  or the little growth space  $\mathcal{H}_\mu^0(B_X)$  determined as follows:

$$\mathcal{H}_\mu^\infty(B_X) = \left\{ f \in \mathcal{H}(B_X) : \sup_{z \in B_X} \mu(z)|f(z)| < \infty \right\},$$

$$\mathcal{H}_\mu^0(B_X) = \left\{ f \in \mathcal{H}_\mu^\infty(B_X) : \lim_{\|z\| \rightarrow 1} \mu(z)|f(z)| = 0 \right\},$$

where  $\mathcal{H}(B_X)$  is the space of holomorphic functions on  $B_X$  and  $\mu$  is a normal weight on  $B_X$ .

Growth spaces are a significant and intriguing class of Banach spaces of holomorphic functions. They have been investigated in various contexts, with numerous general and specialized references available.<sup>12,13</sup>

Some key properties of these spaces, when  $B_X$  is the unit disk  $\mathbb{B} \subset \mathbb{C}$ , include the following:

- For a normal weight  $\mu$ ,  $\mathcal{H}_\mu^\infty(\mathbb{B}) \supset \mathcal{H}^\infty$  if and only if  $\lim_{|z| \rightarrow 1} \mu(z) = 0$ . On the other hand, if  $\limsup_{|z| \rightarrow 1} \mu(z) > 0$ , then  $\mathcal{H}_\mu^0 = \{0\}$ ;
- The identity map  $I : \mathcal{H}_\mu^\infty(\mathbb{B}) \rightarrow (\mathcal{H}_\mu^\infty(\mathbb{B}), \tau_{co})$  is continuous.
- The bidual  $[\mathcal{H}_\mu^0(\mathbb{B})]''$  isometrically isomorphic to  $\mathcal{H}_\mu^\infty(\mathbb{B})$ ;
- The point evaluation functionals  $\delta_z^{\mathcal{H}}$  on  $\mathcal{H}_\mu^0(\mathbb{B})$  are bounded and can be uniquely extended to point evaluation functionals on  $\mathcal{H}_\mu^\infty(\mathbb{B})$  with the same norms;
- The operator  $\mathcal{B}_\mu^0(\mathbb{B}) \rightarrow \mathcal{H}_\mu^\infty(\mathbb{B})$ ,  $f \mapsto f''$ , is an isometric isomorphism, where  $\mathcal{B}_\mu^0(\mathbb{B})$  is the subspace of the Bloch space  $\mathcal{B}_\mu(\mathbb{B})$  of functions with  $f(0) = 0$ . It is important to note that the Bloch space consists of functions characterized by the growth of their derivatives, making it closely related to growth spaces;
- The maps  $z \mapsto \delta_z^{\mathcal{H}}$  is continuous, and  $\|\delta_z^{\mathcal{H}}\|$  goes to infinity as  $|z| \rightarrow 1$ .

These are just a few of the reasons motivating our research.

In Section 2, we review the key conditions for spaces of holomorphic functions that will be used to establish the boundedness and compactness of these operators, as well as to provide estimates for their essential norms in our context.

To characterize the boundedness and compactness, and building on the ideas from<sup>5,9</sup> with minor modifications, we establish in Section 3 a connection between functions in the growth space  $\mathcal{H}_\mu^\infty(B_X)$  and



their restrictions to finite-dimensional subspaces. Specifically, we show that if the restrictions of a function to  $m$ -dimensional subspaces (for  $m \geq 2$ ) have uniformly bounded growth norms, then the function belongs to the growth space  $\mathcal{H}_\mu^\infty(\mathbb{B}_m)$ , and vice versa.

In Section 4, we characterize the boundedness and the compactness of  $W_{\psi,\varphi}$  from  $\mathcal{B}_\nu(B_X)$  into  $\mathcal{H}_\mu^\infty(B_X)$  and into  $\mathcal{H}_\mu^0(B_X)$  as well as calculate the operator norms. We will show that these characterizations are completely determined by their behaviour on  $\psi^{[m]}$  and on the point evaluation functions  $\delta_{\varphi^{[m]}(z)}^{\mathcal{B}_\nu(B_X)}$  and  $\delta_{\varphi^{(m)}(z)}^{\mathcal{B}_\nu(B_X)}$ , where  $\psi^{[m]}$  and  $\varphi^{[m]}$  are the restrictions of  $\psi$  and  $\varphi$ , respectively, on the  $m$ -dimensional subspaces and  $\varphi^{(m)} = (\varphi_1, \dots, \varphi_m)$ ,  $m \geq 2$ .

Throughout this paper, we use the notions  $a \lesssim b$  and  $a \asymp b$  for non negative quantities  $a$  and  $b$  to mean  $a \leq Cb$  and, respectively,  $C^{-1}b \leq a \leq Cb$  for some inessential constant  $C > 0$ .

## 2. PRELIMINARIES AND AUXILIARY RESULTS

Let  $X$  be a complex Hilbert space of arbitrary dimension,  $Y$  a Banach space. Denote by  $B_X$  the closed unit ball of  $X$ , and use  $\mathbb{B}_n$  instead of  $B_{\mathbb{C}^n}$ . Fix an orthonormal basis  $(e_k)_{k \in \Gamma}$  of  $X$ . Then any  $z \in X$  can be expressed as

$$z = \sum_{k \in \Gamma} z_k e_k, \quad \bar{z} = \sum_{k \in \Gamma} \bar{z}_k e_k.$$

### 2.1. Möbius transformations

The analogues of Möbius transformations on a Hilbert space  $X$  are the mappings  $\Phi_a : B_X \rightarrow B_X$ ,  $a \in B_X$ , defined as follows:

$$\Phi_a(z) = \frac{a - P_a(z) - s_a Q_a(z)}{1 - \langle z, a \rangle}, \quad z \in B_X \quad (2.1)$$

where  $s_a = \sqrt{1 - \|a\|^2}$ ,  $P_a(z) = \frac{\langle z, a \rangle}{\|a\|^2} a$ , and  $Q_a(z) = z - \frac{\langle z, a \rangle}{\|a\|^2} a$  for  $z \in B_X$ .

We define  $\Phi_0(z) = -z$ .

Denote by  $\text{Aut}(B_X)$  the group of automorphisms of the unit ball  $B_X$ .

For details on Möbius transformations, we refer to K. Zhu's book.<sup>14</sup>

### 2.2. Banach spaces of holomorphic functions

$\mathcal{H}(B_X, Y)$  is denoted by the vector space of  $Y$ -valued holomorphic functions on  $B_X$ . An element  $f \in \mathcal{H}(B_X, Y)$  is named locally bounded holomorphic on  $B_X$  if for every  $z \in B_X$  there exists a neighbourhood  $U_z$  of  $0 \in X$  such that  $f(U_z)$  is bounded. Put

$$\begin{aligned} \mathcal{H}_{LB}(B_X, Y) \\ = \{f \in \mathcal{H}(B_X, Y) : f \text{ is locally bounded on } B_X\}. \end{aligned}$$

For  $f \in \mathcal{H}(B_X)$ , its complex gradient and radial derivative are defined by

$$\nabla f(z) = \left( \frac{\partial f}{\partial z_k}(z) \right)_{k \in \Gamma},$$

$$Rf(z) := \sum_{k \in \Gamma} \frac{\partial f}{\partial z_k}(z) (z_k e_k) = \langle z, \overline{\nabla f(z)} \rangle,$$

respectively. Thus,  $\nabla f(z)$  is the unique element in  $X$  representing the linear operator  $f'(z) \in X'$ , hence,

$$\begin{aligned} f'(z)(x) &= \sum_{k \in \Gamma} \frac{\partial f}{\partial z_k}(z) (x_k e_k) \\ &= \langle x, \overline{\nabla f(z)} \rangle, \quad x \in X. \end{aligned}$$

It is clear that

$$|Rf(z)| \leq \|\nabla f(z)\| \|z\|, \quad z \in B_X.$$

Now, let  $\mathcal{E} \subset \mathcal{H}(B_X)$  be a Banach space.

For each  $z \in B_X$ , the point-evaluation functional  $\delta_z^\mathcal{E}$  at  $z$  defined by  $\delta_z^\mathcal{E}(f) := f(z)$  for all  $f \in \mathcal{E}$ .

Thus,

$$|f(z)| \leq \|f\| \|\delta_z^\mathcal{E}\|, \quad f \in \mathcal{E}, z \in B_X, \quad (2.2)$$

where  $\|\delta_z^\mathcal{E}\| = \sup\{|f(z)| : f \in \mathcal{E}, \|f\| \leq 1\}$ .

For all  $\Phi = (\Phi_j)_{j \in \Gamma} \in \text{Aut}(B_X)$ , for every  $j \geq 1$ ,  $m \geq 2$  and all  $f \in \mathcal{E}$ , we write

$$\Phi_{(m)} = (\Phi_1, \dots, \Phi_m),$$

$$f \cdot \Phi_{(m)} = (f \cdot \Phi_1, \dots, f \cdot \Phi_m).$$

Below, we present a comprehensive list of conditions, some of which will be necessary for characterizing the boundedness, compactness, or determining the essential norm of the operators discussed in this work.

- (e1)  $\mathcal{E}$  includes the constant functions.
- (e2) The closed unit ball  $B_{\mathcal{E}}$  is  $\tau_{co}$ -compact.
- (e3) There are  $m \geq 2$  and constant  $C > 0$  such that for all  $\Phi \in \text{Aut}(B_X)$ , for all  $f \in \mathcal{E}$ ,  $\Phi_j \cdot f \in \mathcal{E}$ ,  

$$\|\Phi_j \cdot f\| \leq C\|f\|, \quad j \in \{1, \dots, m\}.$$

**Remark 2.1.** It follows from (e1) that  $\inf_{z \in B_X} \|\delta_z^{\mathcal{E}}\| > 0$ , and in particular, the following equivalent conditions are satisfied:

- (e1a) On each compact set,  $\|\delta_z^{\mathcal{E}}\|$  is bounded from below by a positive constant;
- (e1b) The functions in  $\mathcal{E}$  do not all vanish at each point  $z \in B_X$ .

Indeed, since the function  $1 \in \mathcal{E}$ , for every  $z \in B_X$  we have  $\|\delta_z^{\mathcal{E}}\| \geq \frac{1}{\|1\|}$ . It is obvious that (e1a)  $\Rightarrow$  (e1b). Now, assume that (e1b) holds but (e1a) fails. Then we can find a compact subset  $K$  of  $B_X$  and a sequence  $\{z_n\}_{n \geq 1} \in K$  and  $z_0 \in K$  such that  $z_n \rightarrow z_0$  and  $\|\delta_{z_n}^{\mathcal{E}}\| \rightarrow 0$ . This implies that  $f(z_0) = 0$  for all  $f \in \mathcal{E}$ , which is incompatible with (e1b).

By the uniform boundedness principle, we can easily prove the following:

**Proposition 2.1.** The mapping  $\delta^{\mathcal{E}} : B_X \rightarrow \mathbb{C}$ ,  $z \mapsto \|\delta_z^{\mathcal{E}}\|$ , is bounded on compact subsets of  $B_X$  for every Banach space  $\mathcal{E}$  of holomorphic functions on  $B_X$ .

### 3. GROWTH SPACES AND BLOCH-TYPE SPACES

For a normal weight  $\nu$  on  $B_X$ , we write

$$I_{\nu}^1(z) := \int_0^{\|z\|} \frac{dt}{\nu(t)}.$$

**Remark 3.1.** Since  $\nu$  is positive, continuous,  $m_{\nu, \delta} := \min_{t \in [0, \delta]} \nu(t) > 0$ . Moreover, it follows from  $(W_1)$  that  $\nu$  is strictly decreasing on  $[\delta, 1)$ , hence,  $\max_{t \in [0, 1)} \nu(t) =: M_{\nu} < \infty$ . Then, it is easily seen that

$$\nu(z)I_{\nu}^1(z) < R_{\nu} := \delta \frac{M_{\nu}}{m_{\nu, \delta}} + 1 - \delta < \infty. \quad (3.1)$$

for every  $z \in B_X$ .

We define bounded holomorphic spaces  $\mathcal{H}^{\infty}(B_X)$ , growth holomorphic spaces  $\mathcal{H}_{\mu}^{\infty}(B_X)$ , little growth holomorphic spaces  $\mathcal{H}_{\mu}^0(B_X)$ , Bloch-type spaces  $\mathcal{B}_{\nu}(B_X)$ , and little Bloch-type spaces  $\mathcal{B}_{\nu, 0}(B_X)$  on the unit ball  $B_X$  as follows:

$$\begin{aligned} \mathcal{H}^{\infty}(B_X) &= \left\{ f \in \mathcal{H}(B_X) : \sup_{z \in B_X} |f(z)| < \infty \right\}, \\ \mathcal{H}_{\mu}^{\infty}(B_X) &= \left\{ f \in \mathcal{H}(B_X) : \sup_{z \in B_X} \mu(z)|f(z)| < \infty \right\}, \\ \mathcal{H}_{\mu}^0(B_X) &= \left\{ f \in \mathcal{H}_{\mu}^{\infty}(B_X) : \lim_{\|z\| \rightarrow 1} \mu(z)|f(z)| = 0 \right\}, \\ \mathcal{B}_{\nu}(B_X) &:= \left\{ f \in \mathcal{H}(B_X) : \|f\|_{\mathcal{B}_{\nu}(B_X)} := \sup_{z \in B_X} \nu(z)|Rf(z)| < \infty \right\}. \end{aligned}$$

It is easy to check that  $\mathcal{H}^{\infty}(B_X)$ ,  $\mathcal{H}_{\mu}^{\infty}(B_X)$  and  $\mathcal{B}_{\nu}(B_X)$  are Banach under following norms

$$\begin{aligned} \|f\|_{\infty} &:= \sup_{z \in B_X} |f(z)|, \\ \|f\|_{\mathcal{H}_{\mu}^{\infty}} &:= \sup_{z \in B_X} \mu(z)|f(z)|, \\ \|f\|_{\mathcal{B}_{\nu}(B_X)} &:= |f(0)| + \|f\|_{\mathcal{B}_{\nu}(B_X)}, \end{aligned}$$

respectively.

Now we consider the holomorphic function

$$g(z) := 1 + \sum_{k > k_0} 2^k z^{n_k}, \quad z \in \mathbb{B}_1, \quad (3.2)$$

where  $k_0 = \lceil \log_2 \frac{1}{\nu(\delta)} \rceil$ ,  $n_k = \lceil \frac{1}{1-r_k} \rceil$  with  $r_k = \nu^{-1}(1/2^k)$  for every  $k \geq 1$ . Here, the symbol  $[x]$  represents the greatest integer less than or equal to  $x$ . By Theorem 2.3, <sup>15</sup>  $g(t)$  is increasing on  $[0, 1)$  and

$$\begin{aligned} |g(z)| &\leq g(\|z\|) \in \mathbb{R}, \quad z \in \mathbb{B}_1, \\ 0 &< C_1 := \inf_{t \in [0,1)} \nu(t)g(t) \\ &\leq \sup_{t \in [0,1)} \nu(t)g(t) \\ &\leq \sup_{z \in \mathbb{B}_1} \nu(z)|g(z)| =: C_2 < \infty. \end{aligned} \quad (3.3)$$

**Lemma 3.1.** Let  $\nu$  be a normal weight on  $B_X$ . Then there is  $C > 0$  such that for every  $z \in B_X$ ,

$$|f(z)| \leq \mu(z)^{-1} \|f\|_{\mathcal{H}_\mu^\infty(B_X)}, \quad f \in \mathcal{H}_\mu^\infty(B_X), \quad (3.4)$$

$$|f(z)| \leq C(1 + I_\nu^1(z)) \|f\|_{\mathcal{B}_\nu(B_X)}, \quad f \in \mathcal{B}_\nu(B_X). \quad (3.5)$$

*Proof.* The inequality (3.4) is obvious. The inequality (3.5) was proved in (Proof of Theorem 3.2) <sup>9</sup>.  $\square$

**Lemma 3.2.** For every  $\nu$  normal weight  $\nu$  on  $B_X$ , we have

- (1)  $\|\delta_z^{\mathcal{H}_\nu^\infty(B_X)}\| = 1/\nu(z)$ ;
- (2)  $\|\delta_z^{\mathcal{B}_\nu(B_X)}\| \asymp 1 + I_\nu^1(z)$ .

*Proof.* (1) It is obvious.

(2) It follows easily from the definition of  $\delta_z^{\mathcal{B}_\nu(B_X)}$  and (3.5) that

$$\|\delta_z^{\mathcal{B}_\nu(B_X)}\| \lesssim 1 + I_\nu^1(z).$$

Now we consider the function  $f_0$  defined by

$$f_0(z) = (1 + C_2)^{-1} (1 + \int_0^{\|z\|} g(t) dt), \quad z \in B_X,$$

where  $g$  is defined by (3.2). It is clear that  $f_0 \in \mathcal{B}_\nu(B_X)$  and by (3.3), it is easy to see that  $\|f_0\|_{\mathcal{B}_\nu(B_X)} \leq 1$ . Then, in view of (3.3) again, this yields that

$$\begin{aligned} \|\delta_z^{\mathcal{B}_\nu(B_X)}\| &\geq |f_0(z)| \\ &\geq \max \left\{ \frac{1}{1 + C_2}, \frac{C_1}{1 + C_2} \right\} (1 + I_\nu^1(z)). \end{aligned}$$

$\square$

It is easy to prove the following:

**Corollary 3.3.**  $\mathcal{H}_\nu^\infty(B_X)$ ,  $\mathcal{B}_\nu(B_X)$  satisfy the properties (e1), (e2), (e3).

We will demonstrate below that the analysis of growth spaces on  $\mathcal{B}_X$  can be reduced to studying functions defined on finite-dimensional subspaces. It is worth noting that similar results for Bloch-type spaces were recently studied in <sup>9</sup>.

For each finite subset  $F \subset \Gamma$ , without loss of generality we may assume that  $F = \{1, \dots, m\}$ , we denote by  $\mathbb{B}_m$  the unit ball of  $\text{span}\{e_k, k \in F\}$ . For each  $m \in \mathbb{N}$ , we write

$$\begin{aligned} \mu^{[m]} &:= \mu|_{\text{span}\{e_1, \dots, e_m\}}, \\ z_{[m]} &:= (z_1, \dots, z_m) \in \mathbb{B}_m. \end{aligned}$$

For  $m \geq 2$ , we denote by

$$\begin{aligned} OS_m &:= \{x = (x_1, \dots, x_m), \\ &\quad x_k \in X, \langle x_k, x_j \rangle = \delta_{kj}\} \end{aligned}$$

the family of orthonormal systems of order  $m$ .

For every  $x \in OS_m$  fixed and  $f \in \mathcal{H}(B_X)$ , we define

$$f_x(z_{[m]}) = f\left(\sum_{k=1}^m z_k x_k\right).$$

Then

$$\|\nabla f_x(z_{[m]})\| = \left\| \nabla f\left(\sum_{k=1}^m z_k x_k\right) \right\|. \quad (3.6)$$

**Definition 3.1.** We denote

$$\begin{aligned} \mathcal{H}_{\mu, \text{aff}}^\infty(B_X) \\ := \{f \in \mathcal{H}(B_X) : \|f\|_{\mathcal{H}_{\mu, \text{aff}}^\infty(B_X)} < \infty\} \end{aligned}$$

where

$$\|f\|_{\mathcal{H}_{\mu, \text{aff}}^\infty(B_X)} := \sup_{\|x\|=1} \|f(\cdot x)\|_{\mathcal{H}_\mu^\infty(\mathbb{B}_1)},$$

with  $f(\cdot x) : \mathbb{B}_1 \rightarrow \mathbb{C}$  given by  $f(\cdot x)(\lambda) = f(\lambda x)$  for every  $\lambda \in \mathbb{B}_1$ , and

$$\|f(\cdot x)\|_{\mathcal{H}_\mu^\infty(\mathbb{B}_1)} = \sup_{\lambda \in \mathbb{B}_1} \mu(\lambda x) |f(\lambda x)|.$$

It is easy to verify that  $\|\cdot\|_{\mathcal{H}_{\mu,\text{aff}}^\infty(B_X)}$  is a norm on  $\mathcal{H}_{\mu,\text{aff}}^\infty(B_X)$ , called the *affine norm*, moreover,  $(\mathcal{H}_{\mu,\text{aff}}^\infty(B_X), \|\cdot\|_{\mathcal{H}_{\mu,\text{aff}}^\infty(B_X)})$  is a Banach space.

**Proposition 3.4.** Let  $f \in \mathcal{H}(B_X)$ . The following are equivalent:

- (1)  $f \in \mathcal{H}_\mu^\infty(B_X)$ ;
- (2)  $\sup_{x \in OS_m} \|f_x\|_{\mathcal{H}_{\mu[m]}^\infty(\mathbb{B}_m)} < \infty$  for every  $m \geq 2$ ;
- (3)  $\sup_{x \in OS_m} \|f_x\|_{\mathcal{H}_{\mu[m]}^\infty(\mathbb{B}_m)} < \infty$  for some  $m \geq 2$ .

Moreover, for each  $m \geq 2$

$$\|f\|_{\mathcal{H}_\mu^\infty(B_X)} = \sup_{x \in OS_m} \|f_x\|_{\mathcal{H}_{\mu[m]}^\infty(\mathbb{B}_m)}. \quad (3.7)$$

*Proof.* (1)  $\Rightarrow$  (2): Let  $m \geq 2$  and  $z_{[m]} := (z_1, \dots, z_m) \in \mathbb{B}_m$ . Since  $\|\sum_{j=1}^m z_j e_j\| = \|z_{[m]}\|$ , we get

$$\begin{aligned} \|f_x\|_{\mathcal{H}_{\mu[m]}^\infty(B_X)} &= \sup_{z_{[m]} \in \mathbb{B}_m} \mu^{[m]}(z_{[m]}) |f_x(z_{[m]})| \\ &\leq \sup_{z \in B_X} \mu(z) \left| f\left(\sum_{j \in F} z_j e_j\right) \right| \\ &\leq \|f\|_{\mathcal{H}_\mu^\infty(B_X)} < \infty. \end{aligned} \quad (3.8)$$

In particular, we obtain (2).

(2)  $\Rightarrow$  (3): This is evident.

(3)  $\Rightarrow$  (1): Assume that

$$\sup_{x \in OS_m} \|f_x\|_{\mathcal{H}_{\mu[m]}^\infty(\mathbb{B}_m)} < \infty$$

for some  $m \geq 2$ . We fix  $z \in B_X \setminus \{0\}$ . Consider  $x = (\frac{z}{\|z\|}, x_2, \dots, x_m) \in OS_m$  and  $z_{[m]} := (\|z\|, 0, \dots, 0) \in \mathbb{C}^m$ . Then  $\|z_{[m]}\| = \|z\|$  and

$$|f_x(z_{[m]})| = \left| f\left(\sum_{k=1}^m z_k x_k\right) \right| = |f(z)|.$$

This yields that

$$\begin{aligned} \|f\|_{\mathcal{H}_\mu^\infty(B_X)} &= \sup_{z \in B_X} \mu(z) |f(z)| \\ &\leq \sup_{z \in B_X} \mu^{[m]}(z_{[m]}) |f_x(z_{[m]})| \\ &\leq \sup_{x \in OS_m} \|f_x\|_{\mathcal{H}_{\mu[m]}^\infty(\mathbb{B}_m)} < \infty. \end{aligned} \quad (3.9)$$

Thus  $f \in \mathcal{H}_\mu^\infty(B_X)$ .

On the other hand, it is clear that

$$\sup_{x \in OS_m} \|f_x\|_{\mathcal{H}_{\mu[m]}^\infty(\mathbb{B}_m)} \leq \|f\|_{\mathcal{H}_\mu^\infty(B_X)}$$

for every  $m \geq 2$ . (3.10)

Hence, we obtain (3.7) from (3.8), (3.9) and (3.10).  $\square$

**Remark 3.2.** In the case  $m = 1$ , the proposition is not true. Indeed, we consider  $\mu(z) := 1 - \|z\|^2$ , and  $f : B_X \rightarrow \mathbb{C}$  given by

$$f(z) := \sum_{n=1}^{\infty} \left\langle \frac{e_n}{n} - \frac{z}{\sqrt{n}}, e_n \right\rangle, \quad z \in B_X.$$

Then  $f \in \mathcal{H}(B_X)$  because

$$\begin{aligned} \sum_{n=1}^{\infty} \left| \left\langle \frac{e_n}{n} - \frac{z}{\sqrt{n}}, e_n \right\rangle \right|^2 &\leq \sum_{n=1}^{\infty} \frac{1}{n^2} + \|z\|^2 + \sum_{n=1}^{\infty} \frac{2}{n^{3/2}} < \infty. \end{aligned}$$

For each  $x = \sum_{n=1}^{\infty} \langle x, e_n \rangle e_n \in OS_1$  and for every  $z_{[1]} := z_k \in \mathbb{B}_1$  for some  $k \geq 1$ , we have

$$f_x(z_{[1]}) = f(z_k x_k) = \frac{1}{k} - \frac{z_k x_k}{\sqrt{k}},$$

and thus, since  $|f_x(z_{[1]})| \leq 2$ , we get

$$\begin{aligned} \sup_{x \in OS_1} \|f_x\|_{\mathcal{H}_\mu^\infty(\mathbb{B}_1)} &= \sup_{x \in OS_1} (1 - \|z_{[1]}\|^2) |f_x(z_{[1]})| \leq 2. \end{aligned}$$

However, since

$$\begin{aligned} (1 - \|z\|^2) |f(z)| &= (1 - \|z\|^2) \left| \sum_{n=1}^{\infty} \left\langle \frac{e_n}{n} - \frac{z}{\sqrt{n}}, e_n \right\rangle \right| \\ &\rightarrow \sum_{n=1}^{\infty} \frac{1}{n} \quad \text{as } z \rightarrow 0, \end{aligned}$$

we obtain that  $f \notin \mathcal{H}_\mu^\infty(B_X)$ .

By employing a similar argument as in the proof of Proposition 2.3 in,<sup>9</sup> we can easily obtain the following result, for which the proof will be omitted.

**Proposition 3.5.**  $\mathcal{H}_\mu^\infty(B_X) = \mathcal{H}_{\mu,\text{aff}}^\infty(B_X)$ .  
Moreover,

$$\begin{aligned} \|f\|_{\mathcal{H}_\mu^\infty(B_X)} &\leq \|f\|_{\mathcal{H}_{\mu,\text{aff}}^\infty(B_X)} \\ &\lesssim \|f\|_{\mathcal{H}_\mu^\infty(B_X)}, \quad f \in \mathcal{H}_\mu^\infty(B_X). \end{aligned}$$

#### 4. BOUNDEDNESS AND COMPACTNESS CRITERIA

Consider the weighted composition operator  $W_{\psi,\varphi}$  from  $\mathcal{B}_\nu(B_X)$  into  $\mathcal{H}_\mu^\infty(B_X)$  and into  $\mathcal{H}_\mu^0(B_X)$  defined by

$$(W_{\psi,\varphi}f)(z) := \psi(z) \cdot (f \circ \varphi)(z), \quad z \in B_X.$$

The component operators are the multiplication operator  $M_\psi f = \psi \cdot f$  and the composition operator  $C_\varphi f = f \circ \varphi$ , which correspond to the cases when the composition symbol  $\varphi$  is the identity function on  $\mathbb{B}$  and the multiplication symbol  $\psi$  is the constant function 1, respectively.

**Theorem 4.1.** The following are equivalent:

- (1)  $W_{\psi,\varphi} : \mathcal{B}_\nu(B_X) \rightarrow \mathcal{H}_\mu^\infty(B_X)$  is bounded;
- (2)  $M_{\psi,\varphi,\mu}^{[m]} := \sup_{z \in \mathbb{B}_m} \mu^{[m]}(z) |\psi^{[m]}(z)| \|\delta_{\varphi^{[m]}(z)}^{\mathcal{B}_\nu(B_X)}\| < \infty$  for some  $m \geq 2$ ;
- (3)  $M_{\psi,\varphi,\mu} := \sup_{z \in B_X} \mu(z) |\psi(z)| \|\delta_{\varphi(z)}^{\mathcal{B}_\nu(B_X)}\| < \infty$ .

In this case, we have

$$\|W_{\psi,\varphi}\| = M_{\psi,\varphi,\mu}. \quad (4.1)$$

*Proof.* (3)  $\Rightarrow$  (2): It is clear.

(1)  $\Rightarrow$  (3): Suppose  $W_{\psi,\varphi} : \mathcal{B}_\nu(B_X) \rightarrow \mathcal{H}_\mu^\infty(B_X)$  is bounded. Fix  $z \in B_X$ . For each  $f \in \mathcal{B}_\nu(B_X)$  with  $\|f\|_{\mathcal{B}_\nu(B_X)} \leq 1$ , we have

$$\begin{aligned} \mu(z) |\psi(z)| |f(\varphi(z))| &\leq \|W_{\psi,\varphi}f\|_{\mathcal{H}_\mu^\infty(B_X)} \\ &\leq \|W_{\psi,\varphi}\| \|f\|_{\mathcal{B}_\nu(B_X)} \leq \|W_{\psi,\varphi}\|. \end{aligned}$$

By definition of  $\delta_{\varphi(z)}^{\mathcal{B}_\nu(B_X)}$  (see Proposition 2.1), by taking the supremum over all  $f$  within the closed unit ball of  $\mathcal{B}_\nu(B_X)$ , we obtain:

$$\mu(z) |\psi(z)| \|\delta_{\varphi(z)}^{\mathcal{B}_\nu(B_X)}\| \leq \|W_{\psi,\varphi}\|.$$

Taking the supremum over all  $z \in B_X$  yields

$$M_{\psi,\varphi,\mu} \leq \|W_{\psi,\varphi}\| < \infty. \quad (4.2)$$

(2)  $\Rightarrow$  (1): Assume  $M_{\psi,\varphi,\mu}^{[m]} < \infty$  for some  $m \geq 2$ . Let  $f \in \mathcal{B}_\nu(B_X)$  with  $\|f\|_{\mathcal{B}_\nu(B_X)} \leq 1$ . We write  $z_x := \sum_{k=1}^m z_k x_k$  for each  $x \in OS_m$ . It should be noted that  $\|z_x\| = \|z_{[m]}\|$  and hence  $\mu^{[m]}(z_{[m]}) = \mu^{[m]}(z_x)$ . Then

$$\begin{aligned} \|(W_{\psi,\varphi}(f))_x\|_{\mathcal{H}_{\mu^{[m]}}^\infty(\mathbb{B}_m)} &= \sup_{z_x \in \mathbb{B}_m} \mu^{[m]}(z_x) |\psi^{[m]}(z_x)| (f \circ \varphi)_x(z_{[m]})| \\ &\leq M_{\psi,\varphi,\mu}^{[m]} < \infty \end{aligned}$$

for every  $x \in OS_m$ . By (3.7),  $W_{\psi,\varphi}$  is bounded because

$$\begin{aligned} \|W_{\psi,\varphi}(f)\|_{\mathcal{H}_\mu^\infty(B_X)} &= \sup_{x \in OS_m} \|(W_{\psi,\varphi}(f))_x\|_{\mathcal{H}_{\mu^{[m]}}^\infty(\mathbb{B}_m)} \\ &\leq M_{\psi,\varphi,\mu}^{[m]} < \infty. \end{aligned}$$

(4)  $\Rightarrow$  (2): For  $z \in B_X$ , we have

$$\begin{aligned} \mu(z) |\psi(z)| |f(\varphi(z))| &\leq \mu(z) |\psi(z)| \|\delta_{\varphi(z)}^{\mathcal{B}_\nu(B_X)}\| \\ &\leq M_{\psi,\varphi,\mu} < \infty. \end{aligned}$$

Consequently,

$$\|W_{\psi,\varphi}f\|_{\mathcal{H}_\mu^\infty(B_X)} \leq M_{\psi,\varphi,\mu} < \infty. \quad (4.3)$$

Finally, from (4.2), (4.3) we deduce (4.1).  $\square$

We next characterize the compactness of  $W_{\psi,\varphi}$ . As shown in <sup>10</sup>, we can demonstrate the following:

**Lemma 4.2** (<sup>10</sup>, Lemma 2.10). Let  $\mathcal{E}, \mathcal{F}$  be two Banach spaces of holomorphic functions on  $B_X$ . Assume that

- (1)  $\delta_z^\mathcal{E}$  are continuous for every  $z \in B_X$ ;
- (2) The closed unit ball of  $\mathcal{E}$  is  $\tau_{co}$ -compact.
- (3)  $T : (\mathcal{E}, \tau_{co}) \rightarrow (\mathcal{F}, \tau_{co})$  is continuous.



Then,  $T$  is compact if and only if for each bounded sequence  $\{f_n\}$  in  $\mathcal{E}$  satisfying  $f_n \rightrightarrows 0$  on compact sets, then  $\|Tf_n\|_{\mathcal{F}} \rightarrow 0$ .

**Theorem 4.3.** Assume that the operator  $W_{\psi, \varphi} : \mathcal{B}_\nu(B_X) \rightarrow \mathcal{H}_\mu^\infty(B_X)$  is bounded. Then, the following are equivalent:

(1) There exists  $m \geq 2$  such that

$$\lim_{r \rightarrow 1} \sup_{\|\varphi_{(m)}(z)\| > r} \mu(z) |\psi(z)| \|\delta_{\varphi_{(m)}(z)}^{\mathcal{B}_\nu(B_X)}\| = 0, \quad (4.4)$$

where  $\varphi_{(m)} := (\varphi_1, \dots, \varphi_m)$ .

(2)  $W_{\psi, \varphi}$  is compact.

*Proof.* First, we show that  $\psi \in \mathcal{H}_\mu^\infty(B_X)$ . Indeed, by the boundedness of  $W_{\psi, \varphi}$  and Theorem 4.1, we have  $M_{\psi, \varphi, \mu} < \infty$ . Then, by Remark 2.1,  $\inf_{z \in B_X} \|\delta_{\varphi(z)}^{\mathcal{B}_\nu(B_X)}\| =: \alpha > 0$ . Consequently,

$$\alpha \mu(z) |\psi(z)| < M_{\psi, \varphi, \mu}, \quad z \in B_X.$$

This means  $\psi \in \mathcal{H}_\mu^\infty(B_X)$ .

(2)  $\Rightarrow$  (1): Suppose  $W_{\psi, \varphi} : \mathcal{B}_\nu(B_X) \rightarrow \mathcal{H}_\mu^\infty(B_X)$  is compact. Fix  $m \geq 2$ . It is obvious that (4.4) holds if  $\overline{\varphi_{(m)}(B_X)} \cap \partial B_X \neq \emptyset$ . Then we can find sequence  $\{z^n\}_{n \geq 1} \subset B_X$  such that  $\|\varphi_{(m)}(z^n)\| \rightarrow 1$ . By the definition of  $\delta_{\varphi_{(m)}(z^n)}^{\mathcal{B}_\nu(B_X)}$ , with  $\varepsilon > 0$  is given we can find a sequence  $\{f_n\}_{n \geq 1} \subset \mathcal{B}_\nu(B_X)$  with  $\|f_n\|_{\mathcal{B}_\nu(B_X)} \leq 1$  for every  $n \geq 1$  satisfying

$$|f_n(\varphi_{(m)}(z^n))| > \|\delta_{\varphi_{(m)}(z^n)}^{\mathcal{B}_\nu(B_X)}\| - \varepsilon. \quad (4.5)$$

By the condition (e2), without loss of generality, we may assume that  $f_n \rightrightarrows 0$  in  $\mathcal{B}_\nu(B_X)$  on compact subsets of  $B_X$  and  $\{f_n\}_{n \geq 1}$  is uniformly bounded on compact sets.

For each  $n \geq 1$ , denote  $a^n := \varphi(z^n)$  and consider the automorphism  $\Phi_{a^n} \in \text{Aut}(B_X)$  defined by (2.1). For each  $j \in \{1, \dots, m\}$ , put

$$G_{a^n, j} := (a^n_{(m)})_j \cdot f_n - ((\Phi_{a^n})_{(m)})_j \cdot f_n.$$

By (e3),  $G_{a^n, j} \in \mathcal{B}_\nu(B_X)$ . It is an easy calculation to show that for every  $w \in B_X$ ,

$$\begin{aligned} |G_{a^n, j}(w)| &= |(a^n_{(m)})_j \cdot f_n(w) - ((\Phi_{a^n})_{(m)}(w))_j| \\ &\leq \frac{3\sqrt{1 - \|a^n_{(m)}\|^2}}{1 - \|w\|} |f_n(w)|. \end{aligned}$$

Then, by (2.2),

$$|G_{a^n, j}(w)| \leq \frac{3\sqrt{1 - \|a^n_{(m)}\|^2}}{1 - \|w\|} \|\delta_w\|,$$

consequently, by Proposition 2.1, and since  $\|a^n_{(m)}\| = \|\varphi_{(m)}(z^n)\| \rightarrow 1$  as  $n \rightarrow \infty$ , for each  $j \in \{1, \dots, m\}$ ,  $G_{a^n, j} \rightrightarrows 0$  on compact subsets of  $B_X$ . Now by the condition (e3), there exists  $C > 0$  such that for all  $j \in \{1, \dots, m\}$ , we have

$$\begin{aligned} \|G_{a^n, j}\|_{\mathcal{B}_\nu(B_X)} &\leq \|a^n_{(m)}\| \|f_n\|_{\mathcal{B}_\nu(B_X)} + \|((\Phi_{a^n})_{(m)})_j \cdot f_n\|_{\mathcal{B}_\nu(B_X)} \\ &\leq (C + 1) \|f_n\|_{\mathcal{B}_\nu(B_X)} \leq C + 1 \quad \forall n \geq 1. \end{aligned}$$

Therefore, since  $W_{\psi, \varphi}$  is compact, By (2.2) and Lemma 4.2,  $\|\psi \cdot ((G_{a^n, j}) \circ \varphi)\|_{\mathcal{H}_\mu^\infty(B_X)} \rightarrow 0$  as  $n \rightarrow \infty$  for every  $j \in \{1, \dots, m\}$ . Note that  $\Phi_{a^n}(a^n) = 0$ . Therefore, by (4.5), we have

$$\begin{aligned} &\mu(z^n) |\psi(z^n)| \|\varphi_{(m)}(z^n)\| (\|\delta_{\varphi_{(m)}(z^n)}^{\mathcal{B}_\nu(B_X)}\| - \varepsilon) \\ &\leq \mu(z^n) |\psi(z^n)| \|\varphi_{(m)}(z^n)\| \|f_n(\varphi_{(m)}(z^n))\| \\ &= \mu(z^n) |\psi(z^n)| \sqrt{\sum_{j=1}^m |(G_{a^n, j})(\varphi_{(m)}(z^n))|^2} \\ &= \sqrt{\sum_{j=1}^m \|\psi \cdot ((G_{a^n, j}) \circ \varphi)\|_{\mathcal{H}_\mu^\infty(B_X)}^2} \\ &\rightarrow 0 \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Consequently,

$$\begin{aligned} &\lim_{n \rightarrow \infty} \mu(z^n) |\psi(z^n)| \|\delta_{\varphi_{(m)}(z^n)}^{\mathcal{B}_\nu(B_X)}\| \\ &< \varepsilon \lim_{n \rightarrow \infty} \mu(z^n) |\psi(z^n)| \leq \varepsilon \|\psi\|_{\mathcal{H}_\mu^\infty(B_X)}. \end{aligned}$$

This implies that (4.4) holds because  $\varepsilon$  is arbitrary.

(1)  $\Rightarrow$  (2): Assume that (4.4) holds for some  $m \geq 2$ . By Lemma 4.2, it suffices to prove that if  $\{f_n\}_{n \geq 1} \subset \mathcal{B}_\nu(B_X)$ ,  $\|f_n\|_{\mathcal{B}_\nu(B_X)} \leq 1$  for all  $n \geq 1$  and  $f_n \Rightarrow 0$  on compact subsets of  $B_X$  then  $\|W_{\psi,\varphi} f_n\|_{\mathcal{H}_\mu^\infty(B_X)} \rightarrow 0$  as  $n \rightarrow \infty$ .

Let  $\{f_n\}_{n \geq 1}$  be such a sequence. Given  $\varepsilon > 0$ . Then we can choose a number  $r \in (0, 1)$  such that  $\mu(z)|\psi(z)|\|\delta_{\varphi(m)}(z)\| < \varepsilon$  whenever  $\|\varphi(m)(z)\| > r$ . Since  $|f_n(w)| \leq \|\delta_w^{\mathcal{B}_\nu(B_X)}\|$  for all  $w \in B_X$ , if  $\|\varphi(m)(z)\| > r$ , then  $\mu(z)|\psi(z)|\|f_n(\varphi(m)(z))\| < \varepsilon$ . Thus  $\mu(z)|\psi(z)|\|f_n(\varphi(m)(z))\| < \varepsilon$  when  $\|\varphi(z)\| > r$ , because  $\|\varphi(z)\| \geq \|\varphi(m)(z)\| > r$  for every  $z \in B_X$ .

Now, we consider the case  $\|\varphi(z)\| \leq r$ . Then  $\|\varphi(m)(z)\| \leq r$ . Note that

$$\begin{aligned} B[\varphi(m), r] \\ := \{\varphi(m)(y) : \|\varphi(m)(y)\| < r, y \in B_X\} \\ \subset \mathbb{B}_m \subset \mathbb{C}^m \end{aligned}$$

is relatively compact for every  $0 \leq r < 1$ , by the hypothesis,  $f_n \rightarrow 0$  uniformly on  $\overline{B[\varphi(m), r]}$ . Then, there exists  $N \in \mathbb{N}$  such that  $|f_n(w)| < \varepsilon/\|\psi\|_{\mathcal{H}_\mu^\infty(B_X)}$  for all  $n \geq N$ ,  $w \in \overline{B[\varphi(m), r]}$ . Thus,

$$\mu(z)|\psi(z)|\|f_n(\varphi(m)(z))\| < \varepsilon \quad \text{if } \|\varphi(z)\| \leq r.$$

□

We will now examine the boundedness and compactness of the operator  $W_{\psi,\varphi}$  mapping into  $\mathcal{H}_\mu^0(B_X)$ .

**Theorem 4.4.** The following are equivalent:

- (1)  $\psi \in \mathcal{H}_\mu^0(B_X)$ , and there is  $m \geq 2$ , such that or every  $0 \leq r < 1$ ,  $k \geq 1$ :

$$\varphi(m)(rB_X) \text{ is relatively compact; } (4.6)$$

$$\lim_{\|z\| \rightarrow 1} \mu(z)|\psi(z)|\|\delta_{\varphi(k)}^{\mathcal{B}_\nu(B_X)}\| = 0; \quad (4.7)$$

- (2)  $W_{\psi,\varphi}^0 : \mathcal{B}_\nu(B_X) \rightarrow \mathcal{H}_\mu^0(B_X)$  is compact.

*Proof.* (1)  $\Rightarrow$  (2): Suppose (1) holds. Fix  $f \in \mathcal{B}_\nu(B_X)$ . We show that  $W_{\psi,\varphi} f = \psi \cdot (f \circ \varphi) \in \mathcal{H}_\mu^\infty(B_X)$ . Since  $\mu(z)|\psi(z)|\|f(\varphi(k)(z))\| \rightarrow \mu(z)|\psi(z)|\|f(\varphi(z))\|$  as  $k \rightarrow \infty$  for each  $z \in B_X$ , and  $\mathcal{H}_\mu^0(B_X)$  is closed in  $\mathcal{H}_\mu^\infty(B_X)$ , it suffices to show that  $\psi \cdot (f \circ \varphi(k)) \in \mathcal{H}_\mu^0(B_X)$  for every  $k \geq 1$ . Given  $k \geq 1$ . By the hypothesis (1), for given  $\varepsilon > 0$  there is  $r \in (0, 1)$  such that

$$\begin{aligned} \mu(z)|\psi(z)|\|f(\varphi(k)(z))\| \\ \leq \mu(z)|\psi(z)|\|\delta_{\varphi(k)}^{\mathcal{B}_\nu(B_X)}\| \|f\|_{\mathcal{B}_\nu(B_X)} \quad (4.8) \\ \leq \varepsilon \|f\|_{\mathcal{B}_\nu(B_X)} \quad \text{for } \|z\| > r. \end{aligned}$$

On the other hand, assumption (1) implies that

$$\begin{aligned} \sup_{\|z\| \leq r} \mu(z)|\psi(z)|\|f(\varphi(k)(z))\| \\ \leq \mu(z)|\psi(z)|\|\delta_{\varphi(k)}^{\mathcal{B}_\nu(B_X)}\| \|f\|_{\mathcal{B}_\nu(B_X)} < \infty. \end{aligned} \quad (4.9)$$

Consequently,  $\psi \cdot (f \circ \varphi(k)) \in \mathcal{H}_\mu^\infty(B_X)$ . Moreover, by (4.8),  $\psi \cdot (f \circ \varphi(k)) \in \mathcal{H}_\mu^0(B_X)$ .

We also obtain from (4.8) and (4.9) that  $W_{\psi,\varphi}^0$  is bounded.

The compactness of the operator  $W_{\psi,\varphi}^0$  can now be established by following a similar argument as in the proof of Theorem 4.3 and using condition (4.6).

(2)  $\Rightarrow$  (1): First, since  $W_{\psi,\varphi}^0$  is bounded and  $1 \in \mathcal{B}_\nu(B_X)$  it is easy to check that  $\psi \in \mathcal{H}_\mu^0(B_X)$ .

In order to prove (4.6), first we have to show the following claim:

$$\frac{1}{2}\|z-w\| \leq \|\delta_z^{\mathcal{H}_\mu^\infty(B_X)} - \delta_w^{\mathcal{H}_\mu^\infty(B_X)}\|, \quad z, w \in B_X. \quad (4.10)$$

Indeed, by direct calculation, it is easy to check that

$$\begin{aligned} \frac{1}{2}\|z-w\| \\ \leq \sqrt{1 - \frac{(1-\|z\|^2)(1-\|w\|^2)}{|1-\langle z, w \rangle|^2}} \\ = \varrho_X(z, w), \end{aligned}$$

where  $\varrho_X$  is the pseudohyperbolic metric in  $B_X$  (see <sup>16</sup> p.99). On the other hand, we also

have

$$\varrho_X(z, w) = \sup_{\substack{f \in \mathcal{H}^\infty(B_X) \\ \|f\|_\infty \leq 1}} \varrho(f(z), f(w))$$

(see (3.4) in [5]), where  $\varrho(x, y) = \left| \frac{x-y}{1-\bar{x}y} \right|$  is the pseudohyperbolic metric in  $\mathbb{B}_1$ . Note that, since the function  $\eta \mapsto \frac{\eta}{1-f(z)f(w)}$  is holomorphic from  $\mathbb{B}_1$  into  $\mathbb{B}_1$  and  $f(z) - f(w) \mapsto 0$ , it follows from Schwarz's lemma that

$$\begin{aligned} \varrho_X(z, w) &\leq \sup_{\substack{f \in \mathcal{H}^\infty(B_X) \\ \|f\|_\infty \leq 1}} |f(z) - f(w)| \\ &\leq \sup_{\substack{f \in \mathcal{H}^\infty(B_X) \\ \|f\|_\infty \leq 1}} |\delta_z^{\mathcal{H}^\infty(B_X)}(f) - \delta_w^{\mathcal{H}^\infty(B_X)}(f)| \\ &= \|\delta_z^{\mathcal{H}^\infty(B_X)} - \delta_w^{\mathcal{H}^\infty(B_X)}\|. \end{aligned}$$

Hence, (4.10) is proved.

Next, we prove (4.6). For  $r \in (0, 1)$ , the set  $V_r := \{\delta_z^{\mathcal{H}^\infty(B_X)} : \|z\| \leq r\} \subset (\mathcal{H}^\infty(B_X))'$  is bounded. Then, since  $W_{\psi, \varphi}$  is compact, the set

$$(W_{\psi, \varphi})^*(V_r) = \left\{ \psi(z) \delta_{\varphi(z)}^{\mathcal{B}_\nu(B_X)} : \|z\| \leq r \right\}$$

is relatively compact in  $[\mathcal{B}_\nu(B_X)]'$ .

We know that, for every  $K \subset [\mathcal{B}_\nu(B_X)]'$  and every bounded subset  $D \subset \mathbb{C}$ , if the set  $\{t\eta : t \in D, \eta \in K\}$  is relatively compact in  $\mathcal{B}_\nu(B_X)$  then  $K \subset [\mathcal{B}_\nu(B_X)]'$  is also relatively compact. With this in mind, since the set  $\{\psi(z) : \|z\| \leq r\}$  is bounded, the set  $\{\delta_z^{\mathcal{B}_\nu(B_X)}, \|z\| \leq r\}$  is relatively compact. Then, it follows from (4.10) that  $\varphi(rB_X)$  is relatively compact, so is  $\varphi_{(m)}(rB_X)$  for  $m \geq 2$ .

Finally, we prove (4.7). Assume that there exist  $m \geq 1$ ,  $\varrho > 0$  and  $\{z^n\}_{n \geq 1} \subset B_X$ ,  $\|z^n\| \rightarrow 1$  such that  $\mu(z^n) |\psi(z^n)| \|\delta_{\varphi_{(m)}(z^n)}\| > \varrho$  for all  $n \geq 1$ . Then, we may choose  $\{f_n\}_{n \geq 1} \subset \mathcal{B}_\nu(B_X)$  such that  $\|f_n\|_{\mathcal{B}_\nu(B_X)} \leq 1$  and  $|f_n(\varphi_{(m)}(z^n))| > \|\delta_{\varphi_{(m)}(z^n)}^{\mathcal{B}_\nu(B_X)}\| - \varrho/2$  for every  $n \geq 1$ . Thus

$$\begin{aligned} \mu(z^n) |\psi(z^n)| |f(\varphi_{(m)}(z^n))| \\ > \varrho - \varrho/2 \mu(z^n) |\psi(z^n)|. \end{aligned}$$

Therefore, since  $\psi \in \mathcal{H}_\mu^0(B_X)$ ,  $W_{\psi, \varphi}^0 f_n \notin \mathcal{H}_\mu^0(B_X)$ . This contradicts the boundedness of  $W_{\psi, \varphi}^0$ .  $\square$

**Remark 4.1.** In the case where  $\dim X < \infty$ , and by following the proof of Theorem 4.4, the following statements are equivalent:

- (1)  $\lim_{\|z\| \rightarrow 1} \mu(z) |\psi(z)| \|\delta_{\varphi_{(k)}(z)}^{\mathcal{B}_\nu(B_X)}\| = 0$  for every  $k \geq 1$  and  $\psi \in \mathcal{H}_\mu^0(B_X)$ ;
- (2)  $W_{\psi, \varphi}^0 : \mathcal{B}_\nu(B_X) \rightarrow \mathcal{H}_\mu^0(B_X)$  is compact;
- (3)  $W_{\psi, \varphi}^0 : \mathcal{B}_\nu(B_X) \rightarrow \mathcal{H}_\mu^0(B_X)$  is bounded.

## Acknowledgment

*This research is conducted within the framework of science and technology projects at institutional level of Quy Nhon University under the project code T2024.854.29. The author gratefully acknowledges the many helpful suggestions of Professor Thai Thuan Quang during the preparation of the paper.*

## REFERENCES

1. R. J. Fleming, J. E. Jamison. *Isometries on Banach spaces: function spaces*, Chapman and Hall CRC, New York, 2003.
2. C. Cowen, B. Maccluer. *Composition operators on spaces of analytic functions*, Boca Raton CRC Press, 1995.
3. F. Colonna, M. Tjani. Operator norms and essential norms of weighted composition operators between Banach spaces of analytic functions, *Journal of Mathematical Analysis and Applications*, **2016**, 434(1), 93-124.
4. T. Eklund, P. Galindo, M. Lindström, I. Nieminen. Norm essential norm and weak compactness of weighted composition operators between dual Banach spaces of analytic functions, *Journal of Mathematical Analysis and Applications*, **2017**, 451(1), 1-13.

5. O. Blasco, P. Galindo, A. Miralles. Bloch functions on the unit ball of an infinite dimensional Hilbert space, *Journal of Functional Analysis*, **2014**, *267*, 1188-1204.
6. O. Blasco, P. Galindo, M. Lindström, A. Miralles. Composition operators on the Bloch space of the unit ball of a Hilbert space, *Banach Journal of Mathematical Analysis*, **2017**, *11*(2), 311-334.
7. J. Chen, M. Wang. Integration operators between Bloch type spaces and growth spaces on the unit ball of a complex Banach space, *Annales Polonici Mathematici*, **2021**, *127*, 177-199.
8. J. Laitila, H. O. Tylli, M. Wang. Composition operators from weak to strong spaces of vector-valued analytic functions, *Journal of Operator Theory*, **2009**, *62*, 281-295.
9. T. T. Quang. Banach-valued Bloch-type functions on the unit ball of a Hilbert space and weak spaces of Bloch-type, *Constructive Mathematical Analysis*, **2023**, *6*(1), 6-21.
10. M. Tjani. *Compact composition operators on some Möbius invariant Banach spaces*, Doctoral dissertation, Michigan State University, 1996.
11. L. A. Rubel, A. L. Shields. The second duals of certain spaces of analytic functions, *Journal of the Australian Mathematical Society*, **1970**, *11*, 276-280.
12. K. D. Bierstedt, J. Bonet, J. Taskinen. Associated weights and spaces of holomorphic functions, *Studia Mathematica*, **1998**, *127*, 137-168.
13. J. Bonet, P. Domanski, M. Lindstrom, J. Taskinen. Composition operators between weighted Banach spaces of analytic functions, *Journal of the Australian Mathematical Society*, **1998**, *64*, 101-118.
14. K. Zhu. *Spaces of Holomorphic Functions in the Unit Ball*, Springer-Verlag, New York, 2005.
15. Z. J. Hu, S. S. Wang. Composition operators on Bloch-type spaces, *Proceedings of the Royal Society of Edinburgh Section A*, **2005**, *135*, 1229-1239.
16. K. Goebel, S. Reich. *Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*, Marcel Dekker, New York, 1984.



© 2025 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

## Khảo sát ứng dụng tinh dầu gừng trong bảo quản Bánh ít lá gai

Võ Thị Trọng Hoa<sup>1,\*</sup>, Lê Duy Thanh<sup>1</sup>, Ngô Kim Khuê<sup>1</sup>,  
Nguyễn Thanh Hằng<sup>2</sup>, Lê Cảnh Định<sup>1</sup>

<sup>1</sup>Khoa Khoa học Tự nhiên, Trường Đại học Quy Nhơn, Việt Nam

<sup>2</sup>Phòng QC, Công ty cổ phần Dược - Trang thiết bị Y tế Bình Định, Việt Nam

Ngày nhận bài: 19/08/2024; Ngày sửa bài: 27/09/2024;

Ngày nhận đăng: 30/09/2024; Ngày xuất bản: 28/02/2025

### TÓM TẮT

Bánh ít lá gai là loại bánh đặc sản rất được ưa chuộng của tỉnh Bình Định, được làm chủ yếu từ bột nếp, lá gai, đậu xanh, dừa và đường với phương pháp chế biến truyền thống. Tuy nhiên, các nghiên cứu đã chỉ ra sự có mặt của một số chủng nấm mốc và vi khuẩn là nguyên nhân gây hư hỏng Bánh ít lá gai. Trong nghiên cứu này, tinh dầu gừng đã được bổ sung vào công thức bánh như một chất kháng khuẩn/kháng nấm giúp kéo dài thời gian sử dụng của bánh. Kết quả khảo sát ảnh hưởng của việc bổ sung tinh dầu gừng đến hạn sử dụng của Bánh ít lá gai thành phẩm cho thấy: bổ sung tinh dầu gừng với nồng độ 4 mL/kg bột có thể ngăn ngừa sự phát triển của vi sinh vật, đảm bảo kéo dài hạn sử dụng của Bánh ít lá gai thêm 3 ngày so với bánh không bổ sung tinh dầu gừng. Các chỉ tiêu cảm quan cũng chỉ ra rằng nồng độ tinh dầu bổ sung này phù hợp với thị hiếu của người sử dụng Bánh ít lá gai.

**Từ khóa:** Bánh ít lá gai, vi sinh vật, nấm mốc, tinh dầu gừng.

\*Tác giả liên hệ chính.

Email: vothitronghoa@qnu.edu.vn



# Investigating the application of ginger essential oil in Ramie Leaf Rice Cake preservation

Vo Thi Trong Hoa<sup>1,\*</sup>, Le Duy Thanh<sup>1</sup>, Ngo Kim Khue<sup>1</sup>,  
Nguyen Thanh Hang<sup>2</sup>, Le Canh Dinh<sup>1</sup>

<sup>1</sup>Faculty of Natural Sciences, Quy Nhon University, Vietnam

<sup>2</sup>QC Department, Binh Dinh Pharmaceutical - Medical Equipment Joint Stock Company, Vietnam

Received: 19/08/2024; Revised: 27/09/2024;

Accepted: 30/09/2024; Published: 28/02/2025

## ABSTRACT

Ramie leaf rice cake (RLRC) is a very popular specialty cake of Binh Dinh province, made mainly from glutinous rice flour, ramie leaves, green beans, coconut and sugar with traditional processing methods. However, studies have shown the presence of some strains of mold and bacteria as the cause of RLRC spoilage. In this study, ginger essential oil (GEO) was added to the cake formula as an antibacterial/antifungal agent to help prolong the shelf life of the cake. Survey results on the effects of adding GEO on the shelf life of RLRC show that adding GEO at a concentration of 4mL/kg of flour can prevent the growth of microorganisms, guaranteed to extend the shelf life of RLRC by 3 days compared to cakes without GEO added. Sensory criteria also show that the concentration of this additional essential oil is suitable for the tastes of RLRC consumers.

**Keywords:** *Ramie leaf rice cake, microorganisms, molds, ginger essential oil.*

## 1. INTRODUCTION

Ramie leaf rice cake is a well-known traditional specialty of Binh Dinh province. As the other rice cakes in Asia countries, these cakes made from the basic ingredients such as glutinous sticky rice, sugar, and some types of beans, coconuts, or sesame. The common methods for making these cakes include steaming, frying, and boiling. Because of manufacturing process without using the additives and tropical climate, the shel-life of these cakes is so short, about 2-3 days.<sup>1</sup> Research by Hoa et al has shown that four fungal strains are frequently present and cause damage to Binh Dinh RLRC, which are *Aspergillus niger*, *Aspergillus flavus*, *Aspergillus tamaris* and *Aspergillus fumigatus*.<sup>2</sup>

Thanh et al using extracts of natural compounds from cinnamon, anise, and black cardamom added to RLRC.<sup>3</sup> Research results showed that extracts of natural substances from black cardamom with the addition rate of 5 mL/kg of flour can help limit the deterioration of sensory quality and prevent the growth of microorganisms, thereby helping to prolong the storage time of RLRC up to 5 days in normal temperature, 2-3 days longer than traditional production methods. However, the above natural compounds have a strong aroma and affect the taste and smell of this product. Therefore, it is possible to research the use of some other essential oils that are more suitable for prolonging the preservation time while limiting the change

\*Corresponding author.

Email: vothitronghoa@qnu.edu.vn

in flavor properties of the cake and ensuring that it does not affect the health of consumers.

Ginger (*Z. officinale*) is a species that produces oleoresins and essential oils with chemical compounds of diverse properties. Essential oils have been characterized mainly as gingerols, which have been reported to possess analgesic, anti-inflammatory, anticancer, antipyretic, antioxidant, and antibacterial activities. The antifungal activity of *Z. officinale* essential oil is well documented, mainly in filamentous fungi such as *Penicillium spp.*, *Rhizopus sp.*, *A. flavus*, *A. solani*, *A. oryzae*, *A. niger*, *F. moniliforme*, *F. verticillioides* and in yeasts such as *Saccharomyces cerevisiae* and *C. albicans*.<sup>4-7</sup>

To enhance the shelf life of Ramie Leaf Rice Cake (RLRC) while maintaining consumer safety, we conducted an experiment using ginger essential oil (GEO) as a natural preservative. Our goal was to assess GEO's effectiveness in meeting microbiological standards and maintaining the sensory quality of RLRC over storage time.

## 2. MATERIALS AND METHOD

### 2.1. Material

The research materials included ramie leaves, glutinous rice, peeled mung beans, shredded coconuts, and banana leaves, which were sourced from the Binh Dinh Province. Refined sugar was obtained from TTC Bien Hoa-Dong Nai Sugar One-Member Limited Company, Vietnam.

GEO extracted by Hilltown Essential Oil Co. Ltd. (CC43 Ho Van Hue Street, Ward 9, Phu Nhuan District, Ho Chi Minh City, Vietnam).

Chemicals used in microbiological research: NA (Nutrient Agar), PDA (Potato Dextrose Agar), ethanol, NaCl 0.9%.

Equipment used in microbiological research: Benchtop pH meter BP 3001 (USA); quantitative balance; biological safety cabinets; autoclave sterilizer; Petri dishes; implant; alcohol

lamp; test tube; volumetric flasks, absorbent and waterproof cotton balls,...

Baking equipment such as dough mixer, electric stove,...

### 2.2. Method

#### 2.2.1. Method of counting total aerobic microorganisms

According to TCVN 4884-1:2015 (ISO 4833-1:2013): Microorganisms in the food chain - Methods for the enumeration of microorganisms. Microbiological analysis techniques were conducted according to Man.<sup>8</sup>

Using the pour plate technique, count the colonies on the agar medium after aerobic incubation at  $35 \pm 2$  °C for 24 to 48 hours. The number of aerobic bacteria in 1 g or 1 mL of the food product sample tested is calculated from the number of colonies counted on the culture plates according to the dilution concentrations.

Select plates with 15 to 300 colonies on plates of 2 consecutive dilutions. Calculate according to the formula:

$$N = A/(V.D_f)$$

In which:

N: total number of CFU/g (mL) of substrate sample analyzed

A: average number of CFU per petri dish at a certain dilution

V: volume of culture sample per petri dish

D<sub>f</sub>: sample dilution

#### 2.2.2. Method for determining antimicrobial ability of ginger essential oil

The antimicrobial activity test of GEO was performed using the agar disc diffusion method as described by Euloge et al (2012) with modifications. Different concentrations of essential oils were mixed into molten PDA medium (45 °C), shaken well to disperse the essential oils, and the medium was divided into sterile Petri dishes (90 mm in diameter). Fungal

rings (5 mm in diameter) were cut from the growing edge of the pure fungal culture (after 3 days) and transferred to the center of the Petri dish containing the essential oil medium, the control dishes (without essential oil) were inoculated according to the same procedure. Then the fungal dishes were incubated in the dark at  $28 \pm 2$  °C. Colony diameter was measured after 24 hours per time, lasting up to 120 hours. The percentage of inhibition of mycelial growth by essential oils was calculated according to the formula of Philippe et al.<sup>9</sup>

Inhibition of mycelial growth (%) =  $\frac{d_c - d_t}{d_c}$

Where:

$d_c$ : diameter of colony in control sample

$d_t$ : diameter of colony in sample containing essential oil.

2.2.3. Method of investigating the effect of ginger essential oil on the shelf life of Ramie leaf rice cake

In this experiment, the outer layer of the cake/cake crust was prepared using a ratio of 0.6 kg Ramie leaves and 0.5 kg sugar per 1 kg of glutinous rice flour. Ramie leaves were boiled, ground, and mixed with the flour. Ginger essential oil was incorporated during kneading. The filling was prepared using the traditional recipe. Steamed cakes were allowed to cool.



Figure 1. The crust (left) and filling (right) preparation.

The control (without GEO) and added GEO samples of RLRC were maintained at room temperature. The concentration of GEO selected will be based on the results of the survey on the antimicrobial ability of ginger essential oil as presented above. Assuming the inhibitory

concentration is  $a$ , we have the following table of GEO concentration survey.

Table 1. Essential oil concentration survey.

| GEO concentration | $a-2$ | $a-1$ | $a$ | $a+1$ | $a+2$ |
|-------------------|-------|-------|-----|-------|-------|
| Results           | M1    | M2    | M3  | M4    | M5    |

Step: 1 mL

Samples were taken to determine and evaluate the total number of aerobic microorganisms in the cake periodically after each day (24 hours).

2.2.4. Sensory quality assessment method

Assess the sensory quality of RLRC products by scoring method simulating the TCVN 3215-79 standard using the 20-point system. TCVN 3215-79 is a standard using the 20-point system built on a unified scale of 6 levels of 5 points (0 ÷ 5) in which point 0 corresponds to “damaged” product quality, and points from 1 ÷ 5 correspond to decreasing levels of defects. At point 5, the product is considered to have no defects or errors. In the considered properties, the product has typical and obvious good properties for that indicator. The 20-point system is divided into 6 levels of product quality classification.<sup>10</sup>

Table 2. Weight coefficient of indicators (according to TCVN 3215-79).

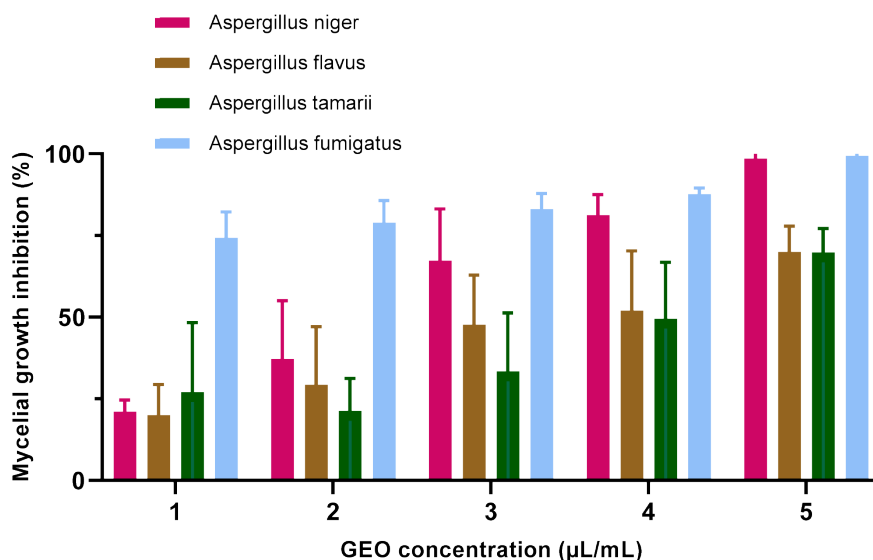
| No | Indicator name | Weight coefficient |
|----|----------------|--------------------|
| 1  | Status         | 1.5                |
| 2  | Color          | 0.6                |
| 3  | Odor           | 0.5                |
| 4  | Taste          | 1.4                |

Table 3. Sensory indicators are applied according to TCCS 01/2017/BILG-ST.

| No | Indicator name | Requirements                                          |
|----|----------------|-------------------------------------------------------|
| 1  | Status         | Pyramid cake, filling made of coconut and green beans |
| 2  | Color          | Black                                                 |
| 3  | Odor           | Aroma of ramie leaves                                 |
| 4  | Taste          | Sweet, rich, fatty                                    |

### 2.2.5. Data collection and processing methods

The reported data are the average values of 3 experimental repetitions. The data were processed and the graph were drawn using Graphpad Prism 8. The average values were analyzed by ANOVA using the Tukey test and the value of  $p < 0.05$  indicates a statistically significant difference.



**Figure 2.** Percentage of mycelial growth inhibition of GEO after 48 hours.

The result in Figure 2 and Figure 3 show that GEO has a strong inhibitory effect on *Aspergillus fumigatus* and *Aspergillus niger* strains but weaker inhibition on *Aspergillus flavus* and *Aspergillus tamarii* strains at almost all concentrations.

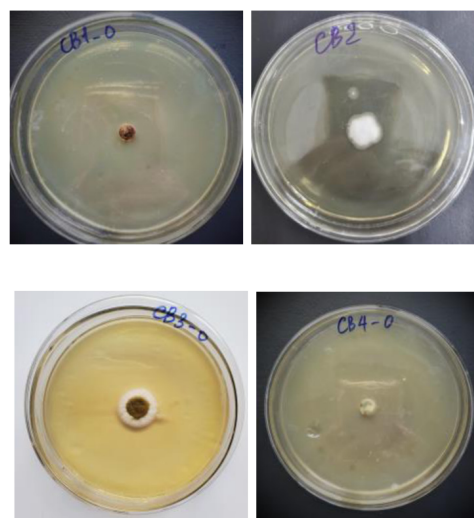
In addition, the survey result shows that the percentage of mycelium growth inhibition was proportional to the GEO concentration. The higher the concentration is, the stronger the inhibition ability is. The percentage of mycelium growth inhibition of GEO was significantly different ( $p < 0.05$ ), comparing the concentration of 3 μL/mL and the concentration of 5 μL/mL. However, the results between the concentration of 4 μL/mL and the concentration of 5 μL/mL showed an insignificant difference ( $p > 0.05$ ). We can rely on this to choose a more optimal concentration of essential oil, ensuring both antifungal effectiveness and the sensory properties of the cake (because high

## 3. RESULT

### 3.1. The ability of GEO to inhibit the growth of microorganisms under culture conditions

The results of investigating the ability of GEO to inhibit fungal mycelium growth with four pathogen fungal strains (*Aspergillus niger*, *Aspergillus flavus*, *Aspergillus tamarii* and *Aspergillus fumigatus*) after 48 hours are shown below.

GEO concentration makes the product have a “pungent” and “spicy” smell).

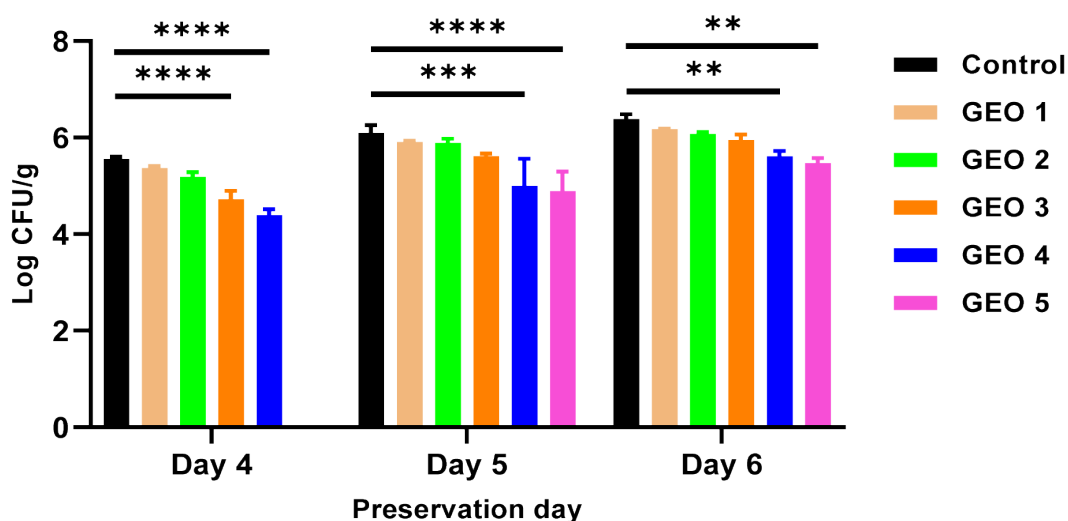


**Figure 3.** Image of the inhibition survey of GEO at a concentration of 5 mL/L on 4 fungal strains after 48 hours of culture (CB1 - *Aspergillus niger*, CB2 - *Aspergillus flavus*, CB3 - *Aspergillus tamarii*, CB4 - *Aspergillus fumigatus*).

### 3.2. Effect of GEO on the growth of microorganisms on RLRC during storage

The spoilage index of finished cakes is determined based on the total number of aerobic microorganisms developed on samples over storage time.

The following Figure 4 presents the results of the study on the growth of aerobic microorganisms in the cake samples during storage.



**Figure 4.** The total number of aerobic microorganisms (Log CFU/g) developed on cake samples over storage time (GEO 1, 2, 3, 4, 5 means the GEO concentration add to the cake with 1, 2, 3, 4, 5 mL/kg of flour, respectively) (\*\*\*\*:  $p < 0.0001$ , \*\*\*:  $p < 0.001$ , \*\*:  $p < 0.01$ , \*:  $p < 0.05$ ).

Initially, in the first 2 days of preservation period time, there is no damage on the RLRC surface. In the GEO5 sample, no aerobic microorganisms were detected on day 4, resulting in a recorded value of zero. After that, the control sample (without adding GEO) began to show viscous flow, while the samples using GEO began to gradually harden. In particular, since day 5, the total number of aerobic microorganisms (TNOAM) increased rapidly and exceeded the allowable limit according to TCVN 12941: 2020 standards, with the total number of aerobic microorganisms of the control sample reaching  $129.5 \times 10^4$  (CFU/g), approximately 6.11 logCFU/g.

Under normal conditions, on the 4<sup>th</sup> day of storage, there was no significant difference of the TNOAM between the samples supplemented with GEO and the control samples (without

adding GEO). However, the difference between the experimental samples and the control samples was significant ( $p < 0.05$ ) on the 5<sup>th</sup> day of storage, which shows that the addition of GEO contributed to limiting the growth of aerobic microorganisms on RLRC according to storage time.

The concentration of GEO added to the cake is inversely proportional to the TNOAM growing in RLRC, in which the GEO concentration of 5 mL/kg of flour shows the most obvious difference. The number of microorganisms in this experimental sample was the smallest and did not change much after 6 days of storage (the 6<sup>th</sup> day was not significantly different from the 4<sup>th</sup> day). The addition of ginger essential oil to the GEO 5 sample resulted in a significant reduction in total aerobic microbial count. On day 5, the



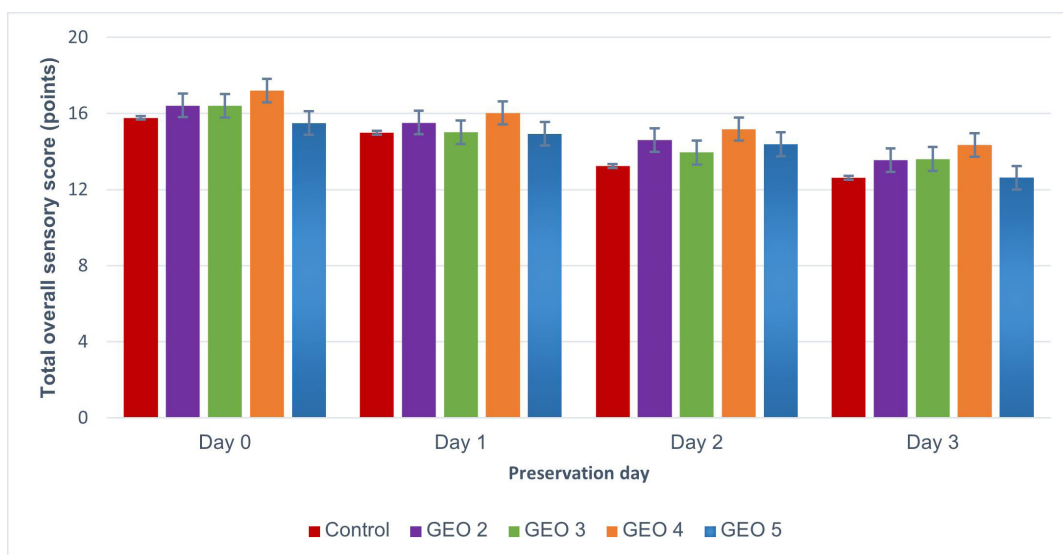
control sample had a count of 6.09 logCFU/g, whereas the GEO5 sample had a significantly lower count of 4.88 logCFU/g. However, there was no significant difference ( $p > 0.05$ ) in the total number of aerobic microorganisms between the sample supplemented with GEO concentration of 5mL/kg and the sample supplemented with GEO 4mL/kg. This shows that adding GEO at a concentration of 4 mL/kg or 5 mL/kg is equally effective in inhibiting the growth of microorganisms. Therefore, to ensure the sensory properties of the product and reduce raw material costs during production, we can choose a lower GEO concentration of 4 mL/kg

and still ensure the effectiveness of extending the preservation time of the cake for 3 more days compared to the previous traditional cake (only kept for a maximum of 3 days).

### 3.3. Effect of GEO on sensory quality of RLRC during storage

To evaluate the influence of GEO on the change in sensory quality of RLRC, a sensory test according to TCVN 3215-79 was performed.

The results of surveying the influence of GEO concentration on the sensory quality of RLRC according to storage time are shown in Figure 5 below.



**Figure 5.** Effect of GEO concentration on the change in sensory quality of RLRC over storage time.

The analysis results show that the overall sensory quality of RLRC reflected in the total average sensory score decreased with storage time. What's more there are differences between cake samples on the same day of storage. After 3 days of storing under normal conditions, the control sample without GEO supplementation show the slight decrease of the overall sensory score from 15.77 points, corresponding good grade, to 12.62 points, corresponding average grade. According to the microbiological test results, after 4 days of storage, the TNOAM in the control sample was under safe food standard, hence we did not continue to perform sensory evaluation after the 3<sup>rd</sup> day of storage.

The cake adding GEO has an aroma in harmony with the typical smell and taste of the traditional RLRC. Especially in the traditional method, people often use ginger in the cake filling as a way to add flavor delicious for cake. Maybe that's why the filling is less damaged than the crust (according to research by Le Duy Thanh).<sup>3</sup>

However, when observing the total sensory score results by day, we found that there was no obvious difference between samples GEO2, GEO3 and GEO4 with additional GEO concentrations of 2, 3 and 4 mL/kg flour, respectively. There is a significant difference between GEO4 and control as well as GEO5,

this shows that adding GEO increases the sensory quality of the cake. However, adding a lot of GEO will reduce the sensory quality of the cake, because it is not suitable for consumer tastes. This can be observed by comparing the sensory score of GEO4 and GEO5.

From the above analysis, it shows that, in terms of sensory quality, the cake sample supplemented with a GEO concentration of 4 mL/kg of flour has a higher sensory quality than the control sample on the first day of storage. Besides, the sensory score of the GEO4 sample decreased slowly over storage time. Therefore, the GEO concentration chosen to add to the traditional RLRC production process is 4 mL/kg of flour.

#### 4. CONCLUSION

The research result indicates that GEO can inhibit all four mold strains that cause RLRC spoilage, in which GEO achieves strong inhibition against *Aspergillus fumigatus* and *Aspergillus niger* strains, weaker inhibition against *Aspergillus flavus* strains and *Aspergillus tamaris* strains at almost all survey concentrations.

The results of surveying the effect of adding GEO on the shelf life of RLRC show that adding GEO concentration of 4 mL/kg of flour can prevent the growth of microorganisms, ensuring the extension of the shelf life of the RLRC product adds 3 days compared to cakes without added GEO. And the GEO concentration of 4 mL/kg of flour added to the cake meets the tastes of consumers.

Within the scope of the research, we have not yet evaluated the minimum inhibitory concentration of GEO against all four mold strains that cause RLRC spoilage. More surveys are needed to determine the minimum inhibitory concentration of GEO against these four strains.

#### Acknowledgements

*I would like to thank Institute of Food and Biotechnology, Industrial University of Ho Chi Minh City and Quy Nhon University for creating*

*favorable conditions for me to use laboratory equipment during this research.*

*This research is conducted within the framework of science and technology projects at institutional level of Quy Nhon University under the project code T2023.800.10.*

#### REFERENCES

1. Binh Dinh Department of Science and Technology. *Report on the results of implementing the science and technology project supporting the development of Binh Dinh province's distinctive products in 2015 (in Vietnamese)*, Binh Dinh, 2015.
2. T. T. H. Vo, T. T. Nguyen, T. M. N. Nguyen, T. M. D. Nguyen. Survival and growth of food borne harmful fungi during storage of Vietnamese ramie leaf rice cake, *Asian Journal of Dairy and Food Research*, **2024**, online first.
3. L. D. Thanh. *Research on the use of natural compounds in preserving Binh Dinh ramie leaf rice cake (Vietnamese)*, Quy Nhon University - Level Project Report, Binh Dinh, 2022.
4. K. S. Nguanpuag, S. Kanlayanarat, V. Srilaong, K. Tanprasert, C. Techavuthiporn. Ginger (*zingiber officinale*) oil as an antimicrobial agent for minimally processed produce: a case study in shredded green papaya, *International Journal of Agriculture and Biology*, **2011**, 13, 895-901.
5. S. Noori, F. Zeynali, H. Almasi. Antimicrobial and antioxidant efficiency of nanoemulsion-based edible coating containing ginger (*Zingiber officinale*) essential oil and its effect on safety and quality attributes of chicken breast fillets, *Food Control*, **2017**, 84, 312-320.
6. S. Munda, S. Dutta, S. Halder. Chemical analysis and therapeutic uses of ginger (*Zingiber officinale* Rosc.) essential oil: a review, *Journal of Essential Oil-Bearing Plants*, **2018**, 21(4), 994-1002.
7. E. I. C. López, M. F. H. Balcázar, J. M. R. Mendoza, A. D. R. Ortiz, M. T. O. Melo, R. S. Parrales, T. H. Delgado. Antimicrobial activity

- of essential oil of zingiber officinale roscoe (zingiberaceae), *American Journal of Plant Sciences*, **2017**, 8, 1511-1524.
8. L. V. V. Man, L. M. Huong. *Food microbiology experimentation (Vietnamese)*, Vietnam National University Ho Chi Minh City Publishing House, Ho Chi Minh City, 2010.
9. S. Philippe, F. Souaïbou, A. Guy, D. T. Sébastien, Y. Boniface, A. Paulin, Y. Issaka1, S. Dominique. Chemical composition and antifungal activity of essential oil of fresh leaves of ocimum gratissimum from Benin against six mycotoxigenic fungi isolated from traditional cheese wagashi, *Research Journal of Biological*, **2012**, 1, 22-27.
10. H. D. Tu. *Food sensory analysis techniques (Vietnamese)*, Science and Technology Publishing House, Hanoi, 2010.



© 2025 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).



# Ảnh hưởng của độ sâu mực nước đến sinh trưởng và tỷ lệ sống của ốc bươu đồng (*Pila polita*) trong điều kiện ương

Ngô Kim Khuê, Võ Văn Chí\*

Khoa Khoa học Tự nhiên, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 02/09/2024; Ngày sửa bài: 03/10/2024;

Ngày nhận đăng: 11/10/2024; Ngày xuất bản: 28/02/2025

## TÓM TẮT

Ốc bươu đồng (*Pila polita*) là một loài có chất lượng thịt ngon và giàu dinh dưỡng, vì vậy chúng đang là đối tượng đầy triển vọng cho nuôi trồng thủy sản. Nghiên cứu này được thực hiện để đánh giá ảnh hưởng của các độ sâu mực nước đến sinh trưởng và tỷ lệ sống của ốc bươu đồng. Ốc được ương thử nghiệm tại 4 độ sâu khác nhau là 15 cm, 20 cm, 25 cm và 30 cm với mật độ 300 con/bể 50 cm x 40 cm và được lặp lại 3 lần. Ốc được cho ăn bèo cám trong 15 ngày đầu và bầu, bí trong 15 ngày tiếp theo. Kết quả thí nghiệm cho thấy, sau 30 ngày ương, ốc được ương tại 30 cm cho giá trị sinh trưởng chiều cao và cân nặng lớn nhất (lần lượt là 12,87 mm và 0,495 g) trong khi chiều rộng đạt giá trị cao nhất tại độ sâu 25 cm và 30 cm (lần lượt là 9,62 mm và 9,59 mm). Tỷ lệ sống của ốc dao động từ 97,17 – 97,83%, trong đó ốc được ương tại 15 cm có tỷ lệ sống thấp hơn 3 nghiệm thức còn lại. Như vậy, xét tổng thể có thể thấy độ sâu mực nước 30 cm là phù hợp nhất đối với việc ương ốc.

**Từ khóa:** Ốc bươu đồng, sinh trưởng, tỷ lệ sống, độ sâu mực nước.

\*Tác giả liên hệ chính.

Email: vovanchi@qnu.edu.vn



# Effects of water depths on growth and survival of black apple snail (*Pila polita*) in rearing condition

Ngo Kim Khue, Vo Van Chi\*

*Faculty of Natural Sciences, Quy Nhon University, Vietnam*

*Received: 02/09/2023; Revised: 03/10/2024;*

*Accepted: 11/10/2024; Published: 28/02/2025*

## ABSTRACT

The black apple snail (*Pila polita*) is a species with delicious meat and rich nutrition, so it is a promising target for aquaculture. This study was conducted to evaluate the effects of water depths on the growth and survival rate of this species. The snails were experimentally reared at 4 different water depths of 15 cm, 20 cm, 25 cm and 30 cm with stocking density of 300 individuals per tank 50 cm x 40 cm and repeated 3 times. The snails were fed duckweed in the first 15 days and fresh calabash in the next 15 days. The experimental results showed that after 30 days of rearing, the snails reared at 30 cm had the highest height and weight (12.87 mm and 0.495 g, respectively) while the width of snails reached the highest values at depths of 25 cm and 30 cm (9.62 mm and 9.59 mm, respectively). The survival rate of snails ranged from 97.17 - 97.83%, in which snails reared at 15 cm had a lower survival rate than the other 3 treatments. Thus, in general, it can be seen that the water depth of 30 cm is the most suitable for rearing snails.

**Keywords:** *Black apple snail, growth, survival, water depths.*

## 1. INTRODUCTION

The black apple snail (*Pila polita*) is a relatively new species belonging to the Gastropod phylum with delicious meat and rich nutrients; therefore, it is a quite promising species for aquaculture. Culturing black apple snails has taken place widely in freshwater bodies across the country, especially in the Mekong Delta because snails have great potential for farming, high consumption demand, and bring stable economic efficiency to farmers. However, the natural resources of black apple snails are decreasing due to the invasion of the golden apple snail, overexploitation and an increasingly polluted environment.<sup>1</sup> There have been some studies

on seed production and rearing,<sup>2</sup> or biological characteristics of black apple snails.<sup>2-4</sup> Some other studies mentioned the effects of densities on growth and survival,<sup>5</sup> effects of different protein levels in diet on growth and survival,<sup>6</sup> or effects of pellet food supplemented Yucca Formulation on the growth and survival rate of this species.<sup>7</sup> In addition, the effects of some environmental factors on black apple snails have been reported in some research, for example, the effects of light conditions on the hatching rate, growth and survival rate,<sup>1</sup> effects of pH on growth and survival.<sup>8</sup> However, in the world and also in Vietnam, there have not been any studies on the effects of water depths on the growth or survival of this species. Therefore, this is the first

---

\*Corresponding author:

Email: vovanchi@qnu.edu.vn

to examine the effects of water depths on black apple snails during the rearing period.

## 2. MATERIALS AND METHODS

### 2.1. Experimental design

After hatched for 1 day, snails were chosen to arrange the experiment on water depth levels. The experiment was repeated 3 times with 4 different water levels (15 cm, 20 cm, 25 cm and 30 cm) used to rear snails in the tank of 50 cm x 40 cm. Water lettuce were used as the substrate with a distribution area of  $\frac{1}{4}$  of the tanks. 300 snails were selected and arranged in each tank. The total rearing time was 30 days. In the first 15 days, the snails were fed duckweed and in the next 15 days, they were fed fresh calabash twice a day (morning and afternoon). Every 3 or 4 days, 50% of the water in the tank was siphoned and changed, and 100% of the water was changed every 10 days.

### 2.2. Measurement of factors

The environment factors were defined. Temperature was measured every day using thermometer AS ONE JC-2211, and pH and DO were measured every day while  $\text{NH}_3$  was measured every 5 days using test Sera – Germany.

Survival and growth (the height, width and weight) of snails were examined every 10 days. Survival of snails was calculated as a percentage of a number of snails at examining time per number of snails at the beginning. The height and width of snails were measured using a caliper with an accuracy of 0.1 mm according to the method of Olena et al.<sup>9</sup> The weight of the snails was measured using a scale with an accuracy of 0.01 g.

### 2.3. Data analysis

Mean values and standard deviation were analysed using Microsoft Excel 2013. SPSS 20.0 using One-way ANOVA with LSD post hot test was applied to compare the differences between means of the treatments with significance  $p < 0.05$ .

## 3. RESULTS AND DISCUSSION

### 3.1. The environmental factors

During the experiment, temperature varied from 27 - 29 °C. Suitable temperature for black apple snails is 20 - 32 °C. When the temperature is below 15 °C or above 40 °C, the snails go into hibernation or aestivation.<sup>10</sup> Aquatic animals grow well at temperatures of 25 - 30 °C,<sup>11</sup> or proper temperature for aquaculture pond water is between 20 and 30 °C.<sup>12</sup> Thus, the temperature in the experiment is suitable for black apple snails.

Water was changed every 3 - 4 days, so pH was stable with a value of 6.5. Thao and Binh<sup>7</sup> reared these snails with 3 levels of pH (6, 7 and 8) and reported that snails grew and survived best at pH of 8. Phu and Ut believed that suitable pH for aquatic creatures is from 6.5 to 9.0.<sup>13</sup> Similarly, the suitable pH range for fish culture is 6.5 - 9.0.<sup>12</sup> According to Wurts and Durborow<sup>14</sup> and Bhatnagar et al.,<sup>15</sup> a pH between 6.5 and 9 is essential for aquaculture pond. Therefore, the pH value in this experiment is in a proper range although it is not the most optimal for black apple snails.

In this study, DO was 5 mg/L and  $\text{NH}_3$  was less than 0.03 mg/L. Bhatnagar et al.<sup>15</sup> and Bhatnagar and Singh reported that  $\text{DO} > 5$  ppm is essential for good aquaculture while Ekubo and Abowei believed that if fish are exposed to less than 0.3 mg/L of DO for a long time period, they can die and 5.0 mg/L is adequate for them.<sup>16,17</sup> The toxic levels of  $\text{NH}_3$  for short-term exposure usually lie between 0.6 and 2.0 mg/L, and sublethal effects may occur at 0.1 to 0.3 mg/L.<sup>18</sup> Maximum limit of ammonia concentration for aquatic animals is 0.05 mg/L.<sup>12</sup> According to Vietnamese standard TCVN 13952:2024 for freshwater aquaculture, proper DO and  $\text{NH}_3$  for freshwater animals are more than 4 mg/L and less than 0.1 mg/L respectively.<sup>19</sup> Thus, these values in the experiment are suitable for black apple snails.

3.2. The growth of black apple snails

At 10 days of rearing, the height of snails reared at 30 cm was highest, with value of 7.23 mm ( $p<0,05$ ) while these values in 3 remaining treatments were not different ( $p>0,05$ ). After 20 days of rearing, snails cultured at 15 cm had the lowest height (8.22 mm), followed by snails reared at 20 cm (8.87 mm), at 25 cm and 30 cm (9.97 mm and 9.90 mm respectively,  $p<0,05$ ).

However, the height of snails reared at 25 cm and 30 cm did not differed at 20 days of rearing ( $p>0.05$ ). At 30 days (Table 1), the height of snails in 4 treatments was totally different, of which, snails reared at 30 cm had the highest value (12.87 mm), followed by the values at 25 cm, 20 cm and 15 cm (12.06 mm, 11.08 mm and 10.54 mm respectively) ( $p<0.05$ ). Thus, the deeper water levels may be more suitable for the height growth of snails.

Table 1. The height (mm) of black apple snails at different water levels.

| Rearing time<br>(days) | Water depths              |                           |                           |                           |
|------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
|                        | 15 cm                     | 20 cm                     | 25 cm                     | 30 cm                     |
| 0                      | 3.62 ± 0.14 <sup>a</sup>  | 3.62 ± 0.14 <sup>a</sup>  | 3.62 ± 0.14 <sup>a</sup>  | 3.62 ± 0.14 <sup>a</sup>  |
| 10                     | 6.72 ± 0.47 <sup>b</sup>  | 6.58 ± 0.42 <sup>b</sup>  | 6.56 ± 0.45 <sup>b</sup>  | 7.23 ± 0.58 <sup>a</sup>  |
| 20                     | 8.22 ± 0.48 <sup>c</sup>  | 8.87 ± 0.85 <sup>b</sup>  | 9.97 ± 0.96 <sup>a</sup>  | 9.90 ± 0.67 <sup>a</sup>  |
| 30                     | 10.54 ± 0.41 <sup>d</sup> | 11.08 ± 0.57 <sup>c</sup> | 12.06 ± 0.44 <sup>b</sup> | 12.87 ± 0.84 <sup>a</sup> |

Note: In the same row, the different superscript letters indicate the significant difference ( $p<0.05$ ),  $n=30$ .

Similar to the height, after 10 days of rearing, the width of snails reared at a water depth of 30 cm reached the highest value (5.83 mm,  $p<0.05$ ) while the width of snails cultured at the 3 remaining water depths (15 cm, 20 cm and 25 cm) was not significantly different ( $p>0.05$ ). After 20 and 30 days of rearing, the widths of snails at water levels of 15 cm and

20 cm were not significantly different ( $p>0.05$ ) and smaller than those at water depths of 25 cm and 30 cm ( $p<0.05$ ). At 30 days of rearing (Table 2), the snail width at 15 cm, 20 cm, 25 cm, and 30 cm were 8.30 mm, 8.51 mm, 9.62 mm, and 9.59 mm, respectively. Therefore, it can be seen that snails adapt better at deeper water levels when they grow up.

Table 2. The width (mm) of black apple snails at different water levels.

| Rearing time<br>(days) | Water depths             |                          |                          |                          |
|------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
|                        | 15 cm                    | 20 cm                    | 25 cm                    | 30 cm                    |
| 0                      | 3.06 ± 0.05 <sup>a</sup> | 3.06 ± 0.05 <sup>a</sup> | 3.06 ± 0.05 <sup>a</sup> | 3.06 ± 0.05 <sup>a</sup> |
| 10                     | 5.28 ± 0.34 <sup>b</sup> | 5.25 ± 0.26 <sup>b</sup> | 5.34 ± 0.41 <sup>b</sup> | 5.83 ± 0.42 <sup>a</sup> |
| 20                     | 6.68 ± 0.43 <sup>b</sup> | 6.75 ± 0.67 <sup>b</sup> | 7.75 ± 0.31 <sup>a</sup> | 7.63 ± 0.56 <sup>a</sup> |
| 30                     | 8.30 ± 0.26 <sup>b</sup> | 8.51 ± 0.49 <sup>b</sup> | 9.62 ± 0.43 <sup>a</sup> | 9.59 ± 0.65 <sup>a</sup> |

Note: In the same row, the different superscript letters indicate the significant difference ( $p<0.05$ ),  $n=30$ .

The weight of snails after 10 days of rearing was highest at a depth of 30 cm ( $p<0.05$ ), while in the remaining 3 treatments there was

no statistical difference ( $p>0.05$ ). After 20 days of rearing, the weight of snails at 25 cm and 30 cm were the highest (0.224 g and 0.226 g,

respectively,  $p<0.05$ ), followed by weight at the depth of 20 cm (0.200 g) while at the depth of 15 cm was still the smallest (0.157 g,  $p<0.05$ ). After 30 days of rearing, the weight of snails in the 4 treatments was completely different (Table 3). Specifically, the weight of snails at the depth of

15 cm was the smallest (0.290 g), followed by this at 20 cm (0.328 g), 25 cm (0.454 g) and the highest value was at the depth of 30 cm (0.495 g). In general, snails tend to grow better when reared at greater depths.

**Table 3.** The weight (gram) of black apple snails at different water levels.

| Rearing time<br>(days) | Water depths               |                            |                            |                            |
|------------------------|----------------------------|----------------------------|----------------------------|----------------------------|
|                        | 15 cm                      | 20 cm                      | 25 cm                      | 30 cm                      |
| 0                      | 0.015 ± 0.005 <sup>a</sup> | 0.015 ± 0.005 <sup>a</sup> | 0.015 ± 0.005 <sup>a</sup> | 0.015 ± 0.005 <sup>a</sup> |
| 10                     | 0.084 ± 0.019 <sup>b</sup> | 0.083 ± 0.017 <sup>b</sup> | 0.080 ± 0.018 <sup>b</sup> | 0.103 ± 0.023 <sup>a</sup> |
| 20                     | 0.157 ± 0.019 <sup>c</sup> | 0.200 ± 0.053 <sup>b</sup> | 0.224 ± 0.027 <sup>a</sup> | 0.226 ± 0.040 <sup>a</sup> |
| 30                     | 0.290 ± 0.033 <sup>d</sup> | 0.328 ± 0.040 <sup>c</sup> | 0.454 ± 0.067 <sup>b</sup> | 0.495 ± 0.095 <sup>a</sup> |

*Note: In the same row, the different superscript letters indicate the significant difference ( $p<0.05$ ),  $n=30$ .*

After 30 days of rearing, with initial height, width, and weight of 3.62 mm, 3.06 mm, and 0.015 g, respectively, the black apple snails reached a height of 10.54 - 12.87 mm, a width of 8.30 - 9.62 mm and a weight of 0.290 - 0.495 g. These results are much higher than those of report by Binh and Thao although these authors reared snails at a lower density and longer time than us.<sup>5</sup> Specifically, Binh and Thao reported that after 35 days of rearing,<sup>5</sup> the height and weight of the snails only reached from 7.77 mm and 0.12 g at the density of 1200 snails/m<sup>2</sup> to 9.81 mm and 0.22 g at the density of 300 snails/m<sup>2</sup>. In another study of Thao and Binh,<sup>8</sup> snails reared at the density of 50 individuals/tank of 80 x 40 cm at different pH values also gave the lower growth at pH of 6 and 7 after 40 days of rearing compared to our results at pH of 6.5 after 30 days of rearing (heights ranging from 7.05 - 11.97 mm and weights ranging from 0.13 - 0.61 g). Or, Thao et al.<sup>7</sup> used pellet food supplemented Yucca formulation to feed black apple snails and the weight and height of snails after 40 days of rearing were only 0.45 - 0.50 g and 12.41 - 12.92 mm, respectively. We reared snails in tanks of 50 x 40 cm at the density of 300 snails/tank (equivalent to 1500 snails/m<sup>2</sup>) and used duckweed and fresh calabash as food

while the above authors fed them with industrial food. This may suggest that, at the early stage, industrial food may not be completely suitable for black apple snails. Another study showed similar results to ours despite only rearing snails at the density of 900 snails/m<sup>2</sup>.<sup>20</sup>

In this study, black apple snails were reared at water depth of 15 - 30 cm. In other studies, the snails were cultured at different water levels such as at 15 cm,<sup>5</sup> 30 cm,<sup>8</sup> or 30 - 40 cm;<sup>7</sup> however, the research results of these authors are lower than ours despite lower rearing densities. It is worth noticing that industrial foods or pellet foods were used in these studies, while we used duckweed and fresh calabash. Thus, it can be speculated that not only water depths effect on growth of snails, but food also plays a very important role, and foods used in this study are suitable for black apple snails while industrial foods or pellet foods may not really proper.

**3.3. Survival of black apple snails**

In general, the survival rate of snails during the experiment was very high, with a range of 97.17 - 98.17% (Table 4). At all examining times (10 days, 20 days, and 30 days), survival of snails at 15 cm was always lower than that of 3 other treatments ( $p<0.05$ ). At the end of the

experiment (30 days), survival of snails at 15 cm, 20 cm, 25 cm, and 30 cm were 97.17%, 98.17%, 97.83%, and 97.67%, respectively. These values are quite similar to the results of Diem et al.<sup>6</sup> when studied on the effects of different protein levels (15%, 20%, 25%, 30%, 35% and 40%) on growth and survival of snails (with 88.4% at 40% protein to 98.0% at 20% protein), or Thao et al. (95.83% - 99.33%).<sup>7</sup> Binh and Thao also reported a similar survival rate compared to our results when rearing snails at a density of 300 individuals/m<sup>2</sup> (97.1%) but lower at other densities (600, 900, and 1200 snails/m<sup>2</sup>).<sup>5</sup> Other studies also showed lower survival rates of snails,

eg. 66.7 - 88.3% or 34.7% - 76.7%.<sup>8,20</sup> Thus, it can be seen that the survival rate of snails in our study is much higher than in other studies,<sup>5,8,20</sup> proving that snails adapt well to the rearing conditions and foods. We reared them in tanks and used duckweed and fresh calabash as food. Although Thao and Binh reported that pH = 8 was a suitable value for black apple snails during the rearing stage,<sup>8</sup> it was clear that the survival rate was quite low compared to our results. This suggests that the authors might use inappropriate food or the addition of aluminum alum to reduce the pH could affect the survival of snails.

**Table 4.** The survival of black apple snails at different water levels.

| Rearing time<br>(days) | Water depths              |                           |                           |                           |
|------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
|                        | 15 cm                     | 20 cm                     | 25 cm                     | 30 cm                     |
| 0                      | 100 ± 0.00                | 100 ± 0.00                | 100 ± 0.00                | 100 ± 0.00                |
| 10                     | 97.67 ± 0.29 <sup>b</sup> | 99.00 ± 0.87 <sup>a</sup> | 98.33 ± 0.58 <sup>a</sup> | 98.67 ± 0.58 <sup>a</sup> |
| 20                     | 97.17 ± 0.29 <sup>b</sup> | 98.17 ± 0.76 <sup>a</sup> | 97.83 ± 0.29 <sup>a</sup> | 97.67 ± 0.29 <sup>a</sup> |
| 30                     | 97.17 ± 0.29 <sup>b</sup> | 98.17 ± 0.76 <sup>a</sup> | 97.83 ± 0.29 <sup>a</sup> | 97.67 ± 0.29 <sup>a</sup> |

Note: In the same row, the different superscript letters indicate the significant difference ( $p < 0.05$ ),  $n = 30$ .

4. CONCLUSION

The environmental factors in the experiment were suitable for the black apple snails. The height and weight of snails were biggest at a water depth of 30 cm (12.87 mm and 0.495 g, respectively), while the width of snails reached the highest value at 25 cm and 30 cm (9.62 mm and 9.59 mm, respectively). The survival rate of snails was very high (97.17 – 97.83%), and the snails reared at 20 cm, 25 cm, and 30 cm had higher survival rates than that of 15 cm. In general, the water depth of 30 cm is the most suitable for black apple snails in rearing conditions.

Acknowledgement

*This research is conducted within the framework of science and technology projects at institutional level of Quy Nhon University under the project code T2024.837.08.*

REFERENCES

1. N. T. T. Thao, N. T. N. Trang. Effects of light conditions on the hatching rate, growth and survival rate of black apple snail (*Pila polita*), *CTU Journal of Science*, **2015**, 38(1), 73-79.

2. V. X. Chu. *Study on some biological characteristics and trial seed production of the black apple snail Pila polita*, Master thesis, Tay Nguyen University, 2011.

3. L. V. Binh, N. T. T. Thao. Some reproductive biological characteristics of black apple snails distributed in the Mekong Delta, *CTU Journal of Science*, **2020**, 56, (2B), 117-126.

4. P. D. Phuc, V. X. Chu. Study on some biological characteristics of black apple snails *Pila polita* in Dak Lak, *Vietnam Journal of Agricultural Sciences*, **2014**, 23, 106-112.

5. L. V. Binh, N. T. T. Thao. Effect of different stocking densities on growth and survival rate of black apple snail larvae (*Pila polita*), *CTU Journal of Science*, **2014**, 1, 83-91.



6. V. T. K. Diem, L. V. Binh, N. T. T. Thao, N. T. Thanh, N. A. Tuan. Effects of different protein contents in diet on the growth and survival rate of black apple snail (*Pila polita*), *CTU Journal of Science*, **2018**, 54(3B), 177-185.
7. N. T. T. Thao, D. V. Danh, L. V. Binh. Effect of supplementation of yucca formulation on growth and survival rate of black apple snail (*Pila polita*), *Vietnam Journal of Agricultural Sciences*, **2023**, 21(7), 834-842.
8. N. T. T. Thao, L. V. Binh. Effect of pH on the results of hatching snails (*Pila polita*), *Journal of Agriculture and Rural Development*, **2018**, 2, 111-175.
9. O. Uvayeva, T. Vakaliuk, G. Shcherbina, E. Shimkovich. Sexual dimorphism in shell morphology of mollusks of the genus *Viviparus* – important objects of water resources of Ukraine, *E3S Web of Conferences*, **2021**, 280, 10011.
10. A. L. Kong, J. S. Kenny. The reproductive biology of the ampullariid snail *Pomacea urceus* (Müller), *Journal of Molluscan Studies*, **1989**, 55, 53-65.
11. C. E. Boyd. *Water quality in pond for aquaculture*, Department of Fisheries and Allied Aquaculture Auburn University, Alabama, 1998.
12. A. Bhatnagar, P. Devi. Water quality guidelines for the management of pond fish culture, *International Journal of Environmental Science*, **2013**, 3(6), 1980-2009.
13. T. Q. Phu, V. N. Ut. *Water quality management in fresh water ponds*, Publishing House of Agriculture, 2012.
14. W. A. Wurts, R. M. Durborow. *Interactions of pH, carbon dioxide, alkalinity and hardness in fish ponds*, Southern regional aquaculture center, SRAC Publication No. 464, 1992.
15. A. Bhatnagar, S. N. Jana, S. K. Garg, B. C. Patra, G. Singh, U. K. Barman. *Water quality management in aquaculture*, In: *Course Manual of summer school on development of sustainable aquaculture technology in fresh and saline waters*, CCS Haryana Agricultural, Hisar (India), 2004.
16. A. Bhatnagar, G. Singh. Culture fisheries in village ponds: a multilocation study in Haryana, India, *Agriculture and Biology Journal of North America*, **2010**, 1(5), 961-968.
17. A. A. Ekubo, J. F. N. Abowei. Review of some water quality management principles in culture fisheries, *Research Journal of Applied Sciences, Engineering and Technology*, **2011**, 3(2), 1342-1357.
18. H. R. Robinette. Effect of selected sublethal levels of ammonia on the growth of channel catfish (*Ictalurus punctatus*), *Progressive Fish Culturist*, **1976**, 38, 26-29.
19. National standards – TCVN 13952:2024, *Water for aquaculture - freshwater - quality requirements*, Ha Noi, 2024.
20. T. V. Tam. *Research to complete the technical process of seed production and commercial farming of black apple snail (Pila polita) in Hung Yen province*, The report of provincial project, 2022.



© 2025 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).



# Một điều kiện tương đương để một họ ma trận chuẩn tắc là chéo hóa tương đẳng đồng thời được

Võ Đức Tư Duy<sup>1</sup>, Lê Thanh Hiếu<sup>2,\*</sup>

<sup>1</sup>Phòng Quản lý Khoa học và Hợp tác Quốc tế, Trường Đại học Phan Thiết, Việt Nam

<sup>2</sup>Khoa Toán và Thống kê, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 25/08/2024; Ngày sửa bài: 29/09/2024;

Ngày nhận đăng: 08/10/2024; Ngày xuất bản: 28/02/2025

## TÓM TẮT

Trong bài báo này, chúng tôi đưa ra một điều kiện cần và đủ để một họ hữu hạn các ma trận chuẩn tắc là chéo hóa đồng thời được qua phép  $*$ -tương đẳng (gọi tắt là SDC). Bên cạnh đó, chúng tôi cũng xây dựng một gói lệnh MATLAB tương ứng để kiểm tra một họ hữu hạn các ma trận chuẩn tắc có là SDC hay không, và xác định ma trận tương đẳng làm chéo hóa đồng thời các ma trận ban đầu nếu các ma trận là SDC. Một số ví dụ minh họa và kiểm tra gói lệnh này cũng được trình bày trong bài báo.

**Từ khóa:**  $*$ -tương đẳng, chéo hóa tương đẳng đồng thời, chéo hóa tương đương đồng thời, ma trận chuẩn tắc.

\*Tác giả liên hệ chính.

Email: lethanhhieus@qnu.edu.vn

# An alternative equivalent condition for a finite family of normal matrices to be simultaneously diagonalizable via congruence

Vo Duc Tu Duy<sup>1</sup>, Le Thanh Hieu<sup>2,\*</sup>

<sup>1</sup>Department of Research Administration and International Relations, University of Phan Thiet, Vietnam

<sup>2</sup>Department of Mathematics and Statistics, Quy Nhon University, Vietnam

Received: 25/08/2024; Revised: 29/09/2024;

Accepted: 08/10/2024; Published: 28/02/2025

## ABSTRACT

This paper gives an alternative equivalent condition for a finite family of normal matrices to be simultaneously diagonalizable via  $*$ -congruence. The matrices do not need pairwise commute. A corresponding MATLAB package is developed. Some numerical tests for this package are also presented.

**Keywords:**  $*$ -congruence, simultaneous diagonalization via congruence, simultaneous diagonalization via similarity, normal matrices.

## 1. INTRODUCTION

The problems of simultaneously diagonalizing a family of matrices (via congruence or similarity) are known to be long-standing due to their applications, for examples, signal processing, data analysis and multi-linear algebra,<sup>1</sup> quadratic equations and optimization,<sup>2,3</sup> ...

There is a relationship between two concepts of simultaneous diagonalizations via similarity (SDS) and via congruence (SDC); see, e.g., in the book of Horn and Johnson.<sup>4</sup> One should, hence, need to distinguish these existing concepts as follows.

**Notations and definitions.** Let  $\mathbb{F}$  denote the field of real numbers  $\mathbb{R}$  or complex ones  $\mathbb{C}$ , and  $\mathbb{F}^{n \times n}$  be the set of all square matrices of order  $n$  with entries in  $\mathbb{F}$ . Let  $\mathcal{S}^n$ ,  $\mathcal{H}^n$ , and  $\mathcal{N}^n$  denote the sets of real symmetric, Hermitian, and normal matrices in  $\mathbb{F}^{n \times n}$ , respectively. By  $*$ ,  $.^T$ , we denote the conjugate transpose and transpose of a matrix, respectively. For  $A \in \mathcal{H}^n$ , we write  $A \succeq 0$  (resp.,  $A \succ 0$ ) for the meaning that  $A$  is positive semidefinite (resp., positive definite). As usual,  $I_{n \times d}$  denotes the  $n \times d$  identity matrix, and we shortly write  $I_n$  if  $n = d$ .

Matrices  $C_1, \dots, C_m \in \mathbb{F}^{n \times n}$  are said to be

- (i) *simultaneously diagonalizable via similarity* on  $\mathbb{F}$ , shortly  $\mathbb{F}$ -SDS, if there exists a nonsingular matrix  $P \in \mathbb{F}^{n \times n}$  such that  $P^{-1}C_iP$ 's are all diagonal matrices in  $\mathbb{F}^{n \times n}$ . When  $m = 1$ , we will say " $C_1$  is  $\mathbb{F}$ -DS", or  $\mathbb{F}$ -diagonalizable as usual;

- (ii) *simultaneously diagonalizable via  $*$ -congruence* on  $\mathbb{F}$ , abbreviated  $*$ -SDC, if there exists a nonsingular matrix  $P \in \mathbb{F}^{n \times n}$  such that  $P^*C_iP$  is diagonal for every  $i = 1, \dots, m$ . When  $m = 1$ , we will say " $C_1$  is  $*$ -DC";

In case  $C_i$ 's are all Hermitian, it is worth mentioning that the diagonal matrices  $P^*C_iP$ 's are always real due to the Hermitianian of  $C_i$ 's. Moreover,  $P$  can be chosen to be real if  $C_i$ 's are all real.<sup>5</sup>

- (iii) *simultaneously diagonalizable via  $T$ -congruence* on  $\mathbb{F}$ , abbreviated  $T$ -SDC, if there exists a nonsingular matrix  $P \in \mathbb{F}^{n \times n}$  such that  $P^TC_iP$  is diagonal for every  $i = 1, \dots, m$ . When  $m = 1$ , we will say " $C_1$  is  $T$ -DC".

Unlike the  $*$ -SDC case, the diagonal matrices  $P^*C_iP$ 's do not need to be real even  $C_i$ 's are real symmetric. The readers are referred to the work by Bustamante et. al.<sup>6</sup> for the  $T$ -SDC properties;

- (iv) *commuting* if they pairwise commute:  $C_iC_j = C_jC_i$  for every  $i, j = 1, \dots, m$ .

In the rest of this paper, the term " $\text{SDC}$ " will mean either "*simultaneous diagonalization via congruence*" or "*simultaneously diagonalizing via congruence*", or "*simultaneously diagonalizable via congruence*", and depending upon the situation, we will recognize the  $*$ - or  $T$ -congruence. It is analogous to the term " $\text{SDS}$ ".

\*Corresponding author.

Email: lethanhhieuv@qnu.edu.vn

**An overview of the SDC problem.** The SDC problem is known that first appeared in 1868 by Weierstrass,<sup>7</sup> in the 1930s by Albert,<sup>8</sup> Finsler,<sup>9</sup> Hertenes,<sup>10</sup> and later studies developed some conditions ensuring that two quadratic forms are SDC (see, e.g., in., works by More<sup>11</sup> and Pong<sup>12</sup> and references therein). However, these works provide only sufficient conditions, except for a few ones.<sup>4,13</sup>

From the practical point of view, Bunse-Gerstne et al.<sup>14</sup> proposed a Jacobi-like algorithm for SDC two commuting normal matrices, and this is numerically extended to several commuting ones by Mendle.<sup>15</sup> Recently, there have been some works<sup>3,5,16,17</sup> that present some (equivalent or sufficient) conditions for the  $*$ -SDC property of collections of either complex or real Hermitian matrices, and another one deals with the  $T$ -SDC problem for complex symmetric matrices.<sup>6</sup> It is noticed that the  $*$ - and  $T$ -congruences coincide only when the initial matrices are real symmetric, which are also real Hermitian. The two  $*$ -SDC and  $T$ -SDC problems are different, even if the initial matrices are symmetric. For example, Bustamante et. al.<sup>6</sup> show that the two real symmetric matrices

$$C_1 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \in \mathcal{S}^2$$

are  $T$ -SDC. But they are not  $*$ -SDC over  $\mathbb{C}$ .<sup>5</sup>

Several works deal with the normal SDC problem, i.e., the simultaneous diagonalization of several normal matrices via  $*$ -congruence.<sup>18,19</sup> However, they sound purely theoretical. There has been no algorithm to detect whether the given normal matrices are  $*$ -SDC.

**Contribution of the paper.** In this paper, we solve the normal SDC problem, i.e., the simultaneous diagonalization of several normal matrices via  $*$ -congruence. We first give a sufficient and necessary condition for a finite family of normal matrices to be simultaneously diagonalizable (via either congruence or similarity). It is noticed that the SDC property of a family of arbitrary square matrices can be checked by splitting the matrices into their Hermitian and skew-Hermitian parts.<sup>5</sup> The SDC property of the matrix family is confirmed if a positive definite matrix exists that solves a system of linear equations defined by the Hermitian and skew-Hermitian parts; see Theorem 7 below. The number of linear equations depends upon the number of Hermitian and skew-Hermitian parts. This may have a big computation complexity. Our (sufficient and necessary) condition in this paper restricts the number of such matrix linear equations.

On the other hand, we develop a MATLAB package to solve the normal SDC problem and its numerical tests.

**Auxiliary results.** We now recall some existing results on SDC that will be frequently used in this paper.

**Lemma 1.**<sup>3</sup> Suppose there is  $0 \neq \lambda \in \mathbb{R}^m$  such that

$\mathcal{C}(\lambda) \succ 0$ , where, without loss of generality, we assume  $\lambda_1 \neq 0$ . Then  $C_1, \dots, C_m \in \mathcal{S}^n$  is SDC if and only if  $P^T C_i P$  and  $P^T C_j P$  commute for all  $2 \leq i \neq j \leq m$ , where  $P$  is determined such that  $P^T \mathcal{C}(\lambda) P = I$  (the identity matrix).

As shown in the paper of Jiang and Li,<sup>3</sup> the matrix  $P$  in Lemma 1 is determined as  $P = U D^{1/2}$ , where  $U$  is orthogonal and  $D^{1/2}$  is the square root of the diagonal matrix  $D$  in an eigenvalue decomposition of  $\mathcal{C}(\lambda)$ :

$$D = U^T \mathcal{C}(\lambda) U.$$

The following results can be found in many books on Linear Algebra; their proofs are hence omitted in this paper.

**Lemma 2.**<sup>4</sup> (i) Every  $A \in \mathcal{H}^n$  can be diagonalized via similarity by a unitary matrix. That is, it can be written as  $A = U \Lambda U^*$ , where  $U$  is unitary,  $\Lambda$  is real diagonal and is uniquely defined up to a permutation of diagonal elements.

Moreover, if  $A \in \mathcal{S}^n$ , then  $U$  is picked to be real.

(ii) Suppose each of  $C_1, \dots, C_m \in \mathbb{F}^{n \times n}$  is  $\mathbb{F}$ -DS. Then, they are  $\mathbb{F}$ -SDS if and only if they are commuting.

(iii) Let  $A \in \mathbb{F}^{n \times n}$ ,  $B \in \mathbb{F}^{m \times m}$ . The matrix  $M = \text{diag}(A, B)$  is diagonalizable via similarity if and only if so are both  $A$  and  $B$ .

(iv) A complex symmetric matrix  $A$  is diagonalizable via similarity, i.e.,  $P^{-1} A P$  is diagonal for some invertible matrix  $P \in \mathbb{C}^{n \times n}$ , if and only if it is complex orthogonally diagonalizable, i.e.,  $Q^{-1} A Q$  is diagonal for some complex orthogonal matrix  $Q \in \mathbb{C}^{n \times n}$ :  $Q^T Q = I$ .

(v) Suppose  $A = \text{diag}(\alpha_1 I_{n_1}, \dots, \alpha_k I_{n_k})$ ,  $\alpha_i$ 's are distinct. If  $AB = BA$  then  $B = \text{diag}(B_1, \dots, B_k)$  with  $B_i \in \mathbb{F}^{n_i \times n_i}$  for all  $i = 1, \dots, k$ . Furthermore,  $B$  is Hermitian (resp., symmetric) if and only if so are  $B_i$ 's.

**Construction of the paper.** Section 2 is devoted to the SDC problem for normal matrices, in which we give a sufficient and necessary condition for a family of normal matrices to be SDC. And then, we propose a corresponding algorithm. Section 3 discusses the numerical experiments with respective our SDC algorithm in Section 2. We also give a numerical example illustrating our algorithm. Section 4 presents the conclusion.

## 2. THE NORMAL SDC PROBLEM

In this section, we deal with the normal SDC problem. Our conditions for a family of normal matrices to be  $*$ -SDC can be viewed as a generalization of that in Theorem 7 below. For convenience to the readers, we revisit these results as follows.

### 2.1. SDC and SDS of Hermitian matrices: revisited

We first summarize some existing results of the SDC and SDS of several Hermitian matrices. The following is presented in the book of Horn and Johnson whose proof does not completely give a nonsingular matrix that simultaneously diagonalizes the given matrices.<sup>4</sup> Our proof leads to an algorithm that may be use-



ful in practice. The idea is to follow that of proving Theorem 9 in the paper of Jiang and Duan for real symmetric matrices.<sup>3</sup>

**Theorem 3.**<sup>5</sup> The matrices  $I, C_1, \dots, C_m \in \mathcal{H}^n, m \geq 1$ , are SDC if and only if they are commuting. Moreover, when this is the case, they are SDC by a unitary matrix, and the resulting diagonal matrices are all real.

**Theorem 4.** Let  $A_1, \dots, A_m \in \mathbb{F}^{n \times n}, m \geq 1$ , be such that each of them is diagonalizable via similarity. Then, these matrices are simultaneously diagonalizable via similarity (shortly, SDS) if and only if they pairwise commute.

The following are not hard to prove, we omit their proofs.

**Lemma 5.** The matrices  $C_1, \dots, C_m \in \mathcal{H}^n$  are SDC if and only if for any  $\lambda \in \mathbb{R}^m$  with a  $\lambda_i \neq 0$ , the matrices  $C_1, \dots, C_{i-1}, \sum_{t=1}^m \lambda_t C_t, C_{i+1}, \dots, C_m$  are SDC.

**Lemma 6.**<sup>5</sup> The matrices  $C_1 = \begin{bmatrix} \hat{C}_1 & 0 \\ 0 & 0_k \end{bmatrix}, \dots, C_m = \begin{bmatrix} \hat{C}_m & 0 \\ 0 & 0_k \end{bmatrix}$  are SDC if and only if so are  $\hat{C}_1, \dots, \hat{C}_m$ .

Using Theorem 3, we comprehensively describe the SDC property of a family of Hermitian matrices as follows.

As a consequence of Theorem 3, every commuting collection of Hermitian matrices can be SDC. However, this is just a sufficient but unnecessary condition. For example, the matrices

$$C_1 = \begin{bmatrix} -1 & -2 & 0 \\ -2 & -28 & 0 \\ 0 & 0 & 5 \end{bmatrix}, C_2 = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 20 & 0 \\ 0 & 0 & -3 \end{bmatrix},$$

$$C_3 = \begin{bmatrix} 2 & 4 & 0 \\ 4 & 1 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

are SDC by

$$P = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

but  $C_1 C_2 \neq C_2 C_1$ . The following provides some equivalent SDC conditions for Hermitian matrices. It turns out that the SDC property of a family of such matrices is equivalent to the feasibility of a positive semidefinite program (SDP). This also allows us to use SDP solvers, for example, "CVX",<sup>20</sup> ... to check the SDC property of Hermitian matrices.

**Theorem 7.**<sup>5</sup> The following conditions are equivalent:

- (i) Matrices  $C_1, \dots, C_m \in \mathcal{H}^n$  are SDC.
- (ii) There exists a nonsingular matrix  $P \in \mathbb{C}^{n \times n}$  such that  $P^* C_1 P, \dots, P^* C_m P$  are commuting.

<sup>\*</sup>In fact, the skew-Hermitian part of  $A$  is usually defined as  $\frac{A-A^*}{2}$ .

- (iii) There exists a positive definite matrix  $X = X^* \in \mathcal{H}^n$  solves the following system of  $\frac{m(m-1)}{2}$  linear equations

$$C_i X C_j = C_j X C_i, \quad 1 \leq i < j \leq m. \quad (1)$$

If  $C_1, \dots, C_m$  are real, then one can pick  $P$  and  $X$  to be real.

## 2.2. The normal SDC problem

Recall that a square matrix  $N \in \mathbb{F}^{n \times n}$  is said to be normal if

$$N^* N = N N^*.$$

It is well-known that (real or complex) Hermitian, unitary, orthogonal matrices are normal, but the converse is not true in general. The readers are referred to, e.g., the work by Grone et al.,<sup>21</sup> for equivalent conditions for a normal matrix.

The third condition of Theorem 7 leads us to a sufficient and necessary condition for the \*-SDC property of a family of arbitrary square matrices. This can be done by splitting the matrices into their Hermitian and skew-Hermitian parts as follows. For square matrices  $A_1, \dots, A_m \in \mathbb{F}^{n \times n}$ , their Hermitian and skew-Hermitian parts<sup>\*</sup> are

$$A_i^h := \frac{A_i + A_i^*}{2}, \quad A_i^s := \frac{A_i - A_i^*}{2i}, \quad (2)$$

where  $i$  is the imaginary unit,  $i^2 = -1$ . Noticing that  $A_i^h$  and  $A_i^s$  are Hermitian and that

$$A_i = A_i^h + iA_i^s, \quad A_i^* = A_i^h - iA_i^s. \quad (3)$$

It is not hard to show that  $A_1, \dots, A_m$  are \*-SDC if and only if so are  $A_i^h, A_i^s, i = 1, \dots, m$ .

**Lemma 8.**<sup>13</sup> The square matrices  $A_1, \dots, A_m \in \mathbb{F}^{n \times n}$  are SDC if and only if so are  $A_i^h, A_i^s, i = 1, \dots, m$ .

Theorem 7 and Lemma 8 lead to a sufficient and necessary condition for a family of arbitrary square matrices to be \*-SDC, in which, after splitting up the initial matrices into Hermitian and skew-Hermitian, there are  $m(m-1)$  matrix equations as in (1). One can apply Theorem 7 and Lemma 8 to normal matrices. Below, we will introduce a smaller number of normal matrix equations; see Theorem 11.

Since any normal matrix is always diagonalizable by a unitary one,<sup>4</sup> it is diagonalizable via both sense similarity and congruence.

It is well-known that any finite family of commuting square matrices can be simultaneously upper triangularized by a unitary matrix.<sup>4</sup> Moreover, if these matrices are normal, then so are the resulting upper triangular matrices, and hence they are diagonal. Theorem 4 thus leads to the following observation.

**Lemma 9.** <sup>4</sup> Normal matrices  $N_1, \dots, N_m$  are SDC by a unitary matrix if and only if they pairwise commute.

Consequently, the normal matrices  $N_1, \dots, N_m$  are SDS if and only if they are SDC by a unitary matrix.

*Proof.* Suppose  $N_1, \dots, N_m$  pairwise commute. There exists a unitary matrix  $U$  such that  $U^*N_iU$  is upper triangular for every  $i = 1, \dots, m$ .<sup>4</sup> Since  $U^*N_iU =: T_i$  is normal due to the normality of  $N_i$ ,  $T_i$  must be diagonal. Thus  $N_1, \dots, N_m$  are SDC by the unitary matrix  $U$ .

Conversely, if  $N_1, \dots, N_m$  are SDC by a unitary matrix  $U$ , then  $U^*N_iU$ 's are diagonal and pairwise commute. This implies the commutativity of  $U^*N_iU$ 's and that of  $N_i$ 's.

The last part is obvious.  $\square$

**Lemma 10.** Let  $M, N$  be normal matrices and  $X$  be a square matrix of the same order  $n$ . The following statements are true:

i) The conditions

$$MXN = NXM \quad (4)$$

$$MXN^* = N^*XM \quad (5)$$

hold if and only if all the following conditions hold:

$$M^h \cdot X \cdot N^h = N^h \cdot X \cdot M^h, \quad (6)$$

$$M^h \cdot X \cdot N^s = N^s \cdot X \cdot M^h, \quad (7)$$

$$M^s \cdot X \cdot N^h = N^h \cdot X \cdot M^s, \quad (8)$$

$$M^s \cdot X \cdot N^s = N^s \cdot X \cdot M^s. \quad (9)$$

ii) Moreover, with the above materials and if  $X$  is Hermitian then (7)&(9) can be replaced by

$$M^h \cdot X \cdot M^s = M^s \cdot X \cdot M^h, \quad (10)$$

$$N^s \cdot X \cdot N^h = N^h \cdot X \cdot N^s. \quad (11)$$

*Proof.* The observation is derived from direct computations, see the Appendix 4, using the expansions (2) and (3) for  $M$  and  $N$ .  $\square$

The following is our main theorem.

**Theorem 11.** Let  $N_1, \dots, N_m \in \mathcal{N}^n, m \geq 2$ . The following conditions are equivalent:

- i)  $N_1, \dots, N_m$  are SDC.
- ii) There exists a nonsingular matrix  $P$  such that the matrices  $P^*N_tP, P^*N_t^*P, t = 1, \dots, m$ , pairwise commute.
- iii) There exists a positive definite matrix  $X$  such that

$$\begin{aligned} N_iXN_j &= N_jXN_i \text{ and} \\ N_iXN_j^* &= N_j^*XN_i, \quad 1 \leq i \leq j \leq m. \end{aligned} \quad (12)$$

- iv) The matrices  $N_t^h, N_t^s, t = 1, \dots, m$ , are SDC.

*Proof.* The equivalence of i) and iv) is obvious due to the authors' work.<sup>5</sup> Theorem 3.1

i)  $\Rightarrow$  ii). Suppose  $N_1, \dots, N_m$  are SDC by a nonsingular matrix  $P$ , that is the matrix  $P^*N_iP$  is diagonal, and so is  $P^*N_i^*P$ , for every  $i = 1, \dots, m$ . It is then obvious  $P^*N_iP, P^*N_i^*P, t = 1, \dots, m$ , pairwise commute.

ii)  $\Rightarrow$  iii). Suppose the  $2m$  matrices  $P^*N_iP, P^*N_i^*P, i = 1, \dots, m$ , pairwise commute, for some nonsingular matrix  $P$ . Then

$$(P^*N_iP) \cdot (P^*N_jP) = (P^*N_jP) \cdot (P^*N_iP),$$

$$(P^*N_iP) \cdot (P^*N_j^*P) = (P^*N_j^*P) \cdot (P^*N_iP),$$

for every  $i \neq j$ . This implies

$$N_i(PP^*)N_j = N_j(PP^*)N_i,$$

$$N_i(PP^*)N_j^* = N_j^*(PP^*)N_i,$$

for every  $i \neq j$ . The conclusion is obvious with  $X = PP^*$ .

iii)  $\Rightarrow$  i). Let  $Q$  be the square root of  $X \succ 0$  satisfying (12). Note that  $Q = Q^*$ . It follows from (12) that the matrices  $QN_tQ, QN_t^*Q$ 's pairwise commute. This implies that, for  $1 \leq t, l \leq m$ ,

$$N_tQ^2N_l = N_lQ^2N_t, \quad N_tQ^2N_l^* = N_l^*Q^2N_t.$$

Applying Lemma 10 to each pair of  $(t, l)$  and  $X = Q^2 = Q^*Q \succ 0$ , one obtains the commutativity of the Hermitian matrices

$$QN_t^hQ, \quad QN_t^sQ, \quad t = 1, \dots, m.$$

By Lemma 9, these latter matrices are SDC by a unitary matrix  $V$ , and hence so are the matrices  $QN_tQ$ 's due to Lemma 8. This yields  $N_1, \dots, N_m$  are SDC by the nonsingular matrix  $U = QV$ .  $\square$

**Algorithm 1.** SDC of normal matrices.

INPUT:  $N_1, \dots, N_m \in \mathcal{N}^n$ .

OUTPUT: A nonsingular matrix  $U$  such that  $U^*N_iU$ 's are diagonal.

*Step 1:* If the system (12) has a positive definite solution  $X$ , go to the next step.  
Otherwise, conclude the initial matrices are not SDC.

*Step 2:* Compute the square root  $X^{\frac{1}{2}}$  of  $X$  by using eigenvalue decomposition of  $X$ .

*Step 3:* Simultaneously diagonalizing the commuting and Hermitian matrices

$$\frac{1}{2}X^{\frac{1}{2}}(N_i + N_i^*)X^{\frac{1}{2}}, \quad \frac{i}{2}X^{\frac{1}{2}}(N_i^* - N_i)X^{\frac{1}{2}},$$

for  $i = 1, \dots, m$ , by applying the Jacobi-like al-

gorithm,<sup>5</sup> to determine a unitary matrix  $V$ . Return  $U = QV$ .

The last step of Algorithm 1 can apply the Jacobi-like algorithm Algorithm 3.1 exploiting the works by Bunse-Gerstner et. al. and by Mendl.<sup>5,14,15</sup>

*Example 1.* The real symmetric matrices

$$C_1 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \quad C_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix},$$

which are normal, are  $\mathbb{C}$ -SDC as shown in the work of Bustamante and collaborators.<sup>6</sup> However, they are not  $\mathbb{R}$ -SDC due to Theorem 7. Indeed, we want to check if there is a positive semidefinite matrix  $X = \begin{bmatrix} x & y \\ y & z \end{bmatrix} \succ 0$ , which is equivalent to  $x > 0$  and  $xz > y^2$ , such that

$$C_1XC_2 = C_2XC_1 (= (C_1XC_2)^*).$$

This is equivalent to

$$\begin{cases} x > 0, & xz > y^2 \\ x + y + z = 0. \end{cases}$$

But the last condition is impossible since there do not exist  $x, z > 0$  such that  $xz > y^2 = (x+z)^2$ . Thus  $C_1$  and  $C_2$  are not SDC on  $\mathbb{R}$ .  $\diamond$

*Example 2.* . Let

$$N_1 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad N_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$N_3 = \begin{pmatrix} 3i & -i & i \\ -i & 5i & -i \\ i & -i & 3i \end{pmatrix}.$$

Theorem 11 leads to finding a positive definite matrix

$$X = \begin{bmatrix} x & y & z \\ \bar{y} & t & u \\ \bar{z} & \bar{u} & v \end{bmatrix} \succ 0, \quad x, t, v \in \mathbb{R}, \quad (13)$$

which is equivalent to that

$$\Leftrightarrow x > 0, xt > |y|^2, \det(X) > 0,$$

such that

$$N_iXN_j = N_jXN_i, \\ N_iXN_j^* = N_j^*XN_i,$$

$1 \leq i < j \leq 3$ . By directly computing, with the help of the expansion  $y = \operatorname{Re}(y) + i\operatorname{Im}(y)$  and similarly to  $u, z$ , the linear system above (in  $X$ ) is equivalent to

$$v = x, \quad t = x - y + z, \quad u = y = \bar{y}, \quad z = \bar{z}.$$

We then pick  $x = 3, z = 2, y = u = 0, t = 5$  and then

$$X = \begin{bmatrix} 3 & 0 & 2 \\ 0 & 5 & 0 \\ 2 & 0 & 3 \end{bmatrix} \succ 0$$

makes  $X^{\frac{1}{2}}N_1X^{\frac{1}{2}}, X^{\frac{1}{2}}N_2X^{\frac{1}{2}}, X^{\frac{1}{2}}N_3X^{\frac{1}{2}}, X^{\frac{1}{2}}N_1^*X^{\frac{1}{2}}, X^{\frac{1}{2}}N_2^*X^{\frac{1}{2}}, X^{\frac{1}{2}}N_3^*X^{\frac{1}{2}}$  to be commuting by Theorem 11. Thus three initial matrices are SDC on  $\mathbb{R}$ , and so are they on  $\mathbb{C}$ .

We will see Example 3 showing the numerical experiment of computing a square root of  $X$  and a non-singular for  $*$ -SDC  $N_1, N_2$  and  $N_3$ .  $\diamond$

### 3. NUMERICAL TESTS

In this section, we perform some numerical tests illustrating our main algorithm implemented in MATLAB R2022a running on a PC with Intel Core i3 CPU 3.3GHz, 8GB RAM, Windows 10 x64 operating system.

It is well known that a matrix  $N$  is normal if and only if it can be written as  $N = A + iB$  with  $A^* = A, B^* = -B$  and  $AB = BA$ . Notice furthermore that  $A = A^*$  has only real eigenvalues, while  $B$  is skew-Hermitian, and hence its eigenvalues are all purely imaginary. As an existing result,<sup>5</sup>  $A$  and  $B$  are  $*$ -SDC by a unitary matrix. This leads us to set up a collection of normal matrices that are for sure  $*$ -SDC as follows. Fix a unitary matrix  $Q$ , and pick  $m$  diagonal matrices  $D_i$  whose diagonal elements are real in  $(1, 1)$ , and  $m$  diagonal ones  $S_i$  whose diagonal elements are purely imaginary in  $(1, 1)$ . Then the corresponding normal matrices are constructed as

$$N_i = Q(D_i + iS_i)Q^*, \quad i = 1, \dots, m,$$

which are  $*$ -SDC by  $Q$ . The first stage of Algorithm 3.2 is implemented with the CVX toolbox [19] calling SDPT3 version 4.0 [36] that solves the following semidefinite program:

$$\min\{s \mid X \succeq 0, s \geq \varepsilon, N_iXN_j = N_jXN_i, \\ N_iXN_j^* = N_j^*XN_i, 1 \leq i < j \leq m\}, \quad (14)$$

where the tolerance  $\varepsilon > 0$  is given. We then exploit the MATLAB function `sqrtn.m`, which executes the algorithm proposed by Deadman and collaborators,<sup>22</sup> to compute the square root  $Q$  of  $X$ . For the second stage, we thank the works of Mendl for executing the Jacobi-like algorithm.<sup>15</sup> In our experiment, we pick  $\varepsilon$  as the floating-point relative accuracy  $\operatorname{eps}(\frac{3}{2})$  of MATLAB for detecting the SDC property as in (12), while we keep their tolerance for the last stage to be  $\operatorname{eps}$  to the power of  $\frac{3}{2}$ .<sup>15</sup> We have performed the tests with the collections of at most 20 normal matrices (of common sizes 5, 10, ..., 30, respectively). All experiments give the backward errors approximately bounded above by  $10^{-8}$ .

*Example 3.* We continue Example 2 with finding a non-singular matrix  $U$  that  $*$ -SDC  $N_1, N_2$  and  $N_3$ . We first numerically compute the square root of  $X$  as

$$X^{\frac{1}{2}} \simeq \begin{bmatrix} 1.6180 & 0 & 0.6180 \\ 0 & 2.2361 & 0 \\ 0.6180 & 0 & 1.6180 \end{bmatrix}.$$

Noticing that  $N_1, N_2$  are real symmetric and  $N_3$  is complex symmetric. Furthermore, for a nonsingular matrix  $P$ ,  $P^*N_3P$  does not need to be normal. So, we cannot apply the extended Jacobi-like algorithm.<sup>15</sup> However, we can apply the SDP-SDC method to the matrices  $X^{\frac{1}{2}}N_1X^{\frac{1}{2}}, X^{\frac{1}{2}}N_2X^{\frac{1}{2}}, X^{\frac{1}{2}}\frac{1}{2}(N_3+N_3^*)X^{\frac{1}{2}}$  and  $X^{\frac{1}{2}}\frac{1}{2}(N_3^*-N_3)X^{\frac{1}{2}}$ ,<sup>5</sup> which are all Hermitian and are commuting, to obtain the nonsingular matrix

$$V \simeq \begin{bmatrix} -0.4082 & -0.7071 & 0.5774 \\ 0.8165 & 0 & 0.5774 \\ -0.4082 & 0.7071 & 0.5774 \end{bmatrix}$$

that simultaneously diagonalizes the latter matrices above. Finally, a nonsingular that simultaneously diagonalizes the initial matrices  $N_1, N_2, N_3$  is

$$U = X^{\frac{1}{2}}V \simeq \begin{bmatrix} -0.9129 & -0.7071 & 1.2910 \\ 1.8257 & 0 & 1.2910 \\ -0.9129 & 0.7071 & 1.2910 \end{bmatrix},$$

where

$$\begin{aligned} U^*N_1U &\simeq \text{diag}(0, 0, 15), \\ U^*N_2U &\simeq \text{diag}(-10, 0, 5), \\ U^*N_3U &\simeq \text{diag}(30i, 2i, 15i). \end{aligned}$$

#### 4. CONCLUSION

We have provided a sufficient and necessary condition for a finite family of normal matrices to be simultaneously diagonalizable via  $*$ -congruence. A corresponding MATLAB package has been developed, and some numerical tests have also been performed.

#### Acknowledgment

*This study is conducted within the framework of science and technology projects at the institutional level of Quy Nhon University under the project code T2024.830.01.*

#### REFERENCES

1. L. D. Lathauwer. A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization, *SIAM Journal on Matrix Analysis and Applications*, **2006**, 28(3), 642–666.
2. J. B. H. Urruty. Potpourri of conjectures and open questions in nonlinear analysis and optimization, *SIAM Review*, **2007**, 49(2), 255–273.
3. R. Jiang, D. Li. Simultaneous diagonalization

of matrices and its applications in quadratically constrained quadratic programming, *SIAM Journal on Optimization*, **2016**, 26(3), 1649–1668.

4. R. A. Horn, C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 2013.
5. T. H. Le, T. N. Nguyen. Simultaneous diagonalization via congruence of hermitian matrices: some equivalent conditions and a numerical solution, *SIAM Journal on Matrix Analysis and Applications*, **2022**, 43, 882–911.
6. M. D. Bustamante, P. Mellon, M. V. Velasco. Solving the problem of simultaneous diagonalisation via congruence, *SIAM Journal on Matrix Analysis and Applications*, **2020**, 41, 1616–1629.
7. K. Weierstrass. Zur theorie der quadratischen und bilinearen formen, *Monatsbericht der Berliner Akademie der Wissenschaften*, **1868**, 19–44.
8. A. A. Albert. A quadratic form problem in the calculus of variations, *Bulletin of the American Mathematical Society*, **1938**, 44, 250–253.
9. P. Finsler. Über das vorkommen der niter und semide niter formen in scharen quadratischer formen, *Commentarii Mathematici Helvetici*, **1937**, 9, 188–192.
10. M. R. Hestenes, E. J. McShane. A theorem on quadratic forms and its application in the calculus of variations, *Transactions of the American Mathematical Society*, **1940**, 47(3), 501–512.
11. J. J. Moré. Generalization of the trust region problem, *Optimization methods and software*, **1993**, 2, 189–209.
12. T. K. Pong, H. Wolkowicz. The generalized trust region subproblem, *Computational Optimization and Applications*, **2014**, 58, 273–322.
13. R. A. Horn, C. R. Johnson. *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1991.
14. A. B. Gerstner, R. Byers, V. Mehrmann. Numerical methods for simultaneous diagonalization, *SIAM Journal on Matrix Analysis and Applications*, **1993**, 14(4), 927–949.
15. C. Mendl., simdiag.m. MATLAB central file exchange. Available at <https://www.mathworks.com/matlabcentral/fileexchange/46794-simdiag-m>, 2020.
16. B. D. Sutton. Simultaneous diagonalization of nearly commuting Hermitian matrices: do-one-then-do-the-other, *IMA Journal of Numerical Analysis*, **2024**, 44(2), 1061–1089.

17. J. F. Watters. Simultaneous quasi-diagonalization of normal matrices, *Linear Algebra and its Applications*, **1974**, 9, 103–117.
18. G. Pastuszak, T. Kamizawa, A. Jamiołkowski. On a criterion for simultaneous block-diagonalization of normal matrices, *Open Systems and Information Dynamics*, **2016**, 23(1), 1650003.
19. M. Grant, S. P. Boyd. CVX: Matlab software for disciplined convex programming, version 1.21. Apr. 2011. <https://cvxr.com/cvx/>.
20. R. Grone, C. R. Johnson, E. M. Sa, H. Wolkowicz. Normal matrices, *Linear Algebra and its Applications*, **1987**, 87, 213–225.
21. E. Deadman, N. J. Higham, R. Ralha. *Blocked Schur algorithms for computing the matrix square root*, Lecture Notes in Computer Science, Springer-Verlag, Heidelberg, 2013.

## Appendix

*Proof of Lemma 10.* i) By applying (3) to  $M$  and  $N$ , one has

$$\begin{aligned}
 MXN &= (M^h XN^h - M^s XN^s) \\
 &\quad + i(M^h XN^s + M^s XN^h), \\
 NXM &= (N^h XM^h - N^s XM^s) \\
 &\quad + i(N^s XM^h + N^h XM^s), \\
 MXN^* &= (M^h XN^h + M^s XN^s) \\
 &\quad - i(M^h XN^s - M^s XN^h), \\
 N^*XM &= (N^h XM^h + N^s XM^s) \\
 &\quad - i(N^s XM^h - N^h XM^s).
 \end{aligned}$$

Substituting the above identities into (4)-(5) one obtains that

$$\begin{aligned}
 M^h XN^h - M^s XN^s &= N^h XM^h - N^s XM^s, \\
 M^h XN^s + M^s XN^h &= N^s XM^h + N^h XM^s, \\
 M^h XN^h + M^s XN^s &= N^h XM^h + N^s XM^s, \\
 M^h XN^s - M^s XN^h &= N^s XM^h - N^h XM^s.
 \end{aligned}$$

Adding side-by-side the first and the third (resp., the second and the fourth) equations one has

$$M^h XN^h = N^h XM^h, \quad M^h XN^s = N^s XM^h.$$

Subtracting side-by-side the first and the third (resp., the second and the fourth) equations one has

$$M^s XN^s = N^s XM^s, \quad M^s XN^h = N^h XM^s,$$

Conversely, from (2), the identities (6)-(9) are equivalent to

$$\begin{aligned}
 (M + M^*)X(N + N^*) &= (N + N^*)X(M + M^*), \\
 (M + M^*)X(N - N^*) &= (N - N^*)X(M + M^*), \\
 (M - M^*)X(N + N^*) &= (N + N^*)X(M - M^*), \\
 (M - M^*)X(N - N^*) &= (N - N^*)X(M - M^*),
 \end{aligned}$$

respectively. Expanding the above identities leads to that

$$\begin{aligned}
 MXN + MXN^* + M^*XN + M^*XN^* &= \\
 NXM + NXM^* + N^*XM + N^*XM^* &, \\
 MXN - MXN^* + M^*XN - M^*XN^* &= \\
 NXM + NXM^* - N^*XM - N^*XM^* &, \\
 MXN + MXN^* - M^*XN - M^*XN^* &= \\
 NXM - NXM^* + N^*XM - N^*XM^* &, \\
 MXN - MXN^* - M^*XN + M^*XN^* &= \\
 NXM - NXM^* - N^*XM + N^*XM^* &.
 \end{aligned}$$

By adding side-by-side the above identities, we have

$$MXN = NXM.$$

Similarly, by adding the first and the third, then subtracting the second and the fourth identities, side-by-side, we additionally obtain

$$MXN^* = N^*XM.$$

ii) This is an immediate consequence of the first part with noting that  $X^* = X$  and

$$\begin{aligned}
 M + M^* &= U(\Lambda_M + \overline{\Lambda_M})U^*, \\
 M - M^* &= U(\Lambda_M - \overline{\Lambda_M})U^*, \\
 N + N^* &= V(\Lambda_N + \overline{\Lambda_N})V^*, \\
 N - N^* &= V(\Lambda_N - \overline{\Lambda_N})V^*,
 \end{aligned}$$

where  $M = U\Lambda_M U^*$ ,  $N = V\Lambda_N V^*$  are eigenvalue decomposition of the normal matrices  $M$  and  $N$  ( $\Lambda_M, \Lambda_N$  are complex diagonal and  $U, V$  are unitary matrices).  $\square$



© 2025 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).



# Phân tích toàn diện phổ của berberin, hợp chất chính trong bài thuốc dân gian Việt Nam “Đại tràng hoàn Bà Giăng”

Nguyễn Hoàng Sa\*

*Trường Đại học Khánh Hòa, Việt Nam*

*Ngày nhận bài: 29/09/2024; Ngày sửa bài: 09/01/2025;  
Ngày nhận đăng: 20/01/2025; Ngày xuất bản: 28/02/2025*

## TÓM TẮT

"Đại tràng hoàn Bà Giăng", một bài thuốc dân gian nổi tiếng của lương y Bà Giăng, đã được sử dụng phổ biến ở Việt Nam hơn một thế kỷ qua để điều trị viêm loét đại tràng và tiêu chảy. Trong nghiên cứu này, các mẫu thảo dược khô được chiết xuất bằng ethanol 70% để thu được dịch chiết (DTBG). Berberin, một hợp chất chính trong DTBG, đã được tách chiết bằng phương pháp sắc ký cột silica gel. Cấu trúc của hợp chất này được xác định qua phân tích phổ và so sánh với dữ liệu phổ đã công bố. Báo cáo cung cấp bình luận chi tiết về phổ của berberin, sử dụng các kỹ thuật phổ UV-Vis, FT-IR, EI-MS, 1D-NMR và 2D-NMR.

**Từ khóa:** Berberin, protoberberine alkaloid, phân tích phổ.

\*Tác giả liên hệ chính.

Email: [nhoangsa@gmail.com](mailto:nhoangsa@gmail.com)

# Complete spectral analysis of berberine, a key compound in Vietnamese traditional remedy “Đại tràng hoàn Bà Giăng”

Nguyen Hoang Sa\*

Khanh Hoa University, Vietnam

Received: 29/09/2024; Revised: 09/01/2025;

Accepted: 20/01/2025; Published: 28/02/2025

## ABSTRACT

The traditional herbal remedy “Đại tràng hoàn Bà Giăng,” attributed to the renowned herbalist Bà Giăng, has been widely used in Vietnam for over a century, particularly for the treatment of ulcerative colitis and diarrhea. In this study, dried herbal materials were extracted using 70% ethanol, resulting in an extract known as DTBG. Berberine, a key alkaloid in DTBG, was isolated using silica gel column chromatography. Its structure was elucidated through comprehensive spectral analyses, including UV-Vis, FT-IR, EI-MS, and 1D/2D-NMR techniques. This report provides a detailed discussion of berberine’s spectral data, contributing valuable insights into the chemistry of this compound.

**Keywords:** Berberine, protoberberine alkaloid, spectral analyses.

## 1. INTRODUCTION

Vietnamese traditional medicine (VTM) has been utilized for centuries, relying heavily on herbal remedies that are perceived as safe, natural, and without significant side effects. Among the various diseases addressed by VTM, ulcerative colitis, marked by symptoms such as abdominal pain, diarrhea, rectal bleeding, and weight loss, stands out as a major health concern.<sup>1</sup> Since the early 20<sup>th</sup> century, the herbal remedy “Đại tràng hoàn Bà Giăng” (BG) has played a pivotal role in treating inflammatory bowel disease and diarrhea in traditional Vietnamese medicine.

The efficacy of BG is attributed to its rich composition of bioactive compounds, particularly alkaloids. Alkaloids are organic

compounds containing nitrogen, and they are recognized for their diverse chemical structures and broad medicinal properties. The biosynthesis of these compounds can vary significantly across plant species, contributing to their complex pharmacological effects.

BG is a traditional herbal formula composed of a variety of medicinal plants, including *Radix glycyrrhizae* (Cam thảo), *Atractylodes macrocephala* (Bạch truật), *Rhizoma coptidis* (Hoàng liên), *Saussurea lappa* (Mộc hương), *Poria cocos* (Bạch linh), *Codonopsis pilosulae* (Đảng sâm), a fermentative herb mixture (Thần khúc), *Myristica fragrans* (Nhục đậu khấu), *Citrus reticulata* (Trần bì), *Hordeum vulgare* (Mạch nha), *Malus doumeri* (Sơn trà), *Dioscorea persimilis* (Hoài sơn), and

\*Corresponding author.

Email: [nhoangsa@gmail.com](mailto:nhoangsa@gmail.com)

*Amomum villosum* (Sa nhân). Each of these plants contributes to the therapeutic potential of BG.

Several of these ingredients have shown remarkable bioactivities. For instance, *Rhizoma coptidis* (Hoàng liên) exhibits cytotoxic effects against a variety of cancer cell lines and possesses anti-proliferative properties, particularly on human esophageal cancer cells.<sup>2</sup> Its primary alkaloid constituents have also demonstrated inhibitory effects on arthritis.<sup>3</sup> Furthermore, *Atractylodes macrocephala* (Bạch truật) has been reported to inhibit nitric oxide (NO) production in activated macrophages, exhibit neuroprotective and anti-inflammatory effects, and induce apoptosis in human leukemia cells.<sup>5-8</sup>

In a recent study, we investigated the effects of the ethanol extract from BG on acute myeloid leukemia (OCI-AML) cells. The extract exhibited significant anti-cancer activity while showing no observable toxicity in acute and subchronic toxicity studies conducted on mice.<sup>9</sup>

In this study, dried powders of BG's herbal components were extracted using 70% ethanol, resulting in an ethanol extract referred to as DTBG. Berberine, a prominent alkaloid isolated from DTBG, was subjected to detailed structural characterization. This study provides a comprehensive spectral analysis of berberine, utilizing UV-Vis, FT-IR, EI-MS, and NMR techniques, including both one-dimensional and two-dimensional <sup>1</sup>H/<sup>13</sup>C NMR experiments in DMSO-d<sub>6</sub>.

## 2. METHODS

### 2.1. General

The EI-MS spectrum was recorded using an Agilent mass spectrometer. Proton nuclear magnetic resonance (<sup>1</sup>H-NMR) at 500.13 MHz and carbon-13 nuclear magnetic resonance (<sup>13</sup>C-NMR) at 125.77 MHz were performed on a Bruker Avance 500 NMR spectrometer at room temperature. Chemical shifts are reported in delta (δ) parts per million (ppm). Coupling constants (*J*) were provided in Hertz (Hz).

### 2.2. HPLC-DAD detection of the concentrations of the DTBG

The concentration of compounds in the DTBG extract was determined using high-performance liquid chromatography with diode-array detection (HPLC-DAD). An Eclipse XDB-C<sub>18</sub> column (5 μm, 4.6 mm × 250 mm) was used at 30 °C. The mobile phase consisted of solvent A (water with 0.1% formic acid) and solvent B (acetonitrile with 0.1% trifluoroacetic acid). A gradient elution was employed, starting with 35% solvent B and increasing to 65% over 45 minutes at a flow rate of 0.5 mL/min. The detection process lasted 60 minutes at a flow rate of 0.25 mL/min, with monitoring at 290 nm and 294 nm. The injection volume was 5 μL. Berberine was detected with a retention time (*R<sub>t</sub>*) of 13.37 minutes in the HPLC chromatogram.

### 2.3. Herb materials, extraction, isolation, and purification

The herb materials used in this study were supplied by Ba Giang Pharmaceutical Company, Thanh Hoa, Vietnam. A total of 1 kg of dried "Đại tràng hoàn Bà Giằng" (BG) powder was subjected to three overnight extractions at 60 - 80°C using 70% ethanol. The resulting crude extract (DTBG) was obtained by evaporating the solvent under reduced pressure, yielding 220 g of extract. The DTBG was fractionated using silica gel column chromatography, eluted with acetone-methanol (100:5), producing eleven fractions. Fraction 5 (3.55 g) was further subjected to silica gel chromatography using acetone-methanol mixtures of increasing polarity, resulting in five subfractions. Subfraction 4 (1.27 g) underwent further purification using acetone-methanol (1:1) as the eluent. Additional purification over silica gel with acetone-methanol-acetic acid (30:30:1) produced 722 mg of raw berberine.

The raw berberine was dissolved in 10 mL of methanol, followed by the addition of five drops of concentrated hydrochloric acid while gently heating the mixture in a water bath. The resulting deep yellow solution was filtered

through a pre-warmed funnel into a pre-warmed flask. After cooling to room temperature, the solution was placed in a refrigerator at +4 °C overnight. Yellow crystals of berberine chloride formed and were collected by filtration, then washed with ice-cold methanol. The yield of berberine chloride was 565 mg, with a melting point of 195 - 197 °C.

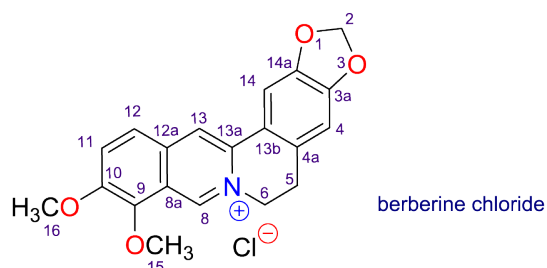
Thin-layer chromatography (TLC) analysis of berberine was performed using *n*-butanol/cold acetic acid/water (12:3:4 v/v) as the eluent. Berberine appeared as a yellow spot with an  $R_f$  value of 0.53, exhibiting strong yellow fluorescence under UV light at 366 nm.

Berberine chloride as the ionic form  $[M^+Cl^-]$  was isolated as a yellow powder, with the following key spectroscopic data: EI-MS  $m/z$  336  $[M]^+$  (rel.% = 2.6),  $^1H$  NMR (DMSO- $d_6$ , 500.13 MHz):  $\delta_H$  9.92 (1H, s, H-8), 8.97 (1H, s, H-13), 8.21 (1H, d,  $J$  = 9.0, H-11), 8.01 (1H, d,  $J$  = 9.0, H-12), 7.81 (1H, s, H-14), 7.12 (1H, s, H-4), 6.15 (2H, s, H-2), 4.95 (1H, t,  $J$  = 6.5, H-6), 4.12 (3H, s, 9-OCH<sub>3</sub>), 4.10 (3H, s, 10-OCH<sub>3</sub>), 3.20 (1H, t,  $J$  = 6.5, H-5);

$^{13}C$  NMR (DMSO- $d_6$ , 125.77 MHz):  $\delta_C$  149.9 (C-10), 149.3 (C-3a), 147.2 (C-14a), 145.0 (C-8), 143.1 (C-9), 137.0 (C-13a), 132.5 (C-12a), 130.2 (C-4a), 126.2 (C-11), 123.0 (C-12), 120.9 (C-13b), 120.0 (C-8a), 119.7 (C-13), 107.8 (C-4), 104.8 (C-14), 101.5 (C-2), 61.5 (C-15), 56.4 (C-16), 54.7 (C-6), 25.8 (C-5).

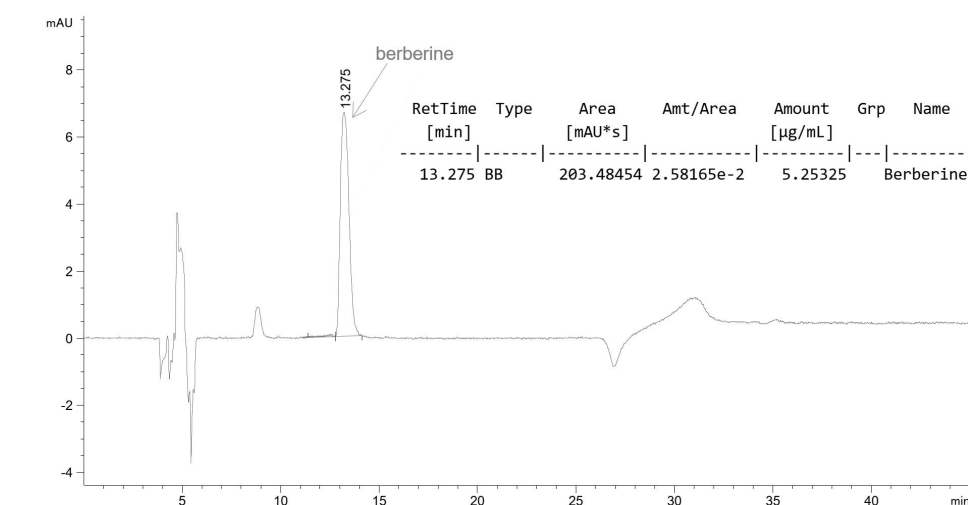
### 3. RESULTS AND DISCUSSION

The isolated compound's structure was identified as berberine by comparing its spectral data with reference literature.<sup>10,11</sup> Berberine belongs to the class of protoberberine alkaloids, which are part of the isoquinoline alkaloid group. These alkaloids often occur as protoberberinium salts, such as berberine or berberine chloride, or as tetrahydroprotoberberines. Berberine is well-known for its bright green-yellow fluorescence, which is utilized in histological staining to visualize heparin in mast cells under fluorescence microscopy.

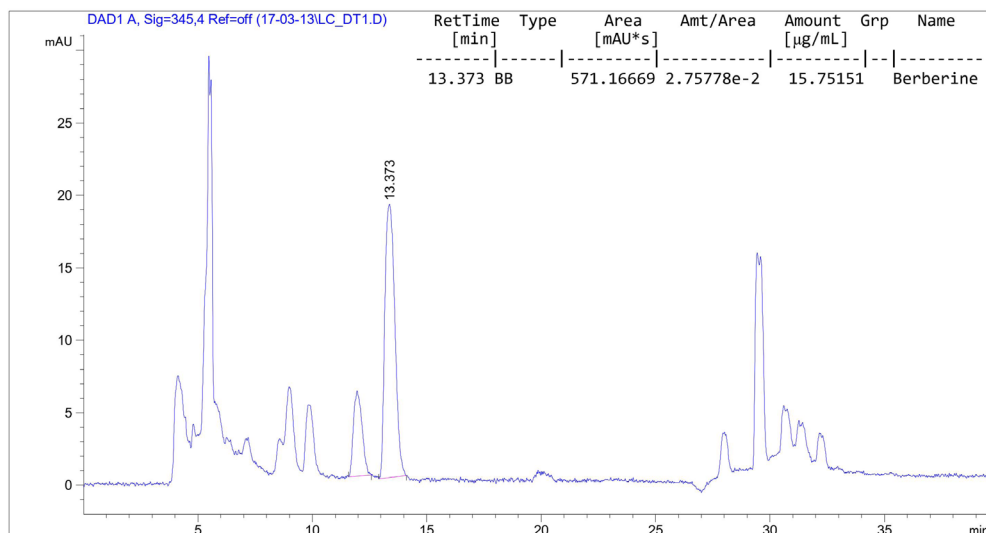


**Figure 1.** Structures of berberine chloride.

Figure 2 presents the HPLC analysis of the purified berberine from DTBG. The purity of the isolated compound was determined to be 98.7%. HPLC analysis was also used to quantify key marker compounds, including berberine, within the DTBG extract. The analysis revealed that berberine content in DTBG was approximately 0.63%, as shown in Figure 3.



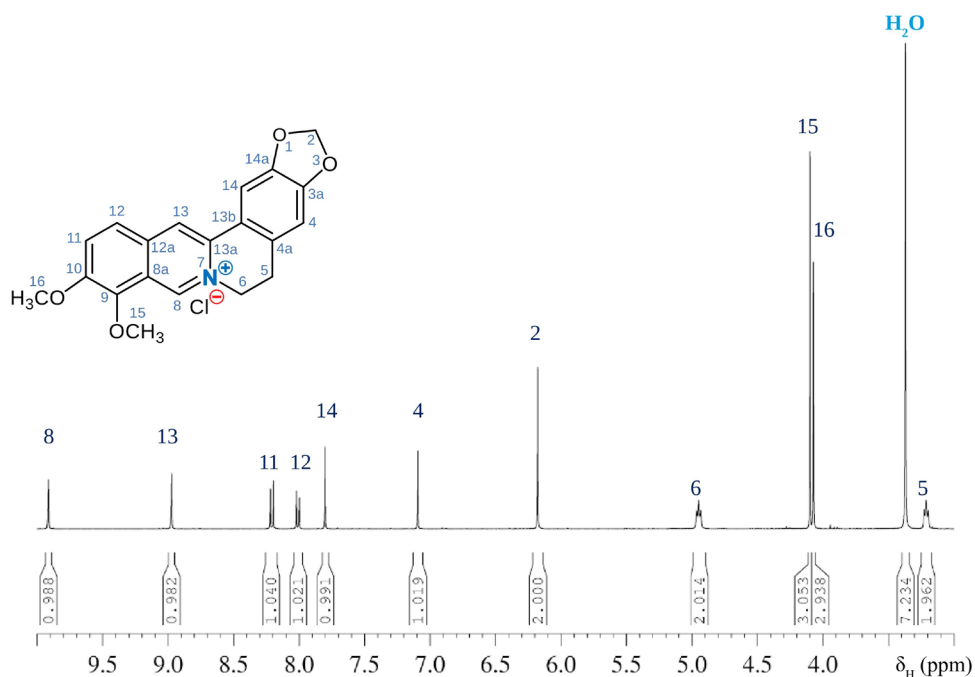
**Figure 2.** HPLC analysis of berberine, with the mobile phase composed of ACN-H<sub>2</sub>O:0.1% TFA in linear gradient mode over 0-45 minutes, at a flow rate of 1.5 mL/min.



**Figure 3.** HPLC chromatogram of berberine from DTBG, using the same mobile phase and flow rate conditions.

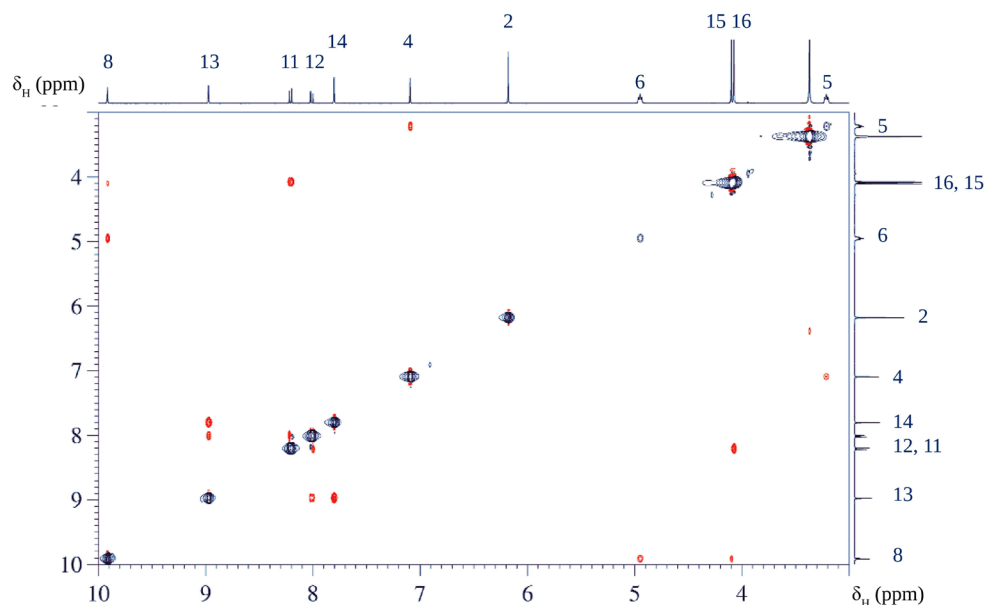
The most deshielded proton signal at  $\delta_H$  9.92 is assigned to H-8 due to its proximity to the positively charged nitrogen atom (N). Similarly, the signal at  $\delta_H$  8.97 is attributed to H-13 in the same aromatic ring (Figure 4). The proton signals at  $\delta_H$  8.21 and 8.01 form an AB-spin system with a coupling constant ( $J$ ) of 9.0 Hz, which is characteristic of aromatic protons H-11 and H-12. The NOESY spectrum (Figure 5) helps differentiate these signals, with  $\delta_H$  8.21 showing a nuclear Overhauser effect

(NOE) contact with a methoxy group at  $\delta_H$  4.10, thereby confirming its assignment to H-11. By deduction, the methoxy signal corresponds to H-16. The singlets at  $\delta_H$  7.81 and 7.12 are assigned to H-14 and H-4, respectively. NOESY analysis confirms this assignment through an NOE contact between H-13 and H-14. The singlet at  $\delta_H$  6.15, corresponding to two protons, is assigned to H-2 based on its chemical shift. Multiple signals at  $\delta_H$  4.95 and  $\delta_H$  3.20 are attributed to methylene protons H-6 and H-5, respectively.



**Figure 4.**  $^1\text{H}$ -NMR of berberine chloride in  $\text{DMSO-d}_6$ .

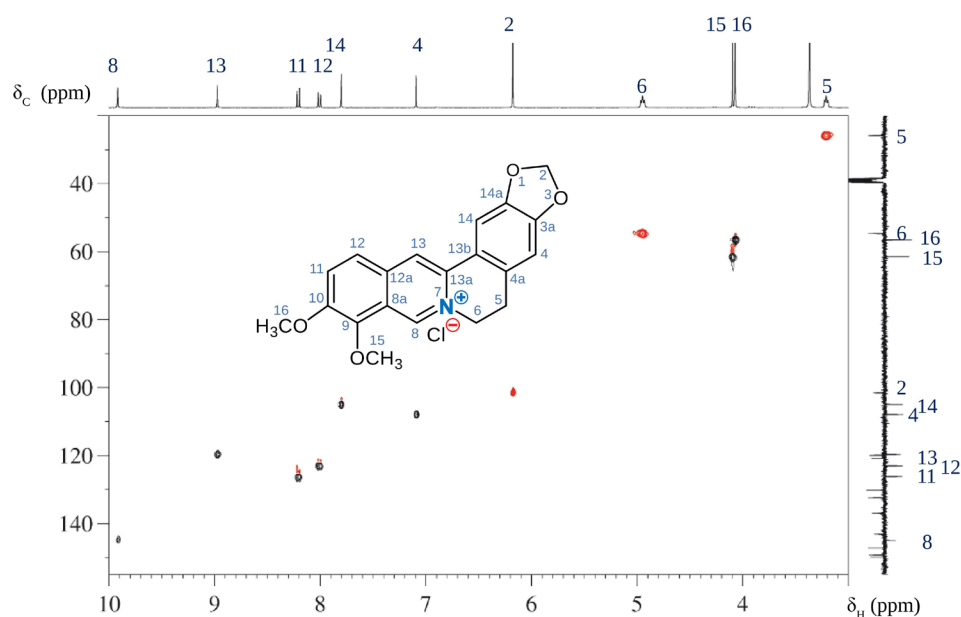




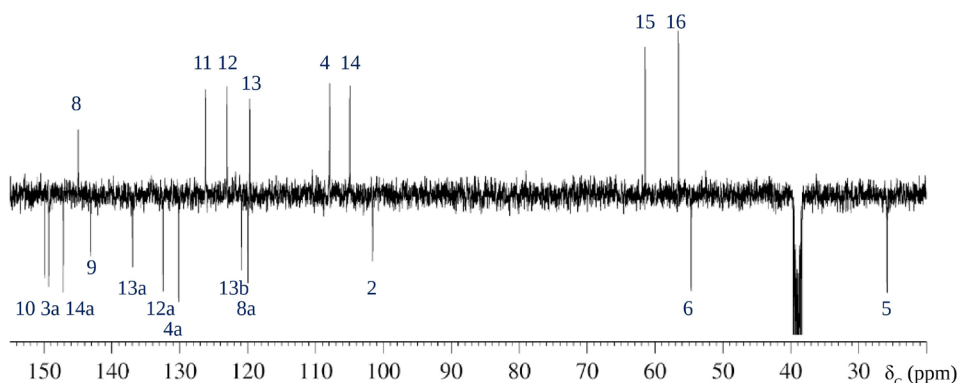
**Figure 5.** NOESY of berberine chloride.

Having fully assigned the proton signals, the carbon atoms carrying these protons can be identified using the HSQC spectrum (Figure 6). However,  $^{13}\text{C}$  chemical shifts are influenced by the solvent, potentially causing minor discrepancies in assignments. The values recorded in DMSO- $d_6$  (Figure 7) are consistent with those reported by Blasko et al.<sup>10</sup> Quaternary carbon atoms, which lack direct proton attachments, are more challenging to assign and often require HMBC analysis for confirmation.

From H-8, we expect five HMBC correlations (Figure 8 and Figure 9): three-bond correlations with C-6, C-13a, C-12a, and C-9, and a two-bond correlation with C-8a. C-9 should also correlate with the methoxy protons (H-15), confirming that the signal at  $\delta_{\text{C}}$  143.1 is assigned to C-9. Similarly, H-11 correlates with C-12a, leading to the assignment of the signal at  $\delta_{\text{C}}$  132.5 for C-12a. C-13a is correlated with H-14, assigning the signal at  $\delta_{\text{C}}$  137.0 to C-13a.



**Figure 6.** HSQC of berberine chloride.

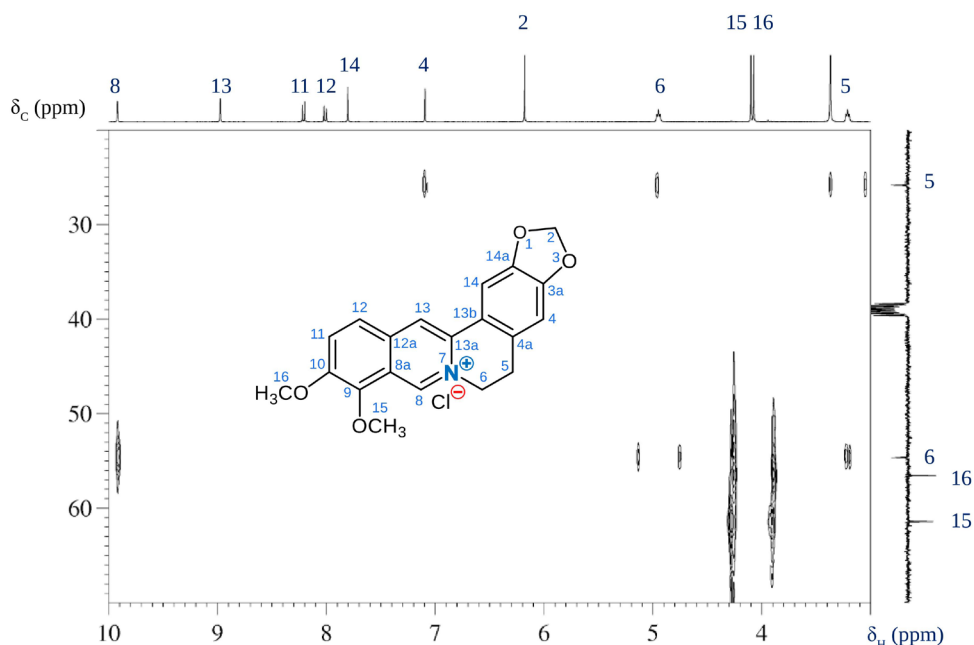


**Figure 7.** APT  $^{13}\text{C}$ -NMR of berberine chloride.

C-8a correlates with both H-12 and H-13, supporting the assignment of the signal at  $\delta_{\text{C}}$  120.0 to C-8a. All five HMBC correlations for H-8 are now assigned. In the linear annulated aromatic ring, the only remaining unassigned quaternary carbon is C-10, which can be identified via correlations with H-16 and H-12, giving the signal at  $\delta_{\text{C}}$  149.9. The remaining quaternary carbon signals are from C-4a, C-13b, C-14a, and C-3a, all part of the isolated aromatic ring. C-3a correlates with H-2 and H-14, corresponding to the signal at  $\delta_{\text{C}}$  149.3. C-14a, linked to H-2 and H-4, corresponds to the signal at  $\delta_{\text{C}}$  147.2. C-13b correlates with H-13 and H-4, producing

the signal at  $\delta_{\text{C}}$  120.9, while C-4a is assigned via correlations with H-14 and H-6 to  $\delta_{\text{C}}$  130.2.

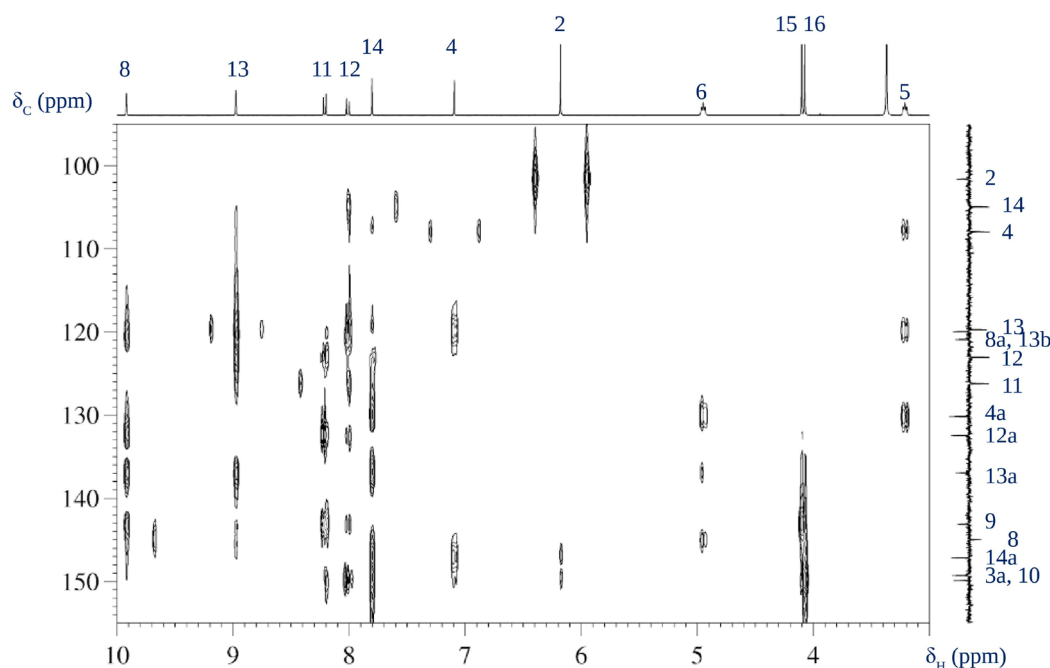
The UV-Vis spectrum of berberine chloride (Figure 10) exhibits three prominent absorption peaks at 227 nm, 252 nm, and 354 nm, with logarithmic molar absorptivity ( $\lg \epsilon$ ) values of 0.750, 0.745, and 0.740, respectively. Additionally, a weaker absorption band is observed at 423 nm with a  $\lg \epsilon$  value of 0.15. These absorption bands, known as B bands, correspond to forbidden  $\pi$ -electron transitions ( $^1\text{A}_{1g} \rightarrow ^1\text{B}_{2u}$ ) commonly associated with aromatic systems containing benzene rings or six-membered heteroaromatic rings.



**Figure 8.** HBMC of berberine chloride in the aliphatic  $^{13}\text{C}$  region.

The absorption observed in the blue region of the visible spectrum contributes to the compound's characteristic yellow coloration. The conjugated  $\pi$ -system of berberine chloride experiences a bathochromic shift, attributed

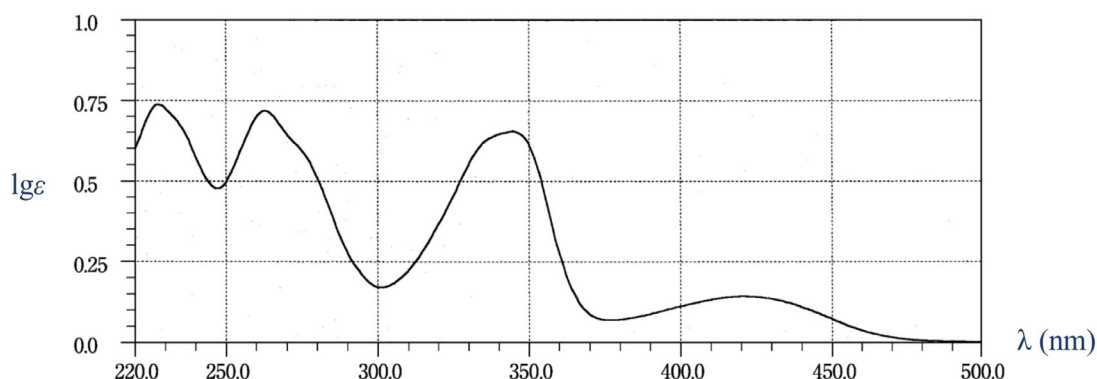
to the electronic effects of four auxochromic oxygen substituents and the positively charged quaternary nitrogen atom. This shift highlights the impact of the molecule's structural features on its electronic transitions and optical properties.



**Figure 9.** HBMN of berberine chloride in the aromatic  $^{13}\text{C}$  region.

The crystallization of berberine chloride with four water molecules explains the strong O—H vibrational band observed at 3500-3200  $\text{cm}^{-1}$  in the FT-IR spectrum (Figure 11). The C—H stretching region reveals the presence

of both  $sp^3$  and  $sp^2$  hybridized carbon atoms, while the broad band at 1630  $\text{cm}^{-1}$  is attributed to the C=N vibration. The sharp bands observed between 1600 and 1450  $\text{cm}^{-1}$  correspond to the C=C stretching of the aromatic rings.



**Figure 10.** UV-Vis of berberine chloride.

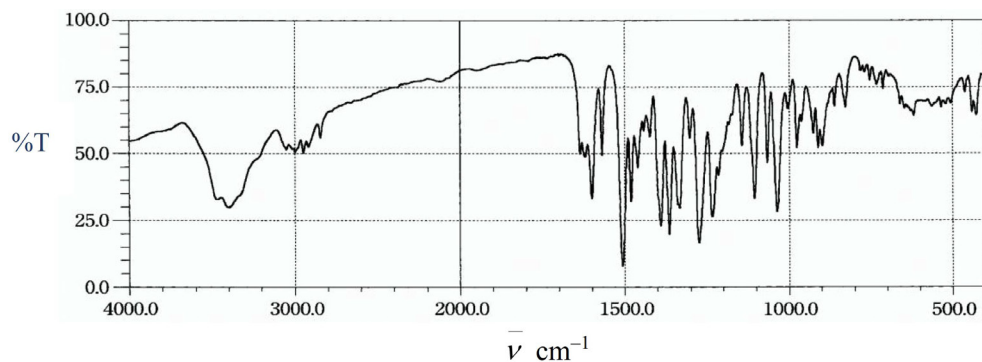


Figure 11. FT-IR of berberine chloride in KBr.

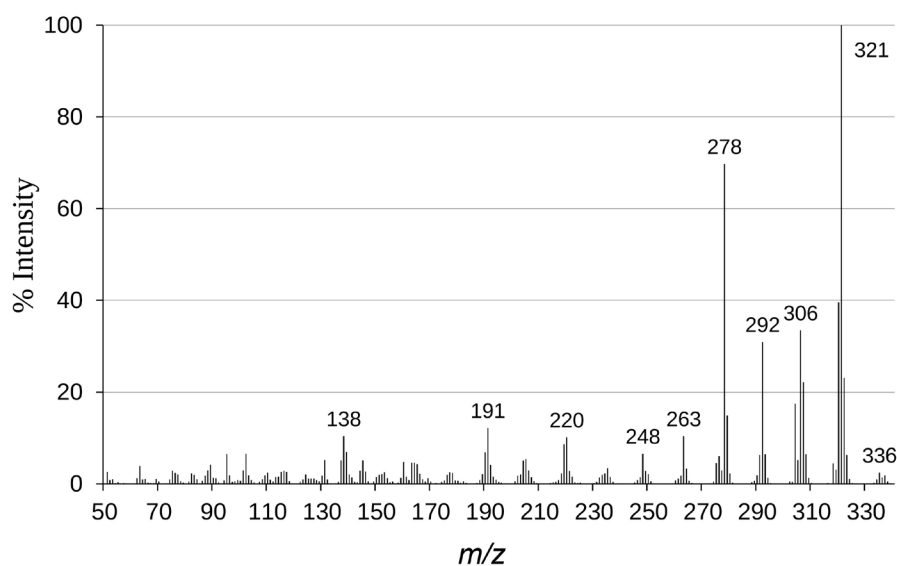


Figure 12. EI-MS of berberine chloride.

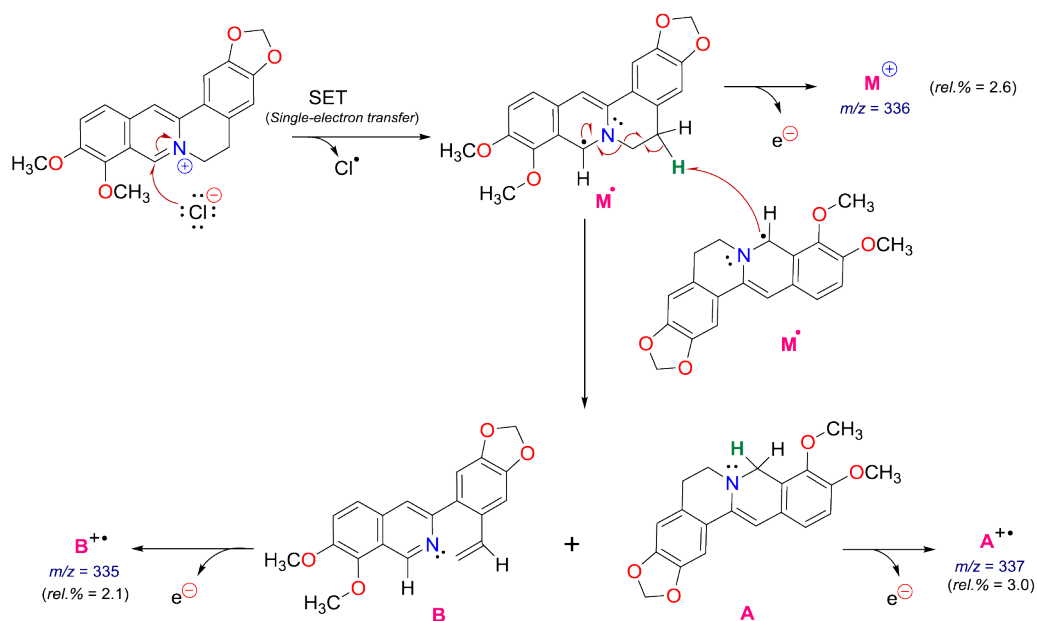
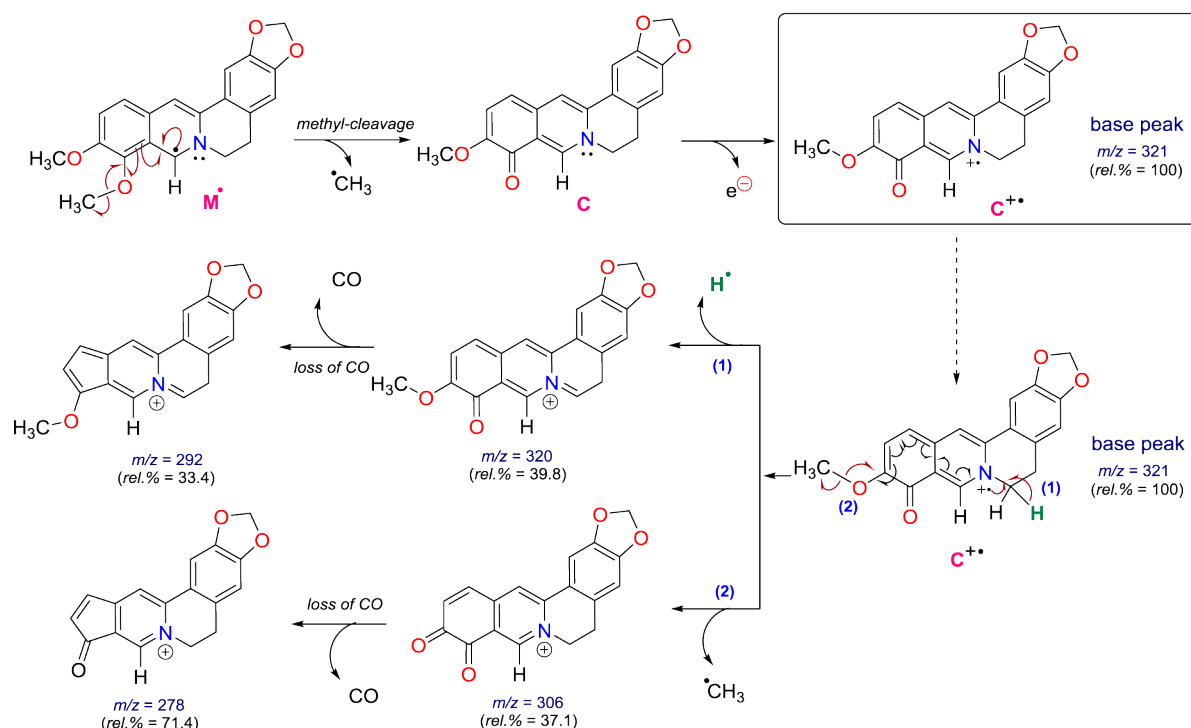


Figure 13. Electron transfer mechanism in the ionization source,  $M^+$  ( $m/z$  336),  $A^+$  ( $m/z$  337), and  $B^+$  ( $m/z$  335).

The mass spectrum (Figure 12) reveals the formation of several ions through an electron transfer process in the inlet system. The cation  $M^+$  ( $m/z$  336) generated from berberine chloride undergoes radical formation ( $M^\bullet$ ), followed by the loss of an electron, resulting in the ion detected at  $m/z$  336. Two additional ions at  $m/z$  335 and 337 are formed through disproportionation reactions,

producing neutral molecules A and B, which are subsequently ionized (Figure 13).

Other prominent ions at  $m/z$  320, 304, 292, and 278 are likely generated through a common cleavage mechanism. These ions, having an even number of electrons, follow similar fragmentation pathways, as outlined in Figure 14.



**Figure 14.** Cleavage of a methyl group from radical  $M^\bullet$ , leading to the formation of a neutral molecule  $C$  and its corresponding fragment ions.

#### 4. CONCLUSION

In summary, the  $^1H$ -NMR and  $^{13}C$ -NMR spectra of berberine chloride have been fully assigned. Comprehensive signal interpretation was conducted using various  $^2D$ -NMR spectroscopies: HSQC, NOESY, and HMBC. Furthermore, supplementary techniques, including UV-Vis, FT-IR, and EI-MS, were utilized, yielding precise data for the thorough structural elucidation of berberine.

#### REFERENCES

1. R. J. Xavier, D. K. Podolsky. Unravelling the pathogenesis of inflammatory bowel disease, *Nature*, **2007**, 448(7152), 427-434.
2. B. L. Ma, Y. M. Ma, R. Shi, T. M. Wang, N. Zhang, C. H. Wang, Y. Yang. Identification of the toxic constituents in *Rhizoma Coptidis*, *Journal of Ethnopharmacology*, **2010**, 128(2), 357-364.
3. N. Iizuka, K. Miyamoto, K. Okita, A. Tangoku, H. Hayashi, S. Yosino, T. Abe, T. Morioka, S. Hazama, M. Oka. Inhibitory effect of *Coptidis Rhizoma* and berberine on the proliferation of human esophageal cancer cell lines, *Cancer Letters*, **2000**, 148(1), 19-25.
4. X. Zhou, X. Lin, Y. Xiong, L. Jiang, W. Li, J. Li, L. Wu. Chondroprotective effects of palmitate on osteo arthritis in vivo and *in vitro*: a possible mechanism of inhibiting the Wnt/ $\beta$ -catenin and Hedgehog signaling pathways, *International Immunopharmacology*, **2016**, 34, 129-138.



5. C. M. Yao, X. W. Yang. Bioactivity-guided isolation of polyacetylenes with inhibitory activity against NO production in LPS-activated RAW264.7 macrophages from the rhizomes of *Atractylodes macrocephala*, *Journal of Ethnopharmacology*, **2014**, 151(1), 791-799.
6. N. Zhang, C. Liu, T. M. Sun, X. K. Ran, T. G. Kang, D. Q. Dou. Two new compounds from *Atractylodes macrocephala* with neuroprotective activity, *Journal of Asian Natural Products Research*, **2017**, 19(1), 35-41.
7. L. S. Hoang, M. H. Tran, J. S. Lee, T. Q. M. Ngo, M. H. Woo, B. S. Min. Inflammatory inhibitory activity of Sesquiterpenoids from *Atractylodes macrocephala* Rhizomes, *Pharmaceutical Society of Japan*, **2016**, 64(5), 507-511.
8. H. L. Huang, T. W. Lin, Y. L. Huang, R. L. Huang. Induction of apoptosis and differentiation by atractylenolide-1 isolated from *Atractylodes macrocephala* in human leukemia cells, *Bioorganic and Medicinal Chemistry Letters*, **2016**, 26(8), 1905-1909.
9. N. H. Sa, T. T. Thuy, T. D. Quan, D. D. Thien, N. T. Tam, L. C. Hoan, L. T. H. Nhung, H. T. N. Ni, S. Adorisio, D. V. Delfino. The alkaloids and bioactivities of ethanol extract from a traditional remedy of Vietnam, *The Natural Products Journal*, **2020**, 10(1), 20-25.
10. G. Blaskó, G. A. Cordell, S. Bhamarapravati, C. W. W. Beecher. Carbon-13 NMR assignments of Berberine and Sanguinarine, *Heterocycles*, **1988**, 27, 911-916.
11. L. Grycová, J. Dostál, R. Marek. Quaternary protoberberine alkaloids, *Phytochemistry*, **2007**, 68(2), 150-175.



© 2025 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).



# Liên kết hydrogen không cổ điển $C_{sp^2}-H...O/Se/Te$ and $O-H...Se/Te$ trong các hệ phức giữa acid formic acid và selenoformaldehyde, telluroformaldehydes

Nguyễn Tiến Trung<sup>1,\*</sup>, Nguyễn Trường An<sup>2</sup>, Nguyễn Ngọc Trí<sup>1</sup>, Vũ Thị Ngân<sup>1</sup>

<sup>1</sup>Phòng Thí nghiệm Hóa học tính toán và Mô phỏng, Khoa Khoa học Tự nhiên,  
Trường Đại học Quy Nhơn, Việt Nam

<sup>2</sup>Khoa Hóa học Tính toán, Viện Hóa lý J. Heyrovsky, Viện Hàn lâm Khoa học Séc, Cộng hòa Séc

Ngày nhận bài: 10/10/2024; Ngày sửa bài: 08/01/2025;  
Ngày nhận đăng: 20/01/2025; Ngày xuất bản: 28/02/2025

## TÓM TẮT

Trong nghiên cứu này, độ bền các phức cũng như các liên kết hydrogen và đặc trưng của chúng trong các hệ giữa acid formic với các dẫn xuất aldehyde được khảo sát một cách đầy đủ. Nhìn chung, độ bền của các phức được tăng cường đáng kể đối với các nhóm thế đẩy hoặc hút electron trong các dẫn xuất aldehyde. Các phức với dẫn xuất thế Se được tìm thấy bền hơn so với các phức Te. Các phức đối với sự thế halogen được đánh giá kém bền hơn so với các phức của dẫn xuất thế  $-CH_3$  và  $-NH_2$ . Các kết quả thu được chỉ ra độ bền đáng kể của liên kết hydrogen  $O-H...Se/Te$  trong việc làm bền các phức so với  $C-H...O$ . Các kết quả tính toán cũng cho thấy rằng độ bền các liên kết hydrogen không cổ điển giảm theo thứ tự  $O-H...Se > O-H...Te \sim C-H...O > C-H...Se/Te$ . Sự chuyển dời xanh của liên kết hydrogen  $C-H...Se$  mạnh hơn so với  $C-H...Te$ . Phân tích NBO thấy rằng xu hướng chuyển dời xanh của tần số dao động liên kết  $C_{sp^2}-H$  phụ thuộc chính vào sự giảm mật độ electron ở  $\sigma^*(C_{sp^2}-H)$  trong quá trình hình thành phức. Sự chuyển dời đỏ trong các liên kết hydrogen không cổ điển  $O-H...Se/Te$  được quyết định bởi sự tăng đáng kể mật độ electron tại  $\sigma^*(O-H)$  so với sự tăng cường  $\%s(O)$  trong quá trình tạo phức.

**Từ khóa:** Liên kết hydrogen không cổ điển  $O/C_{sp^2}-H...O/Se/Te$ ,  $Se/Te$ -aldehydes, acid formic.

\*Tác giả liên hệ chính.

Email: nguyentientrung@qnu.edu.vn

# The $C_{sp^2}$ -H...O/Se/Te and O-H...Se/Te nonconventional hydrogen bonds in the systems of formic acid with selenoformaldehydes and telluroformaldehydes

Nguyen Tien Trung<sup>1,\*</sup>, Nguyen Truong An<sup>2</sup>, Nguyen Ngoc Tri<sup>1</sup>, Vu Thi Ngan<sup>1</sup>

<sup>1</sup>Lab of Computational Chemistry and Modelling, Faculty of Natural Sciences,  
Quy Nhon University, Vietnam

<sup>2</sup>Department of Computational Chemistry, J. Heyrovsky Institute of Physical Chemistry,  
Czech Academy of Sciences, Czech Republic

Received: 10/10/2024; Revised: 08/01/2025;

Accepted: 20/01/2025; Published: 28/02/2025

## ABSTRACT

This study conducts a thorough investigation into the stability of complexes as well as the hydrogen bonds along with their characteristics in the systems between formic acid and chalcone derivatives. Generally, the strength of complexes is enhanced irrespective of electron donating or withdrawing substitution in chalcone derivatives. It is found that the complexes involving Se-substitution are slightly more stable than Te-one. The halogenated complexes are less stable than  $CH_3$ - and  $NH_2$ - substituted ones, in which the largest stability belongs to the complexes involving  $NH_2$ - substituted group. The obtained results show a dominant role of O-H...Se/Te compared to C-H...O in contributing to the stabilization of complexes. Calculated results indicate that the strength of nonconventional hydrogen bonds decreases in the order of O-H...Se > O-H...Te ~ C-H...O > C-H...Se/Te. The larger blue shift of C-H bond in the complexes investigated is observed in C-H...O compared to C-H...Se/Te. The magnitude of blue shift is larger in the case of C-H...Se than in C-H...Te hydrogen bonds. NBO analysis shows that the blue-shifted stretching frequency of  $C_{sp^2}$ -H depends mainly on a reduction of electron density at  $\sigma^*(C_{sp^2}-H)$  orbital. The redshift in O-H...Se/Te hydrogen bonds in these systems is determined by a considerable increase of electron density at  $\sigma^*(O-H)$  orbitals, overcoming an s-character enhancement of the O site upon complexation.

**Keywords:**  $O/C_{sp^2}$ -H...O/Se/Te nonconventional hydrogen bonds, Se/Te-aldehydes, formic acid.

## 1. INTRODUCTION

Hydrogen bonds play a crucial role in molecular information recognition, protein folding, structural rearrangement of nucleic acids, crystallization, polymerization, supramolecular chemistry, solvation, or organic synthesis.<sup>1,2</sup> Many investigations were performed to understand the characteristics of hydrogen bonds

and their importance in the arrangement of molecules, surface phenomena,...<sup>3-6</sup> Noticeably, non-classical hydrogen bonds of the form C-H...O/N/X/ $\pi$  (X= F, Cl, Br) were detected in many systems including proteins, DNA, RNA, and material surfaces. Besides, the activation of C-H groups in intermediate structures is one of the issues of current interest in organic synthesis.<sup>7,8</sup>

\*Corresponding author.

Email: nguyentientrung@qnu.edu.vn

Experimental, Raman and theoretical findings of blue-shifted C-H bonds involved in hydrogen bonds have been reported, of which the most abundant is the C-H $\cdots$ O/N form.<sup>9-17</sup>

Several investigations focused on explaining the nature of blue-shifted hydrogen bonding, but none can be universally applied to other complex systems stabilized by hydrogen bonds.<sup>9,16-20</sup> Given the critical role of hydrogen bonds in biological systems, the strength and properties of classical hydrogen bonds O/S-H $\cdots$ O/S have been studied and reported by theoretical and experimental methods.<sup>21,22</sup> Many nonconventional hydrogen bonds have been discovered recently and play essential roles in proteins and catalysis. Indeed, the hydrogen bonds O/N-H $\cdots$ Se/Te were recently discovered experimentally.<sup>23-24</sup> Besides, the existence of H $\cdots$ Se hydrogen bonds in proteins and nucleobase pairs was determined by spectroscopic evidence.<sup>25,26</sup> However, systematic studies of the O-H $\cdots$ Se/Te hydrogen bond system to understand their strength and properties have not been reported. In addition, the influence of this O-H $\cdots$ Se/Te on the strength and characteristics of the C<sub>sp<sup>2</sup></sub>-H $\cdots$ Se/Te was not considered and evaluated in the literature.

In addition, the aldehydes, carboxylic acid, and their halogenated and amide derivatives are commonly found in the biological structures and biologically active compounds. Theoretical studies indicated that the origin of stabilization in carboxylic acids and amide complexes is charge-delocalization.<sup>2,27</sup> The experimental and theoretical investigations into O-H redshift and C<sub>sp<sup>2</sup></sub>-H blueshift in the hydrogen bonds in formic acid, acetic acid, and acetaldehyde complexes were reported.<sup>28-30</sup> The strength and nature of blue- and red-shifting hydrogen bonds were examined in formamide, aldehydes, hydroxyl derivatives, and carboxylic acids.<sup>31-33</sup> More recently, the C<sub>sp<sup>2</sup></sub>-H blue shifting in the C<sub>sp<sup>2</sup></sub>-H $\cdots$ O hydrogen bonds have been reported for

the complexes of aldehydes with formic acid, as well as the binary systems of aldehyde and carboxylic acids.<sup>16,34,35</sup>

Further, a theoretical study on the nature of hydrogen bonds in NH<sub>2</sub>CYH $\cdots$ XH complexes with Y = O, S, Se, Te; X = F, HO, NH<sub>2</sub> was examined by quantum chemistry analyses.<sup>36</sup> Noticeably, the C<sub>sp<sup>2</sup></sub>-H $\cdots$ O/S/Se/Te hydrogen bonds were evaluated in the XCHO $\cdots$ nH<sub>2</sub>Z and dimers of chalcogenoaldehyde derivatives systems.<sup>37,38</sup> However, the two hydrogen-bond systems between selenoformaldehyde or telluroformaldehyde and carboxylic acid and their derivatives have not yet been observed clearly. In addition, the effects of various X substituents relating to the strength of nonconventional hydrogen bonds and the shifts in C<sub>sp<sup>2</sup></sub>-H or O-H stretching frequency need to be addressed. Furthermore, the existence and role of various interactions following complexation, such as O-H $\cdots$ Se/Te and C<sub>sp<sup>2</sup></sub>-H $\cdots$ O/Se/Te, should be examined in the complexes of Se/Te derivatives to have insights into the nonconventional hydrogen bonds. In this work, the interaction complexes between aldehyde derivatives and formic acid are selected to have thorough insights into the effects of different substituent groups on the stability and characteristics of O-H $\cdots$ Se/Te, C<sub>sp<sup>2</sup></sub>-H $\cdots$ Se/Te nonconventional hydrogen bonds, and the complex strength.

## 2. COMPUTATIONAL METHODS

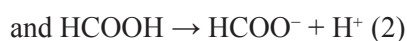
The optimized geometries for monomers and complexes are considered at the MP2/6-311++G(3df,2pd) level of theory with the correction of zero-point vibrational energies (ZPE). The interaction energy ( $\Delta E^*$ ) of each complex is calculated and corrected for basis set superposition errors (BSSE) at the CCSD(T)/6-311++G(3df,2pd)//MP2/6-311++G(3df,2pd) level. Particularly,  $\Delta E^*$  is represented as the difference in total electronic energy between the complexes and the sum of XCHZ and HCOOH monomers given by the expression:

<https://doi.org/10.52111/qnjs.2025.19109>



$$\Delta E^* = E_c - (E_{m1} + E_{m2}).$$

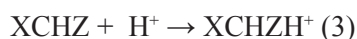
Here,  $E_c$  is the energy of the optimized complexes including ZPE and BSSE corrections.  $E_{m1}$  and  $E_{m2}$  are energy values of the optimized monomers HCHZ and HCOOH, respectively, with the ZPE correction. Moreover, deprotonation enthalpy (DPE) is defined as the enthalpy change for the deprotonation reactions:



and is considered as DPE of C-H and O-H bonds in both XCHZ and HCOOH monomers.

$DPE(C-H) = E^\circ(HCOO^-) - E^\circ(HCOOH) + TCE(HCOO^-) + TCE(H^+) - TCE(HCOOH)$ , in which  $E^\circ$  and TCE values are electronic energy and thermal correction to the enthalpy of species, respectively.

In addition, proton affinity (PA) is defined as the enthalpy change for the protonation reaction:



and is considered PA at the Se and Te sites in XCHZ monomers.

$$PA(Se/Te) = E^\circ(XCHZ) - E^\circ(XCHZH^+) + TCE(XCHZ) + TCE(H^+) - TCE(XCHZH^+).$$

These calculations are performed by the Gaussian 16 package.<sup>39</sup>

The Atoms-in-Molecules (AIM) analysis is observed at the MP2/6-311++G(3df,2pd) level through the AIMAll program.<sup>40,41</sup> The local electron energy density ( $H(r)$ ) is computed from potential ( $V(r)$ ) and kinetic ( $G(r)$ ) terms by the equation:

$$H(r) = G(r) + V(r).$$

The individual energy of each hydrogen bond ( $E_{HB}$ ) is estimated using the formula:

$$E_{HB} = 0.5V(r).^{42}$$

Further, the natural bond orbital (NBO) analysis is examined by using NBO 5.G software

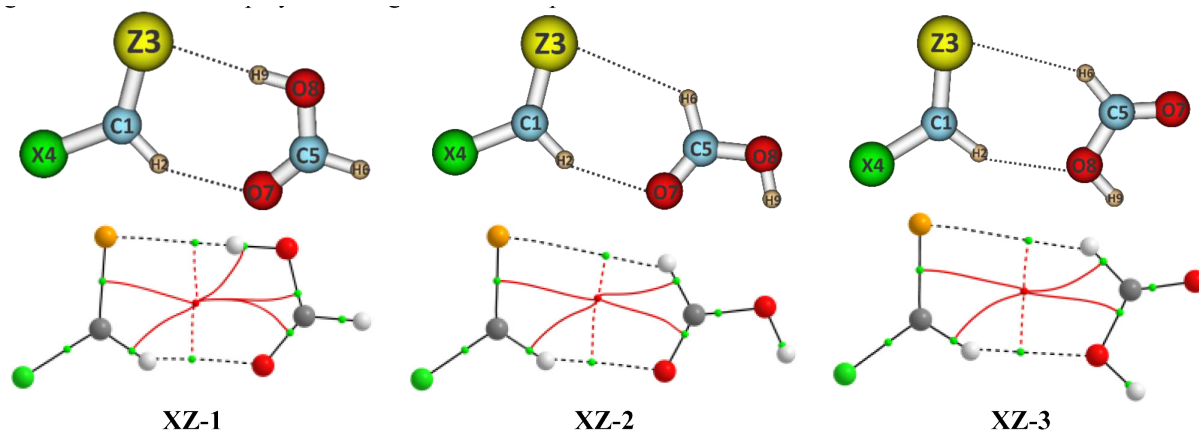
at the range-separated hybrid functional  $\omega$ B97X-D/6-311++G(3df,2pd) level.<sup>43-45</sup> Besides, Symmetry Adapted Perturbation Theory (SAPT) calculations based on the Psi4 program with the def2-TZVPD basis set are applied to unravel the different contributions of energy components, including the electrostatic ( $E_{elec}$ ), induction ( $E_{ind}$ ), and dispersion ( $E_{disp}$ ) to the stability of examined complexes.<sup>46</sup>

### 3. RESULTS AND DISCUSSION

#### 3.1. Stable geometry and AIM analysis

The interaction of XCHZ with HCOOH induces 36 stable complexes, denoted by **XZ-n** (with X = H, F, Cl, Br, CH<sub>3</sub>, NH<sub>2</sub>; Z = Se, Te and n = 1-3) with three geometries **XZ-n** displayed in Figure 1. The sp<sup>2</sup>-hybridized carbon atom in XCHZ is labeled hereafter as C<sub>sp2</sub>. Each complex contains a ring that is stabilized by two intermolecular contacts, including O-H...Se/Te and C<sub>sp2</sub>-H...O in **XZ-1**, C-H...Se/Te and C<sub>sp2</sub>-H...O in **XZ-2** and **XZ-3**. All H...O, H...Se and H...Te distances (Table S1) are in the range of 2.18–2.45 Å, 2.33–3.20 Å and 2.65–3.33 Å, respectively. These values are close to the sums of van der Waals radii of relevant atoms (being 2.72 Å for H...O, 3.00 Å for H...Se and 3.16 Å for H...Te). The presence of H...Se/Te and C<sub>sp2</sub>-H...O/Se/Te nonconventional hydrogen bonds through the bonding critical points (BCPs, green points in Figure S1) is confirmed by the values of electron density, Laplacian, and H(r) at BCPs (Table S1).<sup>41</sup> Moreover, the C<sub>sp2</sub>/O-H...Se/Te distances are also close to those in complexes between NH<sub>2</sub>CHSe and NH<sub>2</sub>CHTe, with HF/H<sub>2</sub>O/NH<sub>3</sub>.<sup>47</sup> The intermolecular contacts with the presence of chalcogen atoms (O, S, Se, Te) in hydrogen bonds were also reported in some studies.<sup>16,34,36,48</sup>

In addition, the decreasing tendency of strength for hydrogen bonds is from O-H...Se to O-H...Te ~ C-H...O and then to C-H...Se/Te (Table S1). Indeed, the  $E_{HB}$  values indicate that the stability of C-H...O (-7.4 ÷ -15.6 kJ.mol<sup>-1</sup>) is 2÷3 times as large as that of C-H...Se/Te

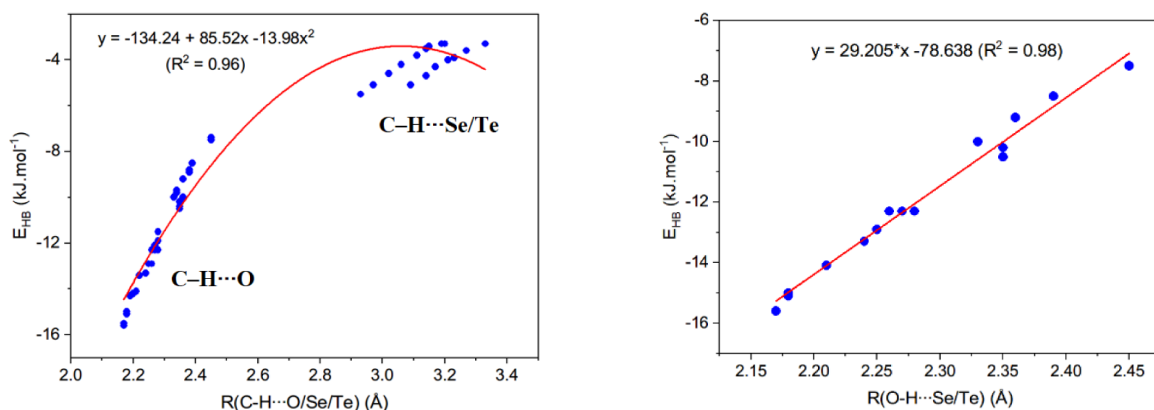


**Figure 1.** The stable structures and topological features of XCHZ and HCOOH complexes (X = H, F, Cl, Br, CH<sub>3</sub>, NH<sub>2</sub> and Z = Se, Te).

( $-3.3 \div -5.5$  kJ.mol<sup>-1</sup>). Similarly, the strength of O–H...Se ( $-16.8 \div -23.1$  kJ.mol<sup>-1</sup>) is 1.5 times as large as that of O–H...Te ( $-12.5 \div -13.8$  kJ.mol<sup>-1</sup>). Along with the results of individual hydrogen bond energies in literature by Khanh et al.<sup>16</sup>, it is found that the hydrogen bond stability decreases in going from O–H...O  $\gg$  O–H...S  $>$  O–H...Se  $>$  O–H...Te. These results are also supported by a previous investigation of hydrogen bonding complexes containing chalcogen atoms.<sup>34</sup> Besides, the second correlation of energy ( $E_{\text{HB}}$ ) versus distances  $R(\text{C–H...O/Se/Te})$  is obtained according to expression  $E_{\text{HB}} = -134.24 + 85.52R(\text{C–H...O/Se/Te}) - 13.98R^2(\text{C–H...O/Se/Te})$  ( $R^2 = 0.96$ ) as given in Figure 2. However, a highly linear correlation of  $E_{\text{HB}}$  with respect to  $R(\text{O–H...Se/Te})$  is observed as expression  $E_{\text{HB}} = 29.205R(\text{O–H...Se/Te}) - 78.638$  ( $R^2 = 0.98$ ). The results indicate that the shorter the intermolecular distance, the stronger the nonconventional hydrogen bond is, and vice versa.

As given in Table S1 of ESI, it can be suggested that the C–H...O hydrogen bonds play a main role in stabilizing **XZ-2** and **XZ-3** along with the addition of the C–H...Se/Te ones. Nevertheless, for **XZ-1** complexes, the C–H...O hydrogen bonds serve as a supplementary factor

along with the substantial importance of the O–H...Se/Te ones in complex stabilization. To clarify the relationship between the strength of nonconventional hydrogen bonds and  $r_{\text{H...OZ}}/\Sigma r_{\text{vdW}}$  ratio, the  $r_{\text{H...Z}}/\Sigma r_{\text{vdW}}$  ratios are gathered and plotted in Table S1 and Figures S2, S3 of ESI. The results indicate a good agreement between the  $r_{\text{H...Z}}/\Sigma r_{\text{vdW}}$  ratio and  $E_{\text{HB}}$  of nonconventional hydrogen bonds in investigated complexes. Indeed, the  $r_{\text{H...Z}}/\Sigma r_{\text{vdW}}$  ratios of C–H...O hydrogen bonds in the **XSe/Te-1** are smaller than those in **XSe/Te-2**, and **XSe/Te-3**. This result is consistent with the high stability of C–H...O hydrogen bonds in **XSe/Te-1** compared to **XSe/Te-2**, and **XSe/Te-3**. The  $r_{\text{H...Z}}/\Sigma r_{\text{vdW}}$  ratio of C–H...O increases in the order of NH<sub>2</sub>  $<$  Cl  $\sim$  Br  $<$  F  $<$  CH<sub>3</sub>  $<$  H substitutions, along with the decrease of C–H...O strength in the order: NH<sub>2</sub>  $>$  Cl  $\sim$  Br  $>$  F  $>$  CH<sub>3</sub>  $>$  H substituted derivatives. It is worth noted that the O–H...Se/Te hydrogen bonds are much more stronger than the C<sub>sp2</sub>-H...Se/Te ones, which is proven by the much smaller  $r_{\text{H...Z}}/\Sigma r_{\text{vdW}}$  ratio of O–H...Se/Te than that of C<sub>sp2</sub>-H...Se/Te hydrogen bonds. For **XZ-2** and **XZ-3**, the  $r_{\text{H...Z}}/\Sigma r_{\text{vdW}}$  ratios of C–H...Se/Te are larger than those of C–H...O, being in line with the weaker strength of the C–H...Se/Te than the C–H...O.



**Figure 2.** The relationships between  $E_{\text{HB}}$  and  $R(\text{C/O-H}\cdots\text{O/Se/Te})$  in the complexes.

**Table 1.** Deprotonation enthalpies of  $\text{C}_{\text{sp}^2}\text{-H}$  (DPE) and proton affinities at Z sites (PA) at the CCSD(T)/6-311++G(3df,2pd) (in  $\text{kJ.mol}^{-1}$ ).

| Monomer                                  | HCHSe  | FCHSe  | ClCHSe | BrCHSe | $\text{CH}_3\text{CHSe}$ | $\text{NH}_2\text{CHSe}$ |
|------------------------------------------|--------|--------|--------|--------|--------------------------|--------------------------|
| PA(Se)                                   | 775.0  | 737.9  | 766.3  | 773.0  | 811.7                    | 860.0                    |
| DPE( $\text{C}_{\text{sp}^2}\text{-H}$ ) | 1623.6 | 1567.2 | 1551.7 | 1539.0 | 1620.7                   | 1595.6                   |
| Monomer                                  | HCHTe  | FCHTe  | ClCHTe | BrCHTe | $\text{CH}_3\text{CHTe}$ | $\text{NH}_2\text{CHTe}$ |
| PA(Te)                                   | 812.2  | 780.7  | 802.7  | 807.9  | 841.0                    | 884.6                    |
| DPE( $\text{C}_{\text{sp}^2}\text{-H}$ ) | 1612.7 | 1561.0 | 1550.7 | 1540.8 | 1615.5                   | 1582.3                   |

Some characteristics of  $\text{XCHZ}$  monomers, including deprotonation enthalpies (DPE) and proton affinities (PA) are examined in Table 1. With the same Z in  $\text{XCHZ}$ , the strength of  $\text{O-H}\cdots\text{Z}$  hydrogen bonds in **XZ-1** decreases in the order of  $\text{NH}_2 > \text{CH}_3 > \text{H} > \text{F/Cl/Br}$  substitutions (*cf.* Table S1), which are consistent with a decrease of gas-phase basicity (PA) at Se/Te site in  $\text{XCHZ}$  (*cf.* Table 1). However, a decrease in the strength of  $\text{C-H}\cdots\text{O}$  hydrogen bonds in **XZ-1** is in the ordering  $\text{NH}_2 > \text{F/Cl/Br} > \text{CH}_3 > \text{H}$  substitutions, which differs from the trend of the polarity of  $\text{C}_{\text{sp}^2}\text{-H}$  bond in the isolated  $\text{XCHZ}$ . As a result, the strength of **XZ-1** is mainly proposed by the presence of  $\text{O-H}\cdots\text{Z}$ . Similarly, a decrease in gas-phase basicity at the Se/Te site in the sequence of  $\text{NH}_2 > \text{CH}_3 > \text{H} > \text{F/Cl/Br}$  derivatives causes a reduction of the strength of  $\text{C-H}\cdots\text{Z}$  hydrogen bonds in **XZ-2**, **XZ-3** decrease in the same tendency. The strengthening trend of  $\text{C-H}\cdots\text{O}$  hydrogen bonds in **XZ-2** and **XZ-3** is

also suitable for an increase in the polarity of the  $\text{C}_{\text{sp}^2}\text{-H}$  bond in  $\text{XCHZ}$  in the order of  $\text{H} < \text{CH}_3 < \text{NH}_2 < \text{F/Cl/Br}$  (*cf.* Table S1). With the identical X in  $\text{XCHZ}$ , the stronger strength of  $\text{O-H}\cdots\text{Se}$  hydrogen bonds compared to  $\text{O-H}\cdots\text{Te}$  ones is observed in **XZ-1**. However, for the same X in the isolated  $\text{XCHZ}$ , proton affinity at the Te site is larger than that at the Se one (*cf.* Table 1). Accordingly, these hydrogen bonds should be assigned by a larger electrostatic attraction of H and Se relative to H and Te.

### 3.2. Interaction energy and SAPT analysis

Interaction energies ( $\Delta E^*$ ) of complexes, including both ZPE and BSSE correction at the CCSD(T)/6-311++G(3df,2pd)//MP2/6-311++G(3df,2pd) level are listed in Table 2, and their relationship is plotted in Figure 3. Generally, the  $\Delta E^*$  values in the range of -7.1 and -37.2  $\text{kJ.mol}^{-1}$  indicate that the obtained complexes are quite stable, in which the **XZ-1** complexes are 2÷3 times as stable as

**XZ-2** and **XZ-3** complexes. The considerable strength of **XZ-1** is suggested by the presence of O-H...Z hydrogen bond respective to the remaining nonconventional C-H...O/Z hydrogen bonds. The larger stability of **XZ-2** relative to **XZ-3** is due to the larger gas phase basicity of the O site of >C=O compared to that of the O site of -OH in formaldehyde acid. Both of them are stablized by the C-H...O/Z hydrogen bonds.

It is found that the most stable structures (**XZ-1**) of XCHZ...HCOOH in the present work have a comparable strength of XCHO...HCOOH complexes and are slightly larger than that

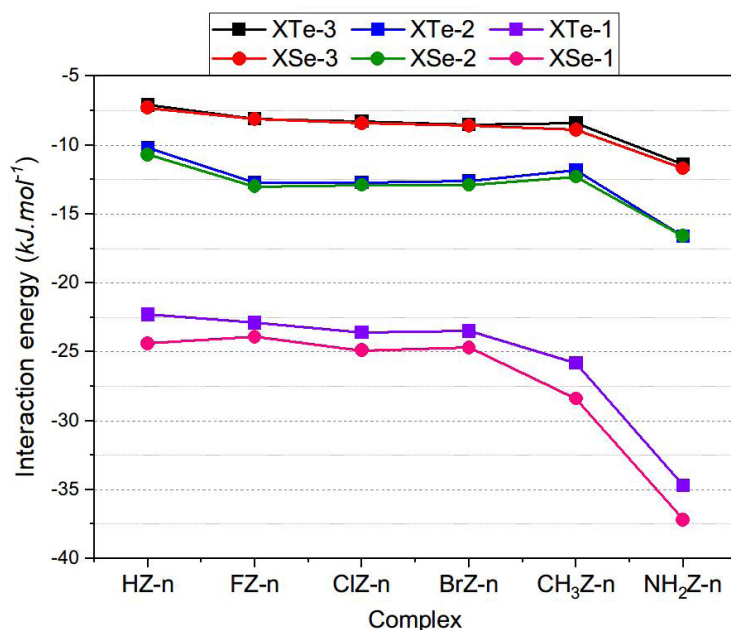
of XCHS...HCOOH.<sup>16</sup> The **XZ-1** complexes are more stable than the binary complexes of small aldehydes and less stable than dimers of typical carboxylic acids. Indeed, the interaction energies of HCHO...HCHO, HCHO...HCHS, HCHS...HCHS are in turn -12.3, -11.7, -10.7 kJ.mol<sup>-1</sup> at CCSD(T)-F12/heavy-aug-cc-pVTZ level of theory.<sup>49</sup> The complexes of HCHO...HCHO and CH<sub>3</sub>CHO...CH<sub>3</sub>CHO have interaction energies of -21.0 and -22.4 kJ.mol<sup>-1</sup> at MP2/aug-cc-pVTZ.<sup>50</sup> The very large interaction energies of CH<sub>3</sub>COOH...CH<sub>3</sub>COOH and HCOOH...HCOOH at MP2/aug-cc-pVTZ are -60.7 kJ.mol<sup>-1</sup> and -56.9 kJ.mol<sup>-1</sup>.<sup>51</sup>

**Table 2.** Interaction energy corrected by both ZPE and BSSE ( $\Delta E^*$ , kJ.mol<sup>-1</sup>) of the complexes RCHZ...HCOOH at CCSD(T)/6-311++G(3df,2pd)//MP2/6-311++G(3df,2pd).

| Complex                   | $\Delta E^*$ | Complex                   | $\Delta E^*$ | Complex                   | $\Delta E^*$ |
|---------------------------|--------------|---------------------------|--------------|---------------------------|--------------|
| <b>HSe-1</b>              | -24.4        | <b>HSe-2</b>              | -10.7        | <b>HSe-3</b>              | -7.3         |
| <b>FSe-1</b>              | -23.9        | <b>FSe-2</b>              | -13.0        | <b>FSe-3</b>              | -8.1         |
| <b>ClSe-1</b>             | -24.9        | <b>ClSe-2</b>             | -12.9        | <b>ClSe-3</b>             | -8.4         |
| <b>BrSe-1</b>             | -24.7        | <b>BrSe-2</b>             | -12.9        | <b>BrSe-3</b>             | -8.6         |
| <b>CH<sub>3</sub>Se-1</b> | -28.4        | <b>CH<sub>3</sub>Se-2</b> | -12.3        | <b>CH<sub>3</sub>Se-3</b> | -8.9         |
| <b>NH<sub>2</sub>Se-1</b> | -37.2        | <b>NH<sub>2</sub>Se-2</b> | -16.6        | <b>NH<sub>2</sub>Se-3</b> | -11.7        |
| <b>HTe-1</b>              | -22.3        | <b>HTe-2</b>              | -10.2        | <b>HTe-3</b>              | -7.1         |
| <b>FTe-1</b>              | -22.9        | <b>FTe-2</b>              | -12.7        | <b>FTe-3</b>              | -8.1         |
| <b>CTe-1</b>              | -23.6        | <b>CTe-2</b>              | -12.7        | <b>CTe-3</b>              | -8.3         |
| <b>BrTe-1</b>             | -23.5        | <b>BrTe-2</b>             | -12.6        | <b>BrTe-3</b>             | -8.5         |
| <b>CH<sub>3</sub>Te-1</b> | -25.8        | <b>CH<sub>3</sub>Te-2</b> | -11.8        | <b>CH<sub>3</sub>Te-3</b> | -8.4         |
| <b>NH<sub>2</sub>Te-1</b> | -34.7        | <b>NH<sub>2</sub>Te-2</b> | -16.6        | <b>NH<sub>2</sub>Te-3</b> | -11.4        |

It is noticeable that the interaction energy of **XSe-n** is *ca.* 0.1–2.6 kJ.mol<sup>-1</sup>, slightly more negative than that of **XTe-n**, about for the same X in XCHZ derivatives. This indicates a slightly lower strength of **XTe-n** as compared to **XSe-n**, which should be resulting from the stronger electrostatic force of Se and H atoms overcoming Te and H atoms.<sup>34</sup> Generally, the strength of **XZ-n** complexes is enhanced irrespective of electron donating or withdrawing X substitution

(cf. Table 2 and Figure 3). For the alike Z, it is found that the halogenated complexes are less stable than CH<sub>3</sub>- and NH<sub>2</sub>- substituted ones. The largest stability belongs to the complexes **NH<sub>2</sub>Z-n** with more negative values of 2.8-13.3 kJ.mol<sup>-1</sup>. It can be suggested that O-H...Se/Te and C-H...O/Se/Te nonconventional hydrogen bonds play a significant role in contributing to the stability of **NH<sub>2</sub>Z-n** compared to other complexes.



**Figure 3.** The relationship of the interaction energies and different X substitutions.

As seen from Table 2, the interaction energy of complexes indeed increases in the substituting order:  $F < H < Br \sim Cl < CH_3 < NH_2$  for **XSe-1** and  $H < F < Br \sim Cl < CH_3 < NH_2$  for **XTe-1**. This result is suitable for the changing trend in the stability of  $O-H \cdots Se/Te$  hydrogen bonds, affirming a main role of  $O-H \cdots Se/Te$  along with an addition of  $C-H \cdots O$  hydrogen bonds in stabilizing **XZ-1**. The energetic components based on the SAPT2+ approach are observed in Table S2. The calculated results imply the stability of complexes is mainly contributed by three terms, including electrostatic, dispersion, and induction, with contributing percentages of *ca.* 8 - 53%, 7 - 46% and 7 - 63%, respectively. The dominating role of electrostatic components contributing to the complex stability is also reported in the complexes of  $XCHO \cdots HCOOH$  and  $XCHS \cdots HCOOH$ .<sup>16</sup>

### 3.3. Characteristics of $C_{sp^2}-H \cdots Se/Te$ and $O-H \cdots Se/Te$ nonconventional hydrogen bonds

The formation of  $C-H \cdots Z$ ,  $C-H \cdots O$  and  $O-H \cdots Z$  interactions in the complexes are further clarified by the change in lengths ( $\Delta r$ , mÅ) and the stretching frequencies ( $\Delta \nu$ ,  $cm^{-1}$ ) of  $O-H$  and  $C_{sp^2}-H$  bond participating in the considered

hydrogen bonds. The results at MP2/6-311++G(3df,2pd) are gathered in Table 3.

Considering the nonconventional hydrogen bonds of  $C-H \cdots Se/Te$ , the  $C-H$  length in **XZ-2** and **XZ-3** decreases slightly in the range of 0.07 – 0.93 mÅ, accompanied by an increase in their stretching frequency of 1.9 – 14.8  $cm^{-1}$  as compared to the corresponding values of isolated monomers. Therefore, these  $C-H \cdots Se/Te$  contacts belong to blue-shifting hydrogen bonds. The larger blue shift of the  $C-H$  bond in the complexes investigated is observed in  $C-H \cdots O$  compared to  $C-H \cdots Se/Te$  and has a much smaller magnitude than that in the complexes of  $XCHO \cdots HCOOH$ . It was found that all  $C-H \cdots Se/Te$  nonconventional hydrogen bonds in dimers of  $RCHZ$  ( $R=H, F, Cl, Br$  or  $CH_3$ ,  $Z=Se, Te$ ) are red-shifted.<sup>16</sup> In the present work, a small  $C_{sp^2}-H$  elongation of 0.10 – 0.52 mÅ along with a decrease in stretching frequency of 0.2 – 6.5  $cm^{-1}$  is found in the case of **NH<sub>2</sub>Se-3**, **BrTe-3**, **HTe-3**, **CH<sub>3</sub>Te-3**, and **NH<sub>2</sub>Te-3**, indicating red shifting hydrogen bonds. These values are quite close to a previous report of hydrogen bonded complexes.<sup>37</sup> The largest shortening of  $C-H$  bond length is observed for **FZ-2** and **FZ-3**, associated with the strongest electron-withdrawing group



being F, whereas the smallest is observed in **NH<sub>2</sub>Z-2** and **NH<sub>2</sub>Z-3**. Besides, the blue shift of C–H6 bond in C5–H6···Z3 interaction in **XZ-2** and **XZ-3** decreases in the substituting order: F > Cl > Br > H > CH<sub>3</sub> > NH<sub>2</sub>. This result is consistent with increased basicity at the Se/Te site in XCHZ (Table 1). The blue shift of the C<sub>sp2</sub>-H bonds with the Z same is more strongly affected by electron-withdrawing groups (F, Cl, Br) than by electron-donating groups (CH<sub>3</sub>,

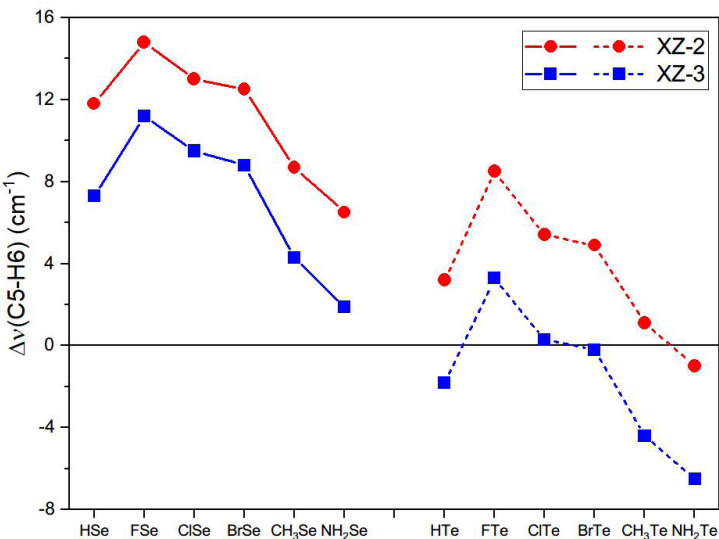
NH<sub>2</sub>). The magnitude of the C5–H6 blue shift is larger for C–H···Se than for C–H···Te, which is in alignment with the increasing tendency of the proton affinity at Te relative to the Se site (cf. Table 1). Furthermore, the larger blueshift of the C–H bond in C–H···Z hydrogen bonds of **XZ-2** in comparison with that of **XZ-3** is observed, which relates to the more crucial role of C–H···O hydrogen bond in **XZ-2** compared to that in **XZ-3**, toward the blue shift of C–H bonds.

**Table 3.** Changes in bond lengths (in mÅ) and their corresponding stretching frequencies (in cm<sup>-1</sup>) at MP2/6-311++G(3df,2pd).

| Complex  | HSe-1  | FSe-1  | ClSe-1 | BrSe-1 | CH <sub>3</sub> Se-1 | NH <sub>2</sub> Se-1 | HTe-1  | FTe-1  | ClTe-1 | BrTe-1 | CH <sub>3</sub> Te-1 | NH <sub>2</sub> Te-1 |
|----------|--------|--------|--------|--------|----------------------|----------------------|--------|--------|--------|--------|----------------------|----------------------|
| Δr(C1H2) | 0.42   | 0.76   | 1.54   | 1.77   | -0.37                | 0.22                 | 1.09   | 0.51   | 1.37   | 1.55   | -0.08                | 0.69                 |
| Δν(C1H2) | 8.4    | 3.8    | -4.5   | -6.8   | 18.9                 | 8.1                  | 0.3    | 7.2    | -2.4   | -4.6   | 14.3                 | 1.6                  |
| Δr(O8H9) | 15.32  | 12.49  | 13.24  | 13.17  | 16.80                | 21.18                | 13.97  | 12.20  | 12.78  | 12.69  | 15.09                | 19.24                |
| Δν(O8H9) | -326.2 | -266.6 | -284.0 | -282.5 | -355.5               | -439.4               | -299.5 | -261.9 | -274.6 | -273.1 | -322.1               | -402.6               |

| Complex  | HSe-2 | FSe-2 | ClSe-2 | BrSe-2 | CH <sub>3</sub> Se-2 | NH <sub>2</sub> Se-2 | HTe-2 | FTe-2 | ClTe-2 | BrTe-2 | CH <sub>3</sub> Te-2 | NH <sub>2</sub> Te-2 |
|----------|-------|-------|--------|--------|----------------------|----------------------|-------|-------|--------|--------|----------------------|----------------------|
| Δr(C1H2) | 0.17  | 0.21  | 0.76   | 0.92   | -0.34                | 0.18                 | 0.84  | 0.04  | 0.72   | 0.89   | -0.20                | 0.53                 |
| Δν(C1H2) | 7.8   | 7.8   | 3.0    | 1.2    | 15.6                 | 6.9                  | -0.3  | 10.0  | 3.3    | 1.0    | 12.8                 | 1.3                  |
| Δr(C5H6) | -0.73 | -0.93 | -0.79  | -0.76  | -0.47                | -0.30                | -0.52 | -0.78 | -0.65  | -0.60  | -0.28                | -0.07                |
| Δν(C5H6) | 11.8  | 14.8  | 13.0   | 12.5   | 8.7                  | 6.5                  | 7.3   | 11.2  | 9.5    | 8.8    | 4.3                  | 1.9                  |

| Complex  | HSe-3 | FSe-3 | ClSe-3 | BrSe-3 | CH <sub>3</sub> Se-3 | NH <sub>2</sub> Se-3 | HTe-3 | FTe-3 | ClTe-3 | BrTe-3 | CH <sub>3</sub> Te-3 | NH <sub>2</sub> Te-3 |
|----------|-------|-------|--------|--------|----------------------|----------------------|-------|-------|--------|--------|----------------------|----------------------|
| Δr(C1H2) | -0.07 | -0.33 | -0.17  | -0.07  | -0.48                | -0.21                | 0.51  | -0.36 | -0.09  | -0.04  | -0.43                | -0.05                |
| Δν(C1H2) | 10.1  | 13.4  | 13.6   | 12.6   | 16.7                 | 12.5                 | 3.5   | 13.6  | 12.5   | 11.8   | 15.2                 | 9.9                  |
| Δr(C5H6) | -0.07 | -0.43 | -0.20  | -0.17  | 0.09                 | 0.24                 | 0.18  | -0.16 | 0.06   | 0.10   | 0.38                 | 0.52                 |
| Δν(C5H6) | 3.2   | 8.5   | 5.4    | 4.9    | 1.1                  | -1.0                 | -1.8  | 3.3   | 0.3    | -0.2   | -4.4                 | -6.5                 |



**Figure 4.** The relationship between the change in C-H stretching frequency and different X substitutions in **XZ-2** and **XZ-3**.



For **XZ-3**, the C-H...O hydrogen bond displays a blue shift with a small contraction and an increase of C-H stretching frequency ( $\Delta r = -0.04 \div -0.48 \text{ m\AA}$  and  $\Delta \nu = 9.9 \div 16.7 \text{ cm}^{-1}$ ), except for a small elongation in **HTe-3**. Besides, the blue shift of the C-H bond observed in **XZ-2** is emphasized by a slight enhancement of its bond length and stretching frequency (except for **HTe-2**, with a slight redshift of C-H). A slight redshift is observed in **ClZ-1** and **BrZ-1**, while most **XZ-1** structures indicate a blue shifting of C-H...O hydrogen bonds. The redshift for hydrogen bonds in **ClZ-1** and **BrZ-1** is based on the high polarity of the  $C_{sp^2}$ -H bonds in ClCHZ and BrCHZ monomers (Z = Se, Te).<sup>16</sup> As shown in Table 3, a C-H bond contraction and an increase in its stretching frequency in **XSe-n** are larger than those in **XTe-n**. This observation is induced by the C-H strong polarity in **XCHTe** compared to **XCHSe**. Meanwhile, with the same Z in **XZ-3**, the larger blueshift in **CH<sub>3</sub>Z-3** in comparison with **F/Cl/BrZ-3** comes from the smaller polarity of the C-H bond in **CH<sub>3</sub>Z-3**. A very small blue shift of  $C_{sp^2}$ -H in C-H...O in XCHZ...HCOOH (Z = Se, Te) as compared to the considerable blue shift of that in XCHO...HCOOH (up to  $100 \text{ cm}^{-1}$ ) was attained.<sup>16</sup> This shows the profound role of O-H...O relative to O-H...Se/Te in increasing the very considerable blue shift of the  $C_{sp^2}$ -H bond. In other words, it implies a noticeable role of O related to S, Se and Te in the blue shifting of C-H...O hydrogen bonds.

On the other hand, for O-H...Se/Te hydrogen bonds in **XZ-1** form, the bond elongation of  $12.5 - 21.2 \text{ m\AA}$  for O-H...Se and  $12.2 - 19.2 \text{ m\AA}$  and O-H...Te is observed which is accompanied by a remarkable decrease of  $266.6 - 439.4 \text{ cm}^{-1}$  and  $261.9 - 402.6 \text{ cm}^{-1}$  in their stretching frequency, respectively (cf. Table 3). These changes show the redshift of O-H stretching frequencies involved in the O-H...Se/Te nonconventional hydrogen bonds. Particularly, with the same X in XCHZ, the redshift phenomenon in the O-H...Se hydrogen bonds is larger than the O-H...Te ones. In the

case of the same Z, the redshift of O-H...Z hydrogen bonds is found in the increasing order of F/Cl/Br < H < CH<sub>3</sub> < NH<sub>2</sub> substitutions and agrees with the enhancement of proton affinity of chalcogen atom in the XCHZ (Table 1). The obtained result of the O-H red shift in the hydrogen bonds is similar to previous reports in complexes of chalcogen derivatives and formic acid.<sup>16,34</sup>

### 3.4. NBO analysis

In order to determine the strength and characteristics of intermolecular interactions and complexes investigated, the NBO analysis is carried out at  $\omega$ B97X-D/6-311++G(3df,2pd), as presented in Table 4. The positive values of electron density transfer (EDT) for most complexes ( $0.0003 - 0.0587 \text{ e}$ ) imply that the electron density is transferred mainly from XCHZ to HCOOH. This tendency is caused by the large energy of electron transfer  $n(Z3) \rightarrow \sigma^*(C5-H6)$  ( $7.02 - 20.57 \text{ kJ.mol}^{-1}$ ) and  $n(Z3) \rightarrow \sigma^*(O8-H9)$  ( $65.42 - 109.35 \text{ kJ.mol}^{-1}$ ) in comparison to  $n(O7/O8) \rightarrow \sigma^*(C1-H2)$  ( $2.88 - 14.59 \text{ kJ.mol}^{-1}$ ), except for **FSe-2**, **ClSe-2** and **BrSe-2** with the negative values of EDT. Hence, the  $n(Z3) \rightarrow \sigma^*(C5-H6)$  and  $n(Z3) \rightarrow \sigma^*(O8-H9)$  electron density transfers play a dominant role in stabilizing most complexes as compared to the  $n(O7/O8) \rightarrow \sigma^*(C-H)$  transfers. A larger role of  $n(O7) \rightarrow \sigma^*(C1-H2)$  ( $10.58 - 11.41 \text{ kJ.mol}^{-1}$ ) related to  $n(Z3) \rightarrow \sigma^*(C5-H6)$  ( $7.02 - 7.48 \text{ kJ.mol}^{-1}$ ) is found in **FSe-2**, **ClSe-2** and **BrSe-2** (cf. Table 4). This result should be due to the stronger polarity of the C-H bond in XCHSe (with X = F, Cl, Br), leading to the larger ability of electron acceptance of  $\sigma^*(C1-H2)$  orbital.

The very large  $E_{\text{inter}}$  values of  $n(Z3) \rightarrow \sigma^*(O8-H9)$  in **XZ-1** compared to  $n(Z3) \rightarrow \sigma^*(C5-H6)$  in **XZ-2** and **XZ-3** indicate that the much stronger electron density transfers from lone pair of electrons from Z3 to  $\sigma^*(O8-H9)$  orbitals are much stronger than from Z3 to  $\sigma^*(C5-H6)$  ones. Moreover,  $E_{\text{inter}}[n(O7/O8) \rightarrow \sigma^*(C1-H2)]$  is higher for **XZ-1** than **XZ-2** and **XZ-3**, implying a stronger electron density

transfer in **XZ-1**. These obtained results of  $E_{inter}$  are consistent with the trend in the stability of hydrogen bonds as analyzed in the AIM part above.

For changing X in XCHZ,  $E_{inter}[n(Z3) \rightarrow \sigma^*(C5-H6)]$  in **XZ-2**, **XZ-3** and  $E_{inter}[n(Z3) \rightarrow \sigma^*(O8-H9)]$  in **XZ-1** increase in the order of  $F/Cl/Br < H < CH_3 < NH_2$  substituents,

suitable for the increasing trend of the proton affinity at Z3 site (Table 1). In other words, the intermolecular charge density transfers from XCHZ to HCOOH are raised by the presence of  $CH_3$  and  $NH_2$  groups and are diminished by halogen substituents. It is noted that the increase in values of  $E_{inter}[n(O7/O8) \rightarrow \sigma^*(C1-H2)]$  is attained by the substitution of electron-donating or electron-withdrawing groups.

**Table 4.** Electron density transfer (EDT), hyperconjugation interaction energies ( $E_{inter}$ , in  $\text{kJ}\cdot\text{mol}^{-1}$ ), and changes of electron density ( $\Delta\sigma^*$ , in electron) and s-character percentage of atom involving hydrogen bond ( $\Delta\%s$ , in %).

| Complex                                        | HSe-1  | FSe-1  | ClSe-1 | BrSe-1 | CH <sub>3</sub> Se-1 | NH <sub>2</sub> Se-1 | HTe-1  | FTe-1  | ClTe-1 | BrTe-1 | CH <sub>3</sub> Te-1 | NH <sub>2</sub> Te-1 |
|------------------------------------------------|--------|--------|--------|--------|----------------------|----------------------|--------|--------|--------|--------|----------------------|----------------------|
| EDT                                            | 0.051  | 0.036  | 0.039  | 0.038  | 0.052                | 0.059                | 0.050  | 0.039  | 0.041  | 0.040  | 0.050                | 0.056                |
| $E_{inter}[n(O7) \rightarrow \sigma^*(C1-H2)]$ | 9.28   | 12.54  | 13.21  | 12.25  | 9.91                 | 14.59                | 7.36   | 9.99   | 10.83  | 10.03  | 8.23                 | 13.54                |
| $E_{inter}[n(Z3) \rightarrow \sigma^*(O8-H9)]$ | 84.06  | 68.05  | 70.77  | 68.76  | 90.54                | 109.35               | 75.32  | 65.42  | 67.51  | 66.00  | 77.16                | 95.35                |
| $\Delta\sigma^*(C1-H2)$                        | -0.002 | -0.002 | -0.002 | -0.002 | -0.002               | -0.002               | -0.001 | -0.001 | -0.001 | -0.001 | -0.001               | -0.001               |
| $\Delta\%s(C1)$                                | 1.51   | 1.68   | 1.71   | 1.71   | 1.52                 | 1.74                 | 1.43   | 1.66   | 1.72   | 1.74   | 1.43                 | 1.71                 |
| $\Delta\sigma^*(O8-H9)$                        | 0.052  | 0.042  | 0.045  | 0.045  | 0.055                | 0.064                | 0.051  | 0.044  | 0.046  | 0.046  | 0.053                | 0.061                |
| $\Delta\%s(O8)$                                | 3.37   | 2.97   | 3.07   | 3.03   | 3.57                 | 4.07                 | 3.05   | 2.78   | 2.85   | 2.83   | 3.23                 | 3.70                 |

| Complex                                        | HSe-2  | FSe-2  | ClSe-2 | BrSe-2 | CH <sub>3</sub> Se-2 | NH <sub>2</sub> Se-2 | HTe-2  | FTe-2  | ClTe-2 | BrTe-2 | CH <sub>3</sub> Te-2 | NH <sub>2</sub> Te-2 |
|------------------------------------------------|--------|--------|--------|--------|----------------------|----------------------|--------|--------|--------|--------|----------------------|----------------------|
| EDT                                            | 0.003  | -0.002 | -0.002 | -0.002 | 0.003                | 0.003                | 0.005  | 0.001  | 0.001  | 0.001  | 0.004                | 0.004                |
| $E_{inter}[n(O7) \rightarrow \sigma^*(C1-H2)]$ | 7.61   | 10.58  | 11.41  | 10.78  | 7.61                 | 10.45                | 6.65   | 8.69   | 9.78   | 9.32   | 6.98                 | 10.7                 |
| $E_{inter}[n(Z3) \rightarrow \sigma^*(C5-H6)]$ | 9.41   | 7.02   | 7.48   | 7.48   | 11.37                | 14.59                | 10.91  | 9.03   | 9.28   | 9.28   | 12.79                | 16.39                |
| $\Delta\sigma^*(C1-H2)$                        | -0.001 | -0.002 | -0.002 | -0.002 | -0.002               | -0.001               | -0.001 | -0.001 | -0.001 | -0.001 | -0.001               | -0.001               |
| $\Delta\%s(C1)$                                | 1.13   | 1.3    | 1.44   | 1.5    | 1.1                  | 1.21                 | 1.12   | 1.31   | 1.46   | 1.52   | 1.11                 | 1.26                 |
| $\Delta\sigma^*(C5-H6)$                        | 0.001  | -0.001 | -0.001 | -0.001 | 0.001                | -0.001               | 0.003  | 0.002  | 0.002  | 0.002  | 0.003                | 0.002                |
| $\Delta\%s(C5)$                                | 0.50   | 0.38   | 0.39   | 0.38   | 0.61                 | 0.86                 | 0.45   | 0.37   | 0.37   | 0.37   | 0.55                 | 0.80                 |

| Complex                                        | HSe-3  | FSe-3  | ClSe-3 | BrSe-3 | CH <sub>3</sub> Se-3 | NH <sub>2</sub> Se-3 | HTe-3  | FTe-3  | ClTe-3 | BrTe-3 | CH <sub>3</sub> Te-3 | NH <sub>2</sub> Te-3 |
|------------------------------------------------|--------|--------|--------|--------|----------------------|----------------------|--------|--------|--------|--------|----------------------|----------------------|
| EDT                                            | 0.007  | 0.002  | 0.002  | 0.002  | 0.006                | 0.007                | 0.009  | 0.005  | 0.005  | 0.005  | 0.008                | 0.010                |
| $E_{inter}[n(O8) \rightarrow \sigma^*(C1-H2)]$ | 3.51   | 5.68   | 5.89   | 5.39   | 3.64                 | 4.85                 | 2.88   | 4.39   | 4.93   | 4.51   | 2.97                 | 4.47                 |
| $E_{inter}[n(Z3) \rightarrow \sigma^*(C5-H6)]$ | 12.5   | 9.45   | 10.03  | 9.82   | 14.3                 | 18.22                | 14.71  | 12.46  | 12.79  | 12.58  | 16.59                | 20.57                |
| $\Delta\sigma^*(C1-H2)$                        | -0.001 | -0.001 | -0.001 | -0.001 | -0.001               | -0.001               | -0.001 | -0.001 | -0.001 | -0.001 | -0.001               | 0.001                |
| $\Delta\%s(C1)$                                | 0.72   | 0.84   | 0.94   | 0.97   | 0.68                 | 0.76                 | 0.72   | 0.85   | 0.96   | 1.00   | 0.68                 | 0.80                 |
| $\Delta\sigma^*(C5-H6)$                        | 0.003  | 0.001  | 0.002  | 0.002  | 0.003                | 0.002                | 0.005  | 0.004  | 0.004  | 0.004  | 0.005                | 0.005                |
| $\Delta\%s(C5)$                                | 0.72   | 0.64   | 0.64   | 0.63   | 0.83                 | 1.15                 | 0.66   | 0.62   | 0.61   | 0.60   | 0.75                 | 1.09                 |

(EDT,  $E_{inter}$  at  $\omega B97X-D/6-311++G(3df,2pd)$ ;  $\Delta\sigma^*$ ,  $\Delta\%s$  at  $MP2/6-311++G(3df,2pd)$ )

The C1-H2 blueshift in C1-H2...O7/O8 is caused by a small decrease of electron density at  $\sigma^*(\text{C1-H2})$  orbital from 0.001e to 0.002e (excepting for a very slight increase of  $\sigma^*(\text{C}_{\text{sp}^2}\text{-H})$  occupation in **NH<sub>2</sub>Te-3**). Generally, the blueshift of C5-H6 in the complexes is emphasized by a slight increase in electron density at  $\sigma^*(\text{C5-H6})$  from -0.001e to 0.005e. The O8-H9 redshift is induced by a large increase in electron density at  $\sigma^*(\text{O8-H9})$  orbital from 0.042e to 0.064e. The larger blueshift of C5-H6 is observed in **XSe-n** compared to **XTe-n**. Indeed, the electron density of the  $\sigma^*(\text{C5-H6})$  orbital gets a larger increase when Se is replaced by Te, and s-character percentage of C5 is smaller for C5-H6...Te3 than C5-H6...Se3. The larger blueshift of C5-H6...Se3 bonds, which relates directly to the smaller values of  $\Delta\sigma^*(\text{C5-H6})$ , is found in **XZ-2** in comparison to **XZ-3**. Moreover, the substitution of the H atom in HCHZ by electron-donating groups ( $\text{CH}_3$ ,  $\text{NH}_2$ ) leads to an enhancement of  $\text{C}_{\text{sp}^2}\text{-H}$  stretching frequency involving  $\text{C}_{\text{sp}^2}\text{-H}\cdots\text{Z}$  hydrogen bonds. At the same time, the corresponding diminishment is caused by the presence of electron-withdrawing groups (F, Cl, Br). Consequently, it can be underlined that the larger contraction of the  $\text{C}_{\text{sp}^2}\text{-H}$  bond and its stretching frequency blueshift is identified by a decrease in the population of electron density at  $\sigma^*(\text{C5-H6})$  orbitals and an increase in the percentage of s-character of  $\text{C}_{\text{sp}^2}$  atom. This observation is similar to results obtained in complexes of aldehydes and carboxylic acids.<sup>16,34,37</sup>

For C1-H2...O7/O8 hydrogen bonds, the stronger blueshift of C1-H2 in **XSe-n** compared to that in **XTe-n** relates to a more significant decrease of electron density in  $\sigma^*(\text{C1-H2})$  orbital. The larger redshift phenomenon in the O8-H9...Se hydrogen bonds compared to the O8-H9...Te ones (cf. Table 4) arises from a significant enhancement in occupation of  $\sigma^*(\text{O8-H9})$  orbital. Besides, the increase in electron density of  $\sigma^*(\text{O8-H9})$  orbital and s-character percentage of the O8 site is smaller for halogenated derivatives than for  $\text{CH}_3$  and

$\text{NH}_2$  substitutions. It can be suggested that the increasing occupation of  $\sigma^*(\text{O8-H9})$  orbital is the main factor in determining red-shifting hydrogen bonds in all examined complexes.

#### 4. CONCLUSION

Thirty-six stable structures of XCHZ-HCOOH complexes with three various geometric structures ( $\text{X} = \text{H}, \text{F}, \text{Cl}, \text{Br}, \text{CH}_3, \text{NH}_2$ ;  $\text{Z} = \text{Se}, \text{Te}$ ) are located on the potential energy surfaces. The binding energy of  $\text{C}_{\text{sp}^2}\text{-H}\cdots\text{Se/Te}$  bonds (from -3.3 to -5.5  $\text{kJ}\cdot\text{mol}^{-1}$ ) is smaller than  $\text{C}_{\text{sp}^2}\text{-H}\cdots\text{O}$  and  $\text{O-H}\cdots\text{Te}$  nonconventional hydrogen bonds (from -7.4 to -15.6  $\text{kJ}\cdot\text{mol}^{-1}$ ) and is *ca.* half of  $\text{O-H}\cdots\text{Se}$  ones (from -12.4 to -23.1  $\text{kJ}\cdot\text{mol}^{-1}$ ). Following complexation, the  $\text{O-H}\cdots\text{Se/Te}$  bond strength is determined by proton affinity at the Z atom of XCHZ, while the polarity of  $\text{C}_{\text{sp}^2}\text{-H}$  covalent bonds in XCHZ plays a decisive role in stabilizing  $\text{C}_{\text{sp}^2}\text{-H}\cdots\text{O/Se/Te}$  hydrogen bonds. The existence and stability of nonconventional hydrogen bonds in the investigated systems are clarified by AIM and NBO analyses. With the same X, the  $\text{C}_{\text{sp}^2}\text{-H}\cdots\text{O/Se/Te}$  hydrogen bonds in **XSe-2** and **XTe-2** are more stable than in **XSe-3** and **XTe-3**. The significantly larger stability of  $\text{O-H}\cdots\text{Se}$  hydrogen bonds compared to  $\text{C}_{\text{sp}^2}\text{-H}\cdots\text{O/Se/Te}$  causes a larger strength of **XSe-1** and **XTe-1**, in which **XZ-1** is twice more stable than **XZ-2** and **XZ-3**. With the same Z, the stability of complexes experiences an enhancement upon substituting one H in XCHZ by a  $\text{CH}_3$  group, an  $\text{NH}_2$  group, or a halogen atom (F, Cl, Br). Remarkably, it is found that the electron-donating groups make the complex more stable than the electron-withdrawing groups. Furthermore, the results from NBO analysis illustrate that all the  $\text{O-H}\cdots\text{Se/Te}$  hydrogen bonds are redshifted, resulting from a noticeable electron density increase of  $\sigma^*(\text{O-H})$  orbital following complexation. The  $\text{C}_{\text{sp}^2}\text{-H}$  bond in the  $\text{C-H}\cdots\text{O/Se/Te}$  is transferred from the blueshift to redshift in ordering O and Se and then to Te substitution. This result is suggested by a slight decrease of electron density at  $\sigma^*(\text{C}_{\text{sp}^2}\text{-H})$  orbital and a rise in s-character percentage of C in  $\text{C}_{\text{sp}^2}\text{-H}$  bonds upon complexation.

## Acknowledgements

This research is funded by Vietnam Ministry of Education and Training (MOET) under grant number B2024-DQN-01.

## REFERENCES

1. G. A. Jeffrey, W. Saenger. *Hydrogen bonding in biological structures*, Springer-Verlag Berlin Heidelberg, 1991.
2. S. J. Grabowski. *Hydrogen bonding – new insights*, In J. Leszczynski (Ed.), *Series challenges and advances in computational chemistry and physics*, New York, NY Springer, 2006.
3. J. D. Olivier, F. Marc, C. Enric. Activation of C–H...halogen (Cl, Br, and I) hydrogen bonds at the organic/inorganic interface in fluorinated tetrathiafulvalenes salts, *Chemistry A European Journal*, **2001**, 7(12), 2635.
4. J. J. J. Dom, B. Michielsen, B. U. W. Maes, W. A. Herrebout, B. J. V. D. Veken. The C–H... $\pi$  interaction in the haloethane/ethene complex: a cryosolution infrared and Raman study, *Chemical Physics Letters*, **2009**, 469, 85.
5. J. M. H. Ramon, A. M. Grana. Blue-shifting hydrogen bond in the benzene–benzene and benzene–naphthalene complexes, *Journal of Computational Chemistry*, **2007**, 28, 540.
6. C. D. Keefe, M. Isenor. Ab initio study of the interaction of CHX<sub>3</sub> (X = H, F, Cl, or Br) with benzene and hexafluorobenzene, *The Journal of Physical Chemistry A*, **2008**, 112, 3127.
7. L. S. Sremaniak, J. L. Whitten, M. J. Truitt, J. L. White. Weak hydrogen bonding can initiate alkane C–H bond activation in acidic zeolites, *The Journal of Physical Chemistry B*, **2006**, 110(42), 20762.
8. R. E. Plata, D. E. Hill, B. E. Haines, D. G. Musaev, L. Chu, D. P. Hickey, M. S. Sigman, J. Q. Yu, D. G. Blackmond. A Role for Pd(IV) in catalytic enantioselective C–H functionalization with monoprotected amino acid ligands under mild conditions, *Journal of the American Chemical Society*, **2017**, 139 (27), 9238.
9. P. R. Shirhatti, S. Wategaonkar. Blue shifted hydrogen bond in 3-methylindole-CHX<sub>3</sub> complexes (X = Cl, F), *Physical Chemistry Chemical Physics*, **2010**, 12, 6650.
10. P. R. Shirhatti, D. K. Maity, S. Wategaonkar. C–H...Y hydrogen bonds in the complexes of p-cresol and p-cyanophenol with fluoroform and chloroform, *The Journal of Physical Chemistry A*, **2013**, 117, 2307.
11. W. A. Herrebout, S. M. Melikova, S. N. Delanoye, K. S. Rutkowski, D. N. Shchepkin, B. J. V. D. Veken. A cryosolution infrared study of the complexes of fluoroform with ammonia and pyridine: evidence for a C–H...N pseudo blueshifting hydrogen bond, *The Journal of Physical Chemistry A*, **2005**, 109, 3038.
12. B. G. Oliveira, R. C. M. U. D. Araújo, M. N. Ramos. A theoretical study of blueshifting hydrogen bonds in  $\pi$  weakly bound complexes, *Journal of Molecular Structure (Theochem.)*, **2009**, 908, 79.
13. R. Gopi, N. Ramanathan, K. Sundararajan. Experimental evidence for blue-shifted hydrogen bonding in the fluoroform–hydrogen chloride complex: a matrix-isolation infrared and ab initio study, *The Journal of Physical Chemistry A*, **2014**, 118, 5529.
14. B. Behera, P. K. Das. Blue-shifted hydrogen bonding in the gas phase CH/D<sub>3</sub>CN...HCCl<sub>3</sub> complexes, *The Journal of Physical Chemistry A*, **2019**, 123, 1830.
15. I. S. Sosulin, E. S. Shiryeva, D. A. Tyurin, V. I. Feldman. Matrix isolation and ab initio study on the CHF<sub>3</sub>...CO complex, *The Journal of Physical Chemistry A*, **2018**, 122, 4042.
16. N. T. Trung, P. N. Khanh, A. J. P. Carvalho, M. T. Nguyen. Remarkable shifts of C<sub>sp2</sub>-H and O-H stretching frequencies and stability of complexes of formic acid with formaldehydes and thioformaldehydes, *Journal of Computational Chemistry*, **2019**, 40, 1387.
17. R. Gopi, N. Ramanathan, N. Sundararajan. Blueshift of the C-H stretching vibration in CHF<sub>3</sub>-H<sub>2</sub>O complex: matrix isolation infrared spectroscopy and ab initio computations, *Chemical Physics*, **2016**, 476, 36.
18. Y. Mao, M. H. Gordon. Probing blue-shifting hydrogen bonds with adiabatic energy



- decomposition analysis, *The Journal of Physical Chemistry Letters*, **2019**, *10*, 3899.
19. C. Wang, D. Danovich, S. Shaik, Y. Mo. A unified theory for the blue- and red-shifting phenomena in hydrogen and halogen bonds, *Journal of Chemical Theory and Computation*, **2017**, *13*, 1626.
  20. Y. Mo, C. Wang, L. Guan, B. Braïda, P. C. Hiberty, W. Wu. On the nature of blueshifting hydrogen bonds, *Chemistry A European Journal*, **2014**, *20*, 8444.
  21. B. J. Mintz, J. M. Parks. Benchmark interaction energies for biologically relevant noncovalent complexes containing divalent sulfur, *The Journal of Physical Chemistry A*, **2012**, *116*, 1086.
  22. S. Sarkar, M. B. Bandyopadhyay. Cooperative nature of the sulfur centered hydrogen bond: investigation of  $(\text{H}_2\text{S})_n$  ( $n = 2-4$ ) clusters using an affordable yet accurate level of theory, *Physical Chemistry Chemical Physics*, **2019**, *21*, 25439.
  23. K. K. Mishra, S. K. Singh, P. Ghosh, D. Ghosh, A. Das. The nature of selenium hydrogen bonding: gas phase spectroscopy and quantum chemistry calculations, *Physical Chemistry Chemical Physics*, **2017**, *19*, 24179.
  24. K. K. Mishra, S. K. Singh, K. Kumar, G. Singh, B. Sarkar, M. S. Madhusudhan, A. Das. Water-mediated selenium hydrogen-bonding in proteins: pdb analysis and gas-phase spectroscopy of model complexes, *The Journal of Physical Chemistry A*, **2019**, *123*, 5995.
  25. V. R. Mundlapati, D. K. Sahoo, S. Ghosh, U. K. Purame, S. Pandey, R. Acharya, N. Pal, P. Tiwari, H. S. Biswal. Spectroscopic evidences for strong hydrogen bonds with selenomethionine in proteins, *The Journal of Physical Chemistry Letters*, **2017**, *8*, 794.
  26. J. Salon, J. Sheng, J. S. Jiang, G. X. Chen, J. C. Williams, Z. Huang. Oxygen replacement with selenium at the thymidine 4-position for the Se base pairing and crystal structure studies, *Journal of the American Chemical Society*, **2007**, *129*, 4862.
  27. R. W. Gora, S. J. Grabowski, J. Leszczynski. Dimers of formic acid, acetic acid, formamide and pyrrole-2 carboxylic acid: an ab initio study, *The Journal of Physical Chemistry A*, **2005**, *109*, 6397.
  28. R. W. Góra, M. Maj, S. J. Grabowski. Resonance-assisted hydrogen bonds revisited - Resonance stabilization vs. charge delocalization, *Physical Chemistry Chemical Physics*, **2013**, *15*, 2514.
  29. K. C. Asit, Z. H. Therese. Theoretical Investigation of the Cooperativity in  $\text{CH}_3\text{CHO}\cdots 2\text{H}_2\text{O}$ ,  $\text{CH}_2\text{FCHO}\cdots 2\text{H}_2\text{O}$ , and  $\text{CH}_3\text{CFO}\cdots 2\text{H}_2\text{O}$  Systems, *Journal of Atomic, Molecular, and Optical Physics*, **2012**, *0*, 1.
  30. K. Tabayashi, O. Takahashi, H. Namatame, M. Taniguchi. Substituent R-effects on the core-electron excitation spectra of hydrogen-bonded carboxylic-acid ( $\text{R}-\text{COOH}$ ) clusters: comparison between acetic-acid and formic-acid clusters, *Chemical Physics Letters*, **2013**, *557*, 1.
  31. T. S. Thakur, M. T. Kirchner, D. Bläser, R. Boese, G. R. Desiraju. Nature and strength of  $\text{C}-\text{H}\cdots\text{O}$  interactions involving formyl hydrogen atoms: computational and experimental studies of small aldehydes, *Physical Chemistry Chemical Physics*, **2011**, *13*, 14076.
  32. K. Damanjit, K. Rajinder. Theoretical characterization of hydrogen Bonding interactions between  $\text{RCHO}$  ( $\text{R} = \text{H}, \text{CN}, \text{CF}_3, \text{OCH}_3, \text{NH}_2$ ) and  $\text{HOR}'$  ( $\text{R}' = \text{H}, \text{Cl}, \text{CH}_3, \text{NH}_2, \text{C}(\text{O})\text{H}, \text{C}_6\text{H}_5$ ), *Journal of Chemical Sciences*, **2015**, *127*, 1299.
  33. Q. Zhang, L. Du. Hydrogen bonding in the carboxylic acid-aldehyde complexes, *Computational and Theoretical Chemistry*, **2016**, *1078*, 123.
  34. N. T. T. Cuc, N. T. An, V. T. Ngan, N. T. Trung. Importance of water and intramolecular interaction governs substantial blue shift of  $\text{C}_{\text{sp}^2}-\text{H}$  stretching frequency in complexes between chalcogenoaldehydes and water, *RSC Advances*, **2022**, *12*, 1998.
  35. N. T. An, V. T. Ngan, N. T. Trung. Profound importance of the conventional  $\text{O}-\text{H}\cdots\text{O}$  hydrogen bond versus a considerable blue shift of the  $\text{C}_{\text{sp}^2}-\text{H}$  bond in complexes of substituted carbonyls and carboxyl, *Physical Chemistry Chemical Physics*, **2024**, *26*, 22775.

36. Y. Yang, W. T. Zhang. Theoretical study on blue-shifted hydrogen bonds in  $\text{CH}_3\text{CHO}$  dimers, *Acta Chimica Sinica*, **2009**, 67, 599.
37. N. T. T. Cuc, P. D. C. Tu, N. T. A. Nhung, M. T. Nguyen, N. T. Trung, V. T. Ngan. Theoretical aspects of nonconventional hydrogen bonds in the complexes of aldehydes and hydrogen chalcogenides, *The Journal of Physical Chemistry A*, **2021**, 125(48), 10291.
38. L. T. T. Quyen, B. N. Tung, P. N. Thach, N. N. Tri, N. T. Trung. Characteristics of nonconventional hydrogen bonds and stability of dimers of chalcogenoaldehyde derivatives: a noticeable role of oxygen compared to other chalcogens, *RSC Advances*, **2024**, 14, 14114.
39. M. J. Frisch, G. W. Trucks, H. B. Schlegel, et al. *Gaussian 16 Rev. A. 03*, Gaussian Inc., Wallingford CT, 2016.
40. T. A. Keith. *AIMAll (Version 19.10.12)*, TK Gristmill Software, Overland Park KS, USA, 2019.
41. P. L. Popelier. *Atoms in molecules: an introduction*, Prentice Hall, London, 2000.
42. I. Mata, I. Alkorta, E. Espinosa, E. Molins. Relationships between interaction energy, intermolecular distance and electron density properties in hydrogen bonded complexes under external electric fields, *Chemical Physics Letters*, **2011**, 507, 185.
43. E. D. Glendening, J. K. Baderhoop, A. E. Read, J. E. Carpenter, J. A. Bohmann and F. Weinhold. *GenNBO 5.G*, Theoretical Chemistry Institute, University of Wisconsin Madison, WI, 2001.
44. F. Weinhold, C. R. Landis. Natural bond orbitals and extensions of localized bonding concepts, *Chemistry Education Research and Practice*, **2001**, 2, 91.
45. J. D. Chai, M. H. Gordon. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections, *Physical Chemistry Chemical Physics*, **2008**, 10, 6615.
46. J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, N. J. Russ, M. L. Leininger, C. L. Janssen, E. T. Seidl, Q. D. Allen, H. F. Schaefer, R. A. King, E. F. Valeev, C. D. Sherrill, T. D. Crawford. Psi4: an open-source ab initio electronic structure program, *WIREs Computational Molecular Science*, **2012**, 2(4), 556.
47. S. Bedoura, H. W. Xi, K. H. Lim. Hydrogen bond nature in formamide ( $\text{CYHNH}_2 \cdots \text{XH}$ ;  $\text{Y} = \text{O}, \text{S}, \text{Se}, \text{Te}$ ;  $\text{X} = \text{F}, \text{HO}, \text{NH}_2$ ) complexes at their ground and low-lying excited states, *Journal of Physical Organic Chemistry*, **2014**, 27(3), 226.
48. N. T. An, N. T. Duong, N. N. Tri, N. T. Trung. Role of  $\text{O}-\text{H} \cdots \text{O/S}$  conventional hydrogen bonds in considerable  $\text{C}_{\text{sp}^2}-\text{H}$  blueshift in the binary systems of acetaldehyde and thioacetaldehyde with substituted carboxylic and thiocarboxylic acids, *RSC Advances*, **2022**, 12(54), 35309.
49. E. V. Dornshuld, C. M. Holy, G. S. Tschumper. Homogeneous and heterogeneous noncovalent dimers of formaldehyde and thioformaldehyde: structures, energetics, and vibrational frequencies, *The Journal of Physical Chemistry A*, **2014**, 118, 3376.
50. I. V. Alabugin, S. Bresch, M. Manoharan. Hybridization trends for main group elements and expanding the Bent's rule beyond carbon: more than electronegativity, *The Journal of Physical Chemistry A*, **2014**, 118, 3663.
51. G. A. Dolgonos. Which isomeric form of formaldehyde dimer is the most stable – a high-level coupled-cluster study, *Chemical Physics Letters*, **2013**, 585, 37.



© 2025 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).





# Dự đoán hiệu quả và có thể giải thích tương tác protein-ligand sử dụng mô hình Gradient Boosting và AI có thể giải thích

Trương Thị Cẩm Mai\*

Khoa Khoa học Tự nhiên, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 25/11/2024; Ngày sửa bài: 24/01/2025;

Ngày nhận đăng: 04/02/2025; Ngày xuất bản: 28/02/2025

## TÓM TẮT

Dự đoán khả năng liên kết của các phân tử nhỏ với các mục tiêu protein là một bước quan trọng trong quá trình khám phá thuốc hiện đại, mở ra tiềm năng đẩy nhanh việc xác định các liệu pháp điều trị hiệu quả đồng thời giảm chi phí thí nghiệm. Trong nghiên cứu này, chúng tôi sử dụng bộ dữ liệu BELKA, một thư viện hóa học mã hóa bằng DNA (DEL) quy mô lớn, để huấn luyện các mô hình học máy nhằm dự đoán khả năng liên kết. Bằng cách áp dụng XGBoost, một thuật toán gradient boosting dựa trên cấu trúc cây quyết định, cùng với các bước tiền xử lý và thiết kế đặc trưng chuyên sâu, chúng tôi đã phát triển các mô hình dự đoán cho ba mục tiêu protein: BRD4, HSA, và SEH để dự đoán khả năng liên kết phân tử cho ba mục tiêu protein. Các mô hình này thể hiện năng lực dự đoán mạnh mẽ, đồng thời cho phép giải thích kết quả thông qua phân tích SHAP nhằm xác định các đặc trưng phân tử quan trọng quyết định khả năng liên kết. Đánh giá trên bộ dữ liệu kiểm tra BELKA cho thấy những thách thức trong việc khái quát hóa, cung cấp những hiểu biết quý giá về sự phức tạp của mô hình dự đoán trong khám phá thuốc. Nghiên cứu này nhấn mạnh tiềm năng của học máy trong việc thúc đẩy quá trình khám phá thuốc bằng máy tính, cho phép khám phá không gian hóa học hiệu quả hơn để tìm kiếm các liệu pháp điều trị tiềm năng.

**Từ khóa:** Khám phá thuốc, học máy, AI có thể giải thích.

\*Tác giả liên hệ chính.

Email: [truongcammai@qnu.edu.vn](mailto:truongcammai@qnu.edu.vn)

# Efficient interpretable prediction of protein-ligand interactions using Gradient Boosting models and explainable AI

Truong Thi Cam Mai\*

*Faculty of Natural Sciences, Quy Nhon University, Vietnam*

*Received: 25/11/2024; Revised: 24/01/2025;*

*Accepted: 04/02/2025; Published: 28/02/2025*

## ABSTRACT

The prediction of small molecule binding affinity to protein targets is a critical step in modern drug discovery, offering the potential to accelerate the identification of effective therapeutics while reducing experimental costs. In this study, we employ the BELKA dataset, a large-scale DNA-encoded chemical library (DEL), to train machine learning models for binding affinity prediction. Using XGBoost, a tree-based gradient boosting algorithm, and extensive preprocessing and feature engineering, we develop predictive models for three protein targets: BRD4, HSA, and sEH to predict whether a given small molecule is a binder or not to one of three protein targets. The models demonstrate strong predictive capabilities, with interpretability achieved through SHAP analysis to identify molecular features driving binding predictions. Evaluation of the BELKA test dataset reveals challenges in generalization, providing valuable insights into the complexities of predictive modelling in drug discovery. This work highlights the promise of machine learning in advancing computational drug discovery by enabling efficient exploration of the chemical space for potential therapeutics.

**Keywords:** *Drug discovery, machine learning, explainable artificial intelligence.*

## 1. INTRODUCTION

The development of machine learning (ML) models to predict the binding affinity of small molecules to specific protein targets holds transformative potential for drug discovery. Predicting these interactions is central to identifying new, effective drug candidates, as small molecule drugs interact with cellular protein machinery to influence disease-associated biological processes.

Traditionally, screening and testing small molecules for binding affinity to protein targets involve labour-intensive and costly

physical experiments, which severely limits the speed and scope of drug discovery efforts.<sup>1,2</sup> The search space for small molecule drugs is estimated to encompass approximately  $10^{60}$  chemical compounds, which is impractical to screen physically.<sup>3</sup>

With the pharmaceutical landscape evolving, the integration of ML-based predictive models offers a promising alternative to these conventional approaches, enabling efficient exploration of the vast chemical space for potential therapeutics. Traditional high-throughput screening (HTS) technologies can

---

\*Corresponding authors.

Email: [truongcammai@qnu.edu.vn](mailto:truongcammai@qnu.edu.vn)

assess libraries of small molecules against protein targets, but they are often restricted to collections of tens of thousands to a few million compounds.<sup>4</sup> In response to this limitation, DNA-encoded chemical libraries (DELs) have emerged as a more scalable solution.<sup>5</sup> DELs use unique DNA barcodes to tag each molecule, allowing the pooling of millions of compounds and simplifying the identification of binders through DNA sequencing. This method has substantially expanded the feasible scale of chemical libraries and presents an attractive foundation for computational models aimed at binding affinity prediction.

Advances in ML architectures and feature representation techniques, such as Simplified Molecular Input Line Entry System (SMILES) and graph-based molecular representations, have made it possible to capture complex chemical properties and interactions computationally.<sup>6</sup> SMILES, as a string-based molecular representation, encodes atom connectivity and stereochemistry, facilitating ML models' application in molecular property prediction, drug discovery, and materials design.

Hence, in this work, we explore the application of a tree-based gradient boosting approach, specifically XGBoost, for predicting binding affinity.<sup>7</sup> In addition to model development, an Explainable AI (XAI) method is integrated to interpret model behaviour, enhancing transparency and interpretability in the prediction of molecular binding. The findings from this study aim to contribute to the broader field of computational drug discovery, leveraging ML to identify promising drug candidates with high precision and potentially reduce the costs associated with traditional drug development methods. By enabling more efficient exploration of chemical space, this work aspires to pave the way toward discovering new lifesaving therapeutics for complex diseases. Conclusively, in this study, we make the following contributions:

- **Dataset Utilization:** We leverage the BELKA dataset<sup>8</sup>, a large-scale DNA-encoded chemical library, providing a comprehensive resource for binding affinity modelling.

- **Predictive Modelling:** We employ the XGBoost model optimized with advanced preprocessing and feature reduction techniques to predict binding affinities for three biologically significant protein targets: BRD4, HSA, and SEH.

- **Interpretability:** Through XAI analysis, we enhance the interpretability of the models, offering molecular-level insights into the features influencing binding predictions.

- **Benchmarking:** We evaluate our methodology on the BELKA dataset, highlighting the challenges of generalization for unseen cases.

## 2. RELATED WORK

### 2.1. Drug discovery and protein-target interactions

The pharmaceutical field relies heavily on understanding and predicting protein-target interactions, as these molecular interactions are critical in developing effective drugs. Small molecule drugs are typically designed to modulate specific protein targets linked to disease mechanisms. Protein-ligand binding is fundamental to this process, as the ability of a drug candidate to bind to a specific protein target determines its efficacy and safety.

Traditional drug discovery methodologies, such as high-throughput screening (HTS), involve synthesizing large libraries of small molecules and testing their affinity with the protein targets. However, HTS is costly, time-intensive, and limited in scope due to physical constraints, allowing only a fraction of potential drug-like compounds to be examined. Innovations, such as DNA-encoded chemical libraries (DELs)<sup>5</sup>, have addressed some of these limitations by enabling more extensive exploration of chemical space. In DELs, small molecules are tagged with unique

DNA barcodes, allowing millions of compounds to be screened in a pooled format. As such, DELs offer a scalable and efficient alternative to traditional HTS. Advances in molecular biology and DNA sequencing have further accelerated DEL technology, facilitating its adoption in both academia and industry.

## 2.2. SMILES and molecular representations

SMILES is one of the most widely adopted formats for encoding chemical structures in computational chemistry.<sup>6</sup> SMILES strings represent molecular structures in a linear form, capturing atoms, bonds, and stereochemistry in a machine-readable format. This notation has become essential for ML applications in drug discovery due to its simplicity and the ease with which it can be integrated into computational pipelines. SMILES can also be converted to other representations, such as 3D structures and molecular graphs, allowing flexibility in model input formats.

Alternative molecular representations, such as molecular fingerprints and molecular graphs, offer distinct advantages. Molecular fingerprints encode the presence or absence of substructures, providing a high-dimensional, fixed-length vector representation suited for various ML tasks.<sup>9–13</sup> Meanwhile, molecular graphs represent the connectivity of atoms in the molecule, capturing spatial information that can be valuable for models like graph neural networks (GNNs).<sup>14–17</sup> Recent studies suggest that combining multiple representations, such as SMILES with molecular graphs, can enhance predictive accuracy by leveraging diverse information formats.

## 2.3. Machine learning in molecular binding prediction

ML has become essential to molecular binding prediction, with recent models achieving high performance by leveraging large datasets and sophisticated algorithms. ML models, especially deep learning (DL) frameworks, can capture complex relationships in chemical and biological

data, allowing them to predict molecular properties with increasing accuracy.

Traditional ML methods, such as quantitative structure-activity relationship (QSAR) models, relied on engineered molecular descriptors to predict binding affinity. Still, recent ML approaches enable the use of raw chemical representations such as SMILES and molecular graphs, reducing the need for extensive feature engineering.<sup>18–20</sup> Convolutional neural networks (CNNs)<sup>21</sup>, graph neural networks (GNNs)<sup>22</sup>, and recurrent neural networks (RNNs)<sup>23</sup> have been widely used to encode molecular structures.

In addition to DL approaches, gradient-boosting algorithms like XGBoost have gained recognition for their efficacy in molecular property prediction. XGBoost suits tasks involving structured, high-dimensional data, such as molecular fingerprints. By leveraging an ensemble of decision trees, XGBoost iteratively refines predictions, minimizing error while maintaining interpretability. Unlike DL models, XGBoost offers a computationally efficient alternative that is well-suited for datasets with tabular or fingerprint-based representations. Recent studies have shown that integrating molecular representations, such as Extended-Connectivity Fingerprints (ECFPs)<sup>24</sup> with XGBoost, yields highly accurate binding affinity predictions while retaining transparency. These models are particularly valuable in scenarios where interpretability is crucial, such as drug discovery pipelines.<sup>25–27</sup> Additionally, XGBoost's robustness to overfitting, especially when combined with appropriate feature selection and regularization, makes it a strong candidate for handling imbalanced datasets often encountered in molecular binding tasks.

## 2.4. Explainability in ML for drug discovery

As ML models become increasingly complex, understanding the decision-making process within these models is critical for their adoption in sensitive fields like drug discovery. XAI methods aim to make the behaviour of complex ML



models more interpretable by providing insights into how input features influence predictions. In drug discovery, XAI can offer insights into which molecular features contribute most significantly to binding affinity, helping chemists understand and validate model predictions.<sup>28–30</sup>

While tree-based models such as XGBoost provide inherent interpretability through their structured decision paths, this interpretability becomes limited when applied to high-dimensional and large-scale datasets. In such cases, the complexity and volume of features make it challenging to discern the specific contributions of individual features to each prediction, thereby reducing the practical utility of feature importance metrics provided by these models. To address this limitation, XAI methods, like SHAP (SHapley Additive exPlanations)<sup>31</sup> values and LIME (Local Interpretable Model-agnostic Explanations)<sup>32</sup>, are often applied to these models, enabling the decomposition of predictions into contributions from individual features. For example, SHAP values, derived from cooperative game theory, were especially used to quantify each feature's influence on the prediction.<sup>33</sup> These explanation methods not only facilitate model interpretation but also foster trust in ML predictions, an essential factor for the integration of AI into pharmaceutical workflows.

### 3. DATASET

The BELKA dataset used in this study comprises training and test samples that detail the interactions between various small molecules and three protein targets: bromodomain-containing protein 4 (BRD4), soluble epoxide hydrolase (EPHX2/sEH), and human serum albumin (ALB/HSA).<sup>8</sup> The dataset presents a binary classification of whether a given small molecule is a binder or not to one of three protein targets.

#### 3.1. Dataset targets

The BELKA dataset encompasses three distinct protein targets: BRD4, EPHX2/sEH, and ALB/

HSA. Each target represents a unique class of biomolecular interactions, selected to provide a diverse benchmarking ground for modelling small molecule-protein interactions. These targets were carefully chosen for their biological significance and existing therapeutic relevance. Their acquisition and preparation followed rigorous protocols to ensure data fidelity and reproducibility.

##### 3.1.1. BRD4

Bromodomain-containing protein 4 is a pivotal member of the BET protein family, involved in recognizing acetylated lysines on histone tails.<sup>34</sup> BRD4 has emerged as a prominent therapeutic target in oncology, with inhibitors designed to disrupt its role in transcriptional regulation, particularly in cancer proliferation pathways. Recombinant BRD4 was acquired through baculovirus expression in insect cells to preserve post-translational modifications critical for its bromodomain function. Protein purity and structural integrity were validated through size-exclusion chromatography and binding assays with known BRD4 inhibitors. These quality-control measures ensured that the BRD4 used in DEL screenings retained its native binding characteristics, enabling high-confidence small molecule-protein interaction studies.

##### 3.1.2. BPHX2/sEH

Soluble epoxide hydrolase is an enzyme involved in metabolizing lipid epoxides, converting them into diols through hydrolysis.<sup>35</sup> This enzymatic activity has been implicated in numerous physiological and pathological processes, including inflammation, pain, and cardiovascular diseases. Recombinant human EPHX2 was expressed in *Escherichia coli* and purified via affinity chromatography. Its activity was verified using substrate-based fluorescence assays to confirm functional integrity before integration into DEL screening assays. By selecting sEH as a target, the BELKA dataset facilitates the evaluation of ligand binding in the context of enzymatic specificity and inhibition.

### 3.1.3. ALB/HSA

Human serum albumin, the most abundant plasma protein, plays a key role in drug pharmacokinetics by binding and transporting a wide range of endogenous and exogenous compounds.<sup>36</sup> For this dataset, HSA was isolated from human plasma and subjected to additional purification to remove potential impurities. Its binding activity was assessed through equilibrium dialysis and competitive ligand-binding assays to confirm its ability to interact with small molecules<sup>37</sup>. Using HSA in the DEL screening enables exploring protein-small molecule interactions that influence drug bioavailability and distribution.

### 3.2. Dataset acquisition

The raw readout acquisition process is visualized in Figure 1. The primary library, AMA014, is a triazine-based shree-cycle library designed to resemble DEL-A. An additional orthogonal DEL, termed *kinase0* (*kin0*), was designed to mimic kinase inhibitor chemistry.

The screening methodology involved combining the DEL with the target protein, isolating DEL/target complexes, eluting the bound DEL through heat application, and repeating the selection with the fresh target protein. This iterative process, conducted over three rounds for AMA014, aimed to enrich high-affinity binders. Each selection series for AMA014 was performed in triplicate to assess reproducibility. In contrast, the smaller *kinase0* library underwent a single selection round, performed in duplicate with a single negative control. Post-selection, the eluted DELs were subjected to sequencing to quantify binding events. The dataset includes both binary binding labels and raw sequencing counts, facilitating diverse analyses, including evaluating hit-calling methods and experimental design parameters. The raw dataset encompasses approximately 4.25 billion physical measurements, with compressed data totalling around 600 GB.

All protein targets underwent rigorous selection and preparation to maintain high experimental reproducibility. For each target, protein binding assays were conducted to confirm the enrichment of small molecule binders across multiple rounds of DEL screening. The screening workflow included initial binding assays with the target protein, iterative selection and amplification of enriched libraries, and sequencing to quantify binding events. These protocols were designed to capture high-affinity interactions and a broad spectrum of molecular binders, ensuring a comprehensive dataset for benchmarking predictive models.

### 3.3. Dataset description

Each row in the dataset encapsulates the chemical composition and binding characteristics of a small molecule with a specific protein target, providing a structured basis for learning binding patterns across different protein targets and molecular configurations.

The training dataset  $\mathcal{D}_{train}$  contains 98,415,610 samples and  $\mathcal{D}_{test}$  contains 878,022 samples. The training dataset  $\mathcal{D}_{train}$  (as shown in Table 1) includes molecular structures represented by SMILES strings, with each sample specifying four chemical building blocks, a complete molecular structure, the protein target, and the binary binding label (1 for binding, 0 for no binding) as the output variable. The test dataset  $\mathcal{D}_{test}$  follows a similar structure without the binding label, providing the molecular structure and target protein only. Each feature of the data is described as follows:

- **id:** A unique identifier for each record. Every unique combination of small molecule features is represented by three consecutive rows, each corresponding to a specific protein target: BRD4, HSA, or sEH. This structure allows for direct comparisons of binding affinity predictions across the three protein targets for the same molecular structure.

• **buildingblock1\_smiles:** A SMILES string representing the first building block of the molecule. This component forms part of the molecular structure and contributes specific chemical properties to the final molecule.

• **buildingblock2\_smiles:** A SMILES string for the second building block. Together with the first and third building blocks, it helps define the molecule's structure and potential binding characteristics.

• **buildingblock3\_smiles:** A SMILES string representing the third building block of the molecule, completing the combination of foundational elements used to form the final molecule.

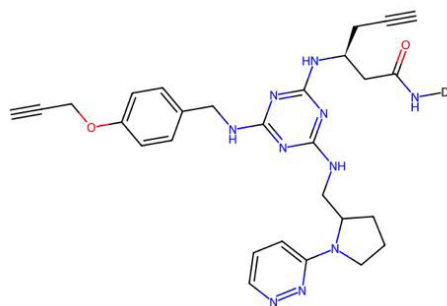
• **molecule\_smiles:** A SMILES string for the entire molecule, constructed from the building blocks and representing the complete molecular structure, including atoms, bonds, and stereochemistry. This column is a primary input for machine learning models to predict binding affinity based on the molecule's overall chemical structure.

• **protein\_name:** The name of the protein target for each molecule, which can be one of three values—BRD4, HSA, or sEH. Each protein

target has a specific biological significance and is used to determine the binding affinity of the molecule to a particular protein. For each unique molecule, the dataset includes rows for all three proteins to allow cross-target comparisons.

• **binds** ( $\mathcal{D}_{train}$  only): A binary label indicating whether the molecule binds to the specified protein target. A value of '1' signifies that the molecule binds to the target, while '0' indicates no binding. This label is used as the output variable  $y$ .

For instance, the 2D representation of a molecule in the BELKA dataset is demonstrated in Figure 1.



**Figure 1.** The 2D representation of a BELKA molecule (C#CCOc1ccc(CNc2nc(NCC3CCCN3c3cccn3)nc(N[C@@H](CC#C)CC(=O)N[Dy])n2)cc1).

**Table 1.** Training dataset excerpts for 3 targets: BRD4, HSA, and sEH.

| id | buildingblock1_smiles                                        | buildingblock2_smiles             | buildingblock3_smiles               | molecule_smiles                                                                    | protein_name | binds |
|----|--------------------------------------------------------------|-----------------------------------|-------------------------------------|------------------------------------------------------------------------------------|--------------|-------|
| 0  | <chem>C#CC[C@@H](CC(=O)O)NC(=O)OCC1c2ccccc2-c2ccccc21</chem> | <chem>C#CCOc1ccc(CN)cc1.Cl</chem> | <chem>Br.Br.NCC1CCCN1c1cccn1</chem> | <chem>C#CCOc1ccc(CNc2nc(NCC3CCCN3c3cccn3)nc(N[C@@H](CC#C)CC(=O)N[Dy])n2)cc1</chem> | BRD4         | 0     |
| 1  | <chem>C#CC[C@@H](CC(=O)O)NC(=O)OCC1c2ccccc2-c2ccccc21</chem> | <chem>C#CCOc1ccc(CN)cc1.Cl</chem> | <chem>Br.Br.NCC1CCCN1c1cccn1</chem> | <chem>C#CCOc1ccc(CNc2nc(NCC3CCCN3c3cccn3)nc(N[C@@H](CC#C)CC(=O)N[Dy])n2)cc1</chem> | HSA          | 0     |
| 2  | <chem>C#CC[C@@H](CC(=O)O)NC(=O)OCC1c2ccccc2-c2ccccc21</chem> | <chem>C#CCOc1ccc(CN)cc1.Cl</chem> | <chem>Br.Br.NCC1CCCN1c1cccn1</chem> | <chem>C#CCOc1ccc(CNc2nc(NCC3CCCN3c3cccn3)nc(N[C@@H](CC#C)CC(=O)N[Dy])n2)cc1</chem> | sEH          | 0     |

## 4. METHODOLOGY

### 4.1. Data preprocessing

In the data preprocessing phase, as demonstrated in Figure 2, the dataset was processed in increments of  $10^4$  rows to manage memory efficiently, given its large size. Each molecule in the dataset, represented by SMILES strings, was processed to create ECFPs, a commonly used molecular representation in cheminformatics. The SMILES strings for each molecule were converted into RDKit molecular objects, and ECFPs were generated with a radius of 2 and a fingerprint size of 2048 bits. The ECFPs were transformed into sparse matrix format to optimize memory usage, and additional bit information for each fingerprint was captured to enhance interpretability.

To reduce the dimensionality and improve computational efficiency, each of the building blocks (as shown in Section 3), namely *buildingblock1\_smiles*, *buildingblock2\_smiles* and *buildingblock3\_smiles* were mapped to unique integer identifiers, with dictionaries created for each set of SMILES strings.

For instance, building blocks in *buildingblock1\_smiles* were mapped to integer values in *blocks\_dict\_1*, while a shared dictionary, *blocks\_dict\_23*, was created for *buildingblock2\_smiles* and *buildingblock3\_smiles* due to the overlap between these blocks. These mappings were then saved, allowing for efficient lookup and reuse. To address the class imbalance, particularly given the scarcity of positive binding cases, the dataset was downsampled for non-binding entries, retaining all rows where the binding was detected and sampling a subset of non-binding cases. This balanced dataset provided an optimal size and improved training stability.

Processed data, including the sparse ECFP matrices and integer-encoded building

blocks, were saved into a training balanced set  $\mathcal{E}_{train}$  and a test set  $\mathcal{E}_{test}$ , in compressed formats for efficient storage and retrieval. This preprocessing pipeline allowed for structured and memory-efficient representation of the dataset, supporting effective model development for binding prediction across the three protein targets.

### 4.2. Model implementation

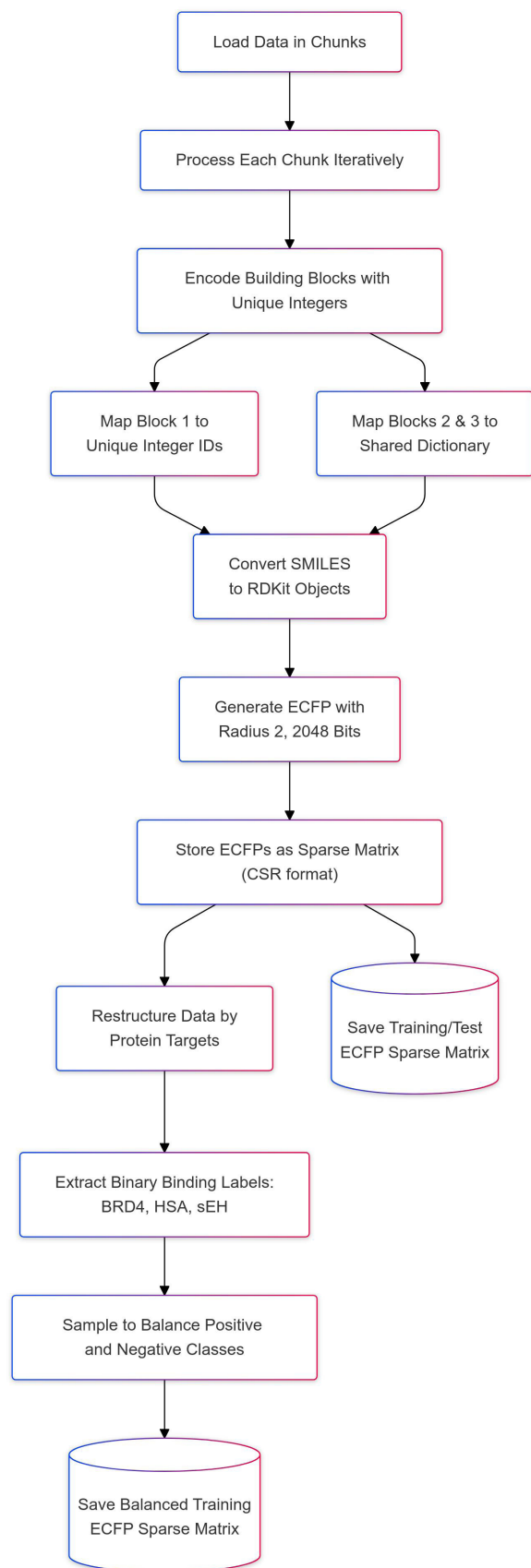
In this section, we described our implementation of a multi-step model training and evaluation process to predict the binding affinity of small molecules to three protein targets: BRD4, HSA, and sEH. This approach involved model selection, feature reduction, model training and evaluation.

#### 4.2.1. Training setup and data partitioning

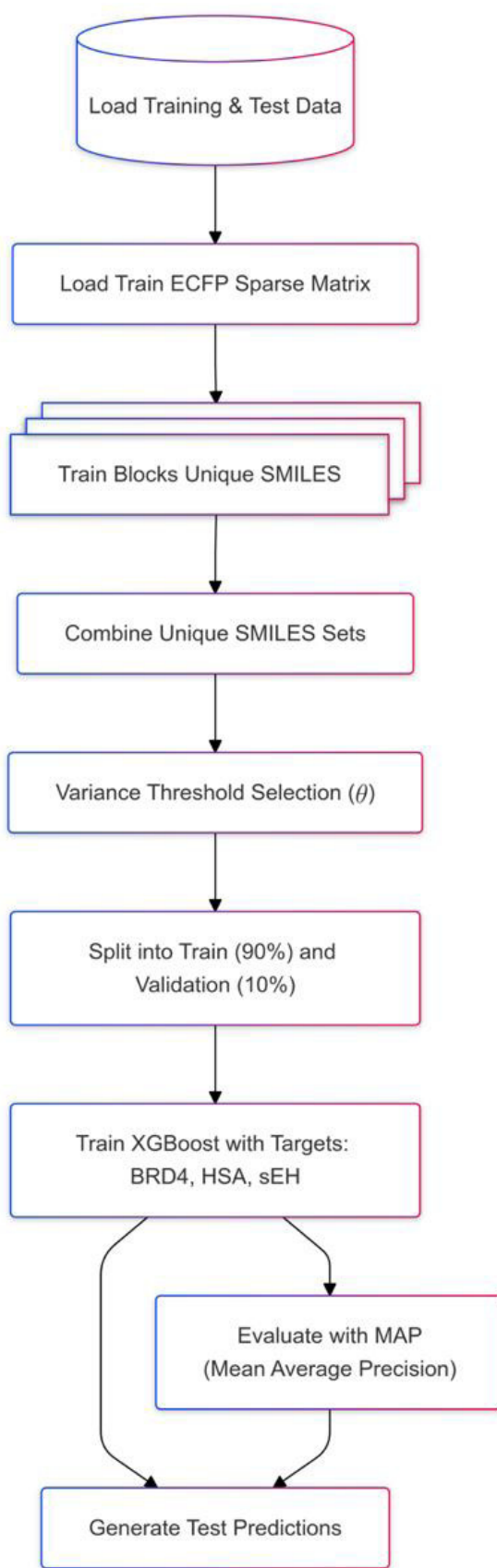
To control randomization across the training process, we initialized a fixed seed as 42, allowing for reproducibility in sampling and shuffling steps. The dataset  $\mathcal{E}_{train}$  was then split into a training set  $\mathcal{A}_{train}$  (90%) and a validation set  $\mathcal{A}_{val}$  (10%) based on a shuffled index of samples. This partitioning enabled model tuning on the training set while using the  $\mathcal{A}_{val}$  to assess model generalization and prevent overfitting.

#### 4.2.2. Feature reduction by variance threshold

The initial input data contained high-dimensional molecular fingerprints generated from ECFPs. To reduce dimensionality and enhance model performance, we applied a variance threshold to the ECFP feature matrix. Features with variance  $\theta$  below 0.005 were removed, as low-variance features contribute minimally to distinguishing between classes. This filtering reduced computational complexity and mitigated overfitting by retaining only the most informative features.



**Figure 2.** The pipeline of preprocessing dataset.



**Figure 3.** The pipeline of model training, validation and inference.



#### 4.2.3. Model training and inference

In this step, we employed XGBoost, a gradient boosting algorithm, to train separate binary classification models for each protein target: BRD4, HSA, and sEH.

For each model, we configured the objective function as binary logistic regression with a learning rate  $r$  of 0.2, and the evaluation metric as the average precision score (AP). Early stopping was applied with a patience  $p$  of 100 rounds to prevent overfitting, and the model was allowed up to 4000 iterations for convergence. To handle imbalanced data, we computed a scale positive weight for each target, defined as

$$\omega = \frac{N_{neg}}{N_{pos}}, \text{ where } N_{neg} \text{ and } N_{pos} \text{ represent the}$$

counts of non-binding and binding samples, respectively.

#### 4.3. Model explanation with XAI

To understand the contribution of specific molecular features to each model's predictions, we applied an interpretability method, namely SHAP.

We utilized the SHAP TreeExplainer for XGBoost, which computes Shapley values efficiently in tree-based models. SHAP summary plots and bar charts were generated to visualize the global importance of features in predicting binding affinity for each target.

### 5. RESULTS

#### 5.1. Model performance

To evaluate the effectiveness of our predictive models on the binding affinity classification task, we conducted a comprehensive performance assessment across the three protein targets: BRD4, HSA, and sEH. The metrics used for evaluation included accuracy, mean average precision (MAP), recall, and the area under the precision-recall curve (AUCPR), which are presented in Table 1. Results were reported separately for the training set  $\mathcal{A}_{train}$  and the validation set  $\mathcal{A}_{val}$  to provide insights into training stability and generalization.

#### 5.1.1. Evaluation metrics

**Accuracy** measures the overall correctness of predictions and is defined as the ratio of correctly classified samples (both positive and negative) to the total number of samples. Mathematically, accuracy is expressed as:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP represents true positives, TN represents true negatives, FP represents false positives, and FN represents false negatives. While accuracy provides a general assessment of the model's classification capability, it can be less informative in imbalanced datasets, as it may overemphasize the correct classification of the majority class.

**Recall (Sensitivity)** measures the proportion of actual positive cases correctly identified by the model. It is defined as:

$$\text{Recall} = \frac{TP}{TP + FN}$$

Recall focuses on the model's ability to capture true binders, which is critical in applications where missing positive cases (e.g., potential drug candidates) could have significant consequences. A high recall ensures that the model effectively identifies most true binding interactions.

**Mean Average Precision (MAP)** evaluates the ranking quality of predictions, particularly the precision of positive cases across various thresholds. It is calculated as the mean of the Average Precision (AP) scores over all classes, where AP combines precision and recall into a single metric that emphasizes the ranking order of positive predictions. MAP is computed as:

$$\text{MAP} = \frac{1}{N} \sum_{i=1}^N \text{AP}(y_{\text{true}}^i, y_{\text{pred}}^i)$$

where  $R_k$  is the recall at rank  $k$  and  $P_k$  is the precision at rank  $k$ . MAP is particularly valuable for imbalanced datasets, as it prioritizes the correct ranking of true positives, making it

sensitive to the quality of predictions for the minority class.

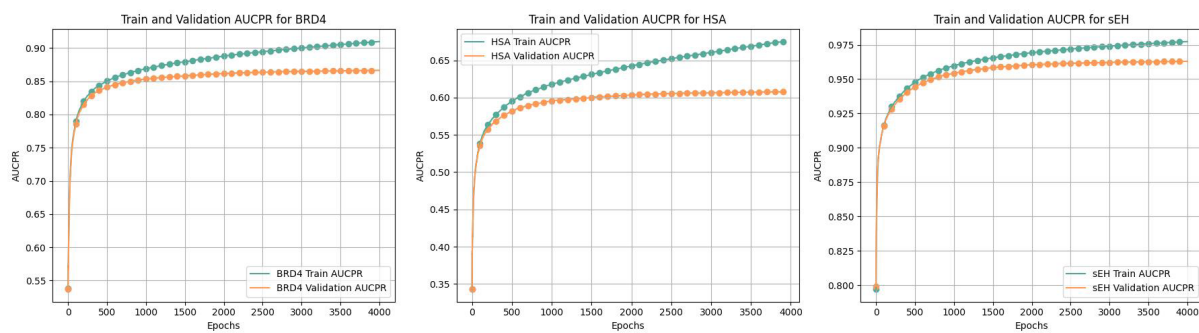
**Area Under the Precision-Recall Curve (AUCPR)** quantifies the trade-off between precision and recall across all decision thresholds. Unlike the Receiver Operating Characteristic (ROC) curve, the Precision-Recall curve is more informative in imbalanced datasets, as it emphasizes the model's ability to classify the positive class correctly. AUCPR is calculated as the area under the curve formed by plotting precision against recall at varying thresholds. A higher AUCPR indicates a better

balance between precision and recall, reflecting the model's ability to maintain high sensitivity (recall) without compromising specificity (precision).

These four metrics were chosen to provide a comprehensive evaluation of the models, capturing their overall classification accuracy, ranking quality, sensitivity to true positives, and the precision-recall trade-off. By analyzing these metrics, we can gain deeper insights into the strengths and limitations of the models for each protein target, enabling targeted improvements in future iterations.

**Table 2.** The model performance on the training and validation set with accuracy, mean average precision (MAP), recall and area under the precision-recall curve (AUCPR).

|                       | BRD4     | HSA    | sEH    | BRD4   | HSA    | sEH    |
|-----------------------|----------|--------|--------|--------|--------|--------|
|                       | Accuracy |        |        | MAP    |        |        |
| $\mathcal{A}_{train}$ | 0.9637   | 0.9164 | 0.9798 | 0.5708 | 0.3341 | 0.7913 |
| $\mathcal{A}_{val}$   | 0.9583   | 0.9082 | 0.9767 | 0.5364 | 0.3006 | 0.7754 |
|                       | Recall   |        |        | AUCPR  |        |        |
| $\mathcal{A}_{train}$ | 0.9910   | 0.9543 | 0.9979 | 0.9098 | 0.6751 | 0.9773 |
| $\mathcal{A}_{val}$   | 0.9275   | 0.8467 | 0.9778 | 0.8663 | 0.6076 | 0.9629 |

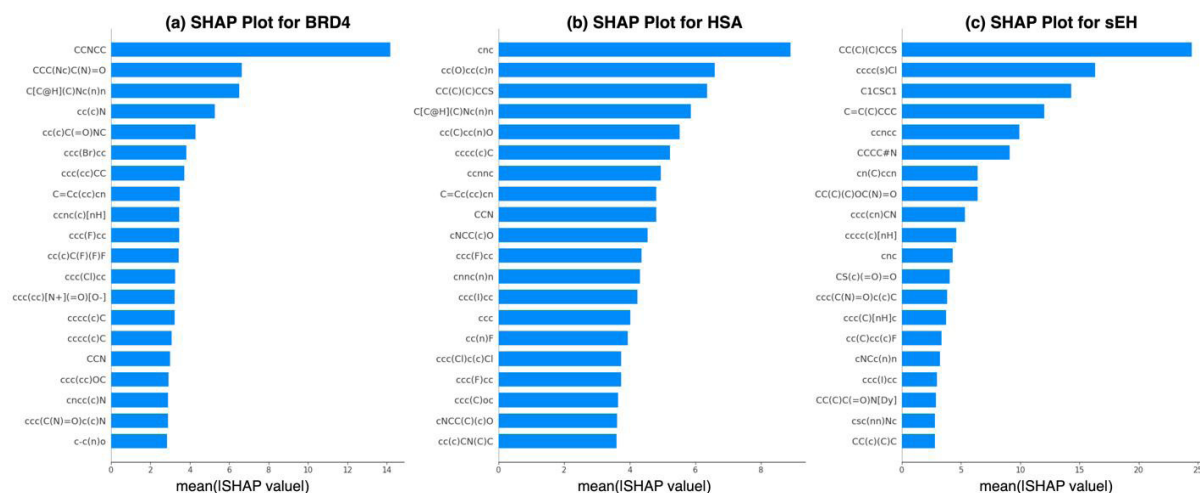


**Figure 4.** The area under the precision-recall curve (AUCPR) visualization of the model on the training (in green) and validation (in orange) set.

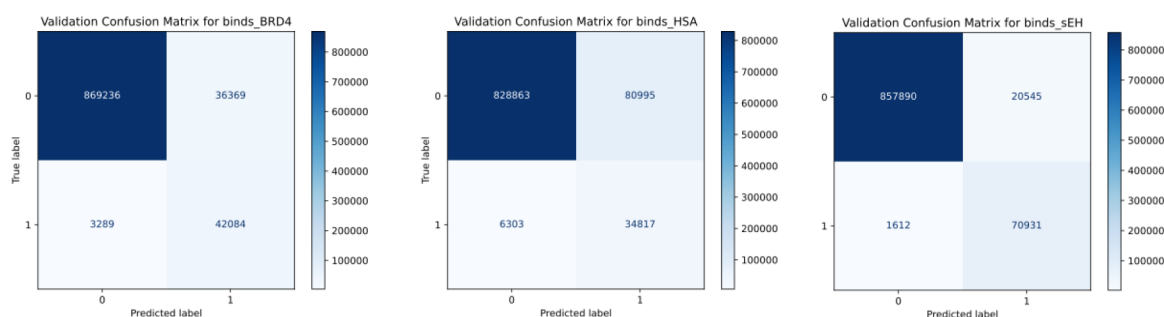
5.1.2. Performance evaluation

For BRD4, the model achieved a high accuracy of 0.9637 on  $\mathcal{A}_{train}$  and 0.9583 on  $\mathcal{A}_{val}$  indicating minimal overfitting and strong predictive performance. However, the MAP values, which assess the ranking quality of positive predictions, were relatively modest at 0.5708 for  $\mathcal{A}_{train}$  and 0.5364 for  $\mathcal{A}_{val}$  . This

reflects the inherent difficulty of ranking positive binders for BRD4. Despite this, the recall values were consistently high, reaching 0.9910 for  $\mathcal{A}_{train}$  and 0.9275 for  $\mathcal{A}_{val}$ , demonstrating the model's capability to identify a significant proportion of true binders. The AUCPR scores, 0.9098 for  $\mathcal{A}_{train}$  and 0.8663 for  $\mathcal{A}_{val}$  , further confirms the model's effectiveness in differentiating binders from non-binders.



**Figure 5.** The global explanation (top influencing features) for the model's prediction of the binding affinity for 3 targets: BRD4, HSA and sEH.



**Figure 6.** The confusion matrices for the model's prediction of the binding affinity for 3 targets: BRD4, HSA and sEH.

For HSA, the model exhibited slightly lower accuracy than BRD4, with values of 0.9164 for  $\mathcal{A}_{train}$  and 0.9082 for  $\mathcal{A}_{val}$ . The MAP scores for HSA, 0.3341 for  $\mathcal{A}_{train}$  and 0.3006 for  $\mathcal{A}_{val}$ , were the lowest among the three targets, indicating challenges in ranking true binders effectively. Nevertheless, the recall metrics for HSA were robust, achieving 0.9543 on  $\mathcal{A}_{train}$  and 0.8467 on  $\mathcal{A}_{val}$ . The AUCPR values, 0.6751 for  $\mathcal{A}_{train}$  and 0.6076 for  $\mathcal{A}_{val}$ , suggest the model's reasonable ability to identify binding patterns, though there is room for improvement in precision-recall balance.

The model's performance on sEH was the strongest overall. The accuracy reached 0.9798 for  $\mathcal{A}_{train}$  and 0.9767 for  $\mathcal{A}_{val}$ , showcasing exceptional classification accuracy. Similarly, the MAP scores, 0.7913 for  $\mathcal{A}_{train}$  and 0.7754

for  $\mathcal{A}_{val}$ , were significantly higher than those for BRD4 and HSA, indicating superior ranking performance. Recall values were near perfect at 0.9979 for  $\mathcal{A}_{train}$  and 0.9778 for  $\mathcal{A}_{val}$ , further emphasizing the model's sensitivity in detecting true binding events. The AUCPR scores, 0.9773 for  $\mathcal{A}_{train}$  and 0.9629 for  $\mathcal{A}_{val}$ , reinforce the robustness of the model for sEH, highlighting its capability to separate binders from non-binders with high confidence effectively.

In addition to evaluating model performance on the  $\mathcal{A}_{train}$  and  $\mathcal{A}_{val}$  sets, the final models were assessed on the BELKA test dataset  $\mathcal{E}_{test}$ , as part of a Kaggle competition. Since the test labels for each target were not made available, the evaluation relied solely on the predictions' final test scores. On the public test set, the model achieved an accuracy of 0.2042,

while the private test set yielded a slightly lower accuracy of 0.1843.

Finally, these results demonstrate that the boosting models, when combined with efficient data preprocessing and dimensionality reduction, can achieve reliable predictions across diverse protein targets. The differences in MAP and AUCPR scores among the targets underscore the varying complexities of binding prediction, with sEH being the most tractable and HSA presenting the greatest challenges. The results on the test datasets highlight the challenges posed by the BELKA dataset, particularly the difficulty in achieving generalizable predictions across unseen data. The gap between validation and test performance underscores the potential for further enhancements in model robustness and generalization.

## 5.2. Model interpretability

To provide insights into the decision-making process of the predictive models, we employed SHAP to quantify the contribution of individual molecular features to the model's output. SHAP explanations are particularly valuable in understanding which molecular substructures, represented as SMILES fragments, had the most significant impact on the binding affinity prediction for each protein target: BRD4, HSA, and sEH. The SHAP summary plots for the three targets are presented in **Error! Reference source not found.**, with the x-axis representing the mean absolute SHAP value, indicative of the average magnitude of a feature's impact on the model's predictions.

**BRD4 (Error! Reference source not found.a):** the most influential molecular feature was the fragment "CCNCC," which exhibited the highest mean SHAP value, highlighting its strong association with binding predictions. Other significant contributors included fragments with nitrogen and aromatic substructures such as "CCC(N)C(N)=O" and "C(c)(H)C(Nc=In)," suggesting that these groups may play a key role in interacting with BRD4's bromodomains.

Notably, the diversity of impactful features underscores the model's ability to capture complex molecular patterns that influence binding specificity.

**HSA (Error! Reference source not found.b):** SHAP analysis revealed "CC(C)(C)CCS" as the most impactful feature. This fragment aligns with HSA's known affinity for hydrophobic and bulky molecular groups, which are critical for its role as a drug carrier protein. Additional significant features included "ccccc1CCl" and "C1CSC1," suggesting a preference for aromatic and cyclic substructures. These insights provide a molecular-level understanding of the interactions influencing the binding of small molecules to HSA.

**sEH (Error! Reference source not found.c):** the SHAP summary plot demonstrated that the fragment "CC(C)(C)CCS" had the largest average impact on model predictions, followed by "ccccc1CCl" and "C=C(C)C(CC)." These features are consistent with known hydrophobic binding pockets in sEH, highlighting the model's ability to identify molecular characteristics critical for binding affinity. Notably, the sEH model exhibited a larger range of SHAP values than the other targets, reflecting a higher sensitivity to specific molecular fragments.

The SHAP analysis across all three targets highlights the models' reliance on chemically meaningful features, providing interpretability and transparency in their predictions. These findings not only enhance confidence in the models but also offer valuable insights for the rational design of small molecules with desired binding properties. Future efforts could involve leveraging these SHAP-derived insights for feature engineering or guiding experimental validation to further refine predictive performance.

## 6. DISCUSSION

This study demonstrates the potential and challenges of ML in molecular binding prediction. The XGBoost models achieved high



performance on training and validation datasets, particularly for sEH, which benefited from its consistent molecular binding patterns. However, BRD4 and HSA presented unique challenges due to more diverse binding chemistries, resulting in slightly lower scores. SHAP analysis revealed chemically meaningful features, providing valuable insights into the molecular determinants of binding and guiding potential drug design efforts. The evaluation of the BELKA test dataset highlights a notable performance drop, with public and private test scores of 0.2042 and 0.1843, respectively. This gap underscores the inherent difficulty of generalizing predictive models to unseen data in large, diverse chemical spaces. It also highlights the importance of robust feature selection, additional data augmentation, and more generalized learning methods to bridge the gap between validation and test performance. Hence, in our future work, we would like to investigate the influence of more feature selection, dimensional reduction and data augmentation techniques.

While tree-based models like XGBoost are interpretable and effective for structured data, the reliance on binary binding labels rather than continuous affinity scores limits their ability to capture nuanced interactions. Future work could integrate graph-based molecular representations or hybrid approaches combining DL with traditional ML to improve prediction accuracy and generalizability.<sup>15,17,38,39</sup> Additionally, leveraging semi-supervised learning or transfer learning could further enhance model robustness in unseen data scenarios.<sup>34-36</sup>

## 7. CONCLUSION

This study highlights the potential of machine learning to revolutionize drug discovery by predicting small molecule binding affinities with high efficiency. Using the BELKA dataset, we demonstrated the capability of XGBoost models to achieve strong predictive performance while providing interpretability through SHAP analysis. However, challenges in generalization, particularly on unseen test datasets, reveal

areas for methodological improvement. By combining robust predictive capabilities with interpretable outputs, this work advances computational approaches for drug discovery, enabling more efficient exploration of chemical space and paving the way for identifying novel therapeutics.

## REFERENCES

1. P. Arora, M. Behera, S. A. Saraf, R. Shukla. Leveraging artificial intelligence for synergies in drug discovery: from computers to clinics, *Current Pharmaceutical Design*, **2024**, 28(30), 2187-2205.
2. K. Sharma, P. Manchikanti. Artificial Intelligence in drug development and healthcare-nature and scope, *Artificial Intelligence in Drug Development: Patenting and Regulatory Aspects*, Springer, 2024, 1-33.
3. J. L. Reymond, R. V. Deursen, L. C. Blum, L. Ruddigkeit. Chemical space as a source for new drugs, *MedChemComm*, **2010**, 1, 30-38.
4. J. P. Landry, Y. Fei, X. Zhu. High throughput, label-free screening small molecule compound libraries for protein-ligands using combination of small molecule microarrays and a special ellipsometry-based optical scanner, *International Drug Discovery*, **2011**, 6(1), 8-13.
5. A. G. Martínez, E. J. Donckele, F. Samain, D. Neri. DNA-encoded chemical libraries: a comprehensive review with successful stories and future challenges, *ACS Pharmacology and Translational Science*, **2021**, 4(4), 1265-1279.
6. D. Weininger. SMILES, a chemical language and information system - introduction to methodology and encoding rules, *Journal of Chemical Information and Computer Sciences*, **1988**, 28, 31-36.
7. T. Chen, C. Guestrin. *Xgboost: a scalable tree boosting system*, Proceedings of the 22<sup>nd</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, 785-794.
8. A. Blevins, I. K. Quigley, B. J. Halverson, N. Wilkinson, R. S. Levin, A. Pulapaka, W. Reade, A. Howard. *NeurIPS 2024 - Predict New Medicines with BELKA*, 2024.



9. S. F. Baygi, D. K. Barupal. IDSL\_MINT: a deep learning framework to predict molecular fingerprints from mass spectra, *Journal of Cheminformatics*, **2024**, 16(1), 8.
10. D. Boldini, D. Ballabio, V. Consonni, R. Todeschini, F. Grisoni, S. A. Sieber. Effectiveness of molecular fingerprints for exploring the chemical space of natural products, *Journal of Cheminformatics*, **2024**, 16(1), 35.
11. D. K. Duvenaud, D. Maclaurin, J. Iparraguirre, R. Bombarell, T. Hirzel, A. A. Guzik, R. P. Adams, *Convolutional networks on graphs for learning molecular fingerprints*, Advances in Neural Information Processing Systems, 2015.
12. J. Wang, L. Zhang, J. Sun, X. Yang, W. Wu, W. Chen, Q. Zhao. Predicting drug-induced liver injury using graph attention mechanism and molecular fingerprints, *Methods*, **2024**, 221, 18-26.
13. N. T. T. Hung. *Analyzing and visualizing chemical molecular properties using K-Dimension tree algorithm*, Bachelor thesis, Frankfurt University of Applied Sciences, 2020.
14. S. Goldman, J. Li, C. W. Coley. Generating molecular fragmentation graphs with autoregressive neural networks, *Analytical Chemistry*, **2024**, 96(8), 3419-3428.
15. S. Kim, J. Woo, W. Y. Kim. Diffusion-based generative AI for exploring transition states from 2D molecular graphs, *Nature Communications*, **2024**, 15(1), 341.
16. K. D. Luong, A. K. Singh. *Fragment-based pretraining and finetuning on molecular graphs*, Advances in Neural Information Processing Systems, 2024.
17. A. Morehead, J. Cheng. Geometry-complete perceptron networks for 3d molecular graphs, *Bioinformatics*, **2024**, 40(2).
18. A. Z. Dudek, T. Arodz, J. Gálvez. Computational methods in developing quantitative structure-activity relationships (QSAR): a review, *Combinatorial Chemistry and High Throughput Screening*, **2006**, 9(3), 213-228.
19. T. A. Soares, A. N. Alves, A. Mazzolari, F. Ruggiu, G. W. Wei, K. Merz. The re-evolution of quantitative structure-activity relationship (QSAR) studies propelled by the surge of machine learning methods, *Journal of Chemical Information and Modeling*, **2022**, 62(22), 5317-5320.
20. M. Honma. An assessment of mutagenicity of chemical substances by (quantitative) structure-activity relationship, *Genes and Environment*, **2020**, 42(1), 23.
21. Y. L. Cun, Y. Bengio. Convolutional networks for images, speech, and time series, *The Handbook of Brain Theory and Neural Networks*, **1995**, 3361(10), 1995.
22. F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, G. Monfardini. The graph neural network model, *IEEE Transactions On Neural Networks*, **2008**, 20(1), 61-80.
23. D. E. Rumelhart, G. E. Hinton, R. J. Williams. Learning internal representations by error propagation, parallel distributed processing, explorations in the microstructure of cognition, *Biometrika*, **1986**, 71(6), 599-607.
24. D. Rogers, M. Hahn. Extended-connectivity fingerprints, *Journal of Chemical Information and Modeling*, **2010**, 50(5), 742-754.
25. T. R. Novianady, G. M. Idroes, I. Hardi. Machine learning approach to predict AXL kinase inhibitor activity for cancer drug discovery using XGBoost and Bayesian optimization, *Journal of Soft Computing and Data Mining*, **2024**, 5(1), 46-56.
26. S. H. Mahmud, W. Chen, H. Jahan, Y. Liu, N. I. Sujan, S. Ahmed. iDTi-CSsmoteB: identification of drug-target interaction based on drug chemical structure and protein sequence using XGBoost with over-sampling technique SMOTE, *IEEE Access*, **2019**, 7, 48699-48714.
27. D. Zhang, H. D. Chen, H. Zulfiqar, S. S. Yuan, Q. L. Huang, Z. Y. Zhang, K. J. Deng. iBLP: an XGBoost-based predictor for identifying bioluminescent proteins, *Computational and Mathematical Methods in Medicine*, **2021**, 2021(1), 6664362.
28. R. Alizadehsani, S. S. Oyelere, S. Hussain, S. K. Jagatheesaperumal, R. R. Calixto, M. Rahouti, M. Roshanzamir, V. H. C. D. Albuquerque. Explainable artificial intelligence for drug discovery and development-a comprehensive survey, *IEEE Access*, **2024**.

29. K. K. Kirboğa, S. Abbasi, E. U. Küçüksille. Explainability and white box in drug discovery, *Chemical Biology and Drug Design*, **2023**, 102(1), 217-233.
30. T. T. H. Nguyen, V. B. Truong, V. T. K. Nguyen, Q. H. Cao, Q. K. Nguyen. Towards trust of explainable ai in thyroid nodule diagnosis, *International Workshop on Health Intelligence*, Washington DC, USA, 2023.
31. S. M. Lundberg, S. I. Lee. *A unified approach to interpreting model predictions*, Proceedings of the 31<sup>st</sup> International Conference on Neural Information Processing Systems, Auckland, NewZealand, 2017.
32. M. T. Ribeiro, S. Singh, C. Guestrin. 'Why should i trust you?' Explaining the predictions of any classifier, Proceedings of the 22<sup>nd</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, USA, 2016.
33. T. Clement, H. T. T. Nguyen, N. Kemmerzell, M. Abdelaal, D. Stjelja. *Coping with data distribution shifts: XAI-Based adaptive learning with SHAP clustering for energy consumption prediction*, Australasian Joint Conference on Artificial Intelligence, Brisbane, Australia, 2023.
34. B. Padmanabhan, S. Mathur, R. Manjula, S. Tripathi. Bromodomain and extra-terminal (BET) family proteins: new therapeutic targets in major diseases, *Journal of Biosciences*, **2016**, 41, 295-311.
35. J. W. Newman, C. Morisseau, B. D. Hammock. Epoxide hydrolases: their roles and interactions with lipid metabolism, *Progress in Lipid Research*, **2005**, 44, 1-51.
36. A. Varshney, P. Sen, E. Ahmad, M. Rehan, N. Subbarao, R. H. Khan. Ligand binding strategies of human serum albumin: how can the cargo be utilized?, *Chirality: The Pharmacological, Biological, and Chemical Consequences of Molecular Asymmetry*, **2010**, 22, 77-87.
37. Y. Sun, Z. Ji, X. Liang, G. Li, S. Yang, S. Wei, Y. Zhao, X. Hu, J. Fan. Studies on the binding of rhaponticin with human serum albumin by molecular spectroscopy, modeling and equilibrium dialysis, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, **2012**, 87, 171-178.
38. F. E. Kelvinius, D. Georgiev, A. Toshev, J. Gasteiger. Accelerating molecular graph neural networks via knowledge distillation, *Advances in Neural Information Processing Systems*, **2024**, 36.
39. V. Fung, J. Zhang, E. Juarez, B. G. Sumpter. Benchmarking graph neural networks for materials chemistry, *npj Computational Materials*, **2021**, 7, 84.
40. Y. Cho, J. Shim, K. Seok. Binding affinity prediction using self-supervised learning and QP-ensemble, *QBS*, **2023**, 42, 57-64.
41. Y. Gu, X. Zhang, A. Xu, W. Chen, K. Liu, L. Wu, S. Mo, Y. Hu, M. Liu, Q. Luo. Protein-ligand binding affinity prediction with edge awareness and supervised attention, *Iscience*, **2023**, 26(1), 105892.
42. B. Tanoori, M. Z. Jahromi, E. G. Mansoori. Binding affinity prediction for binary drug-target interactions using semi-supervised transfer learning, *Journal of Computer-Aided Molecular Design*, **2021**, 35, 883-900.



© 2025 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

**MỤC LỤC**

|     |                                                                                                                                                                                                                                                |     |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 1.  | Nghiên cứu và ứng dụng mạng truyền thông công nghiệp trong hệ thống xử lý nước thải sinh hoạt tại thành phố Quy Nhơn<br><b>Đỗ Văn Cần, Trần Xuân Khoa</b> .....                                                                                | 5   |
| 2.  | Một số tiêu chuẩn mới về ổn định hóa đối với hệ tuyến tính từng phần hai mô hình có trễ thời gian<br><b>Nguyễn Hoàng Tường Vy, Lê Quang Thuận</b> .....                                                                                        | 17  |
| 3.  | Nghiên cứu thu thập dữ liệu và giám sát chất lượng điện áp sử dụng Ad-hoc LoRa và PLC Siemens thông qua mạng truyền thông Modbus TCP/IP<br><b>Trần Xuân Khoa, Huỳnh Đức Hoàn, Phan Gia Trí</b> .....                                           | 29  |
| 4.  | Toán tử hợp có trọng từ không gian kiểu Bloch vào không gian tăng trưởng trên hình cầu đơn vị của không gian Hilbert<br><b>Dương Thanh Vỹ</b> .....                                                                                            | 49  |
| 5.  | Khảo sát ứng dụng tinh dầu gừng trong bảo quản Bánh ít lá gai<br><b>Võ Thị Trọng Hoa, Lê Duy Thanh, Ngô Kim Khuê, Nguyễn Thanh Hằng, Lê Cảnh Định</b> .....                                                                                    | 61  |
| 6.  | Ảnh hưởng của độ sâu mực nước đến sinh trưởng và tỷ lệ sống của ốc bươu đồng ( <i>Pila polita</i> ) trong điều kiện ương<br><b>Ngô Kim Khuê, Võ Văn Chí</b> .....                                                                              | 71  |
| 7.  | Một điều kiện tương đương để một họ ma trận chuẩn tắc là chéo hóa tương đẳng đồng thời được<br><b>Võ Đức Tư Duy, Lê Thanh Hiếu</b> .....                                                                                                       | 79  |
| 8.  | Phân tích toàn diện phổ của berberin, hợp chất chính trong bài thuốc dân gian Việt Nam “Đại tràng hoàn Bà Giăng”<br><b>Nguyễn Hoàng Sa</b> .....                                                                                               | 87  |
| 9.  | Liên kết hydrogen không cổ điển $C_{sp^2}-H...O/Se/Te$ and $O-H...Se/Te$ trong các hệ phức giữa acid formic acid và selenoformaldehyde, telluroformaldehydes<br><b>Nguyễn Tiến Trung, Nguyễn Trường An, Nguyễn Ngọc Trí, Vũ Thị Ngân</b> ..... | 99  |
| 10. | Dự đoán hiệu quả và có thể giải thích tương tác protein-ligand sử dụng mô hình Gradient Boosting và AI có thể giải thích<br><b>Trương Thị Cẩm Mai</b> .....                                                                                    | 115 |

