TRƯỜNG ĐẠI HỌC QUY NHƠN
QUY NHON UNIVERSITY

TẠP CHÍ KHOA HỌC
JOURNAL OF SCIENCE

CHUYÊN SAN KHOA HỌC TỰ NHIÊN VÀ KỸ THUẬT
ISSUE: NATURAL SCIENCES AND ENGINEERING

16 (1)
2022

BÌNH ĐỊNH, 02/2022

MỤC LỤC

TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC QUY NHƠN

Chính quy mêtric và Ứng dụng - Một tổng quan

Huỳnh Văn Ngãi*

Khoa Toán và Thống kê, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 08/10/2021; Ngày nhận đăng: 06/01/2022

TÓM TẮT

Lý thuyết chính quy mêtric và các tính chất liên quan là một công cụ hữu hiệu trong các lĩnh vực Giải tích biến phân và Tối ưu. Trong vài chục năm gần đây, rất nhiều công trình nghiên cứu của nhiều nhà toán học đã đóng góp vào lý thuyết này, trên cả hai phương diện lý thuyết và ứng dụng. Mục đích của bài báo tổng quan này nhằm trình bày một số phát triển nổi bật gần đây, trong đó nhấn mạnh chính vào những đóng góp của nhóm chúng tôi trong lý thuyết này. Đặc biệt, là một số ứng dụng của tính chính quy mêtric trong nghiên cứu sự hội tụ của phương pháp Newton giải phương trình suy rộng. Ngoài ra, chúng tôi áp dụng tính chính quy mêtric để thu được một phiên bản tổng quát của nguyên lý lồi cho ánh xạ đa trị.

Từ khóa: *Tính chính quy mêtric, đối đạo hàm, tính ổn định nhiễu, tính lặp Newton, lồi suy rộng.*

* *Tác giả liên hệ chính.*
Email: ngaivn@yahoo.com

https://doi.org/10.52111/qnjs.2022.16101
Tạp chí Khoa học - Trường Đại học Quy Nhơn, **2022**, *16*(1), 5-21

JOURNAL OF SCIENCE
QUY NHON UNIVERSITY

A survey on the metric regularity and applications

Huynh Van Ngai*

Faculty of Mathematics and Statistics, Quy Nhon University, Vietnam.

Received: 29/06/2021; Accepted: 06/09/2021

ABSTRACT

The theory of metric regularity and related topics plays an important role in variational analysis and have many applications in optimization. In recent decades, it has attracted the study of many researchers in both the theoretical aspects and applications. In the survey paper, we present some recent developments, with emphasis on the results established by our research group in the recent decade in this theory and its applications in the study of the Newton-type methods for solving generalized equations. In addition, based on the metric regularity, we establish a new result on a generalization of the convex principle for paraconvex multimaps.

Keywords: *Metric regularity, coderivative, perturbation stability, Newton iteration, paraconvex.*

1. INTRODUCTION

The study of many mathematical problems originated from practical applications, such as optimization and complementarity problems, variational inequalities, as well as models in equilibrium problems, control theory and design problems, leads to consider inclusions of the type:

$$\text{Find} \quad x \in X \text{ s.t. } y \in F(x) \quad \text{for given } y \in Y, \tag{1.1}$$

here, X, Y are metric spaces and $F : X \rightrightarrows Y$ is a set-valued mapping (also called multimap or multifunction) describing the model under consideration. These inclusions are usually called generalized equations, due to the pioneering work of Robinson.[1,2] The existence of solutions as well as the behavior and stability of the solutions of (1.1) are important principal topics and have attracted many authors working in the fields of variational analysis and optimization. The readers are referred to the monographs,[3–12] to some recent contributions[13–15] and the references therein.

One of the key ingredients to deal with the existence as well as the stability of the solutions of (1.1) is the metric regularity and related properties. Historically, this property goes back to the celebrated Banach open mapping theorem and latter to the Lyusternik theorem (see[5,9,10,16–27] and the references given therein). Recently, many important applications of this property have been found and investigated, especially in the study of stability of variational systems as well as convergence analysis of some algorithms, e.g., as the Newton type methods (see[28–31]). More recently, some generalized metric regularity concepts have been introduced and studied, due to the point of view of applications. For example, in the papers[32,33], some variants of relative metric regularity have used in convergence analysis of some optimization algorithms. An important notion of extended met-

* *Corresponding author.*
Email: ngaivn@yahoo.com

https://doi.org/10.52111/qnjs.2022.16101

ric regularity is the directional metric regularity (see[34–39]). One can be referred to[6,34,35] and references given therein for applications of directional metric regularity in sensitivity analysis and optimization.

In this survey paper, we presents some contributions in the recent decade of our research group to the theory of metric regularity and some applications. This survey is far to be exhaustive on this subject. We recommend the recent book and the survey papers by Ioffe[9,40,41] for the excellent accounts on the recent developments of the theory of metric regularity and divers applications in variational analysis and optimization. Outline of the paper is as follows. In Section 2, we recall the notion of the metric regularity and related notions; some classical results on the metric regularity; some variational characterizations and the perturbation stability of this property. Section 3 is devoted to the relative metric regularity, in which we present some very recent results concerning the directional metric regularity relative to a cone. In the final section, some applications to the Newton methods for generalized equations and to the convex principle for multimaps are reported.

2. METRIC REGULARITY OF MULTIMAPS

We recall firstly some basic notations and notions from set-valued analysis. Throughout the paper, for a metric space X endowed with metric d, denote by $B(x,\rho)$ and $B[x,\rho]$ the open and the closed ball centered at $x \in X$ with radius $\rho > 0$, respectively. The distance function to a subset $C \subseteq X$ is denoted by $d(x,C) := \inf_{u\in C} d(x,u)$. By a set-valued mapping (or a multimap) $F : X \rightrightarrows Y$, it means a correspondance from X to $\mathcal{P}(Y)$, the set of the (possibly empty) subsets of Y. Given a multimap F, the graph of F, the domain of F are the sets gph $F := \{(x,y) \in X \times Y : y \in F(x)\}$ and Dom $F := \{x \in X : F(x) \neq \emptyset\}$, respectively. Inverse of F is denoted by $F^{-1} : Y \rightrightarrows X$, and defined by

$$x \in F^{-1}(y) \iff y \in F(x).$$

A multimap between normed spaces is called a convex (respectively closed) multimap if its graph is convex (respectively closed) graph in the product space.

2.1. Metric regularity: Classical results

Consider an operator equation defined by

$$f(x) = y, \tag{2.1}$$

where $f : X \to Y$ is a mapping acting between metric spaces X, Y .
In practice, one finds out an approximate solution rather than an exact one. The error of some approximate solution x is the quantity

$$d(x, f^{-1}(y)) = \inf\{d(x,u) : \ f(u) = y\}.$$

Naturally, the distance $d(y, F(x))$ is used to judge approximate solutions. One seeks so an error estimate of the form

$$d(x, f^{-1}(y)) \leq \kappa d(y, f(x)) \tag{2.2}$$

for all (x,y) in a suitable domain. If (2.2) is satisfied for (x,y) near a given $(\bar{x},\bar{y})$ with $\bar{y} = F(\bar{x})$, then F is called metrically regular at $\bar{x}$.

The metric inequality can be extended naturally to multimaps. For example, consider a system of inequalities:

$$g_i(x) \leq y_i, i = 1, ..., m. \tag{2.3}$$

This system of inequalities can be investigated via the generalized equation of the form: $y \in G(x)$, where,

$$G(x) := (g_i(x))_{i=1,...,m} + \mathbb{R}^m_+; \ \ y = (y_i)_{i=1,...,m}, \tag{2.4}$$

https://doi.org/10.52111/qnjs.2022.16101

then $G : X \rightrightarrows \mathbb{R}^m$ is a multimap.

Let us now recall the notion of metric regularity.

Definition 1. *A multimap (set-valued mapping) $F : X \rightrightarrows Y$ is said to be metrically regular at $(\bar{x}, \bar{y})$ ($\bar{y} \in F(\bar{x})$) if there are $\kappa, \delta > 0$ such that*

$$d(x, F^{-1}(y)) \leq \kappa d(y, F(x))$$

for all $(x, y) \in B((\bar{x}, \bar{y}), \delta)$. The infimum of such κ denoted by $\operatorname{Reg} F(\bar{x}, \bar{y})$, and is called the regular modulus of F at $(\bar{x}, \bar{y})$.

In the linear case, the metric regularity is closely related to the Banach open mapping principle for bounded linear operators between Banach spaces, restated as follows.

Theorem 2. *(Banach open mapping principle) Let X, Y be Banach spaces and let $A \in \mathcal{L}(X, Y)$, the space of bounded linear operators from X to Y. If the operator A is surjective, that is, $\operatorname{Im} A = Y$, then A is an open mapping,that is, there is $r > 0$ such that $rB_Y \subseteq A(B_X)$. The upper bound of such r is called the Banach constant of A :*

$$C(A) = \inf\{\|A^* y^*\| : \ \|y^*\| = 1\}.$$

Moreover, the following inequality holds

$$d(x, A^{-1}(y)) \leq C(A)^{-1} \|Ax - y\|$$

for all $(x, y) \in X \times Y$.

The Banach open mapping principle tells us that a bounded linear operator between Banach spaces is (locally or equivalently globally) metrically regular if and only if it is surjective. This principle was extended to continuously differentiable mappings by Lyusternik as follows.

Theorem 3. *(Lyusternik) Let X, Y be Banach spaces; and let $f : X \to Y$ be a continuously differentiable mapping at $\bar{x} \in X$ with $f(\bar{x}) := \bar{y}$. Then f is metrically regular at $\bar{x}$ if and only if $Df(\bar{x})$ is onto: If $\operatorname{Im} Df(\bar{x}) = Y$.*

The metric regularity is strongly connected to the Robinson and Mangasarian-Fromovitz constraint qualifications in Mathematical Programming. Consider $F := f - C$, where $f : X \to Y$ is a mapping of $\mathcal{C}^1$ class and $C \subseteq Y$ is a nonempty closed convex subset. Given $(\bar{x}, 0) \in \operatorname{gph} F$, then F is metrically regular at $(\bar{x}, 0)$ *if* the Robinson constraint qualification (RCQ) is satisfied:

$$0 \in \operatorname{int}[f(\bar{x}) + Df(\bar{x})X - C].$$

In particular, for systems of equality and inequality (2.4), one has the equivalence:

$$(\text{RCQ}) \Leftrightarrow (MFCQ) \ (\text{Mangasarian-Fromovitz constraint qualification}).$$

In the case of convex multimaps, a necessary and sufficient condition was given by Robinson-Ursescu, stated in the following theorem.

Theorem 4. *(Robinson-Ursescu) Given Banach spaces X, Y, a closed and convex multimap $F : X \rightrightarrows Y$ (F has a closed and convex graph), F is metrically regular at $(x_0, y_0) \in \operatorname{gph} F$ if and only if $y_0 \in \operatorname{int}(\operatorname{Im} F)$.*

Next we recall the two notions of the openness at a linear rate and of Lipschitz-like (or Aubin) property of multimaps.

Definition 5. *Let $F : X \rightrightarrows Y$ be a multimap acting between metric spaces X, Y and let $(\bar{x}, \bar{y}) \in \operatorname{gph} F$.*

(a) F is said to be open at a linear rate around $(\bar{x}, \bar{y})$ if there exist $s, \varepsilon > 0$ such that

$$B(y, ts) \subseteq F(B(x, t)),$$

$\forall (x, y) \in B((\bar{x}, \bar{y}), \varepsilon) \cap \operatorname{gph} F$. The supremum of such r denoted by $\operatorname{Sur} F(\bar{x}, \bar{y})$ and is called the rate of openness (or surjection) of F at $(\bar{x}, \bar{y})$.

https://doi.org/10.52111/qnjs.2022.16101

(b) We say that F is Lipschitz-like (or Aubin) at $(\bar{x},\bar{y})$ if there exist $L,\varepsilon>0$ such that

$$d(y,F(x))\le Ld(x,u),$$

$\forall x\in B(\bar{x},\varepsilon)$, $(u,y)\in B((\bar{x},\bar{y}),\varepsilon)\cap \mathrm{gph}\,F$. The supremum of such L is the Lipschitz rate (or modulus) of F at $(\bar{x},\bar{y})$, and is denoted by $\mathrm{Lip}\,F(\bar{x},\bar{y})$.

The equivalence of these notions to the metric regularity was given independently by several authors (see, e.g., Borwein-Zuang[42], Kruger[43], Penot[44], Ioffe[45]).

Proposition 6. *For a multimap $F: X\rightrightarrows Y$ between metric spaces X,Y, and for $(\bar{x},\bar{y})\in \mathrm{gph}\,F$, the following three assertions are equivalent.*

(a) F is metrically regular around $(\bar{x},\bar{y})$;

(b) F is open at linear rate around $(\bar{x},\bar{y})$;

(c) F^{-1} is Lipschitz-like around $(\bar{x},\bar{y})$.

Moreover, one has the equality between modulus

$$\mathrm{Reg}\,F(\bar{x},\bar{y})=\mathrm{Lip}\,F^{-1}(\bar{y},\bar{x})=\frac{1}{\mathrm{Sur}\,F(\bar{x},\bar{y})}.$$

2.2 Characterization of the metric regularity via strong slopes

Let $f: X\to\mathbb{R}\cup\{+\infty\}$ be an extended real-valued function defined on a metric space. We make use of the notations: $\mathrm{dom}f:=\{x\in X: f(x)<+\infty\}$, the domain of f; $\limsup\limits_{y\to x,y\neq x}$, $\liminf\limits_{y\to x,y\neq x}$ mean that the limits superior/inferior are taken as $y\to x$ and $y\neq x$; while $\limsup\limits_{y\to x}$, $\liminf\limits_{y\to x}$ allow the case $y=x$. The symbol $[f(x)]_+$ stands for $\max(f(x),0)$. Recall that the local slope of a lower semicontinuous function f at $x\in\mathrm{dom}f$ is denoted by $|\nabla f|(x)$ and defined by $|\nabla f|(x)=0$ if x is a local minimum of f; otherwise

$$|\nabla f|(x)=\limsup_{y\to x,y\neq x}\frac{f(x)-f(y)}{d(x,y)}.$$

For $x\notin\mathrm{dom}f$, we set $|\nabla f|(x)=+\infty$. The non-local slope of f is defined by

$$|\Gamma f|(x):=\sup_{u\neq x}\frac{[f(x)-f(u)]_+}{d(x,u)}.$$

For $x\notin\mathrm{dom}f$, we set $|\Gamma f|(x)=+\infty$.

It is well-known in the literature that if X is a normed space and f is Fréchet differentiable at x then $|\nabla f|(x)=\|f'(x)\|$. Obviously, one always has the relation $|\nabla f|(x)\le|\Gamma f|(x)$ for all $x\in X$.

Recall the lower semicontinuous envelope $(x,y)\mapsto\varphi^F(x,y):=\varphi^F_y(x)$ of the function $(x,y)\mapsto d(y,F(x))$ defined by, for $(x,y)\in X\times Y$,

$$\varphi^F_y(x):=\liminf_{(u,v)\to(x,y)}d(v,F(u))=\liminf_{u\to x}d(y,F(u)).$$

For the simplification of the notation, when one works only with a given mapping F, one denotes $\varphi_y:=\varphi^F_y$. The following theorem established by Ngai-Tron-Théra[46] gives an estimate of the regularity rate $\mathrm{Reg}\,F(\bar{x},\bar{y})$ via strong slopes of the functions φ_y.

Theorem 7. *(Ngai-Tron-Théra[46]) Let X be a complete metric space and let Y be a metric space. For a multimap $F: X\rightrightarrows Y$, and for given $(\bar{x},\bar{y})\in\mathrm{gph}\,F$, one has*

$$\begin{aligned}\mathrm{Reg}\,F(\bar{x},\bar{y})^{-1}&=\mathrm{Sur}\,F(\bar{x},\bar{y})\\&\ge\liminf_{(x,y)\to(\bar{x},\bar{y}),y\notin F(x)}|\nabla\varphi_y|(x).\end{aligned}$$

When Y is a normed space (or more general, a smooth manifold, or a length metric space), the equality holds

$$\mathrm{Reg}\,F(\bar{x},\bar{y})^{-1}=\liminf_{(x,y)\to(\bar{x},\bar{y}),y\notin F(x)}|\nabla\varphi_y|(x).$$

https://doi.org/10.52111/qnjs.2022.16101

2.3 Coderivative characterizations

Firstly we recall the main definitions and results from Variational Analysis necessary and used in the sequel. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be an extended real-valued function defined on a Banach space X, the *Fréchet (regular) subdifferential* of f at $\bar{x} \in \text{dom} f$ is defined by

$$\partial f(\bar{x}) = \left\{x^\star \in X^\star : \liminf_{x\to\bar{x}} \frac{f(x)-f(\bar{x})-\langle x^\star, x-\bar{x}\rangle}{\|x-\bar{x}\|} \geq 0\right\}.$$

When f is a convex function, the Fréchet subdifferential coincides with the subdifferential in the sense of convex analysis. When $\bar{x} \notin \text{dom } f$, one sets $\partial f(\bar{x}) = \emptyset$. We shall also work on an interesting subclass of Banach spaces, called Asplund spaces being spaces such that on which every convex continuous function is generically Fréchet differentiable. Notice that any space with a Fréchet smooth renorming (and therefore any reflexive space) is Asplund. It is well-known that a Banach space is Asplund if and only if each of its separable subspaces has a separable dual.

In the Asplund setting, the Fréchet subdifferential enjoys a fuzzy sum rule which was firstly proved by Fabian[48] (see also[10,Thm.2.33]).

Given a nonempty closed set $C \subseteq X$, the *indicator function* associated to C is the function ι_C defined by $\iota_C(x) = 0$, when $x \in C$ and $\iota_C(x) = \infty$ otherwise. The *Fréchet normal cone* to C at $\bar{x}$ is the set $N(C, \bar{x}) := \partial\iota_C(\bar{x})$ if $\bar{x} \in C$, and $N(C, \bar{x}) := \partial\iota_C(\bar{x}) = \emptyset$ if $\bar{x} \notin C$.

The *limiting subdifferential* (or also called the Mordukhovich subdifferential) is defined by

$$\partial_{\mathcal{L}} f(\bar{x}) = \{x^\star \in X^\star : \exists x_k \in \bar{x},\ f(x_k) \to f(\bar{x}), \text{ and } \exists x_k^\star \in \partial f(x_k),\ x_k^\star \overset{\star}{\to} x^\star\}.$$

The *limiting normal cone* $N_{\mathcal{L}}(C, \bar{x})$ to a closed set C is defined through the indicator function of the set:

$$N_{\mathcal{L}}(C, \bar{x}) := \partial_{\mathcal{L}}\delta_C(\bar{x}).$$

Given a normal cone mapping $\mathbb{N}$, it is associated with a set-valued mapping $F : X \rightrightarrows Y$ a coderivative $D^\star_{\mathbb{N}} : Y^\star \rightrightarrows X^\star$ by the formula

$$D^\star_{\mathbb{N}} F(x,y)(y^\star) := \big\{x^\star \in X^\star \mid (x^\star, -y^\star) \in \mathbb{N}(\text{gph } F, (x,y))\big\}. \tag{2.5}$$

For further the properties and calculus rules for the Fréchet and liniting subdifferentials as well as coderivatives, see[10,49–51] and the references given therein. In what follows, the Fréchet coderivative of F will be denoted by $D^\star_{\mathcal{F}} F$, or simply by $D^\star F$, while the limiting coderivative is noted by $D^\star_{\mathcal{L}} F$.

The following theorem gave an estimate for the slope of φ_y via the coderivative of the multimap in question.

Theorem 8. *(Ngai-Tron-Théra[46]) Let $F : X \rightrightarrows Y$ be a closed multimap acting between Asplund spaces X and Y. Then for any $(x, y) \in X \times Y$ with $y \notin F(x)$, one has the following estimate*

$$|\nabla\varphi_y|(x) \geq \tau(x, y), \tag{2.6}$$

where $\tau(x, y)$ is defined by

$$\tau(x,y) := \liminf_{\xi\downarrow 0} \left\{ \|x^\star\| : \begin{array}{l} x^\star \in \widehat{D}^\star F(u,v)(y^*),\ \|y^*\| = 1, \\ (u,v) \in \text{gph } F,\ u \in B(x,\xi), \\ \|y - v\| \leq \varphi_y(x) + \xi, \\ |\langle y^*, v - y\rangle - \varphi_y(x)| \leq \xi \end{array} \right\}.$$

This theorem yields immediately the following corollary.

Corollary 9. *With the assumptions as in the preceding theorem, for any $(\bar{x}, \bar{y}) \in \text{gph } F$, one has*

$$\liminf_{(x,y)\to(\bar{x},\bar{y}), y\notin F(x)} |\nabla\varphi_y|(x) = \lim_{\varepsilon\to 0} \inf\{\|x^*\| : x^* \in D^*F(u,v)(y^*), (u,v) \in B((\bar{x},\bar{y}),\varepsilon),\ \|y^*\| = 1\}.$$

In view of this corollary, Theorem 7 implies immediately the following characterization of the metric regularity through the coderivatives.

https://doi.org/10.52111/qnjs.2022.16101

Theorem 10. *(Ioffe[52]) With the assumptions as in the preceding corollary, one has*

$$\begin{aligned}&\operatorname{Sur} F(\bar{x},\bar{y}) =\\ &\lim_{\varepsilon\to 0}\inf\{\|x^*\| :\ x^*\in D^*F(u,v)(y^*),\\ &\qquad\qquad (u,v)\in B((\bar{x},\bar{y}),\varepsilon),\ \|y^*\|=1\}.\end{aligned}$$

In the case when X, Y are finite dimensional, one obtains the following nice point-based characterization due to Mordukhovich (e.g.,[10]).

Theorem 11. *(Mordukhovich[25]) Suppose that X, Y are finite dimensional spaces, and $F : X \rightrightarrows Y$ is a closed multimap. For given $(\bar{x},\bar{y}) \in \operatorname{gph} F$, F is metrically regular at $(\bar{x},\bar{y})$ if and only if*

$$\operatorname{Ker} D^\star_{\mathcal{L}} F(\bar{x},\bar{y}) = \{0\},$$

where

$$\operatorname{Ker} D^\star_{\mathcal{L}} F(\bar{x},\bar{y}) = \{y^*\in Y^* :\ \ 0\in D^\star_{\mathcal{L}} F(\bar{x},\bar{y})(y^*)\}.$$

2.4 Stability of the metric regularity and regularity radius

For a mapping $f : X \to Y$, let us denote

$$\operatorname{Lip} f = \sup_{x_1,x_2\in X, x_1\neq x_2} \frac{d(f(x_1),g(x_2))}{d(x_1,x_2)};$$

$$\operatorname{Lip} f(\bar{x}) = \limsup_{x_1,\, x_2\to\bar{x}, x_1\neq x_2} \frac{d(f(x_1),f(x_2))}{d(x_1,x_2)},$$

called the Lipschitz modulus of f on X and near $\bar{x}$, respectively.
The regularity radius of a multimap F at $(\bar{x},\bar{y}) \in \operatorname{gph} F$ is defined by

$$\begin{aligned}\operatorname{Rad} F(\bar{x},\bar{y}) = \inf_{f:X\to Y}\{&\operatorname{Lip} f(\bar{x}) :\ F+f \text{ fails to be}\\ &\text{metrically regular at } (\bar{x},\bar{y}+f(\bar{x}))\}.\end{aligned}$$

The following relation between the regularity modulus and the regularity radius was established in [53,54,40]:

$$\operatorname{Rad} F(\bar{x},\bar{y}) \geq \frac{1}{\operatorname{Reg} F(\bar{x},\bar{y})}. \qquad (2.7)$$

Precisely,

Theorem 12. *(Ioffe[54], Dontchev-Lewis-Rockafellar[53]) Let $F : X \rightrightarrows Y$ be a closed multimap from a completed metric space X to a normed space Y. Assume that F is metrically regular at $(\bar{x},\bar{y}) \in \operatorname{gph} F$ with $\operatorname{reg} F(\bar{x},\bar{y}) := \tau$. Then for any locally Lipschitz mapping $f : X \to Y$ at $\bar{x}$ with Lipschitz constant $L \in (0,\tau^{-1})$, one has*

$$\operatorname{Reg}(F+f)(\bar{x},\bar{y}+f(\bar{x})) \leq 1/(\tau^{-1}-L).$$

Note that, due to Dontchev, Lewis, and Rockafellar[53] and also to Ioffe[54], that the equality holds if one of the following conditions is satisfied:

- X and Y are finite dimensional spaces;
- $F : X \to Y$ is a single-valued mapping.

The following theorem due to Ngai[55] shows the validity of the equality in (2.7) holds for the case of multimaps under suitable assumptions.

Theorem 13. *(Ngai[55]) Let $F : X \rightrightarrows Y$ be a closed multimap acting from a completed metric space X to a uniformly convex space Y. Assume that F is metrical regular around $(\bar{x},\bar{y}) \in \operatorname{gph} F$ with $\operatorname{reg} F(\bar{x},\bar{y}) = \tau \in (0,+\infty)$. Then the equality in (2.7) holds under one of the two following conditions:*

(i) F is upper semicontinuous around $\bar{x}$, $F(u)$ is convex for u near $\bar{x}$ and either $F(\bar{x})$ is singleton or $\bar{y}$ is an interior point of $F(\bar{x})$;

(ii) Y is a finite dimensional space.

3. RELATIVE METRIC REGULARITY

Given a subset W of $X \times Y$ and a point $(x,y) \in X \times Y$, define the following set:

$$W_x := \{z\in Y :\ \ (x,z)\in V\}$$

$$\text{and}\quad W_y := \{u\in X :\ \ (u,y)\in V\}.$$

https://doi.org/10.52111/qnjs.2022.16101

Definition 14. *(Ioffe[38]) For $W \subset X \times Y$, a multimap $F : X \rightrightarrows Y$ is said to be* metrically regular relative to W at $(\bar{x}, \bar{y}) \in W \cap \operatorname{gph} F$ with a modulus $\tau > 0$, *if there is $\delta > 0$ such that*

$$d\big(x, F^{-1}(y) \cap \operatorname{cl} W_y\big) \leq \tau d(y, F(x)) \qquad (3.1)$$

whenever $(x, y) \in \big(B(\bar{x}, \delta) \times B(\bar{y}, \delta)\big) \cap W$ and $d(y, F(x)) < \delta$.

We shall denote $\operatorname{reg}_W F(\bar{x}, \bar{y})$, the infinum over all $\tau > 0$ such that (3.1) is verified.

Given a cone $C \subseteq Y$ in a normed linear space Y, for $\delta > 0$, let us set

$$C(\delta) := \{v \in Y : \ d(v, C) \leq \delta \|v\|\}.$$

and

$$W_F(C, \delta) := \big\{(x, y) \in X \times Y : \ y \in F(x) + C(\delta)\big\}.$$

Let us recall the definition of the metric regularity with respect to a cone.

Definition 15. *F is called metrically regular relative to C at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, if there is $\delta > 0$ such that F is metrically regular relative to $W := W_F(C, \delta)$ at $(\bar{x}, \bar{y})$.*

Due to Definition 14, we see that F is relatively metrically regular with respect to a cone C at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ with a modulus $\tau > 0$, if there is $\varepsilon > 0$ such that

$$d\big(x, F^{-1}(y) \cap \operatorname{cl} W_{F,y}(C, \delta)\big) \leq \tau d(y, F(x)) \qquad (3.2)$$

for all $(x, y) \in \big(B(\bar{x}, \varepsilon) \times B(\bar{y}, \varepsilon)\big) \cap W_F(C, \delta)$ with $d(y, F(x)) < \varepsilon$.

We consider the lower semicontinuous envelope relatively to W of the function $x \mapsto d(y, F(x))$, which is defined as follows.

$$\varphi_{F,W}(x, y) := \begin{cases} = \liminf\limits_{\operatorname{cl} W_y \ni u \to x} d(y, F(u)) & \text{if } x \in \operatorname{cl} W_y \\ +\infty & \text{otherwise.} \end{cases} \qquad (3.3)$$

Note that obviously $\varphi_{F,W}(x, y) \geq 0$ and $\varphi_{F,W}(x, y) \leq d(y, F(x))$ for every $(x, y) \in \operatorname{cl} W_y \times Y$.

The following result established by Ngai-Théra[39], gave a slope characterization of the relative metric regularity.

Theorem 16. *Let $F : X \rightrightarrows Y$ be a closed multimap from a completed metric space X to a metric one Y. For $(\bar{x}, \bar{y}) \in \operatorname{gph} F \cap W$, $W \subset X \times Y$ and $\tau \in (0, +\infty)$, consider the following assertions, one has $(a) \Leftrightarrow (b) \Leftarrow (c)$.*

(a) F is metrically regular relative to W at $(\bar{x}, \bar{y})$ with modulus τ;

(b) There are $\alpha, \beta > 0$ such that

$$|\Gamma \varphi_{F,W}(\cdot, y)|(x) \geq \tau^{-1}$$

for any $(x, y) \in B(\bar{x}, \alpha) \times B(\bar{y}, \alpha)$ with $\varphi_{F,W}(x, y) \in (0, \beta)$.

(c) There are $\alpha, \beta > 0$ such that

$$|\nabla \varphi_{F,W}(\cdot, y)|(x) \geq \tau^{-1}$$

for any $(x, y) \in B(\bar{x}, \alpha) \times B(\bar{y}, \alpha)$ with $\varphi_{F,W}(x, y) \in (0, \beta)$.

As the usual metric regularity, the relatively metric regularity with respect to a cone also possesses the following perturbation stability.

Theorem 17. *(Ngai-Tron-Théra, 2019[56]) Let $F : X \rightrightarrows Y$ be a closed multimap from a completed X to a normed space Y. For a nonempty cone $C \subseteq Y$, if F is relatively metrically regular at $(\bar{x}, \bar{y})$ with a modulus $\tau > 0$ with respect to C, then for any locally lipschitz mapping $g : X \to Y$ around $\bar{x}$ with a sufficiently small Lipschitz constant, the multimap $F + g$ is also relatively metrically regular with respect to C at $(\bar{x}, \bar{y} + g(\bar{x}))$.*

https://doi.org/10.52111/qnjs.2022.16101

The next result is a characterization in terms of coderivatives of the relative metric regularity. For this, associated to $F : X \rightrightarrows Y$, we define the multimap $G : X \rightrightarrows Y \times Y$ as follows

$$G(x) = F(x) \times F(x), \quad x \in X.$$

Theorem 18. *(Ngai - Tron - Théra,* [56]*) Let $F : X \rightrightarrows Y$ be a closed multimap between Asplund spaces X, Y. For $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ and a nonempty cone $C \subsetneq Y$, suppose that F has convex values around $\bar{x}$. Then F is relatively metrically regular with respect to C with modulus $\tau \leq m^{-1}$ at $(\bar{x}, \bar{y})$, provided*

$$\liminf_{\substack{(x,y_1,y_2) \xrightarrow{G} (\bar{x},\bar{y},\bar{y}) \\ \delta \downarrow 0^+}} d(0, D^\star_{\mathcal{F}} G(x, y_1, y_2)(T(C, \delta))) > m > 0, \tag{3.4}$$

Recall that A multimap $F : X \rightrightarrows Y$ is *partially sequentially normally compact* (PSNC, shortly,[10]) at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, if for all sequence $\{(x_k, y_k, x_k^\star, y_k^\star)\}_{n\in\mathbb{N}} \subset \operatorname{gph} F \times X^\star \times Y^\star$ verifying

$$(x_k, y_k) \to (\bar{x}, \bar{y}), x_k^\star \in D^\star_{\mathcal{F}} F(x_k, y_k)(y_k^\star),$$
$$y_k^\star \overset{w^\star}{\to} 0, \|x_k^\star\| \to 0,$$

one has $\|y_k^\star\| \to 0$ as $k \to \infty$. Mention that (PSNC) at $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ holds automatically for any multimap when Y is a finite dimensional space. Under condition (PSNC), we obtain a point-based sufficient condition for the relative metric regularity as follows.

Corollary 19. *(see*[56]*) With the assumptions as in Theorem 18, assume in addition that F^{-1} is PSNC at $(\bar{x}, \bar{y}, \bar{y})$. If*

$$d_\star(0, D^\star_{\mathcal{L}} G(\bar{x}, \bar{y}, \bar{y})(T(C, 0))) > 0,$$

then F is relatively metrically regular with respect to C around $(\bar{x}, \bar{y})$.

In particular for the case $F(x) := g(x) - D$, here $D \subseteq Y$, a closed convex subset, $g : X \to Y$ is a continuous map near a given point $\bar{x} \in X$ such as $g(\bar{x}) \in D$, one obtains the following corollary.

Corollary 20. *(see*[56]*) Let X, Y and $C \subseteq Y$ be as before. Let $D \subseteq Y$ be a closed convex subset and $g : X \to Y$ be a continuous map around $\bar{x} \in X$ with $d_0 := f(\bar{x}) \in D$. Then the multimap $F(x) := g(x) - D$, $x \in X$ is relatively metrically regular with respect to C around $(\bar{x}, \bar{y})$ with a modulus $\tau = m^{-1}$, provided*

$$\liminf_{\substack{(x,d_1,d_2) \to (\bar{x},d_0,d_0) \\ \delta \downarrow 0^+}} d_\star(0, D^\star_{\mathcal{F}} f(x)(T(C, \delta) \cap N(D, d_1) \times N(K, d_2)) > 0, \tag{3.5}$$

4. APPLICATIONS

4.1. Newton methods for solving generalized equations

4.1.1. Newton iteration for equations and the Kantorovich theorem

We recall firstly the classical Newton algorithm for solving equations associate to smooth mappings beween Banach spaces. Consider the equation

$$\text{find } x \in X \text{ such that } \ f(x) = 0,$$

where X, Y are Banach spaces and $f : X \to Y$ is a continuously differentiable map. The Newton method for solving this equation consists of the iterations:

$$x_{k+1} = x_k - Df(x_k)^{-1} f(x_k),$$

here x_0 is a started point, and $Df(x_k)$ is invertible for all k. In the other works, the regular zeros (i.e., at which the derivative is invertible) of f are the fixed points of the following Newton operator:

$$N_f(x) = x - Df(x)^{-1} f(x).$$

Generally, when $Df(x)$ is not necessarily invertible, but assumed just to be surjective, the Newton operator is given by

$$N_f(x) = x - Df(x)^+ f(x),$$

here $Df(x)^+$ denotes the Moore-Penrose generalized inverse which coincides with the usual inverse $Df(x)^{-1}$ when it exists.
The following local and non-local quadratic convergence results are due to Kantorovich (1948) (see[57–59]).

Theorem 21. *(Kantorovich-1948[58]) Let $f : X \to Y$ be a mapping of class $\mathcal{C}^2$. Let $\xi \in U$ be such that $f(\xi) = 0$ and the derivative of f at this point be surjective. For $r > 0$ with $B[\xi, \rho] \subseteq U$, set*

$$M(f, \xi, \rho) = \sup_{\|x-\xi\| \le \rho} \|Df(\xi)^+ D^2 f(x)\|.$$

If $2M(f, \xi, \rho)\rho \le 1$ then for all $x_0 \in B[\xi, \rho]$, then the Newton sequence $x_{k+1} = N_f(x_k)$ is completely defined and converges to ξ, with a quadratically convergence rate,

$$\|x_k - \zeta\| \le \left(\frac{1}{2}\right)^{2^k - 1} \|x_0 - \zeta\|.$$

Set $\beta(f, x_0) = \|Df(x_0)^+ f(x_0)\|$ if $Df(x_0)$ is surjective, and $\beta(f, x_0) = +\infty$, otherwise.

Theorem 22. *(Kantorovich[58]) Let f: $X \to Y$ be a mapping of class $\mathcal{C}^2$. For $x_0 \in X$, suppose that the following conditions are satisfied:*

- *$Df(x_0)$ is surjective,*
- *$2\beta(f, x_0) \le \rho$,*
- *$2\beta(f, x_0)M(f, x_0, \rho) \le 1$.*

Then the Newton sequence $x_{k+1} = N_f(x_k)$ is well-defined and converges to some ξ with $f(\xi) = 0$, and one has the estimation

$$\|x_k - \xi\| \le 1.63281... \left(\frac{1}{2}\right)^{2^k - 1} \|x_0 - \xi\|,$$

with

$$1.63281... = \sum_{k=0}^{\infty} \frac{1}{2^{2^k - 1}}.$$

4.1.2. Newton type method of generalized equations

Given Banach spaces X, Y, consider the following generalized equation of the form of the sum of a single-valued map and a set-valued one:

$$0 \in f(x) + F(x), \quad x \in X \tag{4.1}$$

here $f : X \to Y$ is a function of class C^1, and $F : X \rightrightarrows Y$ is a closed multimap.

The Newton type method for solving (4.1) can be described as follows (see[2,60,61]): Given a starting point x_0, the sequence (x_k) is iteratively defined in solving the auxiliary generalized equation: x_{k+1} is a suitable solution of

$$0 \in f(x_k) + Df(x_k)(x - x_k) + F(x) \quad \text{for } k = 0, 1, ... \tag{4.2}$$

Equivalently,

$$x_{k+1} \in (Df(x_k) + F)^{-1}(Df(x_k) - f(x_k)).$$

This Newton type method means of the use of *"partial linearization"* of the single-valued part f.

The following theorem due to Adly-Ngai-Vu[62] gives a local version of convergence for Newton's iteration (4.2).

Theorem 23. *(Adly - Ngai - Vu[62]) Let $f : X \longrightarrow Y$ be a second-order continuously differentiable function on an open subset U of X and let $F : X \rightrightarrows Y$ be a closed multimap. Let $\xi \in U$ be a solution of (4.1) and set $\zeta := Df(\xi)(\xi) - f(\xi) \in Y$. Suppose that the multimap $\Phi = Df(\xi) + F$ is metrically regular on*

https://doi.org/10.52111/qnjs.2022.16101

$V = \mathbb{B}[\xi, r] \times \mathbb{B}[\zeta, \rho]$ of (ξ, ζ) with modulus $\tau > 0$ with $r > 0$, $\mathbb{B}[\xi, r] \subset U$. Set

$$M(\tau, \xi, r) := \tau \sup_{\|u-\xi\| \leqslant r} \left\| D^2 f(z) \right\|,$$

$$\delta = \min\{r, \rho, \tau\rho\}.$$

If $2M(\tau, \xi, r)r \leqslant 1$, then for any $x_0 \in \mathbb{B}[\xi, \delta]$, the sequence $(x_n)_{n \geqslant 0}$ in (4.2) is well-defined and converges quadratically to ξ. More precisely, one has

$$\|x_n - \xi\| \leqslant \left(\frac{1}{2}\right)^{2^n - 1} \|x_0 - \xi\|. \tag{4.3}$$

In order to establish a non-local of the Kantorovich theorem, we need the following definition of the metric regularity on a suitable domain.

Definition 24. *Let a multimap $F : X \rightrightarrows Y$, $x_0 \in X$ and positive constants $r > 0$, $s > 0$, we define*

$$V(F, x_0, r, s) = \{(x, y) \in X \times Y : x \in \mathbb{B}[x_0, r], \; d(y, F(x)) < s\}. \tag{4.4}$$

The multimap F is called metrically regular on $V(F, x_0, r, s)$ with a modulus $\tau > 0$ if for all $(x, y) \in V(F, x_0, r, s)$, one has

$$d\big(x, F^{-1}(y)\big) \leqslant \tau d\big(y, F(x)\big). \tag{4.5}$$

When the multimap F verifies Definition 24, denote the quantity $\operatorname{reg}(F, x_0, r, s)$, the infimum over all $\tau > 0$ satisfying (4.5). Otherwise, set $\operatorname{reg}(F, x_0, r, s) = \infty$.

The next result states a global convergence result of the Newton iterations. The assumptions for this are based on the classical Kantorovich's theorem (see[59]).

Theorem 25. *(Adly-Ngai-Vu[62]) Let $f : X \longrightarrow Y$ and $F : X \rightrightarrows Y$ be maps acting between Banach spaces X, Y. Suppose that f is a $\mathcal{C}^2-$mapping on an open subset $U \subset X$ and F is a closed multimap. Consider problem (4.1), and define the following quantities*

$$\beta(\tau, x) := \tau d\big(0, (f + F)(x)\big),$$

$$M(\tau, x, \rho) := \tau \sup_{\|u-x\| \leqslant \rho} \left\| D^2 f(u) \right\|.$$

Let $x \in U$, $\delta \in (0, 1]$ and $\rho > 0$, $t > 0$ such that the following statements hold.

1. *$G := Df(x) + F$ is metrically regular on $V := V(G, x, 4\rho, t)$ with a modulus $\tau > 0$,*
2. *$d\big(0, f(x) + F(x)\big) < s$,*
3. *$2\beta(\tau, x)M(\tau, x, \rho) \leqslant \delta$,*
4. *$2\eta\beta(\tau, x) \leqslant \rho$, with $\eta = \frac{1-\sqrt{1-\delta}}{\delta} = \frac{1}{1+\sqrt{1-\delta}}$.*

Then the generalized equation (4.1) has a solution ξ satisfying

$$\|x - \xi\| \leqslant 2\eta\beta(\tau, x) \leqslant \rho. \tag{4.6}$$

Moreover, there exists a sequence (x_n) generated by the Newton iterations (4.2) with the starting point x and converges to ξ, and the following error estimate holds

- *if $\delta < 1$, then*

$$\|x_n - \xi\| \leqslant \frac{4\sqrt{1-\delta}}{\delta} \frac{\theta^{2^n}}{1 - \theta^{2^n}} \beta(\tau, x), \tag{4.7}$$

  where $\theta = \frac{1-\sqrt{1-\delta}}{1+\sqrt{1-\delta}}$;

- *if $\delta = 1$, then*

$$\|x_n - \xi\| \leqslant 2^{-n+1}\beta(\tau, x). \tag{4.8}$$

When the mapping f is analytic, an extension of Smale's $(\alpha, \gamma)-$theory to generalized equation was also extensively developed in[62].

https://doi.org/10.52111/qnjs.2022.16101

4.1.3. Newton methods under relativemetric regularity

We consider now the case in which the multimap is not longer metrically regular but merely relatively metrically regular. For a cone $C \subseteq Y$, consider the inexact Newton iterations with respect to C defined as follows.

$$0 \in f(x_k) + Df(x_k)(x - x_k) + F(x) + \varepsilon_k C \cap \mathbb{S}_Y \tag{4.9}$$

for $k = 0, 1, ...$, where $\mathbb{S}_Y$ stands for the unit sphere in Y, and (ε_n) is a suitable sequence of positive reals, converging to zero.
To analyze the convergence of the iterative process (4.9), we established a non-local version of the relative metric regularity as was stated in the following theorem.

Theorem 26. *(Ngai [63]) Let $F : X \rightrightarrows Y$ be a closed multimap acting from a completed metric space X to a normed one Y. For a nonempty cone $C \subseteq Y$, suppose that $F : X \rightrightarrows Y$ is metrically regular on $V(F, \bar{x}, r, s)$ for some $r, s > 0$ with a modulus $\tau > 0$, relatively to C, i.e., there is a real $\delta \in (0, 1)$ such that for all $(x, y) \in V(F, \bar{x}, r, s) \cap W_F(C, \delta)$, one has*

$$d(x, F^{-1}(y) \cap \operatorname{cl} W_{F,y}(C, \delta)) \leq \tau d(y, F(x)). \tag{4.10}$$

Then with a map $f : X \to Y$ being locally Lipschitz on $B[\bar{x}, r]$ with a Lipschitz constant $L > 0$, the multimap $G := F + g$ is relatively metrically regular on $V(G, \bar{x}, \frac{r}{4}, R)$ with respect to C with modulus

$$\operatorname{Reg}_C(G, \bar{x}, \tfrac{r}{4}, R) \leq \kappa := \left(\tfrac{1-\beta(1+\theta)}{\tau(1+\beta(1+\theta))} - L\right)^{-1},$$

provided that

$$\beta \in (0, 1), 0 < \theta < \frac{\delta(1-\beta)}{1+\beta\delta}, \text{ and}$$

$$L < \frac{\theta(1-\theta)\beta}{\tau\left[1+\beta(1+\theta)\right](1+\theta)}, \; R := \min\left\{s, \frac{r}{5\kappa}\right\}.$$

The convergence result of the iterative process (4.9) is stated as follows.

Theorem 27. *(Ngai [63]) Given a nonempty cone $C \subseteq Y$, a function $f : X \longrightarrow Y$ between two Banach spaces X, Y being a $\mathcal{C}^2-$mapping on an open subset $U \subset X$ and a closed multimap $F : X \rightrightarrows Y$. Consider problem (4.1), and define the quantities*

$$\beta(\tau, x) := \tau d\big(0, (f + F)(x)\big),$$

$$M(\tau, x, \rho) := \tau \sup_{\|u-x\| \leqslant \rho} \left\| D^2 f(u) \right\|.$$

Let $x \in U$, $\delta \in (0, 1]$ and $\rho > 0$, $t > 0$, such that the following statements hold.

1. *$G = Df(x) + F$ is metrically regular on $V := V(G, x, 4r, s)$ relatively to C with a modulus $\tau > \operatorname{Reg}_C(G, x, 4r, s)$,*
2. *$0 \in f(x) + F(x) + C$,*
3. *$d\big(0, f(x) + F(x)\big) < t$,*
4. *$2\beta(\tau, x)M(\tau, x, \rho) \leqslant \delta$,*
5. *$2\eta\beta(\tau, x) \leqslant \rho$,*
   with $\eta = \frac{1-\sqrt{1-\delta}}{\delta} = \frac{1}{1+\sqrt{1-\delta}}$.

Then the generalized equation (4.1) has a solution $\xi \in U$ satisfying

$$\|x - \xi\| \leqslant 2\eta\beta(\tau, x) \leqslant r. \tag{4.11}$$

Moreover, there are constants $0 < b_1, b_2 < 1$ and $\varepsilon > 0$ such that for any sequence of positive reals (ε_n) satisfying

$$\varepsilon_0 \in (0, \varepsilon); \;\; \varepsilon_{n+1} \leq b_1\varepsilon_n; \;\; \varepsilon_n^2 \leq b_2\varepsilon_{n+1}, \; n \in \mathbb{N},$$

there exists a sequence (x_n) generated by the iterative process (4.9) with the starting point x and converges to ξ. In addition, if $\delta < 1$ and $\limsup_{n\to\infty} \frac{\varepsilon_{n+1}}{\varepsilon_n^2} < +\infty$, then the convergence rate of such sequence (x_n) is quadratic.

https://doi.org/10.52111/qnjs.2022.16101

4.2. Convex principle for multimaps

In[64,65], Polyak has established an interesting convex principle with numerous applications in linear algebra, optimization and control theory. This convex principle states that a nonlinear image of a small ball in a Hilbert space is convex, provided the mapping is of class $\mathcal{C}^{1,1}$ and the center of the ball is a regular point of the mapping. In this final subsection, based on the metric regularity theory, we present a generalization of the Polyak convex principle for multimaps. Recall[66] that a multimap $F : X \rightrightarrows Y$ between normed spaces X, Y is paraconvex with a modulus $L > 0$ if for all $x_1, x_2 \in X$, $t \in [0, 1]$,

$$tF(x_1) + (1-t)F(x_2) \subseteq F(tx_1 + (1-t)x_2) + \\ + Lt(1-t)\|x_1 - x_2\|^2.$$

Obviously, a single-valued mapping of class $\mathcal{C}^{1,1}$ is paraconvex.

Theorem 28. *Let X be a Hilbert space and Y be a normed space. Let $F : X \rightrightarrows Y$ be a paraconvex multimap with modulus $L > 0$. For given $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, $\varepsilon_1, \varepsilon_2 > 0$, if F is metrically regular with a constant $\kappa > 0$ on $B(\bar{x}, \varepsilon_1) \times B(\bar{y}, \varepsilon_2)$, then for $0 < \delta < \delta_0 := \min\{\varepsilon, (2\kappa L)^{-1}\}$, the set $F(B[\bar{x}, \delta]) \cap B(\bar{y}, \varepsilon_2)$ is a convex set in Y.*

We need the following lemma.

Lemma 29. *Let X be a Hilbert spaces. For $\delta > 0$, $\bar{x} \in X$, $x_1, x_2 \in B[\bar{x}, \delta]$, $t \in (0, 1)$, $x_t := tx_1 + (1-t)x_2$, one has*

$$B[x_t, \delta - \sqrt{\delta^2 - s}] \subseteq B[\bar{x}, \delta],$$

where $s := t(1-t)\|x_1 - x_2\|^2$.

Proof. Using the equality

$$\|ta + (1-t)b\|^2 = t\|a\|^2 + (1-t)\|b\|^2 - t(1-t)\|a-b\|^2,$$

validated in any Hilbert space X, $a, b \in X$, and $t \in \mathbb{R}$, one has the estimation

$$\begin{aligned}\|x_t - \bar{x}\|^2 &= t\|x_1 - \bar{x}\|^2 + (1-t)\|x_2 - \bar{x}\|^2 - \\ &\quad - t(1-t)\|x_1 - x_2\|^2 \\ &\le \delta^2 - t(1-t)\|x_1 - x_2\|^2 = \delta^2 - s.\end{aligned}$$

Hence for $\gamma := \delta - \sqrt{\delta^2 - s}$, $z \in B[x_t, \gamma]$, one has

$$\begin{aligned}\|z - \bar{x}\|^2 &\le (\|z - x_t\| + \|x_t - \bar{x}\|)^2 \\ &= \|z - x_t\|^2 + 2\|z - x_t\|\|x_t - \bar{x}\| + \|x_t - \bar{x}\|^2 \\ &\le \gamma^2 + 2\gamma\sqrt{\delta^2 - s} + \delta^2 - s = \delta^2,\end{aligned}$$

that is, $z \in B[\bar{x}, \delta]$, and the lemma is proved. □

Proof of Theorem 28. Let $y_1, y_2 \in F(B[\bar{x}, \delta]) \cap B(\bar{y}, \varepsilon_2)$, $t \in (0, 1)$ be given. We want to show $y_t := ty_1 + (1-t)y_2 \in F(B[\bar{x}, \delta]) \cap B(\bar{y}, \varepsilon_2)$. As obviously $y_t \in B(\bar{y}, \varepsilon_2)$, for $y_1, y_2 \in B(\bar{y}, \varepsilon_2)$, $t \in (0, 1)$, it suffices to show $y_t \in F(B[\bar{x}, \delta])$. Let $x_1, x_2 \in B[\bar{x}, \delta]$ be such that $y_i \in F(x_i)$, $i = 1, 2$. Then $x_t := tx_1 + (1-t)x_2 \in B[\bar{x}, \delta] \subseteq B(\bar{x}, \varepsilon_1)$, according to the metric regularity with constant κ of F on $B(\bar{x}, \varepsilon_1) \times B(\bar{y}, \varepsilon_2)$, one has

$$d(x_t, F^{-1}(y_t)) \le \kappa d(y_t, F(x_t)).$$

Since F is paraconvex with modulus L,

$$y_t \in tF(x_1) + (1-t)F(x_2) \subseteq F(x_t) + Lt(1-t)\|x_1 - x_2\|^2,$$

therefore, $d(y_t, F(x_t)) \le Lt(1-t)\|x_1 - x_2\|^2$, which along with the previous inequality yields

$$d(x_t, F^{-1}(y_t)) \le \kappa Lt(1-t)\|x_1 - x_2\|^2 = \kappa L s.$$

As by assumption, $\kappa L < \delta/2$, one can find $z_t \in F^{-1}(y_t)$ such that $\|x_t - z_t\| \le \delta s/2$. Then by

$$\delta s/2 \le \delta - \sqrt{\delta^2 - s},$$

one has $z_t \in B[x_t, \delta - \sqrt{\delta^2 - s}]$, and thanks to Lemma 29, $z_t \in B[\bar{x}, \delta]$, consequently $y_t \in F(z_t) \subseteq F(B[\bar{x}, \delta])$. □

https://doi.org/10.52111/qnjs.2022.16101

5. CONCLUSIONS

The paper gives a survey on some recent contributions of our research group to the theory of metric regularity and its applications. Some important characterizations of the metric regularity via the strong slopes as well as the coderivatives have established. We have presented some extended variants of the notion of metric regularity and related properties. Regarding the radius of metric regularity which is an interesting topic in this theory, we have established a relationship between the regularity radius and the regularity modulus for general multimaps under some suitable assumptions. It remains still some open questions related to estimate the regularity radius for future studies. Many applications to numerical algorithms for solving optimization problems and generalized equations have found out recently, especially the applications to the Newton type methods have been investigated in this paper. We have used the metric regularity to obtain a new result related to the convexity of an image of balls through paraconvex multimaps. Many open problems related to the theory of metric regularity and applications are devoted to future researches.

Acknowledgement

I would like to thank the two anonymous referees for reading carefully the paper and for the valuable suggestions. I would like to thank my collaborators from University of Quy Nhon: Dr. Nguyen Huu Tron, Dr. Nguyen Van Vu, and Dao Ngoc Han for the useful discussions and their collaborations.

REFERENCES

1. S. M. Robinson. *Generalized equations and their solutions, part I: Basic theory*, Point-to-Set Maps and Mathematical Programming, Springer Berlin Heidelberg, 1979.
2. S. M. Robinson. *Generalized equations and their solutions, part II: Applications to nonlinear programming*, Optimality and Stability in Mathematical Programming, Springer Berlin Heidelberg, 1982.
3. D. Azé, J. N. Corvellec, R. E. Lucchetti. Variational pairs and applications to stability in nonsmooth analysis, *Nonlinear Analysis*, **2002**, Series A, *49* (5), 643–670.
4. J. P. Aubin, H. Frankowska. *Set-Valued Analysis* , Birkhäuser Boston Inc., 1990.
5. J. M. Borwein, Q. J. Zhu. *Techniques of Variational Analysis*, Springer New York, 2005.
6. J. F. Bonnans, A. Shapiro. *Perturbation Analysis of Optimization Problems*, Springer New York, 2000.
7. R. S. Burachik, A. N. Iusem. *Set-Valued Mappings and Enlargements of Monotone Operators*, Springer New York, 2008.
8. A. L. Dontchev, R. T. Rockafellar. *Implicit Functions and Solution Mappings: A View from Variational Analysis*, 2rd edition, Springer New York, 2016.
9. A. D. Ioffe. *Variational Analysis of Regular Mappings: Theory and Applications*, Springer Monographs in Mathematics, Switzerland, 2017.
10. B. S. Mordukhovich. *Variational analysis and generalized differentiation. I. Basic theory*, Springer Berlin, 2006.
11. J. P. Penot. *Calculus without Derivatives*, Graduate Texts in Mathematics, Springer New York, 2013.
12. R. T. Rockafellar, R. J. B. Wets. *Variational Analysis*, Grundlehren der mathematischen Wissenschaften 317, Springer New York, 1998.

13. A. V. Dmitruk, A. Y. Kruger. Metric regularity and systems of generalized equations, *Journal of Mathematical Analysis and Applications*, **2008**, *342* (2), 864–873.

14. R. Cibulka, J. Preininger, T. Roubal. On uniform regularity and strong regularity, *Optimization*, **2018**, *68* (2-3), 1–29.

15. M. Durea, M. Panțiruc, R. Strugariu. A new type of directional regularity for mappings and applications to optimization. *SIAM Journal on Optimization*, **2017**, *27*, 1204–1229.

16. D. Azé. A survey on eror bounds for lower semicontinuous functions. *ESAIM Proceedings*, **2003**,*13*, 1–17.

17. A. L. Dontchev, R. T. Rockafellar. Regularity and conditioning of solution mappings in variational analysis, *Set-Valued Analysis*, **2004**, *12* (1-2), 79–109.

18. A. Jourani. Metric regularity and second-order necessary optimality conditions for minimization problems under inclusion constraints, *Journal of Optimization Theory and Applications*, **1994**, *81* (1), 97–120.

19. A. Jourani, L. Thibault. Verifiable conditions for openness and regularity of multivalued mappings, *Transactions of the American Mathematical Society*, **1995**, *347*, 1255–1268.

20. Y. S. Ledyaev, Q. J. Zhu. Implicit multifunction theorems, *Set-Valued Analysis*, **1999**, *7* (3), 209–238.

21. H. Frankowska. Some inverse mapping theorems, *Annales de l'Institut Henri Poincaré Analyse Non Linéaire*, **1990**, *7* (3), 183–234.

22. X. Y. Zheng, K. F. Ng. Metric regularity and constraint qualifications for convex inequalities on Banach spaces, *SIAM Journal on Optimization*, **2003**, *14* (3), 757–772.

23. H.V. Ngai, M. Théra. Error bounds in metric spaces and application to the perturbation stability of metric regularity, *SIAM Journal on Optimization*, **2008**, *19* (1), 1–20.

24. B. S. Mordukhovich. Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions, *Transactions of the American Mathematical Society*, **1993**, *340* (1), 1–35.

25. B. S. Mordukhovich. Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis, *Transactions of the American Mathematical Society*, **1994**, *343* (2), 609–657.

26. J. P. Penot. Metric regularity, openness and Lipschitzian behavior of multifunctions, *Nonlinear Analysis*, **1989**, *13* (6), 629–643.

27. J. P. Penot. Compactness properties, openness criteria and coderivatives, *Set-Valued Analysis*, **1998**, *16* (4), 363–380.

28. S. Adly, H. V. Ngai, V. V. Nguyen. Stability of metric regularity with set-valued perturbations and application to Newton's method for solving generalized equations, *Set-Valued and Variational Analysis*, **2017**, *25* (3), 543–567.

29. S. Adly, R. Cibulka, H. V. Ngai. Newton's method for solving inclusions using set-valued approximations, *SIAM Journal on Optimization*, **2015**, *25* (1), 159–184.

30. F. J. Aragón Artacho, A. Belyakov, A. L. Dontchev, M. López. Local convergence of quasi-Newton methods under metric regularity, *Computational Optimization and Applications*, **2014**, *58* (1), 225–247.

31. D. Klatte, B. Kummer. Approximations and generalized Newton methods, *Mathematical Programming*, **2018**, Series B, *168* (1-2), 673–716.

https://doi.org/10.52111/qnjs.2022.16101

32. D. R. Luke, N. H. Thao, M. K. Tam. Quantitative convergence analysis of iterated expansive, set-valued mappings, *Mathematics of Operations Research*, **2018**, *43* (4), 1143–1176.

33. T. Aspelmeier, C. Charitha, D. R. Luke. Local linear convergence of the ADMM/Douglas-Rachford algorithms without strong convexity and application to statistical imaging, *SIAM Journal on Imaging Sciences*, **2016**, *9* (2), 842–868.

34. A. V. Arutyunov, A. F. Izmailov. Directional Stability Theorem and Directional Metric Regularity, *Mathematics of Operations Research*, **2006**, *31* (3), 526–543.

35. A. V. Arutyunov, E. R. Avakov, A. Izmailov. F. Directional regularity and metric regularity, *SIAM Journal on Optimization*, **2007**, *18* (3), 810–833.

36. H. Gfrerer. On directional metric subregularity and second-order optimality conditions for a class of nonsmooth mathematical programs, *SIAM Journal on Optimization*, **2013**, *23* (1), 632–665.

37. H. Gfrerer. On Directional Metric Regularity, Subregularity and Optimality Conditions for Nonsmooth Mathematical Programs, *Set-Valued and Variational Analysis*, **2013**, *21* (2), 151–176.

38. A. D. Ioffe. On regularity concepts in variational analysis, *Journal of Fixed Point Theory and Applications*, **2010**, *8* (2), 339–363.

39. H. V. Ngai, M. Théra. Directional metric regularity of multifunctions, *Mathematics of Operations Research*, **2015**, *40* (4), 969–991.

40. A. D. Ioffe. Metric regularity- a survey. Part I. Theory, *Journal of the Australian Mathematical Society*, **2016**, *101* (2), 188–243.

41. A. D. Ioffe. Metric regularity- a survey. Part II. Applications, *Journal of the Australian Mathematical Society*, **2016**, *101* (3), 376–417.

42. J. M. Borwein, D. M. Zhuang. Verifiable necessary and sufficient conditions for openness and regularity for set-valued and single-valued mapps. *Journal of Mathematical Analysis and Applications*, **1988**, *134*, 441–459.

43. A. Y. Kruger. A covering theorem for set-valued mappings, *Optimization*, **1988**, *19* (6), 763–780.

44. J. P. Penot. On regularity conditions in mathematical programming, *Mathematical Programming Studies*, **1982**, *19*, 167–199.

45. Ioffe, Alexander D. Nonsmooth analysis: differential calculus of nondifferentiable mappings, *Transactions of the American Mathematical Society*, **1981**, *266* (1), 1–56.

46. H. V. Ngai, N. H. Tron, M. Théra. Implicit multifunction theorems in complete metric spaces, *Mathematical Programming*, **2013**, *139*, 301–326.

47. D. Yost. Asplund spaces for beginners, *Acta Universitatis Carolinae. Mathematica et Physica*, **1993**, 34 (2), 159–177.

48. J. Fabian Marián. Subdifferentiability and trustworthiness in the light of a new variational principle of Borwein and Preiss, *Acta Universitatis Carolinae*, **1989**, *30*. 51–56.

49. B. S. Mordukhovich. Coderivatives of set-valued mappings: calculus and applications, *Nonlinear Analysis*, **1997**, *30* (5), 3059–3070.

50. B. S. Mordukhovich, J.V. Outrata. Coderivative analysis of quasi-variational inequalities with applications to stability and optimization, *SIAM Journal on Optimization*, **2007**, *18* (2), 389–412.

https://doi.org/10.52111/qnjs.2022.16101

51. D. Klatte, B. Kummer. *Nonsmooth Equations in Optimization. Regularity, Calculus, Methods and Applications*, Kluwer Academic Publishers, 2002.

52. A. D. Ioffe. Metric regularity and subdifferential calculus, *Russian Mathematical Surveys*, **2000**,*55* (3), 501–558.

53. A. L. Dontchev, A. S. Lewis, R. T. Rockafellar. The radius of metric regularity, *Transactions of the American Mathematical Society*, **2003**, *355* (2), 493–517.

54. A. D. Ioffe. On regularity estimates for mappings between embedded manifolds, *Control and Cybernetics*, **2007**, *36* (3), 659-668.

55. H. V. Ngai. Regularity Modulus and Radius of Set-Valued Mappings, *Vietnam Journal of Mathematics*, **2014**, *42*, 567–578.

56. H. V. Ngai, N. H. Tron, Théra M. Metric Regularity Relative to a Cone, *Vietnam Journal of Mathematics*, **2019**, *47*, 733–756.

57. K. Kantorovich, G. Akilov. *Analyse Fonctionnelle*, Moscou, 1981.

58. K. Kantorovich, V. Leonid. On Newton's method for functional equation, *Doklady Akademii Nauk SSSR*, **1948**, *59*, 1237–1240.

59. Jean-Pierre Dedieu. *Points Fixes, Zéros et la Mèthode de Newton*, Mathématiques & Applications, Springer, 2006.

60. A. L. Dontchev. Local analysis of Newton–type method based on partial linearization, *Lectures in Applied Mathematics*, **1996**, *32*, 295–306.

61. A. L. Dontchev. Local convergence of the Newton method for generalized equations, *Comptes Rendus de l'Académie des Sciences Paris Série I Mathematics*, **1996**, *322* (4), 327–331.

62. S. Adly, H. V. Ngai, V. V. Nguyen. Newton-type method for solving generalized equations: Kantorovich's and Smale's approaches, *Journal of Mathematical Analysis and Applications*, **2016**, *439* (1), 396–418.

63. H.V. Ngai. Directional metric regularity and Newton methods for generalized equations, In preparation.

64. B. Polyak, Convexity of nonlinear image of a small ball with applications to optimization, *Set-Valued Analysis*, **2001**, *9* (1-2), 159–168.

65. B. Polyak. Local programming, *Computational Mathematics and Mathematical Physics*, **2001**, *41* (9), 1259–1266.

66. S. Rolewicz. Paraconvex Analysis, *Control and Cybernetics*, **2005**, *34* (3), 951–965.

https://doi.org/10.52111/qnjs.2022.16101

TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC QUY NHƠN

Tổng hợp cấu trúc opal nghịch đảo Co_3O_4 có độ xốp cao ứng dụng làm vật liệu xúc tác điện hóa hiệu suất cao

Nguyễn Văn Nghĩa[1], Phan Thế Vinh[2], Nguyễn Thị Lệ Thanh[3], Nguyễn Thị Hồng Trang[1,*]

[1]Khoa Khoa học tự nhiên, Trường Đại học Quy Nhơn, Việt Nam
[2]Cao học ngành Vật lý chất rắn khóa 22, Khoa Khoa học tự nhiên, Trường Đại học Quy Nhơn, Việt Nam
[3]Cao học ngành Vật lý chất rắn khóa 20, Khoa Khoa học tự nhiên, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 27/09/2021; Ngày nhận đăng: 02/12/2021

TÓM TẮT

Bằng phương pháp dùng khuôn là các quả cầu polystyrene, kết hợp với quá trình nhiệt phân trong không khí, chúng tôi đã chế tạo thành công vật liệu nano Co_3O_4 cấu trúc xốp có độ trật tự cao. Hình thái và cấu trúc tinh thể của vật liệu được khảo sát bằng phương pháp kính hiển vi điện tử quét (SEM) và phương pháp nhiễu xạ tia X (XRD). Kết quả ảnh SEM cho thấy khoảng cách giữa các lỗ xốp cỡ vài đến vài chục nano mét và đường kính trung bình của các lỗ xốp trong khoảng 260 – 300 nm. Kết quả phân tích phổ XRD của mẫu cho thấy vật liệu được tổng hợp có cấu trúc tinh thể dạng lập phương. Tính chất xúc tác điện hóa của vật liệu được khảo sát thông qua phản ứng tổng hợp oxi từ nước. Kết quả cho thấy hiệu suất vượt trội của cấu trúc xốp nano so với cấu trúc hạt nano.

Từ khóa: *Co_3O_4, nano xốp, xúc tác điện hóa, phản ứng tổng hợp oxi.*

**Tác giả liên hệ chính.*
Email: nguyenthihongtrang@qnu.edu.vn

JOURNAL OF SCIENCE
QUY NHON UNIVERSITY

Facile synthesis of highly ordered mesoporous Co_3O_4 inverse opals as a high-performance electrocatalyst

Nguyen Van Nghia,[1] Phan The Vinh,[2]
Nguyen Thi Le Thanh,[3] Nguyen Thi Hong Trang[1,*]

[1]Faculty of Natural Sciences, Quy Nhon University, Vietnam
[2]22nd Master Class of Solid State Physics, Faculty of Natural Sciences, Quy Nhon University, Vietnam
[3]20st Master Class of Solid State Physics, Faculty of Natural Sciences, Quy Nhon University, Vietnam

Received: 27/09/2021; Accepted: 02/12/2021

ABSTRACT

In this study, highly ordered mesoporous Co_3O_4 inverse opals are synthesized using polystyrene beads as a hard template, followed by a calcination in air. The morphology and crystalline structure of the mesoporous Co_3O_4 are investigated by scanning electron microscopy and X–ray diffraction techniques. The characterizations reveal that the thickness of the pore walls of the inverse opals is in the range of a few to a few tens of nanometers, the average diameter of the pores is in the range of 260 – 300 nm, and the mesoporous Co_3O_4 has a spinel crystalline structure. The electrocatalytic properties of the materials are investigated by the oxygen evolution reaction. The results show a superior activity of the mesoporous structure with a small Tafel slope and a low overpotential in the alkaline medium compared to the nanoparticle counterpart.

Keywords: *Co_3O_4, mesoporous inverse opal, electrocatalysis, oxygen evolution reaction.*

1. INTRODUCTION

Renewable energy sources have been developed rapidly due to the world's growing energy demands and serious environmental threat caused by the depletion of natural fossil resources such as coal, oil, and natural gas. Even so, currently renewable energy sources only account for a very small part of the total global energy consumption. The renewable energy is still not popular in the world because the extraction technology is still limited, making product cost still high.

Catalytic splitting of water into hydrogen and oxygen provides a potential path to product clean H_2 and O_2 for human society as the input materials for fuel cells, and metal–air batteries such as lithium–air batteries, zinc–air batteries. However, one of major hurdles of water electrolysis is anodic oxygen evolution reaction (OER) which needs high onset potential and shows slow sluggish kinetics due to four-electron transfer process. Consequently, extensive efforts have been undertaken to develop highly efficient catalysts with low onset potential and promoted reaction kinetics. Many materials with different structures have been studied recently to improve OER efficiency.

Typically, ruthenium oxide (RuO_2) and iridium oxide (IrO_2) have been proven to be

**Corresponding author.*
Email:nguyenthihongtrang@qnu.edu.vn

https://doi.org/10.52111/qnjs.2022.16102

highly efficient OER catalysts.[1] Unfortunately, these noble metal oxide catalysts suffer from poor chemical stability in alkaline media and high price, which limit their practical large-scale application as water splitting anodes. Therefore, design and development of OER catalysts with low cost and high activity has attracted considerable attention, and lots of efforts have been made.[2,3]

Cobalt oxide (Co_3O_4) has been intensely studied for electrochemical energy conversion and storage, including as an electrocatalyst for the water oxidation/OER, in lithium-ion batteries, and as supercapacitors, due to its excellent catalytic properties and high corrosion stability. The low cost and earth abundance of Co_3O_4 have suggested it to be a highly promising substitute for rare metal based catalysts.[2,3] Moreover, Co_3O_4 with different morphologies, such as hollow fluffy cage,[3] three-dimensional ordered nanoporous,[4] core-shell,[5] has been used as efficient OER catalyst. As the surface area and electronic states of Co_3O_4 are two of the main contributors for its catalytic performance, substantial work has been focused on developing new synthetic approaches to realize a variety. Additionally, mesoporous nanostructures are beneficial for promoting electrochemical performance of electrodes due to their interconnected pores, large specific surface area, controllable pore size and pore wall composition.[6,7] Template method has been considered as one of the most convenient and effective methods to prepare mesoporous inverse opal structures.

In this paper, Co_3O_4 with mesoporous inverse opal structure (Co_3O_4 IO) is fabricated on nickel foam (NF) substrates via chemical method with polystyrene (PS) beads as hard template, followed by the calcination process in air. The influence of fabrication conditions such as solution concentration, calcinated temperature on morphology and crystal structure are studied by SEM and XRD analytic techniques. The electrocatalytic properties for OER of as-synthesized materials are also studied.

2. EXPERIMENTAL

2.1. Nickel foam substrate preparation

Nickel foam susbtrate with an area of 1.0 cm x 1.0 cm was ultrasonically cleaned for 5 minutes in a 0.5 m hydrochloric acid (HCl) and for 30 minutes in ethanol (EtOH) and distilled (DI) water to remove any contaminants and amorphous surface oxides, and was dried in box furnace at 60 ^oC for 12 h in air.

2.2. Synthesis of PS nanobeads

PS spheres were prepared by emulsion polymerization using sodium dodecyl Sulfafe (SDS, Sigma Aldrich), sotassium persulfafe (PPS, Sigma Aldrich), and aqueous distilled water as emulsifier, initiator, and dispersion medium, respectively. A mixture of 8 mg SDS surfactants, 100 mg PPS initiators and 30 ml deionized water was kept to 70 °C for 30 minutes under N_2 flow in a three-necked reaction flask. 6 ml of styrene monomers, which had been filtered with alumina, was quickly injected to the solution. After polymerization at 70 °C for 4 h, the obtained PS dispersion was dialysated filtered in deionized water for 3 days to remove unreacted reagents. PS beads of 280 nm average diameter were obtained.

2.3. Co_3O_4 IO fabrication

The substrates (NF or glass subtrates) as electrodes were vertically immersed into 50 mL of PS solution which was prepared by diluting 1 ml as-prepared PS bead solution with 35 ml of distilled water and 14 ml of anhydrous ethanol. By keeping the sustrates in the PS solution at 65 °C for 48 h until the solution was completely evaporated, PS beads were uniformly coated on the electrodes. As a cobalt oxide source, cobalt nitrate hexahydrate ($Co(NO_3)_2.6H_2O$, Sigma Aldrich) was slowly added into anhydrous ethanol, and stirred at ambient condition for 1 h. After stirring, $Co(NO_3)_2.6H_2O$ solution was dropped on PS coated substrate (20 µl cm^{-2}) by micropipette. The electrodes were dried overnight in ambient conditions, followed by thermal annealing at 90 °C for 1 h in air. Finally,

these $Co(NO_3)_2.6H_2O$-treated PS/substrates were transferred to box furnace, and annealed at 450 °C for 2 h (ramping rate of 5 °C min^{-1}) in air to form Co_3O_4 IO with a removal of PS template. The synthesized process of Co_3O_4 IO is summarized in Figure 1.

In this paper, we studied different samples by varying the synthetic parameters with $Co(NO_3)_2.6H_2O$ concentration of 0.1 M; 0.2 M and 0.4 M and calcinated temperatures of 300 ºC, 450 ºC and 600 ºC. The samples with different fabricated conditions denoted as M01.450, M02.450, M04.450, M02.300, and M02.600 respectively.

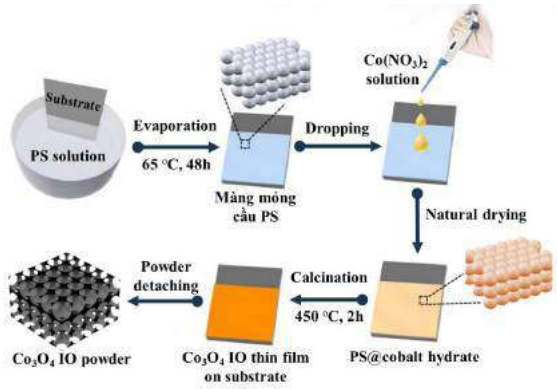

Figure 1. The fabrication process of Co_3O_4 IO

2.4. Characterization and investigation of electrochemical properties of Co_3O_4 IO electrode

Scanning electron microscopic (SEM) measurement was performed using a HITACHI S–4800 microscope to observe the morphology of Co_3O_4 IO samples. The crystal structure of as-synthesized materials was characterized by using X-ray diffraction (D8 Advance Bruker). The crystallinity and chemical structure of the Co_3O_4 IO were studied by a Raman spectroscopy (XploRA–Horiba).

To evaluate the electrochemical performance of the electrodes, experiments were carried out in a three-electrode electrochemical cell using potentiostat (DY2300). The as-prepared Co_3O_4 IO on substrate were used as working electrode, electrode, Platinum (Pt) wire and Ag/AgCl (saturated in KCl) were used as the counter and reference electrode. The electrolyte was 1.0 M KOH. For OER performamce, LSV polarization curves were recorded from 0.2 - 0.8 V versus saturated Ag/AgCl at a scan rate of 10 mV s^{-1} under N_2 medium.

3. RESULTS AND DISCUSSION

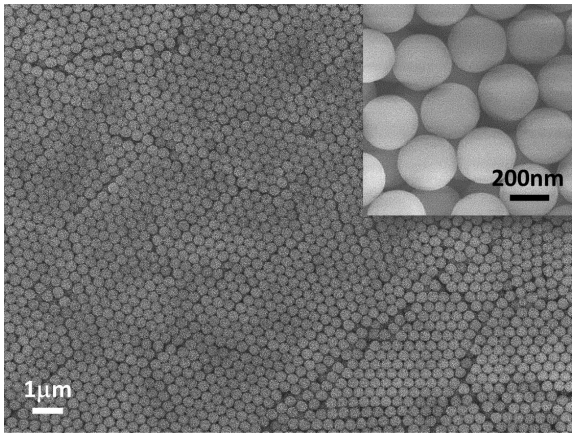

Figure 2. SEM image of PS beads deposited on glass substrate (the inset is the enlarged image of corresponding sample).

The SEM image was represented to investigate the morphology of the nanostructures involved in each step of the entire fabrication process in Figure 1. Figure 2 shows the SEM image of PS beads deposited on the glass substrate, evidencing the relatively uniform diameter, smooth surface, and well-ordered arrangement. The average diameter of the PS beads is about 270 nm.

SEM images of Co_3O_4 IO at different prepared conditions with $Co(NO_3)_2.6H_2O$ concentrations of 0.1 M; 0.2 M; 0.4 M (at the calcination temperature of 450 ºC) are shown in Figure 3.

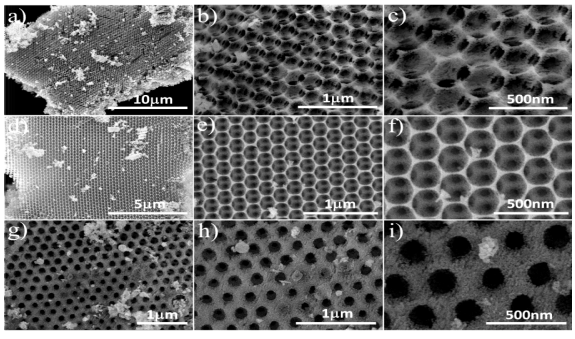

Figure 3. SEM images of Co_3O_4 IO with concentrations of $Co(NO_3)_2.6H_2O$: 0.1 M (a, b, c); 0.2 M (d, e, f); 0.4 M (g, h, i).

https://doi.org/10.52111/qnjs.2022.16102

At concentration of 0.1 M, the wall between the pores of IO structure is thin, the average diameter of the pore is approximately 300 nm (Fig. 3 a, b, c). With this morphology, the electrode has a large surface area, high porosity, but weak inter-pore adhesion leading to a fragile structure. When increasing the concentration to 0.2 M, the obtained Co_3O_4 IO structure has a fairly homogeneous and clear shape, with the honeycomb-like morphology. The thickness between the pores increases, the diameter of the pores decreases with the average value of about 270 nm (Fig. 3d, e, f). Although the surface area of material in this condition is slightly reduced, the morphology is more uniform.

Continuing to increase the concentration to 0.4 M, the thickness between the pores increased, so the surface area dramatically decreased, and the material morphology was not uniform anymore. The diameter of the pores is now greatly reduced, averaging only 170 nm (Fig. 3g, h, i). From the SEM images, it can be seen that changing the concentration of $Co(NO_3)_2.6H_2O$ has affected the morphology as well as the texture of the Co_3O_4 IO material. Based on the above analysis results, we found that 0.2 M is the optimal condition for material synthesis and research.

Calcination temperature is also an important parameter that can affect the morphology of the material. Figure 4 shows the SEM image of the Co_3O_4 IO material calcined at 600 ºC (at $Co(NO_3)_2.6H_2O$ concentration of 0.2 M). In comparison with the samples synthesized at the calcination temperature of 450 ºC (Fig. 3d, e, f), the thickness of the wall between the pores is thinner, so the structure of material easily gets colaped. We predict that if the temperature continues to increase, it is very likely that this structure will not be maintained, and the honeycomb shape will be broken. If the sample is calcinated at a temperature lower than 450 ºC, this temperature is not high enough to completely ignite the PS[8] to form a porous nanostructure. Combined with the above analysis results, the limit calcination temperature for the IO sample can be selected as 600 ºC or less, and the optimal temperature of 450 ºC was chosen in this study.

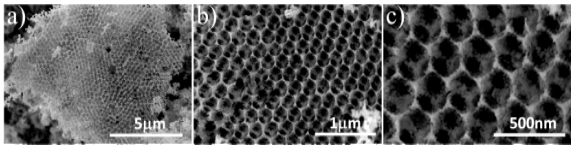

Figure 4. The SEM image of the Co_3O_4 IO material calcined at 600 ºC at different magnifications.

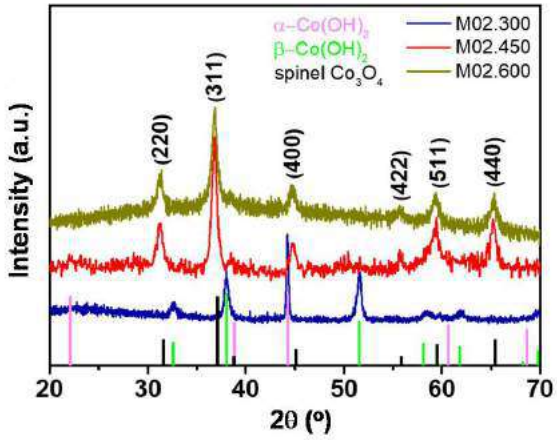

Figure 5. The XRD patterns of the Co_3O_4 IO samples at $Co(NO_3)_2.6H_2O$ concentration of 0.2 M and calcination temperature of 300 ºC, 450 ºC and 600 ºC, respectively.

The XRD patterns of the Co_3O_4 IO structures at $Co(NO_3)_2.6H_2O$ concentration of 0.2 M and calcination temperature of 300 ºC, 450 ºC and 600 ºC are shown in Figure 5. XRD results reveal that, at the 300 ºC, the material only exists phase α, β of $Co(OH)_2$ (JCPDS No. 30-443). At the 450 ºC and 600 ºC, XRD patterns exhibit diffraction peaks at 31.2º, 36.8º, 44.6º, 55.5º, 59.3º and 65.2º, which are assigned to (220), (311), (400), (422), (511) and (440) faces of cubic crystallite Co_3O_4 (JCPDS No.42–1467). No peaks ascribed to other impurities can be found, revealing the high purity of as-synthesized spinel oxides. This result indicates that cobalt precursor has been completely transformed into crystalline Co_3O_4 at the temperature higher 450ºC.

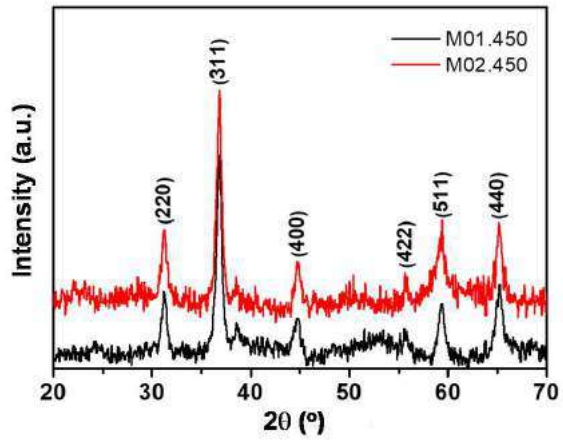

Figure 6. The XRD patterns of Co_3O_4 IO at calcination temperature of 450 °C and $Co(NO_3)_2.6H_2O$ concentrations of 0.1 M and 0.2 M.

Figure 6 shows the XRD profiles of samples synthesized at $Co(NO_3)_2.6H_2O$ concentration of 0.1 M and 0.2 M and calcination temperature of 450 ºC (M01.450 and M02.450). The XRD patterns of two samples at different concentrations has exactly the same peak positions and are confirmed to agree with reported spinel cobalt oxide. This result indicates that $Co(NO_3)_2.6H_2O$ concentration does not affect the crystal structure of the as-synthesized material.

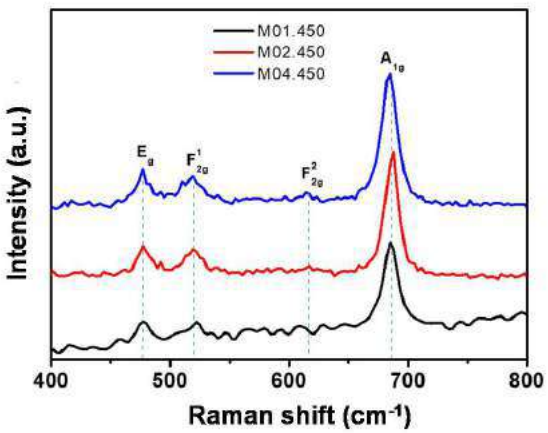

Figure 7. Raman spectra of samples synthesized at different $Co(NO_3)_2.6H_2O$ concentrations and calcination temperature of 450 °C (M01.450, M02.450, and M04.450).

Raman spectra of samples samples synthesized at different $Co(NO_3)_2.6H_2O$ concentrations and calcination temperature of 450 °C (M01.450, M02.450, and M04.450) are presented in Figure 7. As revealed by these results, there are four Raman-active modes at 482, 521, 618, and 690 cm^{-1} attributed to the E_g, F^1_{2g}, F^2_{2g}, and A^1_g vibration modes of the cubic structure of Co_3O_4,[9] further confirms the formation of Co_3O_4 spinel, which is corroborated by the XRD. From the XRD and raman results, it is clearly revealing that cobalt salt concentration affects only the morphology (SEM results) without changing the crystal structure of Co_3O_4 IO.

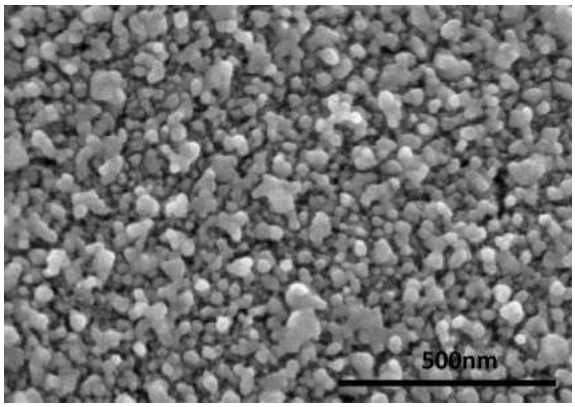

Figure 8. SEM image of the Co_3O_4 nanoparticle was synthesized by hydrothermal method.

To evaluate the electrochemical property of Co_3O_4 IO, we also synthesized Co_3O_4 nanoparticle materials (denoted as Co_3O_4 NP) as the reference by hydrothermal method at 120 ºC, 12 h. The cobalt precursor is also a cobalt nitrate salt, but without the PS template. The morphology of the Co_3O_4 nanoparticles is shown in Figure 8.

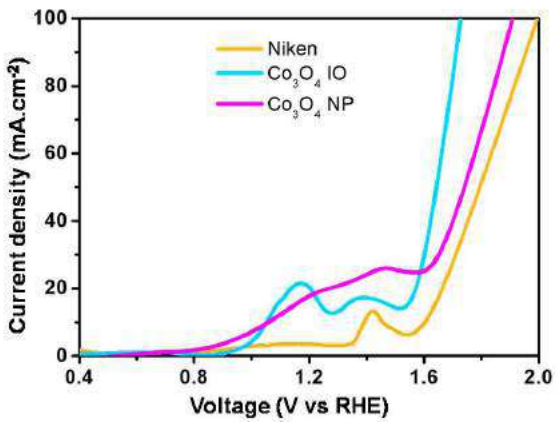

Figure 9. LSV characteristics of Co_3O_4 IO (sample M02.450) and Co_3O_4 NP as the reference.

https://doi.org/10.52111/qnjs.2022.16102

Figure 9 compares LSV characteristics of Co_3O_4 IO and Co_3O_4 NP materials fabricated on NF substrates. The graphs show that the overpotential values (η) of the NF substrate, Co_3O_4 IO, and Co_3O_4 NP (at the current density of 30 $mA.cm^{-2}$) is 1.71 (V), 1.65 (V), and 1.60 (V), respectively. Thus, compared with the Co_3O_4 NP, the Co_3O_4 IO material needs a much lower overpotential value. Moreover, the LSV characteristic curve of the Co_3O_4 IO shows higher slope, demonstrating the OER reaction rate of the Co_3O_4 IO sample is much faster than that of than that of Co_3O_4 NP counterpart. The above results confirm the outstanding catalytic performance of the Co_3O_4 IO sample.

This result can be explained as follows: The NF substrate has a three-dimensional (3D) porous structure with high surface area, whereas the distribution of the catalyst on the entire substrate is uniform, which greatly increases the catalytic active sites of the material. Meanwhile, the regular empty spaces, surrounded by the mesoporous walls composes of tiny nanoparticles (size 5 - 7 nm), are interconnected by the pores/voids where the removed PS spheres have resided in contact with each other. The interconnected voids can act as diffusion channels, and hence, facilitating the charge exchange process, increasing the conductivity of the Co_3O_4 IO catalyst. Thus, the high catalytic performance of Co_3O_4 IO is demonstrated to be mainly attributed to both high surface area and good electrical conductivity. These results seem to hint at the interest in Co_3O_4 IO as an electrochemical catalyst for the OER process.

4. CONCLUSION

The Co_3O_4 ordered porous architecture, so-called inverse-opal structure has been successfully fabricated by chemical method using PS spheres as hard template, followed by calcination process in air. The morphology of the IO structute changes in terms of pore diameter and wall thickness of interconnected voids with the concentration of $Co(NO_3)_2.6H_2O$ and the calcination temperature. XRD patterns confirm the as-prepared IO material has a cubic crystal structure of Co_3O_4. We believe this work represents a valuable addition focusing on the importance of the IO structure to improve the electrocatalytic performance of Co_3O_4-based electrodes in OER. This contribution is therefore expected to further encourage studies of electrode design in electrochemical water-splitting, and metal-air batteries such as lithium-air batteries, and zinc-air batteries.

Acknowlegment

This work is funded by the Vietnam Ministry of Education and Training, under grant number B2020-DQN-06.

REFRENCES

1. E. G. L. Michael, F. Stephane. Mechanism of oxygen reactions at porous oxideelectrodes. Part 2—Oxygen evolution at RuO_2, IrO_2 and $Ir_xRu_{1-x}O_2$ electrodes in aqueous acid and alkaline solution, *Physical Chemistry Chemical Physics,* **2011**, *13*, 5314-5335.
2. M. Bajdich, M. Garcia-Mota, A. Vojvodic, J. K. Norskov, A. T. Bell. Theoretical investigation of the activity of cobalt oxides for the electrochemical oxidation of water, *Journal of the American Chemical Society*, **2013**, *135*, 13521–13530.
3. J. Rosen, G. S. Hutchings, F. Jiao. Ordered mesoporous cobalt oxide as highly efficient oxygen evolution catalyst, *Journal of the American Chemical Society*, **2013,** *135,* 4516–4521.
4. W. Y. Xia, N. Li, Q. Y. Li, K. H. Ye, C. W. Xu. Au-$NiCo_2O_4$ supported on threedimensional hierarchical porous graphene-like material for highly effective oxygen evolution reaction, *Scientific Reports*, **2016,** *6*.
5. Z. Zhuang, W. Sheng, Y. Yan. Synthesis of monodispere Au@Co_3O_4 core-shell nanocrystals and their enhanced catalytic activity for oxygen evolution reaction, *Advanced Materials*, **2014,** *26*, 3950 - 3955.

6. X. Zou, J. Su, R. Silva, A. Goswami, B. R Sathe, T. Asefa. Efficient oxygen evolution reaction catalyzed by low-density Ni-doped Co_3O_4 nanomaterials derived from metalembedded graphitic C_3N_4, *Chemical Communications*, **2013**, *49*, 7522 - 7524.

7. Seol A Cho, Yu Jin Jang, Hee-Dae Lim, Ji-Eun Lee, Yoon Hee Jang, Trang Thi Hong Nguyen, Filipe Marques Mota, David Fenning, Kisuk Kang, Yang Shao-Horn, Dong Ha Kim. Hierarchical Porous Carbonized Co_3O_4 Inverse Opals via Combined Block Copolymer and Colloid Templating as Multifunctional Electrocatalysts and Cathodes in LiO_2 Battery, *Advanced Energy Materials,* **2017***, 7*.

8. C. Konstantinos, E. EleniPavlidou, D. B. Vouvoudi. Decomposition kinetic and mechanism of syndiotactic polystyrene nanocomposites with MWCNTs and nanodiamonds studied by TGA and Py-GC/MS, *Thermochimica Acta*, **2014,** *583*, 15-24.

9. F. Gu, C. Li, Y. Hu, L. Zang. Synthesis and optical characterization of Co_3O_4 nanocrystals, *Journal of Crystal Growth*, **2007**, *304,* 369-373.

https://doi.org/10.52111/qnjs.2022.16102

TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC QUY NHƠN

Xác định dư lượng kháng sinh amoxicillin trong nước thải nuôi trồng thủy sản bằng phương pháp điện hóa sử dụng điện cực nano platin trên nền glassy cacbon

Nguyễn Thị Liễu[1,*], Lê Thu Hương[1], Võ Thị Hoa Trâm[1], Phạm Thị Hải Yến[2]

[1]Khoa Khoa học tự nhiên, Trường Đại học Quy Nhơn, Việt Nam
[2]Viện Hóa học, Viện Hàn lâm Khoa học và Công nghệ Việt Nam, Cầu Giấy, Hà Nội, Việt Nam

Ngày nhận bài: 27/09/2021; Ngày nhận đăng: 21/12/2021

TÓM TẮT

Platin nano hình hoa (Platinum nanoflowers) được kết tủa lên bề mặt điện cực glassy cacbon (GC) để hình thành điện cực mới (PtNFs/GCE). Đặc tính của điện cực PtNFs/GCE cho việc xác định Amoxicillin (AMX) bằng phương pháp Von - Ampe hòa tan anot sóng vuông (SWASV) đã được nghiên cứu. Sau khi tối ưu hóa các điều kiện thí nghiệm, cường độ dòng đỉnh hòa tan AMX cao nhất thu được trong dung dịch đệm photphat 0,2 M (pH = 7,5) với thời gian làm giàu 150 s, biên độ xung 60 mV và tốc độ quét là 0,25 $V.s^{-1}$. Giới hạn phát hiện và giới hạn định lượng của phương pháp phân tích được xác định lần lượt là 0,113 µM và 0,372 µM. Trong điều kiện thích hợp, phương pháp đạt được độ tái lặp cao: RSD = 9,65% (n=5). Ảnh hưởng của một số kháng sinh nhóm β-lactam và các chất gây nhiễu có thể có khác đã được nghiên cứu. Cuối cùng, điện cực PtNFs/GCE đã được áp dụng để xác định dư lượng kháng sinh trong nước thải nuôi trồng thủy sản với kết quả hoàn toàn phù hợp về mặt thống kê với phương pháp sắc ký lỏng hiệu năng cao (HPLC).

Từ khóa: *Glassy cacbon, platinum nanoflowers, kết tủa điện hóa, Von – Ampe hòa tan anot sóng vuông, Amoxicillin.*

**Tác giả liên hệ chính.*
Email: nguyenthilieu@qnu.edu.vn

JOURNAL OF SCIENCE
QUY NHON UNIVERSITY

Determination of Amoxicillin antibiotic residues in aquaculture wastewater by the electrochemical method using platinum nanoflowers modified glassy carbon electrode

Nguyen Thi Lieu[1,*], Le Thu Huong[1], Vo Thi Hoa Tram[1], Pham Thi Hai Yen[2]

[1]Faculty of Natural Sciences, Quy Nhon University, Vietnam
[2]Institute of Chemistry, Vietnam Academy of Science and Technology, Cau Giay, Hanoi, Viet Nam

Received: 27/09/2021; Accepted: 21/12/2021

ABSTRACT

The glassy carbon electrode electrodeposited platinum nanoflowers on the surface to form a new electrode (PtNFs/GCE). The behavior of PtNFs/GCE has been studied for Amoxicillin (AMX) determination by square wave anodic stripping voltammetry (SWASV). After optimizing the experimental and voltammetric conditions, the best stripping peak obtained current intensity in 0.2 M phosphate buffer solution (pH = 7.5) with 150 s accumulation time, 60 mV pulse amplitude, and 0.25 $V.s^{-1}$ scan rate. The limit of detection and quantity was observed as 0.113 µM and 0.372 µM, respectively. Under suitable conditions, the method gained high reproducibility: RSD = 9.65% (n=5). Interference effects of some β-lactam antibiotics and other possible interferent were studied. Finally, the PtNFs/GCE was successfully applied to analyze Amoxicillin antibiotic residues in aquaculture wastewater, resulting in satisfactory statistical agreement with high-performance liquid chromatography (HPLC).

Keywords: *Glassy carbon, platinum nanoflowers, electrodeposition, square wave anodic stripping voltammetry, Amoxicillin.*

1. INTRODUCTION

Antibiotics are one of the most important drugs used in medicine because they can inhibit or kill bacteria that cause infectious diseases in humans and animals.[1,2] Amoxicillin (AMX) is the most used antibiotic from the β-lactam group to treat bacterial infections that affect humans and other animals. After its action mechanism, 60 - 70% of the drug is excreted and unaltered through urine and feces because of its low metabolism;[3,4] thus, Amoxicillin is rapidly released into the environment.[5] Because of its long persistence in the environment, the drug is considered a severe contaminant is leading to adverse effects on living organisms.

Accordingly, it is necessary to design a quantitative detection approach for AMX. Amoxicillin has been determined by spectrophotometry,[6] capillary electrophoresis,[7] and high-performance liquid chromatography.[8] However, they still have some disadvantages such as complicated procedures, time-consuming, expensive instrument requirement, and low detection sensitivity. Recently, the electrochemical technique has become a promising method used in medicine

Corresponding author.
Email: nguyenthilieu@qnu.edu.vn

https://doi.org/10.52111/qnjs.2022.16103

determination because of its high sensitivity and selectivity, fast measurement, low-cost instrument, and simple operation.

In this paper, the authors used the electrochemical technique to investigate the electrochemical behavior of Amoxicillin which is also the first use of the electrochemical method in Vietnam. This research reported the application of platinum nanoflowers modified glassy carbon electrodes (PtNFs/GCE) to detect AMX. It has numerous advantages in electrochemical analysis such as stability, easy fabrication, high conductivity, low background current, and reproducibility.[9,10] By using the electrodes, we obtained good electrochemical signals of AMX.

2. MATERIALS AND METHODS

2.1. Reagents and Apparatus

Reagents: Hexachloroplatinic (IV) acid hexahydrate ($H_2PtCl_6.6H_2O$) used for fabrication of PtNFs/GCE was of pure grade (KGaA, 64271 Darmstadt, Germany). The electrolyte solution comprises 1.0 mM H_2PtCl_6, 0.1 M H_2SO_4 solution, and a $[Fe(CN)_6]^{3-}$ used redox system in 0.2 M phosphate buffer solution pH = 7 for the study of electrochemical properties. Amoxicillin trihydrate ($C_{16}H_{19}N_3O_5S{\cdot}3H_2O$, AMX) was purchased from Merck (Germany). AMX was prepared daily from a 1 mM stock solution. Phosphate buffer solution (PBS) is the supporting electrolyte of electrochemical measurements.

Apparatus: Square wave anodic stripping voltammetry (SWASV) using PtNFs/GCE were performed on a system Autolab Electrochemical CPA-HH5 (Hanoi, Vietnam), with a three-electrode configuration (GCE and PtNFs/GCE as a working electrode, Ag/AgCl reference electrode, and platinum wire as an auxiliary electrode).

2.2. Preparation of modified electrode

The glassy carbon (GC, d = 3.0 mm) electrode was polished with a water slurry of 0.2 µm Al_2O_3, rinsed with ethanol and water, and then dried. The electrodeposition of platinum nanoparticles on the bare glassy carbon electrode was carried out in an H_2SO_4 0.1 M solution containing 1.0 mM H_2PtCl_6 at a constant potential of -0.2 V and deposition time of 150 s. Following that, the Pt/GCE was gently cleaned with distilled water before use.

2.3. Electrochemical measurements

The formation of PtNFs/GCE and their properties were investigated by the cyclic voltammetric method (CV). Measurements performed in 5 mM $[Fe(CN)_6]^{3-}$ from -0.3 V to 0.7 V at a scan rate of 0.1 $V.s^{-1}$ were used to predict electrochemical features of electrodes and compare their electrochemically active surface areas. The analytical solution containing AMX, 0.2 M phosphate buffer (PBS, pH = 7.5) was put into a cell with the three electrodes. It was determined by square wave voltammetry (SWV) with (0.3, 1.1 V) of the potential range.

All experiments described in this section were performed at room temperature (25 ± 1 ^oC).

3. RESULTS AND DISCUSSION

3.1. Morphological characterization of the GCE, PtNFs/GCE by SEM

The surface morphology of the bare GCE, PtNFs/GCE are investigated by SEM with different scales. The SEM image shows that the GCE surface has a smooth and homogeneous morphology (Fig. 1a). However, the surface of the electrode's appearance changed dramatically after PtNFs/GCE (Fig. 1b). According to the SEM images (Fig. 1b), the GCE surface is densely developed with Pt nanoflowers. After electrodeposition, many Pt nanoflowers are produced on the GCE surface; the surface of PtNFs/GCE was rougher with irregular cubic nanoflowers shape whose size varies in the range (50 – 400 nm). Electrochemical methods also characterize the PtNFs/GCE.

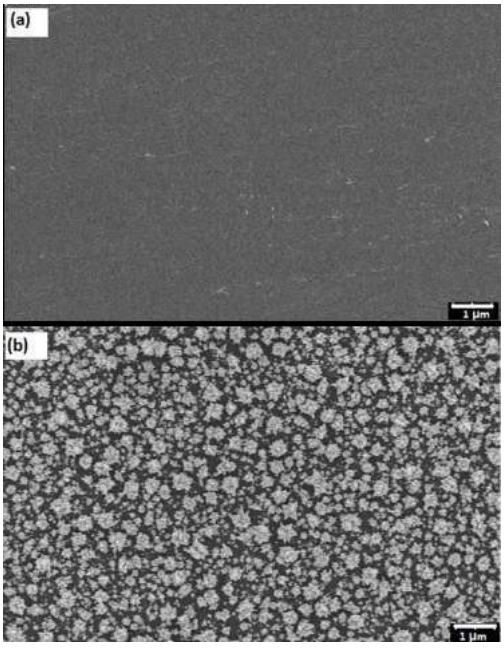

Figure 1. SEM images of GCE (a), PtNFs/GCE(b)

3.2. Formation and characterization of PtNFs/GCE

The PtNFs modified the GCE in 0.1 M H_2SO_4 containing 1.0 mM H_2PtCl_6 solution by a cyclic voltammetry sweep in the range between -0.3 and 1.4 V. Fig. 2 shows the cyclic voltammetric responses of a PtNFs/GCE and bare GCE in 0.2 M phosphate buffer solution (PBS), pH, 7.0 containing 5 mM $K_3Fe(CN)_6$ at a scan rate of 0.1 $V.s^{-1}$.

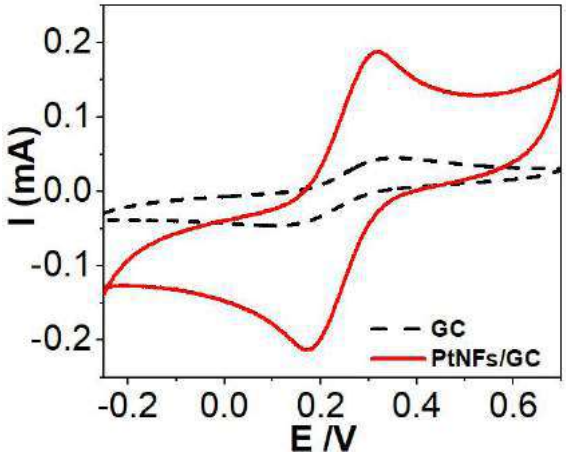

Figure 2. Cyclic voltammograms obtained at a bare GCE and PtNFs/GCE in 0.2 M phosphate buffer solution (PBS), pH, 7.0 containing 5 mM $K_3Fe(CN)_6$ at a scan rate of 0.1 $V.s^{-1}$.

It could be found that the electrochemical peaks of the bare GCE in the PBS are low. Their electrochemical peaks could also be observed at the PtNFs/GCE, with the intensity increasing significantly. The SEM images that well-dispersed nanoscale Pt particles were deposited on the GCE.

We determined the electrochemically active surface areas (AEAS) of GCE and the PtNFs/GCE with cyclic voltammetry method based on Randles – Sevcik equation[11]:

$$I_p = 2.69\times10^5\times n^{3/2}\times A_{EAS}\times D^{1/2}\times v^{1/2}\times C$$

According to the above equation, the A_{EAS} of the PtNFs/GCE fabricated with the electrodeposition time 150 s under stirring 50 rpm was calculated as 0.184 cm^2 larger than the bare GCE (0.044 cm^2).

The results clearly explain that the modified electrode with platinum nanoflowers is beneficial for improved sensitivity in analysis thanks to the larger active surface area. Thus, the intensity of the stripping peak is increased by almost 4 times that of the bare GCE; AMX could be detected sensitively at the PtNFs/GCE. Therefore, the PtNFs/GC electrode allows for improved selectivity and sensitivity.

3.3. Optimization of experimental parameters

3.3.1. Effect of solution pH

For the development of the analytical method, the influence of the pH of the supporting electrolyte (hydrogen ion concentration) on the oxidation reaction of AMX (20 µM) was simultaneously evaluated by the SWV technique using the PtNFs/GC electrode.

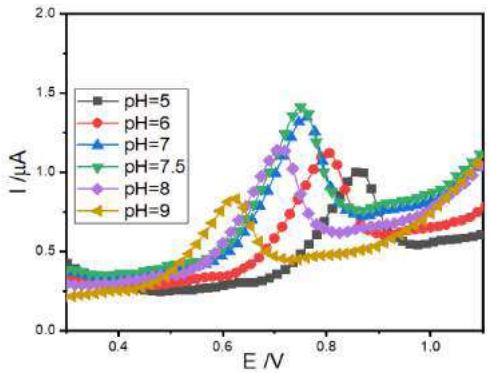

Figure 3. Effects of pH on the stripping peak current of 20 µM AMX on PtNFs/GCE

https://doi.org/10.52111/qnjs.2022.16103

The pH was investigated in the range of 5.0 - 9.0 with 0.2 M phosphate buffer. The SW potential plots obtained in this investigation are shown in Fig 3. It can be seen that the oxidation peak current for AMX increased until it peaked at pH 7.5 when the pH increased. Therefore, we chose pH 7.5 as the optimal pH value for subsequent experiments.

3.3.2. Effect of accumulation time

There was a linear correlation between the AMX peak current (I_p) and accumulation time (t_{dep}) in the range of 60 - 240 s. The result (Fig. 4) shows that the peak current increased with an increase in accumulation time from 60 to 150 s, which can be attributed to the rapid adsorption of AMX on the surface of the electrode. However, the oxidation peak current was leveled off with a further increase in accumulation time beyond 150 s. It can be ascribed to the saturation of AMX on the surface of the electrode. Consequently, the accumulation time of 150 s was suitable.

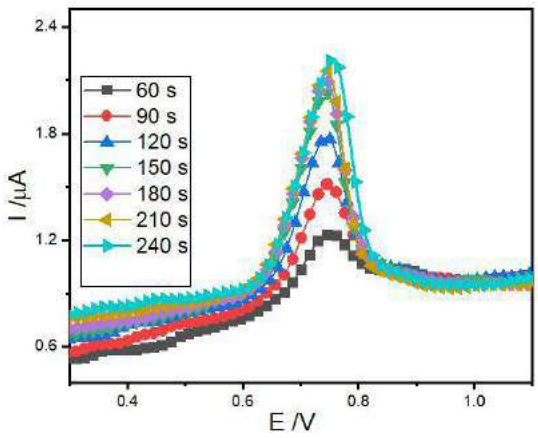

Figure 4. Effects of accumulation time on the stripping peak current of 20 µM AMX on PtNFs/GCE

3.3.3. Effect of scan rate

The Effect of scan rate on the AMX electrochemical responses was investigated by the SWASV technique, varying the potential scan rate from 0.05 to 0.5 V/s. The obtained results show that, as the scan rate increases, the oxidation peak current for AMX increases until it peaks at 0.25 V/s scan rate, the oxidation peak current for AMX declines slightly. Hence, recorded the best peak definition at when 0.25 V/s scan rate.

3.4. Reproducibility, linear range, and limit of detection

Reproducibility: Under the suitable conditions (0.2 M phosphate buffer (pH = 7.5); 20 µM AMX; accumulation time = 150 s: ΔE = 60 mV; v = 0.25 V/s; E_{range} = 0.3 V to 1.1 V, the SWASV with PtNFs/GCE gained high reproducibility of AMX stripping responses. The RSD value of reproducibility was calculated to be 9.65% indicating that the fabrication procedure was reliable (n = 5).

Linear range: The wide linearity range was good in the range of 0.5 – 60.0 µM, with a correlation coefficient of 0.996. The SWASVs and linear regression lines/equations for the AMX are shown in Fig. 5a and 5b, respectively.

Limit of detection (LOD): The limit of detection (LOD) was calculated as 3σ. The detection limits of 0.113 µM were estimated. With the low detection limit mentioned above, the SWASV using PtNFs/GCE can directly detect AMX in the local aquaculture wastewater without a concentration step before analysis.

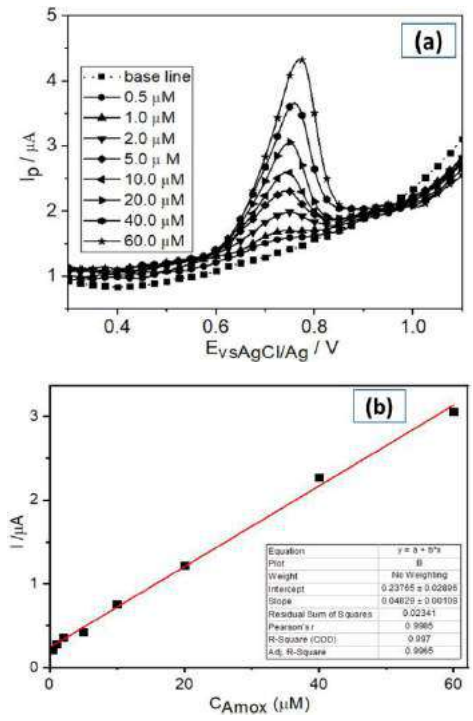

Figure 5. SWVs of AMX samples ranging in concentration from 0.5 to 60.0 µM (a) and relationship between AMX concentrations with corresponding peak current (b)

3.5. Interference study

Some β-lactam antibiotics and another possible interference in the local aquaculture wastewater samples may be adsorbed onto the electrode surface and then influence the stripping response of AMX. We conducted experiments with the presence of variable concentrations of some β-lactam antibiotics and another possible interferent in a solution containing 20 µM AMX in 0.2 M phosphate buffer (pH 7.5). The results show that the maximum concentration of interfering species causes an error of less than ± 12,6% to determine AMX. The results showed that the concentration of ampicillin, penicillin G exceeded 50 times, and the concentration of K^+, Na^+, Ca^{2+}, Mg^{2+}, Fe^{2+} exceeded 100 times did not affect the determination of the 20 µM AMX. The proposed sensor selectivity for AMX is based on the fact that AMX is the only phenolic penicillin and a medium-spectrum β-lactam antibiotic and that electro-oxidation occurs at the p-hydroxy substituent of AMX[13,14] that is absent from other members of the β-lactam group and also from all other possible inhibitors.

3.6. Analysis of real samples

The SWASV procedure using PtNFs/GCE with suitable conditions was applied to determine AMX in the local aquaculture wastewater at Binh Dinh province, Central Viet Nam.

Some organic compounds affecting the determination of AMX in aquaculture wastewater are removed by the extraction method. Then, all water samples were added to 0.2 M phosphate buffer (pH = 7.5) and determined optimal conditions were. The results in experimental values are shown in Table 1. We also compared the results with the same samples determined from high-performance liquid chromatography (HPLC). As shown in Table 1, the comparative results suggest that the concentrations of AMX were in good agreement with the results found in the HPLC measurement. The method reported in this work is also highly sensitive, reproducible, and accurate. Therefore, the PtNFs/GC electrode is potentially helpful for the analysis of AMX in real samples.

Table 1. Analytical results for the determination of AMX (n = 3) in real water samples

Sample	AMX (µM)			
	Sample location	SWASV	HPLC	Error (%)
Samples of wastewater from Tuy Phuoc shrimp farming	13°50'15.9"N 109°12'46.9"E	1.32 ±0.33	1.21±ND	9.09
Samples of wastewater from Tuy Phuoc shrimp farming	13°50'17.9"N 109°12'50.3"E	1.04±0.36	1.16 ±ND	10.34
Samples of wastewater from Phu My shrimp farming	14°21'51.4"N 109°07'18.3"E	0.76±0.42	0.88±ND	13.64
Samples of wastewater from Phu My shrimp farming	14°23'12.2"N 109°06'45.6"E	0.93±0.27	0.85±ND	9.41

SWASV: square wave anodic stripping voltammetry; HPLC: high-performance liquid chromatography. ND: Not determined.

4. CONCLUSIONS

In this study, a new electrochemical sensor for the determination of AMX antibiotic residues was fabricated simply by the electrochemical method. Traditional electrochemistry and microscopic analysis methods have been applied to characterize platinum nanoflowers electrodeposited on GCE. The electrochemical signals of AMX increased by using PtNFs/GCE. It helps produce the electrochemical Sensor for the determination of AMX at trace concentration.

https://doi.org/10.52111/qnjs.2022.16103

A simple, sensitive, and inexpensive method for the determination of AMX is proposed. The PtNFs/GCE demonstrated better detection sensitivity and higher SWASV signals than the bare GCE, with a detection limit of about 0.113 μM.

Acknowledgements

This research is conducted within the framework of science and technology projects at the institutional level of Quy Nhon University under the project code T2021.702.05.

REFERENCES

1. J. Wang, J. D. MacNeil, J. F. Kay, Chemical Analysis of Antibiotic Residues in Food, *John Wiley & Sons*, New York, **2011**, 1-60.
2. P. B. Deroco, R. C. Rocha-Filho, and O. Fatibello-Filho. A new and simple method for the simultaneous determination of Amoxicillin and nimesulide using carbon black within a di-hexadecyl phosphate film as an electrochemical sensor, *Talanta*, **2018**, *179*, 115-123.
3. G. Yang and F. Zhao. Molecularly imprinted polymer grown on multiwalled carbon nanotube surface for the sensitive electrochemical determination of Amoxicillin, *Electrochimica Acta*, **2015**, *174*(1), 33-40.
4. M. G. P. Valenga, M. L. Felsner, C. F. de Matos, E. G. de Castro, and A. Galli. Development and validation of a voltammetric method for determination of Amoxicillin in river water, *Analytica Chimica Acta*, **2020**, *1138,* 79–88.
5. N. Ul Ain, I. Anis, F. Ahmed, M. R. Shah, S. Parveen, S. Faizi, S. Ahmed. Colorimetric detection of Amoxicillin based on querecetagetin coated silver nanoparticles, *Sensors and Actuators B: Chemical*, **2018**, *265*, 617-624.
6. A. H. Hamed, S. K. Ibrahim, S. A. Dhahir, and A. H. Mhemeed. Indirect Spectrophotometric Determination of Amoxicillin by Oxidative with Cerium ammonium Nitrate Using Arsenazo III as a Reagent, *Chemical Science Transactions*, **2019**, *8*(2), 1617-1620.
7. Zhou X, Xing D, Zhu D, Tang Y, Jia L. Development and application of a capillary electrophoresis-electrochemiluminescent method for the analysis of enrofloxacin and its metabolite ciprofloxacin in milk, *Talanta*, **2008**, 75, 1300-1306.
8. S. M. Foroutan, A. Zarghi, A. Shafaati, A. Khoddam & H. Movahed. Simultaneous determination of amoxicillin and clavulanic acid in human plasma by isocratic reversed-phase HPLC using UV detection, *Journal of Pharmaceutical and Biomedical Analysis*, **2007**, *45*(3), 531-534.
9. L. Su, W. Jia, L. Zhang, C. Beacham, H. Zhang, Y. Lei. Facile synthesis of a platinum nanoflower monolayer on a single-walled carbon nanotube membrane and its application in glucose detection, *The Journal Physical Chemistry C,* **2010**, *114(42),* 18121-18125.
10. C. M. Welch, R. G. Compton. The use of nanoparticles in electroanalysis: a review, *Analytical and Bioanalytical Chemistry*, **2006**, *384(3)*, 601-619.
11. Neto, B. B., Scarminio, I. S., and Bruns, R. E. Chapter 6 Exploring the response surface, Data Handling in Science and Technology, 2005, 245-312.
12. Di Zhao, Y., De Zhang, W., and Li, S. F. Y. The oxidation and reduction behavior of nitrite at carbon nanotube powder microelectrodes, *Microchemical Journal,* **2003**, *75,* 189-198.
13. B. Rezaei and S. Damiri. Electrochemistry and adsorptive stripping voltammetric determination of Amoxicillin on a multiwalled carbon nanotubes modified glassy carbon electrode, *Electroanalysis*, **2009**, *21*(14), 1577-1586.
14. D. P. Santos, M. F. Bergamini, and M. V. B. Zanoni. Voltammetric Sensor for amoxicillin determination in human urine using polyglutamic acid/glutaraldehyde film, *Sensors and Actuators B Chemical*, **2008**, *133*(2), 398-403.

TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC QUY NHƠN

Điều khiển các bộ điện tử công suất của máy phát điện không đồng bộ nguồn kép không chổi than sử dụng máy biến áp quay

Lê Thái Hiệp*, Nguyễn An Toàn

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

Ngày nhận bài: 20/08/2021; Ngày nhận đăng: 01/11/2021

TÓM TẮT

Các loại máy phát điện không đồng bộ nguồn kép không chổi than (BDFIG) đã được chứng minh là giảm thiểu được chi phí bảo dưỡng và chi phí vận hành, trong khi lại tăng độ tin cậy so với máy phát điện không đồng bộ nguồn kép dùng trong điện gió. Bài báo này đã đề xuất phương pháp điều khiển các bộ điện tử công suất của máy phát điện không đồng bộ nguồn kép không chổi than sử dụng máy biến áp quay (BDFIG-RT) trong vai trò làm máy phát cho tuabin điện gió loại tốc độ thay đổi. Bộ điều khiển được đề xuất dựa trên hệ trục dq gắn trên rôto, và được mô phỏng trên phần mềm Matlab Simulink. Các kết quả mô phỏng cho thấy rằng đáp ứng động học hoàn toàn tốt trong trường hợp tốc độ gió thay đổi, kể cả có xảy ra sự cố ngắn mạch trên lưới. Nhờ phương pháp điều khiển này mà trong quá trình xảy ra ngắn mạch trên lưới thì máy phát vẫn kết nối với lưới và hoạt động ổn định.

Từ khóa: *Máy phát điện không đồng bộ nguồn kép không chổi than sử dụng máy biến áp quay, bộ biến đổi phía lưới, bộ biến đổi phía rôto, bộ điện tử công suất, máy biến áp quay.*

**Tác giả liên hệ chính.*
Email: lethaihiep@qnu.edu.vn

https://doi.org/10.52111/qnjs.2022.16104

JOURNAL OF SCIENCE
QUY NHON UNIVERSITY

Control of power converters in the brushless doubly fed induction generator with rotary transformer

Le Thai Hiep*, Nguyen An Toan

Faculty of Engineering and Technology, Quy Nhon University, Vietnam

Received: 20/08/2021; Accepted: 01/11/2021

ABSTRACT

Brushless doubly fed induction generators (BDFIGs) show commercial promise for wind-power generation due to their lower capital and operational costs and higher reliability as compared with doubly fed induction generators. This paper proposed a control method of power converters in the brushless doubly fed induction generator with rotary transformer (BDFIG-RT), which operate as a variable-speed generator. The proposed controller is designed on the rotating dq reference frame, and simulated in Matlab Simulink. The simulation results show that the proposed controller has good dynamic performance when changes in wind speed and short circuit fault on grid. Thanks to this control system, the generator is still connected to the grid, and operates stably during a short circuit fault.

Keywords: *Brushless doubly fed induction generator with rotary transformer, grid-side converter, rotor-side converter, power converter, rotary transformer.*

1. INTRODUCTION

Currently, there are many types of wind generators, but the type of doubly fed induction generator (DFIG) is being used commonly. This type of generator is more popularly used due to the advantages of installation cost and wide wind energy exploitation range from 4 m/s to 24 m/s. This range is from less than 30% to over 30% of the synchronous speed, respectively. However, every DFIG needs to use carbon brushes to make electrical connections between the rotor windings and power supply from the power electronics. Using carbon brushes will cause sparks, contact resistance, wear of brushes and slip rings, heat generation at the point of contact, etc. These problems reduce the reliability of a DFIG. There are many methods of manufacturing Brushless Doubly Fed Induction Generators (BDFIGs). Particularly, they are divided into main groups:

- Unified BDFIG is in just one machine. This BDFIG type has two stators (the electrical parts and also the magnetic parts are separated) and a common rotor (the electrical parts are connected together, but the magnetic parts are separated).[1-4] Similar to this type is a brushless doubly-fed reluctance generator.[5] In these types, windings of a stator generating power are directly connected to the grid, but the windings of the other stator for control are connected to the grid through power electronics.[1-8]

- Another type of BDFIG is Brushless Doubly Fed Induction Generator with Rotary Transformer (BDFIG-RT). The main part of this type is DFIG. But instead of the three-phase

* *Corresponding author.*
Email: lethaihiep@qnu.edu.vn

https://doi.org/10.52111/qnjs.2022.16104

windings of the generator rotor are connected to the slip rings and carbon brushes, they are connected to the rotor windings of a three -phase rotating transformer. The stator windings of this transformer are powered by a power electronics unit.[1,9-12]

- A handful of BDFIGs are combinations of two blocks from two machines. This BDFIG type, stator windings of main generator are directly connected to the grid, and rotor windings of this generator are supplied by an auxiliary generator via a rotated inverter.[1,13] This auxiliary generator can be a wound rotor induction generator, the rotor windings connected to a rotated inverter.[13] Or the BDFIG type is formed from two wound rotor induction generators with their rotor parts connected both mechanically and electrically.[1,14]

Among the BDFIG types, the BDFIG-RT type has the simplest structure, very similar to the traditional DFIG type. Therefore, BDFIG-RT has the prospect of commercialization and application to replace DFIG in the near future.

There have been many studies on BDFIG-RT, but no detailed study focused on the control of this machine. Therefore, the purpose of this study is to analyze and propose control loops for the power converters of BDFIG-RT.

2. THE PROPOSED CONTROL STRUCTURES

2.1. Operational characteristics of the BDFIG-RT

The BDFIG type that uses a rotary transformer to supply electric power to the rotor windings (as shown in Figure 1). The generator stator windings in the BDFIG-RT are connected directly to the grid, but the generator rotor windings are connected to the rotor windings of a rotary transformer. Besides, the stator windings of this transformer are powered by a rotor-side converter (RSC) combined with a grid-side converter (GSC) connected to the grid (as shown in Figure 2). The rotary transformer is designed to be suitable for generators in the form (a) or the form (b) as shown in Figure 1.[9-12]

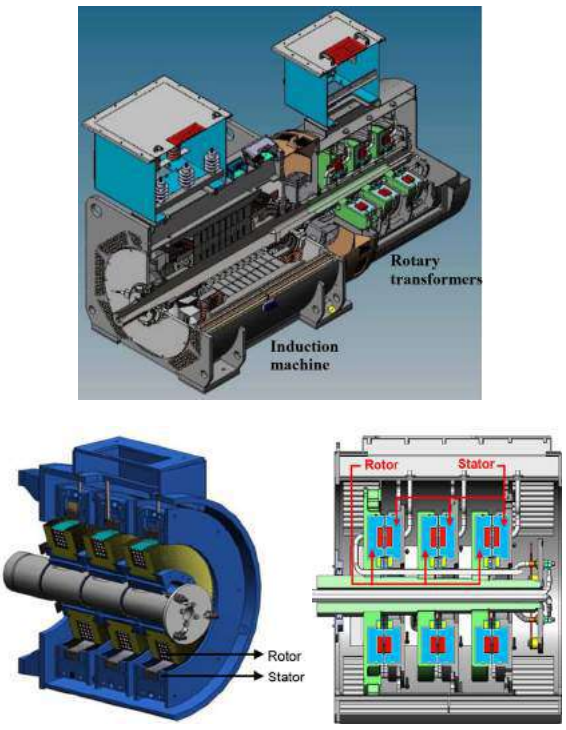

Figure 1. Sectional view of the BDFIG-RT structure, with either a rotary transformer of type (a) or type (b).[9-12]

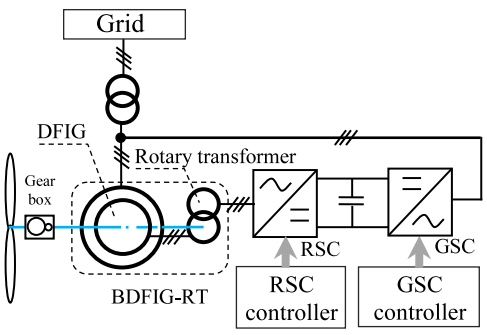

Figure 2. Electrical connection diagram in BDFIG-RT

The main part of BDFIG-RT is composed from DFIG, so it has the same power characteristics as DFIG. So BDFIG-RT can operate in sub- and super-synchronous modes, and there is also a synchronous mode in the short-term. Specifically, during operation, BDFIG-RT usually operates in two modes:

- When the rotor rotational speed is less than the synchronous speed of the stator magnetic field, the generator works in sub-synchronous mode. In this case, the rotation of the rotor magnetic field generated by the rotor currents is the same direction as the rotation of

the mechanical torque. So in this mode, rotor receives energy from the grid (Figure 3.a). With that, the GSC works as a rectifier and the RSC works as an inverter to generate rotor currents with an appropriate frequency. This frequency is suitable for the speed that needs to be supplemented with the rotational speed of magnetic poles of the rotor to equal the synchronous speed. In this mode, the generator rotor receives energy from the grid, only the generator stator generates power (Figure 4.a).

- Conversely, when the rotor rotates at a speed more than the synchronous speed of the stator magnetic field, the generator operates in super-synchronous mode. In this case, the rotation of the rotor magnetic field is opposite to the rotation of mechanical torque. So electricity is generated from the rotor to the grid (Figure 3.b). Then the RSC works as a rectifier and the GSC works as an inverter to generate currents with the same frequency as the grid frequency. In this mode, both stator and rotor sides of the generator supply active power to grid (Figure 4.c).

In the special case, when the rotor rotational speed is equal to the synchronous speed of the stator magnetic field, the generator works in synchronous mode. In which case, the rotor neither generates nor receives active power (Figure 4. b). But this mode only exists temporarily and is an intermediate mode in the above two modes, because the wind speed is always variable.

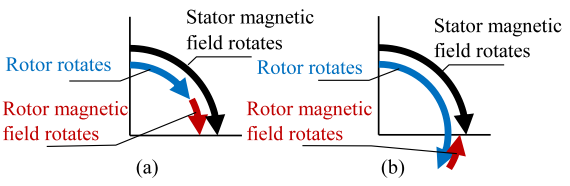

Figure 3. Direction and rotational speed of component magnetic fields and rotor speed in BDFIG-RT.

The active power generated in the stator (P_s) is calculated according to (1). The active power generated or received in the rotor (P_r) is calculated according to (2).

$$P_s = \frac{P_m}{1-s} \quad (1)$$

$$P_r = \frac{s}{1-s} P_m = sP_s \quad (2)$$

Where P_m is mechanical power.

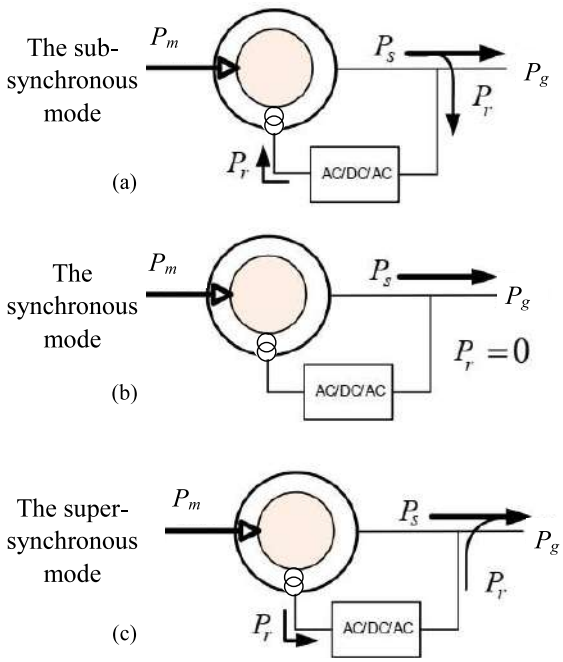

Figure 4. Power flows in BDFIG-RT.

With: $s = 1 - \frac{\omega}{\omega_s}$

Where: s is the slip ratio; ω_s is the speed of rotating magnetic field in the stator; ω is the rotational speed of the rotor.

2.2. Modeling BDFIG-RT

With a DFIG, the current and voltage can be measured in both stator and rotor sides. But the similar parameters can be only measured in the stator of the BDFIG-RT. The value of current and voltage in the rotor of this generator must be measured indirectly through the stator windings of the rotating transformer (Figure 5).

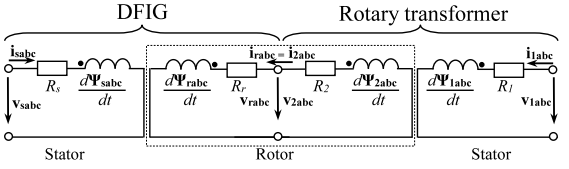

Figure 5. Single-phase equivalent circuit diagram of BDFIG-RT.

https://doi.org/10.52111/qnjs.2022.16104

The equations describe the relationship between voltage and other parameters in a DFIG:

$$\mathbf{v}_{\mathbf{sabc}} = R_s \mathbf{i}_{\mathbf{sabc}} + \frac{d\boldsymbol{\psi}_{\mathbf{sabc}}}{dt} \quad (3)$$

$$\mathbf{v}_{\mathbf{rabc}} = R_r \mathbf{i}_{\mathbf{rabc}} + \frac{d\boldsymbol{\psi}_{\mathbf{rabc}}}{dt} \quad (4)$$

$$\mathbf{v}_{\mathbf{1abc}} = R_1 \mathbf{i}_{\mathbf{1abc}} + \frac{d\boldsymbol{\psi}_{\mathbf{1abc}}}{dt} \quad (5)$$

$$\mathbf{v}_{\mathbf{2abc}} = -R_2 \mathbf{i}_{\mathbf{2abc}} + \frac{d\boldsymbol{\psi}_{\mathbf{2abc}}}{dt} \quad (6)$$

Where, the parameters of generator part: R_s and R_r are the stator and rotor resistance; $\mathbf{v}_{\mathbf{sabc}}$ and $\mathbf{v}_{\mathbf{rabc}}$ are the three-phase stator and rotor voltage matrix; $\mathbf{i}_{\mathbf{sabc}}$ and $\mathbf{i}_{\mathbf{rabc}}$ are the three-phase stator and rotor current matrix; $\boldsymbol{\psi}_{\mathbf{sabc}}$ and $\boldsymbol{\psi}_{\mathbf{rabc}}$ are the three-phase stator and rotor flux matrix. The parameters of rotary transformer part: R_1 and R_2 are the stator and rotor resistance; $\mathbf{v}_{\mathbf{1abc}}$ and $\mathbf{v}_{\mathbf{2abc}}$ are the three-phase stator and rotor voltage matrix; $\mathbf{i}_{\mathbf{1abc}}$ and $\mathbf{i}_{\mathbf{2abc}}$ are the three-phase stator and rotor current matrix; $\boldsymbol{\psi}_{\mathbf{1abc}}$ and $\boldsymbol{\psi}_{\mathbf{2abc}}$ are the three-phase stator and rotor flux matrix.

The rotary transformer rotor and the generator rotor are connected in the same phase with each other. Therefore, the current and voltage of the same phase of them are equal (as shown in Figure 5).

$$\mathbf{v}_{\mathbf{2abc}} = \mathbf{v}_{\mathbf{rabc}} \quad (7)$$

$$\mathbf{i}_{\mathbf{2abc}} = \mathbf{i}_{\mathbf{rabc}} \quad (8)$$

The flux in the stator of generator part is created by the inductance in the phases of the stator itself, plus the mutual inductance of the phases in the rotor to the stator. The same goes for the rotor of generator part. The phenomenon is similar for rotary transformer.

$$\boldsymbol{\psi}_{\mathbf{sabc}} = \boldsymbol{\psi}_{\mathbf{sabc(s)}} + \boldsymbol{\psi}_{\mathbf{sabc(r)}} \quad (9)$$

$$\boldsymbol{\psi}_{\mathbf{rabc}} = \boldsymbol{\psi}_{\mathbf{rabc(s)}} + \boldsymbol{\psi}_{\mathbf{rabc(r)}} \quad (10)$$

$$\boldsymbol{\psi}_{\mathbf{1abc}} = \boldsymbol{\psi}_{\mathbf{1abc(1)}} + \boldsymbol{\psi}_{\mathbf{1abc(2)}} \quad (11)$$

$$\boldsymbol{\psi}_{\mathbf{2abc}} = \boldsymbol{\psi}_{\mathbf{2abc(1)}} + \boldsymbol{\psi}_{\mathbf{2abc(2)}} \quad (12)$$

Where, $\boldsymbol{\psi}_{\mathbf{sabc(s)}}$, $\boldsymbol{\psi}_{\mathbf{rabc(r)}}$, $\boldsymbol{\psi}_{\mathbf{1abc(1)}}$, $\boldsymbol{\psi}_{\mathbf{2abc(2)}}$ are inductive flux components, and mutual inductance flux components include $\boldsymbol{\psi}_{\mathbf{sabc(r)}}$, $\boldsymbol{\psi}_{\mathbf{rabc(s)}}$, $\boldsymbol{\psi}_{\mathbf{1abc(2)}}$, $\boldsymbol{\psi}_{\mathbf{2abc(1)}}$.

In order to analyze the control process, it is necessary to convert the three-phase voltage, current, and magnetic flux parameters in the *abc* reference frame to the *dq* reference frame attached to the rotor. These parameters are converted according to the Park transition. The voltage equations (3)-(6) are rewritten in the *dq* reference frame:

$$\underline{v}_{sdq} = R_s \underline{i}_{sdq} + \frac{d\underline{\psi}_{sdq}}{dt} + j\omega_s \underline{\psi}_{sdq} \quad (13)$$

$$\underline{v}_{rdq} = R_r \underline{i}_{rdq} + \frac{d\underline{\psi}_{rdq}}{dt} + j\omega_{sl} \underline{\psi}_{rdq} \quad (14)$$

$$\underline{v}_{2dq} = -R_2 \underline{i}_{2dq} + \frac{d\underline{\psi}_{2dq}}{dt} + j\omega_{sl} \underline{\psi}_{2dq} \quad (15)$$

$$\underline{v}_{1dq} = R_1 \underline{i}_{1dq} + \frac{d\underline{\psi}_{1dq}}{dt} + j\omega_{sl} \underline{\psi}_{1dq} \quad (16)$$

With $\underline{i}_{rdq} = \underline{i}_{2dq}$, $\underline{v}_{rdq} = \underline{v}_{2dq}$.

Besides, $\omega_{sl} = \omega_s - \omega_r$ is the slip speed, which is also the rotational speed of the rotor magnetic field.

The stator and rotor fluxes of generator and rotary transformer can be written in the *dq* reference frame as follows:

$$\underline{\psi}_{sdq} = L_s \underline{i}_{sdq} + L_m \underline{i}_{rdq} \quad (17)$$

$$\underline{\psi}_{rdq} = L_m \underline{i}_{sdq} + L_r \underline{i}_{rdq} \quad (18)$$

$$\underline{\psi}_{1dq} = L_1 \underline{i}_{1dq} + L_{tm} \underline{i}_{2dq} \quad (19)$$

$$\underline{\psi}_{2dq} = L_{tm} \underline{i}_{1dq} + L_2 \underline{i}_{2dq} \quad (20)$$

Where

$$L_s = L_m + L_{s\sigma} \quad (21)$$

$$L_r = L_m + L_{r\sigma} \quad (22)$$

$$L_1 = L_{tm} + L_{1\sigma} \quad (23)$$

$$L_2 = L_{tm} + L_{2\sigma} \quad (24)$$

Where L_m is mutual inductance, L_s, $L_{s\sigma}$ are stator inductance and stator leakage inductance, L_r, $L_{r\sigma}$ are rotor inductance and rotor leakage inductance of generator part. Similarly, L_{tm} is mutual inductance, L_1, $L_{1\sigma}$ are stator inductance and stator leakage inductance, L_2, $L_{2\sigma}$ are rotor inductance and rotor leakage inductance of rotary transformer part.

The active and reactive power in the stator of the generator P_s, Q_s are calculated according to (25), (26). Likewise, the active and reactive power in the stator of the rotary transformer P_1, Q_1 are calculated by equations (27), (28), respectively.

$$P_s = \frac{3}{2}\mathrm{Re}\left\{\underline{v}_{sdq}\hat{\underline{i}}_{sdq}\right\} = \frac{3}{2}\left(v_{sd}i_{sd} + v_{sq}i_{sq}\right) \quad (25)$$

$$Q_s = \frac{3}{2}\mathrm{Im}\left\{\underline{v}_{sdq}\hat{\underline{i}}_{sdq}\right\} = \frac{3}{2}\left(v_{sq}i_{sd} - v_{sd}i_{sq}\right) \quad (26)$$

$$P_1 = \frac{3}{2}\mathrm{Re}\left\{\underline{v}_{1dq}\hat{\underline{i}}_{1dq}\right\} = \frac{3}{2}\left(v_{1d}i_{1d} + v_{1q}i_{1q}\right) \quad (27)$$

$$Q_1 = \frac{3}{2}\mathrm{Im}\left\{\underline{v}_{1dq}\hat{\underline{i}}_{1dq}\right\} = \frac{3}{2}\left(v_{1q}i_{1d} - v_{1d}i_{1q}\right) \quad (28)$$

Where $\hat{\underline{i}}_{sdq}$ and $\hat{\underline{i}}_{1dq}$ are complex conjugate $\underline{i}_{sdq}$ and $\underline{i}_{1dq}$.

2.3. Model and propose control loops of GSC connected to an inductive filter

In the circuit shown in Figure 6, R_g is the resistance of the filter reactor L_g.

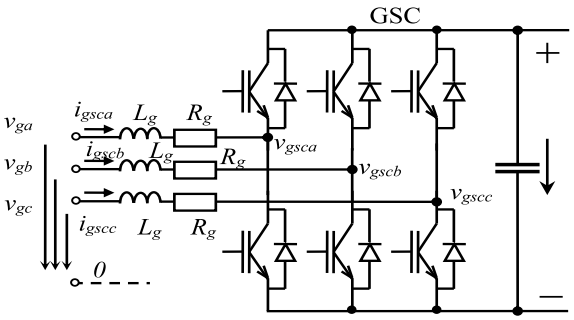

Figure 6. The GSC connected to an three-phase inductive filter.

On the basis of the circuit shown in Figure 6, the voltage matrix on the grid is calculated:

$$\mathbf{v}_{\mathbf{gabc}} = L_g\frac{d\mathbf{i}_{\mathbf{gscabc}}}{dt} + R_g\mathbf{i}_{\mathbf{gscabc}} + \mathbf{v}_{\mathbf{gscabc}} \quad (29)$$

Convert (29) to the dq reference frame with the d axis coincident with the grid voltage vector, and this reference frame rotates at synchronous speed.

$$\underline{v}_{gdq} = L_g\frac{d\underline{i}_{gscdq}}{dt} + R_g\underline{i}_{gscdq} + j\omega_s L\underline{i}_{gscdq} + \underline{v}_{gscdq} \quad (30)$$

The power that is received from the grid into the GSC is determined as follows:

$$P_{gsc} = -\frac{3}{2}\mathrm{Re}\left\{\underline{v}_{gdq}\hat{\underline{i}}_{gscdq}\right\} = -\frac{3}{2}\left(v_{gd}i_{gscd} + v_{gq}i_{gscq}\right) \quad (31)$$

$$Q_{gsc} = -\frac{3}{2}\mathrm{Im}\left\{\underline{v}_{gdq}\hat{\underline{i}}_{gscdq}\right\} = -\frac{3}{2}\left(v_{gq}i_{gscd} - v_{gd}i_{gscq}\right) \quad (32)$$

Where $\hat{\underline{i}}_{gscdq}$ is complex conjugate of $\underline{i}_{gscdq}$.

From equation (29), a control loop for the GSC set in the Laplace is determined as follows:

$$\underline{v}_{gscdq} = \underline{v}_{gdq} - \underline{v}^*_{gscdq} - j\omega_s L\underline{i}_{gscdq} \quad (33)$$

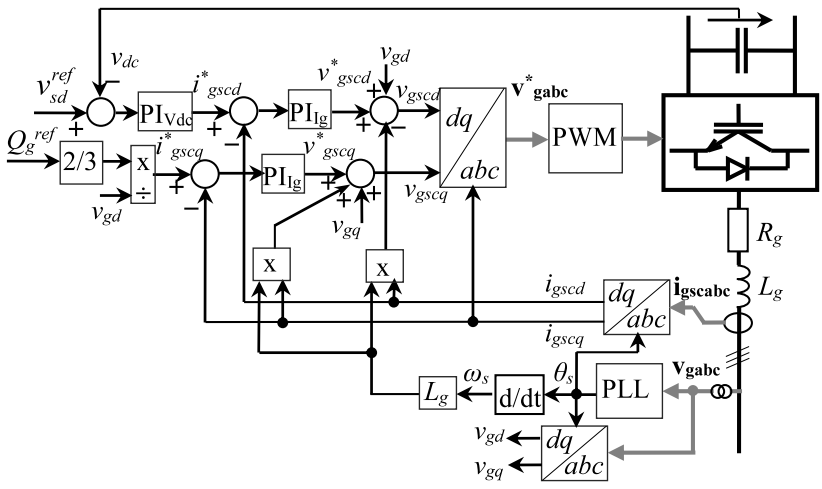

Figure 7. Control diagram of GSC.

https://doi.org/10.52111/qnjs.2022.16104

With $\underline{v}^*_{gscdq} = (L_g s + R_g)\underline{i}_{gscdq}$

Where $\underline{v}^*_{gscdq}$ is determined from the current control loop $\underline{i}_{gscdq}$ with the PI_{Ig} regulator.

$$\underline{v}^*_{gscdq} = \left(k_{pIg} + \frac{k_{iIg}}{s}\right)(\underline{i}^*_{gscdq} - \underline{i}_{gscdq}) \quad (34)$$

Where $\underline{i}_{gscdq}$ is determined from the DC voltage control loop v_{dc} with the PI_{Vdc} controller.

$$i^*_{gscd} = \left(k_{pVdc} + \frac{k_{iVdc}}{s}\right)(v^{ref}_{dc} - v_{dc}) \quad (35)$$

With v^{ref}_{dc} is the DC voltage setting value in the GSC and RSC.

And i^*_{gscd} is determined through reactive power setting value (Q^{ref}_{gsc}) at the GSC:

$$i^*_{gscq} = \frac{2}{3}\frac{Q^{ref}_{gsc}}{v_{gd}} \quad (36)$$

From equations (33)-(36), this study has established the control diagram for the GSC as shown in Figure 7.

2.4. Propose RSC control loops

From equation (18), $\underline{i}_{rdq}$ can be determined:

$$\underline{i}_{rdq} = \frac{1}{L_m}\underline{\psi}_{sdq} - \frac{L_s}{L_m}\underline{i}_{sdq} \quad (37)$$

Substitute (37) into the equation (19) to get:

$$\underline{\psi}_{1dq} = L_1\underline{i}_{1dq} + \frac{L_{tm}}{L_m}\underline{\psi}_{sdq} - \frac{L_{tm}L_s}{L_m}\underline{i}_{sdq} \quad (38)$$

Then, substitute (38) into the equation (16) to get:

$$\underline{v}_{1dq} = R_1\underline{i}_{1dq} + \frac{d}{dt}\left(L_1\underline{i}_{1dq} + \frac{L_{tm}}{L_m}\underline{\psi}_{sdq} - \frac{L_{tm}L_s}{L_m}\underline{i}_{sdq}\right) + j\omega_{sl}\left(L_1\underline{i}_{1dq} + \frac{L_{tm}}{L_m}\underline{\psi}_{sdq} - \frac{L_{tm}L_s}{L_m}\underline{i}_{sdq}\right) \quad (39)$$

Convert equation (39) into the Laplace:

$$\underline{v}_{1dq} = \underline{v}^*_{I1dq} + \underline{v}^*_{\psi sdq} - \underline{v}^*_{Isdq} + j\omega_{sl}\left(L_1\underline{i}_{1dq} + \frac{L_{tm}}{L_m}\underline{\psi}_{sdq} - \frac{L_{tm}L_s}{L_m}\underline{i}_{sdq}\right) \quad (40)$$

With: $\underline{v}^*_{I1dq} = (R_1 + L_1 s)\underline{i}_{1dq}$, $\underline{v}^*_{Isdq} = \frac{L_{tm}L_s}{L_m}s\underline{i}_{sdq}$, $\underline{v}^*_{\psi sdq} = \frac{L_{tm}}{L_m}s\underline{\psi}_{sdq}$.

Where: $\underline{v}^*_{I1dq}$ is determined from the current control loop $\underline{i}_{1dq}$ with the PI_{I1} regulator; $\underline{v}^*_{Isdq}$ is determined from current control loop $\underline{i}_{sdq}$ with the PI_{Is} regulator; $\underline{v}^*_{\psi sdq}$ is determined from the magnetic flux control loop $\underline{\psi}_{sdq}$ with the regulator $PI_{\psi s}$.

$$\underline{v}^*_{I1dq} = \left(k_{pI1} + \frac{k_{iI1}}{s}\right)(\underline{i}^*_{1dq} - \underline{i}_{1dq}) \quad (41)$$

$$\underline{v}^*_{Isdq} = \left(k_{pIs} + \frac{k_{iIs}}{s}\right)(\underline{i}^*_{sdq} - \underline{i}_{sdq}) \quad (42)$$

$$\underline{v}^*_{\psi sdq} = \left(k_{p\psi s} + \frac{k_{i\psi s}}{s}\right)(\underline{\psi}^*_{sdq} - \underline{\psi}_{sdq}) \quad (43)$$

From equations (25), (26), they are combined with $v^{ref}_{sq} = 0$, $v^{ref}_{sd} = v^{ref}_s$, (because the *dq* reference frame is attached to the rotor, with the d-axis is the same direction as the stator voltage vector, and this reference frame rotates at synchronous speed ω_s), leading to i^*_{sd} and i^*_{sq} and are determined:

$$i^*_{sd} = \frac{2}{3}\frac{P_s}{v^{ref}_{sd}} \quad (44)$$

$$i^*_{sq} = -\frac{2}{3}\frac{Q_s}{v^{ref}_{sd}} \quad (45)$$

Transient fluctuations of stator flux in the *dq* reference frame can be ignored when determining a reference value, so the equation (13) is rewritten:

$$\underline{v}^{ref}_{sdq} \approx R_s\underline{i}^*_{sdq} + j\omega_s\underline{\psi}^*_{sdq} \quad (46)$$

From (46), ψ^*_{sd}, ψ^*_{sq} are determined:

$$\psi^*_{sd} \approx -\frac{R_s i^*_{sq}}{\omega_s} \quad (47)$$

$$\psi^*_{sq} \approx \frac{R_s i^*_{sd}}{\omega_s} - \frac{v^{ref}_{sd}}{\omega_s} \quad (48)$$

To determine according to the torque control method (corresponding to the active power), it can be based on the electromagnetic torque of the DFIG:

$$T_e = \frac{3}{2}(\psi_{sd}i_{sq} - \psi_{sq}i_{sd}) \approx -\frac{3}{2}\psi_{sq}i_{sd} \quad (49)$$

With $\psi_{sd} = L_s i_{sd} + L_m i_{rd} \approx 0$

Therefore $i_{sd} \approx -\frac{L_m}{L_s} i_{rd}$

Thanks to that, from (49) can determine i_{rd}:

$$i_{rd} \approx \frac{2}{3} \frac{L_s T_e}{L_m \psi_{sq}} \quad (50)$$

If the power losses on the windings and steel cores of the rotary transformer are neglected, the power transmitted through the windings of the rotating transformer can be considered to be approximately equal.

$$\frac{3}{2}\left(v_{1d} i_{1d} + v_{1q} i_{1q}\right) \approx \frac{3}{2}\left(v_{2d} i_{2d} + v_{2q} i_{2q}\right) \quad (51)$$

Thus, $v_{1d} i_{1d} \approx v_{2d} i_{2d}$, $v_{1q} i_{1q} \approx v_{2q} i_{2q}$.

Thereby, i_{1d}^* can be determined based on combination with the equation (50) and $\underline{i}_{2dq} = \underline{i}_{rdq}$:

$$i_{1d}^* \approx \frac{i_{rd}}{k_t} \approx \frac{2}{3} \frac{L_s T_e^{ref}}{k_t L_m \psi_{sq}} \quad (52)$$

Where $k_t = v_1/v_2$ is the voltage variation ratio of rotary transformer; the torque T_e^{ref} is determined by looking up the wind turbine characteristics corresponding to the rotor rotation speed.

From equation (17), i_{sq} is determined:

$$i_{sq} = \frac{1}{L_s} \psi_{sq} - \frac{L_m}{L_s} i_{rq} \quad (53)$$

Equation (48) and equation (53) are combined together, and if the stator resistance is ignored, then i_{sq} is determined as follows:

$$i_{sq} \approx -\frac{v_{sd}}{\omega_s L_s} - \frac{L_m}{L_s} i_{rq} \quad (54)$$

From (26) is combined with (54), $v_{sq}^{ref} = 0$, $v_{sd}^{ref} = v_s^{ref}$, it is rewritten as below:

$$Q_s \approx -\frac{3}{2} v_{sd}^{ref} i_{sq} \approx \frac{3}{2} \frac{(v_{sd}^{ref})^2}{\omega_s L_s} + \frac{3}{2} \frac{L_m}{L_s} v_{sd}^{ref} i_{rq} \quad (55)$$

Thanks to (55) combined with $\underline{i}_{2dq} = \underline{i}_{rdq}$, i_{1q}^* can be determined as follows:

$$i_{1q}^* \approx \frac{i_{2q}}{k_t} \approx \frac{2}{3} \frac{L_s Q_s}{k_t L_m v_{sd}^{ref}} - \frac{v_{sd}^{ref}}{k_t \omega_s L_m} \quad (56)$$

On the basis of equation (40) combined with other mathematical expressions presented above, the RSC control diagram is proposed as shown in Figure 8.

3. SIMULATION

3.1. The simulated system

In order to check the operation of the control system proposed above, an application of BDFIG-RT in the power grid with the structure shown in Figure 9 is simulated in this study. Due to the generator is an asynchronous generator, it is usually consumed a large amount of reactive power. Therefore, the generator is usually connected with a three-phase compensator capacitor to directly supply a basic amount of reactive power to it (Figure 9.a). In addition, for the purpose of converting the characteristics of generator from a DFIG to a wound-rotor asynchronous generator connected the auxiliary resistor, an auxiliary resistor (R_crowbar) is connected to the stator of the rotating transformer (Figure 9.a).[15] As a result, the generator can be still connected to the grid in a certain time when there is a short circuit fault on the grid.[15]

https://doi.org/10.52111/qnjs.2022.16104

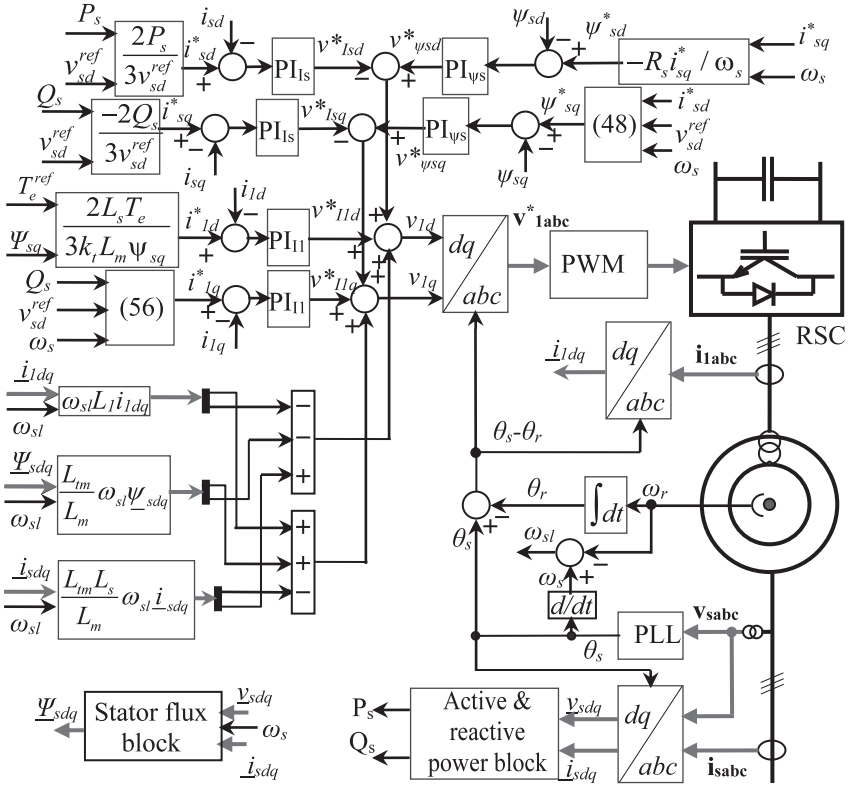

Figure 8. Control diagram of RSC

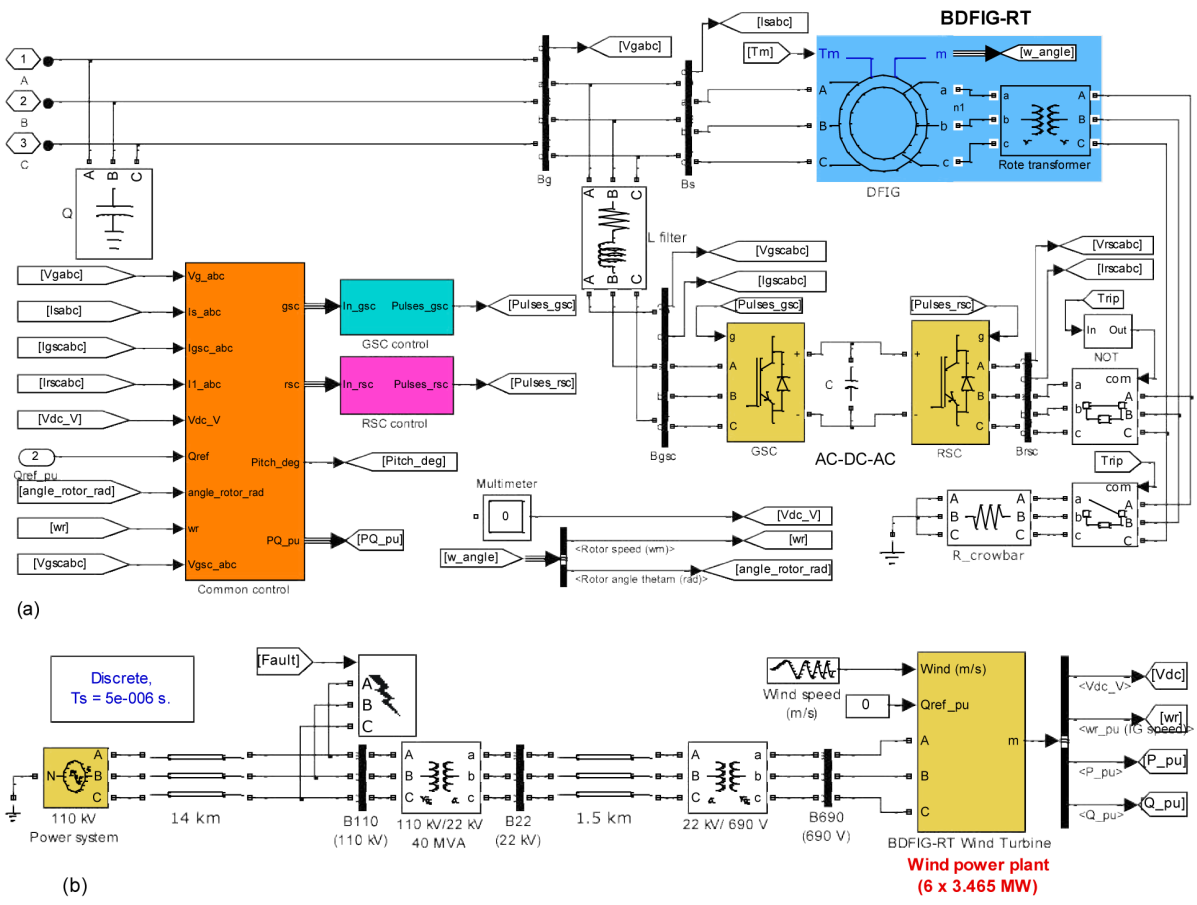

Figure 9. The structure of BDFIG-RT (a), and it is connected to the grid (b) in Matlab Simulink

The output voltage of each generator in the wind power plant is 690 V. This output is connected to a transformer to boost voltage from 690 V to 22 kV at each wind turbine. Next, the output of all transformers are connected together at a busbar in the 22kV/110kV power station (Figure 9.b). At this power station, the wind power plant is connected to the 110 kV power grid of the area.

3.2. Parameters of simulated system

The parameters of components of the wind power system using BDFIG-RT are presented in Table 1 – Table 3. In which, the wind turbine part (with parameters as Table 1) provides mechanical energy to the generator (with parameters as Table 2). The BDFIG-RT has the stator windings of the rotary transformer connected to the back-to-back power converter (with parameters as Table 3). This converter is connected to the grid via a three-phase filter inductor (with parameters as Table 3).

Table 1. Wind turbine parameters

Parameter	Value	Unit
Rated wind speed	12	m/s
Wind turbine inertia constant	3.32	s
Shaft mutual damping	0.75	pu
Shaft spring constant refer to high-speed shaft	1.11	pu

Table 2. BDFIG-RT parameters in the simulation

Parameter	Value	Unit
Generator data		
Rated power	3.465	MW
Rated stator voltage	690	V
Rated rotor voltage	1975	V
R_s (stator resistance)	0.023	pu
$L_{s\sigma}$ (stator leakage inductance)	0.18	pu
R_r (rotor resistance)	0.016	pu
$L_{r\sigma}$ (rotor leakage inductance)	0.16	pu
L_m (mutual inductance)	2.9	pu
Inertia constant	0.685	s
Pairs of poles	3	
Rotary transformer data		
Rated stator voltage	690	V
Rated rotor voltage	1975	V
R_1 (stator resistance)	0.1005	pu
$L_{1\sigma}$ (stator leakage inductance)	0.0022	pu
R_2 (rotor resistance)	0.0899	pu
$L_{2\sigma}$ (rotor leakage inductance)	0.0023	pu
L_{tm} (mutual inductance)	0.0545	pu

Table 3. Parameters of back-to-back power converter and filter inductor

Parameter	Value	Unit
v_{dc} (DC voltage)	1150	V
L_g (inductance of filter inductor)	0.3	pu
R_g (resistance of filter inductor)	0.003	pu

4. RESULTS AND DISCUSSION

Some different cases are simulated in this paper in order to check the suitability of the control system proposed above. First, it is the case of variable wind speed, the dynamic response of this system is examined. Second, this system is surveyed when there is a short circuit fault on the grid.

4.1. Case of variable wind speed

In this section, the operation of the system is simulated and considered when wind speed varies (as shown in Figure 10). The results of simulation shown that the generator speed varies within the appropriate range. This range is from less to over than 30% of the synchronous speed (Fig. 11.a). As a result, the variation of active power that is generated or received in the generator rotor via the rotary transformer is suitable for the control law presented (Figure 11.b). The active power that is generated in stator of the generator also varies dependently wind speed (Figure 11.b).

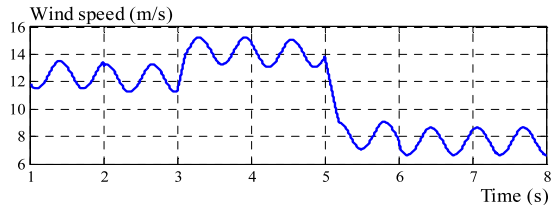

Figure 10. Wind speed chart.

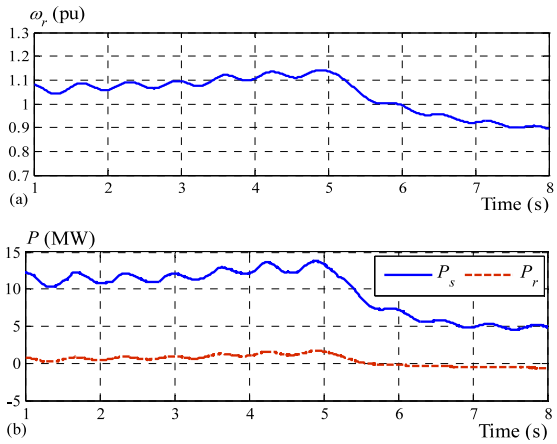

Figure 11. Chart of speed (a) and active power (b) of BDFIG-RT.

https://doi.org/10.52111/qnjs.2022.16104

4.2. Case of short circuit fault on the grid

The 1st case study: there was a three-phase short circuit fault on the 110 kV transmission grid and near the wind power plant. This fault has happened at 1.3 s, and it has been cleared at 1.45 s. The fault duration time was 150 ms, and the voltage has dropped to about 17.33% of the rated voltage during the this fault (as shown in Figure 12). The simulation results are shown in Figure 13 and Figure 14. According to the results in Figure 13, the control system has taken less than a current cycle to regulate the short-circuit current value to the permissible value (Figure 13.d). This control method has been based on the principle of converting the generator from a DFIG to a wound-rotor asynchronous generator connected the auxiliary resistors. According to this method, the active power that has been generated in the generator stator has been reduced to a minimum throughout the time of short circuit fault (as shown in Figure 13.a). Thereby, the overall power flow transfer to the grid has been reduced during the short circuit as shown in Figure 14.a. In the time of short circuit fault, the RSC has been controlled to neither transmit nor receive active power, this has been shown by the currents shown in Figure 13.b. For the purpose of stabilizing the voltage, the GSC has supplied reactive power into the grid (Figure 13.c). In addition, in order to recover the voltage after the fault, a large amount of reactive power has been supplied into the grid immediately after the short circuit fault has been cleared (Figure 14.b). Thanks to good control, the generator speed has been only slightly increased (as shown in Figure 14.c).

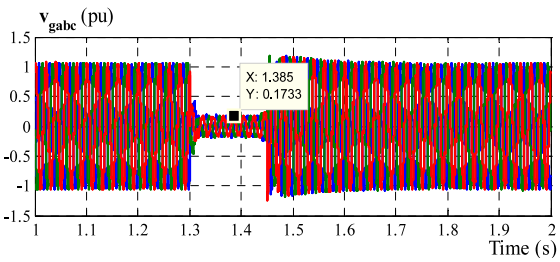

Figure 12. The grid voltage with 1st case study.

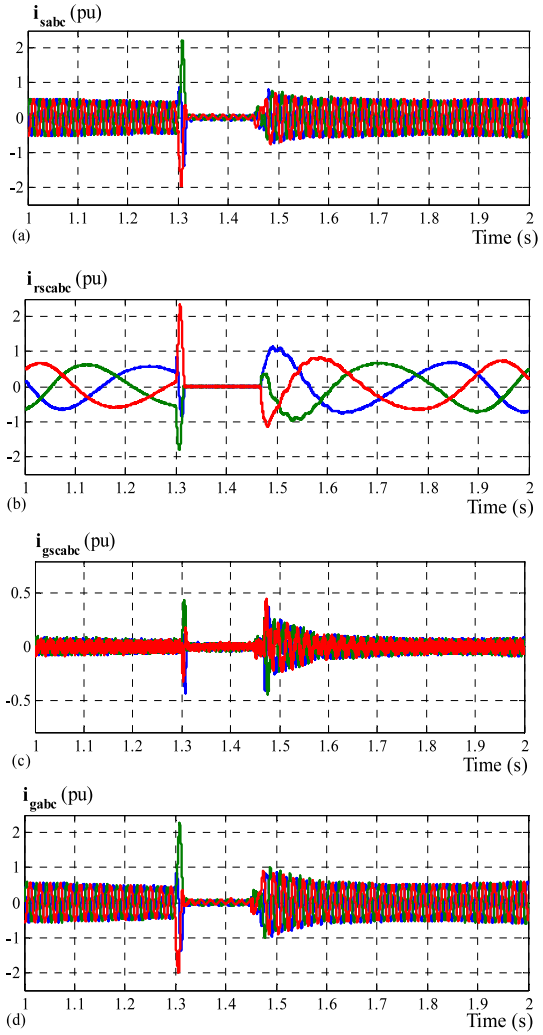

Figure 13. Currents in *abc* reference frame with 1st case study.

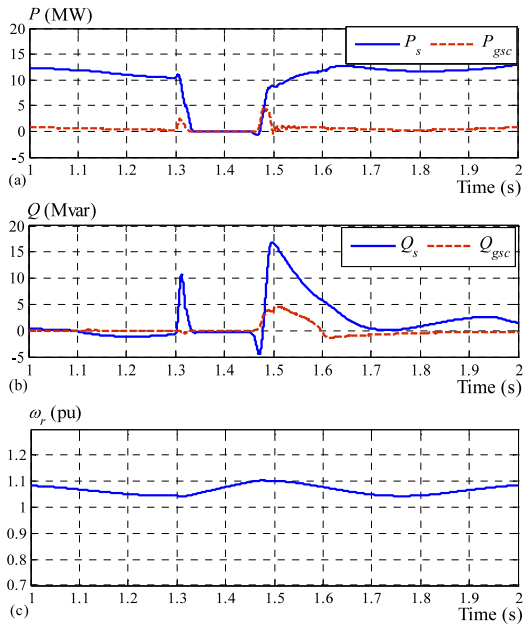

Figure 14. Power and speed of BDFIG-RT with 1st case study.

The 2nd case study: there was a single phase-to-ground fault (C phase to ground) on the 110 kV transmission grid and near the wind power plant. This fault has happened at 1.3 s and has been cleared at 1.45 s. During the fault, the phase C voltage has dropped to about 19.55% of the rated value (Figure 15). At the same time, an overvoltage has occurred in the remaining 2 phases to about 1.33 pu (Figure 15). This value is smaller than the limited ground fault of 1.4 in the Circular No. 25/2016/TT-BCT of the Ministry of Industry and Trade. The simulation results of this case are shown in Figure 16 and Figure 17. According to the results in Figure 16, the control system has regulated the short-circuit currents within the normal operating range (Figure 16.d). In this case, the control method has not changed, even though the short circuit has been asymmetrical. In the time of short circuit fault, according to this method, the currents in the generator stator have been reduced (Figure 16.a). Therefore, the active power that has been generated by the generator has been reduced during the short circuit fault (Figure 17.a). In order to support voltage stability, the GSC (with a small contribution) together with the generator stator have consumed the reactive power of the grid (Figure 16.c and Figure 17.b). After that, the voltage of the grid has been reduced as soon as the fault has been cleared (Figure 15). Then, a large amount reactive power has been supplied by generator into the grid to support voltage stabilization at this time (Figure 17.b). During the short circuit fault, the RSC has been disabled (Figure 16.b). Thanks to the appropriate control, the generator speed has been increased only slightly (as shown in Figure 17.c).

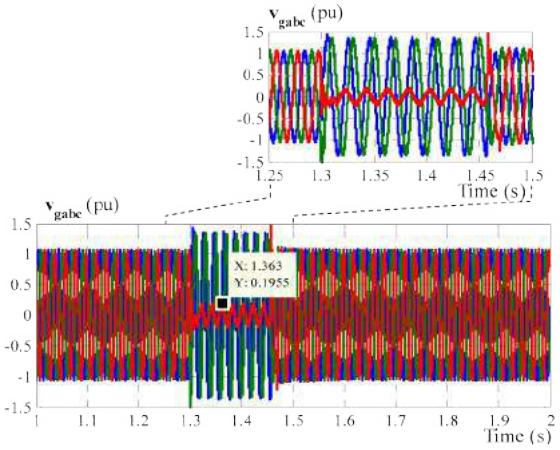

Figure 15. The grid voltage with 2nd case study.

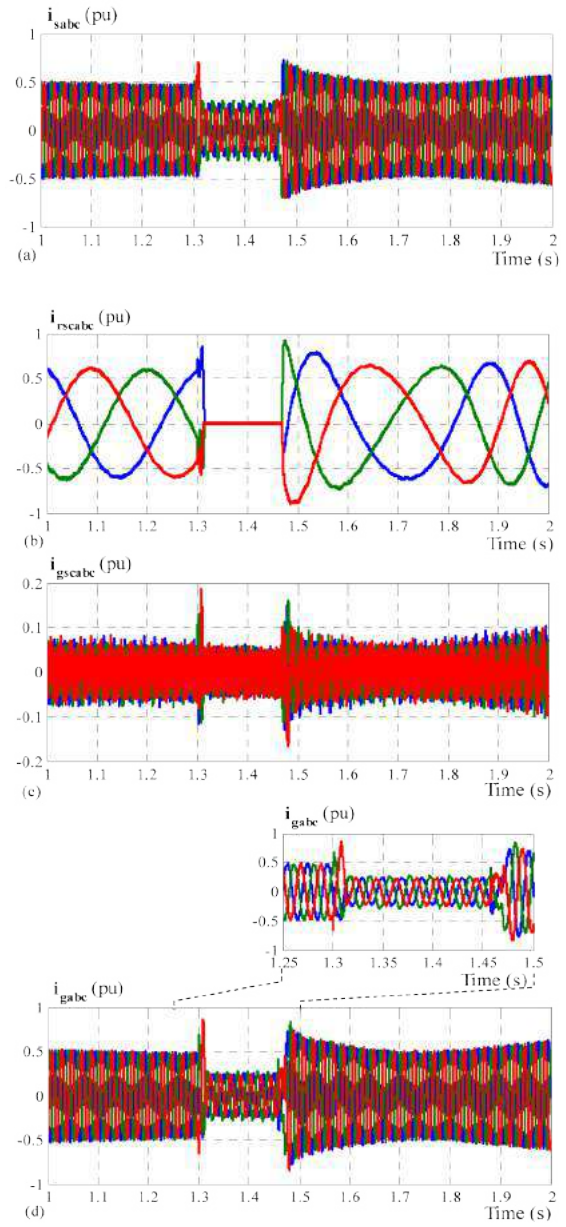

Figure 16. Currents in *abc* reference frame with 2nd case study.

https://doi.org/10.52111/qnjs.2022.16104

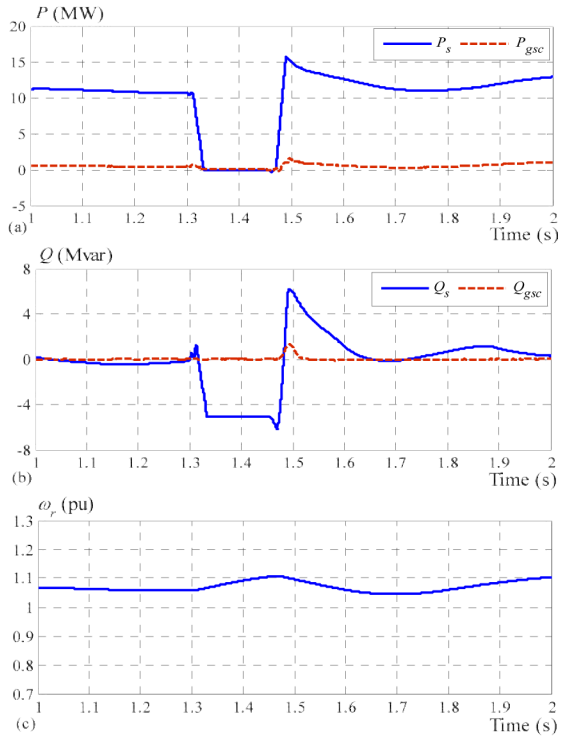

Figure 17. Power and speed of BDFIG-RT with 2nd case study.

Thus, through checking the dynamic process in the above cases, the control system proposed above has operated stably. The response time of the system was very short. Current and voltage parameters have conformed to the required values.

4. CONCLUSION

BDFIG-RTs solve the restrictions of DFIGs, which must be done maintenance of carbon brushes and slip rings, in wind power. This paper analyzed and proposed control loops for the power converters of a BDFIG-RT. This paper analyzed and proposed control loops for the power converters of BDFIG-RT. Through simulation in MatLab software, this research tested the proposed control loops. The simulation results of the system with a wind power station using BDFIG-RT showed that the operation of this system was stable with varying wind speed or a short circuit fault on the grid.

The results of this study serve as a premise for experimental studies and practical applications of BDFIG-RT in wind power.

Acknowledgment

This study is conducted within the framework of science and technology projects at institutional level of Quy Nhon University under project code T2020.670.18.

REFERENCES

1. M. Cheng, P. Han, G. Buja, M. G. Jovanović. Emerging multiport electrical machines and systems: Past developments, current challenges, and future prospects, *IEEE Transaction on Industrial Electronics*, **2018**, *65*(7), 5422 – 5435.
2. P. Han, M. Cheng, Y. Jiang and Z. Chen. Torque/power density optimization of a dual-stator brushless doubly-fed induction generator for wind power application, *IEEE Transaction on Industrial Electronics*, **2017**, *64*(12), 9864-9875.
3. H. Liu, Y. Zhang, F. Zhang, S. Jin, H. Zhang, H. Nian. Design and performance analysis of dual-stator brushless doubly-fed machine with cage-barrier rôtor, *IEEE Transactions on Energy Conversion*, **2019**, *34*(3), 1347 – 1357.
4. P. Han, M. Cheng, R. Luo. Design and analysis of a brushless doubly-fed induction machine with dual-stator structure, *IEEE Transactions on Energy Conversion*, **2016**, *31*(3), 1132 – 1141.
5. Min-Fu Hsieh, Yao-Hsin Chang, David G. Dorrell. Design and analysis of brushless doubly-fed reluctance machine for renewable energy applications, *IEEE Transactions on Magnetics*, **2016**, *52*(7).
6. S. Shao, E. Abdi, F. Barati, R. McMahon. Stator-flux-oriented vector control for brushless doubly fed induction generator, *IEEE Transaction on Industrial Electronics*, **2009**, *56*(10), 4220-4228.
7. R. Zhao, A. Zhang, Y. Ma, X. Wang, J. Yan, Z. Ma. The dynamic control of reactive power for the brushless doubly fed induction machine with indirect stator-quantities control scheme, *IEEE Transactions on Power Electronics,* **2015**, *30*(9), 5046 - 5057.
8. T. Taluo, L. Ristić, M. Jovanović. Dynamic modeling and control of BDFRG under

unbalanced grid conditions, *Energies,* **2021**, *14*(14), 1-25.

9. M. Ruviaro, F. Rüncos, N. Sadowski, I. M. Borges. Analysis and test results of a brushless doubly fed induction machine with rotary transformer*, IEEE Transaction on Industrial Electronics,* **2012**, *59*(6), 2670-2677.

10. M. Ruviaro, F. Rüncos. A brushless doubly fed induction machine with flat plane rotary transformers, *2012 XXth International Conference on Electrical Machines, IEEE*, 23-29, 2012.

11. M. Ruviaro, F. Rüncos. Wound rôtor doubly fed induction machine with radial rotary transformer, *Journal of Microwaves, Optoelectronics and Electromagnetic Applications*, **2013**, *12*(2), 411-426.

12. Hui Zhong, Chao Wu, Yingjie Wang. Design study on novel three-phase rotary transformer used for brushless doubly fed induction generators, *2017 20th International Conference on Electrical Machines and Systems (ICEMS), IEEE*, 1-5, 2017.

13. Naveed-ur-Rehman Malik, Chandur Sadarangani. Dynamic modeling and control of a brushless doubly-fed induction generator with a rotating power electronic converter, *2012 XXth International Conference on Electrical Machines, IEEE*, 900-906, 2012.

14. J. Hu, J. Zhu, D. G. Dorrell. A new control method of cascaded brushless doubly fed induction generators using direct power control, *IEEE Transactions on Energy Conversion*, **2014**, *29*(3), 771 - 779.

15. Tran Duong Hoang Phuc, Le Thai Hiep. Study on operating modes of doubly fed induction generator with a short circuit fault on grid near the wind power plant, *Journal of science – Quy Nhon university*, **2021**, *15*(1), 37-44.

https://doi.org/10.52111/qnjs.2022.16104

TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC QUY NHƠN

Khảo sát chu kỳ sinh sản và theo dõi một số chỉ số thành thục của sò mía (*Tapes dorsatus*) ở vùng đầm Thị Nại, tỉnh Bình Định

Đặng Thị Ngọc Hà[1,*], Đặng Thị Ngọc Dung[2]

[1]Khoa Khoa học tự nhiên, Trường Đại học Quy Nhơn, Việt Nam
[2]Trường Đại học Sư phạm Kỹ thuật Thành phố Hồ Chí Minh, Việt Nam

Ngày nhận bài: 08/11/2021; Ngày nhận đăng: 02/12/2021

TÓM TẮT

Sò mía *(Tapes dorsatus)* được thu từ tự nhiên ở khu vực ven bờ đầm Thị Nại, tỉnh Bình Định nhằm khảo sát chu kỳ sinh sản và bố trí thí nghiệm nuôi vỗ thành thục. Kết quả khảo sát chu kỳ sinh sản tự nhiên cho thấy sò mía đạt đỉnh cao về thể trạng và thành thục sinh dục vào tháng 10 đến tháng 2 trong năm. Thực nghiệm theo dõi một số chỉ tiêu thành thục sinh dục sò mía được thực hiện trong 7 đợt, với các loại hình thức nuôi khác nhau: nuôi vỗ trong các rổ nhựa không có cát đáy, nuôi trực tiếp trên nền đáy cát ở bãi triều tự nhiên, có lưới bao quanh; nuôi vỗ trong các rổ nhựa có đáy cát dày 10 - 15 cm. Kết quả thực nghiệm cho thấy, sò được nuôi vỗ trong các rổ nhựa có đáy cát dày 10 - 15 cm cho sinh trưởng khối lượng, chiều dài và tỷ lệ sống cao nhất (các giá trị này đạt lần lượt 51,6g, 49,2 mm và 96,5% sau 60 ngày nuôi). Tương tự như vậy, các chỉ số thành thục CI và GI cũng đạt giá trị cao nhất khi sò được nuôi ở nghiệm thức này. Sức sinh sản của sò đạt cao nhất, với 3.537±1.102 trứng/g thịt tươi và 28.339±10.556 trứng/khối lượng tổng khi sò được nuôi trong các rổ nhựa có đáy cát 10 - 15 cm.

Từ khóa: *Tapes dorsatus, sò, sinh sản, nuôi vỗ thành thục, tỷ lệ sống.*

**Tác giả liên hệ chính.*
Email: dangthingocha@qnu.edu.vn

JOURNAL OF SCIENCE
QUY NHON UNIVERSITY

Investigating the reproductive cycle and some maturation indexes of clams (*Tapes dorsatus*) collected from Thi Nai lagoon, Binh Dinh province

Dang Thi Ngoc Ha[1,*], Dang Thi Ngoc Dung[2]

[1]*Faculty of Natural Sciences, Quy Nhon University, Vietnam*
[2]*Ho Chi Minh City University of Technology and Education, Vietnam*

Received: 08/11/2021; Accepted: 02/12/2021

ABSTRACT

Clams *Tapes dorsatus* were collected from the coastal areas of Thi Nai lagoon to examine some reproductive parameters and conduct the maturation culture experiment. The results of examining the reproductive cycle showed that clams get the highest value of condition index and gonad index from October to February of years. The gonado index GI the trial test of maturing mangrove clam was carried out in 7 periods with different treatments: the clams reared in the plastic baskets without sandy bottom; the clams directly reared on the sandy bottom of sandbank covered by net; the clams reared in the plastic baskets with sandy bottom of 10 - 15 cm. The results of the experiment indicated that clams reared in the plastic baskets with sandy bottom of 10 - 15 cm obtained the highest weight, length and survival rate (these values were 51.6g, 49.2mm and 96,5% respectively after 60 days of rearing). Similarly, CI and GI were also highest for clams cultured in this treatment. The fecundity of the clams was highest, with 3,537±1,102 eggs/g fresh muscle and number of eggs/g total weight of clams 28,339±10,556 when they were reared in the plastic baskets with sandy bottom of 10 - 15 cm.

Keywords: *Tapes dorsatus*, *clam, reproduction, maturation culture, survival rate.*

1. INTRODUCTION

Binh Dinh has a coastline of 134 km, a water surface area of more than 7,600 hectares, 33 large and small islands, 3 coastal lagoons. The coastal area has many typical ecosystems, such as estuaries and rivers lagoons, mangroves, coral reefs, seagrass beds. Considering such favorable natural conditions, it is very suitable to develop aquaculture and seafood.[1,2] However, for many years, outside of the planned shrimp farming area, other aquatic and seafood products have not been paid enough attention, especially with molluscs in mangrove areas, coastal areas such as oysters, snails, abalone, clam. Especially, mangrove clam is a new farming species being collected and sold. At present, some households are doing experiments for commercial farming (with seeds imported from Nha Trang City, Khanh Hoa Province) which is bringing very high economic benefits.

Clam *(Tapes dorsatus)* is also known as two-headed clam, two-pronged clams, sweet clams, sweet silk scallops.[3] The clam is a bivalve mollusk. This filter-eating species is widely distributed, but there are very few studies on mangrove clam in the world as well as in Vietnam. The scientific studies on this species are too small, and do not meet the development of the

**Corresponding author.*
Email: dangthingocha@qnu.edu.vn

https://doi.org/10.52111/qnjs.2022.16105

mangrove clam rearing industry. Moreover, due to the lack of strict management of resources in the localities, fishermen have over-exploited this resource, making the exploitation yield decreases and the seed source is gradually scarce.[2]

Currently, at some tidal areas in the coastal area of Binh Dinh province, some households have recruited seed produced in Nha Trang for rearing testing. Initially, some rearing households said that the seed breed rearing in the trial condition had a fast growth rate and high survival rate. The product is sold with a high price ranging from 90 to 150 thousand VND per kilogram. Moreover, the market is more stable in comparison with some local aquaculture species.

Facing to those challenges, however, there are not many studies on clam's biological characteristics and reproduction. The mangrove clam seed used as a brood clam to produce breeding stock is now mainly applied similar to the production process of oysters, clams, and clams. Therefore, in this study, we conducted a survey on the reproductive cycle within 24 months, and then we tested the maturation of mangrove clam collected from experimental location in the coastal tidal area of Binh Dinh province within 2 months. The initial results of successful examining the reproductive cycle and some maturation indexes of clams will contribute to the expansion of the commercial mangrove clam rearing model to many mangrove clam rearing households in Binh Dinh province with the different conditions of type sandy bottom. Furthermore, the goal of this study may contribute to improving the efficiency of using water surface area, diversifying cultured objects, transferring scientific and technical advances in rearing process, timely responding to current production requirements.[3-8]

2. OBJECT AND METHODS

2.1. Object and study sites

The object of this study was clam individuals *Tapes dorsatus* collected from the wild along the coastal areas of Thi Nai lagoon to examine some reproductive parameters and conduct the maturation culture experiment.

The sites to do the experiment were the Thi Nai's tidal zone (Cua Bien: *Seagate in Hoi Son village, Nhon Hoi commune, Quy Nhon city*) and mangrove intertidal zone (Diem Van village, *Phuoc Thuan commune, Tuy Phuoc district, Binh Dinh province: 1528928.000; 605218.000 (Vietnam coordinates, 2000)*) Binh Dinh province in Thi Nai lagoon. These areas are from a half to three kilometers far from the seagate, that have high daily tidal fluctuation of 0.8 - 1.5 m. The salinity in this area is in range of 25‰.

Figure 1. The wild clams collecting sites and the experiment conducting sites in Thi Nai lagoon (Diem Van village, *Phuoc Thuan commune, Tuy Phuoc district, Binh Dinh province:1528928.000; 605218.000 (Vietnam coordinates, 2000)* and Cua Bien: *Seagate in Hoi Son village, Nhon Hoi commune, Quy Nhon c*ity).

2.2. Research methods

2.2.1. Sample collection

A total of 700 clam individuals of *Tapes dorsatus* (100 individuals for every 4 months) were randomly collected from the coastal areas of Thi Nai lagoon to examine some reproductive parameters as gender, condition index (CI) and gonad index (GI). In addition, a total 2100 healthy individuals were selected to conduct the maturation culture experiment.

2.2.2. The experiment of maturation culture

The experiment included 3 treatments, with 100 individuals for each treatment. The experiment was carried out in 7 periods as 7 replicates, with 2 months/period. There were 3 plastic baskets with size of 60 x 80cm used to rear clams for each period in Treatment 1 and Treatment 3. Therefore, there were a total of 2100 clam individuals used for this experiment. The details of the experimental treatments were described as the followings:

- Treatment 1 (NT1): The clams were reared in the plastic baskets without sandy bottom, that were suspended in the wild water;

- Treatment 2 (NT2): The clams were directly reared on the sandy bottom of sandbank covered by net;

- Treatment 3 (NT3): The clams were reared in the plastic baskets with sandy bottom of 10-15 cm, that were suspended in the wild water

The clams were reared in natural environment so they only used natural food.

NT1: The clams reared in the plastic baskets without sandy bottom

NT2: The clams was irectly on sandy bed

NT3: The clams reared in the plastic baskets with sandy bottom

Figure 2. The illustration of the experimental treatments

2.2.3. Determination of the biological and environmental parameters

The environmental parameters as salinity, temperature, pH were measured twice a day (at 7:30 and 14:00) by handness environment meter Horiba U-52, while NO_2^- and kH (mg $CaCO_3$) were tested every 7 days by Sera test kit – Germany.[1,6,7]

The biological parameters of length, width, total weight, muscle weight, gonad index (GI), condition index (CI) and survival rate of clam were determined at the beginning and end of the experiment.

- The biological parameters of length and width were determined by length ruler, precision 1mm.

- The biological parameters of total weight and muscle weight were determined by digital scale, 0.01 accuracy.

- The gonad indiex (GI) is an indicator to assess the maturity of the gonads in clams. The GI genital index was determined as per the method of Walker RL, Heffernar PB.[8] The GI varies from 0 to 4 (with 0: Resting phase; 1: Early development phase; 2: Complete development phase; 3: Maturity phase; 4: Spawn phase).[3,9-12]

- Condition index (CI) is determined as:

$$CI\ (mg/g) = (W_m/W_t)*1000$$

Where: W_m: Weight of fresh muscle of clam (g); W_t: Total weight of clam (g).

- Survival rate (SR) of clam was determined as:

https://doi.org/10.52111/qnjs.2022.16105

$$SR = (Ne/Ns)*100$$

Where: Ns - number of clam individual at the beginning of the experiment; Ne - number of clam individual at the end of the experiment.

At the end of the experiment, we selected 100 matured clam individuals to determine their fecundity. The fecundity was presented as number of eggs/g fresh muscle of clams and number of eggs/g total weight of clams.

To examine the reproductive biology of the wild clams, we dissected these individuals to get the gonads, then observed the gonads to determine gender of clams as well as CI and GI.

2.2.4. Data analysis

The data were presented as mean plus/minus standard deviation (MEAN ± SD) based on Microsoft Office Excel 2010. One-way ANOVA was used to test the difference of mean values between treatments using SPSS 17.0. Evaluation of the difference of mean values was conducted after analysis of variance (Post Hoc Test) by Duncan test method. The difference between the values was determined at the significance level of $p < 0.05$.

3. RESULTS AND DISCUSSION

3.1. Experimental results

3.1.1. Environmental parameters in the experiment

The environmental factors in the treatments are shown in Table 1. It can be seen that the NO_2^- content was from 1.0 to 2.0 mg/L of NT1, NT2 and NT3. The alkalinity (mg $CaCO_3$) of water in the treatments ranged from 89.5 to 107.4 mg/L, of NT1, NT2 and NT3. The other parameters such as temperature, pH and salinity did not largely differ between treatments.

In general, the water environmental parameters are within the suitable range for the growth and development of the clams.

Table 1. Variation of environmental factors, hydration during rearing in conditions

Parameter		NT1	NT2	NT3
NO_2^- (mg/L)		1.0 – 2.0	1.0 – 2.0	1.0 – 2.0
kH (mg $CaCO_3$)		89.5 – 107.4	89.5 – 107.4	89.5 – 107.4
Salinity (‰)	*Morning*	25.6±3.4	25.4±3.5	25.4±3.4
	Afternoon	26.3±3.3	25.5±3.6	26.6±3.3
Temperature (°C)	*Morning*	27.5±2.7	27.5±2.6	27.9±2.5
	Afternoon	28.9±2.3	28.8±2.2	28.6±2.5
pH	*Morning*	7.9±0.3	7.9±0.2	7.9±0.3
	Afternoon	7.8±0.2	7.8±0.2	7.8±0.3

3.1.2. Size, weight and survival rate of the clams

The growth parameters and survival rate of the clams in the treatments are presented in Table 2. Generally, growth values such as weight, length and width of the clams increased after 60 days of raising. The survival rate of the clams decreased but still reached the high values, from 86.7 - 96.5%.

Table 2. Size, weight and survival rate of the clam in the different treatments

Parameter		Total weight (g)	Muscle weight (g)	Length (mm)	Width (mm)	Survival rate (%)
NT 1	*Initial*	43.6±2.3	6.7±1.21	42.2±0.3	33.8±0.2	100.0
	Finish	44.6±4.2^{c}	6.8±1.16^{c}	46.8±0.7^{c}	37.8±0.3^{a}	86.7±1.5^{c}
NT 2	*Initial*	43.6±2.3	6.7±1.21	42.2±0.3	33.8±0.2	100.0
	Finish	49.6±5.5^{b}	7.6±1.23^{b}	47.9±0.5^{b}	36.8±0.4^{b}	90.4±1.2^{b}
NT3	*Initial*	43.6±2.3	6.7±1.21	42.2±0.3	33.8±0.2	100.0
	Finish	51.6±4.3^{a}	9.3±1.26^{a}	49.2±1.1^{a}	37.8±0.4^{a}	96.5±1.8^{a}

Remark: For each parameter, the different letters indicate the significant difference ($p < 0.05$)

In comparison, it was found that clams reared in NT3 (reared in plastic baskets with bottom sand of 10 - 15 cm) had the highest values of weight and length while clams raised in NT1 (reared in plastic baskets without bottom sand) obtained the lowest values. Specifically, after 60 days of rearing, clams raised in treatment NT3 had the weight of 51.6 g and the length of 49.2 mm while these values in NT1 were 44.6g and 46.8 mm, respectively. The width of clams in treatment NT1 and NT3 was not different and higher than this in NT2. The survival rate of clams in NT3 was highest (96.5%) while this value in NT1 was lowest (86.7%). NT2 has the survival rate of clams reaching 90.4%.

It can be said that *Tapes dorsatus* is species that usually bury themselves in the sand, so rearing them in the suspended baskets without bottom sand affects the natural behaviors of this species. This make them obtain the lowest values of size, weight and survival rate. In contrast, reared in the baskets with bottom sand (NT3) with the water circulation, the growth and survival rate of clams are higher because such condition is suitable for both the biological behaviors of the species and creates a good environment for them. In treatment NT2, although clams were raised in the natural bottom, the water circulation was worse than in NT3. Consequently, the growth and survival values of clams in NT2 were lower than these in NT3. Therefore, in fact, such raising method in NT3 is currently the popular commercializing rearing method in Khanh Hoa, Phu Yen, Ben Tre, Hai Phong, Quang Ninh.[1-7,13-15]

3.1.3. Condition index (CI) and gonad index (GI) of the wild clams

The results of examining the condition index (CI) and gonad index (GI) of clams after 60 days of rearing are shown in Table 3.

Table 3. Condition index and gonad index in the different treatments

Parameter		CI (mg/g)	GI
NT 1	*Initial*	156.2±17.7	2.6±0.5
	Finish	154.2±4.1^{c}	2.7±0.6^{b}
NT 2	*Initial*	156.2±17.7	2.6±0.5
	Finish	156.3±5.2^{b}	2.7±1.2^{b}
NT3	*Initial*	156.2±17.7	2.6±0.5
	Finish	183.7±4.9^{a}	3.2±0.5^{a}

Remark: For each parameter, the different letters indicate the significant difference ($p < 0.05$)

The initial condition index of the clams was 156.2±17.7 mg/g. After 60 days of raising, this index in NT1 decreased to 154.2 mg/g while this index in NT2 did not almost change (156.3 ± 5.2 mg/g). In contrast, the CI in NT3 increased. In comparison, it can be seen that the CI in NT3 was highest (183.7±4.9 mg/g), and the lowest value was in NT1 (154.2±4,1). This indicates that the raising condition in Treatment NT3 was suitable. Hua Thai Nhan *et al.*[5] did study on

https://doi.org/10.52111/qnjs.2022.16105

Ben Tre clams *M. lyrata* and suggested that the condition index of clams is related to maturity and changes in water salinity during the rearing process.[5]

The gonad index (GI) at the beginning of the experiment was 2.6±0.5. After 60 days of rearing, this index in NT3 obtained the higher value (3.2) than this value in NT1 and NT2 (2.7) ($p < 0.05$). Chu Chi Thiet and Martin S Kumar [15] found that when food was scarce (0.25 mg organic matter/g fresh weight/day), the gonads of *Ruditapes decussates* developed very slowly (taking about 70 days for females to reach the stage 4 of gonad). The CI and GI reaching the highest value in NT3 implicate the suitable rearing condition for the clams.[16]

3.1.4. The fecundity of clams in the experiment

After 60 days of rearing, 100 matured clam individuals in each treatment were selected to determine the fecundity. The results showed that the number of eggs/g fresh muscle of clams were highest in NT3 (28,339 ± 10,556) followed by NT2 (21,378 ± 12,981) and lowest in NT1 (20,221 ± 12,34) ($p < 0,05$). The number of eggs/g total weight in NT3 (3,537 ± 1,102) was higher than this in NT1 (2,256 ± 1,573) and NT2 (2,221 ± 1,012) but there was no difference of this between NT1 and NT2 ($p > 0,05$). The fecundity of the clams in this study was different from *M. lyrrata* with (19,322 ± 1,884) eggs/g total weight and (91,743 ± 7,769) eggs/g fresh muscle (Chu Chi Thiet and Martin S Kumar).[15]

Table 4. Percentage of female mangrove clam spawning and actual fecundity in rearing conditions

Parameter	NT1 (n=100)	NT2 (n=100)	NT3 (n=100)
Number of eggs/g total weight	2,256 ± 1,573 [b]	2,221 ± 1,012 [b]	3,537 ± 1,102 [a]
Number of eggs/g fresh muscle	20,221 ± 12,34 [c]	21,378 ± 12,981 [b]	28,339 ± 10,556 [a]

Remark: In the same row, the different letters indicate the significant difference ($p < 0.05$)

3.2. The results of examining the reproductive cycle of the wild clams

3.2.1. Biological parameters of the clams

It can be seen that number of female is always higher than that of male in most of the studied month, excepting February 2021 (Table 5). The percentage of female clams varied from 67%, males 29%, and 5% are non-sexual (hermaphrodite). The average percentage of females to 56% with an average value of 39%. This indicates that female clams are predominant in the wild population of this species in Thi Nai lagoon.

Table 5. Biological parameters of the wild clams

Month	Number of individuals	Length (mm)	Width (mm)	Total weight (g)	Muscle weight (g)	Gender (%)		
						Male	Female	Hermaph-rodite
06/2019	100	52.1 ± 0.3	45.2 ± 0.2	57.3 ± 1.28	4.9 ± 1.26	32	65	3
10/2019	100	51.2 ± 0.4	44.3 ± 0.3	59.4 ± 1.34	4.6 ± 1.22	38	53	9
02/2020	100	56.1 ± 0.3	47.2 ± 0.3	59.2 ± 1.50	5.1 ± 1.31	56	44	0
06/2020	100	52.4 ± 0.4	44.3 ± 0.5	55.1 ± 1.16	4.6 ± 1.24	44	49	7
10/2020	100	59.6 ± 0.5	48.4 ± 0.4	59.3 ± 1.28	6.5 ± 1.32	36	59	5
02/2021	100	53.8 ± 0.2	42.2 ± 0.3	55.3 ± 1.28	6.6 ± 1.27	29	67	4
06/2021	100	54.4 ± 0.2	43.3 ± 0.4	55.3 ± 1.28	4.2 ± 1.21	39	55	6
Average	700	54.4 ± 0.3	45.3 ± 0.3	57.3 ± 1.3	5.2 ± 1.30	39	56	5

Remark, L: lenght; R: width; H: height; W-total: meat total, W-meat: meat. There are 3 sexes of sugarcane scallops: male, female and hermaphrodite.

The results in Table 5 showed that in the same year (2020), the clams obtained the highest length and width, total weight, muscle weight with the values of 59.6 mm, 48.4 mm, 59.3 g, 6.5 g, respectively in October 2020, and the clams collected in February also obtained the high values of these parameters while these were lowest in June. The pooled data of 700 clam individuals showed that average length, width, total weight, muscle weight were 54.4 mm, 45.3 mm, 57.3 g and 5.2 g respectively.

3.2.2. Spawning season identification of the wild clam

The spawning season identification of the clams is based on GI and CI. The results of the study showed that GI was from 2.1 to 3.7. The clams in Stages 3 and 4 (genital and reproductive stages) fluctuated from 2.9 to 3.7 mg/g.

Table 6. Gonad index (GI) and condition index (CI) of the wild clams

Month	CI (mg/g)	GI
06/2019	68.9	2.1
10/2019	76.4	2.8
02/2020	80.3	2.9
06/2020	98.4	2.8
10/2020	110.4	3.7
02/2021	120.7	3.5
06/2021	70.4	2.7
Average	89.4	2.9

This percentage reached the highest value from October to February yearly (highest peak in October at 3.7 mg/g), the lowest in June is only 2.1 – 2.7 - 2.8 mg/g.

Similarly, CI of the wild clams also varied from 68.9 – 120.7 mg/g in the same tendency with GI (Table 6), that reached the high values in October and February of the year compared to June. Based on such variation of GI and CI of the clams, it can be identified that the spawning season of clams in Thi Nai lagoon - Binh Dinh is from October to February. Therefore, to preserve the resource of this species, it is necessary to avoid exploiting wild clams in this period of year.

4. CONCLUSIONS

- The environmental parameters in the experiment are suitable for growth and development of the clams *Tapes dorsatus*.

- The clams reared in the plastic baskets with sandy bottom of 10 - 15 cm obtain the highest weight, length, survival rate, CI, GI and fecundity.

- The spawning season of the wild clams is from October to February of years.

Acknowledgements

This research is conducted within the framework of science and technology projects at institutional level of Quy Nhon University under the project code T2019.619.14.

REFERENCES

1. Phung Bay. *Experimental study on hatchery production and commercial oyster farming in the estuary area of Binh Dinh province*, *Final Report*, 2012.

2. Phung Bay and Nguyen Thi Xuan Thu. *Cultivation of molluscs in Vietnam: History of formation, current status and development orientations for the future*, Collection of proceedings at the 3rd International Conference on Oysters, Taipei, Taiwan, 2009.

3. Nguyen Chinh. *Some species of bivalve mollusc (Bivalve Mollusc) have economic value in the sea of Vietnam*, Science and Technology Publisher, Hanoi, 1996.

4. Pham Viet Nam. *Experimental study on hatchery production and commercial oyster culture in the coastal estuary area of Tien Giang province*, final report, 2016.

5. Hua Thai Nhan, Ngo Thi Thu Thao, Huynh Han Chau and Tran Ngoc Hai. *Trial of maturation of clams Meretrix lyrata in blue water-tilapia system*. Collection of scientific reports of the

5th national molluscs conference-Nha Trang, September 17-18, 2007, Agricultural publisher, 2008, 375-383.

6. Summary report of the central agricultural extension project: Breeding of bivalves and molluscs, under the national agricultural extension program, Project implementation organization - Aquaculture Research Institute III, 2011 - 2013.

7. Phung Bay. *Production trial of the Sydney oyster, Saccostrea glomerata (Gould, 1850)*, Collection of scientific proceedings of the 5th national molluscs conference – Nha Trang, September 17-18, 2007, Agricultural publisher, 2008, 357-365.

8. K. E. Carpenter & V. H. Niem (Eds). *The living marine resources of the Western Central Pacific*, Rome: FAO Species Identification Guide for Fishery Purposes, 1998.

9. R. L. Walker, P. B. Heffernan. *Temporal and spatial effects of intertidal exposure on the gametogenetic cycle of the northern quahog,* Mercenaria mercenaria (Linnaeus, 1758), in coastal Georgia, J. Shellfish Res 13, 1994, 479-486.

10. R. D. Braley. *Reproductive condition and season of the giant clams Tridacna gigas and T. derasa utilising a gonad biopsy technique*, In Giant, Clams in Asia and the Pacific, Australian Centre For International Agricultural Research, 1988, 98-103.

11. Francis x. O'beirn, Peter B. Heffernan, Randal I,. Walker, Michelle L. Jansen. Young-of-the-Year Oyster, Crassostrea Reproduction in Coastal Georgia, *Estuaries,* **1996,** *19*(3), 651-658.

12. D. B. Quayle & G. F. Newkirk. *Farming bivalve molluscs methods for study and development advances in World Aquaculture*, Canada, The World Aquaculture Society in association with the International Development Research Center, 1989.

13. Ngo Anh Tuan, Chau Van Thanh and Vu Trong Dai. *Some reproductive characteristics of oysters (Crassostrea belcheri Sowerby, 1871) in Cha River and Ba Ria - Vung Tau province*, Collection of scientific proceedings of the 4th national conference on molluscs - Nha Trang (September 2005), Agricultural publisher, 2007, 263-273.

14. Ngo Thi Thu Thao and Truong Trong Nghia. *Effect of different salt concentrations on feed filtration rate, growth, survival rate and stress tolerance of Anadara granosa blood clam*, Collection of scientific proceedings of the 2nd national conference on mollusks – Nha Trang, August 2001, Agriculture Publishing House, 2001, 137-142.

15. Chu Chi Thiet and Martin S Kumar. *Document on production techniques of Ben Tre clams Meretrix lyrata Sowerby*, 1851, Aquaculture Research Sub-Institute of North Central Vietnam (ARSINC), 2008.

16. Truong Quoc Phu. *Research on some biological, biochemical and technical characteristics of clam Meretrix lyrata (Sowerby) with high yield.* Thesis of Doctorate in Agriculture, Nha Trang University of Fisheries, 1999.

TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC QUY NHƠN

Mật độ và đặc điểm của vi nhựa ở ngao Bộp (*Mactra grandis*) phân bố ở đầm Cù Mông, tỉnh Phú Yên

Lê Quốc Hội[1,2], Võ Văn Chí[1,*]

[1]Khoa Khoa học tự nhiên, Trường Đại học Quy Nhơn, Việt Nam
[2]Trường TH&THCS Cù Chính Lan, Sông Cầu, Phú Yên, Việt Nam

Ngày nhận bài: 24/08/2021; Ngày nhận đăng: 11/11/2021

TÓM TẮT

Mục tiêu của nghiên cứu này là đánh giá ô nhiễm vi nhựa ở ngao Bộp sinh sống ở đầm Cù Mông, tỉnh Phú Yên. Ngao Bộp thu được ở đầm được xử lý bằng KOH 10% để xác định các loại vi nhựa tồn tại trong ống tiêu hóa. Kết quả cho thấy có hai dạng vi nhựa trong ống tiêu hóa của ngao Bộp là vi nhựa dạng sợi và vi nhựa dạng mảnh, với tổng mật độ trung bình là 1,53 vi nhựa/cá thể ở mùa nắng và 8,93 vi nhựa/cá thể ở mùa mưa. Chiều dài các sợi vi nhựa chủ yếu nằm trong khoảng 300 – 2500 µm trong khi diện tích các mảnh vi nhựa nằm trong khoảng 45.000 - 600.000 μm^2. Màu sắc vi nhựa dạng sợi đa dạng hơn so với dạng mảnh, tuy nhiên nhìn chung màu trắng và màu vàng là hai màu chiếm ưu thế của các mẫu vi nhựa, ngoài ra còn có màu xanh lá, tím và xanh biển.

Từ khóa: *Ngao Bộp, ống tiêu hóa, vi nhựa, đầm Cù Mông, ô nhiễm.*

**Tác giả liên hệ chính.*
Email: vovanchi@qnu.edu.vn

JOURNAL OF SCIENCE
QUY NHON UNIVERSITY

Concentration and charateristics of microplastic in Big brown mactra clam (*Mactra grandis*) distributed in Cu Mong lagoon, Phu Yen province

Le Quoc Hoi[1,2], Vo Van Chi[1,*]

[1] *Faculty of Natural Sciences, Quy Nhon University, Vietnam*
[2] *Cu Chinh Lan junior and secondary school, Song Cau, Phu Yen, Vietnam*

Received: 24/08/2021; Accepted: 11/11/2021

ABSTRACT

The objective of this study is to assess microplastic contamination in Big brown mactra clam inhabiting in Cu Mong lagoon, Phu Yen province. The individuals of this clam collected from the lagoon were treated in KOH 10% to determine microplastic particles in their digestive system. The results disclosed two shapes of microplastic in the digestive system of Big brown mactra clam called fibers and fragments with average microplastic concentration of 1.53 particles per individual in sunny season and 8.93 ones per individual in rainy season. The length of the fibers was mostly from 300 to 2500 µm while the area of the fragments was dominant in the range of 45.000 to 600.000 µm^2. The colour of fibers was more diverse than that of fragments; however, in general, white and yellow were the predominant colours of the microplastic particles, followed by green, purple and blue.

Keywords: *Big brown mactra clam, digestive tract, microplastic, Cu Mong lagoon, contamination.*

1. INTRODUCTION

Microplastic pollution is a globally concerned problem. Microplastics, especially synthetic fibers and artificial fibers, have been found in many water bodies such as lakes, rivers, estuaries, seas, etc. These microplastics can be transported through food webs in freshwater, saltwater environments and can affect human health.[1-3] Because of such hazards from microplastics, a lot of researches on microplastics have been done in many different areas in the world. Lot of researchers have conducted many researches on microplastic in the environments. For example, Free et al.[4] did the microplastic research in Hovsgol lake – Mongolia; Cabernard et al.[5] compared Raman and Fourier transform infrared spectroscopy to quantify microplastics in the aquatic environment; Wenfeng aet al.[6] carried out the research of microplastic in the domestic freshwater in China; Zhao et al.[7] studied on microplastic in surface water in Yangtze river system – China. In addition, studies on microplastic accumulation in molluscs have been considered and done by lots of researchers such as Li et al.[8] for *Mytilus edulis* in coastal areas – China, with a density of 1.5 – 7.6 particles/individual; Abolfazl et al.[9] for *Amiantis umbonella, Amiantis purpuratus, Pinctada radiate, Cerithidea cingulata, Thais mutabilis* in coastal areas of Persia bay – Iran, with density of 3.9 – 6.9 particles/individual; Li et al.[10] for 9 bivalve species as *Sc. subcrenata, T. granosa,*

**Corresponding author.*
Email: vovanchi@qnu.edu.vn

https://doi.org/10.52111/qnjs.2022.16106

My. Galloprovincialis, P. yessoensis, A. plicatula, Si. Constricta, R. philippinarum, Me. lusoria, C. sinensis collected from seafood markets in China, with 4.3 – 57.2 particles/individual; Sami et al.[11] for 6 species as *Mytilus galloprovincialis, Ruditapes decussatus, Crassostrea gigas, Hexaplex trunculus, Bolinus brandaris, Sepia officinalis* in Bizerte lagoon –Tunisia, with a density of 703.95 ± 109.80 to 1482.82 ± 19.20 particles/kg fresh body weight. However, although Vietnam is the fourth largest plastic emitting country in the world,[12] so far, researches on microplastic has been rare. It is noted that there are a few studies on the accumulation of microplastics in organisms carried out by some researchers such as the study of Nam et al.[13] on green mussels in Tinh Gia - Thanh Hoa.

Phu Yen as well as many other provinces in Vietnam is being seriously affected by plastic waste. In particular, this is one of the provinces with a border adjacent to the sea, a long coastline, and terrain sloping from West to East, so most of plastic wastes discharged eventually drift to the coastal areas such as bays, lagoons and then flow to seas and oceans. However, there has been no study on microplastic carried out in Phu Yen to assess the pollution level of this waste.

Cu Mong, the habitats of many valuable seafood species as seahorse, oyster, Big brown mactra clam, is one of two important lagoons of Phu Yen, that brings to benefits for local residents. Big brown mactra clam is one of the molluscs widely distributed in this lagoon and is a favorite seafood species of local residents well as widely consumed in other provinces. Based on the field surveys, it can be seen that Big brown mactra clam is a benthic species and often bury itself in the bottom. Like many other bivalve molluscs, Big brown mactra clam is the filter feeder with feeding mainly on algae and organic detritus. With such feeding habits, this species can eat microplastic particles together with their natural food. However, Arapov et al.[14] suggested that although bivalve species are filter feeders, they are able to select food basing on size, shape, nutritional value or chemical composition on the surface of food particles. Therefore, in this study, we chose Big brown mactra clam to study to see whether this species is contaminated microplastics in the digestive tract.

2. METHODS

2.1. Sample collection

The clams were collected from fishermen fishing in the range of V1 ($13^0$35'59"N - $109^0$13'54"E) and V2 ($13^0$32'52"N - $109^0$16'26"E (Figure 1), then were put into the zip bags, frozen and brought to the lab for later analysis.

Figure 1. The sampling sites in Cu Mong lagoon

Samples were collected every three months during 3 months of wet season (October to December) and 3 months of dry season (March to May). A total of 60 individuals of clams were collected (10 clams/month) to examine.

2.2. Sample treatment and analysis

Each individual of clam was separately treated and analysed. At the lab, clams were rinsed by water filtered through the glass fiber filters (GF/A, pore sizes of 1.6 µm), then weighed (weight of shell and muscle).

Next, the clams were put on petri disk to dissect and got the digestive tracts.

The digestive tracts were contained in glass beakers and treated in KOH 10% at 60 ^{0}C within 24 hours according to Alexandre.[15]

The samples after treated in KOH were filtered through 1mm mesh size sieve to discharge extraneous matters and collect microplastic particles from 1- 5 mm.

The samples filtered through the 1 mm sieve were filtered through 250 µm mesh size sieve, then the samples on the sieve were rinsed into a beaker to do the overflow technique using saturated NaCl solution. Finally, sample solution was filtered through a 1.6 µm GF/A filters according to the method of Emilie et al.[16]

The filters were observed under the Leica S9i stereomicroscope to determine the shape of microplastics based on the classification of Free et al.,[4] that includes 5 shapes such as fibers, fragments), pellet, film and foam. The LASX software of stereomicroscope was used to take pictures, measure the size and determine color of each microplastic. Then, all details of analysis were saved as Excel format for later data analysis.

2.3. Data analysis

Data for analysis include shape, concentration, size (length of fibers and area of fragments) and color of microplastic. Because of lack of facilities to determine the chemical origin of microplastics, we only considered fibers from 300 – 5000 µm in length and fragments from 45,000 – 25,000,000 μm^2 in area according to suggestion of GESAMP.[17]

Microsoft Excel 2013 was used to calculate the necessary parameters and make diagrams. Anova single factor in Microsoft Excel 2013 was applied to check the concentration difference of microplastic.

2.4. Microplastic contamination control

To control microplastic contamination from the surrounding environment during sample treatment and analysis, we followed suggestions of GESAMP[17] such as cleaning the working area with alcohol before analysing and treating samples, wearing cotton clothes and rubber gloves, rinsing equipment with water filtered through GF/A glass fiber filter (pore size 1.6 µm) before use. In addition, at each period of sample analysis and treatment, we placed a new filter paper in a petri dish nearby places we were working to check microplastic contamination. After finishing each process, we observed this filter paper under the Leica S9i stereo microscope to check microplastic contamination from the surrounding environment.

During sample analysis and processing, we did not detect any microplastics on the control filter papers.

3. RESULTS AND DISCUSSION

3.1. Microplastic concentration

Two shapes of microplastic found in the digestive tracts of Big brown mactra clams were fibers and fragments. The results in Table 1 showed that in the same season, concentration of fibers was higher than that of gragments. The study results of Li et al.[8] on *Mytilus edulis*, Li et al.[10] on 8 mollusc species (*Sc. subcrenata, T. granosa, My. galloprovincialis, P. yessoensis, Si. constricta, R. philippinarum, Me. lusoria, C. sinensis*) and Sami et al.[11] on 6 mollusc species (*Mytilus galloprovincialis, Ruditapes decussatus, Crassostrea gigas, Hexaplex trunculus, Bolinus brandaris, Sepia officinalis*) also indicated the dominance of fibers compared to fragments.

For different seasons, microplastic concentration (fibers, fragments and total) in wet season was higher than that in dry season ($p < 0,05$). It can be seen that weight of clams was not significantly different between two seasons. The results also indicated that the correlation between weight of clams and number of microplastic in wet season was very low ($r = 0.03$) while that in dry season was negative ($r = -0.55$). Therefore, it can be speculated that the difference of microplastic concentration in the digestive tract of clams in two seasons may be due to other factors instead of clam size. Clearly, there is no river going to Cu Mong lagoon and water in the lagoon is exchanged with the sea

https://doi.org/10.52111/qnjs.2022.16106

through an unique gate. Thus, in rainy season, wastes including plastic waste are washed away from the surrounding areas into the lagoon, but because there is no strong flows from the river, microplastics are easily deposited in the lagoon instead of being washed into the sea. In contrast, in the dry season, the rainfall is low, so there may be a small amount of waste around the lagoon going into the lagoon and hence the amount of microplastics in the dry season will be less than that in rainy season. Therefore, with bottom living habit and filter feeding, Big brown mactra clams can ingest a higher number of microplastic in rainy season compared to dry season.

Table 1. Microplastic concentration in the digestive tracts of Big brown mactra clams

Season	Average weight of clams (g)	Number of microplastic fibers per individual	Number of microplastic fragments per individual	Total number of microplastic per individual	Correlation coefficient between clam weight and number of microplastic (r)
Dry season (n=30)	14.97 ± 1.28^{a}	0.87 ± 0.53^{b}	0.67 ± 0.27^{b}	1.53 ± 0.64^{b}	-0.55
Wet season (n=30)	12.80 ± 4.77^{a}	7.30 ± 3.28^{a}	1.63 ± 1.54^{a}	8.93 ± 3.72^{a}	0.03

Remark: For each parameters (each column), the different letters indicate the significant difference ($p < 0.05$)

Some researchers also present microplastic contamination level as number of microplastic particles per individual such as Li et al.,[8] Abolfazl et al.[9] Although these authors examined the microplastic with size much smaller than this in our study, microplastic concentration in those studies just fluctuated from 1.5 to 7.6 particles/individual while this number in our study was from 1.53 to 8.93 particles/individual (Table 2). This indicates that microplastic contamination level in Cu Mong lagoon is pretty high, which is a bad sign of environment for this ecosystem.

Table 2. Microplastic contamination in some mollusc species

Authors	Study sites/regions	Studied species	Microplastic concentration	Microplastic size (µm)
Li et al., 2016[8]	Coastal areas in China	*Mytilus edulis*	1.5 – 7.6 (particles per individual)	5 – 5000
Abolfazl et al., 2018[9]	Coastal areas in Persia bay, Iran	*Amiantis umbonella, Amiantis purpuratus, Pinctada radiate, Cerithidea cingulata Thais mutabilis*	3.9 – 6.9 (particles per individual)	10 – 5000
This study	Cu Mong lagoon, Phu Yen, Vietnam	*Mactra grandis*	1.53 – 8.93 (particles per individual)	300 - 5000

3.2. Microplastic size

Microplastic fibers found in this study had the different ranges of size but there was difference between wet season and dry season. In the wet season, microplastics almost existed in all size groups of 300 - 5000 µm (except the size group of 4500 - 4700 µm) while in the dry season, microplastics were only in the size group of 500 - 3100 µm (Figure 2).

In general, the microplastic fibers were mainly from 300 - 2500 µm in length, but there was different distribution between dry season and wet season. Specifically, in wet season, the size classes of 1300 - 1500, 1500 - 1700 µm and 1700 - 1900 µm had the largest number, accounting for 10.68%, 9.26% and 10.19% of the total, respectively while the size classes of 1500 - 1700, 1700 - 1900, 1900 - 2100 and 2100 - 2300 µm accounted for the highest proportion in the dry season, with 11.54%, 19.23%, 15.38% and 11,54% of the total respectively.

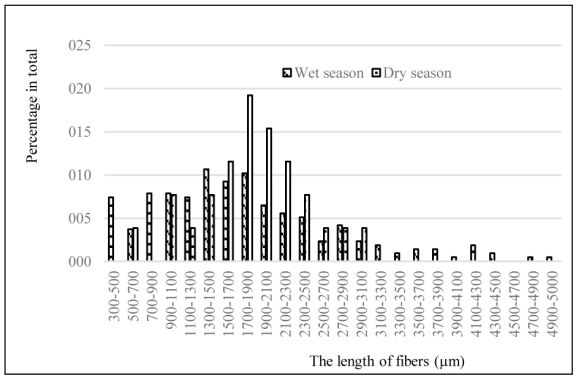

Figure 2. The length of fibers (µm) according to occurrence rate (%)

The size distribution of microplastic fragments also differed between the two seasons. In the dry season, fragments were mostly in the size classes of 200,000 - 400,000 µm^2 and 400,000 - 600,000 µm^2 (accounting for 45.00% and 25.00% of the total, respectively) and there was no fragments in the size group 800,000 – 1,000,000 µm^2. In wet season, there was no microplastic in the size group of 1,000,000 - 1,200,000 µm^2 and mainly in the size group of 45,000 - 200,000 m^2 and 200,000 - 400,000 µm^2, with 47.92% and 27.08% of the total, respectively (Figure 3).

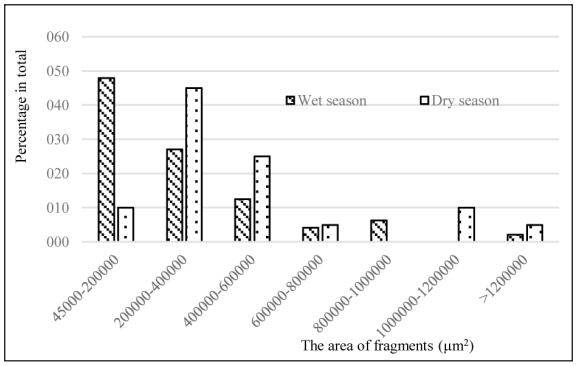

Figure 3. The area of fragments (µm^2) according to occurrence rate (%)

In general, our study results show that the small size class is dominant in total number of microplastic found. Specifically, most of fibers have the length of 300 - 2500 µm, accounting for 88.46% in total in dry season and 83.59% in total in wet season; fragments are predominant in area class of 45,000 – 600,000 µm^2, with 80% in total in dry season and 86.5% in total in dry season (Figure 4). Cabernard et al.,[5] Wenfeng et al.,[6] Zhao et al.[7] also indicated the dominance of small size microplastics. Tayler et al. reported that water currents and UV radiation are responsible for producing a large number of small size microplastics.[18]

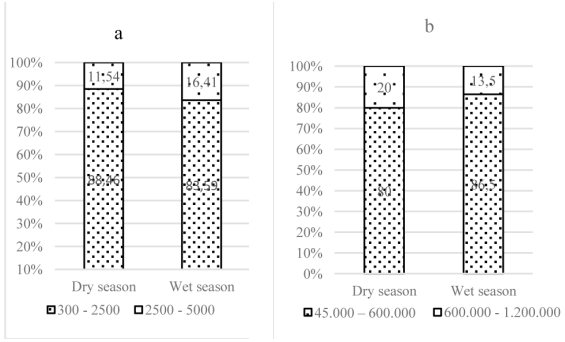

Figure 4. The pooled percentage of microplastic according to length (µm) of fibers (a) and area (µm^2) of fragments (b)

3.3. Colors of microplastics

The colors of microplastics found in the digestive tracts of Big brown mactra clams were quite diverse. Of which, the colors of microplastic fibers were more diverse than this of microplastic fragments. Specifically, the fibers had 10 colors (blue, black, brown, green, orange, pink, purple, white, yellow, gray) while the fragments only

https://doi.org/10.52111/qnjs.2022.16106

comprised 5 colors (white, purple, yellow, orange, green) (Figure 5 and Figure 6).

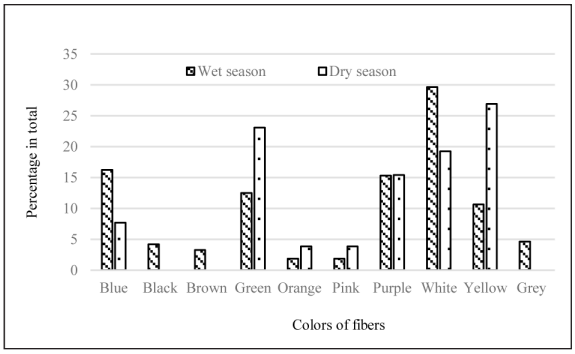

Figure 5. The colors of microplastic fibers according to occurrence rate (%)

The color of fibers changed between the seasons. In dry season, yellow predominated (26.92%), followed by green (23.08%), white (19.23%), and purple (15.38%) while white dominated (29.63%), followed by blue (16.20%), purple (15.28%) and green (12.50%) in rainy season.

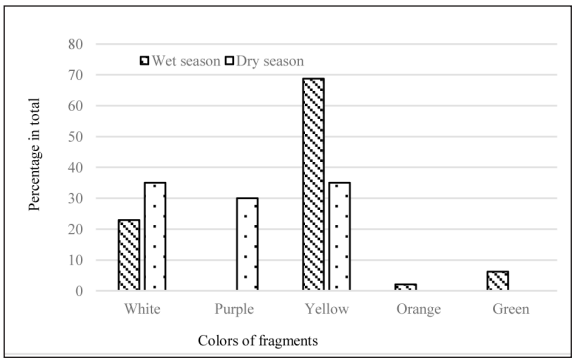

Figure 6. The colors of microplastic fragments according to occurrence rate (%)

For fragments, white and yellow were the main colors, and purple also contributed a large proportion in total. However, there is a difference between two seasons. In wet season, the fragments had 4 colors, of which yellow accounted for the highest percentage (68.75%), followed by white (22.92%) while there were only 3 colors, that are yellow, purple and white with the approximately equal proportions (Figure 6).

In general, in this study, white and yellow are the predominant colors, in addition to green, blue and purple. Ayu et al.[19] reported that white is also the dominant color found in the digestive tracts of fish (accounting for 79.2%) but followed by blue (7.03%) instead of yellow as in our study. In another study on six mollusc species (*Mytilus galloprovincialis, Ruditapes decussatus, Crassostrea gigas, Hexaplex trunculus, Bolinus brandaris and Sepia officinalis*), black, blue, and white are the dominant colors for microplastic fibers and blue, red and black are the main colors of fragments.[11] Thus, the colors of microplastics found in aquatic animals are quite different, which might be due to the different sources of microplastics in the study locations. This is also reported by Gallagher et al.[20]

4. CONCLUSION

Microplastics accumulate in the digestive tracts of Big brown mactra clams distributed in Cu Mong lagoon, Phu Yen province with the concentration of 1.53 to 8.93 particles/individual. The clams collected in the wet season are contaminated with higher number of microplastics than those collected in dry season.

Microplastic fibers as well as microplastic fragments found in the digestive tracts of this clam species are mostly in the small size class (300 - 2500 µm for fibers and 45,000 - 600,000 μm^2 for fragments).

The color of microplastics in the digestive tracts of the clams is quite diverse, in which the color of fibers is more diverse than this of fragments. In addition, the color of microplastics also changes in two seasons.

REFERENCES

1. M. C. M. Blettler, E. Abrial, F. R. Khan, N. Sivri, L.A. Espinola. Freshwater plastic pollution: recognizing research biases and identifying knowledge gaps, *Water Research,* **2018,** *143,* 416 – 424.

2. A. L. Andrady. Microplastics in the marine environment, *Marine Pollution Bulletin*, **2011**, *62*, 1596–1605.

3. F. Wang, C. S. Wong, D. Chen, X. Lu, F. Wang, E.Y. Zeng. Interaction of toxic chemicals with

microplastics: a critical review, *Water Research*, **2018,** *139*, 208–219.

4. C. M. Free, O. P. Jensen, S. A. Mason, M. Eriksen, N. J. Williamson, B. Boldgiv. High-levels of microplastic pollution in a large, remote, mountain lake, *Marine Pollution Bulletin*, **2014**, *85*, 156–163.

5. L. Cabernard, L. Roscher, C. Lorenz, G. Gerdts and S. Primpke. Comparison of Raman and fourier transform infrared spectroscopy for the quantification of microplastics in the aquatic environment, *Environmental Science & Technology*, **2018**, *52*, 13279–13288.

6. W. Wang, A. W. Ndungu, Z. Li, J. Wang. Microplastics pollution in inland freshwaters of China: A case study in urban surface waters of Wuhan, China, *Science of the Total Environment*, **2017,** *575*, 1369 –1374.

7. S. Zhao, L. Zhu, T. Wang, D. Li. Suspended microplastics in the surface water of the Yangtze estuary system, China: first observations on occurrence, distribution, *Marine Pollution Bulletin,* **2014**, *86*, 562 –568.

8. J. Li, X. Qu, L. Su, W. Zhang, D. Yang, P. Kolandhasamy, D. Li, H. Shi. Microplastics in mussels along the coastal waters of China, *Environmental Pollution,* **2016**, *214*, 177–184.

9. A. Naji, M. Nuri, A. D. Vethaak. Microplastics contamination in molluscs from the northern part of the Persian Gulf, *Environmental Pollution*, **2018**, *235*, 113-120.

10. J. Li. D. Yang, L. Li, K. Jabeen, H. Shi. Microplastics in commercial bivalves from China, *Environmental Pollution,* **2015**, *207*, 190–195.

11. S. Abidli, Y. Lahbib, N. T. El Menif. Microplastics in commercial molluscs from the lagoon of Bizerte (Northern Tunisia), *Marine Pollution Bulletin*, **2019**, *142*, 243–252.

12. J. R. Jambeck, R. Geyer, C. Wilcox, T. R. Siegler, M. Perryman, A. Andrady, R. Narayan, K. L. Law, Plastic waste inputs from land into the ocean, *Science*, **2015**, *347*(6223), 768–771.

13. P. N. Nam, P. Q. Tuan, D. T. Thuy, P. Quynh & F. Amiard. Contamination of microplastic in bivalve: first evaluation in Vietnam, *Vietnam Journal of Earth Sciences*, **2019**, *41*(3), 252–258.

14. J. Arapov, D. Ezgeta–Balic, M. Peharda, Z. Nincevic Gladan. Bivalve feeding — how and what they eat?, *Ribarstvo*, **2010**, *68*(3), 105-116.

15. A. Dehaut, Anne-Laure Cassone, L. Frere, L. Hermabessiere, C. Himber, E. Rinnert, G. Riviere, C. Lambert, P. Soudant, A. Huvet, G. Duflos, Ika Paul-Pont. Microplastics in seafood: Benchmark protocol for their extraction and characterization, *Environmental Pollution*, **2016**, *215*, 223-233.

16. Emilie Strady, Thi Ha Dang, Thanh Duong Dao, Hai Ngoc Dinh, Thi Thanh Dung Do, Thanh Nghi Duong, Thi Thuy Duong, Duc An Hoang, Thuy Chung Kieu-Le, Thi Phuong Quynh Le, Huong Mai, Dang Mau Trinh, Quoc Hung Nguyen, Quynh Anh Tran-Nguyen, Quoc Viet Tran, Tran Nguyen Sang Truong, Van Hai Chu, Van Chi Vo. Baseline assessment of microplastic concentrations in marine and freshwater environments of a developing Southeast Asian country, Viet Nam, *Marine Pollution Bulletin*, **2021,** *162*, 111870 (1-10).

17. GESAMP, In: P. J. Kershaw, A. Turra, F. Galgani, (Eds). *Guidelines or the monitoring and assessment of plastic litter and microplastics in the ocean, GESAMP Joint Group of experts on the scientific aspects of marine environmental protection*, London, UK, 2019.

18. T. S. Hebner, M. A. Maurer-Jones. Characterizing microplastic size and morphology of photodegraded polymers placed in simulated moving water conditions, *Environmental Science: Processes & Impacts*, **2020**, *22*, 398-407.

19. A. R. Hastuti, T. F. Djamar, T.F. Lumbanbatu, Y. Wardiatno. The presence of microplastics in the digestive tract of commercial fishes off Pantai Indah Kapuk coast, Jakarta, Indonesia, *Biodiversitas,* **2019**, *20*(5), 1233-1242.

20. A. Gallagher, A. Rees, R. Rowe, J. Stevens, P. Wright. Microplastics in the Solent estuarine complex, UK: an initial assessment, *Marine Pollution Bulletin,* **2016**, 243 –249.

https://doi.org/10.52111/qnjs.2022.16106

TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC QUY NHƠN

Phương pháp Newton chỉnh hóa không chính xác cho tối ưu không ràng buộc với tốc độ hội tụ nhanh

Nguyễn Văn Vũ, Trần Ngọc Nguyên*

Khoa Toán và Thống kê, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 10/09/2021; Ngày nhận đăng: 22/11/2021

TÓM TẮT

Bài báo đề xuất một phương pháp Newton chỉnh hóa không chính xác để giải các bài toán tối ưu không ràng buộc. Thuật toán được đề xuất thuộc vào lớp lược đồ lặp trong - ngoài. Thay vì giải các hệ tuyến tính một cách chính xác, các chương trình giải các hệ tuyến tính lặp sẽ được áp dụng để tìm ra các hướng tìm kiếm xấp xỉ. Chúng tôi sẽ chứng minh rằng thuật toán không chính xác sẽ bảo toàn tính chất hội tụ địa phương nhanh của các thuật toán chính xác. Một số thực nghiệm số sẽ được thực hiện để chỉ ra những điểm tốt của thuật toán được đề xuất.

Từ khóa: *Tối ưu không ràng buộc, phương pháp Newton không chính xác, chỉnh hóa, chặn sai số địa phương.*

* *Tác giả liên hệ chính.*
Email: tranngocnguyen@qnu.edu.vn

https://doi.org/10.52111/qnjs.2022.16107

JOURNAL OF SCIENCE
QUY NHON UNIVERSITY

Inexact regularized Newton method for unconstrained optimization with rapid rate of convergence

Van Vu Nguyen and Tran Ngoc Nguyen*

Faculty of Mathematics and Statistics, Quy Nhon University, Vietnam

Received: 10/09/2021; Accepted: 22/11/2021

ABSTRACT

This paper proposes an inexact regularized Newton method for solving unconstrained optimization problems. The proposed algorithm belongs to the class of outer-inner iteration scheme. Instead of solving exactly linear systems, iterative linear solver will be applied to find approximate search directions. We will show that the inexact algorithm preserved the fast local convergence property of exact algorithms. Some numerical experiments are also conducted to show the benefits of our proposed algorithm.

Keywords: *Unconstrained optimization, inexact Newton method, regularization, local error bound.*

1. INTRODUCTION

Let us consider the following unconstrained minimization problem

$$\min_{x\in\mathbb{R}^n} f(x), \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a twice differentiable function. This type of problem has many applications in machine learning, engineering, economics,....[1,2] Let $\mathcal{S}$ be the solution set of the problem (1). We may assume that the minimization problem (1) has a local solution $x^* \in \mathcal{S}$. A conventional assumption which is used when solving this problem is the second order sufficient condition (SOSC), i.e., the Hessian matrix $\nabla^2 f(x^*)$ is positive definite, see, e.g. Theorem 2.4 in[2]. This assumption is very important because it implies that x^* is the unique local solution of (1). This is crucial for the fast local convergence of Newton algorithm, see, e.g. Theorem 3.5 in[2]. However, in practice, this assumption is somewhat strict and it limits the applicability of numerical algorithm. With the lack of this assumption, behavior of numerical methods can be very bad and sometimes algorithms cannot work. Recently, regularized optimization methods emerge as good alternatives for resolving problems of type (1), in both contexts of convex optimization[3–5] and of nonconvex optimization.[6,7] The main idea of these regularized methods is tend to solve a sequence of regularized problem of the form

$$\min_{x\in\mathbb{R}^n} F_k(x) := f(x) + \frac{\theta_k}{2}\|x - x_k\|^2, \tag{2}$$

where x_k is the iterate at iteration k, $\theta_k > 0$ is a regularization parameter. Instead of the SOSC, in these papers, the authors assume that the gradient of objective provides a local error bound condition at some $x^* \in \mathcal{S}$. Such a condition means that the distance from a point to the solution set of the problem $\mathcal{S}$ can be upper bounded by some term related to the gradient at that point (see (10) below).

* *Corresponding author.*
Email: tranngocnguyen@qnu.edu.vn

https://doi.org/10.52111/qnjs.2022.16107

Very recently, authors in[8] proposed a Newton method applied to the first order optimality conditions of (2). This algorithm belongs to the class of outer-inner iteration scheme. The main role of outer iterations is to compute a trial iterate x_k^+ such that

$$(H_k + \theta_k I)(x_k^+ - x_k) = -\nabla f(x_k), \qquad (3)$$

where H_k is an approximation of the Hessian matrix $\nabla^2 f(x_k)$, x_k is the current iterate and θ_k is updated at the beginning of each iteration. If the trial iterate makes a sufficient reduction on the objective function $F_k(x_k^+)$ and its gradient $\nabla F_k(x_k^+)$, it will be set as the starting point of the next iteration, i.e., $x_{k+1} = x_k^+$. If this is not the case, a sequence of inner iterations for minimizing F_k will be applied for obtaining a sufficient reduction on $F_k(x_{k+1})$ and $\nabla F_k(x_{k+1})$. This outer-inner iterations are also applied in the framework of constrained optimization, see[9–15] for further reading. In both outer and inner iterations, the main computational cost lies in solving exactly the linear system (3). In practical applications where the size of problems are very enormous, factorizing a large matrix to solve such a linear system may take a long time to execute the algorithm. To deal with this problem, inxact Newton method has been proposed for solving nonlinear equation[16] and unconstrained optimization.[17] This method is also applied in the framework of constrained optimization, e.g..[14,18,19] However, in these papers, one still resorts to the second order sufficient conditions for the fast local convergence of their algorithms. In[20], the authors proposed a regularized trust-region Newton method for solving (1). In numerical experiments, authors considered the possibility of applying an inexact solver for solving subproblem which is somewhat similar to (3). However, convergence analysis for their algorithm in this inexact case has not been studied.

In this paper, we will propose an inexact regularied Newton method for solving problem (1). Our algorithm has the same vain of outer-inner algorithm scheme as.[8] However, instead of solving exactly linear systems at each iteration which maybe too expensive, we will introduce an application of inexact method to their algorithm. In this case, the tolerance of the inexactness will be considered carefully so that the fast convergence of the algorithm is still preserved. In particular, at each iteration k, an inexact linear solver, e.g. conjugate gradient, will be applied to solve the "inexact" solution x_k^+ that satisfies

$$\|(H_k + \theta_k I)(x_k^+ - x_k) + \nabla f(x_k)\| \leq \eta_k, \quad (4)$$

where $\{\eta_k\}$ is a sequence of positive number which must be chosen. Many iterative linear solvers can be applied for solving (4), e.g., conjugate gradient method,[21] MINRES,[22] LSQR,[23] GMRES,[24] LSMR.[25] Since the coefficient matrix $H_k + \theta_k I$ is positive definite, we will choose the conjugate gradient method for solving (4). We will prove that this inexactness does not affect to the fast local convergence of the algorithm in[8] even for degenerate cases. More specifically, our proposed algorithm attains a superlinear convergence under a local error bound condition which is milder than the usual SOSC. These good theoretical results will be verified by some numerical experiments. In addition, numerical results also show us that the proposed inexact algorithm can help to reduce the computational time compared with exact algorithms.

The paper is organized as follows. Some notations and description of algorithm will be introduced in Section 2. Section 3 is devoted to the convergence analysis of the proposed algorithm. Some numerical experiments are reported in Section 4 to verify

https://doi.org/10.52111/qnjs.2022.16107

theoretical results and to demonstrate the effectiveness of our proposed algorithm. The paper ends with some conclusion.

2. NOTATION AND ALGORITHM DESCRIPTION

Algorithm 1: (kth iteration)

Input: $m \in \mathbb{N}$, $\gamma > 0, \sigma \in (0,1)$, $\bar{\theta} > 0$, $\kappa > 0$, $\epsilon > 0$, $\eta_{-1} > 0$ and $x_0 \in \mathbb{R}^n$

Output: an optimal solution x_k

1 If $\|\nabla f(x_k)\| \leq \epsilon$, then terminate the algorithm.

2 Choose $\theta_k > 0, \delta_k > 0$ such that $H_k = \nabla^2 f(x_k) + \delta_k I \succeq 0$ and set $\theta_k = \min\{\gamma \|\nabla f(x_k)\|^\sigma, \bar{\theta}\}$, $\eta_k = \kappa \min\{\|\nabla f(x_k)\|^{1+\sigma}, \eta_{k-1}\}$.

3 Compute an trial iterate x_k^+ which satisfies the condition (4).

4 Choose $\zeta_k > 0$ such that $\{\zeta_k\} \to 0$. If $\|\nabla f(x_k^+)\| \leq \zeta_k$, then set $x_{k+1} = x_k^+$. Otherwise, apply a sequence of inner iterations to find x_{k+1} such that

$$\|\nabla f(x_{k+1})\| \leq \zeta_k. \qquad (5)$$

5 Set $k \leftarrow k+1$ and go to Step 1.

Notation For two real vectors $x = [x_1, x_2, \ldots, x_n]^\top$ and $y = [y_1, y_2, \ldots, y_n]^\top$ in the vector space $\mathbb{R}^n$, $x^\top y$ is used to denote the Euclidean scalar product. The associated norm is the ℓ_2-norm, i.e., $\|x\| = (x^\top x)^{1/2}$. The notation $x \leq y$ ($x \geq y$) indicates that $x_i \leq y_i$ (resp. $x_i \geq y_i$) for all $i = \overline{1,n}$. For a vector $x \in \mathbb{R}^n$, the notation $\mathrm{diag}(x)$ stands for the diagonal matrix whose diagonal entries are components of vector x. The induced norm of a rectangular matrix M is defined by $\|M\| = \max\{\|Mx\| : \|x\| \leq 1\}$. Let M be a square symmetric matrix, i.e. $M = M^\top$. The smallest eigenvalue of the matrix M is denoted by $\lambda_{\min}(M)$. The notation $M \succeq 0$ means that M is a positive semi-definite matrix, i.e. $\lambda_{\min}(M) \geq 0$. The open ball of radius r and center x is denoted by $B(x,r) = \{y | \|y - x\| < r\}$. The solution set of the problem (1) will be denoted by $\mathcal{S}$. For every $x \in \mathbb{R}^n$, the notation $\mathrm{d}(x) = \inf_{y \in \mathcal{S}} \|x - y\|$ denotes the distance from x to the solution set $\mathcal{S}$. In this case, if the solution set is nonempty, the notation $\bar{x}$ will be used to denote the projection of x onto $\mathcal{S}$, i.e., $\|x - \bar{x}\| = \mathrm{d}(x)$.

We now introduce our proposed algorithm in this paper. Let m be a natural number and $\gamma > 0, \sigma \in (0,1)$, $\kappa > 0$, $\eta_{-1} > 0$, $\bar{\theta} > 0$ and $\epsilon > 0$. At the beginning of the algorithm, a starting point $x_0 \in \mathbb{R}^n$ should be defined. The details of our algorithm is given in Algorithm 1.

The first step introduces the stopping condition of the algorithm. In Step 2, the regularization parameter δ_k will be chosen such that the approximation of the Hessian matrix is positive semi-definite. In particular, this parameter is chosen such that

$$\delta_k \leq \beta_1 \max\{0, -\lambda_{\min}(\nabla^2 f(x_k))\}, \forall k \in \mathbb{N}, \qquad (6)$$

for some $\beta_1 \geq 1$. This choice means that δ_k is not less than the absolute value of the minimum eigenvalue of $\nabla^2 f(x_k)$. This implies that the matrix $H_k := \nabla^2 f(x_k) + \delta_k I$ is positive semi-definite. In[8] and ,[20] authors proposed to choose $\delta_k = \beta_1 \max\{0, -\lambda_{\min}(\nabla^2 f(x_k))\}$. Apparently, this choice validates the requirement (6). In our algorithm, we adopt the same strategy to choose δ_k. It worths to note that recently, authors in[26] propose a simple search algorithm based on the indefinite factorization method MA57[27] to find an δ_k satisfying (6). The regularization parameter θ_k and the "forced" parameter will be defined based on the norm of the gradient of

https://doi.org/10.52111/qnjs.2022.16107

the current iteration x_k. Step 3 is devoted to calculate an approximation solution x_k^+ with the tolerance η_k defined in Step 2. In Step 4, we choose a tolerance for the condition to trigger the inner iteration algorithm. If the trial iterate x_k^+ creates a sufficient reduction on the first order optimality condition, we set it as the starting point of the next iteration. Otherwise, we will apply a sequence of inner iteration to find an iterate x_{k+1} such that the condition (5) holds true.

Because our main concern in this paper is the local behavior of an inexact regularized Newton method, the globalization scheme is not mentionned here. Nevertheless, we introduce Step 4 to show the possibility of applying globalization strategies. It should be of interest to develop our local scheme to global one by using globalization techniques as in.[8,26]

3. ASYMPTOTIC ANALYSIS OF THE ALGORITHM

Asymptotic analysis of Algorithm 1 is conducted under the following assumptions.

Assumption 1. The function f is twice differentiable, $\nabla^2 f$ is locally Lipschitz continuous and the set $\mathcal{S}$ of minimizers to (1) is nonempty.

From the closeness of $\mathcal{S}$ and the coerciveness of the norm, for every $x \in \mathbb{R}^n$, there exists $\bar{x} \in \mathcal{S}$ such that

$$\mathrm{d}(x) = \|x - \bar{x}\|. \tag{7}$$

Assumption 2. The gradient provides a local error bound condition at some $x^* \in \mathcal{S}$.

From the two above assumptions, there exist postive numbers ℓ, L, r, τ such that for all $x, y \in B(x^*, r)$,

$$\|\nabla f(x) - \nabla f(y)\| \le \ell \|x - y\|, \tag{8}$$
$$\|\nabla^2 f(x) - \nabla^2 f(y)\| \le L\|x - y\|, \tag{9}$$
$$\mathrm{d}(x) \le \tau \|\nabla f(x)\|, \tag{10}$$
$$\gamma \|\nabla f(x)\|^\sigma \le \bar{\theta}. \tag{11}$$

From the definition of θ_k and η_k in Step 2 of Algorithm 1 and inequalities (8), (10) and (11) for all $k \in \mathbb{N}$ such that $x_k \in B(x^*, r)$,

$$\theta_k = \gamma \|\nabla f(x_k)\|^\sigma \ge \gamma b^\sigma \mathrm{d}(x_k)^\sigma, \tag{12}$$
$$\eta_k = \kappa \|\nabla f(x_k)\|^{1+\sigma} \le \kappa \ell^{1+\sigma} \mathrm{d}(x_k)^{1+\sigma}, \tag{13}$$

where $b = 1/\tau$.

Firstly, we recall a result about the upper bound of the regularization parameter δ_k satisfying (6) via the distance function.

Lemma 1 (Lemma 2 in[8])**.** *For all $k \in \mathbb{N}$ such that $x_k \in B(x^*, r/2)$, we then have $\delta_k \le \beta_1 L \mathrm{d}(x_k)$.*

Next lemma to demonstrate that the search direction of the inexact Newton method, i.e. $x_k^+ - x_k$, obtained from (4) will be upper bounded by the distance function evaluated at the current iterate x_k.

Lemma 2. *Let*

$$C_1 := \left(\kappa \ell^{1+\sigma} + \left(\frac{1}{2} + \beta_1\right) L r^{1-\sigma}\right) \frac{1}{\gamma b^\sigma} + 2.$$

For all $k \in \mathbb{N}$ such that $x_k \in B(x^*, r/2)$,

$$\|x_k^+ - x_k\| \le C_1 \mathrm{d}(x_k). \tag{14}$$

Proof. Let $k \in \mathbb{N}$ such that $x_k \in B(x^*, r/2)$. Firstly, let us select $\bar{x}_k \in \mathcal{S}$ with $\mathrm{d}(x_k) = \|x_k - \bar{x}_k\|$. We then have $\nabla f(\bar{x}_k) = 0$. This fact and the Lipschitz continuity of f imply that

$$\begin{aligned}
&\nabla f(x_k) \\
=&\nabla f(x_k) - \nabla f(\bar{x}_k) \\
=&\int_0^1 \nabla^2 f(\bar{x}_k + t(x_k - \bar{x}_k))(x_k - \bar{x}_k)\mathrm{d}t \\
=&\int_0^1 \left[\nabla^2 f(\bar{x}_k + t(x_k - \bar{x}_k)) - \nabla^2 f(x_k)\right] \times \\
&\times (x_k - \bar{x}_k)\mathrm{d}t + \nabla^2 f(x_k)(x_k - \bar{x}_k).
\end{aligned} \tag{15}$$

https://doi.org/10.52111/qnjs.2022.16107

By noting that $\nabla^2 f(x_k) = (H_k+\theta_k I)-(\delta_k+\theta_k)I$, one gets

$$\begin{aligned}&(H_k+\theta_k I)^{-1}\nabla^2 f(x_k)(x_k-\bar{x}_k)\\ =&x_k-\bar{x}_k-(\delta_k+\theta_k)(H_k+\theta_k I)^{-1}(x_k-\bar{x}_k).\end{aligned} \tag{16}$$

From (4), (15), (16) and using (9), we get

$$\begin{aligned}&\|x_k^+-x_k\|\\ \le&\left\|(H_k+\theta_k I)^{-1}\left[(H_k+\theta_k I)(x_k^+-x_k)\right.\right.\\ &\left.\left.+\nabla f(x_k)\right]\right\|+\left\|(H_k+\theta_k I)^{-1}\nabla f(x_k)\right\|\\ \le&\left\|(H_k+\theta_k I)^{-1}\right\|\left(\eta_k+\frac{L}{2}\|x_k-\bar{x}_k\|^2\right)\\ &+\|x_k-\bar{x}_k\|+(\delta_k+\theta_k)\left\|(H_k+\theta_k I)^{-1}\right\|\times\\ &\times\|x_k-\bar{x}_k\|.\end{aligned} \tag{17}$$

Since $H_k\succeq 0$ and $\theta_k>0$, we have $\|(H_k+\theta_k I)^{-1}\|\le\frac{1}{\theta_k}$. Using this inequality, Lemma 1, (12) and noting that $\|x_k-\bar{x}_k\|=\mathrm{d}(x_k)\le r$, we then deduce from (17) that

$$\begin{aligned}&\|x_k^+-x_k\|\\ \le&\frac{1}{\theta_k}\left(\kappa l^{1+\sigma}\mathrm{d}(x_k)^{1+\sigma}+\frac{L}{2}\mathrm{d}(x_k)^2\right)\\ &+2\mathrm{d}(x_k)+\frac{\beta_1 L}{\theta_k}\mathrm{d}(x_k)^2\\ \le&\left(\left(\kappa\ell^{1+\sigma}+\left(\frac{1}{2}+\beta_1\right)Lr^{1-\sigma}\right)\frac{1}{\gamma b^\sigma}+2\right)\times\\ &\times\mathrm{d}(x_k),\end{aligned}$$

from which completes the proof. □

We now show that the sequence of distances from iterates generated by Algorithm 1 to the solution set $\mathcal{S}$ will decrease with a superlinear rate.

Lemma 3. *Let*

$$C_2=\tau\left(\kappa\ell^{1+\sigma}+\left(\left(\beta_1 L+\frac{L}{2}C_1\right)r^{1-\sigma}+\gamma l^\sigma\right)C_1\right).$$

For all $k\in\mathbb{N}$ such that $x_k\in B(x^*,\frac{r}{2(1+C_1)})$, we have

$$\mathrm{d}(x_k^+)\le C_2\mathrm{d}(x_k)^{1+\sigma}.$$

Proof. Let $k\in\mathbb{N}$ be such that $x_k\in B(x^*,\frac{r}{2(1+C_1)})$. Since $C_1>2$, by virtue of Lemma 2, one gets

$$\begin{aligned}\|x_k^+-x^*\|&\le\|x_k^+-x_k\|+\|x_k-x^*\|\\ &\le C_1\mathrm{d}(x_k)+\|x_k-x^*\|\\ &\le(C_1+1)\|x_k-x^*\|\\ &<\frac{r}{2}.\end{aligned}$$

This means that $x_k^+\in B(x^*,\frac{r}{2})$. Hence, local error bound condition (10) holds at x_k^+, i.e.,

$$\mathrm{d}(x_k^+)\le\tau\|\nabla f(x_k^+)\|. \tag{18}$$

With the notation $u_k^+:=x_k^+-x_k$ we deduce from (4) that

$$\begin{aligned}&\|\nabla f(x_k)+\nabla^2 f(x_k)u_k^+\|\\ \le&\left\|(H_k+\theta_k I)(u_k^+)+\nabla f(x_k)\right\|\\ &+\|(\theta_k+\delta_k)u_k^+\|\\ \le&\eta_k+(\theta_k+\delta_k)\|u_k^+\|.\end{aligned} \tag{19}$$

The differentiability of f gives us

$$\begin{aligned}&\nabla f(x_k^+)\\ =&\nabla f(x_k)+\int_0^1\nabla^2 f(x_k+tu_k^+)u_k^+\mathrm{d}t\\ =&\nabla f(x_k)+\nabla^2 f(x_k)u_k^+\\ &+\int_0^1\left[\nabla^2 f(x_k+tu_k^+)-\nabla^2 f(x_k)\right]u_k^+\mathrm{d}t.\end{aligned}$$

Taking the norm on both sides, using (9), (19) and Lemma 2, we obtain

$$\begin{aligned}&\|\nabla f(x_k^+)\|\\ \le&\eta_k+(\delta_k+\theta_k)\|x_k^+-x_k\|\\ &+\frac{L}{2}\|x_k^+-x_k\|^2\\ \le&\eta_k+(\delta_k+\theta_k)C_1\mathrm{d}(x_k)+\frac{L}{2}C_1^2\mathrm{d}(x_k)^2.\end{aligned} \tag{20}$$

Combining (18) with (20) and using (8), (13) Lemma 1, we then get

$$\begin{aligned}&\mathrm{d}(x_k^+)\\ \le&\tau\left(\kappa l^{1+\sigma}\mathrm{d}(x_k)^{1+\sigma}+(\beta_1 L\mathrm{d}(x)+\gamma l^\sigma\mathrm{d}(x)^\sigma)\times\right.\end{aligned}$$

$$\times C_1 \mathrm{d}(x_k) + \frac{L}{2} C_1^2 \mathrm{d}(x_k)^2 \Big)$$
$$\leq \tau \left(\kappa \ell^{1+\sigma} + \left(\left(\beta_1 L + \frac{L}{2} C_1 \right) r^{1-\sigma} + \gamma l^{\sigma} \right) \times \right.$$
$$\left. \times C_1 \right) \mathrm{d}(x_k)^{1+\sigma},$$

which completes the proof. □

By following the same argument as in Section 2 of ,[26] we can prove that if the sequence of iterates is close enough to the region of the solution set $\mathcal{S}$, the inner iteration algorithm will never be triggered. Moreover, Algorithm 1 will converge to some solution of the problem (1) with a superlinear rate of converence. The main result of this paper is summarized in the theorem below.

Theorem 4. *Let Assumptions 1 and 2 hold at $x^* \in \mathcal{S}$. Assume that Algorithm 1 generates an infinite sequence of iterates $\{x_k\}$. There exists $R > 0$ such that if at an iteration $k_0 \in \mathbb{N}$, $x_{k_0} \in B(x^*, R)$, then for all $k \geq k_0$, $x_{k+1} = x_k^+$, $\{x_k\}$ converges to $\widehat{x} \in \mathcal{S}$ and*

$$\lim_{k\to\infty} \frac{\|x_{k+1} - \widehat{x}\|}{\|x_k - \widehat{x}\|} = 0.$$

4. NUMERICAL EXPERIMENTS

In this section, we will make some numerical experiments to show the advantages of the our proposed algorithm. Algorithm 1 is implemented in MATLAB R2012a. Parameters of this algorithm are chosen as below: $\epsilon = 10^{-8}$, $\sigma = 0.5$, $\bar{\theta} = 0.1$, $\gamma = 10^{-2}$, $\kappa = 0.99$, $\eta_{-1} = 0.1$ $\beta_1 = 2$. The conjugate gradient method[2] will be applied to solve system (4). Because we are only interested in the local behavior of Algorithm 1, the globalization strategy is not implemented. In particular, we do not invoke Steps 4 in Algorithm 1. Instead, we will choose starting points which are sufficiently close to the optimal solution of the problem for which Theorem 4 can be applied. The investigation related to global behavior of this algorithm is out of the current work and should be the topic of another research in the future.

4.1. Superlinear convergence of Algorthm 1

This section is devoted to verify the theoretical research developed in this paper. In particular, we will show that Algorithm 1 attains the superlinear rate of convergence in some neighborhood of an optimal solution in which the local error bound condition holds true.

Let us consider the problem (1) in $\mathbb{R}^2$, where

$$f(x) = \begin{cases} \frac{1}{2}(x_2 - 1)^2 & \text{if } x_1 \in [1, 11] \\ \frac{1}{8}(x_1 - 1)^4 (x_1 - 11)^4 & \\ +\frac{1}{2}(x_2 - 1)^2 & \text{otherwise.} \end{cases} \tag{21}$$

The first and second derivatives of f are

$$\nabla f(x) = \begin{cases} \begin{pmatrix} 0 \\ x_2 - 1 \end{pmatrix} & \text{if } x_1 \in [1, 11] \\ \begin{pmatrix} (x_1 - 1)^3 \times \\ \times (x_1 - 11)^3 (x_1 - 6) \\ x_2 - 1 \end{pmatrix} & \text{otherwise.} \end{cases}$$

and

$$\nabla^2 f(x) = \begin{cases} \mathrm{diag}([0, 1]^\top) & \text{if } x_1 \in [1, 11] \\ \mathrm{diag}([(x_1 - 1)^2 (x_1 - 11)^2 \times \\ \times (7x_1^2 - 84x_1 + 227), 1]^\top), \\ & \text{otherwise.} \end{cases}$$

The function f is twice continuously differentiable and the second derivative $\nabla^2 f$ is Lipschitz continuous on $\mathbb{R}^2$. The solution set is $\mathcal{S} = [1, 10] \times \{1\}$. The local error bound condition (10) holds at any $x^* = (x_1^*, 1) \in \mathcal{S}$ such that $1 < x_1^* < 10$. Indeed,

let $r = \min\{x_1^* - 1, 10 - x_1^*\} > 0$. For all $x = (x_1, x_2) \in B(x^*, r)$, we have

$$\begin{aligned} d(x, \mathcal{S})^2 &= (x_2 - 1)^2 \\ &= \|\nabla f(x)\|^2 \end{aligned}$$

which implies that the error bound condition (10) is validated at x^* with $\tau = 1$. We note that, however, the matrix $\nabla^2 f$ is singular for all $x = (x_1, x_2)$ such that $x_1 \in [1, 10]$ which means that the SOSC does not holds in this example. From the starting point $x_0 = (9.0, -50)$, our algorithm converges to solution $\widehat{x} = (9.0, 1.0)$ after 4 iterates. The behavior of Algorithm 1 in this example is showed in Figure 1. In this figure, we plot norms of $x_k - x^*$, where $x^* = [9, 1]$ is the optimal solution. From this figure, we can see that the slope of the graph after each iteration will be more negative. This means that the sequence $\{\frac{\|x_{k+1}-x^*\|}{\|x_k-x^*\|}\}$ tends to zero. In other word, Algorithm 1 attains the superlinear convergence in this case.

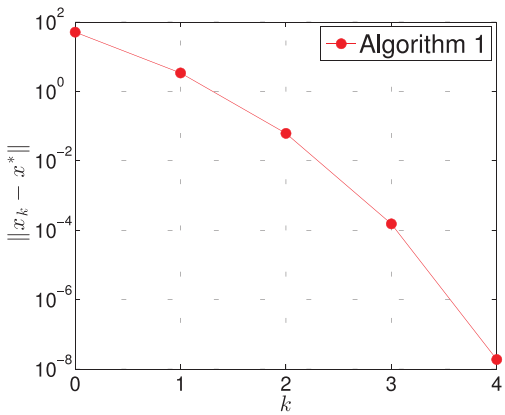

Figure 1: *Behavior of Algorithm 1 when solving* (1) *with* f *is given by* (21).

4.2. Execution time on large-scale problems

As we have mentioned in Section 1, when solving large-scale problems, an exact algorithm may take a long time to solve the linear system (3) which is the main burden in Newton method. Our aim in this section is to show the advantage of our inexact scheme when solving large-scale optimization problems. In particular, we will implement an exact version of Algorithm 1 in which the linear system 3 will be solved (exactly) instead of Step 3 of Algorithm 1. To solve the linear sytem (3), the coefficient matrix $H_k + \theta_k I$ will be factorized by an LDL decomposition, see, e.g.,.[28] Because there is no computation of the square roots of the diagonal elements is needed, this decomposition is more stable and more efficient than Gaussian elimination or Cholesky decomposition.

Problem	Size (n)	Inexact algorithm	Exact algorithm
`Arwhead`	5000	**58.02**	70.99
`bdqrtic`	5000	**2.43**	2.54
`broydn7d`	1000	**19.91**	19.95
`brybnd`	5000	**87.30**	94.86
`dqdrtic`	5000	52.08	**44.12**
`edensch`	5000	**6.48**	6.83
`engval1`	5000	88.27	**87.85**
`freuroth`	5000	**165.98**	169.00
`noncvxun`	1000	4.20	**3.44**
`penalty1`	1000	**7.65**	8.86
`sensors`	1000	**116.56**	122.10

Table 1: *Execution times (in second) of the inexact and the exact algorithms on large-scale problem in CUTEst collection (n is the number of variables)*

In this section, we will consider some large-scale unconstrained problems under form (1) in the CUTEst collection.[29] We will compare CPU times to solve each problem by the inexact and the exact algorithms. Table 1 shows us the numerical results when applying these two algorithms in solving some problems in CUTEst. We will collect problems which satisfy two requirements:

https://doi.org/10.52111/qnjs.2022.16107

the size of a problem (the number of variables) is greater than 1000, and both algorithms are succeed in solving the problem within 200s. We compare only the problems with $n \geq 1000$ because these two algorithms solve the others very fast. This makes the comparisons unfair. From this table, we can see that in most of problems, the inexact algorithm take less time to solve than the exact one. This demonstrates the benefit when using an inxact algorithm instead of the exact algorithm to solve unconstrained optimization problem.

5. CONCLUSION

In this paper, we propose an inexact regularized Newton method for unconstrained optimization. The algorithm is a variant of the algorithm in[8] where linear systems for finding search direction are solved approximately with a suitable tolerance. Assymptotic convergence analysis is performed to show that under some local error bound condition, the algorithm attains a superlinear rate of convergence. Some numerical experiments are conducted to verify theoretical results and to show the advantage of our proposed algorithm. In the future, some researches for globalizing our proposed algorithm should be interested.

Acknowledgment

The authors would like to thank the anonymous referee(s) for his/her careful reading and insightful comments. *This research is conducted within the framework of science and technology projects at institution level of Quy Nhon University under the project code T2021.699.02.*

REFERENCES

1. L. T. Biegler. *Nonlinear Programming.* Society for Industrial and Applied Mathematics, 2010.
2. J. Nocedal and S. J. Wright. *Numerical optimization*, Springer Series in Operations Research and Financial Engineering, New York: Springer, second ed., 2006.
3. C. Humes, Jr. and P. J. S. Silva. Inexact proximal point algorithms and descent methods in optimization, *Optimization and Engineering*, **2005**, *6*(2), 257–271.
4. D.-H. Li, M. Fukushima, L. Qi, and N. Yamashita. Regularized Newton methods for convex minimization problems with singular solutions, *Computational Optimization and Applications*, **2004**, *28*(2), 131–147.
5. Y.-J. Li and D.-H. Li. Truncated regularized Newton method for convex minimizations, *Computational Optimization and Applications*, **2009**, *43*(1), 119–131.
6. W. W. Hager and H. Zhang. Self-adaptive inexact proximal point methods, *Computational Optimization and Applications*, **2008**, *39*(2), 161–181.
7. M. Fuentes, J. Malick, and C. Lemaréchal. Descentwise inexact proximal algorithms for smooth optimization, *Computational Optimization and Applications*, **2012**, *53*(3), 755–769.
8. P. Armand and I. Lankoandé. An inexact proximal regularization method for unconstrained optimization, *Mathematical Methods of Operations Research*, **2017**, *85*(1), 43–59.

9. P. Armand, J. Benoist, and D. Orban. From global to local convergence of interior methods for nonlinear optimization, *Optimization Methods and Software*, **2013**, *28*(5), 1051–1080.

10. P. Armand, J. Benoist, R. Omheni, and V. Pateloup. Study of a primal-dual algorithm for equality constrained minimization, *Computational Optimization and Applications*, **2014**, *59*(3), 405–433.

11. P. Armand and R. Omheni. A globally and quadratically convergent primal-dual augmented Lagrangian algorithm for equality constrained optimization, *Optimization Methods and Software*, **2017**, *32*(1),1–21.

12. P. Armand and R. Omheni. A mixed logarithmic barrier-augmented Lagrangian method for nonlinear optimization, *Journal of Optimization Theory and Applications*, **2017**, *173*(2), 523–547.

13. P. Armand and N. N. Tran. An augmented lagrangian method for equality constrained optimization with rapid infeasibility detection capabilities, *Journal of Optimization Theory and Applications*, **2019**, *181*, 197–215.

14. S. Arreckx and D. Orban. A Regularized Factorization-Free Method for Equality-Constrained Optimization, *SIAM Journal on Optimization*, **2018**, *28*(2), 1613–1639.

15. P. Armand and N. N. Tran. Rapid infeasibility detection in a mixed logarithmic barrier-augmented lagrangian method for nonlinear optimization, *Optimization Methods and Software*, **2019**, *34*(5), 991–1013.

16. R. S. Dembo, S. C. Eisenstat, and T Steihaug. Inexact newton methods, *SIAM Journal on Numerical Analysis*, **1982**, *19*(2), 400–408.

17. R. S. Dembo and T. Steihaug. Truncated-newtono algorithms for large-scale unconstrained optimization, *Mathematical Programming*, **1983**, *26*, 190–212.

18. J.-S. Pang. Inexact newton methods for the nonlinear complementarity problem, *Mathematical Programming*, **1986**, *36*, 54–71.

19. P. Armand, J. Benoist, and J.-P. Dussault. Local path-following property of inexact interior methods in nonlinear programming, *Computational Optimization and Applications*, **2012**, *52*, 209–238.

20. K. Ueda and N. Yamashita. A regularized Newton method without line search for unconstrained optimization, *Computational Optimization and Applications*, **2014**, *59*(1-2), 321–351.

21. M. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems, *Journal of Research of the National Institute of Standards and Technology*, **1952**, *49*, 409–435.

22. C. C. Paige and M. A. Saunders. Solution of sparse indefinite systems of linear equations, *SIAM Journal on Numerical Analysis*, **1975**, *12*(4), 617–629.

23. C. C. Paige and M. A. Saunders. Lsqr: An algorithm for sparse linear equations and sparse least squares, *ACM Transactions on Mathematical Software*, **1982**, *8*, 43–71.

https://doi.org/10.52111/qnjs.2022.16107

24. Y. Saad and M. H. Schultz. Gmres: A generalized minimal residual algorithm for solving nonsymmetric linear systems, *SIAM Journal on Scientific and Statistical Computing*, **1986**, 7(3), 856–869.

25. D. C.-L. Fong and M. Saunders. Lsmr: An iterative algorithm for sparse least-squares problems, *SIAM Journal on Scientific Computing*, **2011**, *33*(5), 2950–2971.

26. P. Armand and N. N. Tran. Local convergence analysis of a primal–dual method for bound-constrained optimization without sosc, *Journal of Optimization Theory and Applications*, **2021**, *189*, 96–116.

27. I. S. Duff, MA57—a code for the solution of sparse symmetric definite and indefinite systems, *ACM Transactions on Mathematical Software*, **2004**, *30*(2), pp. 118–144.

28. G. H. Golub and C. F. van Loan. *Matrix Computations*. JHU Press, fourth ed., 2013.

29. N. I. M. Gould, D. Orban, and P. L. Toint. Cutest: a constrained and unconstrained testing environment with safe threads for mathematical optimization, *Computational Optimization and Applications*, **2015**, *60*, 545–557.

https://doi.org/10.52111/qnjs.2022.16107