Saturation-Allowed Zeroing Neural Networks activated by Various Functions for Time-varying Quadratic Programming

Xiaoyu Guo\textsuperscript{a}, Mei Liu\textsuperscript{b,c}, Qinglin Zhao\textsuperscript{b}, Bin Hu\textsuperscript{b}, Huiyan Lu\textsuperscript{b}, Long Jin\textsuperscript{b,c}

\textsuperscript{a}Department of Electrical Engineering, City University of Hong Kong, Kowloon, Hong Kong
\textsuperscript{b}School of Information Science and Engineering, Lanzhou University, Lanzhou 730000, China
\textsuperscript{c}Key Laboratory of IoT of Qinghai Province, Qinghai Normal University, Xining 810000, China

Abstract. Zeroing neural networks (ZNN) approach, has been presented to solve a lot of time-varying problems activated by monotonically increasing functions. However, the existing ZNN models for time-varying quadratic programming based on ZNN approach may be different from each other in structures, but share two common restrictions, i.e., the function must be convex and unbounded. In order to relax the above restrictions in solving time-varying quadratic programming (TVQP) problems, this paper proposes a saturation-allowed zeroing neural networks (SAZNN) model based on the ZNN approach. Comparing with existing models, the activation function (AF) of SAZNN model tolerates more kinds of functions, e.g., saturation function, non-convex function and unbounded function. Finally, this paper provides simulation results synthesized by the proposed SAZNN model activated by various AFs and verifies the superiority of the proposed SAZNN model in terms of convergence, efficiency and stability.

1. Introduction

As a fundamental optimization problem, quadratic programming (QP), is a process solving a type of mathematical optimization problem [1]. Specifically, a linearly constrained quadratic optimization problem, in fact, is to optimize a quadratic function with several variables that subject to linear constraints. Moreover, it has wide applications in many areas, such as scientific and control fields [2–7]. Note that, time is of great importance in practical application. Hence, there is greatly significant to solve time-varying quadratic programming (TVQP) problems in accurate and rapid manners. Many researchers put efforts in order to solve the above problem efficiently and accurately [8].

According to the reference [9], with time-varying parameters which are obtained from the prediction of control, the existing recurrent neural network (RNN) models can solve time-varying problems [10, 11]. To solve the aforementioned problem, there is an existing recurrent neural network (RNN) proposed to solve linear constrained QP problems [12]. This RNN model provides a systematic approach to exploiting control
techniques in a robust and accurate way for algebraic equations. However, because of some weaknesses on
the activated functions activated by existing RNN model, it is needed to remedy these weaknesses [13–15].
In addition, it is noteworthy that aforementioned neural networks are designed to deal static QP prob-
lem with time-invariant parameters. Nevertheless, those problems and parameters are time-varying.
Furthermore, time-varying QP (TVQP) problems are used in scientific researches and engineering applica-
tions such as signal processing and deep learning [16, 17]. In order to remedy the weakness of the existing
models for time-varying problem, a neural network method, termed zeroing neural networks (ZNN), is
reviewed in this article [18]. ZNN can be used to zero out each element of the error function and generalized
different dimensions.
Although, ZNN approach has some achievements in time-varying problems solving, there are still some
shortcomings in existing ZNN model [19–22]. The first constraint is that the AF of the model must be
monotonically increasing odd [23], i.e., the AF should be an unbounded function. Similarly, the fact that
the AF of the model must be a convex function can be deemed as the second constraint. This paper breaks
these constraints by proposing a new modified model based on the zeroing neural networks approach
called saturation-allowed zeroing neural networks (SAZNN) [24–26].
There are four sections in the remainder of this paper. The problem formulation and construction of
SAZNN model are presented in Section 2 with the theorem proving the convergence of the new model.
Next, various activation functions (AFs) activated by SAZNN model are shown in details. Section 4
mainly provides illustrative examples and computes the simulations of different AFs. We also discuss the
convergence velocity, the advantages and disadvantages of these results under various situations in this
section. Finally, Section 5 provides conclusions on the primary works and achievements of this paper. In
the end of the introduction part, three main contributions of this paper are presented as follows:
1) A new SAZNN model, which is of stability and accuracy, is proposed for solving TVQP problems.
2) In addition to the unbounded and convex function, the proposed SAZNN model can be activated by
saturation functions and non-convex functions, which are restrictions of the existing ZNN model.
3) The convergence speed is accelerated by the new AFs.
4) Simulation results with different AFs are used to verify the correctness.

2. Problem formulation and related work
In order to investigate quadratic programming problem with time-dependent parameters, we revisit
the previous relevant works as well as point out two restrictions on existed ZNN models in this section.

2.1. Preliminaries and problem formulation
We can get formulation from [1] as follows:

\[
\begin{align*}
\text{minimize} & \quad x^T(t)M(t)x(t)/2 + d^T(t)x(t) \\
\text{subject to} & \quad B(t)x(t) = a(t),
\end{align*}
\]

where the superscript \(^T\) denotes the transpose of a vector or a matrix, matrix \(M(t) \in \mathbb{R}^{n \times n}\), vector \(d(t) \in \mathbb{R}^n\), the full-row-rank matrix \(B(t) \in \mathbb{R}^{m \times n}\), and vector \(a(t) \in \mathbb{R}^m\) are all smoothly time-varying. We calculate
\(x(t) \in \mathbb{R}^n\) by solving (1) in real time and in a minimal error manner. The Lagrangian function \(L(x(t), p(t), t) = x^T(t)M(t)x(t)/2 + d^T(t)x(t) + p^T(t)(B(t)x(t) - a(t))\) with \(p(t) \in \mathbb{R}^m\) denoting the Lagrange-multiplier vector contributing is designed to solve problem (1) with the aid of the following equation:

\[
\Theta(t)z(t) = v(t)
\]
where

\[ \Theta(t) = \begin{bmatrix} M(t) & B^T(t) \\ B(t) & 0 \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)} \]

\[ z(t) = \begin{bmatrix} x(t) \\ p(t) \end{bmatrix} \in \mathbb{R}^{n+m}, \quad v(t) = \begin{bmatrix} -d(t) \\ a(t) \end{bmatrix} \in \mathbb{R}^{n+m}. \]

To lay a basis for further research, a saturation-allowed ZNN (SAZNN) model is proposed for solving time-varying quadratic programming (TVQP) problem (1) in the ensuing part.

2.2. Saturation-allowed ZNN model

In the previous sections, we have pointed out that there are two weaknesses in the existing ZNN model when we are solving the TVQP problem. In order to improve the performances, we propose an SAZNN model, which is based on the dynamic error analysis that requests error function converging to zero. We can obtain the following process from [1]. An error function is defined as

\[ \varphi(t) = \Theta(t)z(t) - v(t). \]  

Forcing \( \varphi(t) \) to be zero, an SAZNN design formula is presented as follows:

\[ \dot{\varphi}(t) = -\gamma \Psi_r(\varphi(t)), \]

with \( \Psi_r(V) \) denoting a mapping that is from a set \( V \) to a set \( r \), which is defined as follows:

\[ \Psi_r(V) = \arg \min_{Y \in r} \|Y - V\|_F \]

with \( 0 \in \Psi_r \). Based on the design formula (4), an SAZNN model is proposed for TVQP as follows:

\[ \Theta(t)z(t) = -\Theta(t)z(t) - \gamma \Psi_r(\Theta(t)z(t) - v(t)) + v(t). \]

It is noteworthy that the proposed SAZNN model (5) is equivalent to new design formula (4). As a result, we prove the following theorem based on the design formula (4).

**Theorem 1.** The SAZNN model (5) globally converges to the theoretical solution of TVQP problem (1).

**Proof.** According to the newly designed SAZNN model (5), the \( i \)th evolution of new design formula (4) can be written as

\[ \dot{\varphi}_i(t) = -\gamma \Psi_r(\varphi_i(t)), \]

\( \forall i \in 1, 2, ..., n \). Then, we can define a Lyapunov function candidate for (4) as follows:

\[ u_i(t) = \varphi_i(t)^2/2. \]

From the equation (7), it can be concluded that when \( \varphi_i(t) \neq 0, u_i(t) > 0 \) and when \( \varphi_i(t) = 0, u_i(t) = 0 \). Then, \( \dot{u}_i(t) \), the time derivative of \( u_i(t) \), can be obtained as

\[ \dot{u}_i(t) = -\varphi_i^2(t) \Psi_r(\varphi_i(t)). \]

According to the definition of \( \Psi_r(t) \), a formula can be obtained as follows:

\[ \|\Psi_r(\varphi_i(t)) - \varphi_i(t)\|_F^2 \leq \|Y - \varphi_i(t)\|_F^2, \forall Y \in r. \]

Selecting \( Y = 0 \) leads to \( \|\varphi_i(t)\|_F^2 \leq \|\varphi_i(t)\|_F^2, \) and we can further get

\[ 0 \leq \Psi_r^2(\varphi_i(t)) \leq 2\Psi_r(\varphi_i(t))\varphi_i(t). \]

Finally, the formula is obtained as follows:

\[ 0 \geq -\Psi_r^2(\varphi_i(t))/2 \geq \dot{u}_i(t). \]

According to the above discussions, \( \dot{u}_i(t) \) is the time derivative of \( u_i(t) \geq 0 \) and \( \dot{u}_i(t) \leq 0 \). Moreover, based on Lyapunov theory, we can also summarize that \( \varphi_i(t) \) globally converges to zero. The SAZNN model (5) is globally convergent to the theoretical solution of TVQP problem (1). Consequently, the proof is well accomplished.
3. Activation Function and Theoretical Analysis

From the above analyses, we can summarize that $\Psi_r(\cdot)$ includes some special cases of the SAZNN model. Furthermore, there are still other situations that we are able to show to display superiority compared other models. Some special cases can be used in SAZNN model as activation functions (AF) are provided as:

- The linear AF:
  $$\Psi_r(\phi_i(t)) = \phi_i(t).$$  \hfill (12)

- The powersigmoid AF:
  $$\Psi_r(\phi_i(t)) = \frac{1 + \exp(-4) \cdot (1 - \exp(-4\phi_i(t)))}{1 - \exp(-4) \cdot (1 + \exp(-4\phi_i(t)))}$$  \hfill (13)

- The exponent AF:
  $$\Psi_r(\phi_i(t)) = \exp(10\phi_i(t)) - \exp(10\phi_i(t))$$  \hfill (14)

- The bound AF, $\Psi_r = \{V \in \mathbb{R}^n, a^- \leq V \leq a^+\}$, $a^- < 0$ and $a^+ > 0$:
  $$\Psi_r(V_i) = \begin{cases} a_i^+, & V_i > a_i^+; \\ V_i, & a_i^- \leq V_i \leq a_i^+; \\ a_i^-, & V_i < a_i^-; \end{cases}$$  \hfill (15)

- The boundjp AF, $\Psi_r = \{V \in \mathbb{R}^n, a^- \leq V \leq a^+\}$, $a^- < 0$ and $a^+ > 0$:
  $$\Psi_r(V_i) = \begin{cases} a_i^+, & V_i > a_i^+; \\ 10V_i, & a_i^- \leq V_i \leq a_i^+; \\ a_i^-, & V_i < a_i^-; \end{cases}$$  \hfill (16)
Figure 2: Residual errors of SAZNN model (5) for solving TVQP problem activated by different AFs with two values of $\gamma$. (a) with $\gamma = 1$ and ball AF. (b) with $\gamma = 10$ and ball AF. (c) with $\gamma = 1$ and bound AF. (d) with $\gamma = 10$ and bound AF. (e) with $\gamma = 1$ and generalized step AF. (f) with $\gamma = 10$ and generalized step AF.
• The ball AF, $\Psi_r = \{V \in \mathbb{R}^n, \|V\|_F \leq R_0\}$, and $R_0 > 0$:

$$
\Psi_r(V) = \begin{cases} 
V, & \|V\|_F \leq R_0, \\
R_0/\|V\|_F, & \|V\|_F > R_0.
\end{cases}
$$

(17)

• The generalized step AF:

$$
\Psi_r(\varphi_i(t)) = \begin{cases} 
c, & \varphi_i(t) > \delta^+,
\varphi_i^2(t) + 10\varphi_i(t), & \delta^- < \varphi_i(t) < \delta^+,
-c, & \varphi_i(t) < \delta^-,
\end{cases}
$$

(18)

where $c > 0, \delta^+ > 0, \delta^- < 0$ and the last two has a tiny value.

**Remark 1.** When $\delta^+$ and $\delta^-$ both have tiny values, the range $[\delta^-, \delta^+]$ is just as an integral domain in the generalized step function, and thus, the result of $\Psi_r(\varphi_i(t))$ is equal to $c$ or $-c$. Therefore, we can conclude that $\varphi_i(t)$ is a constant value. By solving the first-order differential equation $\dot{\varphi}(t) = -\gamma\Psi_r(\varphi(t))$, the convergence time can be obtained.
4. Illustrative Examples

In previous sections, various kinds of AFs are presented with saturated properties. Therefore, we respectively simulate every AF and then verify these saturated properties of SAZNN model (5) in this section. Moreover, we also compare superiorities and inferiorities of AFs under different convergence conditions. The matrix $\Theta$ and vector $V$ is shown as:

$$\Theta(t) = \begin{bmatrix} 0.5s(t) + 2 & c(t) & s(4t) \\ c(t) & 0.5c(t) + 2 & c(4t) \\ s(4t) & c(4t) & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 3},$$

(19)

$$v(t) = \begin{bmatrix} -s(3t) & -c(3t) & c(2t) \end{bmatrix}^T \in \mathbb{R}^3,$$

(20)

where $s(\cdot)$ and $c(\cdot)$ denote $\sin(\cdot)$ and $\cos(\cdot)$ respectively.

Considering that the existing ZNN model dissatisfies the non-convex and saturation AFs, SAZNN model (5) is proposed to remedy these qualities. As illustrated in the above figures, Fig. 1 presents the simulation results of SAZNN model (5) with different AFs include bound LP AF, linear AF and powersigmoid AF. In Fig. 1, the convergence time of residual error simulated in two situations with $\gamma = 1$ and $\gamma = 10$, which is infinite in this case. The condition of convergence in Fig. 1(b) is similar as in Fig. 1(a). As shown in Fig.
1(a), with \( \gamma = 1 \), all the residual errors \( \| \psi_i(t) \|_2 \) converge to zero in finite time with three AFs. It is obviously shown that the time for convergence of boundp AF is probably 0.5 s, and those of powersigmoid AF and linear AF respectively about 3 s and 6 s. We can find that the convergence time of linear AF is the longest in them which is about 12 times larger than of boundp AF. As a result, compared with ZNN model, the excellent quality of SAZNN model (5) offers good containment including non-convex function, unbounded function and saturation function. Then, Fig. 2 shows three AFs belonging to the above functions. Each of AFs is in two conditions with \( \gamma = 1 \) and \( \gamma = 10 \). In Fig. 2(a), with \( \gamma = 1 \), residual error of SAZNN model (5) of ball AF converges to zero in nearly 5 s but in the case of \( \gamma = 10 \), residual error of SAZNN model activated by ball AF converges to zero in 0.5 s presented in Fig. 2(b). Furthermore, Fig. 2(c), in which SAZNN model (5) is activated by bound AF, shows that convergence time to zero is about 5 s. Compared to the latest figure, residual error of SAZNN model (5) with bound AF converges rapidly to zero in 0.5 s presented by Fig. 2(d). In Fig. 2(e) and Fig. 2(f), obviously, they clearly illustrate that the convergence time to zero is very tiny because the error with step AF converges rapidly from the initial value to zero.

As described previously, the different AFs and values of \( \gamma \) are believed to have different simulations of convergence situations. From the theorem (1), AFs in the SAZNN model (5) are allowed to be non-convex and Fig. 3(a) is the demonstration for that. Furthermore, Fig. 3(b) shows that \( \dot{\psi}(t) = -1 \) with random initial state less than -1 and converge to zero in about 0.4 s illustrating the non-convex quality of SAZNN model (5). Moreover, Fig. 4 and Fig. 5 respectively show the convergence situations of residual error and initial state vectors \( \dot{\psi}(t) \) which have similar characteristics with the Fig. 3.

**Remark 2.** Generally speaking, the step function can be used to construct the activation function for accelerating the convergence of ZNN model, which is defined as

\[
\Psi_s(\psi_i(t)) = \begin{cases} 
1, & \psi_i(t) > 0, \\
0, & \psi_i(t) = 0, \\
-1, & \psi_i(t) < 0.
\end{cases}
\]

The above equation is not accomplished because when \( \psi_i(t) = 0 \), it contravenes the definition of \( \Psi_s(\cdot) \). Due to the existing of round-off errors and truncation errors, residual error of SAZNN model (5) will oscillate strongly near zero, if we simply have \( \Psi_s(\psi_i(t)) = 0 \) when \( \psi_i(t) = 0 \). Finally, we expand zero input action into a small range \([-0.1, 0.1]\). As shown in the following figures, Fig. 6(a) and Fig. 6(b) give simulations of residual errors of the SAZNN model when AF is the step function. Obviously, when the residual error converges to zero, we are able to recognize that oscillation occurs and continues to oscillate in the figure period.

5. Conclusions

In this paper, an saturation-allowed ZNN (SAZNN) model (5) has been proposed and activated by several kinds of AFs, e.g., saturation function, non-convex function and unbounded function. Therefore, comparing with existing ZNN model, the SAZNN model (5) is verified to solve the TVQP problem activated by the above AFs. Moreover, we illustrate an example and make comparisons for different AFs activated the SAZNN model (5) through the simulation results of residual errors.

**References**

[1] L. Jin, Y. Zhang, S. Li, and Y. Zhang, Modified ZNN for time-varying quadratic programming with inherent tolerance to noises and its applications to kinematic redundancy resolution of robot manipulators, IEEE Trans. Ind. Electron, vol. 63, no. 11, pp. 6978–6988, Nov. 2016.

[2] L. Jin, Y. Zhang, Continuous and discrete Zhang dynamics for real-time varying nonlinear optimization, Numer. Algor., vol. 73, no. 1, pp. 115–140, 2016.

[3] L. Jin, B. Liao, M. Liu, L. Xiao, D. Guo, and X. Yan, Different-level simultaneous minimization scheme for fault tolerance of redundant manipulator aided with discrete-time recurrent neural network, Frontiers in Neurorobotics., In Press with DOI 10.3389/fnbot.2017.00050, 2017.

[4] L. Jin, Y. Zhang, Discrete-time Zhang neural network for online time-varying nonlinear optimization with application to manipulator motion generation, IEEE Trans. Neural Netw. Learn. Syst., vol. 26, no. 7, pp. 1525–1531, Jul. 2015.
[5] L. Jin, Y. Zhang, G2-type SRMPC scheme for synchronous manipulation of two redundant robot arms, IEEE Trans. Cybern., vol. 45, no. 2, pp. 153–164, 2015.

[6] L. Jin, S. Li, Distributed task allocation of multiple robots: A control perspective, IEEE Trans. Syst., Man, Cybern., Syst., In Press with DOI 10.1109/TSMC.2016.2627579, 2016.

[7] L. Jin, S. Li, J. Yu, and J. He, Robot manipulator control using neural networks: A survey, Neurocomputing, In Press with DOI 10.1007/978-1-4471-3244-8-8, 2018.

[8] M. Forti, A. Tesi, New conditions for global stability of neural networks with application to linear and quadratic programming problems, IEEE Trans. Circuits Syst. I, vol. 42, pp. 354–366, Feb. 1995.

[9] L. Jin, S. Li, and B. Hu, RNN models for dynamic matrix inversion: a control-theoretical perspective, IEEE Trans. Ind. Inf., vol. 14, no. 1, pp. 189–199, 2018.

[10] L. Jin, S. Li, H. Wang, and Z. Zhang, Nonconvex projection activated zeroing neurodynamic models for time-varying matrix pseudoinversion with accelerated finite-time convergence, Appl. Soft Comput. (2018).

[11] L. Jin, Y. Zhang, and S. Li, Integration-enhanced Zhang neural network for real-time-varying matrix inversion in the presence of various kinds of noises, IEEE Trans. Neural Netw. Learn. Syst., vol. 27, no. 12, pp. 2615–2627, Dec. 2016.

[12] Y. Zhang, J. Wang, Recurrent neural networks for nonlinear output regulation, Automatica, 37 (8) (2001), pp. 1161–1173.

[13] Y. Liu, J. Li, S. Tong, and C. L. P. Chen, Neural network control-based adaptive learning design for nonlinear systems with full state constraints, IEEE Trans. Neural Netw. Learn. Syst., vol. 27, no. 7, pp. 1562–1571, Jul. 2016.

[14] F. A. Pazos, A. Bhaya, Control Lyapunov function design of neural networks that solve convex optimization and variational inequality problems, Neurocomputing, vol. 72, no. 16–18, pp. 3863–3872, Oct. 2009.

[15] Srivastava, Hari M., F. Muge Sakar Sakar, and H. Ozlem Guney. Some general coefficient estimates for a new class of analytic and Bi-Univalent functions defined by a linear combination, Filomat 32.4 (2018).

[16] Y. Zhang, Y. Wang, L. Jin, B. Mu, and H. Zheng, Different ZFs leading to various ZNN models illustrated via online solution of time-varying underdetermined systems of linear equations with robotic application, Lecture Notes Comput. Sci., vol. 7952, pp. 481–488, 2013.

[17] Z. Zhang, S. Li, and X. Zhang, Simulink comparison of varying-parameter convergent-differential neural-network and gradient neural network for solving online linear time-varying equations, Proceedings of the 12th World Congress on Intelligent Control and Automation, IEEE (2016), pp. 887–894.

[18] L. Jin, S. Li, B. Liao, and Z. Zhang, Zeroing neural networks: A survey, Neurocomputing, vol. 267, pp. 597–604, 2017.

[19] Y. Xia, J. Wang, A recurrent neural network for solving nonlinear convex programs subject to linear constraints, IEEE Trans. Neural Netw., vol. 16, no. 2, pp. 379–386, Mar. 2005.

[20] L. Jin, S. Li, H.M. La, and X. Luo, Manipulability optimization of redundant manipulators using dynamic neural networks, IEEE Trans. Ind. Electron., 64 (6) (2017), pp. 4710–4720.

[21] L. Jin, Y. Zhang, S. Li, and Y. Zhang, Noise-tolerant ZNN models for solving time-varying zero-finding problems: a control-theoretic approach, IEEE Trans. Autom. Control, 62 (2) (2017), pp. 992–997.

[22] B. Liao, Y. Zhang, and L. Jin, Taylor ($\phi^3$) discretization of ZNN models for dynamic equality-constrained quadratic programming with application to manipulators, IEEE Trans. Neural Netw. Learn. Syst., vol. 27, no. 2, pp. 225–237, 2016.

[23] Y. Zhang, Z. Li, Zhang neural network for online solution of timevarying convex quadratic program subject to time-varying linear-equality constraints, Phys. Lett. A, vol. 373, nos. 18/19, pp. 1639–1643, 2009.

[24] P. Zhao, Y. Shen, Y. Huang, and Y. Wang, Solving time-varying quadratic programs based on finite-time Zhang neural networks and their application to robot tracking, Neural Comput. Appl., vol. 26, no. 3, pp. 683–703, Apr. 2015.

[25] L. Jin, A note on successive coefficients of spirallike functions, Filomat 32.4 (2018).

[26] Bataineh, Ahmad. Bernstein polynomials method and it’s error analysis for solving nonlinear problems in the calculus of variations: convergence analysis via residual function, Filomat 32.4 (2018).