

Cancer Treatment Precision Strategies Through Optimal Control Theory

Ahmed J. Abougarair ^{1*}, Abdulhamid A. Oun ², Salah I. Sawan ³, T. Abougard ⁴, H. Maghfiroh ⁵

^{1,2,3,4}Electrical and Electronics Engineering, University of Tripoli, Tripoli, Libya

⁵Department of Electrical Engineering, Universitas Sebelas Maret, Surakarta, Indonesia

Email: ¹ a.abougarair@uot.edu.ly, ² a.oun@uot.edu.ly, ³ s.sawan@uot.edu.ly, ⁵ hari.maghfiroh@gmail.com

*Corresponding Author

Abstract—Lung cancer is a highly heterogeneous disease, with diverse genetic, molecular, and cellular drivers that can vary significantly between individual patients and even within a single tumor. Though combination therapy is becoming more common in the treatment of cancer, it can be challenging to predict how various treatment modalities will interact and what negative effects they may have on a patient's health, such as increased gastrointestinal toxicities, or neurological problems. This paper aims to regulate immunity to tumor therapy by utilizing optimal control theory (OCT). This research suggests a malignant tumor model that can be regulated with a combination of immunological, vaccine, and chemotherapeutic therapy. The optimal control variables are employed to support the best possible treatment plan with the fewest potential side effects by reducing the production of new tumor cells and keeping the number of normal cells above the average carrying capacity. Also, the study addresses patient heterogeneity, individual variations in tumor biology, and immune responses for both young and old cancer patients. Finding the right doses for a treatment that works is the main goal. To do this, we conducted a comparative analysis of two optimum control approaches: the Single Network Adaptive Critic (SNAC) approach, which directly applies the notion of reinforcement learning to the essential conditions for optimality and the Linear Quadratic Regulator (LQR) methodology. Although the study's results show the promise of precision treatment plans, a number of significant obstacles must be overcome before these tactics can be successfully applied in clinical settings. It will be necessary to make considerable adjustments to the healthcare system's infrastructure in order to successfully offer personalized treatment regimens. This includes enhanced interdisciplinary care coordination methods, safe data management systems.

Keywords—Cancer Treatment; SNAC; NCO; LQR; OCP; OPT.

I. INTRODUCTION

Cancer is a group of diseases characterized by abnormal cell growth and division, often leading to death worldwide. Cancer develops when the body's natural control mechanisms malfunction, allowing cells to proliferate uncontrollably instead of undergoing programmed cell death [1]. This unregulated cell growth can result in the formation of a mass of abnormal cells, known as a tumor.

Cancer is a diverse disease with many subtypes, including lung, prostate, breast, and colorectal cancer [2]. A hallmark of cancer cells is their ability to proliferate uncontrollably, disregard normal boundaries, and invade neighboring organs, leading to the spread of the disease. The primary goal of

cancer treatment is to eliminate all cancer cells from the body while minimizing harm to healthy cells [3][4]. The primary and common treatments for cancer therapy include surgical procedures, hormone therapy, radiation therapy, organic therapy, and chemotherapy. These treatments are often combined in various ways to enhance the likelihood of tumor regression [5]-[7]. However, it is crucial to consider both the synergistic and antagonistic effects of these treatments when combined to ensure optimal outcomes.

Chemotherapy is a fundamental part of cancer treatment, using potent chemical agents designed to target and eliminate rapidly dividing cells, a hallmark of cancer cells [6][8]. However, it's essential to consider the potential side effects that chemotherapy may bring about, as their severity can vary from mild and manageable to significant complications.

Immunotherapies are increasingly becoming a crucial element of comprehensive strategies to treat specific cancer types. Immunotherapy aims to enhance the body's natural ability to combat cancer by boosting the effectiveness of the immune system. Additionally, it's worth noting that individuals with compromised immune systems, such as those with AIDS, are more susceptible to certain rare types of cancer [9]-[11].

Here are some of the key specific gaps and challenges in current cancer traditional treatment approaches [12]-[20]:

- Cancer is an extremely heterogeneous disease, tumors can evolve and adapt over time, developing resistance to various treatments through genetic and epigenetic changes. Current treatment strategies often fail to effectively manage this tumor heterogeneity and adaptability, leading to disease relapse and progression.
- Most cancer treatments are still based on broad, one-size-fits-all approaches, rather than being tailored to individual patient and tumor characteristics. Challenges remain in accurately interpreting complex multi-omic data and translating it into optimal, personalized treatment decisions.
- Current treatment schedules and sequences are often based on empirical or historical data, rather than being dynamically optimized based on real-time tumor response and patient-specific factors.
- While the number of targeted therapies and immunotherapies has increased, they are often limited in



availability and accessibility, particularly for patients in resource-constrained settings.

- The ability to effectively integrate and leverage the vast and diverse datasets (e.g., clinical, genomic, imaging, pharmacological) that can inform personalized cancer treatment decisions remains a significant challenge.

The complexity of cancer dynamics and the unique immune-tumor interactions in individual patients make this a significant problem. Current treatments often lack the precision necessary for optimal effectiveness, necessitating a more refined approach [21]-[24]. Several mathematical models for predicting tumor growth and halting disease progression during therapy are being developed and improved as we speak. This may be achieved by reducing the number of cancer cells, cutting down on medication dosages, and reducing side effects by using optimization techniques, where determining the best way to administer treatment is thought of as an OCP of a dynamic system [25]-[28]. The paper addresses the critical challenge of optimizing cancer treatment protocols using chemotherapy and immunotherapy based on OCT.

OCT provides a mathematical framework to formulate personalized treatment strategies that can adapt to the unique characteristics and dynamic responses of each patient's tumor. By modeling the complex tumor growth dynamics and the effects of various treatment modalities, optimal control can help identify the most effective and individualized treatment plans. This adaptive and personalized approach is particularly crucial in addressing the challenge of tumor heterogeneity and the need for precision medicine in cancer care [29]-[31]. OCT enables the dynamic optimization of treatment schedules, dosing, and sequencing based on real-time tumor response and patient-specific factors. This is a significant advancement over the current empirical or standardized treatment approaches, which often fail to account for the temporal and adaptive nature of tumor growth and treatment response. By optimizing the timing, intensity, and combinations of various therapies, optimal control can help maximize the therapeutic efficacy while minimizing adverse side effects and toxicities. OPT and AI provides a flexible and comprehensive framework to integrate diverse data sources, including clinical, genomic, imaging, and pharmacological information, to guide personalized treatment decisions [32]. OCT is grounded in rigorous mathematical and computational principles, allowing for systematic optimization of treatment strategies. This systematic approach can help identify the most effective treatment plans by exploring a wide range of possible interventions and their consequences, ultimately leading to more robust and reliable treatment recommendations [33]-[35].

OCT has the potential for scalable implementation and integration into clinical workflows. This scalability and ease of integration can help address the challenges of transitioning novel computational approaches from research to routine clinical practice [36]-[38].

The paper is organized as follows: section two provides highlights of previous research in combating cancer. Section three introduces the studied model, elaborating on its key

parameters and terms. Section four focuses on the mathematical foundations of OCT and presents the necessary conditions for optimality. It further discusses each of the proposed techniques, providing their mathematical formulations for solving the OCP of the quadratic regulator type. Section five delves into the manipulation of the cancer model and the formulation of the OCP. It presents the proposed solutions for each case study, including both continuous and dosed approaches, utilizing all the proposed techniques. The section concludes with a comparison of the results obtained from these solutions.

II. LITERATURE REVIEW

The application of control engineering principles within the realm of cancer treatment has garnered significant attention in academic literature [39]-[43]. This attention has spurred the exploration of various control methodologies tailored to cancer models [44] [45]. Several studies, such as those referenced as [30][31], have adopted the model initially introduced in [45] as their foundation. These studies have predominantly focused on optimal control treatment through chemotherapy [46]. They have harnessed the capabilities of a well-established nonlinear robust state feedback technique known as SDRE (State-Dependent Riccati Equation) and extended the optimal control synthesis to encompass non-measurable states by employing the extended Kalman filter (EKF) [47][48]. These investigations have yielded promising mathematical outcomes [49].

However, it is noteworthy that the treatment protocols proposed in [30][39] were continuous in nature, lacking the discrete dosing format that would enhance their practical applicability [50]. This limitation suggests that further refinement is necessary for real-world implementation. In [51], another approach to optimal control therapy for the model originally outlined in [45] was presented. This approach introduced a delay parameter in the differential equation to account for the time required for immune cells to be stimulated by tumor cells. In addition to chemotherapy [52], this study also integrated immunotherapy treatment, specifically involving the injection of tumor-activated CD8+ T cells [53]. However, in the process of addressing the OCP, the author resorted to an open-loop method akin to the variational of extremes (shooting method) [54]. Regrettably, this choice resulted in outcomes that were unsatisfactory and challenging to implement practically [55][56].

Moving on to [57], the author adopted a fractional-order rendition of the model originally delineated in [45]. In contrast to the previous studies, the focus here was solely on immunotherapy treatment, employing adaptive sliding mode control [58][59]. These various studies collectively underscore the diverse array of approaches undertaken in the realm of optimal control therapy for cancer models. They integrate different treatment modalities and control methods, showcasing the multifaceted nature of this field [60]. However, it is important to acknowledge that there remain limitations and challenges related to the practical applicability and effectiveness of the proposed treatment protocols [61]. Addressing these limitations will be crucial for advancing the field and ultimately translating these methods into meaningful clinical practices [62].

Moving forward, researchers in [63] introduced a novel framework called control theory for therapy design (CT4TD) [64]. This framework leverages OCT and applies it to patient-specific models of pharmacokinetics (PK) and pharmacodynamics (PD) to create optimized therapeutic strategies for individual patients [65]. CT4TD takes into account the physiological variations among individuals and tailors therapy accordingly [66]. It utilizes the efficient dCRAB/RedCRAB optimization algorithm and has undergone extensive testing on synthetic data [67]. Its application to Imatinib administration in Chronic Myeloid Leukemia demonstrated diverse and improved therapeutic strategies among patients compared to standard regimens [68].

Furthermore, in [69], the authors explored optimal control in Metastatic Castrate Resistant Prostate Cancer [70]. This study employed evolutionary game theory to model the dynamics of patients with metastatic castrate-resistant prostate cancer (mCRPC) undergoing abiraterone therapy [71]. Utilizing an optimal control theory approach, the study identified enhanced treatment schedules capable of minimizing or eradicating resistant cancer cell subpopulations, thereby preserving the effectiveness of abiraterone and enhancing the quality of life for patients [72][73]. In 2018, researchers in reference [74][75] developed a Computer-Aided Diagnosis system inspired by AI to detect various lung diseases from chest X-ray images [76]-[79]. In 2021, a study (reference [80][81]) introduced the use of Bayesian Machine Learning algorithms for accurate breast cancer prediction. Similarly, another study [82] developed a diagnostic system for breast cancer that incorporates computer assistance and achieved promising results with an AUC (Area Under the Curve) value of 0.836.

Additionally, in the same year, employed deep learning techniques to detect malignant cervical cells and demonstrated the model's effectiveness in categorizing various cell morphologies [83]. Their adaptable model, called DGCA region-oriented CNN, improved the classification accuracy, achieving an AUC of 0.670. Notable advantages of this model included scalability, the ability to identify cells with different morphologies, and integration of spatial context details. It achieved peak precision with an AUC of 0.789 when distinguishing between high-grade and low-grade central nervous system tumors [84]. These findings open up possibilities for utilizing AI in lesion identification and classification, guiding treatment strategies, and monitoring responses through image-guidance in oncology [85]. In 2022, a research paper [22] showcased the role of AI in diagnosing prostate cancer by employing an ensemble model with remarkable predictive accuracy [86]. The study emphasized the significance of radiomics in non-invasive diagnostics. In the same year, the same reference [22] highlighted AI's potential in colorectal cancer screening, diagnosis, and prognosis [87-89]. They emphasized advancements in risk stratification, therapy response prediction, and survival outcome assessment. The study underscored AI-driven precision in screening, classification of colorectal cancer subtypes, and prediction of distant hepatic metastases [90]. Various AI tools, including support vector machines (SVM), random forest, artificial neural

networks (ANN), convolutional neural networks (CNN), and k-nearest neighbors (k-NN), were explored for predicting complete pathological responses to therapy, emphasizing the significance of prognostic models in enhancing staging accuracy and guiding colorectal cancer treatment [91]. In 2022, [23] introduced the structure of multimodal integration (MMI) for molecular intelligent diagnostics, surpassing conventional methods [92]. They outlined emerging applications of AI in predicting mutational and molecular profiles in prevalent cancers, covering radiology and histology imaging [93-96].

However, the study acknowledged obstacles in practical AI implementation in the medical field, including challenges related to data organization, feature integration, model interpretation, and compliance with practice regulations [97]-[100]. In a recent study by the authors [101], a novel approach to cancer treatment is introduced, focusing on adaptive drug therapies, which have demonstrated superior efficiency when compared to traditional continuous maximum tolerated dose (MTD) methods [96]. This adaptive strategy tailors drug dosages according to the evolving state of the tumor, rather than adhering to a fixed treatment schedule. The innovation of this method lies in its systematic optimization of adaptive policies, achieved through the utilization of an evolutionary game theory model of cancer dynamics combined with dynamic programming [98]. Specifically, the study optimizes for two primary objectives: the reduction of total drug usage and the acceleration of the time to recovery [83]. This optimization process is accomplished by solving the Hamilton-Jacobi-Bellman equation [94]. The research findings reveal that in comparison to MTD-based strategies, these optimized adaptive treatments significantly reduce the quantity of drugs administered while simultaneously increasing the likelihood of recovery across a spectrum of initial tumor conditions [101]. For further details and a comprehensive summary of various cancer treatment research studies, please refer to Table I.

The limitations of treating cancer using the algorithms which presented in the literature review include the following:

Continuous Dosing Format: Many proposed treatment protocols are continuous in nature, which can be challenging to implement practically in real-world settings where discrete dosing is often preferred for better patient adherence and management of side effects [60].

Open-Loop Methods: Some studies have used open-loop methods, such as the shooting method, which can result in unsatisfactory and difficult-to-implement outcomes [55][56].

Practical Applicability: The proposed treatment protocols may not be directly applicable to real-world scenarios due to the complexity of cancer progression and treatment responses [61].

Clinical Translation: The translation of AI-driven diagnostic systems and optimal control methods into clinical practice can be challenging due to the need for extensive testing and validation [62].

Physiological Variations: Patient-specific models of pharmacokinetics and pharmacodynamics can be complex and require consideration of physiological variations among individuals [66].

Optimization Objectives: Optimizing for multiple objectives, such as reducing total drug usage and accelerating time to recovery, can be challenging and may require trade-offs between these objectives [83].

Model Interpretation: AI models can be difficult to interpret, making it challenging to understand the underlying mechanisms and decision-making processes [97].

Model Complexity: Complex models, such as those incorporating evolutionary game theory and dynamic

programming, can be difficult to implement and require significant computational [98].

Data Organization and Integration: AI-driven diagnostic systems face challenges related to data organization, feature integration, model interpretation, and compliance with practice regulations [97]-[100].

To address the challenges and limitations in the application of control engineering principles to cancer treatment, more robust optimization techniques are needed by leveraging advanced optimization techniques. Also, rigorous validation studies, preferably conducted using clinical data, are essential to assess the robustness and generalizability of these methods across different patient populations and cancer types.

TABLE I. SUMMARY OF CANCER TREATMENT RESEARCH STUDIES

| Reference | Technique Used | Description |
|--------------|--|---|
| [30, 31, 39] | SDRE, EKF for non-measurable states | Used the model from [45] with chemotherapy only; extended optimal control synthesis with EKF, results mathematically promising but not practical due to continuous treatment protocol. |
| [51] | Open loop method, variational of extremes | Added delay parameter to model of [45], used chemotherapy and immunotherapy; results were not satisfying nor practical. |
| [57] | Fractional order model, adaptive sliding mode | Used a fractional order version of the model in [45] focusing on immunotherapy treatment. |
| [63] | CT4TD framework, dCRAB/RedCRAB optimization | Developed CT4TD framework for therapy design using patient-specific PK and PD models; focused on Imatinib in Chronic Myeloid leukemia, resulted in diversified and improved strategies. |
| [69] | OCT | Used evolutionary game theory for mCRPC with abiraterone therapy, identified better treatment schedules for minimizing resistant cancer cells. |
| [75] | Control theory, machine learning methods | Applied control theory and machine learning, like reinforcement learning, for designing cancer treatment schedules using various computational models. |
| [81] | Berberine: A novel therapeutic strategy for cancer | Investigated how BBR's anticancer properties are regulated by several molecular mechanisms. |
| [82] | OCT | Discussed optimal control theory in radiation and systemic therapy; emphasized integrating patient-specific mathematical models for improving outcomes. |
| [87] | Evolutionary game theory, optimal control theory | Studied treating mCRPC with abiraterone; proposed a controlled strategy for abiraterone use to maintain effectiveness and improve patient quality of life. |
| [88] | Pontryagin's principle | Developed a model for breast cancer treatment combining chemotherapy and a ketogenic diet; focused on the effectiveness of treatment combinations. |
| [93] | Bock's direct multiple shooting method | Analyzed four chemotherapy models; emphasized the potential of optimally controlled therapy in altering tumor outcomes and potential benefits of optimizing chemotherapy schedules. |
| [94] | OCT | Outlined the use of optimal control theory in radiation and systemic cancer therapies; focused on personalizing therapy plans and integrating patient-specific models. |
| [100] | Evolutionary game theory, dynamic programming | Described an adaptive drug therapy approach using evolutionary game theory and dynamic programming; focused on optimizing drug dosages and reducing total drug usage. |
| [101] | Adaptive therapy Based on Darwinian evolution theory | Reviewed adaptive therapy in cancer treatment, focusing on overcoming drug resistance by integrating evolutionary dynamics into treatment regimens. |

III. SYSTEM MODEL

Mathematical modeling is a crucial tool for understanding biological processes, especially in the context of disease progression [102]. These models use mathematical equations and analysis to unravel the complex dynamics during different stages of diseases. Specifically, they are instrumental in exploring the interactions and temporal changes between cancer tumor growth, the immune system, and treatment methods like chemotherapy and

immunotherapy. These models shed light on both immediate and long-term impacts of treatments on disease progression.

While our grasp of how the immune system combats cancer is not yet complete, mathematical models using empirical data can significantly enhance our understanding. These models depict the interaction between tumors and the immune system, offering valuable insights [103]. Notably, the model developed by de Pillis has gained prominence for its comprehensive approach and is a focal point in this study [104]. This model includes:

- Immune Response: It incorporates the role of immune cells, particularly $CD8^+$ and NK T cells, which proliferate in response to tumor cells and can destroy them.
- Competition Terms: The model addresses the competition for resources between NK cells and tumor cells, and a predator-prey dynamic between $CD8^+$ immune cells and tumor cells.
- Chemotherapy & Immunotherapy: It considers the impact of a well-known chemotherapy drug, doxorubicin, on all cells in the model, as well as the effects of injected tumor-activated $CD8^+$ T cells on immune cells.

Subsequent studies have expanded on de Pillis's model, adding components like IL-2 cytokines and other lymphocyte cells, thus broadening the scope beyond NK and $CD8^+$ T cells [105]. However, due to the increased complexity of these expanded models, this study opts to use the original model described by de Pillis [106]-[108]. This chosen model is a fourth-order ordinary differential equation (ODE), encompassing four states and two controls, which are critical in understanding the interactions and effects in the context of cancer treatment and immune response [106]-[109].

States:

$E(t)$ represent the $CD8^+T$ cells.

$T(t)$ represent the tumor cells.

$N(t)$ represent the Natural Killer (NK) cells.

$M(t)$ Depict the level or amount of the chemotherapy drug present.

Controls:

$w(t)$: represent the injected tumor-activated $CD8^+T$ cells .

$v(t)$: represent the injected doxorubicin drug.

The mathematical model under discussion is detailed in source [45]. This model is crucial for understanding the dynamics and interactions in the specified context, and its formulation is comprehensively laid out in this reference. For a more concise overview, the key parameters of the model are summarized in Table II. This table serves as a quick reference guide, offering a clear description of each parameter included in the model. Such a summary is useful for readers who need to grasp the essentials of the model without delving into the more complex details found in the full text of the references [110][111].

$$\left. \begin{aligned} \dot{E}(t) &= s + \frac{\rho \cdot E(t) \cdot T(t)}{\alpha + T(t)} - c_1 \cdot E(t) \cdot T(t) \\ &\quad - d_1 \cdot E(t) - a_1 \cdot (1 - e^{-M(t)}) \cdot E(t) + w(t) \\ \dot{T}(t) &= r_1 \cdot T(t) \cdot (1 - b_1 \cdot T(t)) - c_2 \cdot E(t) \cdot T(t) \\ &\quad - c_3 \cdot T(t) \cdot N(t) - a_2 \cdot (1 - e^{-M(t)}) \cdot T(t) \\ \dot{N}(t) &= r_2 \cdot N(t) \cdot (1 - b_2 \cdot N(t)) - c_4 \cdot T(t) \cdot N(t) \\ &\quad - a_3 \cdot (1 - e^{-M(t)}) \cdot N(t) \\ \dot{M}(t) &= v(t) - d_2 \cdot M(t) \end{aligned} \right\} \quad (1)$$

IV. OPTIMAL CONTROL THEORY (OCT)

OCT is an advanced field within control theory that focuses on finding the best possible trajectories for control

and state variables in a system. The primary goal of OCT is to either minimize or maximize a certain criterion, often referred to as a cost function or objective function, while staying within the bounds of the system's physical constraints [30].

OCT has shown great potential in the field of cancer treatment optimization. Here's a more detailed overview of how this theoretical framework can be applied in practical cancer treatment protocols:

- By modeling the evolution of drug resistance in cancer cells, optimal control theory can be used to design treatment strategies that delay the onset of resistance or even reverse it. This can involve switching between different therapies or using adaptive treatment approaches.
- OCT can be integrated with tumor imaging techniques, such as magnetic resonance imaging (MRI) or positron emission tomography (PET), to monitor the response to treatment and adjust the treatment plan accordingly.

TABLE II. MODEL PARAMETERS DESCRIPTIONS

| Parameter | Description |
|-----------|--|
| ρ | Rate of $CD8^+T$ -lysed tumor cell debris activation of $CD8^+T$ cells |
| α | Tumor size for half-maximal $CD8^+T$ -lysed debris $CD8^+T$ activation |
| c_1 | Rate of $CD8^+T$ -cell death due to tumor interaction |
| d_1 | Rate of activated $CD8^+T$ -cell turnover |
| a_1 | Rate of $CD8^+T$ depletion from medicine toxicity |
| r_1 | Growth rate of tumor |
| b_1 | Inverse of carrying capacity of tumor |
| c_2 | Immune system strength coefficient |
| c_3 | Rate of NK-induced tumor death |
| a_2 | Rate of chemotherapy-induced tumor death |
| r_2 | Growth rate of NK cells |
| b_2 | Rate of NK cell turnover |
| c_4 | Rate of NK cell death due to tumor interaction |
| a_3 | Rate of NK depletion from medicine toxicity |
| d_2 | Rate of excretion and elimination of doxorubicin |
| s | $CD8^+T$ -cell normal growth rate |

This can help ensure that the treatment remains optimal throughout the course of therapy [82].

- OCT can be used to design more efficient and informative clinical trials by optimizing the selection of patient cohorts, treatment dosages, and trial endpoints. This can lead to faster and more cost-effective drug development processes.
- In some cases, optimal control theory can be applied in real-time to adjust treatment parameters based on the patient's response, allowing for dynamic and adaptive treatment strategies that adapt to changes in the tumor and patient's condition.

A. Mathematical Preliminaries

In many real-world scenarios the problems can be framed as nonlinear programming (NLP) problems, it deals with optimizing a nonlinear objective function subject to a set of nonlinear constraints [112]. However, when the decision variables are dynamic, changing continuously over time, the

problem becomes one of continuous optimization, also known as an optimal control problem (OCP) [113]. The mathematical formulation of an OCP typically involves defining an objective function, which is a function of both the control and state variables [114]. The solution to an OCP provides the optimal path or trajectory of the control variables over time, which in turn determines the trajectory of the state variables, ensuring that the objective function is optimized while respecting all constraints. The calculus of variations offers techniques to handle such problems, involving concepts like the Hamiltonian function and the Pontryagin's Maximum Principle [115]. These provide a systematic way to derive the necessary conditions that the optimal control and state trajectories must satisfy [31][116].

$$\left. \begin{aligned} & \text{Minimize}_{\mathbf{u}} \quad J(\mathbf{u}) = \varphi(\mathbf{x}(t_f), t_f) + \int_{t_0}^{t_f} L(\mathbf{x}(t), \mathbf{u}(t), t) dt \\ & \text{subject to: } \dot{\mathbf{x}}(t) = \mathbf{a}(\mathbf{x}(t), \mathbf{u}(t), t), \text{ where: } \mathbf{x}(t_0) = \mathbf{x}_0 \\ & \quad \mathbf{g}_1(\mathbf{x}(t)) \leq \mathbf{0} \\ & \quad \mathbf{g}_2(\mathbf{u}(t)) \leq \mathbf{0} \\ & \quad t \in [0, \infty] \end{aligned} \right\} \quad (1)$$

\mathbf{x} is the dependent state variable vector in

$$\mathcal{R}^N \times [t_0, t_f]$$

\mathbf{u} is the independent control variable vector in

$$\mathcal{R}^M \times [t_0, t_f]$$

$J: \mathcal{R}^L \times [t_0, t_f] \rightarrow R$ is the optimization criterion

$L: \mathcal{R}^L \times [t_0, t_f] \rightarrow [t_0, t_f]$ is the bath cost function

$\varphi: \mathcal{R} \times \mathcal{R} \rightarrow \mathcal{R}$ is the final state cost function

$\mathbf{g}_1: \mathcal{R}^N \rightarrow \mathcal{R}^{M_1}$, are the state inequality constraints

$\mathbf{g}_2: \mathcal{R}^M \rightarrow \mathcal{R}^{N_1}$, are the control inequality constraints

$\mathbf{a}: \mathcal{R}^N \rightarrow \mathcal{R}^N$, are the dynamical constraints of the system

B. Necessary Conditions for Optimality (NCO)

The NCO in optimal control theory is fundamental for solving a general OCP as outlined in a referenced equation ((2) [117]). These conditions are derived under specific assumptions and constraints. For the OCPs under consideration, it's common to start without including state or control inequality constraints. State inequality constraints, if present, can either be converted into a new set of dynamic constraints or integrated into the optimization criterion itself. This approach allows for the elimination of certain constraints, such as $g_1(x)$ in (2), simplifying the problem.

There are complexities associated with handling control inequality constraints in Optimal Control Problems (OCPs) and this is a critical challenge that requires a comprehensive exploration of potential solutions and alternative methodologies [118].

- Direct transcription methods, such as direct collocation are widely used to handle control inequality constraints in OCPs. The advantage of direct transcription is that it can handle a wide range of control constraints, including inequalities, by incorporating them as explicit constraints

in the discretized problem. Recent advancements in direct transcription include the use of adaptive mesh refinement, improved constraint handling techniques, and the integration of advanced nonlinear programming algorithms.

- Indirect methods, such as those based on Pontryagin's Maximum Principle, can also be used to handle control inequality constraints, though they may face additional challenges. One approach is to relax the control inequality constraints by introducing a penalty or barrier function in the Hamiltonian, effectively transforming the constrained problem into an unconstrained one. This can be done using techniques like the Augmented Lagrangian or Interior-Point methods, which can help ensure feasibility and convergence.
- Combining direct and indirect methods can lead to hybrid approaches that leverage the strengths of both.

For example, a nested approach where an indirect method is used to solve the inner problem (e.g., optimal control) and a direct method is used to handle the outer problem (e.g., parameter optimization) can be effective in handling control inequality constraints.

- Adaptive and dynamic approaches to constraint handling can be useful in OCPs with control inequality constraints. This can involve techniques like active-set methods, where the active constraints are identified and updated during the optimization process. Alternatively, dynamic constraint relaxation or adaptive constraint scaling can be employed to improve the convergence and robustness of the solution process.

Control inequality constraints, however, present a different challenge. Unlike state constraints, they cannot be transformed or eliminated because the control vector u is independent of equality constraints, as noted in references [118] and [119]. This independence makes the handling of control constraints more complex in the context of OCPs. To derive the NCO, one common approach is to initially assume the absence of any control inequality constraints. This simplification allows for the application of Pontryagin's Minimum Principle, a cornerstone of modern control theory, as referenced in [120]. Pontryagin's principle provides a framework to find the necessary conditions that the control and state trajectories must satisfy for optimality.

Under this principle, a new function called the Hamiltonian is defined. The Hamiltonian is central to deriving the NCO as it encapsulates both the system dynamics and the optimization objective. It is typically a function of the state variables, control variables, and additional variables known as co-state variables. The Hamiltonian integrates the objective function with the system dynamics, providing a comprehensive expression that must be optimized. To incorporate control inequality constraints, after deriving the NCO using the Hamiltonian, additional techniques or modifications may be needed. These could involve augmenting the Hamiltonian with penalty terms or employing other mathematical strategies to account for the constraints on the control variables. By doing so, the solution can be guided to satisfy both the optimality conditions and

the control constraints. This approach is vital for ensuring that the resulting control strategy is not only optimal but also feasible within the given physical and operational limits of the system.

$$\begin{aligned} \mathcal{H}(\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), t) \\ = L(\mathbf{x}(t), \mathbf{u}(t), t) + \boldsymbol{\lambda}^T(t) \\ \cdot \mathbf{a}(\mathbf{x}(t), \mathbf{u}(t), t) \end{aligned} \quad (3)$$

In the given context, $\boldsymbol{\lambda}(t)$ is a vector variable with the same dimensions as $\mathbf{x}(t)$ (i.e., $R^N \times [t_0, t_f]$). It is referred to as the costate variable or Lagrange multiplier. The purpose of introducing this variable is to convert the dynamic constrained Optimal Control Problem (OCP) into an unconstrained OCP. This conversion comes at the cost of doubling the number of state variables. For a more detailed understanding of the derivation of the Hamiltonian equation in (3) and the NCO, please refer to [121] and [122].

From (3), the NCO can be expressed as follows:

$$\left. \begin{aligned} \dot{\mathbf{x}}^*(t) &= \frac{\partial \mathcal{H}(\mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t), t)}{\partial \boldsymbol{\lambda}(t)} \\ \dot{\boldsymbol{\lambda}}^*(t) &= -\frac{\partial \mathcal{H}(\mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t), t)}{\partial \mathbf{x}(t)} \\ 0 &= \frac{\partial \mathcal{H}(\mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t), t)}{\partial \mathbf{u}(t)} \end{aligned} \right\} \quad (4)$$

With the associated boundary conditions:

$$\left. \begin{aligned} \mathbf{x}^*(t_0) &= \mathbf{x}_0 \\ \left[\frac{d\varphi}{d\mathbf{x}}(\mathbf{x}^*(t_f), t_f) - \boldsymbol{\lambda}^*(t_f) \right]^T \cdot \delta \\ + \left[\mathcal{H}^*|_{t_f} + \frac{d\varphi}{d\mathbf{x}}(\mathbf{x}^*(t_f), t_f) \right] \cdot \delta t_f &= 0 \end{aligned} \right\} \quad (5)$$

In this context, φ represents the final state function as defined in (2). $\delta x_f, \delta t_f$ denote the differentials of the terminating state $\mathbf{x}(t_f)$ and final time t_f , respectively. Equation (5) depends on the choice of $\mathbf{x}(t_f)$ and t_f , which can be either free or specified. If either one is free, the corresponding coefficient in (5) is set to zero. For instance, if the final state $\mathbf{x}(t_f)$ is unspecified but t_f is specified (which is often the case in many OCP), the second boundary condition in (10) can be expressed as follows:

$$\boldsymbol{\lambda}^*(t_f) = \frac{d\varphi}{d\mathbf{x}}(\mathbf{x}^*(t_f), t_f) \quad (6)$$

It is evident from the first equation in (5) and (6) that the NCO cannot be directly solved using conventional numerical techniques. This is because the boundary conditions for the state and costate variables do not occur at the same time. In other words, we have knowledge of $\mathbf{x}^*(t)$ at t_0 but $\boldsymbol{\lambda}^*(t)$ at t_f . This disparity in the initial boundary values creates a complex problem to solve, known formally as a split boundary value problem (SBVP). Due to the nonlinear nature of the necessary conditions in the OCP described in equation (4), along with the challenges posed by the SBVP, the combined conditions are often referred to as the curse of complexity. It should be noted that the control inequality constraints, which were previously disregarded, have been addressed in [117] and [120]. It has been demonstrated that

the inclusion of these constraints only affects the third necessary condition in (4).

This section highlights a fundamental issue in solving OCPs, namely the disparity in initial boundary values between state and costate variables. This discrepancy leads to a split boundary value problem (SBVP), which significantly complicates the solution process. The following are some concrete approaches and practical recommendations to tackle this challenge [119][122]:

- Shooting methods are a common approach to resolving SBVPs in optimal control problems. The idea is to "shoot" from the known initial state and iteratively adjust the initial costate values until the final boundary conditions are satisfied. This can be done using techniques like Pontryagin's Maximum Principle, the Indirect Multiple Shooting Method, or the Receding Horizon Control approach. Recent studies have explored advanced shooting methods, such as the Augmented Lagrangian Shooting Method, which can improve convergence and robustness.
- Collocation methods, such as the Gauss-Lobatto collocation or the Legendre-Gauss-Lobatto collocation, can be effective in addressing SBVP challenges. Collocation methods can handle complex dynamics and boundary conditions more flexibly than shooting methods, and they can leverage powerful nonlinear programming solvers.
- Studies have demonstrated the effectiveness of collocation methods in solving a wide range of optimal control problems, including those in cancer treatment.
- Developing adaptive and automatic SBVP resolution methods can be beneficial, especially for complex optimal control problems in cancer treatment. These methods can automatically adjust the discretization, the initialization, or the solution strategy based on the problem characteristics and the intermediate results. Adaptive methods can help mitigate the need for manual tuning and provide more robust and reliable solutions.
- Introducing regularization or penalty terms in the optimal control problem formulation can help address the SBVP challenge. By adding penalty terms or regularization functions to the objective or constraint functions, the problem can be transformed into a more well-posed form, which can improve the convergence and stability of the solution process.
- Leveraging the specific structure or properties of the optimal control problem can sometimes facilitate the SBVP resolution.

For example, if the problem exhibits certain symmetries or has a specific form of dynamics or constraints, specialized methods or simplifications may be applicable.

Pontryagin proposed a more general condition to replace this specific optimal control necessary condition, and it is as follows:

$$\mathcal{H}(\mathbf{x}^*(t), \mathbf{u}^*(t), \boldsymbol{\lambda}^*(t), t) \leq \mathcal{H}(\mathbf{x}^*(t), \mathbf{u}(t), \boldsymbol{\lambda}^*(t), t) \quad (7)$$

The equation denoted as (7) represents the Pontryagin Minimum Principle (PMP). This principle states that the Hamiltonian of the optimal control, $u^*(t)$, must be less than or equal to the Hamiltonian of any other feasible control, $u(t)$, at all points in time.

C. Quadratic Regulation Problem Formulation (QRP)

In advancing the discussion on Optimal Control Problems (OCPs), the focus shifts to a specific format known as a Quadratic Regulator Problem (QRP). The QRP is a well-studied form of OCP, particularly noted for its structured and tractable nature. The primary goal in a QRP is to steer the state variables, denoted as $x(t)$, from a given initial condition $x(0) = x_0$ to a desired equilibrium state. In the context of this discussion, the equilibrium state is set to $x(t_f) = 0$ as $t_f \rightarrow \infty$, where t_f represents the final time. However, in real-world scenarios, the desired equilibrium point might not always be zero (i.e., $x = 0$). If the target equilibrium is a different state $x = x_f$ a change of variables can be applied to transform this target equilibrium point to the origin. This transformation simplifies the problem formulation and analysis, as noted in reference [121].

$$\left. \begin{aligned} \text{Min.}_u \quad J(u) &= \frac{1}{2} \cdot \int_0^{\infty} [x(t)^T \cdot Q(t) \cdot x(t) + u(t)^T \cdot R(t) \cdot u(t)] dt \\ \text{subject to: } \dot{x}(t) &= a(x(t), u(t)) \\ u_l &\leq u(t) \leq u_u \end{aligned} \right\} \quad (8)$$

The NCO for the OCP of (8), can be written using the conditions in (4) and the PMM equating of (7) as:

$$\left. \begin{aligned} \dot{x}^*(t) &= a(x^*(t), u^*(t))n \\ \dot{\lambda}^*(t) &= -Q(t) \cdot x^*(t) - \left[\frac{\partial a(x^*(t), u^*(t))}{\partial x(t)} \right]^T \cdot \lambda^*(t) \\ u^*(t) &= -R(t)^{-1} \cdot \left[\frac{\partial a(x(t), u(t))}{\partial u(t)} \right]^T \cdot \lambda^*(t) \end{aligned} \right\} \quad (9)$$

$\dot{x}^*(t)$ is the state equation. $\dot{\lambda}^*(t)$ is the costate equation. $u^*(t)$ is the optimal control equation. Where the boundary conditions can be written from (5) as:

$$x^*(0) = x_0 \text{ and } \lambda^*(\infty) = 0 \quad (10)$$

If the third equation of (9) violated any control inequality constraints, the PMM equation of (7) is used to find the correct optimal control.

When reformulating the OCP described in (2) into a QRP, specific modifications are made to accommodate this change in objectives and constraints. In a QRP, the objective function typically involves minimizing a quadratic cost function, which represents a balance between the state deviation from the desired equilibrium and the effort (or control energy) required to achieve this state. This quadratic cost function can be expressed as an integral over the control horizon, with terms penalizing both the state deviation and the magnitude of the control input.

The constraints in the QRP are also adapted accordingly. While the original OCP might have had diverse constraints, in a QRP, these are often simplified or reformulated to align with the quadratic nature of the problem. This can involve redefining the system dynamics, state constraints, and control

constraints in a manner that is compatible with the quadratic cost structure.

The transformation to a QRP can offer several advantages, especially in terms of solvability and computational efficiency. QRPs are a well-understood class of control problems with established solution methods, particularly in linear systems where the state dynamics and control functions are linear. These solutions often involve using the Linear Quadratic Regulator (LQR) approach, which provides an optimal control strategy based on solving a set of algebraic Riccati equations. This approach is highly effective for systems where a quadratic cost function accurately captures the trade-offs and objectives of the control problem.

Note that to solve the OCP of (9) (i.e., to find the continuous time functions $x^*(t)$, $u^*(t)$, and $\lambda^*(t)$), one requires a symbolic solver of differential equations, however in practice we do not often solve for the continuous time function, but rather we always seek to use numerical algorithms to solve these OCPs, since computers can run these algorithms significantly fast, and with very high precision.

The choice of the most suitable discretization scheme for optimal control problems in cancer treatment will depend on the specific characteristics of the problem, such as the complexity of the system dynamics, the geometry of the treatment domain, the required accuracy and stability, and the available computational resources. In many cases, a combination of these methods, such as using finite element methods for spatial discretization and finite difference or spectral methods for temporal discretization, can provide a good balance of accuracy, stability, and computational efficiency. we will dive deeper into the different discretization schemes and their comparative analysis in the context of optimal control problems for cancer treatment.

Finite difference methods are relatively simple to implement and understand, can handle a wide range of problem geometries and boundary conditions, and are computationally efficient. But prone to numerical instabilities, especially for high-order derivatives or stiff problems.

Finite element methods can achieve high-order accuracy through the use of higher-order basis functions, and Robust handling of heterogeneous material properties and anisotropic systems. But relatively more complex to implement and require more computational resources.

Spectral methods have excellent accuracy, and high convergence rates, and are computationally efficient, but restricted to problems with simple geometries and periodic or nearly periodic solutions.

Hence in order to solve OCPs numerically they must first be formulated in discrete form. The OCP of (8) can be discretized by using a proper sampling time h as:

$$\left. \begin{aligned} \text{Min.}_u \quad J(u) &= \frac{1}{2} \cdot \sum_{k=0}^{\infty} [x(k)^T \cdot Q(k)x(k) + u(k)^T R(k)u(k)] \\ \text{subject to: } x(k+1) &= f(x(k), u(k)) \\ u_l &\leq u(k) \leq u_u \end{aligned} \right\} \quad (11)$$

Where, $\mathbf{x}(t)$ is replaced with $\mathbf{x}(k)$ for the $t \in [k \cdot h, (k + 1) \cdot h]$, the same is true for $Q(k)$, $R(k)$, $u(k)$, however the dynamic constraints are discretized using any of the famous methods of forward Euler, trapezoidal, 4th Order Runge-Kutta ...etc.

To write the necessary conditions for optimality, we first need to write the Hamiltonian:

$$\mathcal{H}(\mathbf{x}(k), u(k), \lambda(k)) = \frac{1}{2} \cdot [\mathbf{x}(k)^T \cdot Q(k) \cdot \mathbf{x}(k) + u(k)^T \cdot R(k) \cdot u(k)] + \lambda^T(k+1) \cdot f(\mathbf{x}(k), u(k)) \quad (12)$$

Then applying the necessary conditions of (4) on (12) results in

$$\left. \begin{aligned} \mathbf{x}^*(k+1) &= f(\mathbf{x}^*(k), u^*(k)) \\ \lambda^*(k) &= Q(k)x^*(k) + \left[\frac{\partial f(\mathbf{x}(k), u(k))}{\partial \mathbf{x}(k)} \right]^T \cdot \lambda^*(k+1) \\ u^*(k) &= -R(k)^{-1} \cdot \left[\frac{\partial f(\mathbf{x}(k), u(k))}{\partial u(k)} \right]^T \cdot \lambda^*(k+1) \end{aligned} \right\} \quad (13)$$

II. TECHNIQUES FOR SOLVING OCP

A. State Dependent Coefficient Method

The State-Dependent Coefficient (SDC) parameterization is a technique used to handle nonlinear systems by transforming them into a form that is more akin to linear systems. This approach is particularly useful in control theory, especially when dealing with systems that exhibit nonlinear dynamics but can be approximated or represented in a linear-like structure. The fundamental idea behind SDC parameterization is to decompose the nonlinear dynamics of the system into components that include state vectors, control vectors, and matrices that vary depending on the state variables. This decomposition allows for a more manageable representation and analysis of the system, especially when applying linear control techniques. Given the system dynamics represented in (14) [121]:

$$\dot{\mathbf{X}}(t) = a(\mathbf{x}(t), u(t)) \quad (14)$$

Applying the SDC parameterization yields the following expression for the resultant system:

$$\dot{\mathbf{X}} = A(\mathbf{x}) \cdot \mathbf{x} + B(\mathbf{x}) \cdot u \quad (15)$$

The dynamic system representation of (15), is often called nonlinear state dependent affine control system, the term affine means linear in controls i.e., terms such as $u_1, u_2, \cos(u_1), e^{-u_3}$...etc. may not appear in the system dynamics, however terms such as $(5 \cdot e^{-x_1} - e^{-x_2}) \cdot \cos(x_1) \cdot u_1$ may appear since it is linear in the control input u_1 . The SDC representation of (15), is similar to the following linear system representation in state space:

$$\dot{\mathbf{x}} = A \cdot \mathbf{x} + B \cdot u \quad (16)$$

Before delving into the algorithm, it is important to introduce the Linear Quadratic Regulator (LQR) method and its solution for Linear Time-Invariant (LTI) systems. The LQR is an optimization technique used for linear systems

represented by (16). The primary objective of the LQR is to drive the states, denoted as \mathbf{x} , and the control input, u , to the origin starting from any initial conditions $\mathbf{x}(0) = \mathbf{x}_0$. This is achieved by minimizing the quadratic cost associated with the states and control inputs. The problem formulation for the LQR can be expressed as follows:

$$\left. \begin{aligned} \text{Min.}_u \quad J(u) &= \frac{1}{2} \cdot \int_0^\infty [\mathbf{x}(t)^T \cdot Q \cdot \mathbf{x}(t) + u^T(t) \cdot R \cdot u(t)] dt \\ \text{subject to: } \dot{\mathbf{x}}(t) &= A \cdot \mathbf{x}(t) + B \cdot u(t) \end{aligned} \right\} \quad (17)$$

The necessary conditions for optimality follow from (9) as:

$$\left. \begin{aligned} \dot{\mathbf{x}}^*(t) &= A \cdot \mathbf{x}^*(t) + B \cdot u^*(t) \\ \dot{\lambda}^*(t) &= -Q \cdot \mathbf{x}^*(t) - A^T \cdot \lambda^*(t) \\ u^*(t) &= -R^{-1} \cdot B^T \cdot \lambda^*(t) \end{aligned} \right\} \quad t \in [0, \infty] \quad (18)$$

With the same boundary conditions of (10), i.e.:

$$\mathbf{x}^*(0) = \mathbf{x}_0 \text{ and } \lambda^*(\infty) = 0$$

The LQR solution relies on the fact, that the optimal value of the constate variable vector λ is represented as:

$$\lambda^*(t) = P \cdot \mathbf{x}^*(t) \quad (19)$$

Using (19), and the set of (18), it is possible to write:

$$P \cdot A + A^T \cdot P - P \cdot B \cdot R^{-1} \cdot B^T \cdot P + Q = 0 \quad (20)$$

Equation (20) is widely recognized as the Algebraic Riccati Equation (ARE), which plays a crucial role in the LQR. In this equation, the only unknown is the P matrix. By solving the ARE, we can determine the optimal control, denoted as $u^*(t)$, as a state-dependent control using the third equation from (18) as follows:

$$\left. \begin{aligned} u^*(t) &= -R^{-1} \cdot B^T \cdot \lambda^*(t) = -R^{-1} \cdot B^T \cdot P \cdot \mathbf{x}^*(t) \\ u^*(t) &= -K \cdot \mathbf{x}^*(t) \end{aligned} \right\} \quad (21)$$

The K matrix represents the optimal weights obtained from the LQR problem. It is important to note that the LQR method effectively solves the OCP stated in (17) by transforming a differential equation (i.e., (16)) into an algebraic equation (i.e., equation (20)). This transformation simplifies the problem, making it easier to solve.

B. Adaptive Dynamic Programming

In the context of Optimal Control Problems (OCPs), especially when considering approaches like Quadratic Regulator Problems (QRPs), certain preliminary concepts and equations are essential for understanding and implementing the technique. Two main equations are pivotal in this scenario: one for evaluating the optimal cost function and the other for calculating the optimal control. These can be generally reformulated as follows [122][123]:

$$J_{k,N}^*(\mathbf{x}_k) = \text{Min.}_{u_k} \left(U(\mathbf{x} = A\mathbf{x} + B\mathbf{u}, u_k) + J_{k+1,N}^*(\mathbf{x}_{k+1}) \right) \quad (22)$$

$$u_k^* = \arg \text{Min.}_{u_k} \left(U(\mathbf{x}_k, u_k) + J_{k+1,N}^*(\mathbf{x}_{k+1}) \right) \quad (23)$$

Where, $J_{k,N}^*/(\mathbf{x}_k)$ is defined as the cost-to-go function (i.e., it evaluates to the total cost to go from the current state \mathbf{x}_k to

the final state x_N). $U(x_k, u_k)$ is often called the utility function, it evaluates the cost of applying the current control u_k to the current state x_k . In QRP's the utility function U is defined as:

$$U(x_k, u_k) = \frac{1}{2} [x_k^T \cdot Q_k \cdot x_k + u_k^T \cdot R_k \cdot u_k] \quad (24)$$

$\arg \min_{u_k} (\cdot)$ means the value of the argument u_k that caused the minimum value of the expressing in the brackets.

C. Heuristic Action-Critic

Dynamic programming DP is another powerful approach to solve OCPs, it employs the principle of optimality, which asserts that an optimal solution to a problem can be achieved by merging the optimal solutions to its constituent subproblems but the main drawback in DP approach was the curse of dimensionality that arises when we try to account for all states and controls at each time step k , however in Adaptive Dynamic Programming (ADP) an approach involves training two Neural Networks (heuristically in an iterative manner until eventually they converge to the optimal solution) is used [124]. These two NNs are often called Action network and Critic network and their combined structure is called Action-Critic (AC) [125]. The learning strategy used to train NN in classification and function approximation is commonly known as supervised learning. This type learning strategy is often referred to in the machine learning field as Reinforcement learning (RL). RL is fundamentally based on rewarding, i.e., when the NN is doing well it receives a reward meaning that it is in the correct direction and vice versa. The most famous structure of RL comprises of two blocks namely Agent and Environment as in Fig. 1. The agent represents the NN to be trained, while the environment represents the place where the actions of the NN will be performed and judged.

In a RL system, there are typically four main components: $\{X, U, R, F\}$. X represents the set of states, U represents the set of actions, R represents the set of scalar reinforcement signals or rewards, and F represents the function that describes the transition from one state to another under a specific action, denoted as $F: X \times U \rightarrow X$. A policy π is defined as a mapping from states X to actions U , such that at any given time k , the system can be in a state x_k from the set X .

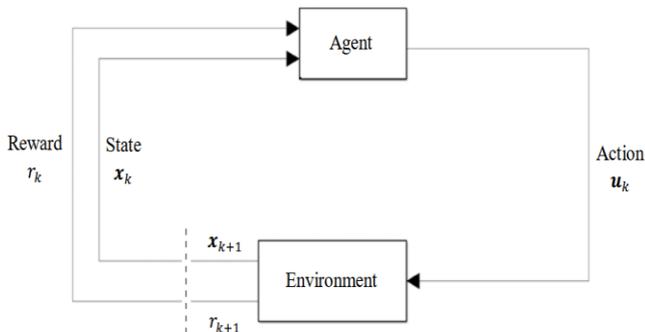


Fig. 1. Reinforcement Learning (RL) Agent-Environment Structure

It takes an action u_k from the set U , which is determined by the policy π , i.e., $u_k = \pi(x_k)$. The system then transitions to the next state x_{k+1} , denoted as $x_{k+1} = F(x_k, u_k)$.

Simultaneously, it receives a reward signal r_{k+1} from the set R , which is determined by the current state x_k , action u_k , and the next state x_{k+1} , i.e., $r_{k+1} = r(x_k, u_k, x_{k+1})$. The objective of RL is to determine a policy that minimizes the accumulated reward starting from the initial state x_0 at $t = 0$. In RL tasks, the estimation of value functions is always involved. A value function provides an estimate of how advantageous it is to be in a particular state x_k . It is defined as follows [125]:

$$J^\pi(x_k) = \sum_{n=0}^{\infty} r_{k+n} = \sum_{n=0}^{\infty} r(x_{k+n}, u_{k+n}, x_{k+1+n}) \quad (25)$$

The optimal value function $J^*(x_k)$ can be written in a recursive form as:

$$J^*(x_k) = \min_{u_k} r(x_{k+n}, u_{k+n}, x_{k+1+n}) + J^*(x_{k+1}) \quad (26)$$

Where, $r(x_{k+n}, u_{k+n}, x_{k+1+n}) = \frac{1}{2} [x_{k+n}^T \cdot Q_k \cdot x_{k+n} + u_{k+n}^T \cdot R_k \cdot u_{k+n}] U(x_k, u_k)$ is often called the utility function, it evaluates the cost of applying the current control u_k to the current state x_k . In order for the control u_k to be optimal, the following condition must be satisfied for all k :

$$\frac{\partial J(x_k)}{\partial u_k} = 0 \quad (27)$$

Then

$$\frac{\partial J(x_k)}{\partial u_k} = \frac{\partial u_k}{\partial u_k} + \frac{\partial J(x_{k+1})}{\partial u_k} = \frac{\partial u_k}{\partial u_k} + \frac{\partial x_{k+1}}{\partial u_k} \cdot \frac{\partial J(x_{k+1})}{\partial x_{k+1}} = 0 \quad (28)$$

The new definition of the costate variable λ_k follows as

$$\lambda_k^* = \frac{\partial J(x_k)}{\partial x_k} \quad (29)$$

The costate variables λ_k has the interpretation of the sensitivity vector of the state variables x_k with respect to the optimization criterion $J(x_k)$. Substituting the definitions of λ_k and U_k into (28) we get

$$u_k^* = -R_k^{-1} \cdot \frac{\partial x_{k+1}}{\partial u_k} \cdot \lambda_{k+1}^* = -R_k^{-1} \cdot \left[\frac{\partial F(x_k, u_k)}{\partial u_k} \right]^T \cdot \lambda_{k+1}^* \quad (30)$$

Similarly

$$\lambda_k^* = \frac{\partial J(x_k)}{\partial x_k} = \frac{\partial u_k}{\partial x_k} + \frac{\partial J(x_{k+1})}{\partial x_k} = Q_k \cdot x_k^* + \frac{\partial x_{k+1}}{\partial x_k} \cdot \lambda_{k+1}^* \quad (31)$$

$$\lambda_k^* = Q_k \cdot x_k^* + \left[\frac{\partial F(x_k, u_k)}{\partial x_k} \right]^T \cdot \lambda_{k+1}^*$$

The action network is trained to find the optimal state feedback control u_k^* which is often written as $u_k^* = u_k^*(x_k)$ while the critic network is trained to find the optimal relation of the costate variable λ_k^* which is also written as $\lambda_k^* = \lambda_k^*(x_k)$. The most famous training technique for the AC system is heuristic training, this technique is also called Heuristic Action Critic (HAC). The HAC technique uses the optimal control equation of (30) along with the costate equation of (31) to train the action and critic networks [124]. Fig. 2 shows the block diagram of the HAC technique. The HAC technique

starts by initializing the action and critic networks [126], then the training process of the action network is as follows:

1. Generate P random states vectors $\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,P}$.
2. Feed these states into the action network, to get the estimated control vectors $\hat{\mathbf{u}}_{k,1}, \hat{\mathbf{u}}_{k,2}, \dots, \hat{\mathbf{u}}_{k,P}$.
3. These controls along with the states are fed to the system model (or the NN version of the system model) the resultant next states are $\mathbf{x}_{k+1,1}, \mathbf{x}_{k+1,2}, \dots, \mathbf{x}_{k+1,P}$.
4. These next states are fed to the critic network to get the estimated costates variables $\hat{\lambda}_{k+1,1}, \hat{\lambda}_{k+1,2}, \dots, \hat{\lambda}_{k+1,P}$.
5. This is the heuristic step in the training of action network: use the optimal control equation of (30) with the current states $\mathbf{x}_{k,i}$ and next costates $\hat{\lambda}_{k+1,i}$ (for $i = 1 \dots P$) as inputs to obtain the better controls $\mathbf{u}_{k,1}, \mathbf{u}_{k,2}, \dots, \mathbf{u}_{k,P}$, note that (30) is only defined for the optimal control \mathbf{u}_k^* only, however as was proven in [21] this equation can be used to generate better control sequence $\mathbf{u}_{k,i}$ than the estimated one $\hat{\mathbf{u}}_{k,i}$ (for $i = 1 \dots P$).
6. Finally, this new better control sequence is used to train the action network by minimizing the error difference $e_{a,i} = \|\mathbf{u}_{k,i} - \hat{\mathbf{u}}_{k,i}\|$ (for $i = 1 \dots P$).

Hence the updating iterative rule for the training of the action network is:

$$\mathbf{u}_k^{[j]} = -\mathbf{R}_k^{-1} \cdot \left[\frac{\partial \mathbf{F}(\mathbf{x}_k^{[j-1]}, \mathbf{u}_k^{[j-1]})}{\partial \mathbf{u}_k^{[j-1]}} \right]^T \cdot \lambda_{k+1}^{[j-1]} \quad (32)$$

Where j is the iteration index. At the same time the critic network is trained as follows:

7. The same states $\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,P}$ are fed to the critic network to generate the estimated costate variables $\hat{\lambda}_{k,1}, \hat{\lambda}_{k,2}, \dots, \hat{\lambda}_{k,P}$ (for $i=1 \dots P$).
8. This is the heuristic step in the training of critic network: use the costate of (31) with the current states $\mathbf{x}_{k,1}$ and next costates $\lambda_{k+1,i}$ (for $i = 1 \dots P$) as inputs to obtain the better costates $\lambda_{k,1}, \lambda_{k,2}, \dots, \lambda_{k,P}$, which follows the same heuristic argument used in (30).
9. Finally, this new better costate sequence is used to train the critic network by minimizing the error difference $e_{c,i} = \|\lambda_{k,i} - \hat{\lambda}_{k,i}\|$, (for $i=1 \dots P$).
10. This completes one iteration of training; the training keeps oscillating between the two networks until the stopping criterion is met.

Similarly, the updating iterative rule for the training of the critic network can be written as:

$$\lambda_k^{[j]} = \mathbf{Q}_k \cdot \mathbf{x}_k^{[j-1]} + \left[\frac{\partial \mathbf{F}(\mathbf{x}_k^{[j-1]}, \mathbf{u}_k^{[j-1]})}{\partial \mathbf{x}_k^{[j-1]}} \right]^T \cdot \lambda_{k+1}^{[j-1]} \quad (33)$$

In case the of lack of system model, we can train a NN to approximate the model of the system using the Radial Basis Function (RBF). The first layer in RBF network consists of N neurons that represents the input vector \mathbf{x} , these N neurons

are connected to the Nh hidden neurons of the second layer, which in turn connected to the single neuron in the output layer [21]. The output equation of this RBF model can be written as (34).

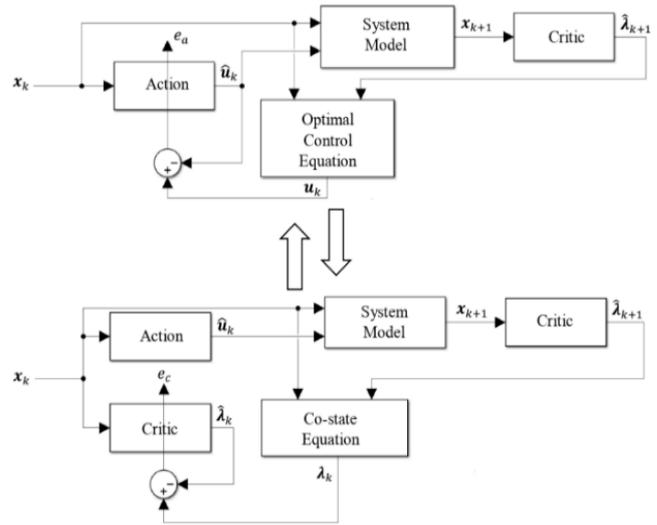


Fig. 2. Heuristic Action-Critic technique block diagram

$$\mathbf{x}_{k+1} = \mathbf{W}_o + \tanh(\mathbf{W}_i \cdot \mathbf{z}_k) \quad (34)$$

Where, $\mathbf{z}_k = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{u}_k \\ \mathbf{1} \end{bmatrix}$ is a vector of size $N + M + 1$, and

$\mathbf{W}_o, \mathbf{W}_i$ represent the wights of input and output respectively. The partial derivative used in (30 and 31) can also be obtained from (34) as

$$\frac{\partial \mathbf{x}_{k+1}}{\partial \mathbf{u}_k} = \mathbf{W}_o \cdot \frac{\partial \tanh(\mathbf{W}_i \cdot \mathbf{z}_k)}{\partial (\mathbf{W}_i \cdot \mathbf{z}_k)} \cdot \mathbf{W}_i \cdot \frac{\partial \mathbf{z}_k}{\partial \mathbf{u}_k} \quad (35)$$

$$\frac{\partial \mathbf{x}_{k+1}}{\partial \mathbf{x}_k} = \mathbf{W}_o \cdot \frac{\partial \tanh(\mathbf{W}_i \cdot \mathbf{z}_k)}{\partial (\mathbf{W}_i \cdot \mathbf{z}_k)} \cdot \mathbf{W}_i \cdot \frac{\partial \mathbf{z}_k}{\partial \mathbf{x}_k} \quad (36)$$

Where:

$$\frac{\partial \mathbf{z}_k}{\partial \mathbf{u}_k} = \begin{bmatrix} \mathbf{0}_{N \times N} \\ \mathbf{I}_{M \times M} \\ \mathbf{0} \end{bmatrix}, \quad \frac{\partial \mathbf{z}_k}{\partial \mathbf{x}_k} = \begin{bmatrix} \mathbf{I}_{N \times N} \\ \mathbf{0}_{M \times M} \\ \mathbf{0} \end{bmatrix}$$

$$\frac{\partial \tanh(\mathbf{W}_i \cdot \mathbf{z}_k)}{\partial (\mathbf{W}_i \cdot \mathbf{z}_k)} = \text{diag} (1 - \mathbf{W}_i \cdot \mathbf{z}_k \cdot (\mathbf{W}_i \cdot \mathbf{z}_k)^T)$$

D. The Single Network Adaptive Critic

The SDRE technique relies fundamentally on the concept of extended linearization. Extended linearization is the action of generalizing the linear systems-based technique to be applied to the class of nonlinear systems. The SDRE technique redefines the system of (8) as:

$$\left. \begin{aligned} \text{Min. } J(u) &= \frac{1}{2} \cdot \int_0^{\infty} [x^T \cdot Q(x) \cdot x + u^T \cdot R(x) \cdot u] dt \\ \text{subject to: } \dot{x} &= A(x) \cdot x + B(x) \cdot u \\ u_l &\leq u \leq u_u \end{aligned} \right\} \quad (37)$$

The SNAC is a special case of the AC system, that eliminates the oscillatory training between the action and critic networks in HAC technique. The SNAC technique can be used only if the system model can be represented in a

nonlinear affine format as in the discrete OCP of (37). The action network can be eliminated in the SNAC, since from (30) and using the system model of (37) we can write (38).

$$\begin{aligned} \mathbf{u}_k^* &= -\mathbf{R}_k^{-1} \cdot \left[\frac{\partial \mathbf{F}(\mathbf{x}_k, \mathbf{u}_k)}{\partial \mathbf{u}_k} \right]^T \cdot \boldsymbol{\lambda}_{k+1}^* \\ &= -\mathbf{R}_k^{-1} \cdot \mathbf{B}_k^T(\mathbf{x}_k) \cdot \boldsymbol{\lambda}_{k+1}^* \end{aligned} \quad (38)$$

Note that the optimal control in (38) is only a function of the current state and next costate variables, hence we can train the critic network and use its results to compute the control action. Furthermore, the critic network used in SNAC relates the next constate variable to the current state variable i.e., $\boldsymbol{\lambda}_{k+1} = \boldsymbol{\lambda}_{k+1}(\mathbf{x}_k)$. The SNAC technique (Fig. 3) can be summarized in the following steps:

1. Generate P random states vectors $\mathbf{x}_{k,1}, \mathbf{x}_{k,2}, \dots, \mathbf{x}_{k,P}$.
2. Feed these states into the critic network, to get the estimated next costate vectors $\hat{\boldsymbol{\lambda}}_{k+1,1}, \hat{\boldsymbol{\lambda}}_{k+1,2}, \dots, \hat{\boldsymbol{\lambda}}_{k+1,P}$.
3. These next costates along with the states are fed to the modified optimal control equation of (35) to get the control vectors $\mathbf{u}_{k,1}, \mathbf{u}_{k,2}, \dots, \mathbf{u}_{k,P}$.
4. These control vectors along with the current states are fed to the system model of (15), to obtain the next states $\mathbf{x}_{k+1,1}, \mathbf{x}_{k+1,2}, \dots, \mathbf{x}_{k+1,P}$.
5. These next states are fed to the critic network to get the estimated next-next costates variables $\hat{\boldsymbol{\lambda}}_{k+2,1}, \hat{\boldsymbol{\lambda}}_{k+2,2}, \dots, \hat{\boldsymbol{\lambda}}_{k+2,P}$. (note that the critic network gives the costate vector with is one step ahead in time with respect to the input state vector).
6. If the next-next costates were used along with the next states in the costate equation of (31) we obtain the next costates $\hat{\boldsymbol{\lambda}}_{k+1,1}, \hat{\boldsymbol{\lambda}}_{k+1,2}, \dots, \hat{\boldsymbol{\lambda}}_{k+1,P}$, which in turn used to train the critic network.
7. The SNAC technique is also applied iteratively until the stopping criterion is met or the squared error $e_{c,i} = \|\boldsymbol{\lambda}_{k+1,i} - \hat{\boldsymbol{\lambda}}_{k+1,i}\|$ (for $i = 1 \dots P$) is less than a specified limit.

The update iterative equations for the SNAC are the one in (33) and the following one:

$$\mathbf{u}_k^{[j]} = -\mathbf{R}_k^{-1} \cdot \mathbf{B}_k^T(\mathbf{x}_k^{[j-1]}) \cdot \boldsymbol{\lambda}_{k+1}^{[j-1]} \quad (39)$$

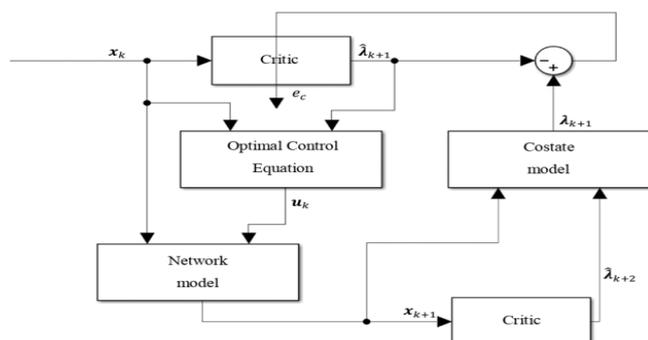


Fig. 3. Heuristic SNAC technique block diagram

Important observation: (33), (34) and (39) presents the updating rule for the training of action and critic networks, these equations differ from the ones used in supervised learning in the sense that they do not represent the optimal final values (i.e. targets) but rather a better target than the one previously computed until eventually they converge to the optimal targets, which is the fundamental idea of RL.

E. Policy Iteration and Value Iteration

There are two famous strategies used in RL, namely policy iteration and value iteration. In policy iteration (PI), the policy (control) is first initialized to an admissible policy, and this policy is improved at each iteration. The AC and SNAC structures used in Fig. 2 and 3 respectively both follow the PI strategy, since the action network was initialized to produce an admissible control for all x_k before the training process starts. Note that it may not be easy to find an admissible control to initialize the system with, however fortunately the most famous admissible control used in policy iteration initialization is the optimal control for the linearized system. Another heuristic trick used in the training of the action and critic networks, is to randomly generate the P state vectors $x_{k,1}, x_{k,2}, \dots, x_{k,P}$, near their final value (which is 0 for the QRP), i.e. $|x_{k,i}| \leq c_i$ and gradually increase the value of c_i until it reaches the specified limits by the state constraints.

While such heuristic approaches can sometimes expedite convergence or improve performance, they also introduce a level of arbitrariness that should be carefully considered. Let's discuss some of the potential issues with these heuristic techniques:

- Heuristic tricks, such as randomly generating state vectors near their final values and gradually increasing them, may not have a strong theoretical foundation or be grounded in the underlying principles of the optimal control problem.
- The effectiveness of these heuristics is often based on empirical observation or intuition rather than a rigorous mathematical analysis of their properties and convergence guarantees. Without a solid theoretical justification, the reliability and generalizability of these heuristics may be limited, and their application may lead to suboptimal or even incorrect solutions.
- The performance and robustness of heuristic approaches can be highly dependent on the specific formulation of the optimal control problem, including the system dynamics, constraints, and cost function. What works well for one problem may not necessarily be effective within the same domain of cancer treatment.
- By biasing the initial conditions or intermediate states towards specific target values, the heuristic tricks may limit the exploration of the entire solution space and potentially lead to suboptimal or even locally optimal solutions.
- The use of heuristic tricks can make the optimization process less transparent and more difficult to interpret, as the reasoning behind the specific choices and their impact on the final solution may not be readily apparent. This

lack of transparency can make it challenging to understand the underlying mechanisms and validate the robustness of the optimal control strategy, which is crucial for building trust and acceptance in the medical community.

- The use of heuristic tricks can introduce variability and non-determinism in the optimization process, making it difficult to establish a consistent benchmark for comparing different optimal control approaches. This lack of standardization can hinder the ability to rigorously evaluate and compare the performance of different methods, which is essential for advancing the field of optimal control in cancer treatment.

On the other hand, the value iteration (VI) strategy uses the value (cost-to-go) function $J_{k,N}$ of (22) to train the action and critic networks. The VI initializes the value function for all x_k , and as the VI algorithm proceeds, this value function decreases (in minimization OCP) for all the states x_k , until eventually converges to the optimal solution (i.e., $J_{k,N} \rightarrow J_{k,N}^*$), the following section discusses how the VI strategy can be used in the Heuristic AC technique. The VI algorithm is described in the following steps, and Fig. 4 shows its block diagram:

1. Generate P random states vectors $x_{k,1}, x_{k,2}, \dots, x_{k,P}$.
2. Feed these states into the action network, to get the estimated control vectors $\hat{u}_{k,1}, \hat{u}_{k,2}, \dots, \hat{u}_{k,P}$, and to the critic network to get the value functions $\lambda_{k,N,1}, \lambda_{k,N,2}, \dots, \lambda_{k,N,P}$. (Note that $\lambda_{k,N,i}(x_{k,i})$ represents the estimated accumulated value function of being at $x_{k,i}$ until the system converges to the final state $x_N = 0$).
3. These controls along with the states are fed to the system model (or the NN version of the system model) the resultant next states are $x_{k+1,1}, x_{k+1,2}, \dots, x_{k+1,P}$.
4. These next states are fed to the critic network to get the estimated value functions $\hat{J}_{k+1,N,1}, \hat{J}_{k+1,N,2}, \dots, \hat{J}_{k+1,N,P}$.
5. Next, to train the critic network we use (22) heuristically, to produce the better value functions $J_{k,N,1}, J_{k,N,2}, \dots, J_{k,N,P}$, where the values of the utility functions $U_{k,1}, U_{k,2}, \dots, U_{k,P}$ of (22) can be found using the current states $x_{k,i}$ along with the estimated controls $u_{k,i}$ (for $i = 1 \dots P$).
6. To train the action network, we can use (30) as it was done in the case of PI version of the AC technique, however the value of the costate variable λ_{k+1} is not radially available, but rather needs to be computed using the estimated value function $\hat{J}_{k+1,N}$ as follows:

Since the RBF was used to describe the relation between the value function $J_{k,N}$ and the state variable x_k then we can write:

$$J_{k,N}(x_k) = W_o \cdot \tanh(W_i \cdot x_k) \quad (40)$$

Then following the same reasoning used in (35), (36) we can find:

$$\lambda_k = \frac{\partial J_{k,N}(x_k)}{\partial x_k} = W_o \cdot \frac{\partial \tanh(W_i \cdot x_k)}{\partial (W_i \cdot x_k)} \cdot W_i \quad (41)$$

Where:

$$\frac{\partial \tanh(W_i \cdot x_k)}{\partial (W_i \cdot x_k)} = \text{diag}(1 - W_i \cdot x_k)$$

As in the case of PI strategy, the VI strategy must also initialize (pretrain) the action and critic networks first before applying the preceding algorithm, however the admissibility test of the value function is:

Let $J_{k,N}^{[0]}(x_k)$ be the value function used in pretraining, and $J_{k,N}^{[1]}(x_k)$ is the refined version obtained from (27), then it can be showed that if $J_{k,N}^{[1]}(x_k) \leq J_{k,N}^{[0]}(x_k)$ (i.e. better than $J_{k,N}^{[0]}(x_k)$) then the value iteration algorithm will asymptotically converge [125].

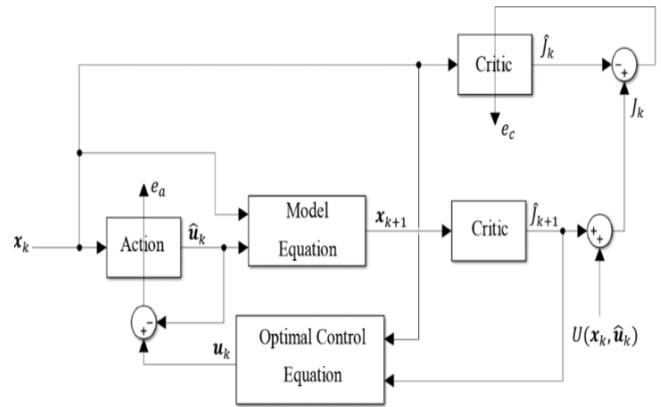


Fig. 4. Value Iteration Heuristic AC technique block diagram

The underlying assumption of this pretraining strategy is that the system dynamics can be adequately represented by a nonlinear affine model structure. However, in practical applications, the actual system dynamics may deviate from this assumption, potentially due to model uncertainties, nonlinearities, or other complexities. The robustness of the pretraining strategy to deviations in system dynamics such as model uncertainties, nonlinearities, or other complexities is an important factor to consider, as it can impact the convergence, solution quality, and stability of the optimal control problem. Exploring alternative cost function structures, such as state-dependent or control-dependent weighting matrices, may enhance the adaptability of the pretraining strategy to diverse system configurations. However, as the system complexity or the problem scale increases, the computational demand may also grow, potentially limiting the practical applicability of the pretraining strategy. Exploring ways to improve the computational efficiency and scalability of the pretraining strategy, such as through the use of model reduction techniques or parallel computing, can enhance its robustness and practical viability [126].

In the context of cancer treatment, there are numerous sources of uncertainty, such as patient-specific parameters, disease progression, and the response to treatment interventions. The robustness of the pretraining strategy to these uncertainties is crucial, as it can determine the

reliability and adaptability of the optimal control solutions in real-world scenarios. Incorporating techniques like robust optimization, adaptive control, or stochastic optimal control into the pretraining strategy may help improve its resilience to uncertainties and disturbances [127].

A common value function used in the pretraining of nonlinear affine systems is:

$$J_{k,N}^{[0]}(\mathbf{x}_k) = \frac{1}{2} \cdot (\mathbf{x}_k^T \cdot \mathbf{P}_0 \cdot \mathbf{x}_k) \quad (42)$$

Then using (38), which is rewritten here for convenience:

$$\begin{aligned} \mathbf{u}_k &= -R_k^{-1} \cdot B_k^T \cdot \frac{\partial J_{k,N}^{[0]}(\mathbf{x}_{k+1})}{\partial \mathbf{x}_{k+1}} = -R_k^{-1} \cdot B_k^T \cdot P_0 \cdot \mathbf{x}_{k+1} \\ &= -R_k^{-1} \cdot B_k^T \cdot P_0 \cdot [A_k \cdot \mathbf{x}_k + B_k \cdot \mathbf{u}_k] \end{aligned}$$

Which can be solve for \mathbf{u}_k as:

$$\mathbf{u}_k = -(\mathbf{B}_k^T \cdot \mathbf{P}_0 \cdot \mathbf{B}_k + \mathbf{R}_k)^{-1} \cdot \mathbf{B}_k^T \cdot \mathbf{P}_0 \cdot \mathbf{A}_k \quad (43)$$

The quadratic cost matrix \mathbf{P}_0 of equation (42), is often chosen to be in the form:

$$\mathbf{P}_0 = \alpha \cdot \mathbf{I}_{N \times N} \quad (44)$$

where $\mathbf{I}_{N \times N}$ is the identity matrix of size $N \times N$. The parameter α is varied until the convergence condition of $J_{k,N}^{[0]}(\mathbf{x}_k)$ is satisfied for all \mathbf{x}_k . The PI strategy is often convergences faster compared to the VI strategy when applied to the same system.

The main difference between VI and PI strategies when applied to the Heuristic AC technique, is the definition of the critic network [128]. In PI the critic network is trained to find the relation between the current costate λ_k and the current state \mathbf{x}_k (i.e., $\lambda_k = \lambda_k(\mathbf{x}_k)$), whereas in the VI, the critic network is trained to find the relation between the current value function $J_{k,N}$ and the current state \mathbf{x}_k (i.e., $J_{k,N} = J_{k,N}(\mathbf{x}_k)$).

Larger values of α :

- Tend to put more weight on the terminal cost, which can lead to faster convergence of the value function and control policy.
- May lead to solutions that prioritize the terminal cost over the stage-wise cost, potentially resulting in suboptimal performance during the intermediate stages of the control problem.
- Tend to make the system more sensitive to perturbations, as the terminal cost becomes the dominant factor in the overall cost function.

Smaller values of α :

- may result in slower convergence, as the algorithm relies more on the stage-wise cost to update the value function and control policy.
- tend to produce solutions that balance the stage-wise cost and the terminal cost, leading to a more optimal overall performance.

- make the system more robust to disturbances, as the stage-wise cost becomes more influential in the overall cost function.

In the domain of cancer treatment optimization, the use of PI can be a powerful tool, but it is not without its challenges, especially when dealing with the complexity of biological systems and the multitude of treatment options [129]-[131].

- The number of possible states and actions in a comprehensive cancer treatment model can quickly become overwhelming, as the state and action spaces grow exponentially with the number of relevant factors.
- This makes it computationally intractable to systematically evaluate all possible policies through brute-force search.
- The human body and the dynamics of cancer are inherently complex, with many interacting processes and nonlinear relationships. Accurately modeling these systems and their response to various treatments is a significant challenge, which can limit the ability to reliably evaluate the performance of candidate policies.
- Uncertainty and Personalization: Cancer treatment often involves a high degree of uncertainty, as the individual patient's response to therapy can vary significantly. Accounting for this uncertainty and developing personalized treatment strategies add additional complexity to the optimization problem.
- Cancer treatment frequently involves a combination of interventions, such as chemotherapy, radiation therapy, and immunotherapy. Coordinating these various treatment modalities and finding the optimal integration of these approaches further increases the computational burden.

To address these challenges, researchers in the field of optimal control theory for cancer treatment have explored various techniques, such as:

- Employing value function approximation, policy parameterization, and other approximate dynamic programming techniques to make the problem more tractable.
- Leveraging detailed cancer treatment simulation models and advanced optimization algorithms to efficiently explore the policy space and identify effective treatment strategies.
- Leveraging the expertise of medical professionals and incorporating their insights into the optimization process to guide the search for suitable policies.

These approaches aim to mitigate the computational complexities and overcome the limitations in finding admissible initial policies for cancer treatment optimization, thereby improving the practical applicability of policy iteration and other optimal control techniques in this domain.

The convergence and stability of the solution can be affected by several factors, which can lead to the failure or

instability of the optimization process. Here are some key considerations [132]-[136]:

- Convergence and stability of the optimal control solution are heavily dependent on the accuracy and completeness of the underlying cancer treatment model. Factors such as nonlinearities, time delays, and uncertainties in the model can introduce instabilities and prevent the optimization from reaching a stable, optimal solution.
- The way the optimal control problem is formulated, including the objective function and the constraints, can significantly impact the convergence and stability of the solution. Poorly defined or overly restrictive constraints, such as strict limits on treatment dosages or schedules, can lead to infeasible or unstable solutions.
- The inherent complexity and dynamic nature of cancer progression and treatment response can pose significant challenges for optimal control techniques. Factors such as tumor heterogeneity, the evolving nature of cancer cells, and the interplay between various treatment modalities can introduce instabilities and make it difficult to find a globally optimal solution. Many of the parameters in the cancer treatment model, such as tumor growth rates, treatment efficacy, and side effects, are subject to significant uncertainty due to inter-patient variability and limited data.

To address these challenges in using optimal control theory for cancer treatment, value function approximation, policy parameterization, and other approximate dynamic programming techniques can be used to make the problem more solvable.

III. RESULT AND DISCUSSION

A. Equilibrium Points

Equilibrium points are the states where the system remains stable if left undisturbed. In the context of cancer therapy, these points might represent states where the disease is under control or in remission. Identifying the equilibrium points is crucial as it helps in understanding the long-term behavior of the system under various treatment strategies. This analysis is essential for designing control strategies that can steer the system towards a desired equilibrium state, thereby optimizing the treatment outcome within the framework of the QRP [137].

Each of the differential equations of cancer model describes the rate of change of the corresponding state variable over time, considering the effects of both the disease dynamics and the treatment controls [138].

To find the equilibrium points of this system, all the dynamic equations in (1) are set to zero (note that at equilibrium points the controls signals are also set to zero), and the state values are solved accordingly [139]-[141].

$$\dot{E}(t) = 0, \dot{T}(t) = 0, \dot{N}(t) = 0 \text{ and } \dot{M}(t) = 0$$

$$\begin{aligned} 0 &= s + \frac{\rho \cdot E(t) \cdot T(t)}{\alpha + T(t)} - c_1 \cdot E(t) \cdot T(t) \\ &\quad - d_1 \cdot E(t) - a_1 \cdot (1 - e^{-M(t)}) \cdot E(t) + w(t) \\ 0 &= r_1 \cdot T(t) \cdot (1 - b_1 \cdot T(t)) - c_2 \cdot E(t) \cdot T(t) \\ &\quad - c_3 \cdot T(t) \cdot N(t) - a_2 \cdot (1 - e^{-M(t)}) \cdot T(t) \\ 0 &= r_2 \cdot N(t) \cdot (1 - b_2 \cdot N(t)) - c_4 \cdot T(t) \cdot N(t) \\ &\quad - a_3 \cdot (1 - e^{-M(t)}) \cdot N(t) \\ 0 &= v(t) - d_2 \cdot M(t) \end{aligned}$$

From the fourth the above equation), the value of $M(t)$ is zero, hence substituting this value in the remaining equations, and solving for each state variable we get:

$$E(t) = f_1(T(t)) = \frac{s \cdot (\alpha + T(t))}{(c_1 \cdot T(t) + d_1) \cdot (\alpha + T(t)) - \rho \cdot T(t)}$$

$$T(t) = \begin{cases} 0 \\ \frac{1}{b_1} - \left(\frac{c_2}{r_1 \cdot b_1}\right) \cdot E(t) - \left(\frac{c_3}{r_1 \cdot b_1}\right) \cdot N(t) \end{cases}$$

$$N(t) = \begin{cases} 0 \\ f_2(T(t)) = \frac{1}{b_2} - \left(\frac{c_4}{r_2 \cdot b_2}\right) \cdot T(t) \end{cases}$$

The solutions can be classified into three categories namely:

Tumor-free: In this group, normal cells persist when the tumor cell population $T(t)$ is zero. The equilibrium point's solution looks like this:

$$\left(\frac{s}{d_1}, 0, \frac{1}{b_2}\right)$$

Dead: When the normal cell population at an equilibrium point is zero, the point is categorized as dead. Two different kinds of dead equilibria may be inferred from the $T(t)$ equation:

Type 1: when both the normal $N(t)$ and tumor $T(t)$ cell populations are zero, i.e.:

$$\left(\frac{s}{d_1}, 0, 0\right)$$

Type 2: when the normal $N(t)$ cell population is zero, while the tumor cells have survived with a population of T' , and the effector cells steady state population is given by $f_1(T')$. The resultant equilibrium point is: $(f_1(T'), T', 0)T'$ is the nonnegative solution of the following equation:

$$T' - \frac{1}{b_1} + \left(\frac{c_2}{r_1 \cdot b_1}\right) \cdot f_1(T') = 0$$

Coexisting: This equilibrium point has nonzero population of both the tumor and normal cells, and is represented as:

$$(f_1(T''), T'', f_2(T''))$$

Where $f_1(T'')$ is the survived population of the effector cells while $f_2(T'')$ is the survived population of the normal cells, and T'' is the survived tumor population and it is the nonnegative solution of the following equation:

$$T'' - \frac{1}{b_1} + \left(\frac{c_2}{r_1 \cdot b_1}\right) \cdot f_1(T'') + \left(\frac{c_3}{r_1 \cdot b_1}\right) \cdot f_2(T'') = 0$$

Among the multiple equilibrium point of the system, the Tumor-free equilibria, is the desired point, since the remaining two equilibrium points do not solve the tumor problem.

Next, in order to formulate the cancer therapy system as a QRP, we first need to shift the equilibrium point to the origin, and this can be easily done by defining the following new set of variables [142]:

$$\theta(t) = \begin{bmatrix} \theta_1(t) \\ \theta_2(t) \\ \theta_3(t) \\ \theta_4(t) \end{bmatrix} = \begin{bmatrix} x_1(t) - x_{f,1} \\ x_2(t) - x_{f,2} \\ x_3(t) - x_{f,3} \\ x_4(t) - x_{f,4} \end{bmatrix}$$

Where

$$x(t) = [x_1(t) \ x_2(t) \ x_3(t) \ x_4(t)]^T \\ = [E(t) \ T(t) \ N(t) \ M(t)]^T$$

and $x_{f,i}$ is the i^{th} component of x_f , which is given by:

$$x_f = \begin{bmatrix} s/d_1 \\ 0 \\ 1/b_2 \\ 0 \end{bmatrix}$$

The variable θ_i is merely a placeholder, and for convenience, it will be replaced by x_i for $i = 1 \dots 4$. As a result, the dynamic equations of the shifted system can be expressed as follows:

$$\dot{x}_1(t) = \frac{\rho \cdot \left(x_1(t) + \frac{s}{d_1}\right) \cdot x_2(t)}{\alpha + x_2(t)} - c_1 \cdot \left(x_1(t) + \frac{s}{d_1}\right) \cdot x_2(t) \\ - d_1 \cdot x_1(t) - a_1 \cdot (1 - e^{-x_4(t)}) \\ \cdot \left(x_1(t) + \frac{s}{d_1}\right) + u_1(t)$$

$$\dot{x}_2(t) = r_1 \cdot x_2(t) \cdot (1 - b_1 \cdot x_2(t)) - c_2 \cdot \left(x_1(t) + \frac{s}{d_1}\right) \cdot x_2(t) \\ - c_3 \cdot x_2(t) \cdot \left(x_3(t) + \frac{1}{b_2}\right) - a_2 \\ \cdot (1 - e^{-x_4(t)}) \cdot x_2(t)$$

$$\dot{x}_3(t) = -r_2 \cdot x_3(t) \cdot (1 + b_2 \cdot x_3(t)) - c_4 \cdot x_2(t) \\ \cdot \left(x_3(t) + \frac{1}{b_2}\right) - a_3 \cdot (1 - e^{-x_4(t)}) \\ \cdot \left(x_3(t) + \frac{1}{b_2}\right)$$

$$\dot{x}_4(t) = u_2(t) - d_2 \cdot x_4(t)$$

B. Problem Formulation

We can proceed by defining the optimization criterion and the inequality constraints for the system [143]. As the cancer therapy system is being formulated as a QRP, the optimization criterion will take the following form:

$$J(u) = \frac{1}{2} \cdot \int_0^{t_f} [x(t)^T \cdot Q(t) \cdot x(t) + u(t)^T \cdot R(t) \cdot u(t)] dt$$

It is important to note that for practical implementation purposes, the upper limit of integration is not set to infinity. Instead, it is set to a large finite number, denoted as t_f .

The control inequality constraints define the upper and lower limits of the drug doses. In this case, the upper limit is represented by 1 and the lower limit by 0, as stated in reference [47]. These constraints can be expressed mathematically as follows:

$$0 \leq u(t) \leq 1, \quad t \in [0, \infty]$$

The $R(t)$ matrix is simply chosen to be constant diagonal matrix of the form: $R(t) = [R_1 \ 0 \ 0 \ R_2]$.

The $Q(t)$ matrix is also chosen to constant diagonal matrix, the resultant $Q(t)$ matrix is:

$$Q(t) = [Q_1 \ 0 \ 0 \ 0; \ 0 \ Q_2 \ 0 \ 0; \ 0 \ 0 \ Q_3 \ 0; \ 0 \ 0 \ 0 \ Q_4]$$

Please bear in mind that since the selected function $Q(t)$ cannot be expressed as a continuous mathematical function, the QRP must be solved in discrete time. There are several numerical techniques available for transcribing the continuous-time OCP, such as Runge-Kutta, Trapezoidal, Simpson's rule, among others.

However, for simplicity, the commonly used forward Euler approximation was employed, with an appropriate sampling time denoted as h .

The Euler methods, while simpler compared to higher-order numerical integration techniques like Runge-Kutta, can offer some advantages when solving optimal control problems for cancer treatment:

- The Euler methods require fewer function evaluations per time step compared to Runge-Kutta methods, making them computationally more efficient. This can be particularly beneficial for optimal control problems with complex cancer treatment models, where the computational cost of the numerical integration is a significant factor.
- Implicit Euler methods are generally more stable and can handle stiff ODE systems more robustly than Runge-Kutta methods. Cancer treatment models often involve stiff ODE systems due to the complex interactions between tumor dynamics, drug pharmacokinetics, and various biological processes. The improved stability of implicit Euler methods can make them more suitable for these types of problems.
- The Euler methods, particularly the implicit versions, can better capture discontinuities in the control inputs or the state variables, which are common in optimal control problems for cancer treatment (e.g., on-off drug administration, maximum drug dose constraints).

Improved handling of discontinuities can lead to more accurate solutions and better convergence of the optimal control problem.

- Runge-Kutta methods are not self-starting, meaning they require an initial condition or starting values to begin the integration process. In cancer treatment optimization, the initial state of the patient (e.g., tumor size, and drug concentrations) may have a significant impact on the optimal control strategy. The need for accurate initial conditions can make the optimal control problem more

sensitive to uncertainties in the patient's initial state, which can be a challenge in real-world applications.

To make sure the calculated control is correct and dependable, one can validate the computed optimum control by comparing it to experimental data or using other computational techniques. Consequently, the discrete QRP formulation for the cancer chemo-immune therapy system is as follows:

$$\begin{aligned} \min. u \quad J(u) &= \frac{1}{2} \cdot \sum_{k=0}^{N-1} x(k)^T \cdot Q(k) \cdot x(k) + u(k)^T \cdot R \\ &\cdot u(k) \text{ subject to: } x(k+1) \\ &= x(k) + h \\ &\cdot (A(k) \cdot x(k) + B(k) \cdot u(k)) \quad 0 \leq u(k) \\ &\leq 1 \} x(0) = x_0 \end{aligned}$$

In applying the QRP framework to the field of cancer therapy, we consider two distinct case studies, each with different therapeutic objectives based on patient profiles. These case studies help demonstrate how QRPs can be adapted to address various clinical scenarios by adjusting the parameters within the model, specifically the weighting matrices in the optimal control problem [144]. Here's a restatement and elaboration of these case studies:

Case Study 1: Young Patient with Cancer

- Patient Profile: A young individual diagnosed with cancer, with no other resistant diseases.
- Therapeutic Focus: The primary goal in this scenario is to reduce the population of cancerous cells. The young age of the patient implies a better ability of their body to compensate for any decrease in normal and immune cells that might occur due to chemotherapy.
- Optimal Control Objective: From an optimal control standpoint, the emphasis is on minimizing the number of tumor cells. This is achieved by adjusting the weighting matrices in the QRP formulation to prioritize the reduction of cancerous cells, while taking into account the resilience of the patient's body to withstand the treatment's adverse effects [145].

Case Study 2: Elderly Patient with Cancer and Co-morbidities

- Patient Profile: An elderly individual diagnosed with cancer, who also suffers from other refractory conditions, such as cardiac disease.
- Therapeutic Focus: Due to the heightened risk associated with the destruction of normal and immune cells in elderly patients, particularly those with co-morbidities, the focus shifts towards immunotherapy rather than chemotherapy. In this case, preserving normal and immune cells is prioritized over aggressively reducing cancerous cells.
- Optimal Control Objective: The optimal control problem is formulated to reflect this shift in priorities. The weighting matrices are adjusted to emphasize the preservation of normal and immune cells. This approach aims to balance the need to control the cancer with the

need to maintain overall health and quality of life, given the patient's age and additional health concerns [146].

In both cases, the QRP framework provides a systematic way to tailor the control strategies (therapeutic interventions) according to patient-specific factors. By adjusting the weights in the cost function of the QRP, the model can prioritize different aspects of the treatment, such as reducing tumor size or preserving healthy cells, in alignment with the clinical objectives for each patient. This demonstrates the flexibility and applicability of the QRP approach in creating personalized treatment strategies in cancer therapy [147].

C. Problem Solution

The application of OCP techniques to solve the Quadratic Regularization Problem (QRP) in the context of cancer therapy involves a systematic approach, utilizing predefined parameters and evaluating different treatment strategies. Here's an overview of how this process unfolds:

1) Setting Up the Parameters

- a. Parameter Values: The problem is defined with specific numerical values for various parameters, which are dimensionless. These values are listed in Table III.
- b. Normalized Parameters: Table IV provides both absolute and normalized values of these parameters. The normalization helps in standardizing the parameters, making them more comparable and manageable [148].
- c. Fixed Parameters Across Techniques: To facilitate a fair comparison of different techniques, these parameter values remain constant for all the methods applied [149].

2) Application of Techniques

- a. Technique Repetition: Each OCP technique is applied four times in total, with two applications per case study.
- b. Treatment Categories:
 - Continuous Treatment: Represents scenarios where treatment is administered daily.
 - Dosed Treatment: Encompasses cases where treatment is administered periodically, for instance, every 21 or 16 days.

3) Visualization and Analysis

- a. System States in Plots: For each application of a technique, a plot or figure is generated. This figure includes the trajectories of the four states of the system, which might be CD8+ and NK T cells, tumor cells, and chemotherapy drug concentration [150].
- b. Control Therapies Representation: The two control therapies (immunotherapy and chemotherapy) are also represented in these plots, showing how they vary over time under different strategies.
- c. Optimization Criterion Visualization: The plots also illustrate the optimization criterion of the OCP, providing insights into how the chosen control strategy aims to optimize the treatment outcome [151].

4) Comparative Analysis

- The performance of each technique is compared based on these plots and the optimization outcomes [152][153]. This comparison is crucial for understanding the effectiveness of different control strategies in the two distinct case studies (younger patient with continuous chemotherapy vs. elderly patient with periodic dosed treatment) [154].
- This structured approach allows for a comprehensive analysis of how different OCP techniques can be employed in cancer therapy. It provides valuable insights into the efficacy of various treatment strategies, taking into account the specific needs and conditions of different patient profiles. The visual representation of the system states and control therapies under different scenarios is instrumental in understanding the dynamics of cancer treatment and the role of optimal control in improving patient outcomes [155].

D. SNAC Technique

To implementation of the SNAC technique as discussed is quite involved. The first step is to determine the pre training policy to initialize the weights and biases of the critic network, the pre training policy must be able to generate the next constate variable λ_{k+1} for the current state variable x_k , and it must be able to do this to any arbitrary value of the state x_k in other words it should be possible to write this policy as:

$$\lambda_{k+1} = \lambda_{k+1}(x_k)$$

A common pre training policy that is easy to compute for any state x_k is the one generated by applying the LQR technique as in (19). To apply this technique the dynamic equations of the shifted system must first be linearized around the desired equilibrium point using the famous Jacobian relations to find the A and B matrices as follows:

$$A(x) = \begin{bmatrix} a_{11} & a_{12} & 0 & a_{14} \\ a_{21} & a_{22} & 0 & 0 \\ 0 & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}, B(x) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$a_{11} = -d_1, a_{12} = \left(\frac{\rho}{\alpha} - c_1\right) \cdot \frac{s}{d_1}, a_{14} = -a_1 \cdot \frac{s}{d_1}$$

$$a_{22} = r_1 - \frac{c_3}{b_2} - \frac{c_2 \cdot s}{d_1}, a_{32} = -\frac{c_4}{b_2}$$

$$a_{33} = -r_2, a_{34} = -\frac{a_3}{b_2}, a_{44} = -d_2$$

The R and Q needed for the LQR technique were set to the static matrices of in Table III, the resultant ARE of (22) were solved for the constant P matrix of (21) which was directly used for the training of the critic network of the SNAC technique. Four different RBFs were used to construct the critic network, one for each costate variable. Note that four networks each having a single output is generally better than single network with four outputs, this is due to the fact that the convergence of the costate variables can take different rates which we cannot know since they have no physical meaning, hence the use of single network for each costate variable eliminates the effect of different rate

convergence problem. Each RBF network has as inputs the four states, and twenty nodes in the hidden layer. The costate and optimal control equations needed for the implementation of the SNAC technique are derived below, which can be found by direct application of the equations of (20), the costate equations are summarized in the following form (where for convience the time dependency was dropped).

$$\dot{\lambda}^*(t) = -Q(t) \cdot x(t) - \left[\frac{\partial a(x^*(t), u^*(t))}{\partial x(t)} \right]^T \cdot \lambda^*(t)$$

$$\lambda_1^* = -q_{11} \cdot x_1 + \lambda_1^* \cdot \left(c_1 \cdot x_2 + d_1 + a_1 \cdot (1 - e^{-x_4}) - \frac{\rho \cdot x_2}{\alpha + x_2} \right) + \lambda_2^* \cdot (c_2 \cdot x_1)$$

$$\lambda_2^* = -q_{22} \cdot x_2 + \lambda_2^* \cdot \left(2 \cdot r_1 \cdot b_1 \cdot x_2 + c_2 \cdot \left(x_1 + \frac{s}{d_1} \right) + c_3 \cdot \left(x_3 + \frac{1}{b_2} \right) + a_2 \cdot (1 - e^{-x_4}) - r_1 \right) + \lambda_1^* \cdot \left(c_1 \cdot x_1 - \frac{\alpha \cdot \rho \cdot x_1}{(\alpha + x_2)^2} \right) + \lambda_3^* \cdot \left(c_4 \cdot \left(x_3 + \frac{1}{b_2} \right) \right)$$

$$\lambda_3^* = -q_{33} \cdot x_3 + \lambda_2^* \cdot (c_3 \cdot x_2) + \lambda_3^* \cdot (r_2 + 2 \cdot r_2 \cdot b_2 \cdot x_3 + c_4 \cdot x_2 \cdot a_3 \cdot (1 - e^{-x_4}))$$

$$\lambda_4^* = -q_{44} \cdot x_4 + \lambda_1^* \cdot \left(a_1 \cdot \left(x_1 + \frac{s}{d_1} \right) \cdot e^{-x_4} \right) + \lambda_2^* \cdot (a_2 \cdot x_2 \cdot e^{-x_4}) + \lambda_3^* \cdot \left(a_3 \cdot \left(x_3 + \frac{1}{b_2} \right) \cdot e^{-x_4} \right) + \lambda_4^* \cdot (d_2)$$

The optimal control equations (without taking the inequality constraints limits into consideration):

$$u^*(t) = -R(t)^{-1} \cdot \left[\frac{\partial a(x(t), u(t))}{\partial u(t)} \right]^T \cdot \lambda^*(t)$$

$$u_1^*(t) = \frac{1}{r_{11}} \cdot \lambda_1^*(t) \text{ and } u_2^*(t) = \frac{1}{r_{22}} \cdot \lambda_4^*(t)$$

In discrete form (and taking the inequality constraints into consideration):

$$u_1^*(k) = \begin{cases} 0 & -\frac{1}{r_{11}} \cdot \lambda_1^*(k+1) < 0 \\ -\frac{1}{r_{11}} \cdot \lambda_1^*(k+1) & \text{otherwise} \\ 1 & -\frac{1}{r_{11}} \cdot \lambda_1^*(k+1) > 1 \end{cases}$$

$$u_2^*(k) = \begin{cases} 0 & -\frac{1}{r_{22}} \cdot \lambda_4^*(k+1) < 0 \\ -\frac{1}{r_{22}} \cdot \lambda_4^*(k+1) & \text{otherwise} \\ 1 & -\frac{1}{r_{22}} \cdot \lambda_4^*(k+1) > 1 \end{cases}$$

The training process took a total of thirty iterations. At each iteration a new dataset containing 2000 state vectors (in the admissible region) is randomly generated, and all the networks were trained to recognize this set for 1000 epochs. The final performance measure for all the networks were less than 10^{-7} . The results of applying this technique on the two case studies are summarized in Fig. 5, Fig. 6, Fig. 7, and Fig. 8.

TABLE III. THE EXACT NUMERICAL VALUES OF THE PARAMETERS OF THE CANCER THERAPY OCP

| Parameter | Description | Case 1 | Case 2 |
|------------|---------------------------------|--|---|
| N | Number of sample points | 2000 | 2000 |
| h | Sampling period | 0.05 [day] | 0.05 [day] |
| x_0 | Initial states | $[0.15, 1, 1, 0.1]^T$ | $[0.15, 1, 1, 0.1]^T$ |
| N_m | Lower limit of NK cells | 0.3 | 0.6 |
| R | Control weighing matrix | $diag([1, 1])$ | $diag([1, 1])$ |
| Q static | Static state weighting matrix | $diag([10, 100, 10, 0.01])$ | $diag([10, 20, 10, 0.01])$ |
| μ | NK dynamic weight | 100 | 100 |
| δ | NK maximum constraint deviation | 0.05 | 0.05 |
| - | Dose period | Q21D. (Latin abbreviation for once every 21 days) [88] | Q16D. (Latin abbreviation for once every 16 days) |

The SNAC approach is a promising method for solving the optimization problem in cancer treatment, but its complexity and computational requirements raise concerns about its practical feasibility and efficiency in real-time applications. The SNAC approach has several advantages for solving the optimization problem in cancer treatment:

- SNAC can handle complex and dynamic systems, allowing it to adapt to changing conditions and optimize treatment strategies accordingly.
- The stochastic nature of SNAC makes it more robust to uncertainties and noise in the data, which is common in cancer treatment.
- SNAC can be applied to various types of cancer and treatment modalities, making it a versatile approach for optimizing cancer treatment.
- SNAC can be tailored to individual patients, taking into account their unique characteristics and treatment responses.
- SNAC can adapt to changing conditions in real-time, allowing for more effective and efficient treatment strategies.
- SNAC has been shown to improve treatment outcomes by optimizing treatment schedules and reducing side effects.

Reduced Drug Usage: SNAC can reduce the amount of drugs used in treatment, which can lead to fewer side effects and improved patient quality of life.

- SNAC can enhance patient outcomes by optimizing treatment strategies and improving treatment efficacy.
- SNAC can identify optimal treatment strategies by considering multiple objectives and constraints, leading to more effective treatment plans.

E. LQR Technique

The final technique applied in this context is the LQR, which is implemented by applying (19) to the system dynamics represented by matrices A and B, and incorporating the static weighting matrices Q and R, as detailed in Table III. The results of this application are summarized in Fig. 9 to Fig. 12. The final technique applied in this context is the LQR, which is implemented by applying (19) to the system dynamics represented by matrices A and B, and incorporating the static weighting matrices Q and R, as detailed in Table V.

In order to fully understand and interpret these results, it's important to distinguish between the two types of therapies applied in the studies [23][56]:

Continuous Therapy: This is more of a theoretical approach, included in the study for mathematical purposes. It's not practically feasible in real-world applications, but it represents the actual solution to the OCP of the discrete Quadratic Regularization Problem (QRP) for the cancer chemo-immune therapy system. Continuous therapy serves as a benchmark for comparing the performance of various studied techniques.

Dosed Therapy: This is the more practical approach used in actual treatment scenarios, involving periodic administration of therapy (e.g., every 21 or 16 days). It does not represent the solution to the OCP of the QRP as precisely because it lacks the dosed constraint on the control signals in the QRP formulation.

The analysis of the results revealed that all optimal control techniques produced almost identical control sequences for both control variables u_1 and u_2 . The SNAC technique stood out by generating the best response in terms of the optimization criterion, owing to its use of advanced heuristic techniques for computing the optimal solution. The LQR techniques, while not achieving the optimal solution, provided a highly stable and robust suboptimal solution due to their reliance on extended linearization theory.

When examining the continuous therapy solutions for both case studies, an interesting observation was noted in the treatment policies generated by the SNAC technique. Despite having identical control weighting matrices R for both cases, the optimal control techniques exhibited noticeable differences in response. In Case 1, focusing on a young patient, the techniques prioritized rapid tumor eradication.

In Case 2, involving an elderly patient with additional health issues, there was a shift towards balancing tumor control with the preservation of immune cells, reflected in the different formulations of the state weighting matrix Q for each case. Table V provides a comprehensive summary of the final values of the state variables and the optimization criterion for all techniques studied, covering both continuous therapy after 20 days and dosed therapy after 90 days, offering a clear comparison of outcomes under different treatment approaches.

TABLE IV. MODEL PARAMETERS VALUES

| Equation | Parameter | Value | Unit | Normalized Value |
|-----------|-----------|------------------------|-------------------------|------------------|
| \dot{E} | s | – | Day^{-1} | 0.33 |
| | ρ | $1.245 \cdot 10^{-2}$ | Day^{-1} | 0.01 |
| | α | $2.5036 \cdot 10^{-3}$ | lU/l^{-1} | 0.3 |
| | c_1 | $3.422 \cdot 10^{-10}$ | $cells^{-1}/Day^{-1}$ | 1 |
| | d_1 | $9 \cdot 10^{-3}$ | Day^{-1} | 0.2 |
| | a_1 | $4.86 \cdot 10^{-2}$ | Day^{-1} | 0.2 |
| \dot{T} | r_1 | $4.31 \cdot 10^{-1}$ | Day^{-1} | 1.5 |
| | b_1 | $1.02 \cdot 10^{-9}$ | Day^{-1} | 1.0 |
| | c_2 | – | Day^{-1} | 0.5 |
| | c_3 | $2.9077 \cdot 10^{-3}$ | $l/cells^{-1}/Day^{-1}$ | 1.0 |
| | a_2 | $9 \cdot 10^{-1}$ | Day^{-1} | 0.3 |
| \dot{N} | r_2 | – | Day^{-1} | 1.0 |
| | b_2 | $1.25 \cdot 10^{-2}$ | Day^{-1} | 1.0 |
| | c_4 | $2.794 \cdot 10^{-13}$ | $cells^{-1}/Day^{-1}$ | 1.0 |
| | a_3 | $6.75 \cdot 10^{-2}$ | Day^{-1} | 0.1 |
| \dot{M} | d_2 | $5.199 \cdot 10^{-1}$ | Day^{-1} | 0.2 |

TABLE V. TERMINATING VALUES OF THE OPTIMIZATION CRITERION AND THE STATE VARIABLES FOR THE TWO TREATMENT PROTOCOLS. (C) FOR CONTINUOUS AND (D) FOR DOSED

| Variable | Therapy | Case 1 | | | Case 2 | | |
|------------------------|---------|----------|----------|-------|----------|----------|-------|
| | | SNAC | LQR | x_f | SNAC | LQR | x_f |
| Optimization criterion | C | 52.3585 | 56.6678 | - | 20.0454 | 20.369 | - |
| | D | 584.1334 | 686.5033 | - | 241.4729 | 261.7581 | - |
| CD8+ T cells | C | 1.6514 | 1.6499 | 1.65 | 1.6502 | 1.6499 | 1.65 |
| | D | 1.6503 | 1.6488 | 1.65 | 1.6498 | 1.65 | 1.65 |
| Tumor cells | C | 0.0007 | 0.001 | 0 | 0.0009 | 0.001 | 0 |
| | D | 0 | 0.0001 | 0 | 0 | 0 | 0 |
| NK T cells | C | 0.999 | 0.9985 | 1 | 0.9987 | 0.9985 | 1 |
| | D | 1 | 0.9999 | 1 | 1 | 1 | 1 |
| Chemotherapy drug | C | 0 | 0 | 0 | 0 | 0 | 0 |
| | D | 0 | 0 | 0 | 0 | 0 | 0 |

SNAC and LQR are two distinct techniques used in clinical settings to develop treatment policies for patients. Both methods have their own strengths and limitations, which are crucial to understand in the context of patient care, particularly when considering the differences in treatment policies for young and elderly patients.

5) SNAC Clinical Significance and Potential Impact on Patient Care

- Personalized Treatment: SNAC is a data-driven approach that uses machine learning to develop personalized treatment policies based on individual patient data. This allows for more effective and efficient treatment, as it takes into account the unique characteristics and needs of each patient.
- Adaptability: SNAC's adaptive nature enables it to adjust treatment policies in real-time as new data becomes available. This adaptability is particularly important in the context of elderly patients, who may have complex and changing health needs.
- Improved Patient Outcomes: By using SNAC, clinicians can develop treatment policies that are tailored to the specific needs of each patient, leading to improved patient outcomes and enhanced quality of life.

6) LQR Clinical Significance and Potential Impact on Patient Care

- Linear Control: LQR is a linear control technique that uses a mathematical model to develop treatment policies. While it is effective for simple systems, its limitations become apparent when dealing with complex, non-linear systems like those encountered in clinical practice.
- Limited Adaptability: LQR is a fixed-policy approach that does not adapt to changing patient conditions. This can lead to suboptimal treatment outcomes, particularly for elderly patients who may have rapidly changing health needs.
- Limited Generalizability: LQR is typically developed using a specific dataset and may not generalize well to other patient populations or settings. This can limit its applicability and effectiveness in clinical practice.

7) Comparison of SNAC and LQR in the Context of Young and Elderly Patients

- Young Patients: For young patients, both SNAC and LQR can be effective in developing treatment policies. However, SNAC's adaptability and personalized approach may be particularly beneficial in this population, as they are more likely to have complex and changing health needs.
- Elderly Patients: For elderly patients, SNAC's adaptability and personalized approach are crucial in

developing effective treatment policies. LQR's fixed-policy approach may not be suitable for this population, as their health needs can change rapidly and unpredictably.

The key findings from the analysis are:

1. **Convergence Rates:** The results shows that the Policy Iteration (PI) strategy typically converges faster than the Value Iteration (VI) strategy. The maximum absolute change in the value function decreases more rapidly for PI compared to VI over the time horizon.
2. **Computational Efficiency:** While the convergence rate is faster for PI, it requires additional computations for the policy evaluation and policy improvement steps at each iteration. VI, on the other hand, has a simpler update rule and may be more computationally efficient, especially for larger problem sizes.
3. **Solution Quality:** The final solutions obtained from both VI and PI are expected to be of similar quality, as they both converge to the optimal value function and control policy under the given problem formulation. However, the rate of convergence may impact the practical implementation, especially for real-time applications with time constraints.

The choice between VI and PI for a specific cancer treatment application would depend on factors such as the problem size, the required solution accuracy, the available computational resources, and the time constraints for the decision-making process. In some cases, a hybrid approach that combines the strengths of both methods may be beneficial. Empirical evaluation on a range of realistic scenarios would be necessary to provide a more

comprehensive assessment of the relative performance of VI and PI strategies.

To address the challenges of computational complexity, sensitivity to initial conditions, and convergence problems in the practical implementation of numerical solvers for OCT in cancer treatment, the following strategies can be employed:

1. Develop reduced-order models that capture the essential dynamics of the cancer treatment problem while significantly reducing the computational complexity. Use techniques like proper orthogonal decomposition (POD) or balanced truncation to construct low-dimensional models.
2. Design the optimal control-based treatment protocols to seamlessly integrate with the existing cancer treatment infrastructure and practices and engage with cancer treatment experts to ensure the practical feasibility and clinical relevance of the optimal control solutions.
3. Discretize the control and state variables, and solve the resulting nonlinear programming problem. Also, derive the necessary optimality conditions and solve the resulting boundary value problem. Employ min-max optimization to find solutions that are optimal under the worst-case realization of the uncertain parameters.
4. Run the numerical solver from multiple, randomly generated initial guesses to increase the chances of finding the global optimum. Use clustering techniques to identify distinct locally optimal solutions and select the most promising one. Also, dynamically adjust the step size of the numerical solver to ensure stable and reliable convergence.

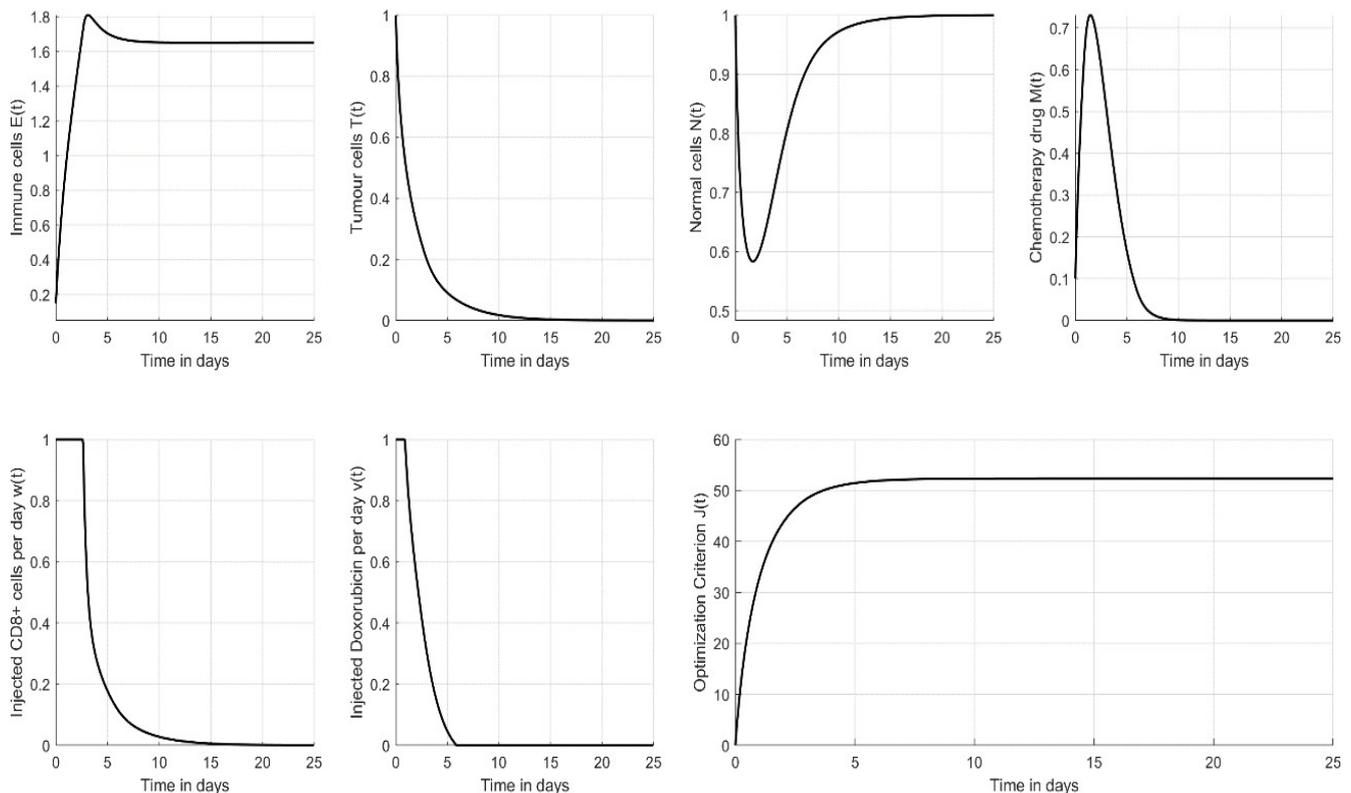


Fig. 5. Continuous therapy solution using SNAC technique for case 1

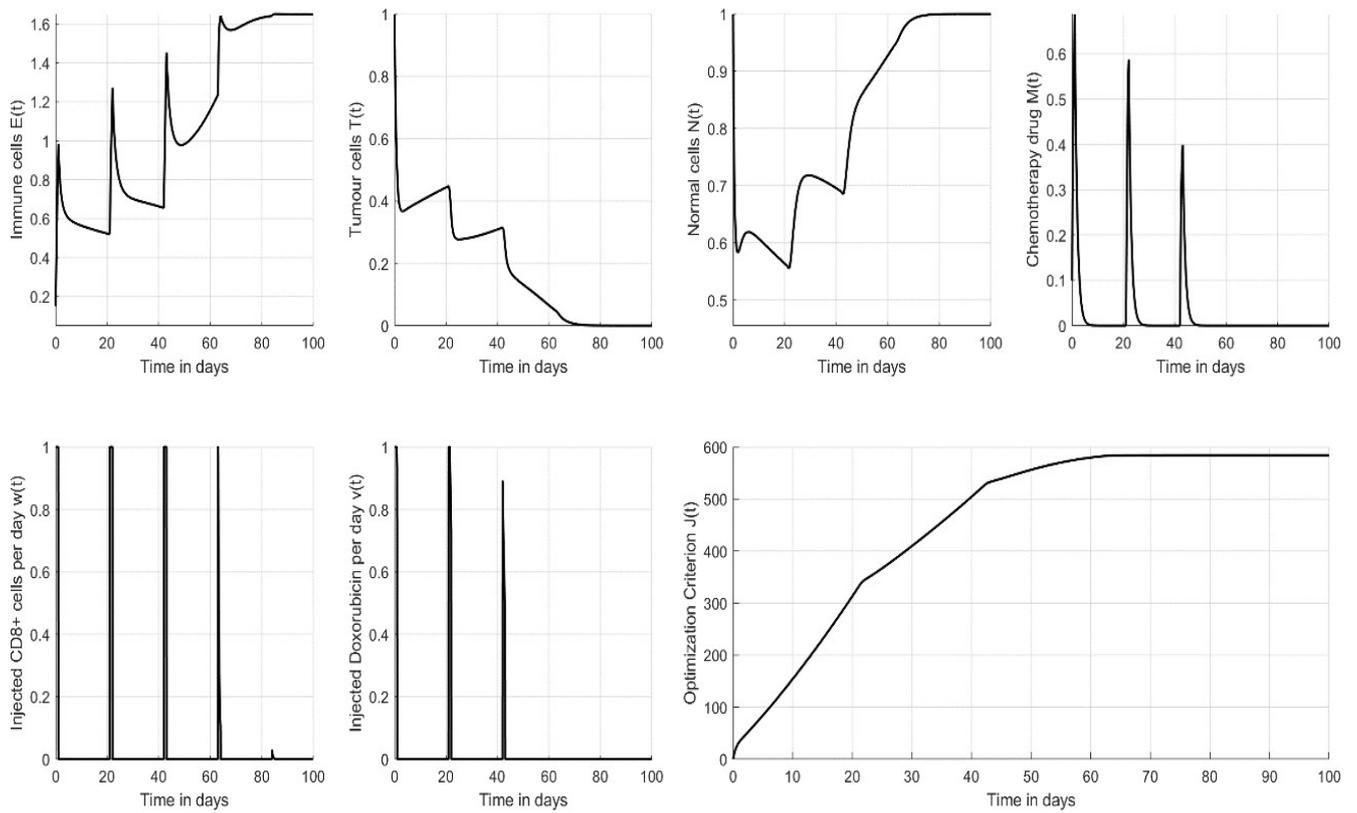


Fig. 6. Dosed therapy solution using SNAC technique for case 1

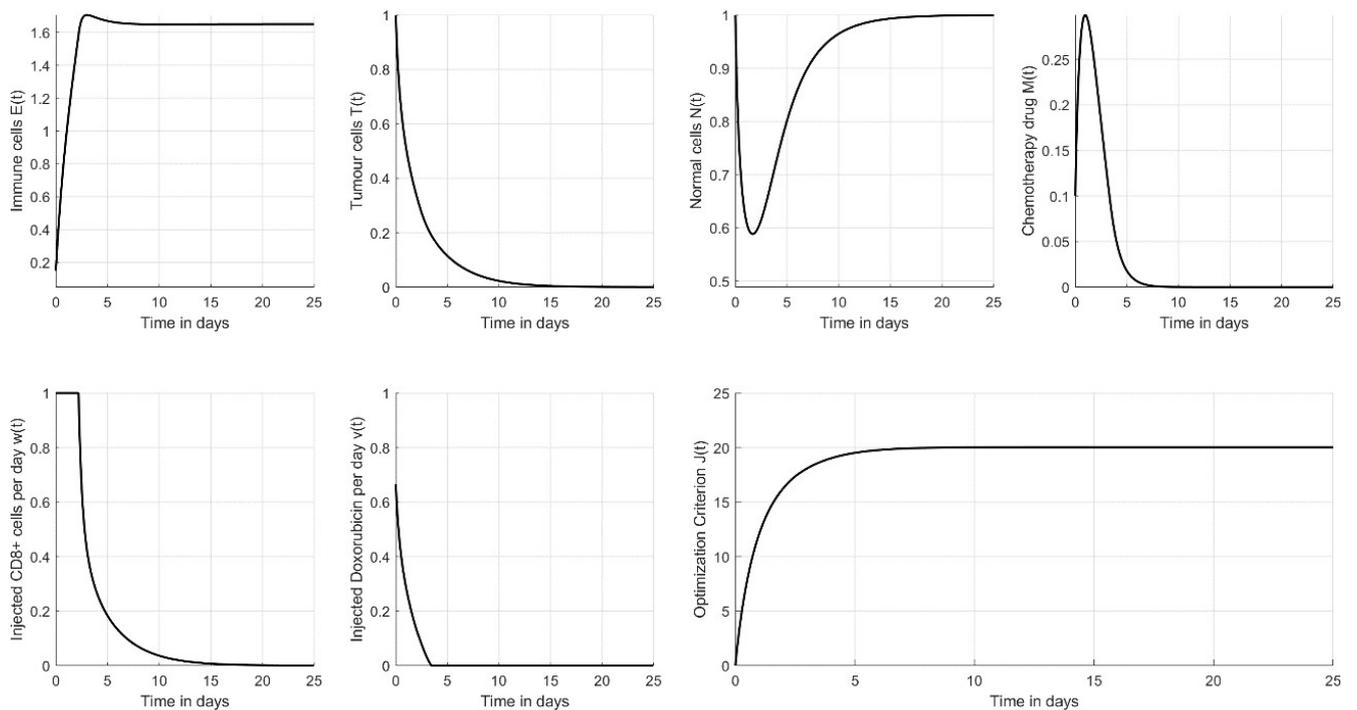


Fig. 7. Continuous therapy solution using SNAC technique for case 2

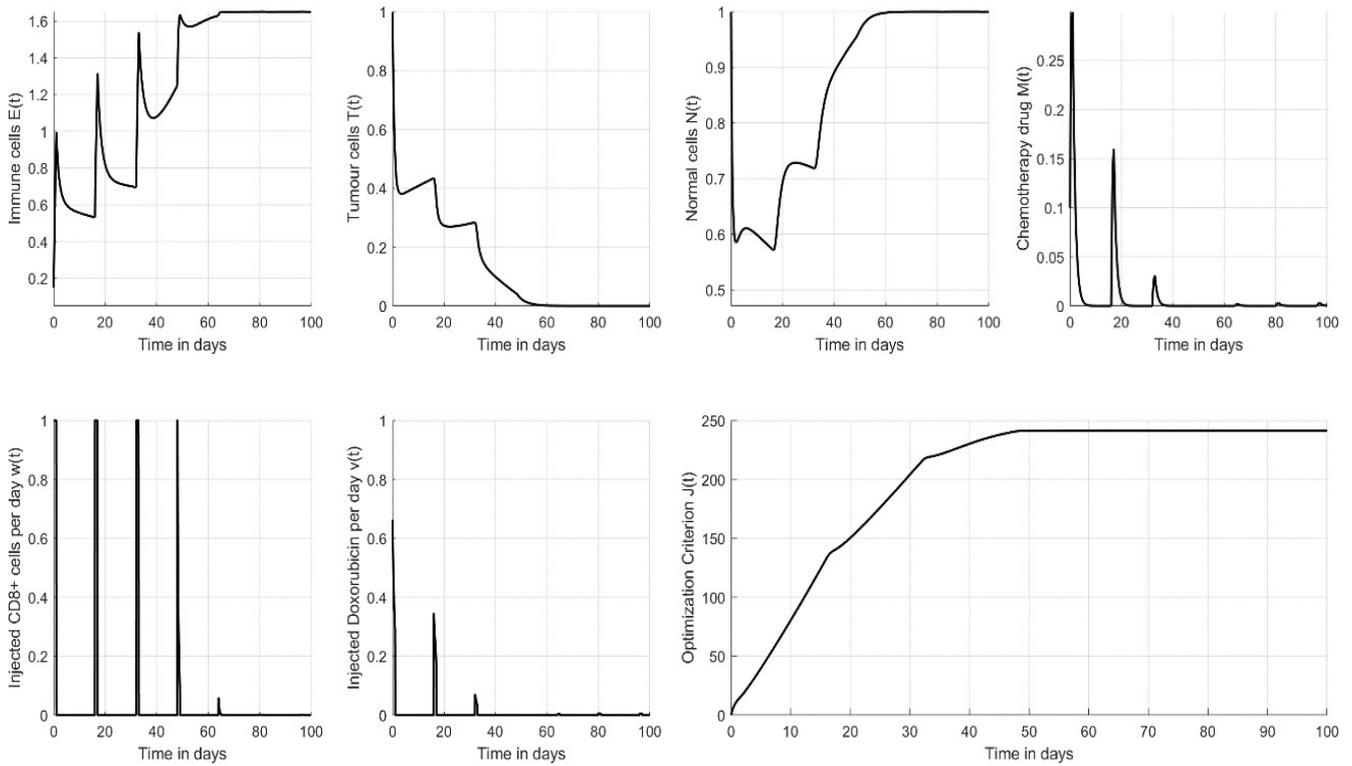


Fig. 8. Dosed therapy solution using SNAC technique for case 2

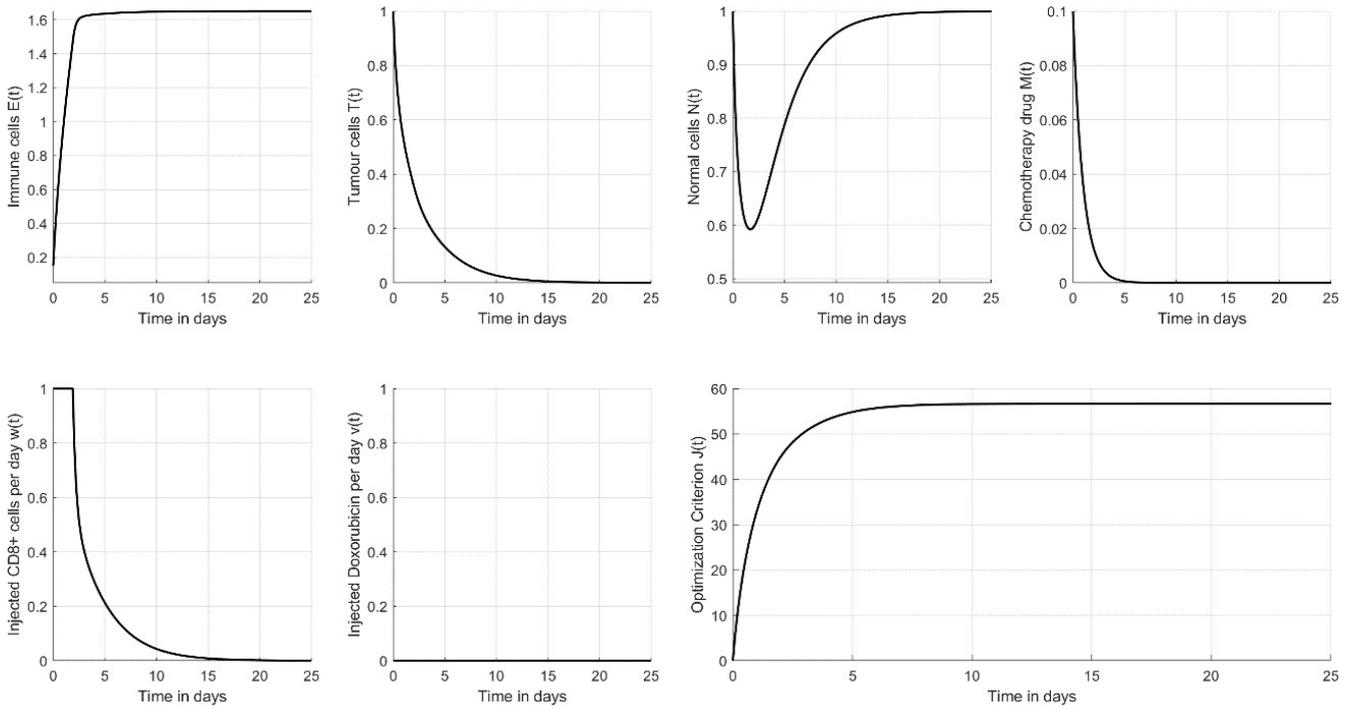


Fig. 9. Continuous therapy solution using LQR technique for case 1

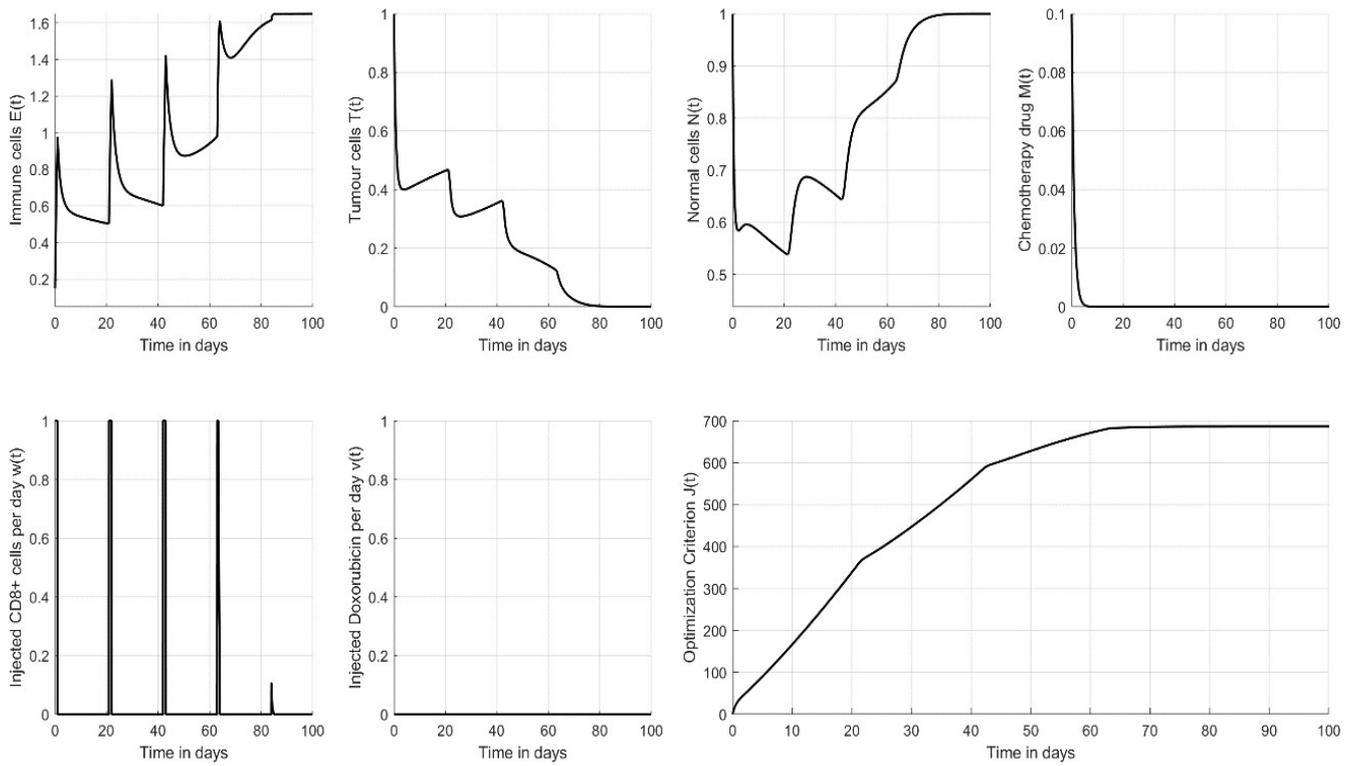


Fig. 10. Dosed therapy solution using LQR technique for case 1

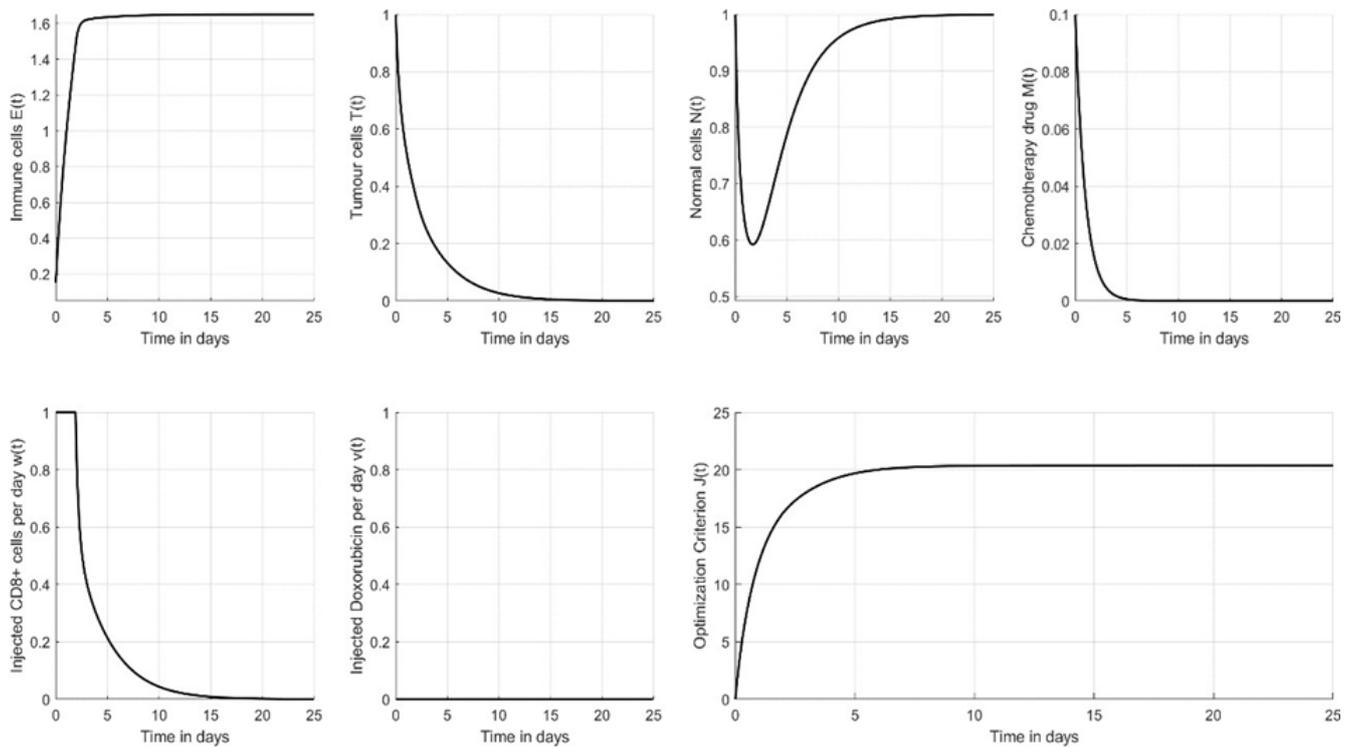


Fig. 11. Continuous therapy solution using LQR technique for case 2

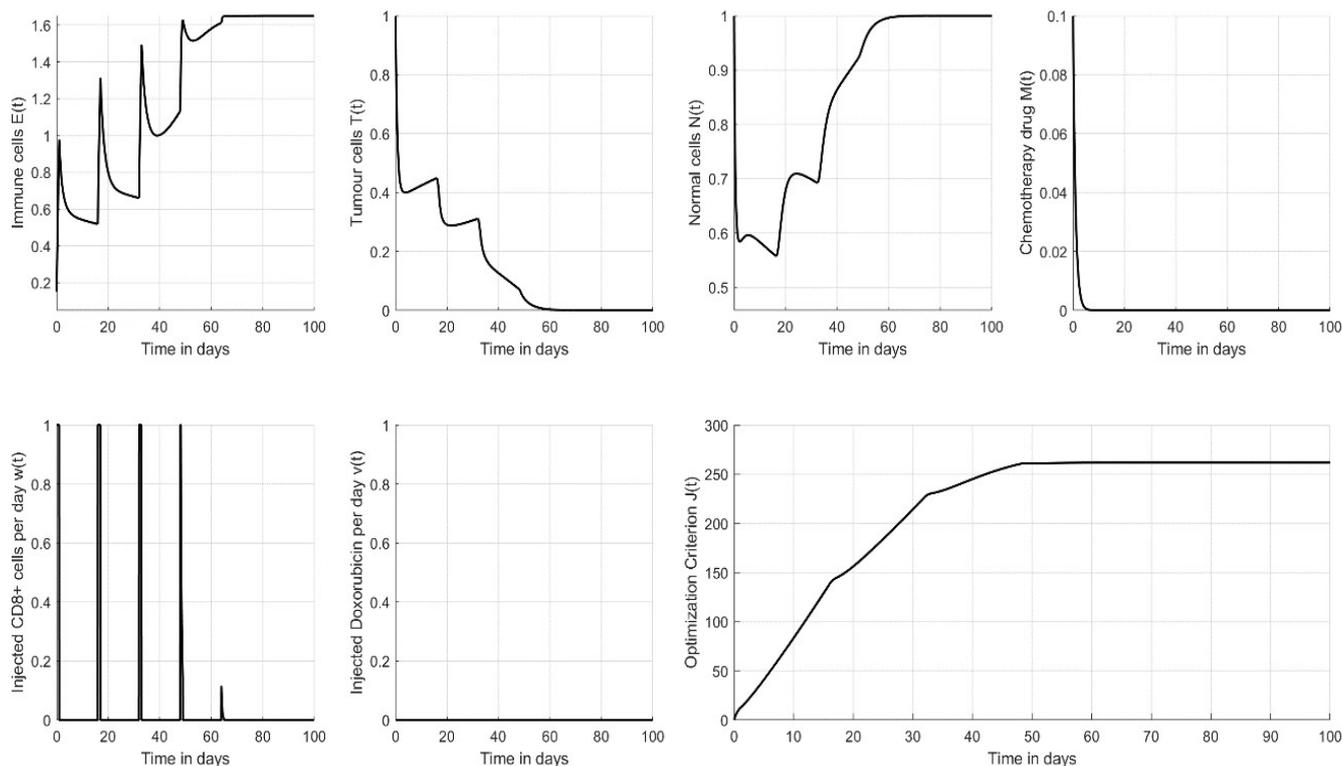


Fig. 12. Dosed therapy solution using LQR technique for case 2

In general, to address the lack of experimental validation, some potential steps could include:

- Collaborate with clinicians and researchers to design and conduct meaningful preclinical and clinical studies to validate the optimal control-based treatment approach.
- Revisit the underlying optimal control theory and mathematical models, and refine them based on available experimental data and feedback from the clinical community. This iterative process can help improve the model's accuracy and relevance to real-world scenarios.
- Stepwise approach to validation, starts with in vitro experiments, then progressing to animal studies, and eventually moving towards early-phase clinical trials. This gradual validation process can help build confidence in the approach and identify any limitations or challenges before large-scale clinical trials.
- Engage with experts in various fields, such as systems biology, computational biology, and control engineering, to leverage their expertise and gain new perspectives on the optimal control-based approach to cancer treatment.

IV. CONCLUSION

Cancer is an extremely complex and heterogeneous disease and can vary significantly between patients and even within the same tumor. This paper makes significant strides in modeling cancer progression and the immune system's response, integrating both chemotherapy and immunotherapy, with a particular focus on LQR and SNAC, establishing foundational principles of OCT. These principles encompass problem formulation, tackling inequality constraints, and elucidating the necessary conditions for achieving optimal solutions.

In summary, SNAC and LQR techniques have strengths and weaknesses in cancer treatment for young and elderly patients. Low side effects of SNAC, low cost, high patient compliance, high treatment flexibility, and high real-time adaptability are advantages but in return, long training time, amount of high-quality data for training, high patient-specific modeling, and high physiological variations are shortcomings. In this technique, model validation and clinical implementation can be challenging due to the complexity of the model and the need for extensive testing and validation. However, LQR's fixed treatment duration, increased side effects, higher cost, lower patient compliance, low treatment flexibility, low model complexity, low computational requirements, low real-time adaptability, low patient-specific modeling, low physiological variations, and low treatment personalization make it less effective for elderly patients.

Through this comparison and the results obtained, we conclude that SNAC technique has better effectiveness than LQR in treating cancer, and we look in the future forward to comparing this technique with other new techniques to overcome the shortcomings. As the healthcare landscape continues to evolve, it is essential to consider the clinical significance and potential impact of these techniques on patient care. This may be achieved through more research and effort in the future.

V. FUTURE WORK

The future work for cancer treatment involves several directions, including:

Future research should focus on developing more robust optimization techniques that consider the dynamic nature of cancer progression and treatment responses. These techniques should be based on mathematical models that

accurately capture the dynamic interactions between cancer cells and treatment modalities.

Incorporating spatial characteristics of tumors using advanced methods like Linear Time-Varying (LTV) approximation to better capture the dynamics of cancer growth and treatment response.

Exploring immunotherapy approaches that can enhance the body's natural immune response to cancer, such as checkpoint inhibitors and cancer vaccines.

Developing targeted therapies that can specifically target cancer cells and minimize harm to healthy cells, such as precision medicine approaches.

Developing real-time monitoring systems that can track the progression of cancer and treatment response, allowing for more effective and timely interventions.

ACKNOWLEDGMENT

This research was funded by Universitas Sebelas Maret, Research Grant: Penelitian Disertasi Doktor (PDD) with contract number 194.2/ UN27.22/PT.01.03/2024.

REFERENCES

- [1] J. Z. Shing *et al.*, "HPV-Associated Cancer Incidence by Disaggregated Asian American, Native Hawaiian, and Other Pacific Islander Ethnicity," *JNCI Cancer Spectrum*, vol. 7, no. 2, 2023, doi: 10.1093/jncics/pkad012.
- [2] E. Miloud *et al.*, "A Perfect Challenge to Select the Ideal Virus Vaccine Using a New Potent Hierarchical Algorithm in a Smart Laboratory," *IEEE access*, vol. 11, pp. 79890–79910, 2023, doi: 10.1109/access.2023.3295329.
- [3] J. Galon and D. Daniela, "Tumor Immunology and Tumor Evolution: Intertwined Histories," *Immunity*, vol. 52, no. 1, pp. 55–81, 2020, doi: 10.1016/j.immuni.2019.12.018.
- [4] S. Morgan *et al.*, "Hypofractionated Radiation Therapy for Localized Prostate Cancer: Executive Summary of an ASTRO, ASCO and AUA Evidence-Based Guideline," *Journal of Urology*, vol. 201, no. 3, pp. 528–534, 2019, doi: 10.1097/JU.000000000000071.
- [5] M. Alqudah, "Cancer treatment by stem cells and chemotherapy as a mathematical model with numerical simulations," *Alexandria Engineering Journal*, vol. 59, pp. 1953–1957, 2020, doi: 10.1016/j.aej.2019.12.025.
- [6] W. Wu *et al.*, "Competitive risk analysis of prognosis in patients with cecum cancer: A population-based study," *Cancer Control*, vol. 28, 2021, doi: 10.1177/1073274821989316.
- [7] X. Wang, G. Yu, and Z. Yan, "Lung cancer subtype diagnosis by fusing image-genomics data and hybrid deep networks," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 20, no. 1, pp. 512–523, 2023, doi: 10.1109/TCBB.2021.3132292.
- [8] L. Pang *et al.*, "Mathematical modeling and dynamic analysis of anti-tumor immune response," *Journal of Applied Mathematics and Computing*, vol. 62, pp. 473–488, 2020, doi: 10.1007/s12190-019-01292-9.
- [9] H. Song *et al.*, "Survival stratification for colorectal cancer via multi-omics integration using an autoencoder-based model," *Experimental Biology and Medicine*, vol. 247, no.11, pp. 898–909, 2022, doi: 10.1177/15353702211065010.
- [10] G. Rajput, S. Agrawal, K. Biyani, and S. Vishvakarma, "Early breast cancer diagnosis using coget activation function-based deep learning implementation on screened mammograms," *Internal Journal Imaging System Technology*, vol. 32, no. 4, 2022, doi: 10.1002/ima.22701.
- [11] A. Abougarair, A. Oun, S. Sawan, and A. Ma'arif, "Deep Learning-Based Automated Approach for Classifying Bacterial Images," *International Journal of Robotics and Control Systems*, vol. 4, no. 2, pp. 849–876, 2024, doi: 10.31763/ijrcs.v4i2.1423.
- [12] L. Huang, W. Yang, Z. Huang, C. Tang, and H. Li, "Artificial intelligence technique in detection of early esophageal cancer," *World Journal Gastroenterol*, vol.26, no. 39, pp. 5959–5969, 2020, doi: 10.3748/wjg.v26.i39.5959.
- [13] D. Herremans, "aiSTROM—A Roadmap for Developing a Successful AI Strategy," *IEEE Access*, vol. 9, pp. 155826–155838, 2021.
- [14] E. Malone *et al.*, "Molecular profiling for precision cancer therapies," *Genome Med.*, vol. 12, no. 8, 2020, doi: 10.1186/s13073-019-0703.
- [15] J. Shing *et al.*, "HPV-Associated Cancer Incidence by Disaggregated Asian American, Native Hawaiian, and Other Pacific Islander Ethnicity," *JNCI Cancer Spectrum*, 2023, doi: 10.1093/jncics/pkad012.
- [16] E. Ar-Reyouchi *et al.*, "A Perfect Challenge to Select the Ideal Virus Vaccine Using a New Potent Hierarchical Algorithm in a Smart Laboratory," *IEEE access*, vol. 11, pp. 79890–79910, 2023, doi: 10.1109/access.2023.3295329.
- [17] E. Ar-Reyouchi *et al.*, "Protocol Wireless Medical Sensor Networks in IoT for the Efficiency of Healthcare," *IEEE Internet of Things Journal*, pp. 10693–10704, 2021, doi: 10.1109/jiot.2021.3125886.
- [18] M. Cascella, M. Rajnik, A. Cuomo, S. C. Dulebohn, and R. Napoli. *Features, Evaluation and Treatment Coronavirus (COVID-19)*. National Center for Biomedical Information, 2020.
- [19] S. E. Elwefati, A. J. Abougarair, and M. M. Bakush, "Control of Epidemic Disease Based Optimization Technique," *2021 IEEE 1st International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering MI-STA*, pp. 52–57, 2021, doi: 10.1109/MI-STA52233.2021.9464453.
- [20] M. Nazar, M. Alam, E. Yafi, and M. Mazliham, "A Systematic Review of Human-Computer Interaction and Explainable Artificial Intelligence in Healthcare with Artificial Intelligence Techniques," *IEEE Access*, vol. 9, pp. 153316–153348, 2021, doi: 10.1109/access.2021.3127881.
- [21] B. Farhood, M. Najafi, and K. Mortezaee, "CD8 + cytotoxic T lymphocytes in cancer immunotherapy: A review," *Journal of Cellular Physiology*, vol. 234, no. 6, pp. 8509–8521, 2018, doi: 10.1002/jcp.27782.
- [22] J. Zhou *et al.*, "Optimal modeling of anti-breast cancer candidate drugs screening based on multi-model ensemble learning with imbalanced data," *Mathematical biosciences and engineering*, vol. 20, no. 3, pp. 5117–5134, 2023.
- [23] A. Bukkuri, "Optimal control analysis of combined chemotherapy-immunotherapy treatment regimens in a PKPD cancer evolution model," *Biomath.*, vol. 9, no. 1, 2020.
- [24] P. Das *et al.*, "Optimal control strategy for cancer remission using combinatorial therapy: a mathematical model-based approach," *Chaos Soliton Fract.*, vol. 145, 2021, doi: 10.1016/j.chaos.2021.110789.
- [25] H. Wei, "Mathematical modeling of tumor growth: the MCF-7 breast cancer cell line," *Mathematical Biosciences and Engineering*, vol. 16, no. 6, pp. 6512–6535, 2019, doi: 10.3934/mbe.2019325.
- [26] J. Shao, J. Ma, Q. Zhang, W. Li, and C. Wang, "Predicting gene mutation status via artificial intelligence technologies based on multimodal integration (MMI) to advance precision oncology," *Seminars in Cancer Biology*, vol. 91, pp. 1–15, 2023, doi: 10.1016/j.semcancer.2023.02.006.
- [27] F. Cui, Q. Cui, and Y. Song, "A Survey on Learning-Based Approaches for Modeling and Classification of Human–Machine Dialog Systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 4, pp. 1418–1432, 2021.
- [28] Y. Wang, L. Zhang, Y. Li, F. Wu, S. Cao, and F. Ye, "Predicting the prognosis of HER2-positive breast cancer patients by fusing pathological whole slide images and clinical features using multiple instance learning," *Mathematical biosciences and engineering*, vol. 20, no. 6, pp. 11196–11211, 2023, doi: 10.3934/mbe.2023496.
- [29] B. Kann, A. Hosny, and H. Aerts, "Artificial Intelligence for Clinical Oncology," *Cancer Cell*, vol. 39, no. 7, 2021.
- [30] A. Abougarair, "Neural Networks Identification and Control of Mobile Robot Using Adaptive Neuro Fuzzy Inference System," *ICEMIS'20: Proceedings of the 6th International Conference on Engineering & MIS*, 2020.

- [31] A. Abougarair, "Adaptive Neural Networks Based Optimal Control for Stabilizing Nonlinear System," *IEEE (MI-STA2023)*, 2023, doi: 10.1109/MI-STA57575.2023.10169340
- [32] K. Bao, "An elementary mathematical modeling of drug resistance in cancer," *Mathematical Biosciences and Engineering*, vol. 18, no. 1, pp. 339–353, 2021.
- [33] H. Zhang and J. Lei, "Optimal treatment strategy of cancers with intratumor heterogeneity," *Mathematical biosciences and engineering*, vol. 19, no. 12, pp. 13337–13373, 2022.
- [34] M. Najafi and H. Basirzadeh, "Optimal control homotopy perturbation method for cancer model," *International Journal of Biomathematics*, vol. 12, no. 3, 2019.
- [35] A. Abougarair and S. Elwefati, "Identification and Control of Epidemic Disease Based Neural Networks and Optimization Technique," *International Journal of Robotics and Control Systems*, vol. 3, no. 4 pp. 780-803, 2023.
- [36] T. Li and Y. Xiao, "Optimal strategies for coordinating infection control and socio-economic activities," *Mathematics and Computers in Simulation*, vol. 207, pp. 533–555, 2023.
- [37] K. Exarchos *et al.*, "Review of Artificial Intelligence Techniques in Chronic Obstructive Lung Disease," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 5, pp. 2331–2338, 2022.
- [38] O. Tutsoy, "Pharmacological, Non-Pharmacological Policies and Mutation: An Artificial Intelligence Based Multi-Dimensional Policy Making Algorithm for Controlling the Casualties of the Pandemic Diseases," *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 12, pp. 9477–9488, 2022.
- [39] H. Zhou, Y. Liu, Z. Wang, and H. Song, "Linear stability for a free boundary problem modeling the growth of tumor cord with time delay," *Mathematical biosciences and engineering*, vol. 21, no. 2, pp. 2344–2365, 2024, doi: 10.3934/mbe.2024103.
- [40] C. Jiang, H. Xu, C. Huang, Y. Chen, R. Zou, and Y. Wang, "Research on knowledge dissemination in smart cities environment based on intelligent analysis algorithms: a case study on online platform," *Mathematical Biosciences and Engineering*, vol. 18, no. 3, pp. 2632–2653, 2021, doi: 10.3934/mbe.2021134.
- [41] S. Muniyappan, A. X. A. Rayan, and G. T. Varieth, "DTiGNN: Learning drug-target embedding from a heterogeneous biological network based on a two-level attention-based graph neural network," *Mathematical Biosciences and Engineering*, vol. 20, no. 5, pp. 9530–9571, 2023, doi: 10.3934/mbe.2023419.
- [42] Y. Zhou, Y. Niu, Q. Luo, and M. Jiang, "Teaching learning-based whale optimization algorithm for multi-layer perceptron neural network training," *Mathematical Biosciences and Engineering*, vol. 17, no. 5, pp. 5987–6025, 2020, doi: 10.3934/mbe.2020319.
- [43] S. Huang, S. Zheng, and R. Chen, "Multi-source transfer learning with Graph Neural Network for excellent modelling the bioactivities of ligands targeting orphan G protein-coupled receptors," *Mathematical biosciences and engineering*, vol. 20, no. 2, pp. 2588–2608, Jan. 2022, doi: 10.3934/mbe.2023121.
- [44] J. Guadalupe, J. Ruiz, P. Rodriguez, J. Ramirez, and J. Gabriel, "Multi-Stroke handwriting character recognition based on sEMG using convolutional-recurrent neural networks," *Mathematical Biosciences and Engineering*, vol. 17, no. 5, pp. 5432–5448, 2020, doi: 10.3934/mbe.2020293.
- [45] H. Mashayekhi and M. Nazari, "Reinforcement learning based feedback control of tumor growth by limiting maximum chemo-drug dose using fuzzy logic," *Journal of Control*, vol. 15, no. 4, pp. 13–23, Jan. 2022, doi: 10.52547/joc.15.4.13.
- [46] J. Qiu *et al.*, "Research on motion recognition based on multi-dimensional sensing data and deep learning algorithms," *Mathematical Biosciences and Engineering*, vol. 20, no. 8, pp. 14578–14595, Jan. 2023, doi: 10.3934/mbe.2023652.
- [47] F. Rihan and H. J. Alsakaji, "Analysis of a stochastic HBV infection model with delayed immune response," *Mathematical Biosciences and Engineering*, vol. 18, no. 5, pp. 5194–5220, 2021, doi: 10.3934/mbe.2021264.
- [48] Y. Xia, W. Zhou, and Z. Yang, "Global analysis and optimal harvesting for a hybrid stochastic phytoplankton-zooplankton-fish model with distributed delays," *Mathematical Biosciences and Engineering*, vol. 17, no. 5, pp. 6149–6180, 2020, doi: 10.3934/mbe.2020326.
- [49] Y. Fu, T. Lu, M. Zhou, D. Liu, Q. Gan, and G. Wang, "Effect of color cross-correlated noise on the growth characteristics of tumor cells under immune surveillance," *Mathematical Biosciences and Engineering*, vol. 20, no. 12, pp. 21626–21642, Jan. 2023, doi: 10.3934/mbe.2023957.
- [50] E. Agyingi, T. I. Wiandt, L. U. Buxbaum, and B. N. Thomas, "Modeling the immune system response: an application to leishmaniasis," *Mathematical Biosciences and Engineering*, vol. 17, no. 2, pp. 1253–1271, 2020, doi: 10.3934/mbe.2020064.
- [51] C. Treeratayapun, A. Jonathan, and N. Suyaroj, "Reinforcement learning optimal control with semi-continuous reward function and fuzzy-rules networks for drug administration of cancer treatment," *Soft computing*, vol. 27, no. 22, pp. 17347–17356, 2023, doi: 10.1007/s00500-023-08068-1.
- [52] L. Li and W. Zhao, "Deterministic and stochastic dynamics of a modified Leslie-Gower prey-predator system with simplified Holling-type IV scheme," *Mathematical Biosciences and Engineering*, vol. 18, no. 3, pp. 2813–2831, 2021, doi: 10.3934/mbe.2021143.
- [53] X. He, M. Liu, and X. Xu, "Analysis of stochastic disease including predator-prey model with fear factor and Lévy jump," *Mathematical Biosciences and Engineering*, vol. 20, no. 2, pp. 1750–1773, 2023, doi: 10.3934/mbe.2023080.
- [54] S. Sriram, H. Natiq, K. Rajagopal, O. Krejcar, and H. Namazi, "Dynamics of a two-layer neuronal network with asymmetry in coupling," *Mathematical biosciences and engineering*, vol. 20, no. 2, pp. 2908–2919, Jan. 2022, doi: 10.3934/mbe.2023137.
- [55] H. Zhu, "A graph neural network-enhanced knowledge graph framework for intelligent analysis of policing cases," *Mathematical Biosciences and Engineering*, vol. 20, no. 7, pp. 11585–11604, 2023, doi: 10.3934/mbe.2023514.
- [56] G. Song, T. Tian, and X. Zhang, "A mathematical model of cell-mediated immune response to tumor," *Mathematical Biosciences and Engineering*, vol. 18, no. 1, pp. 373–385, 2021, doi: 10.3934/mbe.2021020.
- [57] P. Das, S. Das, A. Rihan, M. Uzuntarla, and D. Ghosh, "Optimal control strategy for cancer remission using combinatorial therapy: A mathematical model-based approach," *Chaos, Solitons & Fractals*, vol. 145, p. 110789, Apr. 2021, doi: 10.1016/j.chaos.2021.110789.
- [58] W. Wang, Y. Gao, Z. Hong, and C. Ahn, "Reinforcement Learning-Based Optimal Tracking Control of an Unknown Unmanned Surface Vehicle," *IEEE transactions on neural networks and learning systems*, vol. 32, no. 7, pp. 3034–3045, Jul. 2021, doi: 10.1109/tnnls.2020.3009214.
- [59] M. Piccinini, S. Taddei, M. Larcher, M. Piazza, and F. Biral, "A Physics-Driven Artificial Agent for Online Time-Optimal Vehicle Motion Planning and Control," *IEEE Access*, vol. 11, pp. 46344–46372, Jan. 2023, doi: 10.1109/access.2023.3274836.
- [60] L. Zhang, Q. Zhao, L. Wang, and L. Zhang, "Urban Intersection Signal Control Based on Time-Space Resource Scheduling," *IEEE Access*, vol. 9, pp. 49281–49291, 2021, doi: 10.1109/access.2021.3059496.
- [61] D. Wang, N. Gao, D. Liu, J. Li, and F. L. Lewis, "Recent progress in reinforcement learning and adaptive dynamic programming for advanced control applications," *IEEE/CAA Journal of Automatica Sinica*, pp. 1–19, 2023, doi: 10.1109/jas.2023.123843.
- [62] N. Chen, S. Luo, J. Dai, B. Luo, and W. Gui, "Optimal Control of Iron-Removal Systems Based on Off-Policy Reinforcement Learning," *IEEE Access*, vol. 8, pp. 149730–149740, Jan. 2020, doi: 10.1109/access.2020.3015801.
- [63] U. Mohite and H. G. Patel, "Optimization assisted Kalman filter for cancer chemotherapy dosage estimation," *Artificial Intelligence in Medicine*, vol. 119, p. 102152, Sep. 2021, doi: 10.1016/j.artmed.2021.102152.
- [64] Y. Lin, J. Phee, and N. Azad, "Comparison of Deep Reinforcement Learning and Model Predictive Control for Adaptive Cruise Control," *IEEE Transactions on Intelligent Vehicles*, vol. 6, no. 2, pp. 221–231, 2021, doi: 10.1109/tiv.2020.3012947.
- [65] J. Pakdeeto, S. Wansungnoen, and K. Areerak, "Optimal Speed Controller Design of Commercial BLDC Motor by Adaptive Tabu Search Algorithm," *IEEE Access*, vol. 11, pp. 79710–79720, Jan. 2023, doi: 10.1109/access.2023.3300233.

- [66] X. Zhang and Z. Wang, "Simple Robust Model Predictive Current Control for PMSM Drives without Flux linkage Parameter," *IEEE Transactions on Industrial Electronics*, pp. 3515–3524, 2022, doi: 10.1109/tie.2022.3176288.
- [67] G. Hartmann, Z. Shiller, and A. Azaria, "Model-Based Reinforcement Learning for Time-Optimal Velocity Control," *IEEE Robotics & Automation Letters*, vol. 5, no. 4, pp. 6185–6192, Oct. 2020, doi: 10.1109/lra.2020.3012128.
- [68] E. Schiassi, A. Ambrosio, and R. Furfaro, "Bellman Neural Networks for the Class of Optimal Control Problems with Integral Quadratic Cost," *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 3, pp. 1016–1025, 2024, doi: 10.1109/tai.2022.3206735.
- [69] M. Sarhaddi and M. Yaghoobi, "A new approach in cancer treatment regimen using adaptive fuzzy back-stepping sliding mode control and tumor-immunity fractional order model," *Biocybernetics and Biomedical Engineering*, vol. 40, no. 4, pp. 1654–1665, 2020, doi: 10.1016/j.bbe.2020.09.003.
- [70] J. He, Y. Cen, S. Alelaumi, and D. Won, "An Artificial Intelligence-Based Pick-and-Place Process Control for Quality Enhancement in Surface Mount Technology," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 12, no. 10, pp. 1702–1711, 2022, doi: 10.1109/tcpmt.2022.3215109.
- [71] J. Boisvert, J. Lafontaine, A. Glory, S. Coulombe, and P. Wong, "Comparison of Three Radio-Frequency Discharge Modes on the Treatment of Breast Cancer Cells in Vitro," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 4, no. 5, pp. 644–654, Sep. 2020, doi: 10.1109/trpms.2020.2994870.
- [72] R. Lupat, R. Perera, S. Loi, and J. Li, "Moanna: Multi-Omics Autoencoder-Based Neural Network Algorithm for Predicting Breast Cancer Subtypes," *IEEE Access*, vol. 11, pp. 10912–10924, Jan. 2023, doi: 10.1109/access.2023.3240515.
- [73] Y. Tang, L. Chen, A. Zhang, C. Liao, M. E. Gross, and E. S. Kim, "In Vivo Non-Thermal, Selective Cancer Treatment with High-Frequency Medium-Intensity Focused Ultrasound," *IEEE Access*, vol. 9, pp. 122051–122066, 2021, doi: 10.1109/access.2021.3108548.
- [74] J. Sun, Y. Yang, Y. Wang, L. Wang, X. Song, and X. Zhao, "Survival Risk Prediction of Esophageal Cancer Based on Self-Organizing Maps Clustering and Support Vector Machine Ensembles," *IEEE Access*, vol. 8, pp. 131449–131460, 2020.
- [75] A. Mollahosseini, B. Hasani, and M. Mahoor, "AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild," *IEEE Transactions on Affective Computing*, vol. 10, no. 1, pp. 18–31, Jan. 2019, doi: 10.1109/TAFFC.2017.2740923.
- [76] C. Wu *et al.*, "Towards Patient-Specific Optimization of Neoadjuvant Treatment Protocols for Breast Cancer Based on Image-Guided Fluid Dynamics," *IEEE transactions on bio-medical engineering/IEEE transactions on biomedical engineering*, vol. 69, no. 11, pp. 3334–3344, Nov. 2022, doi: 10.1109/tbme.2022.3168402.
- [77] M. Semmler *et al.*, "Molecular Mechanisms of the Efficacy of Cold Atmospheric Pressure Plasma (CAP) in Cancer Treatment," *Cancers*, vol. 12, no. 2, p. 269, Jan. 2020, doi: 10.3390/cancers12020269.
- [78] D. Nie *et al.*, "Multi-Channel 3D Deep Feature Learning for Survival Time Prediction of Brain Tumor Patients Using Multi-Modal Neuroimages," *Scientific Reports*, vol. 9, no. 1, p. 1103, Jan. 2019, doi: 10.1038/s41598-018-37387-9.
- [79] Y. Oh, G. Bae, K. Kim, M. Yeo, and J. C. Ye, "Multi-Scale Hybrid Vision Transformer for Learning Gastric Histology: AI-Based Decision Support System for Gastric Cancer Treatment," *IEEE journal of biomedical and health informatics*, vol. 27, no. 8, pp. 4143–4153, Aug. 2023, doi: 10.1109/jbhi.2023.3276778.
- [80] Z. Hou, T. Lee, and M. Keidar, "Reinforcement Learning with Safe Exploration for Adaptive Plasma Cancer Treatment," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 6, no. 4, pp. 482–492, 2022, doi: 10.1109/trpms.2021.3094874.
- [81] F. Angaroni *et al.*, "An Optimal Control Framework for the Automated Design of Personalized Cancer Treatments," *Frontiers in Bioengineering and Biotechnology*, vol. 8, May 2020, doi: 10.3389/fbioe.2020.00523.
- [82] C. Qin, D. Yao, Y. Shi, and Z. Song, "Computer-aided detection in chest radiography based on artificial intelligence: a survey," *Bio Medical Engineering on Line*, vol. 17, no. 1, Aug. 2018, doi: 10.1186/s12938-018-0544-y.
- [83] Y. Tian, Y. Du, Q. Zhang, J. Cheng, and Z. Yang, "Depth estimation for advancing intelligent transport systems based on self-improving pyramid stereo network," *IET Intelligent Transport Systems*, vol. 14, no. 5, pp. 338–345, 2020, doi: 10.1049/iet-its.2019.0462.
- [84] S. Badawi, "Data Augmentation for Sorani Kurdish News Headline Classification Using Back-Translation And Deep Learning Model," *Kurdistan Journal of Applied Research*, pp. 27–34, Jun. 2023, doi: 10.24017/science/2023.1.4.
- [85] C. Shorten, T. Khoshgoftaar, and B. Furht, "Deep Learning applications for COVID-19," *Journal of Big Data*, vol. 8, no. 1, Jan. 2021, doi: 10.1186/s40537-020-00392-9.
- [86] C. Campbell and R. Kahwash, "Will Complement Inhibition Be the New Target in Treating COVID-19-Related Systemic Thrombosis," *Circulation*, vol. 141, no. 22, pp. 1739–1741, Jun. 2020, doi: 10.1161/circulationaha.120.047419.
- [87] M. Baptiste, S. Moinuddeen, Courtney Lace Soliz, H. Ehsan, and G. Kaneko, "Making Sense of Genetic Information: The Promising Evolution of Clinical Stratification and Precision Oncology Using Machine Learning," *Genes*, vol. 12, no. 5, pp. 722–722, May 2021, doi: 10.3390/genes12050722.
- [88] A. Lagree *et al.*, "Assessment of Digital Pathology Imaging Biomarkers Associated with Breast Cancer Histologic Grade," *Current oncology*, vol. 28, no. 6, pp. 4298–4316, Oct. 2021, doi: 10.3390/curroncol28060366.
- [89] M. Almahfuz *et al.*, "Clinically Applicable Machine Learning Approaches to Identify Attributes of Chronic Kidney Disease (CKD) for Use in Low-Cost Diagnostic Screening," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 9, pp. 1–11, 2021, doi: 10.1109/jtehm.2021.3073629.
- [90] E. Kim, H. Cho, J. Kwon, Y. Oh, Eun Sook Ko, and H. Park, "Tumor-Attentive Segmentation-Guided GAN for Synthesizing Breast Contrast-Enhanced MRI Without Contrast Agents," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 11, pp. 32–43, 2023, doi: 10.1109/jtehm.2022.3221918.
- [91] Y. Dwivedi *et al.*, "So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy," *International Journal of Information Management*, vol. 71, no. 0268-4012, p. 102642, 2023, doi: 10.1016/j.ijinfomgt.2023.102642.
- [92] C. Shorten, T. M. Khoshgoftaar, and B. Furht, "Text Data Augmentation for Deep Learning," *Journal of Big Data*, vol. 8, no. 1, Jul. 2021, doi: 10.1186/s40537-021-00492-0.
- [93] J. Liao *et al.*, "Artificial intelligence assists precision medicine in cancer treatment," *Frontiers in Oncology*, vol. 12, 2023, doi: 10.3389/fonc.2022.998222.
- [94] C. Secasani *et al.*, "Artificial Intelligence System for Predicting Prostate Cancer Lesions from Shear Wave Elastography Measurements," *Current Oncology*, vol. 29, no. 6, pp. 4212–4223, 2022, doi: 10.3390/curroncol29060336.
- [95] K. Jee, T. Bortfeld, I. Elnaqa, and L. Dong, "Advanced Topics in Particle Radiotherapy," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 6, no. 3, pp. 247–251, Mar. 2022, doi: 10.1109/trpms.2022.3150218.
- [96] P. Vatiwutipong, S. Vachmanus, T. Noraset, and S. Tuarob, "Artificial Intelligence in Cosmetic Dermatology: A Systematic Literature Review," *IEEE Access*, vol. 11, pp. 71407–71425, 2023, doi: 10.1109/access.2023.3295001.
- [97] M. Finocchiaro *et al.*, "Physical Simulator for Colonoscopy: A Modular Design Approach and Validation," *IEEE Access*, vol. 11, pp. 36945–36960, 2023, doi: 10.1109/access.2023.3266087.
- [98] A. Borkowski, "Using Artificial Intelligence for COVID-19 Chest X-ray Diagnosis," *Federal Practitioner*, vol. 37, no. 9, 2020, doi: 10.12788/fp.0045.
- [99] E. Prezioso *et al.*, "Predictive Medicine for Salivary Gland Tumours Identification Through Deep Learning," in *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 10, pp. 4869–4879, Oct. 2022, doi: 10.1109/JBHI.2021.3120178.
- [100] H. Qiu, S. Ding, J. Liu, L. Wang, and X. Wang, "Applications of Artificial Intelligence in Screening, Diagnosis, Treatment, and Prognosis of Colorectal Cancer," *Current Oncology*, vol. 29, no. 3, pp. 1773–1795, 2022, doi: 10.3390/curroncol29030146.

- [101] R. Singh *et al.*, "Precision Oncology: Artificial Intelligence and DNA Methylation Analysis of Circulating Cell-Free DNA for Lung Cancer Detection," *Frontiers in Oncology*, vol. 12, 2022.
- [102] L. Wei *et al.*, "Artificial intelligence (AI) and machine learning (ML) in precision oncology: a review on enhancing discoverability through multiomics integration," *British Journal of Radiology*, vol. 96, no. 1150, Oct. 2023.
- [103] L. Chang, J. Wu, N. Moustafa, and K. Yu, "AI-Driven Synthetic Biology for Non-Small Cell Lung Cancer Drug Effectiveness-Cost Analysis in Intelligent Assisted Medical Systems," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 10, pp. 5055–5066, Dec. 2021.
- [104] I. Haq, M. Ahmed, M. Assam, Y. Ghadi, and A. Algarni, "Unveiling the Future of Oral Squamous Cell Carcinoma Diagnosis: An Innovative Hybrid AI Approach for Accurate Histopathological Image Analysis," *IEEE Access*, vol. 11, pp. 118281–118290, Jan. 2023.
- [105] T. Thakur, I. Batra, A. Malik, D. Ghimire, S. Kim, and A. S. M. Sanwar Hosen, "RNN-CNN Based Cancer Prediction Model for Gene Expression," *IEEE Access*, vol. 11, pp. 131024–131044, Jan. 2023.
- [106] M. Shamim, "Hardware Deployable Edge-AI Solution for Prescreening of Oral Tongue Lesions Using Tiny ML on Embedded Devices," *IEEE Embedded Systems Letters*, vol. 14, no. 4, pp. 183–186, Dec. 2022.
- [107] T. Pham, V. Ravi, C. Fan, B. Luo, and X.-F. Sun, "Classification of IHC Images of NATs With ResNet-FRP-LSTM for Predicting Survival Rates of Rectal Cancer Patients," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 11, pp. 87–95, Jan. 2023.
- [108] T. Kausar, Y. Lu, and A. Kausar, "Breast Cancer Diagnosis Using Lightweight Deep Convolution Neural Network Model," *IEEE Access*, vol. 11, pp. 124869–124886, 2023.
- [109] J. Cunningham, J. Brown, R. Gatenby, and K. Staňková, "Optimal control to develop therapeutic strategies for metastatic castrate resistant prostate cancer," *Journal of Theoretical Biology*, vol. 459, pp. 67–78, Dec. 2018.
- [110] P. Wang, R. Liu, Z. Jiang, Y. Yao, and Z. Shen, "The Optimization of Combination Chemotherapy Schedules in the Presence of Drug Resistance," *IEEE transactions on automation science and engineering*, vol. 16, no. 1, pp. 165–179, 2019.
- [111] P. Samadi *et al.*, "Berberine: A novel therapeutic strategy for cancer," *IUBMB Life*, vol. 72, no. 10, pp. 2065–2079, Jul. 2020.
- [112] X. Wang *et al.*, "Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis," *IEEE Transactions on Cybernetics*, vol. 50, no. 9, pp. 3950–3962, Sep. 2020.
- [113] Z. Hasan *et al.*, "Nanomedicine: Treatment of Chronic Disease Using Gold Nano Thermo Robot (GNTR) Empowered with Nanotechnology Approaches," *IEEE access*, vol. 12, pp. 8552–8584, Jan. 2024.
- [114] K. Mridha, M. Uddin, J. Shin, S. Khadka, and M. Mridha, "An Interpretable Skin Cancer Classification Using Optimized Convolutional Neural Network for a Smart Healthcare System," *IEEE Access*, vol. 11, pp. 41003–41018, 2023.
- [115] P. Manganelli, G. Lazzini, P. Russo, and M. Acunto, "Raman spectroscopy and AI applications in cancer grading. An overview," *IEEE Access*, pp. 54816–54852, 2024.
- [116] B. Huang *et al.*, "3D Lightweight Network for Simultaneous Registration and Segmentation of Organs-at-Risk in CT Images of Head and Neck Cancer," *IEEE Transactions on Medical Imaging*, vol. 41, no. 4, pp. 951–964, Apr. 2022.
- [117] A. Jarrett *et al.*, "Optimal Control Theory for Personalized Therapeutic Regimens in Oncology: Background, History, Challenges, and Opportunities," *Journal of Clinical Medicine*, vol. 9, no. 5, 2020.
- [118] J. Cunningham, J. Brown, R. Gatenby, and K. Staňková, "Optimal control to develop therapeutic strategies for metastatic castrate resistant prostate cancer," *Journal of Theoretical Biology*, vol. 459, pp. 67–78, Dec. 2018.
- [119] S. Oke, M. Matadi, and S. Xulu, "Optimal Control Analysis of a Mathematical Model for Breast Cancer," *Mathematical and Computational Applications*, vol. 23, no. 2, 2018.
- [120] F. Subhan *et al.*, "Cancerous Tumor Controlled Treatment Using Search Heuristic (GA)-Based Sliding Mode and Synergetic Controller," *Cancers*, vol. 14, no. 17, 2022.
- [121] N. Singha, "Implementation of fractional optimal control problems in real-world applications," *Fractional Calculus and Applied Analysis*, vol. 23, no. 6, pp. 1783–1796, 2020.
- [122] M. Gluzman, J. Scott, and A. Vladimirovsky, "Optimizing adaptive cancer therapy: dynamic programming and evolutionary game theory," *Proceedings of the Royal Society B: Biological Sciences*, vol. 287, no. 1925, 2020.
- [123] M. Elkaf, A. Meskaf, and K. Allali, "Dynamics of cancer cells with immunotherapy and virotherapy," *Communications in Mathematical Biology and Neuroscience*, vol. 76, 2021.
- [124] R. Pal, N. Afroz, A. Khan, and M. Padder, "Stability and Dynamical Analysis of Generalized Tumor-Immune Interaction Model with Conformable Fractional Derivative," *Research Square (Research Square)*, 2021.
- [125] M. Wu *et al.*, "Improvement of the anticancer efficacy of PD-1/PD-L1 blockade via combination therapy and PD-L1 regulation," *Journal of Hematology & Oncology*, vol. 15, no. 1, 2022.
- [126] F. Rihan and G. Velmurugan, "Dynamics of fractional-order delay differential model for tumor-immune system," *Chaos, solitons & fractals*, vol. 132, pp. 109592–109592, 2020.
- [127] C. Nanditha and M. Rajan, "an Adaptive Pharmacokinetic Optimal Control Approach in Chemotherapy for Heterogeneous Tumor," *Journal of Biological Systems*, vol. 30, no. 3, pp. 529–551, 2022.
- [128] D. Amilo, B. Kaymakzade, and E. Hincal, "A fractional-order mathematical model for lung cancer incorporating integrated therapeutic approaches," *Scientific Reports*, vol. 13, no. 1, 2023.
- [129] A. Ibrahim, N. Maan, K. Jemon, and A. Abidemi, "Global Stability and Thermal Optimal Control Strategies for Hyperthermia Treatment of Malignant Tumors," *Mathematics*, vol. 10, no. 13, pp. 2188–2188, 2022.
- [130] S. Jawad, M. Winter, Z. Rahman, Y. Al-Yasir, and A. Zeb, "Dynamical Behavior of a Cancer Growth Model with Chemotherapy and Boosting of the Immune System," *Mathematics*, vol. 11, no. 2, 2023.
- [131] K. Dehingia, H. Sarmah, K. Hosseini, K. Sadri, S. Salahshour, and C. Park, "An Optimal Control Problem of Immuno-Chemotherapy in Presence of Gene Therapy," *AIMS Mathematics*, vol. 6, no. 10, pp. 11530–11549, 2021.
- [132] A. J. Abougarair, H. M. Gnan, A. Oun, and S. O. Elwarshfani, "Implementation of a Brain-Computer Interface for Robotic Arm Control," *2021 IEEE 1st International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering MI-STA*, pp. 58–63, 2021, doi: 10.1109/MI-STA52233.2021.9464359.
- [133] A. Abougarair and A. Elmolhi, "Robust Control and Optimized Parallel Control Double Loop Design for Mobile Robot," *IAES International Journal of Robotics and Automation (IJRA)*, vol. 9, no. 3, 2020.
- [134] A. Abougarair, "Optimal Control Synthesis of Epidemic Model," *IJEIT International Journal on Engineering and Information Technology*, vol. 10, no. 2, 2022.
- [135] M. Edardar and A. Abougarair, "Tracking Control with Hysteresis Compensation Using Neural Networks," *IEEE (MI-STA2021)*, 2021, doi:10.1109/MI-STA52233.2021.9464365.
- [136] M. Aburakhis *et al.*, "Adaptive Neural Networks Based Robust Output Feedback Controllers for Nonlinear Systems," *International Journal of Robotics and Control Systems*, vol. 2, no. 1, pp. 37–56, 2022.
- [137] F. Rihan and S. Almekhlafi, "A fractional-order delay differential model with optimal control for cancer treatment based on synergy between anti-angiogenic and immune cell therapies," *Discrete & Continuous Dynamical Systems*, vol. 13, no. 9, pp. 2403–2424, 2020.
- [138] H. Alsakaji, F. Rihan, K. Udhayakumar, and F. Elktaibi, "Stochastic tumor-immune interaction model with external treatments and time delays: An optimal control problem," *Mathematical Biosciences and Engineering*, vol. 20, no. 11, pp. 19270–19299, 2023.
- [139] A. Kumar, U. Dubey, and B. Dubey, "The impact of radio-chemotherapy on tumour cells interaction with optimal control and

- sensitivity analysis,” *Mathematical Biosciences*, vol. 369, pp. 1091466–109146, 2024.
- [140] A. Zrigan *et al.*, “ADRC Robustness for Boiler Turbine Unit at Equilibrium Operation Point,” *IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA)*, pp. 166–170, 2023, doi: 10.1109/MI-STA57575.2023.10169524.
- [141] A. Zrigan *et al.*, “Optimized PID Controller and Generalized Inverted Decoupling Design for MIMO System,” *2023 IEEE International Conference on Advanced Systems and Emergent Technologies (IC_ASET’2023)*, 2023.
- [142] R. Baxter, J. Fann, J. DiMaio, and K. Lobdell, “Digital Health Primer for Cardiothoracic Surgeons,” *The Annals of Thoracic Surgery*, vol. 110, no. 2, pp. 364–372, 2020.
- [143] G. Hull *et al.*, “Cancer control with radical prostatectomy alone in 1,000 consecutive patients,” *The Journal of Urology*, vol. 167, no. 2, pp. 528–534, Feb. 2002.
- [144] A. Jarrett *et al.*, “Optimal Control Theory for Personalized Therapeutic Regimens in Oncology: Background, History, Challenges, and Opportunities,” *Journal of Clinical Medicine*, vol. 9, no. 5, p. 1314, May 2020.
- [145] F. Heydarpour, E. Abbasi, M. Ebadi, and S. Karbassi, “Solving an Optimal Control Problem of Cancer Treatment by Artificial Neural Networks,” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 4, 2020.
- [146] R. Coletti, L. Leonardelli, S. Parolo, and L. Marchetti, “A QSP model of prostate cancer immunotherapy to identify effective combination therapies,” *Scientific Reports*, vol. 10, no. 1, 2020.
- [147] R. Coletti, A. Pugliese, and L. Marchetti, “Modeling the effect of immunotherapies on human castration-resistant prostate cancer,” *Journal of Theoretical Biology*, vol. 509, p. 110500, 2021.
- [148] H. Ma, H. Jian, and Y. Shi, “A sufficient maximum principle for backward stochastic systems with mixed delays,” *Mathematical Biosciences and Engineering*, vol. 20, no. 12, pp. 21211–21228, 2023.
- [149] C. Li, Y. Chen, and Z. Zhao, “Frequency hopping signal detection based on optimized generalized S transform and ResNet,” *Mathematical Biosciences and Engineering*, vol. 20, no. 7, pp. 12843–12863, 2023.
- [150] Y. Ding, G. Liu, and Y. An, “Stability and bifurcation analysis of a tumor-immune system with two delays and diffusion,” *Mathematical Biosciences and Engineering*, vol. 19, no. 2, pp. 1154–1173, 2021, doi: 10.3934/mbe.2022053.
- [151] P. Khalili, S. Zolatash, R. Vatankhah, and S. Taghvaei, “Optimal control methods for drug delivery in cancerous tumor by anti-angiogenic therapy and chemotherapy,” *IET Systems Biology*, vol. 15, no. 1, pp. 14–25, 2021.
- [152] W. Arebi and A. Abougarair, “Smart Glove for Sign Language Translation,” *International Robotics & Automation Journal*, vol. 8, no. 3, pp. 109–117, 2022, doi: 10.15406/iratj.2022.08.00253.
- [153] A. Emhemmed *et al.*, “Simulation Analysis and Control of Wireless Power Transfer for Implantable Medical Devices,” *IEEE 21st international Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, 2022, doi: 10.1109/STA56120.2022.10019046.
- [154] E. Dee, D. Dao, R. Patel, P. Santos, and F. Chino, “Disaggregation of Asian American and Pacific Islander Women With Stage 0-II Breast Cancer Unmasks Disparities in Survival and Surgery-to-Radiation Intervals: A National Cancer Database Analysis From 2004 to 2017,” *JCO Oncology Practice*, vol. 18, no. 8, pp. e1255–e1264, 2022.
- [155] J. Choi *et al.*, “Disparities in Colorectal Cancer Incidence among Asian and Pacific Islander Populations in Guam, Hawai’i, and the United States,” *International Journal of Environmental Research and Public Health*, vol. 21, no. 2, pp. 170–170, 2024.