



Politecnico di Bari

Repository Istituzionale dei Prodotti della Ricerca del Politecnico di Bari

Advanced modeling and control of biomedical systems with application to diabetes progression

This is a PhD Thesis

Original Citation:

Advanced modeling and control of biomedical systems with application to diabetes progression / De Paola, Pierluigi Francesco. - ELETTRONICO. - (2026).

Availability:

This version is available at <http://hdl.handle.net/11589/295800> since: 2026-01-17

Published version

DOI:

Publisher: Politecnico di Bari

Terms of use:

(Article begins on next page)



Italian National Ph.D. Program in Autonomous Systems

ACADEMIC DISCIPLINE: SYSTEMS AND CONTROL ENGINEERING (IINF-04/A)

Final Dissertation

Advanced Modeling and Control of Biomedical Systems with Application to Diabetes Progression

by

Pierluigi Francesco De Paola

Administrative Headquarters:

Politecnico di Bari – Department of Electrical and Information Engineering

Hosting University:

Consiglio Nazionale delle Ricerche – Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni (IEIT), Istituto di Analisi dei Sistemi e Informatica "A. Ruberti" (IASI)

Referees:

Prof. Chiara Toffanin

Prof. Constantino Lagoa

Supervisors:

Prof. Fabrizio Dabbene_

Prof. Alessandro Borri

Alessandro Borri

Coordinator of Ph.D Program

Prof. Mariagrazia Dotoli

LIBERATORIA PER L'ARCHIVIAZIONE DELLA TESI DI DOTTORATO

Al Magnifico Rettore
del Politecnico di Bari

Il sottoscritto DE PAOLA PIERLUIGI FRANCESCO nato a TRICASE (LE) il 26/07/1995
residente a TAURISANO (LE) in via_Dispersi in Russia, 48 e-mail pierluigifrancescodepaola@gmail.com
iscritto al 3° anno di Corso di Dottorato di Ricerca in AUTONOMOUS SYSTEMS ciclo 38°
ed essendo stato ammesso a sostenere l'esame finale con la prevista discussione della tesi dal titolo:
Advanced Modeling and Control of Biomedical Systems with Application to Diabetes Progression

DICHARA

- 1) di essere consapevole che, ai sensi del D.P.R. n. 445 del 28.12.2000, le dichiarazioni mendaci, la falsità negli atti e l'uso di atti falsi sono puniti ai sensi del codice penale e delle Leggi speciali in materia, e che nel caso ricorressero dette ipotesi, decade fin dall'inizio e senza necessità di nessuna formalità dai benefici consequenziali provvedimento emanato sulla base di tali dichiarazioni;
- 2) di essere iscritto al Corso di Dottorato di ricerca AUTONOMOUS SYSTEMS ciclo 38°, corso attivato ai sensi del "Regolamento dei Corsi di Dottorato di ricerca del Politecnico di Bari", emanato con D.R. n.286 del 01.07.2013;
- 3) di essere pienamente a conoscenza delle disposizioni contenute nel predetto Regolamento in merito alla procedura di deposito, pubblicazione e autoarchiviazione della tesi di dottorato nell'Archivio Istituzionale ad accesso aperto alla letteratura scientifica;
- 4) di essere consapevole che attraverso l'autoarchiviazione delle tesi nell'Archivio Istituzionale ad accesso aperto alla letteratura scientifica del Politecnico di Bari (IRIS-POLIBA), l'Ateneo archiverà e renderà consultabile in rete (nel rispetto della Policy di Ateneo di cui al D.R. 642 del 13.11.2015) il testo completo della tesi di dottorato, fatta salva la possibilità di sottoscrizione di apposite licenze per le relative condizioni di utilizzo (di cui al sito <http://www.creativecommons.it/Licenze>), e fatte salve, altresì, le eventuali esigenze di "embargo", legate a strette considerazioni sulla tutelabilità e sfruttamento industriale/commerciale dei contenuti della tesi, da rappresentarsi mediante compilazione e sottoscrizione del modulo in calce (Richiesta di embargo);
- 5) che la tesi da depositare in IRIS-POLIBA, in formato digitale (PDF/A) sarà del tutto identica a quelle **consegnate**/inviata/da inviarsi ai componenti della commissione per l'esame finale e a qualsiasi altra copia depositata presso gli Uffici del Politecnico di Bari in forma cartacea o digitale, ovvero a quella da discutere in sede di esame finale, a quella da depositare, a cura dell'Ateneo, presso le Biblioteche Nazionali Centrali di Roma e Firenze e presso tutti gli Uffici competenti per legge al momento del deposito stesso, e che di conseguenza va esclusa qualsiasi responsabilità del Politecnico di Bari per quanto riguarda eventuali errori, imprecisioni o omissioni nei contenuti della tesi;
- 6) che il contenuto e l'organizzazione della tesi è opera originale realizzata dal sottoscritto e non compromette in alcun modo i diritti di terzi, ivi compresi quelli relativi alla sicurezza dei dati personali; che pertanto il Politecnico di Bari ed i suoi funzionari sono in ogni caso esenti da responsabilità di qualsivoglia natura: civile, amministrativa e penale e saranno dal sottoscritto tenuti indenni da qualsiasi richiesta o rivendicazione da parte di terzi;
- 7) che il contenuto della tesi non infrange in alcun modo il diritto d'Autore né gli obblighi connessi alla salvaguardia di diritti morali ed economici di altri autori o di altri aventi diritto, sia per testi, immagini, foto, tabelle, o altre parti di cui la tesi è composta.

Bari, 15/01/2026

Firma



Il sottoscritto, con l'autoarchiviazione della propria tesi di dottorato nell'Archivio Istituzionale ad accesso aperto del Politecnico di Bari (POLIBA-IRIS), pur mantenendo su di essa tutti i diritti d'autore, morali ed economici, ai sensi della normativa vigente (Legge 633/1941 e ss.mm.ii.),

CONCEDE

- al Politecnico di Bari il permesso di trasferire l'opera su qualsiasi supporto e di convertirla in qualsiasi formato al fine di una corretta conservazione nel tempo. Il Politecnico di Bari garantisce che non verrà effettuata alcuna modifica al contenuto e alla struttura dell'opera.
- al Politecnico di Bari la possibilità di riprodurre l'opera in più di una copia per fini di sicurezza, back-up e conservazione.

Bari, 15/01/2026

Firma





Dr. Pierluigi Francesco De Paola

Advanced Modeling and Control of Biomedical Systems with Application to Diabetes Progression

Thesis submitted for the degree of Philosophiae Doctor

Italian National Ph.D. Program in Autonomous Systems
Politecnico di Bari
Consiglio Nazionale delle Ricerche

Tutors

Prof. Engr. *Fabrizio Dabbene*

Prof. Engr. *Alessandro Borri*



**Consiglio Nazionale
delle Ricerche**

2026



**Consiglio Nazionale
delle Ricerche**

Dissertation submitted for the degree of *Philosophiae Doctor*
Italian National Ph.D. Program in Autonomous Systems

Cycle:

38th

Administrative Headquarters:

Politecnico di Bari

Hosting University:

Consiglio Nazionale delle Ricerche

Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni (IEIIT)

Istituto di Analisi dei Sistemi e Informatica "A.Ruberti" (IASI)

Title:

Advanced Modeling and Control of Biomedical Systems with Application to Diabetes Progression

Ph.D Candidate:

Pierluigi Francesco De Paola, Politecnico di Bari (Bari, Italy) - Consiglio Nazionale Delle Ricerche (Turin/Rome, Italy)

Tutors:

Prof. Engr. Fabrizio Dabbene, Consiglio Nazionale delle Ricerche (IEIIT) (Turin, Italy)

Prof. Engr. Alessandro Borri, Consiglio Nazionale delle Ricerche (IASI) (Rome, Italy)

Coordinator:

Prof. Engr. Mariagrazia Dotoli, Politecnico di Bari (Bari, Italy)

External Reviewers:

Prof. Engr. Chiara Toffanin, Università degli Studi di Pavia (Pavia, Italy)

Prof. Engr. Costantino Lagoa, The Pennsylvania State University (State College, PA, USA)

Last version:

January 12, 2026

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission.

Abstract

The growing intersection between engineering and the life sciences is reshaping modern healthcare, enabling quantitative and predictive approaches to disease understanding and management. Within this context, modeling and control of biomedical systems have emerged as essential tools for interpreting physiological mechanisms, designing personalized interventions, and supporting decision-making in complex, dynamic environments.

Among chronic disease, Type 2 diabetes poses a major and urgent challenge, in light of its rising prevalence worldwide and in light of its aetiology, inherent dependence on lifestyle and long-term behavioral factors.

This Thesis contributes to the advancement of model-based and control-oriented methodologies for Type 2 diabetes management, addressing both physiological modeling and theoretical control aspects and develops along two main research directions.

The first research contribution of this Thesis concerns the mathematical modeling and control of Type 2 diabetes progression through physical activity. While existing glucose–insulin models and control strategies have primarily focused on insulin infusion for short-term regulation, this work targets the crucial long-term, preventive dimension of disease management. Specifically, the first research contribution proposes a mechanistic model of Type 2 diabetes progression suitable to provide an explicit representation of the long-term physiological processes modulated by regular physical activity. The model integrates exercise-induced adaptations in insulin sensitivity and beta-cell function, providing a quantitative framework to evaluate and predict the preventive role of exercise on diabetes course. Moreover, building upon this model, the Thesis introduces novel control strategies that leverage physical activity as control input, enhancing conventional insulin-based methods and promoting long-term, preventive regulation of glucose metabolism.

The second contribution of this Thesis is motivated by the aim of extending the methodological foundations of Type 2 diabetes and, more broadly, disease control through the introduction of a theoretical framework that embeds counterfactual reasoning within a systems-theoretic formulation. Counterfactual reasoning, widely employed in artificial intelligence for personalized decision support, offers a natural paradigm for evaluating alternative “*what-if*” scenarios. By translating this concept into the formal language of control and dynamical systems, the proposed framework provides a principled means to explore hypothetical scenarios that are dynamically feasible, physiologically consistent, and interpretable. Specifically, counterfactual reasoning is framed in control theory by means of optimal control and optimal transport problems, enabling the systematic assessment of intervention and possible control-driven strategies. The proposed framework is indeed general and applicable to a broad class of biomedical systems where reasoning about hypothetical interventions and long-term outcomes is essential.

By contributing to both applied and theoretical advancement in quantitative and control-oriented methods for healthcare in biomedical and healthcare systems, this Thesis set the basis for future development of decision support tools able to assist patients and clinicians in tailoring preventive lifestyle interventions.

*Coloro che vivono sono coloro che lottano,
sono quelli il cui progetto colma l'anima e la mente,
sono coloro che con un alto destino raggiungono l'aspra cima,
coloro che passeggiano pensosi,
innamorati di un gusto sublime,
avendo davanti agli occhi costantemente,
giorno e notte,
un qualche lavoro o un qualche grande amore.*
Victor Hugo

*To God, Who has never left my side.
May He always grant me the grace to seek Him,
and to perceive in my life the gentle echo of His Word.*

Contents

Preface	ix
List of Papers Written by the Author	x
Acronyms	xiii
Introduction	
1 Introduction	2
1.1 Mathematical Modeling in Biomedical Systems and Disease Prevention	2
1.2 Control Applications in Healthcare Systems and Disease Prevention	4
1.3 An Emerging Trend in Biomedical Systems Applications	5
1.4 Thesis Positioning and Motivation	5
1.5 Thesis Structure	11
Part 1	
2 Modeling the Long-Term Effect of Physical Activity on Blood Glucose Regulation	19
2.1 Introduction	19
2.2 Model Formulation	20
2.3 Macro Corpus of Simulations	26
2.4 Results	31
2.5 Discussion	34
2.6 Conclusion	36
3 On the Qualitative Analysis of the Model	45
3.1 Introduction	45
3.2 Formal Analysis of the Model	45
3.3 Properties of the Solutions	55
3.4 Conclusion	58
4 Model Predictive Control Approaches for Long-term Glucose Control via Physical Activity	60
4.1 Introduction	60
4.2 Preliminary Control Approach	61
4.3 Comprehensive Control Approach	65
4.4 Conclusion	71
Part 2	
5 A General Overview About Counterfactuals	78
5.1 General Context	78
5.2 Beyond the Philosophical Framework	78
5.3 A Control System Framework for Counterfactuals	80
6 An Optimal Control Based Formulation for Counterfactuals via Occupation Measures	84
6.1 Introduction	84
6.2 Fully Parametrized Systems	84

6.3	Uncertain Systems	94
6.4	Conclusion	99
7	An Optimal Transport Based Formulation for Counterfactuals via Occupation Measures	103
7.1	Introduction	103
7.2	Problem Statement	104
7.3	Counterfactual Extraction	111
7.4	Results	112
7.5	Conclusion	113
7.6	Background and Additional Contents	115
	Conclusions	
8	Conclusions	123
	Appendices	125
A	Preliminaries	126
A.1	Model description	126
B	Preliminaries on Occupation Measures and Moment-SOS Hierarchy	130
C	Additional Contribution	136
C.1	Stability Analysis of the model by Topp	136

List of Figures

2.1	Time trend of insulin sensitivity over a five-year horizon when exercise intervention is considered (green line) and without exercise (magenta line). Exercise program: 60 minutes/sessions, 3 sessions/week at moderate exercise intensity $u = 50\%$	24
2.2	Top panel: plasma glucose concentration in the fast-slow dynamics as a function of time observed in Case 1 (blue), in Case 2 (yellow) and Case 3 (red). Bottom panel: zoom showing the effect of single training sessions in the same cases.	27
2.3	Top panel: Plasma basal glucose concentration as a function of time observed in Case 1 (blue), Case 2 (yellow), Case 3 (red). Center panel: Serum basal insulin concentration in the same simulated cases. Bottom panel: beta-cell mass in the same simulated cases.	40
2.4	Basal glucose concentration (top panel) and insulin concentration (bottom panel) as a function of u over a five-year simulation horizon (simulations: 60 minutes/session, 3 sessions/week, $\tau_{SI} = 150$ days).	41
2.5	Basal glucose concentration with vigorous exercise ($u = 75\%$) for 75 minutes/week and moderate exercise ($u = 50\%$) as a function of exercise duration, from 150 to 500 minutes/week, using $\tau_{SI} = 180$ days.	42
2.6	Distribution of six out of 18 parameters related to genetic predisposition and environmental effects (i.e., τ_{SI} , $S_{I,target}$, τ_{β} , P_{max} , A_{max} , A_b) in Simulation Set GEP. Variation is expressed as percentage change from the nominal value.	42
2.7	Distribution of R_0 variation in <i>Type 2 diabetes</i> (T2D) subjects with high ($\tau_{SI} = 150$) and moderate ($\tau_{SI} = 180$) baseline risk in Simulation Set DIMB. Variation is expressed as percentage deviation from the nominal value of R_0	43
2.8	Distribution of the variations of k_s , $k_{n,SI}$, ζ_3 in simulated individuals with T2D at the end of the simulation window in Simulation Set IRAI-a. Variations are expressed as percentage deviation from the nominal values of the parameters.	43
2.9	Distribution of the <i>Cumulative Exercise Dose</i> (CED) variation for simulated individuals with T2D at the end of the simulation window in Simulation Set IRAI-b. Variation is expressed as percentage deviation from the dose of the nominal training program, for $\tau_{SI} = 150$, $Pr = 0.1$ (left panel) and $\tau_{SI} = 150$, $Pr = 0.15$ (center panel) and $\tau_{SI} = 180$, $Pr = 0.35$ (right panel).	44
3.1	Stationary plasma insulinemia as a function of stationary plasma glycemia, at different values of u_{eq}	49
3.2	Stationary γ as a function of stationary plasma glycemia.	50
3.3	Stationary insulin secretion capability versus stationary plasma glycemia.	52
4.1	Trajectory of basal glucose concentration over time: open-loop (dashed line) versus controlled case (solid line).	64
4.2	Evolution of the recommended single-session exercise duration over time, as computed through the inverse map in the <i>Model Predictive Control</i> (MPC)-controlled case ($\bar{u} = 50\%$, $T = 2$ days).	65

4.3	Top panel: Basal glucose concentration in the comparison between the true state (plant, blue line), the model predictions with wrong initial conditions (red line) and the observer estimation (yellow line). Bottom panel: Insulin concentration in the comparison between the true state (plant, blue line), the model predictions with wrong initial conditions (red line) and the observer estimation (yellow line).	69
4.4	Top panel: trend of the basal glucose concentration over the 20-year horizon. Bottom panel: trend of the control input in terms of recommended exercise per week.	72
4.5	Comparison between the basal glucose concentration in the undisturbed case (red line) and when a random disturbance is introduced.	73
4.6	Scatter plots of G and I for the 200 virtual subjects simulated by means of log-normal parameter perturbations at 5,10,15,20 years. The red dashed line represents the diabetic threshold (126 mg/dl).	74
5.1	Graphical representation of possible factual-counterfactual pairs in a classification problem. The picture is from [8].	81
6.1	Factual-counterfactual pairs as obtained by solving the optimization problem in the moment space with relaxation order $d = 4$	93
6.2	Factuals (red dots) and counterfactuals (blue dots) obtained by solving the optimization problem in the moment space with relaxation order $d = 4$ (top panel) and $d = 6$ (bottom panel).	94
6.3	Example of counterfactual trajectory as retrieved for the factual (250,0,0) by applying Algorithm 6.2.	95
6.4	Factual-counterfactual pairs as obtained by solving the optimization problem in the moment space with relaxation order $d = 4$ (top panel) and $d = 6$ (bottom panel). Robust counterfactuals lie at lower values of G with increased average distance from the boundary.	100
6.5	Factuals vs Counterfactuals obtained for $d = 4$ (top panel) and $d = 6$ (bottom panel). Both the plots highlight the same dense region of counterfactuals within the interval of values of I between 0 and 10. . . .	101
7.1	Graphical representation of the counterfactual problem in the measure framework.	107
7.2	Factuals vs Counterfactuals obtained for $d = 6$	114
7.3	Factuals vs Counterfactuals obtained for $d = 6$	115
7.4	Factuals vs Counterfactuals obtained for $d = 6$ and $d = 8$	116
B.1	Intuitive representation of the concept of occupation measure	132
C.1	The blue dots denote the set of initial conditions in the three-dimensional space $x_0 = [G_0 \ I_0 \ \beta_0]^T$ that converge to the equilibrium at $G = 100$ mg/dl (ROA of $G = 100$), whereas the red dots correspond to the set of initial conditions that converge to the equilibrium at $G = 600$ mg/dl (ROA of $G = 600$). The ROAs are box-shaped and separated by a surface nearby $\beta=300$ mg.	139
C.2	Outer approximation of the ROA for $G = 100$ mg/dl obtained via occupation measures for $d = 5$ and $d = 11$	140

List of Tables

2.1	Parameter ranges used to simulate inter-individual variability in genetic and environmental predisposition.	30
2.2	Parameter ranges used to simulate variability in individual response to exercise.	31
2.3	Randomization ranges for simulating variability in exercise program. . . .	31
2.4	Basal glucose and insulin concentration observed over a 20-year horizon for $u = 50\%$, 60% , and 70% and $\tau_{SI} = 150$ and 180 days.	32
2.5	Basal glucose levels observed in the simulated <i>Finnish Diabetes Prevention Study</i> (FDPS) and <i>Chinese Da Qing Diabetes Prevention Study</i> (CDQDPS) programs for $u = 30\%$, 40% , and 50% and $\tau_{SI} = 150$ and 180 days. . . .	33
2.6	Percentage of simulated individuals in which G remains below the diabetic threshold in the whole time horizon (no T2D) when parameters related to inter-individual variability are varied in the considered simulation sets. . .	34
4.1	Parameters of the control model with values and units for Preliminary Control Approach.	62
4.2	Validation results for the cohort of 400 virtual subjects.	71
6.1	Coefficients for $v_{d=2}^*(x)$ and $u_{d=2}^*(x)$	92
7.1	Parameter value	113
A.1	Description of the model state variables	127
A.2	Description of the state - depending functions	127
A.3	Parameter values used in the model.	128

Preface

This Thesis is submitted in partial fulfillment of the requirements for the degree of *Philosophiae Doctor* in Autonomous Systems at the *Politecnico di Bari*.

The research presented in this dissertation was carried out at the *Consiglio Nazionale delle Ricerche*, specifically at *Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni* (CNR-IEIIT) and at *Istituto di Analisi di Sistemi ed Informatica "A. Ruberti"* (CNR-IASI), under the supervision of Professors **Fabrizio Dabbene** and **Alessandro Borri**, between November 2022 and October 2025.

The content of this dissertation is, to the best of my knowledge, original, except when using excerpts from other sources that are appropriately acknowledged and referenced. The majority of the content presented in this Thesis consists of research works in which I feature as first author, that have been published in International Journal Articles [1], [2], [3], International Conference Proceedings [6], [7], [8] or currently available as preprints ready for submission to international journals [4]. In the case of unpublished contents (Chapter 3 and Chapter 7), these sections are derived from consolidated research outcomes that are in the final stages of preparation for submission to international peer-reviewed journals.

My supervisors, Professors **Fabrizio Dabbene** and **Alessandro Borri**, to whom I owe my deepest and most sincere gratitude, are co-authors of almost all the works cited herein, which form the core of the content presented in this Thesis. They have contributed in an invaluable manner, providing continuous and exceptionally fruitful mentorship, guidance and support throughout all stages of the research.

I wish to acknowledge with profound gratitude the invaluable contribution of Professors **Alessia Paglialonga**, from CNR-IEIIT, and **Pasquale Palumbo**, from *Università degli Studi di Milano-Bicocca*, whose collaboration has significantly enriched the conceptual and methodological rigor of my research, with particular reference to the content presented in Chapters 2, 3, 4 and related published works.

With respect to the contents in Chapter 6 and Chapter 7, they result from an international research collaboration and I acknowledge with sincere appreciation the support of Dr. **Jared Miller** (*University of Stuttgart, formerly at ETH in Zurich*), in the preliminary stages of the conceptualization, and the immeasurable contribution of Professor **Didier Henrion** (LAAS-CNRS, Toulouse), whose inspiring guidance has markedly advanced the theoretical side of the research related with the contents presented therein.

I proudly mention that my work in [1] was distinguished with an International Award, namely the **IEEE TC-HMS Outstanding Student Paper Award 2024**, that I had the honor of receiving during the 63rd IEEE Conference on Decision and Control, held in Milan on December 2024.

Beyond the content discussed in this dissertation, the research activities carried out during the triennium encompassed further studies, resulting in works that I am co-authoring, those in [5] and in [9].

I have made every effort to ensure a consistent notation throughout all the chapters. To support the reader with the necessary general background material, I have included two appendices, Appendix A and Appendix B, at the end of the Thesis. These are intended to provide general methodological or theoretical foundations relevant respectively to each of the two main parts, Part I and Part II, of this work. Whereas, Appendix C summarizes the content of my additional contribution presented in [6].

Moreover, whenever a specific chapter required supplementary theoretical or methodological background, I have introduced dedicated background sections within that chapter itself, in order to preserve the self-contained and localized nature of the presentation.

The full list of papers written during the triennium is reported hereafter.

List of Papers Written by the Author

International Journal Articles

- [1] De Paola, P. F., Paglialonga, A., Palumbo, P., Keshavjee, K., Dabbene, F., and Borri, A., “The long-term effects of physical activity on blood glucose regulation: A model to unravel diabetes progression,” *IEEE Control Systems Letters*, vol. 7, pp. 2916–2921, 2023 (**Distinguished with the IEEE TC-HMS Outstanding Student Paper Award 2024**). DOI: [10.1109/LCSYS.2023.3290774](https://doi.org/10.1109/LCSYS.2023.3290774).
- [2] De Paola, P. F., Borri, A., Dabbene, F., Keshavjee, K., Palumbo, P., and Paglialonga, A., “Modeling the cumulative benefits of regular physical activity on type 2 diabetes progression,” *Computers in Biology and Medicine*, vol. 198, p. 111 194, 2025. DOI: <https://doi.org/10.1016/j.combiomed.2025.111194>.
- [3] De Paola, P. F., Borri, A., Paglialonga, A., Dabbene, F., and Palumbo, P., “Long-term diabetes prevention via physical activity: An output-feedback MPC approach,” *IEEE Control Systems Letters*, 2025. DOI: [10.1109/LCSYS.2025.3578031](https://doi.org/10.1109/LCSYS.2025.3578031).
- [4] De Paola, P. F., Miller, J., Borri, A., Paglialonga, A., and Dabbene, F., “A control system framework for counterfactuals: An optimization based approach,” *arXiv preprint arXiv:2501.12914*, 2025.
- [5] Multerer, L., De Paola, P. F., Lenatti, M., Paglialonga, A., and Azzimonti, L., “Homogenization of ordinary differential equations for the fast prediction of diabetes progression,” *arXiv preprint arXiv:2412.16261*, 2024.

International Conference Proceedings

- [6] De Paola, P.F. and Borri, A. and Paglialonga, A. and Palumbo, P. and Dabbene, F., “Polynomial approximation of regions of attraction via occupation measures: An application to a biological autonomous system,” in *Proceedings of the IEEE 20th International Conference on Automation Science and Engineering (CASE)*, 2024. DOI: [10.1109/CASE59546.2024.10711611](https://doi.org/10.1109/CASE59546.2024.10711611).
- [7] De Paola, P. F., Borri, A., Paglialonga, A., Palumbo, P., and Dabbene, F., “A model-based approach for glucose control via physical activity,” in *Studies in health technology and informatics*, vol. 327, 2025, pp. 27–31. DOI: [10.3233/SHTI250267](https://doi.org/10.3233/SHTI250267).
- [8] De Paola, P. F., Miller, J., Borri, A., Paglialonga, A., and Dabbene, F., “Robust control-driven counterfactual generation for uncertain systems,” in *11th IFAC Symposium on Robust Control Design ROCOND 2025*, vol. 59, 2025, pp. 116–121. DOI: <https://doi.org/10.1016/j.ifacol.2025.10.089>.
- [9] Lenatti, M. and Zaffalon, M. and Antonucci, A. and De Paola, P.F. and Multerer, L. and Mongelli, M. and Paglialonga, A. and Azzimonti, L., “Counterfactual inference using ordinary differential equations to assess the effect of physical activity on type 2 diabetes onset,” in *International Conference on Artificial Intelligence in Medicine*, Springer, 2025, pp. 212–221. DOI: [10.1007/978-3-031-95838-0_21](https://doi.org/10.1007/978-3-031-95838-0_21).

Acknowledgements

I have always believed, and still firmly believe, that research is the unbreakable bond between human beings and life itself. Perhaps because, at its core, it mirrors our inherent sense of incompleteness, the ceaseless desire, the longing, for truth. The ancient Greeks named the research as *ζήτησις* (*zētēsis*), a word that Aristotle, in the *Metaphysics*, invoked to mark the cornerstone in the path toward the truth, the question that precedes science. The word itself is profoundly revealing, for it shares its root with *ζάω* (*zaō*), “to live”, thus sealing a wondrous and intimate union between seeking, truth, and life.

Along my own journey into research, I have had the privilege and the honor of meeting, as Gurdjieff might have said, extraordinary people who have marked the most meaningful steps of these past three years of my life.

Among all, my deepest gratitude goes first and foremost to my two supervisors, Prof. Fabrizio Dabbene and Prof. Alessandro Borri. To both I owe an immense debt of gratitude: they have been true mentors who introduced me to the path of scientific research with wisdom and kindness, each contributing, in their own way, to reshaping my understanding and direction within research. I have truly regarded them, and always will, as my scientific fathers, from whom I sought every spark of knowledge within my reach; for even their smallest gesture of insight remains, to me, a guiding light along the path of learning. Though they introduced me to the theoretical wonders of control science, they also taught me never to lose sight of its call to application. Perhaps this has been their greatest lesson: that theory, and science, is not beautiful in itself, but in its intrinsic vocation to serve. Through the intensity of their passion for research, a passion I have sought to inherit and to make my own, I have come to glimpse what, in my own understanding, is the true meaning of St. Augustine’s “*quasi horizon*” precept: the awareness that research is never a possession, but a perpetual approach, a horizon forever receding, where truth is not reached, but endlessly sought.

I also wish to express my heartfelt gratitude to Prof. Alessia Paglialonga, whose contribution to my studies has been truly significant. I am deeply thankful for the patience with which she guided me, often setting me on the right path and helping me grow both scientifically and professionally. There were countless occasions in which she offered invaluable insights on how to advance the research, suggesting the next steps, and motivating me whenever things did not unfold as expected.

The same gratitude I owe to Prof. Pasquale Palumbo. Many of the pages that follow bear the quiet trace of our conversations, long exchanges in which he shared, with rare elegance, a glimpse of his boundless knowledge. He is, undoubtedly, a brilliant researcher, yet perhaps even more a person of remarkable kindness. His influence runs through this work like a silent thread, and for that, and for much more, I shall remain forever grateful to him.

In June and July 2025, I had the rare privilege of visiting Prof. Didier Henrion in Toulouse, days filled with conversations, insight, and pure wonder about occupation measures and optimal transport. His capacity to reveal simplicity within complexity left a lasting impression, and his teachings will continue to guide me long after this Thesis. His words remain to me as moments of intellectual delight of my doctoral path, and for them, I carry to Prof. Henrion gratitude beyond measure.

To my friends, near and far, whose presence has woven light and meaning into these years, my gratitude is boundless. I think especially of Marta and Davide, fellow travellers who shared with me the wholeness of this journey: its hopes, its fatigue, its joy.

Dulcis in fundo, my most profound and tender gratitude will always belong to all my family, the steadfast ground on which I have always stood, whose love has carried me through every storm.

To my mother Elisabetta, my father Salvatore and to my beloved Simonetta, the primary authors of the chapters of my life, of the pages written so far, and of the finest ones yet to come.

Acronyms

AI	<i>Artificial Intelligence.</i>	2
AP	<i>Artificial Pancreas.</i>	9
CDQDPS	<i>Chinese Da Qing Diabetes Prevention Study.</i>	viii, 7
CED	<i>Cumulative Exercise Dose.</i>	vi, 30
DPP	<i>Diabetes Prevention Program.</i>	7
FDPS	<i>Finnish Diabetes Prevention Study.</i>	viii, 7
IL-6	<i>Interleukin-6.</i>	19
LMI	<i>Linear Matrix Inequalities.</i>	87
LP	<i>Linear Program.</i>	85
ML	<i>Machine Learning .</i>	2
moment-SOS	<i>Moment - Sum of Squares.</i>	81
MPC	<i>Model Predictive Control.</i>	vi, 4
NMPC	<i>Nonlinear Model Predictive Control.</i>	4
ODE	<i>Ordinary Differential Equations.</i>	2
PDE	<i>Partial Differential Equations.</i>	2
PID	<i>Proportional Integral Derivative.</i>	4
PWC	<i>Piecewise-constant.</i>	24
QSSA	<i>Quasi-steady State Representation.</i>	45
SDP	<i>Semidefinite Programming.</i>	91
SMC	<i>Sliding Mode Control.</i>	4
T1D	<i>Type 1 diabetes.</i>	4
T2D	<i>Type 2 diabetes.</i>	vi, 4
WHO	<i>World Health Organization.</i>	6

Introduction

Chapter 1

Introduction

Life sciences involve a broad spectrum of scientific disciplines that deal with living organisms and their interaction with the environment within complex biological systems. Life sciences, spanning from biology, biochemistry, medicine and wide engineering fields, form the backbone of scientific exploration into the mechanisms that govern life. A striking trend in modern research is the convergence of life sciences with engineering and physical sciences, that is progressively contributing to blurring the boundaries between engineering, biology and medicine and spawning new hybrid disciplines and approaches. One of the most impactful challenge within this cross-cutting, interdisciplinary, framework lies at the intersection between engineering and healthcare [1]. Modern healthcare challenges, from chronic disease management to personalized medicine and complex treatment optimization problems, highlight the need to extend biological insights with quantitative engineering approaches. Engineering in healthcare can produce invaluable benefits, among the others identification of patients at risk and of better therapies, clinical decision support, quantitatively based clinical trial design and personalized medicine blending medical expertise with digital tools [2].

A particularly promising advancement emerging from this convergence is the use of high-fidelity computational models of biomedical systems for real-time simulation and optimization. These models can be employed to build *digital twins*, that constitute high-fidelity virtual representations of a physical biological system (such as an individual patient), built by integrating mathematical models with multi-modal data sources like electronic health records, genomics, medical imaging, and real-time sensor readings [3]. These digital counterparts act as dynamical replicas of the underlying biological system [4], encapsulating its state and behavior in silico [5]. Within this framework, models of biomedical systems constitute powerful tools for predictive diagnostics, treatment optimization, and personalized medicine by allowing clinicians to explore therapeutic strategies and forecast disease progression in a virtual environment before applying interventions in real life. However, capturing the full complexity of living systems in such models is non-trivial. Biomedical systems are governed by intricate physiological processes and exhibit significant patient-specific variability, making accurate modeling and control a formidable challenge that demands interdisciplinary expertise and rigorous validation [1].

1.1 Mathematical Modeling in Biomedical Systems and Disease Prevention

Across applications like disease prevention, personalized treatment planning, and system optimization, mathematical modeling provides the quantitative backbone for biomedical systems to run simulations and generate insights. Whether using differential equations for biological processes or *Artificial Intelligence* (AI) and *Machine Learning* (ML) methodologies for data-driven predictions, these models try to capture physiological mechanisms at the population level and inter-individual patient variability, making biomedical systems powerful frameworks for precision medicine and decision support [3], [6].

1.1.1 Mechanistic Models of Physiological Systems and Chronic Diseases

Ordinary Differential Equations (ODE) and *Partial Differential Equations* (PDE) models represent essential tools for describing complex physiological processes and chronic disease dynamics. By encoding mechanistic knowledge (e.g., interaction between biological agents,

mass balance, biochemical reaction kinetics) into a mathematical representation, these models allow to simulate multi-scale systems ranging from cellular pathways to whole organs and organisms. For instance, large-scale integrative models of human physiology can incorporate hundreds of variables across cardiovascular, respiratory, renal, neural, endocrine, and metabolic subsystems. These mechanistic simulations serve as “virtual laboratories” for probing normal and pathological physiology *in silico*. By grounding disease progression into physical and biochemical mechanisms, mathematical models can complement empirical research with a causal understanding of disease processes entailing significant advantages:

- *interpretability*: models are physics informed, with equations and parameters carrying biological meaning, which lends these models a transparency and interpretability often lacking in purely statistical or black-box approaches [7]. Model interpretability has practical meaning, allowing to reveal the patient’s underlying condition;
- *validation*: since model outputs typically correspond to measurable physiological quantities, models can be challenged and refined through direct comparison with experiments or clinical observations; the model validation provides a quantitative assessment to the hypothesis behind some modeling design choices. Furthermore, by explicitly representing the causal physiological mechanisms, mechanistic models can reliably predict system behavior even in simulation scenarios that differ from those used for model calibration.
- *personalization*: as the model parameters are related with physiological aspects, they can be tuned to individual data, effectively representing patient-specific conditions. Patient-calibrated models can then be used to forecast disease trajectories or treatment responses for that specific individual, enabling personalized decision-making.

In light of these distinctive features, mechanistic modeling has thus become a cornerstone in biomedical systems modeling, allowing researchers to simulate chronic disease dynamics under controlled *in silico* conditions.

1.1.2 Individual-level Modeling: Disease Progression and Treatment Optimization

At an individual patient level, ODE- and PDE-based models can be used, for example, to track disease progression over time, simulate therapies, optimize treatment strategies. Mechanistic ODE/PDE models have been developed for a wide range of chronic diseases. For the sake of an example, in diabetes research, compartmental ODE models of glucose–insulin dynamics capture key feedback loops in glycemic regulation [8]. In cancer research, reaction–diffusion PDE models and structured ODE models describe tumor growth and invasion, often incorporating angiogenesis or immune interactions: these models allow the investigation of the mechanisms driving microscopic cell dynamics to scale up to a macroscopic tumor behavior [9]. Similarly, ODE/PDE frameworks exist for chronic respiratory diseases modeling (e.g., models of lung mechanics and gas exchange in chronic obstructive pulmonary disease)[10].

In clinical pharmacology, classical pharmacokinetic/pharmacodynamic models use ODEs to describe drug absorption, distribution, metabolism, and excretion, coupled to drug effect sites. These mechanistic pharmacokinetic/pharmacodynamic models can predict how a specific treatment translates into drug concentrations and therapeutic effects (or toxicities) in a given patient [11]. For instance, multi-compartment ODE models of therapies kinetics and tumor cell kill dynamics have been employed to design optimal dosing schedules in oncology, balancing efficacy against toxicity [12]. As such models explicitly represent how interventions influence biological state variables, they provide a rational basis for treatment optimization. Mechanistic disease models allow patient specific predictions by calibrating model parameters to individual patient data. In this

regard, notable examples can be found in *Type 1 diabetes* (T1D) and *Type 2 diabetes* (T2D) care [13], [14], [15]: in these works researchers validate the models on patients' data with the aim of identifying specific physiological parameters. Validated models can be employed for forecasting disease progression and to inform personalized treatment planning. Similar approaches have been used in a wide spectrum of disease modeling, spanning from autoimmune myocarditis [16], to psoriasis [17], inflammatory bowel diseases [18], phenylketonuria [19], osteoporosis [20], chronic hepatitis C [21], atherosclerosis [22].

1.1.3 Population and Healthcare System Level

Mechanistic modeling approaches also operate at the population and healthcare system level, where they inform public health decisions and resource planning. In epidemiology, compartmental ODE models (such as the well-acknowledged SIR models) have long been used to simulate infectious disease outbreaks [23]. These models divide populations into compartments (susceptible, infectious, recovered, etc.) and use differential equations to capture how disease spreads through contacts over time. The strength of ODE epidemic models lies in their ability to incorporate mechanistic features like transmission rates, incubation periods, and intervention effects. During the COVID-19 pandemic, for example, governments and scientists relied on mechanistic epidemic models to forecast case trajectories and evaluate control strategies. In this regard, a prominent study modeled COVID-19 transmission in 11 European countries and estimated that stringent non-pharmaceutical interventions (e.g. lockdowns) significantly reduced transmission rates [24]. This illustrates how mechanistic models can serve as “virtual laboratories” also at a population level: public health officials can simulate the outcome of various intervention scenarios (school closures, vaccination roll-outs) to inform real-world decisions. Beyond the application in epidemics, system-level modeling is applied to chronic disease burdens [25], [26].

1.2 Control Applications in Healthcare Systems and Disease Prevention

Control theory offers a formal framework for the analysis and regulation of biomedical systems and medical therapies. In human physiology, homeostatic functions (e.g., glucose regulation, blood pressure) are inherently feedback systems [27], thus allowing reproducibility by engineered control systems to stabilize critical variables and prevent disease progression [28], [29]. Control theory provides important tools, such as linear and nonlinear techniques, stability analysis, optimization, to interpret complex biomedical systems, identify disregulated feedback loops in disease and address proper intervention strategies.

1.2.1 Individual-level Control Applications

Control applications at an individual level are widely exploited and include glucose-insulin regulation, cardiovascular regulation, respiratory system and neurological system control. More in detail, control techniques have been successfully employed for the challenging treatment of severe chronic disease as T1D and T2D [28], [30], [31], [32], hypertension [33], cancer [34], acute respiratory failures [35], neurological and psychiatric disorders [36]. In this regard, the most common and employed control techniques are *Model Predictive Control* (MPC) [30] and *Nonlinear Model Predictive Control* (NMPC) [33], *Proportional Integral Derivative* (PID) control [28], although a wide range of methods have been employed across *Sliding Mode Control* (SMC) [37], robust control [38], and adaptive control [39].

1.2.2 Population-level and Public Health Control Applications

Control theory applications in biomedical systems extend to population-level health and resource system. For the sake of an example, with COVID-19 outbreak, several works have investigated control-based approaches to propose mitigation strategies. In the work in [40], the authors exploits a NMPC approach to minimize the impact of non-pharmaceutical interventions while guaranteeing the working capacity of the healthcare systems. Analogously, in [41] an adaptive robust MPC is designed in order to reliably contain the COVID-19 outbreak and show how such control strategy significantly reduced the number of fatalities. Moreover, optimal control has been leveraged to plan vaccine allocation. For instance, in [42] vaccination strategies are planned solving an optimal control problem relying on the daily vaccination rate and the daily testing rate as control variables. Recently, control methods have been used to optimize the resource allocation in hospital emergency department enhancing emergency response and medical services, while ensuring patients' safety [43].

All these control-theoretic design strategies yield notable advantages in healthcare interventions and enhance personalization and individual tailoring. Moreover control-based techniques are inherently designed to ensure stability properties, robustness guarantees and to handle structural constraints.

1.3 An Emerging Trend in Biomedical Systems Applications

An emerging trend in biomedical systems is the use of *counterfactual reasoning*, which concerns formulating hypothetical 'what-if' alternatives about interventions or conditions to explore alternative scenarios. Researchers and clinicians naturally engage in this kind of conditional reasoning when making treatment decisions by envisioning potential outcomes for patients, essentially counterfactual thinking about alternative care scenarios [44]. In the context of disease modeling and control, counterfactual analysis is emerging as a very promising tool mainly in the framework of AI applications. Counterfactual approaches are increasingly being applied at the individual patient level to support personalized decision-making. For instance, in the context of explainable AI, *counterfactual explanations* have been used in the field of T2D control to suggest minimal changes in a patient's risk factors that would alter a model prediction or diagnosis: this might mean finding the smallest adjustment in biomarkers that could change a patient's classification from "at-risk" to "healthy", thus providing actionable insights for prevention [45], [46], [47], [48]. Beyond individual care, counterfactual analysis holds comparable value at the population and policy level. In this regard, frameworks have been developed to simulate hypothetical 'what-if' simulation scenarios by essentially re-running the epidemic dynamics under alternate policy inputs [49]. Together, these examples show how counterfactual reasoning is gaining prominence in biomedical systems, while also signaling its promise to extend into other domains like control theory.

1.4 Thesis Positioning and Motivation

1.4.1 Type 2 Diabetes

T2D is a chronic disease characterized by impairment in blood glucose regulation, entailing chronic hyperglycemia and severe related issues [50].

T2D has become a fast-growing global epidemic, rising from around 108 million cases in 1980 to 422 million in 2014. As of 2024, an estimated 589 million adults worldwide (1 in 9) have diabetes, over 90% of which are diagnosed with T2D, and this is projected to reach 853 million by 2050 if current trends continue. Prevalence has climbed sharply in every region, driven by aging populations, urbanization, obesity, and sedentary lifestyles [51]. Low- and middle-income countries have seen the fastest growth, but even in Europe diabetes rates are increasing. In Europe about 48 million adults ($\approx 11\%$) have diabetes, with all European Union countries having all experienced a rising prevalence since 1990

[52], though the increase is generally slower than in developing regions. Some countries (e.g. Luxembourg, Ireland) saw over a 2-fold relative jump in prevalence over 30 years [52].

The economic burden of T2D is enormous and rising. Currently, North America, Western Pacific and Europe are estimated to spend respectively around 415, 240 and 190 billion USD on diabetes, accounting for around the 43%, 25% and 20% of the total healthcare expenditures on diabetes [53]. The total amount of social expenditures on diabetes is forecasted to exceed 1000 billion USD by 2045, with an increase of the 9% compared to those in 2021 [53].

T2D is a leading cause of blindness, kidney failure, heart attacks, stroke, and limb amputations, in light of severe damages to blood vessels heart, eyes, kidneys and nerves. Over 50% of people with T2D develop at least one complication, with cardiovascular disease being the primary cause of diabetes-related morbidity and mortality [54].

The rising prevalence of diabetes and obesity has already placed a substantial strain on healthcare systems and this burden is expected to grow in the coming years. While T2D remains more widespread in the general population, T1D has traditionally been the most common chronic condition among children. However, with the increasing incidence of T2D in younger populations, this trend may be reversing [55].

T2D develops from a combination of peripheral insulin resistance and progressive loss of pancreatic beta cell mass and function, resulting in chronic hyperglycemia. Central to its aetiology are unhealthy lifestyle habits, that in the long term promote low-grade systemic inflammation that affects insulin signaling, driving insulin resistance and impairing beta cells integrity [56], [57]. Indeed, the global surge in T2D is tightly linked to sedentary lifestyles, unhealthy diets, and associated weight gain, compounded in many regions by socioeconomic disparities that limit access to healthy food and safe spaces for exercise [58]. Chronic psychosocial stress and other urbanization-related factors (such as inadequate sleep and pollution) are also thought to exacerbate insulin resistance and systemic inflammation, highlighting the complex social determinants of T2D risk.

1.4.1.1 The Effect of Physical Activity on T2D

Although therapeutic interventions, such as through insulin administration or metformin administration, play an important role in the management of the disease, the growing prevalence highly recommend the need to prioritize T2D prevention as a primary public health objective [57]. In this regard, extensive clinical evidence acknowledges the primary role that physical activity can have in T2D prevention. Regular exercise can enhance insulin sensitivity and can help reducing low grade system inflammation [59], [60], thereby potentially counterbalancing insulin resistance and beta-cell mass degradation [60], [61]. The mechanisms by which physical activity exerts these benefits are wide: increased glucose uptake by skeletal muscle, enhanced muscle mass and oxidative capacity, reductions in visceral fat mass, and, predominantly, release of anti-inflammatory myokines that contribute to improved insulin sensitivity, beta cell mass function and glucose homeostasis [60], [61], [62].

In light of this, *World Health Organization* (WHO) suggests at least 150 minutes of moderate-intensity exercise per week or alternatively 75 minutes of vigorous-intensity exercise per week for adults as a minimum target to improve metabolic health and reduce risk of developing chronic diseases [63], highly recommending to go beyond the suggested minimum as larger benefits are expected with larger amount of exercise. Indeed, evidence shows a clear dose-response relationship between physical activity and diabetes risk [64], [65]: for instance, large scale trials in [64], [65], involving up to nearly 100.000 participants, showed that an increase in physical exercise, for example in terms of total minutes per week or in terms of intensity of the exercise session, can contribute to proportionally lower the risk of developing T2D. Both structured exercise programs and improvements in everyday lifestyle activity (such as walking more) have shown preventive efficacy: a low intensity 20-min daily brisk walk could reduce the risk of developing T2D by almost

20% [65]. Several trials have assessed the long-term benefits promoted by regular exercise and lifestyle interventions: the *Diabetes Prevention Program* (DPP), the *Finnish Diabetes Prevention Study* (FDPS) and *Chinese Da Qing Diabetes Prevention Study* (CDQDPS), trial showed a 30–60% reduction in progression from prediabetes to T2D with intensive lifestyle interventions (dietary changes and regular exercise) [57], [66], [67], [68]. The protective effect of physical activity appears durable: extended follow-ups of lifestyle intervention cohorts show sustained diabetes prevention years later. In the CDQDPS trial, for instance, individuals subject to exercise and diet interventions retained a significantly lower T2D incidence (around 40% relative reduction) even after 23–30 years [57], [68] and similar results were outlined in the FDPS trial after seven years of intervention [57], [67]. Notably, exercise benefits are observed independently on weight loss, as assessed in the CDQDPS [57].

1.4.2 Mathematical Models of Glucose–insulin Homeostasis and Diabetes Progression

Understanding and predicting glucose–insulin dynamics in diabetes has motivated the development of mathematical models ranging from simple “minimal” models to complex multi-organ physiological simulations. This section is devoted to the discussion of the main models of glucose–insulin regulation and diabetes progression. More in detail, minimal models and maximal compartmental models are discussed, as well as models capturing long-term T2D progression.

1.4.2.1 Minimal Models of Glucose–insulin Homeostasis

The paradigm of minimal modeling was introduced by Bergman *et al.* [13] in 1979, with the development of a model consisting of two ODEs, describing the dynamics of glucose and insulin concentration. The insulin dynamics is described by means of a remote compartment representing in a minimal way the biological machinery underlying the delay in insulin action and liver glucose output suppression [13], [69].

The idea of defining a “minimal” model was motivated by the goal of quantitatively estimating insulin sensitivity fitting the model to intravenous glucose tolerance test data, thus also allowing the possibility of validation of the results with glucose clamp methods in humans.

Moreover, the model also allowed to explore the relationship between insulin sensitivity and insulin secretion, thus providing valuable insights on the quantitative assessment of the Disposition Index. The Disposition Index is inherently related to the genetic of the simulated individual and results to be a valuable parameter for diabetes diagnosis.

1.4.2.2 Maximal Models of Glucose–insulin Homeostasis

In addition to minimal models, the literature proposes also maximal, full-scale models of glucose insulin homeostasis. Among these, the most comprehensive and widely used ones are summarized in the following.

- Model by Sorensen *et al.*: this model [70] is one of the more complex, including more than 20 nonlinear ODEs and more than 130 parameters, representing glucose and insulin dynamics across organs and tissues: brain, heart, lungs, liver, gut, kidneys, muscle and adipose tissues. The model explicitly includes processes such as insulin-dependent glucose uptake in peripheral tissue, hepatic glucose production, renal glucose excretion, and insulin secretion and clearance, with parameter values selected from careful literature research [70], [71]. The model provides a detailed mechanistic representation of glucose homeostasis, making it valuable for understanding system-wide effects and was originally formulated to simulate normal physiology and insulin therapy in diabetes, and it laid the groundwork for later patient simulators [8], [71]. Despite the model shows, as valuable strength, a notable completeness, its high

complexity limits the applicability for instance in control algorithms design because of the computational burden required to simulate it [71].

- UVA/Padova model: the UVA/Padova simulator is the most known and widely employed model for glucose-insulin system [72], [73], [74], [75]. It is structured as a series of interacting compartmental sub-models representing key processes: gastric emptying and intestinal glucose absorption, the distribution and utilization of glucose in the body, insulin kinetics (secretion or exogenous absorption, and clearance), and insulin action on target tissues, also accounting for multiple daily meals and glycemic variability according to circadian variations. Each subsystem is described by differential equations with parameters originally tuned to human data (e.g. using triple-tracer meal studies), ensuring that the simulator reproduces the major glucose fluxes (meal appearance, endogenous production, utilization) and insulin responses observed experimentally [73], [74], [75]. The high level of detail and validation has made the UVA/Padova simulator a de facto standard for in-silico testing: notably, the simulator accuracy and physiological realism led the U.S. Food and Drug Administration to accept it as a substitute for tests in preclinical testing of insulin therapies and artificial pancreas algorithms.
- Model by Hovorka *et al.*: among the maximal models, the model by Hovorka *et al.* is the simplest and primarily developed for the design of control algorithms [76], [77], [78]. It includes two differential equation for the glucose dynamics, associated with plasma and tissue glucose compartments. Moreover, it includes a representation of an insulin dynamics subsystem and an additional insulin action subsystem involving three compartments related with insulin effect on distribution/transport of glucose, insulin effect on glucose disposal and insulin effect on endogenous glucose production.

1.4.2.3 Mathematical Models of Diabetes Progression

T2D is characterized by persistent impairment in glucose-insulin dynamics, produced by chronic progressive changes, notably increased insulin resistance and decreased beta cell function over years. Mathematical models have been developed to capture these long-term processes, often by augmenting glucose-insulin models with slower “disease state” variables (e.g. beta cell mass). Such models illustrate how T2D develops from normoglycemia in the long term and how interventions might alter the disease trajectory.

- Model by Topp *et al.*: the model developed by Topp *et al.* [79] is acknowledged to be a true landmark in the field of diabetes modeling and control and has been the first model to account for the effect of beta-cell mass and diabetes evolution over years. The model consists of four state variables, including blood glucose and blood insulin concentration, beta-cell mass, and insulin sensitivity. Insulin sensitivity decay dictates diabetes risk evolution and sets the modeling basis for a key core of the model that lies in the description of the effect of glucotoxicity. More in detail, the model is able to reproduce the detrimental effect that chronic hyperglycemia produces on beta-cell mass (reducing beta cell replication and increasing beta cell apoptosis), promoting a positive feedback that drives T2D onset. Despite being relatively simple in terms of state space representation, the model exhibits multiple steady-states, consistently with T2D evolution, represented by two equilibrium points, associated respectively with normoglycemic and hyperglycemic conditions. The transitions between these states explain how gradual insulin resistance could initially be compensated by beta cell mass expansion (i.e., dynamical compensation), until an irreversible condition arises due to glucotoxicity. Moreover, the model highlights how an intervention improving insulin sensitivity or reducing glucotoxic stress on beta cell mass could potentially delay or revert T2D progression. Although this model relies on some simplified modeling assumption (e.g. in the beta-cell dynamics) it was pioneering and laid the ground for further diabetes progression models.

- Models by De Gaetano *et al.*: relying on the main findings and concepts developed in the model in [79], De Gaetano *et al.* [14], [80] developed a series of models to quantitatively simulate T2D progression over years. A first version of the model [80] presented a comprehensive description of short-term glucose–insulin dynamics with slower processes for insulin resistance and beta-cell mass adaptation. A distinctive feature of the model lies in the ability to model the effect of age of diabetes progression and of the dynamics of glycosylated hemoglobin. Moreover, the model introduced the concept of pancreatic reserve. The model in [80] has been further elaborated and extended in [14]. The version in [14] is built on a multi-temporal scale and multi-compartmental model fashion, including subsystems accounting for daily glucose–insulin dynamics, slow variables dynamics, nephron dynamics. Moreover, this model allows the description of interventions or therapies like metformin administration or lifestyle intervention. Notably, this model has been validated against real clinical data from the DPP study [14].
- Model by Ha *et al.*: this model [81] proposes an evolution of the model by Topp *et al.* [79]. Specifically, the model extends the concept of dynamical compensation also to beta-cell function (in terms of insulin secretion capability) and elaborates the description of beta-cell replication signaling addressing the modeling of beta-cell workload as secretion rate per cell. Moreover, it investigates the role of homeostatic regulation of workload on the general homeostatic glucose regulation. The model is able to represent the distinctive bistability property of T2D, analogously with respect to the model in [79] and involves six state variables and additional algebraic state-depending functions. In this regard, the model represents a valuable compromise, offering sufficient physiological detail to ensure accurate and meaningful biological description while showing a level of complexity that remains tractable for analysis and control design.

1.4.3 Control Applications in Glucose Regulation

The wide interest for diabetes modeling promoted the rise of a fruitful emerging line of research named *Artificial Pancreas* (AP), for both T1D and T2D. AP refers to closed-loop systems for blood glucose regulation that integrate continuous glucose sensors, control algorithms and insulin infusion pumps [82]. Over the last decades, numerous groups worldwide have contributed to AP development, advancing both theoretical control strategies and enabling technologies. Importantly, AP encompasses more than just the first commercial closed-loop devices introduced in recent years, but it refers to the entire corpus of closed-loop control approaches for automated glucose regulation.

Within this framework, over years, both model-free and model-based control techniques have been leveraged to design control algorithms that set the core of glucose regulation. In this regard, model-free techniques do not rely explicitly on glucose–insulin physiological model and allow for control design based on simplified strategies (e.g., PID controllers) [83], [84], [85]. On the other hand, model-based approaches explicitly incorporate mathematical models of glucose–insulin dynamics into the controller design [86]. Under this perspective, model-based techniques have shown undisputable advantages in the physiological interpretability of model parameters (e.g., insulin sensitivity) and the ability to test algorithms *in silico* on virtual patient populations. Among the others, control methods exploiting the model-based paradigm have been investigated in the context of:

- adaptive control, to automatically adjust controller parameters over time to adapt to the patient’s evolving condition [87];
- nonlinear and nonlinear time-delay control, to explicitly handle the nonlinear insulin–glucose dynamics and inherent delays in insulin action [32], [88], [89], [90];
- robust control, to maintain stability and performance despite uncertainties in patient parameters and to guarantee safety in worst-case scenarios [86];

- model predictive control, for constraint-handling capability of predicting future blood glucose and optimize therapy delivery over a moving horizon [91], [92].

Since these model-based techniques have demonstrated reliable performance across a wide range of simulation and preclinical studies, they have paved the way for the development of commercial artificial pancreas devices and large-scale cohort trials. In T2D, pilot randomized controlled trials have confirmed that fully closed-loop systems can significantly increase time-in-range and reduce hyperglycemia [93], [94], [95]. These advances highlights how the maturation of model-based control strategies can be directly translated into clinically approved closed-loop systems.

1.4.4 Thesis Rationale and Scientific Contributions

Within this framework, this Thesis addresses two primary lines of research, as detailed below.

1.4.4.1 Mathematical Modeling and Control of T2D Progression via Physical Activity

The increasing prevalence of T2D and the limitations of conventional treatment regimens have motivated significant research into more precise, personalized strategies for disease management. While artificial pancreas systems have achieved notable success in T2D short-term control, their application to long-term, preventive, disease progression control remains a largely underexplored area.

More in detail, although exercise is universally recommended as a cornerstone of diabetes prevention and therapy, as discussed above, most of the existing models of glucose–insulin dynamics and T2D progression do not explicitly account for how structured physical activity affects disease progression over time (e.g., by improving insulin sensitivity, altering beta-cell function decline) and do not provide interpretable, mechanistic, descriptions of the involved physiological machinery underlying exercise benefits. Moreover, as a consequence, control strategies to date have been focusing mainly on insulin infusion as control input, considered to be the last mile for T2D treatment, when the course of the diseases is yet not reversible, with exercise typically treated as an external disturbance or an unmodeled source of variability. This Thesis aims at bridging these gaps and, in this regard, its contribution is two-fold:

- the development of a novel mathematical model of T2D progression suitable to explain and quantitatively assess the long-term effect of exercise, incorporating a mechanistic description of the involved biological processes that drive these benefits.
- the design of model-based control methods that explicitly leverage physical activity as control input for long-term T2D prevention.

By advancing modeling and control techniques, the Thesis contributes new tools to support personalized, physics-informed T2D management and prevention that account for the crucial role played by physical activity.

1.4.4.2 Counterfactuals for Dynamical Systems

From a methodological perspective, to broaden the theoretical foundations of T2D control, this Thesis introduces a proposal for extending the *counterfactual reasoning* framework within a rigorous system-theoretic setting. The motivation relies on a precise rationale. As discussed, counterfactual reasoning has become a central tool in AI for T2D decision support, enabling personalized “what-if” analysis of therapies, timing, and adherence; yet, AI-driven counterfactuals, mainly based on data-driven approaches, are agnostic about the dynamics underlying the phenomenon under consideration, thus overlooking dynamical feasibility, safety constraints, or mechanistic interpretability. Hence, inheriting this paradigm into control and dynamical systems theory allows to complement and

enhance AI-based approaches with the physics-informed setting typical of system and control theory, thereby turning counterfactuals into actionable, model-consistent objects for decision-making. While focused here for T2D control, the framework is general: it provides a template for biomedical systems applications that, in a system-level perspective, can complement AI practices, enhance model-based control with a principled “what-if” analysis and extend the design space for interpretable and robust intervention planning in general biomedical and healthcare systems.

1.5 Thesis Structure

This Thesis includes two parts.

Part I is devoted to the illustration of the mathematical modeling and control of T2D progression under the effect of physical activity and involves **Chapter 2**, **Chapter 3** and **Chapter 4**.

In **Chapter 2** we introduce the model, describe its features and assess the validity of the model predictions in a broad range of simulation scenarios.

In **Chapter 3** we perform a rigorous qualitative analysis of the steady-state behavior of the model and show some general formal properties.

In **Chapter 4** we propose a model-based control approach relying on a MPC formulation leveraging the effect of exercise as control for T2D progression.

Part II is devoted to the discussion of the control system framework for counterfactuals and includes **Chapter 5**, **Chapter 6** and **Chapter 7**.

Chapter 5 deals with a general presentation about the classical, philosophical, notion of counterfactuals, its recent introduction in the AI framework and about our formulation in the field of dynamical systems and control theory.

Finally, **Chapter 6** and **Chapter 7** discuss two different proposals for counterfactuals in the control-system framework, respectively based on an optimal control and optimal transport formulation.

References

- [1] Subramaniam, S. et al., “Grand challenges at the interface of engineering and medicine,” *IEEE Open Journal of Engineering in Medicine and Biology*, vol. 5, pp. 1–13, 2024.
- [2] Tresp, V., Overhage, J. M., Bundschuh, M., Rabizadeh, S., Fasching, P. A., and Yu, S., “Going digital: A survey on digitalization and large-scale data analytics in healthcare,” *Proceedings of the IEEE*, vol. 104, no. 11, pp. 2180–2206, 2016.
- [3] Katsoulakis, E. et al., “Digital twins for health: A scoping review,” *NPJ digital medicine*, vol. 7, no. 1, p. 77, 2024.
- [4] Barricelli, B. R., Casiraghi, E., and Fogli, D., “A survey on digital twin: Definitions, characteristics, applications, and design implications,” *IEEE access*, vol. 7, pp. 167 653–167 671, 2019.
- [5] Croatti, A., Gabellini, M., Montagna, S., and Ricci, A., “On the integration of agents and digital twins in healthcare,” *Journal of Medical Systems*, vol. 44, no. 9, p. 161, 2020.
- [6] De Domenico, M. et al., “Challenges and opportunities for digital twins in precision medicine from a complex systems perspective,” *npj Digital Medicine*, vol. 8, no. 1, p. 37, 2025.
- [7] Baker, R. E., Pena, J.-M., Jayamohan, J., and Jérusalem, A., “Mechanistic models versus machine learning, a fight worth fighting for the biological community?” *Biology letters*, vol. 14, no. 5, p. 20 170 660, 2018.

- [8] Pompa, M., Panunzi, S., Borri, A., and De Gaetano, A., “A comparison among three maximal mathematical models of the glucose-insulin system,” *PloS one*, vol. 16, no. 9, e0257789, 2021.
- [9] Lorenzo, G. et al., “Patient-specific, mechanistic models of tumor growth incorporating artificial intelligence and big data,” *Annual Review of Biomedical Engineering*, vol. 26, 2024.
- [10] Kurian, V. et al., “Systems engineering approach to modeling and analysis of chronic obstructive pulmonary disease,” *ACS omega*, vol. 8, no. 23, pp. 20 524–20 535, 2023.
- [11] Zou, H., Banerjee, P., Leung, S. S. Y., and Yan, X., “Application of pharmacokinetic-pharmacodynamic modeling in drug delivery: Development and challenges,” *Frontiers in pharmacology*, vol. 11, p. 997, 2020.
- [12] Vanoni, M., Palumbo, P., Busti, S., and Alberghina, L., “A critical review of multiscale modeling for predictive understanding of cancer cell metabolism,” *Current Opinion in Systems Biology*, p. 100 531, 2024.
- [13] Bergman, R. N., Ider, Y. Z., Bowden, C. R., and Cobelli, C., “Quantitative estimation of insulin sensitivity,” *American Journal of Physiology-Endocrinology And Metabolism*, vol. 236, no. 6, E667, 1979.
- [14] De Gaetano, A. and Hardy, T. A., “A novel fast-slow model of diabetes progression: Insights into mechanisms of response to the interventions in the diabetes prevention program,” *PLoS One*, vol. 14, no. 10, e0222833, 2019.
- [15] Richter, L. R. et al., “Data assimilation on mechanistic models of glucose metabolism predicts glycemic states in adolescents following bariatric surgery,” *Frontiers in Physiology*, vol. 13, p. 923 704, 2022.
- [16] Vegt, S. A. van der, Polonchuk, L., Wang, K., Waters, S. L., and Baker, R. E., “Mathematical modelling of autoimmune myocarditis and the effects of immune checkpoint inhibitors,” *Journal of Theoretical Biology*, vol. 537, p. 111 002, 2022.
- [17] Shmarov, F., Smith, G. R., Weatherhead, S. C., Reynolds, N. J., and Zuliani, P., “Individualised computational modelling of immune mediated disease onset, flare and clearance in psoriasis,” *PLOS Computational Biology*, vol. 18, no. 9, e1010267, 2022.
- [18] Park, A., Kim, S., Jung, I. H., and Byun, J. H., “An immune therapy model for effective treatment on inflammatory bowel disease,” *PLoS One*, vol. 15, no. 9, e0238918, 2020.
- [19] Charbonneau, M. R. et al., “Development of a mechanistic model to predict synthetic biotic activity in healthy volunteers and patients with phenylketonuria,” *Communications Biology*, vol. 4, no. 1, p. 898, 2021.
- [20] Jörg, D. J. et al., “Modeling osteoporosis to design and optimize pharmacological therapies comprising multiple drug types,” *Elife*, vol. 11, e76228, 2022.
- [21] Etzion, O. et al., “Response guided therapy for reducing duration of direct acting antivirals in chronic hepatitis C infected patients: A pilot study,” *Scientific Reports*, vol. 10, no. 1, p. 17 820, 2020.
- [22] Pan, J. et al., “A prediction tool for plaque progression based on patient-specific multi-physical modeling,” *PLoS Computational Biology*, vol. 17, no. 3, e1008344, 2021.
- [23] Giordano, G. et al., “Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy,” *Nature medicine*, vol. 26, no. 6, pp. 855–860, 2020.
- [24] Flaxman, S. et al., “Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe,” *Nature*, vol. 584, no. 7820, pp. 257–261, 2020.

- [25] Clennin, M., Homer, J., Erkenbeck, A., and Kelly, C., “Evaluating public health efforts to prevent and control chronic disease: A systems modeling approach,” *Systems*, vol. 10, no. 4, p. 89, 2022.
- [26] Jones, A. P., Homer, J. B., Murphy, D. L., Essien, J. D., Milstein, B., and Seville, D. A., “Understanding diabetes population dynamics through simulation modeling and experimentation,” *American journal of public health*, vol. 96, no. 3, pp. 488–494, 2006.
- [27] Alon, U., *An introduction to systems biology: design principles of biological circuits*. Chapman and Hall/CRC, 2019.
- [28] Bequette, B. W., “Algorithms for a closed-loop artificial pancreas: The case for model predictive control,” *Journal of diabetes science and technology*, vol. 7, no. 6, pp. 1632–1643, 2013.
- [29] Angarita-Rodríguez, A., González-Giraldo, Y., Rubio-Mesa, J. J., Aristizábal, A. F., Pinzón, A., and González, J., “Control theory and systems biology: Potential applications in neurodegeneration and search for therapeutic targets,” *International Journal of Molecular Sciences*, vol. 25, no. 1, p. 365, 2023.
- [30] Toffanin, C., Messori, M., Di Palma, F., De Nicolao, G., Cobelli, C., and Magni, L., *Artificial pancreas: Model predictive control design from clinical experience*, 2013.
- [31] Doyle III, F. J., Huyett, L. M., Lee, J. B., Zisser, H. C., and Dassau, E., “Closed-loop artificial pancreas systems: Engineering the algorithms,” *Diabetes care*, vol. 37, no. 5, pp. 1191–1197, 2014.
- [32] Borri, A., Pola, G., Pepe, P., Di Benedetto, M. D., and Palumbo, P., “Symbolic control design of an artificial pancreas for type-2 diabetes,” *IEEE Transactions on Control Systems Technology*, vol. 30, no. 5, pp. 2131–2146, 2021.
- [33] Yao, Y. and Kothare, M. V., “Nonlinear closed-loop predictive control of heart rate and blood pressure using vagus nerve stimulation: An in silico study,” *IEEE Transactions on Biomedical Engineering*, vol. 70, no. 10, pp. 2764–2775, 2023.
- [34] Liliopoulos, S. G., Stavrakakis, G. S., and Dimas, K. S., “Linear and non-linear optimal control methods to determine the best chemotherapy schedule for most effectively inhibiting tumor growth,” *Biomedicines*, vol. 13, no. 2, p. 315, 2025.
- [35] Li, H. and Haddad, W. M., “Model predictive control for a multicompartiment respiratory system,” *IEEE Transactions on Control Systems Technology*, vol. 21, no. 5, pp. 1988–1995, 2012.
- [36] Parastarfeizabadi, M. and Kouzani, A. Z., “Advances in closed-loop deep brain stimulation devices,” *Journal of neuroengineering and rehabilitation*, vol. 14, pp. 1–20, 2017.
- [37] Bakouri, M. A., Salamonsen, R. F., Savkin, A. V., AlOmari, A.-H. H., Lim, E., and Lovell, N. H., “A sliding mode-based Starling-like controller for implantable rotary blood pumps,” *Artificial organs*, vol. 38, no. 7, pp. 587–593, 2014.
- [38] Bakouri, M. et al., “An optimal H-infinity controller for left ventricular assist devices based on a Starling-like controller: A simulation study,” *Mathematics*, vol. 10, no. 5, p. 731, 2022.
- [39] Silva Lima, G. da, Amorim Savi, M., and Moreira Bessa, W., “Adaptive control of cardiac rhythms,” *Scientific Reports*, vol. 14, no. 1, p. 23284, 2024.
- [40] Carli, R., Cavone, G., Epicoco, N., Scarabaggio, P., and Dotoli, M., “Model predictive control to mitigate the COVID-19 outbreak in a multi-region scenario,” *Annual Reviews in Control*, vol. 50, pp. 373–393, 2020.
- [41] Köhler, J., Schwenkel, L., Koch, A., Berberich, J., Pauli, P., and Allgöwer, F., “Robust and optimal predictive control of the COVID-19 outbreak,” *Annual Reviews in Control*, vol. 51, pp. 525–539, 2021.

- [42] Olivares, A. and Staffetti, E., “Optimum control-based vaccination and testing strategies for COVID-19,” *Comput. Methods Programs Biomed*, vol. 211, p. 106411, 2021.
- [43] Wang, C.-H., Tian, R., Hu, K., Chen, Y.-T., and Ku, T.-H., “A Markov decision optimization of medical service resources for two-class patient queues in emergency departments via particle swarm optimization algorithm,” *Scientific Reports*, vol. 15, no. 1, p. 2942, 2025.
- [44] Lee, S.-I. and Topol, E. J., “The clinical potential of counterfactual AI models,” *The Lancet*, vol. 403, no. 10428, p. 717, 2024.
- [45] Lenatti, M., Carlevaro, A., Guergachi, A., Keshavjee, K., Mongelli, M., and Paglialonga, A., “A novel method to derive personalized minimum viable recommendations for type 2 diabetes prevention based on counterfactual explanations,” *Plos one*, vol. 17, no. 11, e0272825, 2022.
- [46] Console, D. et al., “Exploring prediabetes pathways using explainable AI on data from electronic medical records,” *Studies in health technology and informatics*, vol. 316, pp. 736–740, 2024.
- [47] Lenatti, M. and Zaffalon, M. and Antonucci, A. and De Paola, P.F. and Multerer, L. and Mongelli, M. and Paglialonga, A. and Azzimonti, L., “Counterfactual inference using ordinary differential equations to assess the effect of physical activity on type 2 diabetes onset,” in *International Conference on Artificial Intelligence in Medicine*, Springer, 2025, pp. 212–221. DOI: [10.1007/978-3-031-95838-0_21](https://doi.org/10.1007/978-3-031-95838-0_21).
- [48] Carlevaro, A., Lenatti, M., Paglialonga, A., and Mongelli, M., “Counterfactual building and evaluation via explainable support vector data description,” *IEEE Access*, vol. 10, pp. 60849–60861, 2022.
- [49] She, B., Smith, R. L., Pytlarz, I., Sundaram, S., and Paré, P. E., “A framework for counterfactual analysis, strategy evaluation, and control of epidemics using reproduction number estimates,” *PLoS computational biology*, vol. 20, no. 11, e1012569, 2024.
- [50] Whelan, M. and Bell, L., “The English national health service diabetes prevention programme (NHS DPP): A scoping review of existing evidence,” *Diabetic Medicine*, vol. 39, no. 7, e14855, 2022.
- [51] Standl, E., Khunti, K., Hansen, T. B., and Schnell, O., “The global epidemics of diabetes in the 21st century: Current situation and perspectives,” *European journal of preventive cardiology*, vol. 26, no. 2_suppl, pp. 7–14, 2019.
- [52] Goodall, R. et al., “Trends in type 2 diabetes mellitus disease burden in european union countries between 1990 and 2019,” *Scientific Reports*, vol. 11, no. 1, p. 15356, 2021.
- [53] Sun, H. et al., “IDF diabetes atlas: Global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045,” *Diabetes research and clinical practice*, vol. 183, p. 109119, 2022.
- [54] Zheng, Y., Ley, S. H., and Hu, F. B., “Global aetiology and epidemiology of type 2 diabetes mellitus and its complications,” *Nature reviews endocrinology*, vol. 14, no. 2, pp. 88–98, 2018.
- [55] Fagot-Campagna, A. et al., “Type 2 diabetes among north american children and adolescents: An epidemiologic review and a public health perspective,” *The Journal of pediatrics*, vol. 136, no. 5, pp. 664–672, 2000.
- [56] Rohm, T. V., Meier, D. T., Olefsky, J. M., and Donath, M. Y., “Inflammation in obesity, diabetes, and related disorders,” *Immunity*, vol. 55, no. 1, pp. 31–55, 2022.
- [57] Lu, X. et al., “Type 2 diabetes mellitus in adults: Pathogenesis, prevention and therapy,” *Signal Transduction and Targeted Therapy*, vol. 9, no. 1, p. 262, 2024.

- [58] Ong, K. et al., “GBD 2021 diabetes collaborators. global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: A systematic analysis for the global burden of disease study 2021,” *Lancet*, vol. 402, pp. 203–234, 2023.
- [59] Bird, S. R. and Hawley, J. A., “Update on the effects of physical activity on insulin sensitivity in humans,” *BMJ open sport & exercise medicine*, vol. 2, no. 1, e000143, 2017.
- [60] Petersen, A. M. W. and Pedersen, B. K., “The anti-inflammatory effect of exercise,” *Journal of applied physiology*, vol. 98, no. 4, pp. 1154–1162, 2005.
- [61] Curran, M. et al., “The benefits of physical exercise for the health of the pancreatic β -cell: A review of the evidence,” *Experimental Physiology*, vol. 105, no. 4, pp. 579–589, 2020.
- [62] Morettini, M., Palumbo, M. C., Sacchetti, M., Castiglione, F., and Mazza, C., “A system model of the effects of exercise on plasma interleukin-6 dynamics in healthy individuals: Role of skeletal muscle and adipose tissue,” *PloS one*, vol. 12, no. 7, e0181224, 2017.
- [63] Bull, F. C. et al., “World Health Organization 2020 guidelines on physical activity and sedentary behaviour,” *British journal of sports medicine*, vol. 54, no. 24, pp. 1451–1462, 2020.
- [64] Boonpor, J. et al., “Dose–response relationship between device-measured physical activity and incident type 2 diabetes: Findings from the UK biobank prospective cohort study,” *BMC medicine*, vol. 21, no. 1, p. 191, 2023.
- [65] Strain, T. et al., “Quantifying the relationship between physical activity energy expenditure and incident type 2 diabetes: A prospective cohort study of device-measured activity in 90,096 adults,” *Diabetes Care*, vol. 46, no. 6, pp. 1145–1155, 2023.
- [66] Diabetes Prevention Program Research Group, “Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin,” *New England journal of medicine*, vol. 346, no. 6, pp. 393–403, 2002.
- [67] Lindström, J. et al., “Sustained reduction in the incidence of type 2 diabetes by lifestyle intervention: Follow-up of the finnish diabetes prevention study,” *The Lancet*, vol. 368, no. 9548, pp. 1673–1679, 2006.
- [68] Pan, X.-R. et al., “Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance: The Da Qing IGT and diabetes study,” *Diabetes care*, vol. 20, no. 4, pp. 537–544, 1997.
- [69] Bergman, R. N., “Origins and history of the minimal model of glucose regulation,” *Frontiers in endocrinology*, vol. 11, p. 583016, 2021.
- [70] Sorensen, J. T., “A physiologic model of glucose metabolism in man and its use to design and assess improved insulin therapies for diabetes,” Ph.D. dissertation, Massachusetts Institute of Technology, 1985.
- [71] Panunzi, S., Pompa, M., Borri, A., Piemonte, V., and De Gaetano, A., “A revised Sorensen model: Simulating glycemic and insulinemic response to oral and intravenous glucose load,” *Plos one*, vol. 15, no. 8, e0237215, 2020.
- [72] Dalla Man, C., Raimondo, D. M., Rizza, R. A., and Cobelli, C., *Gim, simulation software of meal glucose—insulin model*, 2007.
- [73] Dalla Man, C., Rizza, R. A., and Cobelli, C., “Meal simulation model of the glucose-insulin system,” *IEEE Transactions on biomedical engineering*, vol. 54, no. 10, pp. 1740–1749, 2007.
- [74] Man, C. D., Micheletto, F., Lv, D., Breton, M., Kovatchev, B., and Cobelli, C., “The uva/padova type 1 diabetes simulator: New features,” *Journal of diabetes science and technology*, vol. 8, no. 1, pp. 26–34, 2014.

- [75] Visentin, R. et al., “The uva/padova type 1 diabetes simulator goes from single meal to single day,” *Journal of diabetes science and technology*, vol. 12, no. 2, pp. 273–281, 2018.
- [76] Hovorka, R. et al., “Partitioning glucose distribution/transport, disposal, and endogenous production during IVGTT,” *American Journal of Physiology-Endocrinology and Metabolism*, vol. 282, no. 5, E992–E1007, 2002.
- [77] Hovorka, R. et al., “Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes,” *Physiological measurement*, vol. 25, no. 4, p. 905, 2004.
- [78] Hovorka, R., “Closed-loop insulin delivery: From bench to clinical practice,” *Nature Reviews Endocrinology*, vol. 7, no. 7, pp. 385–395, 2011.
- [79] Topp, B., Promislow, K., Devries, G., Miura, R. M., and T Finegood, D., “A model of β -cell mass, insulin, and glucose kinetics: Pathways to diabetes,” *J. of Theoretical Biology*, vol. 206, no. 4, pp. 605–619, 2000.
- [80] De Gaetano, A. et al., “Mathematical models of diabetes progression,” *American Journal of Physiology-Endocrinology and Metabolism*, vol. 295, no. 6, E1462–E1479, 2008.
- [81] Ha, J., Satin, L. S., and Sherman, A. S., “A mathematical model of the pathogenesis, prevention, and reversal of type 2 diabetes,” *Endocrinology*, vol. 157, no. 2, pp. 624–635, 2016.
- [82] Cobelli, C., Renard, E., and Kovatchev, B., “Artificial pancreas: Past, present, future,” *Diabetes*, vol. 60, no. 11, pp. 2672–2682, 2011.
- [83] Huyett, L. M., Dassau, E., Zisser, H. C., and Doyle III, F. J., “Design and evaluation of a robust PID controller for a fully implantable artificial pancreas,” *Industrial & engineering chemistry research*, vol. 54, no. 42, pp. 10 311–10 321, 2015.
- [84] Laxminarayan, S., Reifman, J., and Steil, G. M., “Use of a food and drug administration-approved type 1 diabetes mellitus simulator to evaluate and optimize a proportional-integral-derivative controller,” *Journal of diabetes science and technology*, vol. 6, no. 6, pp. 1401–1412, 2012.
- [85] Steil, G. M., “Algorithms for a closed-loop artificial pancreas: The case for proportional-integral-derivative control,” *Journal of diabetes science and technology*, vol. 7, no. 6, pp. 1621–1631, 2013.
- [86] Kovács, L., Szalay, P., Almássy, Z., and Barkai, L., “Applicability results of a nonlinear model-based robust blood glucose control algorithm,” *Journal of diabetes science and technology*, vol. 7, no. 3, pp. 708–716, 2013.
- [87] Turksoy, K., Quinn, L., Littlejohn, E., and Cinar, A., “Multivariable adaptive identification and control for artificial pancreas systems,” *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 3, pp. 883–891, 2013.
- [88] Palumbo, P., Pepe, P., Panunzi, S., and De Gaetano, A., “Time-delay model-based control of the glucose–insulin system, by means of a state observer,” *European Journal of Control*, vol. 18, no. 6, pp. 591–606, 2012.
- [89] Borri, A., Palumbo, P., Manes, C., Panunzi, S., De Gaetano, A., et al., “Sampled-data observer-based glucose control for the artificial pancreas,” *Acta Polytechnica Hungarica*, vol. 14, no. 1, pp. 79–94, 2017.
- [90] Di Ferdinando, M., Pepe, P., Di Gennaro, S., Borri, A., and Palumbo, P., “Quantized sampled-data static output feedback control of the glucose–insulin system,” *Control Engineering Practice*, vol. 112, p. 104 828, 2021.
- [91] Incremona, G. P., Messori, M., Toffanin, C., Cobelli, C., and Magni, L., “Model predictive control with integral action for artificial pancreas,” *Control engineering practice*, vol. 77, pp. 86–94, 2018.

- [92] Messori, M., Incremona, G. P., Cobelli, C., and Magni, L., “Individualized model predictive control for the artificial pancreas: In silico evaluation of closed-loop glucose control,” *IEEE Control Systems Magazine*, vol. 38, no. 1, pp. 86–104, 2018.
- [93] Thabit, H. et al., “Closed-loop insulin delivery in inpatients with type 2 diabetes: A randomised, parallel-group trial,” *The lancet Diabetes & endocrinology*, vol. 5, no. 2, pp. 117–124, 2017.
- [94] Bally, L. et al., “Closed-loop insulin delivery for glycemic control in noncritical care,” *New England Journal of Medicine*, vol. 379, no. 6, pp. 547–556, 2018.
- [95] Daly, A. B. et al., “Fully automated closed-loop insulin delivery in adults with type 2 diabetes: An open-label, single-center, randomized crossover trial,” *Nature medicine*, vol. 29, no. 1, pp. 203–208, 2023.

Part 1:
Mathematical Modeling and
Control of T2D Progression via
Physical Activity

Chapter 2

Modeling the Long-Term Effect of Physical Activity on Blood Glucose Regulation

Contents

2.1	Introduction	19
2.2	Model Formulation	20
2.3	Macro Corpus of Simulations	26
2.4	Results	31
2.5	Discussion	34
2.6	Conclusion	36

2.1 Introduction

Despite¹ the numerous clinical evidence, the biological mechanisms underlying the exercise effects remain underexplored from a mathematical perspective, as there are still no mathematical models able to suitably account for the physiological machinery that is primarily responsible of the long-term benefits associated with exercise. Defining a proper model of long-term exercise effect on T2D is certainly a complex problem, in light of the several involved biological processes that reflect into variables and parameters to be accounted for by the modeling approach. The literature proposes models that attempt to describe how physical activity influences blood glucose regulation in the short-term [3], [4], [5], [6], [7], [8], i.e. *during exercise*, with variables evolving to their steady-state regime when the transient related to exercise is over, approximately in a time span of hours. Consequently, these approaches are not suitable to explain how the acute effect of single exercise bouts translates into longer-term adaptations and sustained benefits of regular physical activity on diabetes progression. In other words, while these models effectively describe the instantaneous metabolic responses induced by exercise, they do not consider an explanation of the mechanisms bridging together the effect of the single training session with the prolonged metabolic adaptations that unfold over years in response to regular exercise, primarily involving the increase of insulin sensitivity and of beta-cell mass functioning.

These long-term benefits are associated with the action of myokines, a class of signaling proteins released by skeletal muscle during exercise. Among them, *Interleukin-6* (IL-6) plays a pivotal role as one of the primary mediators linking physical activity to improved metabolic regulation. Through complex cross-talk with the immune and endocrine systems, IL-6 contributes to increased insulin sensitivity, anti-inflammatory effects, and beta-cell preservation [9], [10], [11]. These processes together underpin the long-term adaptations associated with regular exercise, including improved glycemic control and delayed progression of T2D. Capturing these mechanisms within a mathematical framework is therefore essential to bridge the gap between the short-term metabolic responses to exercise and the physiological adaptations driving its protective role against diabetes progression. To overcome these critical limitations, the novel modeling approach proposed in this chapter exhibits three fundamental features [1], [2]:

¹The contents in this chapter are based on the results in the works by the author published in [1] and [2].

- it proposes a two-timescale model, by means of the integration of a well-established minimal model of short-term physical activity effects with a T2D progression model [8], [12];
- it introduces the fast dynamics IL-6, a specific protein released under exercise by skeletal muscle, whose anti-inflammatory action is primarily involved in driving the sustained benefits of exercise [9], [10], [11];
- it formalizes the long-term effect exerted by means of IL-6 action [9], [10], [11] on insulin sensitivity and beta cell mass.

The most important contributions of this chapter are two-fold [1], [2]:

- 1) notably, we show that a modeling approach for the exercise effect neglecting IL-6 effect would not be sufficient to track the long-term benefits of exercise;
- 2) we show that model predictions are well aligned with the evidence from real-world case studies, by accurately investigating the capability of the model of generating predictions that reasonably replicate clinical evidence in a broad spectrum of simulated conditions:
 - different courses for T2D evolution, obtained by simulating different time constant governing insulin sensitivity decay;
 - different exercise programs, obtained by simulating varying intensity and weekly duration;
 - simulating the WHO general exercise recommendations for T2D prevention [13];
 - simulating the effect of physical activity on T2D evolution with discontinuation of the exercise intervention as reported in some clinical diabetes prevention trials [14], [15];
 - simulating cohorts of virtual subjects undergoing disease progression, obtained by independently distributed model parameter perturbations, to assess the capability of the model of consistently replicating the benefits of physical activity in an individualized setting, with the aim of extending the analysis to a more comprehensive scenario with respect to the “average” patient case [14], [15], [16], [17], [18].

We remark that the dynamics of low-grade systemic inflammation are highly complex and involve a multitude of agents and underlying biological machinery [9], [10], [11]. While IL-6 constitutes a crucial component of these processes, it is not the only intermediate player. In this regard, within our modeling framework, IL-6 can be interpreted as a representative proxy that captures the cumulative effects of exercise-induced inflammatory and metabolic adaptations in a form sufficiently tractable to support further control-oriented analysis and design. Indeed, the model paves the way to the development of model-based techniques for T2D progression control that could set the basis for the design of quantitatively assessed protocols for T2D control via physical activity management.

2.2 Model Formulation

As abovementioned, to account for both the short- and long-term effect of exercise, we propose a two-timescale model, defined by integrating the glucose-insulin slow dynamics, suitably modified from the model by Ha *et al.* [12], and the fast dynamics of physical activity described by the model by Roy *et al.* [8]. For what concerns the model by Ha *et al.*, it is described by the following system of differential equations [12]:

$$\frac{dG}{dt} = R_0 - (Eg_0 + S_I \cdot I), \quad (2.1a)$$

$$\frac{dI}{dt} = \beta \frac{ISR}{V} - kI, \quad (2.1b)$$

$$\frac{d\beta}{dt} = \frac{(P_{ISR} - A_{M_G})\beta}{\tau_\beta}, \quad (2.1c)$$

$$\frac{d\gamma}{dt} = -\frac{(\gamma - \gamma_{\infty, G})}{\tau_\gamma}, \quad (2.1d)$$

$$\frac{d\sigma}{dt} = -\frac{(\sigma - \sigma_{ISR, M_\infty})}{\tau_\sigma}, \quad (2.1e)$$

$$\frac{dS_I}{dt} = -\frac{(S_I - S_{I, \text{target}})}{\tau_{SI}}, \quad (2.1f)$$

where

- G (mg/dl) is the basal glucose concentration;
- I (μ U/ml) is the basal insulin concentration;
- β (mg) is the beta cell mass; $P_{ISR}(\cdot)$ and $A_{M_G}(\cdot)$ represent the β -cell replication and apoptosis rates, respectively, modeled as sigmoidal functions of the Insulin Secretion Rate (ISR) and the beta-cell metabolism (M);
- γ (unitless) denotes the shifts in glucose dose response;
- σ (μ U/ μ g/d) is the insulin secretion capacity per unit of beta cell mass;
- S_I (ml/ μ U/d) is the insulin sensitivity.

For what concerns the parameters, we have

- R_0 (mg/dl/d) represents basal hepatic glucose production, zero-order glucose uptake by tissues, and average daily effects of meals [12], [19];
- Eg_0 (1/d) is the glucose effectiveness at zero insulin;
- V (L) is the insulin distribution volume;
- k (1/d) is the insulin clearance rate;
- τ_γ (d) is gamma time constant;
- τ_σ (d) is sigma time constant;
- τ_β (d) is beta-cell time constant;
- τ_{SI} (d) is insulin sensitivity time constant.

For what concerns the fast dynamics, we inherit from the model by Roy *et al.* [8] those variables that specifically account for the short-term effect of exercise, that is

$$\frac{dPVO_2}{dt} = -0.8PVO_2 + 0.8u, \quad (2.2a)$$

$$\frac{dG_{\text{prod}}}{dt} = a_1PVO_2 - a_2G_{\text{prod}}, \quad (2.2b)$$

$$\frac{dG_{\text{up}}}{dt} = a_3PVO_2 - a_4G_{\text{up}} \quad (2.2c)$$

$$\frac{dI_e}{dt} = a_5PVO_2 - a_6I_e, \quad (2.2d)$$

involving the following state variables:

- PVO_2 (unitless) denotes the instantaneous oxygen consumption required by the exercise session, i.e. *during exercise*, expressed as a fraction of the maximum oxygen consumption;
- G_{prod} (mg/kg/min) represents the increased rate of endogenous glucose production by glucogenesis to compensate the increased glucose demand under exercise.

- G_{up} (mg/kg/min) denotes the incremental rate of glucose absorption by muscle tissues promoted by exercise;
- I_e ($\mu\text{U/ml/min}$) represents the increased rate of insulin removal under exercise.

Our approach builds upon these existing models by coupling the slow dynamics described in (2.1) with the fast dynamics of (2.2), thereby enabling the simulation of training sessions over a span of years. Moreover, to capture the sustained benefits of exercise over years, we introduce additional dynamics that capture the influence of exercise-induced IL-6 on insulin sensitivity and beta-cell dynamics. Specifically, we introduce

$$\frac{dIL_6}{dt} = SR_{\text{ex}}PVO_2 - k_m IL_6, \quad (2.3a)$$

$$\frac{dV_l}{dt} = IL_6 - k_s V_l, \quad (2.3b)$$

where

- IL_6 (pg/ml) is the concentration of IL-6 in muscle tissues induced by IL-6 release during the undertaken exercise session; this dynamics and related parameters are inherited from [11];
- V_l ((pg/ml)·min) represents the integral action of IL-6 and is used to model the prolonged, cumulative beneficial effect of exercise promoted by IL-6 release in terms of insulin sensitivity and beta-cell mass.

By introducing the state variables in (2.3), in our modeling framework it is possible to restructure equation (2.1c) and (2.1f) of the original model by Ha *et al.* [12] to capture the long-term effect of exercise on insulin sensitivity and beta-cell mass, as described in detail in the following section.

2.2.1 Modeling Core

Building upon the previous equations, the proposed model is described by the following system of differential equations [1], [2]:

$$\frac{dG}{dt} = R_0 - (Eg_0 + S_I \cdot I)G + \frac{W}{Vol_G}(G_{\text{prod}} - G_{\text{up}}), \quad (2.4a)$$

$$\frac{dI}{dt} = \beta \frac{ISR}{V} - kI - I_e \text{Heav}(I), \quad (2.4b)$$

$$\frac{d\beta}{dt} = \frac{(\bar{P}_{ISR} - \bar{A}_{MG})\beta}{\tau_\beta}, \quad (2.4c)$$

$$\frac{d\gamma}{dt} = -\frac{(\gamma - \gamma_{\infty,G})}{\tau_\gamma}, \quad (2.4d)$$

$$\frac{d\sigma}{dt} = -\frac{(\sigma - \sigma_{ISR_{M\infty}})}{\tau_\sigma}, \quad (2.4e)$$

$$\frac{dS_I}{dt} = -\frac{(S_I - S_{I,\text{target}}) \left(1 - \zeta_3 \frac{V_l^n}{k_{n,S_I}^n + V_l^n} \right)}{\tau_{S_I}}, \quad (2.4f)$$

$$\frac{dV_l}{dt} = IL_6 - k_s V_l, \quad (2.4g)$$

$$\frac{dPVO_2}{dt} = -0.8PVO_2 + 0.8u, \quad (2.4h)$$

$$\frac{dG_{\text{prod}}}{dt} = a_1 PVO_2 - a_2 G_{\text{prod}}, \quad (2.4i)$$

$$\frac{dG_{\text{up}}}{dt} = a_3 PVO_2 - a_4 G_{\text{up}} \quad (2.4j)$$

$$\frac{dI_e}{dt} = a_5 PVO_2 - a_6 I_e, \quad (2.4k)$$

$$\frac{dIL_6}{dt} = SR_{\text{ex}} PVO_2 - k_m IL_6, \quad (2.4l)$$

Physical activity enters the system by means of the input u , whose role within the model framework is crucial. It represents the target point for the sovrabasal oxygen consumption dynamics and is defined as a continuous variable indicating a percentage value ranging from 0 to 92% [1], [8]. In this regard, u is the intensity of the physical activity session. As a consequence, by varying u , it is possible to simulate different exercise levels, spanning from low-intensity activities to high-intensity training. For example, $u = 30\%$ corresponds to mild exercise such as walking, whereas $u \geq 50\%$ denotes more strenuous activities like running [1], [8]. When $u = 0$, no sovrabasal oxygen consumption is required to the individual. More precisely, when $u = 0$, basal oxygen consumption for routine activities is experienced by the individual but it does not contribute to the fast dynamics as its contribution is negligible. When $u > 0$, that means, during exercise, the fast variables associated with exercise (i.e., G_{prod} , G_{up} , I_e , IL_6) perturb the slow dynamics, whereas when the exercise session is over ($u = 0$), those variables, after a transient, go to zero. This allows to couple a fast dynamics of minutes associated with single bouts of exercise session with the overall slow progression. With u switching from 0 to a given continuous value (and vice versa) for each of the exercise sessions, the dynamics of the involved variables is governed by the continuous transient of PVO_2 , which accounts for all the intermediate phases of adaptation of the individual to the physical activity [1].

The exercise intensity u triggers the continuous dynamics of the sovrabasal oxygen consumption PVO_2 , that in turns promotes the dynamics of the incremental glucose production G_{prod} and uptake G_{up} and insulin removal I_e . Acting on equations (2.4a)-(2.4b), these variables represent how glucose regulation is affected *during* the undertaken exercise session.

Moreover, as it can be observed from equation (2.4l), the sovrabasal oxygen consumption PVO_2 , during exercise, stimulates IL-6 release by skeletal muscle, and this process embeds the cornerstone of the model: after release, IL-6 exerts its cumulative, prolonged beneficial effect by means of an integral action represented by the state variable V_l . V_l affects the slow variables associated with the progression of the disease, acting on beta cell mass (2.4c) and insulin sensitivity (2.4f). More in detail, for what concerns beta cell mass (2.4c), V_l acts by increasing beta cell proliferation P_{ISR} and reducing apoptosis A_{MG} [2], where:

$$\bar{P}_{ISR} = P_{ISR} \cdot \left(1 + \zeta_1 \frac{V_l^2}{k_{n,P}^2 + V_l^2} \right), \quad (2.5a)$$

$$\bar{A}_{MG} = A_{MG} \cdot \left(1 - \zeta_2 \frac{V_l^2}{k_{n,A}^2 + V_l^2} \right). \quad (2.5b)$$

That is, we restructure equation (2.1c) of the original model in [12], where

$$P_{ISR} = P_{\text{max}} \frac{ISR^{k_P}}{ISR^{k_P} + \alpha_P^{k_P}} \quad (2.6a)$$

$$A_{MG} = \left(A_{\text{max}} \frac{M_G^{k_A}}{M_G^{k_A} + \alpha_A^{k_A}} + A_b \right) \quad (2.6b)$$

$$(2.6c)$$

and model a new beta-cell dynamics introducing the multiplicative contributions in (2.5a)-(2.5b) to account for the effects of V_l (i.e., to capture the long-term effect of exercise on beta cells).

The effect of V_l on insulin sensitivity is modeled through the term $\left(1 - \zeta_3 \frac{V_l^n}{k_{n,S_I}^n + V_l^n} \right)$ in

(2.4f) [2], modifying the original insulin sensitivity dynamics in (2.1f) of [12]. The effect of the introduced term is to delay disease progression as it slows down the rate of the insulin sensitivity decay. In this way, V_I records the “memory effect” that physical activity is able to exert in the long-term by means of the cumulative anti-inflammatory action produced by IL-6. In this regard, the Michaelis-Menten function in (2.4f) is designed to model a larger effect of exercise in the first year, accordingly to the suggestions in [20], [21]. As a result, higher values of V_I due to increased physical activity lead to an improvement of insulin sensitivity S_I and an increase of beta-cell mass β due to increased proliferation and reduced apoptosis and this, in turn, promotes higher level of basal insulin concentration I and lower level of basal glucose concentration G , thus tracking the effect of exercise on blood glucose regulation over the years [1], [2].

We point out that, when $u = 0$, i.e. no physical activity is performed, the fast dynamics (2.4h)-(2.4l) vanishes. As a consequence, $V_I = 0$ and the model represents a classical diabetes progression model according to [12].

The model endows a set of algebraic state-depending functions inherited from [12], that are listed in Appendix A together with the full list of the model parameters and units (Table A.3).

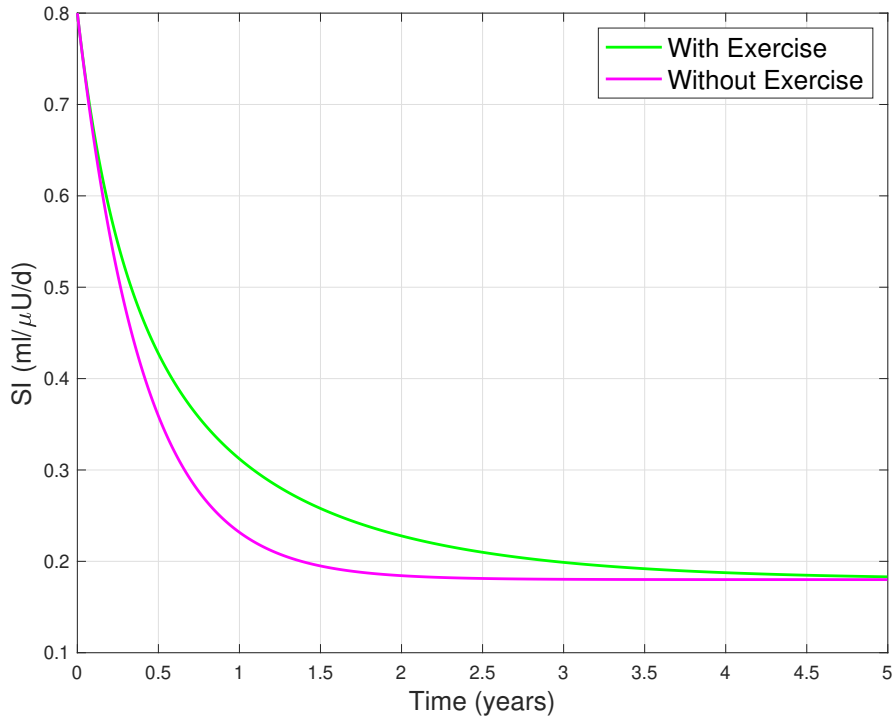


Figure 2.1: Time trend of insulin sensitivity over a five-year horizon when exercise intervention is considered (green line) and without exercise (magenta line). Exercise program: 60 minutes/sessions, 3 sessions/week at moderate exercise intensity $u = 50\%$.

2.2.2 Modeling Assumption on the Input

As abovementioned, physical activity is represented in the model by the input u . In a more comprehensive view, in our modeling framework u is a *Piecewise-constant* (PWC) input defined as:

$$u(t) = \begin{cases} 0 & t \in [kT, (k+1)T - D_k) \\ \bar{u}_k & t \in [(k+1)T - D_k, (k+1)T) \end{cases} \quad k \in \mathbb{N}, \quad (2.7)$$

where D_k and $\bar{u}_k$ defines the duration and intensity levels, respectively, of the k -th exercise session, assuming that sessions repeat with a given period T . Hence, the actual physical activity program is represented in the model by the triple $(\bar{u}_k, D_k, T)$.

2.2.3 Parameter Setting

Among the model parameters, those related to the slow dynamics in equations (2.4a)-(2.4f) are inherited from the model in [12], whereas parameters associated with the fast dynamics in equations (2.4h)-(2.4k) and (2.4l) are inherited from the works in [8] and [11], respectively. Conversely, parameters related to the long-term effect of exercise have been calibrated by simulations with the objective of capturing specific trends for the exercise benefits assessed by real-world studies. Specifically [1], [2]:

- the gains ζ_1 and ζ_2 and the activation coefficients $k_{n,P}$ and $k_{n,A}$ of the Hill functions in equations (2.5a)-(2.5b) are calibrated such that the basal glucose concentration G sets around 126 (mg/dl) on a five-year horizon when considering an exercise program consisting of 60 minutes/sessions, 3 sessions/week at moderate exercise intensity $u = 50\%$, consistently to what described in [16], [22].
- the gain ζ_3 and the activator coefficient k_{n,S_I} of the Michaelis-Menten in (2.4f) are set to reproduce the average improvement at first and fourth year in S_I promoted by exercise as described in [20], [23]. In this regard, an exemplary plot is depicted in Figure 2.1, where the increase in insulin sensitivity at the first and fourth year (around 30% and 8% respectively), parallels the evidence in [23].
- parameter k_s in equation (2.4g) is calibrated to consider the progressive loss of benefit typically observed in case of interruption of physical activity and accounts for 20% decline in the first 8 weeks of detraining, as shown in [24].

2.2.4 Model Behavior

With the aim of conceptually proving the importance of the role of IL-6 in mediating the benefits of exercise over years in our simulation framework and of showing how a modeling approach neglecting the “memory” effect of IL-6 is not capable of capturing the benefits of physical activity in the long term, we compare three simulation cases [1], as detailed below.

- Case 1. This case considers the simulation of the natural evolution of the disease with no exercise intervention; in this case the model reduces to the model in (2.1) by *Ha et al.* [12].
- Case 2. This case involves the simulation of regular exercise sessions, when only the short-term effect of exercise on blood glucose regulation is considered. More precisely, this means neglecting the dynamics in IL-6 (2.3) (and related long-term effects) and considering only the slow dynamics in (2.1), coupled with the fast dynamics in (2.2). The rationale behind this choice is the following: in this way it is possible to simulate repeated training sessions over the progression of the disease and to observe whether this is sufficient to capture the benefits of exercise over years.
- Case 3. This case involves the simulation of the full proposed model in (2.4).

Simulations were performed considering vigorous-intensity exercise ($u = 70\%$), two exercise sessions per week, each lasting 120 minutes.

Figure 2.2 illustrates the simulated trajectories of glucose concentration obtained from the three simulated cases. As observable, Case 1, which does not involve the effects of physical activity, captures only the progression of basal glucose concentration over the five-year simulation period. In contrast, the two cases that account for physical activity display oscillatory behavior due to the fast-slow dynamics induced by repeated exercise

sessions. Each session generates a transient drop in glucose levels, followed by a gradual return to a quasi-stationary state. This pattern is evident in bottom panel of Figure 2.2 [1].

Figure 2.3 (top panel) highlights the basal glycemic trend, obtained sampling glucose concentrations immediately prior to each exercise session [1]. Differently from Figure 2.2, by sampling the state before each exercise session, Figure 2.3 (top panel) is aimed at capturing the slow basal glucose behavior, illustrating how the quasi-stationary glucose concentration evolves over the disease progression. In Case 2 the basal glucose trajectory closely follows that of Case 1 alone, except for the brief transients caused by physical activity. In both cases, the diabetic threshold of 126 mg/dl is exceeded before the end of the third simulation year. This suggests that short-term dynamics alone are insufficient to capture the long-term metabolic benefits of physical activity, even when repeated sessions across a long simulation horizon are considered. Conversely, the proposed model (i.e., Case 3), which includes the IL-6-mediated regulation of beta-cell function and insulin sensitivity, shows basal glucose levels within the normoglycemic range for the entire simulation. The diabetic threshold is only approached toward the fifth year, indicating a delayed progression of the disease. These results suggest that incorporating the effects of IL-6 enables the model to reasonably capture the sustained benefits of regular exercise, with long-term results analogous to the findings from [22]. The results for the basal insulin concentration, shown in the center panel of 2.3, are consistent with the glycemic patterns in Figure 2.3 (top panel) and can be further explained by examining the beta-cell mass dynamics shown in the bottom panel of 2.3. In the proposed model (i.e., Case 3), IL-6 enhances insulin sensitivity, beta-cell proliferation and reduces apoptosis, and this results in overall decreased glucose levels, increased beta-cell mass and sustained insulin concentrations levels [1], [2]. This, in turn, avoids beta-cell mass degradation produced by glucotoxicity. By contrast, in Case 1 and Case 2, the progressively higher glucose levels affect the long-term beta-cell adaptation and lead to progressive beta cell mass deterioration and impaired insulin secretion [1], as originally described in [19].

2.3 Macro Corpus of Simulations

This section is devoted to the discussion and description of the simulation scenarios that we have outlined to assess the validity of the model predictions in a wide range of operating conditions [2]. Specifically a macro corpus of simulations were performed whose rationale is to assess the validity of the model and showing how model predictions are aligned with real-world evidence in a wide range of operating conditions. The methodology followed is detailed below.

Simulations in the following have been performed assuming that conditions predisposing to diabetes arise at $t = 0$, with S_I undergoing the exponential decay according to equation (2.4f). The initial condition for the state variables is assumed to be $[G(0) \ I(0) \ \beta(0) \ \gamma(0) \ \sigma(0) \ S_I(0) \ V_I(0) \ PVO_2(0) \ G_{\text{prod}}(0) \ G_{\text{up}}(0) \ I_e(0) \ IL_6(0)] = [99.7604 \ 9.025 \ 1000.423 \ -0.00666 \ 536.67 \ 0.8 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$; variables related to exercise are initialized to zero to simulate individuals not undergoing exercise interventions before $t = 0$. Unless differently specified, parameters are set to their nominal values (reported in Table A.3) .

2.3.1 Simulation Scenarios

With the aim of assessing the validity of the model in reasonably replicating the beneficial effect of exercise in a wide range of conditions, involving different exercise programs, different risk conditions for the simulated individuals and/or different genetic predispositions to T2D onset, this scenario involves the following set of simulations [2]:

i) **Simulation of a regular exercise program at different intensities**

Well established literature has demonstrated a dose-response relationship between

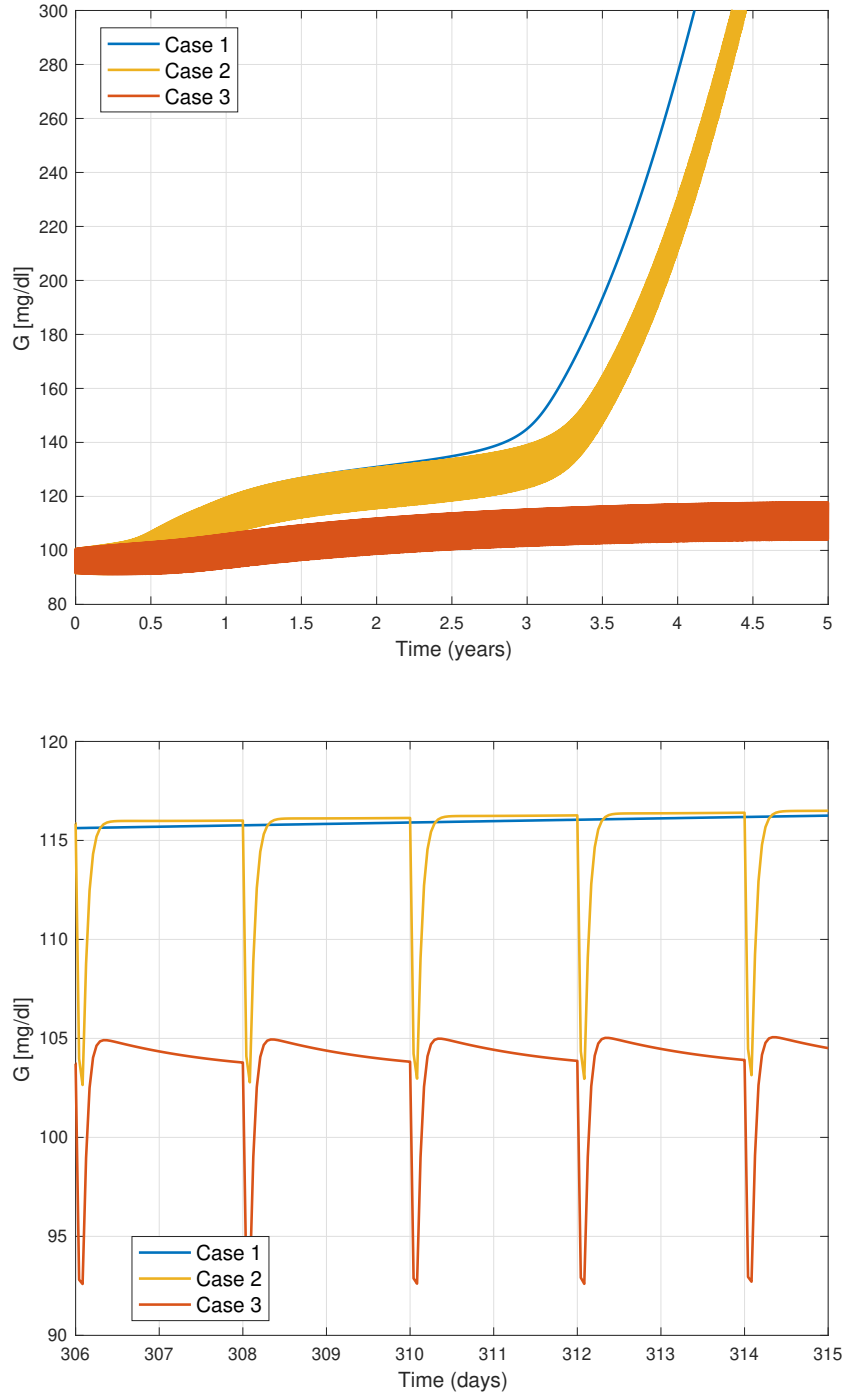


Figure 2.2: Top panel: plasma glucose concentration in the fast-slow dynamics as a function of time observed in Case 1 (blue), in Case 2 (yellow) and Case 3 (red). Bottom panel: zoom showing the effect of single training sessions in the same cases.

the intensity of physical activity and the reduction in T2D incidence [17], [23], [25]. To evaluate the model ability to reproduce the progressively greater benefits associated with higher exercise intensity, we simulated an exercise intervention consisting of three 60-minute sessions per week over a 5-year period, with the intensity u spanning from 0% to 70% to include the whole spectrum from light

to vigorous intensity. Moreover, the same training protocol was simulated over a 20-year horizon, focusing on moderate and vigorous intensities (u from 50% to 70%).

ii) **Simulation of a regular exercise program according to WHO guidelines**

The WHO suggests the following equivalent recommendations for chronic disease prevention: at least 150 min/week of moderate-intensity exercise or at least 75 min/week of vigorous-intensity exercise. Furthermore, prior evidence indicates that the benefits of moderate-intensity exercise tend to plateau once weekly activity reaches approximately 400 minutes [17]. To assess whether the proposed model is capable of reproducing the benefits associated with adherence to WHO recommendations, we simulated two training protocols in which the total weekly exercise duration was distributed across three sessions lasting 50 and 25 minutes each, respectively with the intensity u set to 50% (moderate intensity) and to 75% (vigorous intensity), as described in [8].

Additionally, we evaluated the model capability to capture the saturation effect observed at around 400 min/week of exercise, by performing simulations with weekly exercise durations of 300, 350, 400, 450, and 500 minutes with u set to 50%.

This batch of simulations involved a 5-year horizon.

iii) **Simulation of exercise discontinuation**

Real world evidence indicates that the positive effects of physical activity are sustained even after the exercise intervention being discontinued. For instance, similar evidence has been highlighted by the FDPS [14] and the CDQDPS [15], [26]. These studies enrolled adults with impaired glucose tolerance in a four-year and six-year lifestyle intervention respectively, with a subsequent patient follow-up carried out up to three years and fourteen years after the end of the intervention. The findings demonstrated that both T2D incidence and fasting glucose concentrations remained relatively stable following discontinuation of the intervention up to the end of the follow-up period, suggesting the persistence of beneficial effects over time. To assess the capability of the model in capturing this effect, we simulated the model under similar conditions to those indicated in FDPS and CDQDPS studies. More in detail:

- FDPS: simulations over a seven-year horizon were performed with daily sessions, each lasting 30 minutes, with u set at 30%, 40%, 50%, with exercise intervention stopped at the end of the fourth year [2], [14].
- CDQDPS: simulations over twenty-year horizon were performed with exercise intervention interrupting at the sixth year. The simulated exercise program consists of two sessions per day each lasting 20 minutes at moderate intensity ($u = 30\%, 40\%, 50\%$), analogously to what described in [2], [26].

iv) **Simulation of inter-individual variability**

To assess the capability of the model to reflect the heterogeneity across individuals in terms of T2D progression in response to physical activity, a series of simulations was performed to assess the sensitivity of the model output to perturbations in parameters and physical activity patterns, in three complementary subsets of simulations, as detailed in the followings. Unless otherwise specified, the simulated exercise program consists of regular sessions performed at intensity $u = 60\%$ with three sessions per week each lasting 60 minutes. All the described simulations were run over a 20-year simulation horizon.

- *Genetic and environmental predisposition to diabetes (GEP).*

This simulation campaign addressed variability arising from genetic and environmental factors contributing to T2D onset. Independent uniform random perturbations were applied to a selected subset of model parameters (reported in Table 2.1), associated with such predispositions. Two virtual cohorts were generated:

- * an intervention group (PA group), consisting of 200 virtual individuals engaged in a structured physical activity program;
- * a control group (no PA group), consisting of 200 virtual individuals with no exercise intervention.

The objective was to assess the population-level effectiveness of physical activity and to investigate how baseline physiological differences, associated with inherent predispositions, influence disease progression and intervention outcomes.

- *Dietary intake and metabolic balance (DIMB).*

This set of simulations explored variability related with diet and metabolic balance. Inter-individual differences were introduced by applying uniform random perturbations (within $\pm 10\%$ of the nominal value) to the parameter R_0 , which in diabetes progression models encodes key metabolic processes such as basal hepatic glucose production, zero-order glucose uptake by tissues, and average daily effects of meals [12], [19], [27]. R_0 can be interpreted as a net dietary–metabolic load acting on the glucose subsystem: it aggregates endogenous glucose release and the long-term average contribution of nutritional intake, representing the baseline inflow term that must be counteracted by glucose effectiveness and insulin-mediated uptake to maintain metabolic equilibrium. While R_0 is assumed to be constant for a given individual, reflecting stable dietary habits and intrinsic hepatic function over the considered time scale, it naturally varies across individuals, capturing inter-subject differences in habitual dietary intake, liver glucose output, and overall metabolic balance. Two virtual populations were considered:

- * 200 individuals with a high-risk predisposition ($\tau_{SI} = 150$);
- * 200 individuals with a moderate-risk predisposition ($\tau_{SI} = 180$).

- *Individual response to exercise and adherence to intervention (IRAI).*

This set examined variability in the physiological effects of exercise and in adherence to the intervention program, through two complementary subsets of simulations.

- a) Physiological response variability: random uniform perturbations were applied to parameters governing the physiological response to exercise stimuli (reported in Table 2.2), to simulate inter-individual variability in responsiveness to exercise effect. We highlight that, while parameters used in simulation set GEP are associated with genetic and environmental factors governing the predisposition to T2D, the selected subset of parameters considered here are specifically associated with the inter-individual variability in how subjects respond to and benefit from physical activity. These parameters capture heterogeneity in exercise responsiveness and represent inter-individual, genetically mediated differences in the metabolic effectiveness of physical activity. Two populations were simulated:

- 200 individuals with high-risk predisposition ($\tau_{SI} = 150$);
- 200 individuals with moderate-risk predisposition ($\tau_{SI} = 180$).

These simulations focused specifically on individual differences in physiological adaptation to exercise, independent of adherence or program structure.

- b) Exercise program adherence and variability: in this batch of simulations, inter-individual differences were introduced by randomly varying the intensity, duration, and scheduling of exercise sessions (according to what reported in Table 2.3), and by incorporating session-skipping behavior. The following virtual cohorts were simulated:

- 200 individuals with high-risk predisposition ($\tau_{SI} = 150$), skipping probability $Pr = 10\%$;

- 200 individuals with moderate-risk predisposition ($\tau_{SI} = 180$), skipping probability $Pr = 10\%$;
- 200 individuals with high-risk predisposition ($\tau_{SI} = 150$), skipping probability $Pr = 15\%$;
- 200 individuals with moderate-risk predisposition ($\tau_{SI} = 180$), skipping probability $Pr = 15\%$;
- 200 individuals with moderate-risk predisposition ($\tau_{SI} = 180$), skipping probability $Pr = 35\%$.

To systematically compare these simulations with the reference exercise program, we introduce the notion of *Cumulative Exercise Dose* (CED), defined as:

$$CED = \sum_{i=1}^N u_i \cdot D_i,$$

where u_i and D_i denote the intensity and duration of the i -th session, and N is the total number of sessions across the simulation horizon. This set of simulations was designed to assess the impact of behavioral variability and real-world inconsistencies in program adherence on the long-term benefits of exercise interventions.

In the simulation scenarios discussed above, inter-individual variability is introduced through independent random parameter perturbations. Although physiological parameters in biological systems are often strongly correlated, identifying and accurately characterizing all such correlations is highly complex. As a result, modeling inter-individual variability through independent random parameter perturbations represents a common and well-established practice in the diabetes modeling and control literature [28], [29], [30].

Parameter	Range
W	[50, 150]
$S_{I,target}$	[0.17, 0.20]
Vol_G	[110, 130]
τ_{SI}	[130, 200]
G_{σ_s}	[75, 90]
E_{g0}	[1.3, 1.6]
V	[4.75, 5.25]
k	[500, 900]
τ_β	[7500, 9000]
τ_γ	[2, 3]
τ_σ	[235, 265]
α_M	[142, 158]
α_{ISR}	[1.14, 1.26]
P_{max}	[4.32, 4.77]
α_P	[39, 43]
A_{max}	[8.55, 10]
α_A	[0.42, 0.46]
A_b	[0.76, 0.85]

Table 2.1: Parameter ranges used to simulate inter-individual variability in genetic and environmental predisposition.

Parameter	Range
k_s	$[1.31 \times 10^{-6}, 4.42 \times 10^{-6}]$
SR_{ex}	$[0.04, 0.05]$
k_m	$[0.003, 0.005]$
a_1	$[0.0014, 0.0017]$
a_2	$[0.05, 0.062]$
a_3	$[0.0018, 0.0021]$
a_4	$[0.0437, 0.0534]$
a_5	$[0.0011, 0.0014]$
a_6	$[0.0675, 0.0825]$
ζ_1, ζ_2	$[8.0 \times 10^{-5}, 1.2 \times 10^{-4}]$
$k_{n,P}, k_{n,A}$	$[900000, 1100000]$
$k_{n,SI}$	$[4750000, 5250000]$
ζ_3	$[1.26, 1.54]$

Table 2.2: Parameter ranges used to simulate variability in individual response to exercise.

Parameter	Range
Exercise intensity	$[50\%, 70\%]$
Session duration (minutes)	$[30, 90]$
Period between sessions (days)	$\{1, 2, 3\}$

Table 2.3: Randomization ranges for simulating variability in exercise program.

2.4 Results

2.4.1 Simulations of a Regular Exercise Program at Different Intensities

Figure 2.4 and Table 2.4 illustrate the results obtained simulating the model at different exercise intensity over a five-year and twenty-year horizon respectively. Specifically, Figure 2.4 highlights the capability of the model of reproducing the so call *dose-response* relationship, with the benefits of the exercise increasing as the intensity u increases [2]. As observable, at higher intensity corresponds lower values for glucose concentration G and higher values for insulin concentration I , with the glucose trajectories not exhibiting the marked upward inflection as the intensity exceeds the threshold $u = 40\%$. Notably, at moderate to vigorous intensity ($u = 60\%$ and $u = 70\%$) glucose levels remain consistently below the diagnostic threshold for diabetes (126 mg/dl) throughout the five-year simulation. From Table 2.4 it can be observed that, with $\tau_{SI} = 150$ days and $u = 50\%$, the simulated subject goes diabetic from the tenth year, glucose exceeding 126 mg/dl and rapidly increasing afterward, thus the progression of the disease becoming irreversible. When higher intensities are considered ($u = 60\%$ and $u = 70\%$), physical activity is indeed capable of reversing disease progression and restore normal basal glucose concentration values around 100 mg/dl over the twenty-year simulation horizon. It should be noted that the same occurs even at $u = 50\%$, when assuming a less severe course for the disease (i.e., $\tau_{SI} = 180$), with the simulated subjects keeping basal glucose values within a the healthy range throughout the whole simulation period [2]. The trends for basal insulin concentration I show a progressive increase with u . Indeed, whenever exercise proves to be insufficient to reverse diabetes progression (e.g, $\tau_{SI} = 150$, $u = 50\%$), insulin decreases as a consequence of beta cell mass deterioration. Conversely, higher exercise level ($u = 60\%$, $u = 70\%$) contribute to sustain insulin sensitivity and insulin release, avoiding glucotoxicity and consequently beta cell mass degradation. As a result, glucose values remain below the diabetic threshold [2]. These results suggest that model predictions consistently show a reasonable capability of capturing the association between intensity of exercise and observed benefits and results are well aligned to the literature related to the *dose-response* relationship [13], [17], [25].

2.4.2 Simulations of a Regular Exercise Program According to WHO Guidelines

Figure 2.5 illustrates the simulated trajectories of basal glucose concentration G over a five-year period under exercise programs consistent with the WHO recommendations for cardiovascular and chronic disease prevention [13]. The results indicate that the two minimum recommended simulated programs, namely, 75 minutes per week of vigorous-intensity activity and 150 minutes per week of moderate-intensity activity, yield nearly overlapping glucose profiles throughout the simulation horizon; this evidence suggests that the model consistently captures the equivalence between the two exercise programs [2]. Notably, indeed, the marginal difference in glucose concentrations between these two simulated protocols at fifth year (approximately 4 mg/dl) is negligible and this difference can be further reduced with slight different values of u , for instance by setting slightly higher values for moderate-intensity exercise or slightly lower values for vigorous-intensity exercise. The figure also shows that increasing the weekly duration of moderate-intensity exercise up to 500 minutes results in progressively lower glucose concentrations, as reasonably expected. However, the benefit appears to approach saturation beyond 400 minutes per week [2], a pattern that is consistent with the aforementioned real world evidence about the saturation of the exercise benefit, as reported in [17].

2.4.3 Simulation of Exercise Discontinuation

Results associated with the simulation of the exercise intervention according to FDPS and CDQDPS are reported in Table 2.5 [2]. For what concerns the simulated FDPS, following the interruption of the intervention at the fourth year, glucose values keep below or near the diabetic threshold at the end of the simulation horizon in three of the six cases evaluated. It is worth noting that in these cases glucose levels seem not to irreversibly exceed the diabetic threshold throughout the seven-year horizon. Conversely, the observed trend is that the cases showing an irreversible disease course correspond to those exhibiting diabetic glucose values before the end of the intervention.

Analogous results can be observed in the simulated CDQDPS. Specifically, the discontinuation of the intervention is associated with an irreversible disease course whenever G is almost diabetic when the discontinuation occurs (e.g., for $u = 30\%$). Whereas, exercise effect is sustained in the remaining cases ($u = 40\%$, $\tau_{SI} = 180$ days; $u = 50\%$, $\tau_{SI} = 150$ and 180 days), where physical activity, despite the interruption, proves to be capable of restoring normoglycemia over the twenty-year horizon. In one case, corresponding to a simulated severe course for the disease ($\tau_{SI} = 150$ days) and mild exercise intensity ($u = 40\%$), the glucose trajectory seems to keep stable approaching the diabetic threshold, with the disease not becoming irreversible. Nevertheless, the simulated subject remains in a risk condition [2]. In general, these results show that the model proves to be able to quantitatively capture the prolonged benefits of physical activity even in case of interruption of the intervention and to reasonably replicate the evidence in [14], [15].

		$G_{10-year}$ (mg/dl)	$G_{15-year}$ (mg/dl)	$G_{20-year}$ (mg/dl)	$I_{10-year}$ (μ U/ml)	$I_{15-year}$ (μ U/ml)	$I_{20-year}$ (μ U/ml)
$u = 50\%$	$\tau_{SI} = 150$	128	>>150	>>150	29	< 3	< 3
	$\tau_{SI} = 180$	116	105	101	33	38	40
$u = 60\%$	$\tau_{SI} = 150$	117	108	101	33	36	39
	$\tau_{SI} = 180$	111	101	100	35	39	40
$u = 70\%$	$\tau_{SI} = 150$	112	102	100	35	39	40
	$\tau_{SI} = 180$	108	101	100	36	39	40

Table 2.4: Basal glucose and insulin concentration observed over a 20-year horizon for $u = 50\%$, 60% , and 70% and $\tau_{SI} = 150$ and 180 days.

FDPS		$G_{4\text{th-year}}$ (mg/dl)	$G_{7\text{th-year}}$ (mg/dl)	CDQDPS		$G_{6\text{th-year}}$ (mg/dl)	$G_{20\text{th-year}}$ (mg/dl)
$u = 30\%$	$\tau_{SI} = 150$	131	>> 150	$u = 30\%$	$\tau_{SI} = 150$	133	>> 150
	$\tau_{SI} = 180$	126	136		$\tau_{SI} = 180$	125	>> 150
$u = 40\%$	$\tau_{SI} = 150$	127	>> 150	$u = 40\%$	$\tau_{SI} = 150$	124	123
	$\tau_{SI} = 180$	123	126		$\tau_{SI} = 180$	119	101
$u = 50\%$	$\tau_{SI} = 150$	124	128	$u = 50\%$	$\tau_{SI} = 150$	119	101
	$\tau_{SI} = 180$	120	121		$\tau_{SI} = 180$	115	100

Table 2.5: Basal glucose levels observed in the simulated FDPS and CDQDPS programs for $u = 30\%$, 40% , and 50% and $\tau_{SI} = 150$ and 180 days.

2.4.4 Simulation of Individual Inter-variability

- *Genetic and environmental predisposition to diabetes.*

Table 2.6– Simulation Set GEP summarizes the percentage of simulated individuals in which glucose G remains below the diabetic threshold across the full simulation horizon (no T2D), for both the intervention (PA) and control (no PA) groups. Results show a clear benefit of physical activity: the no T2D outcome occurs in 66% of the PA group versus 35% of the no PA group. This indicates that physical activity exerts a strong protective effect even under simulated inter-individual variability. At the same time, the fact that a subset of individuals in the PA group still develops T2D, despite activity levels consistent with WHO prevention guidelines, highlights the role of individual physiological and genetic predispositions [2].

To further explore these effects, Figure 2.6 presents the distribution of six parameters most directly linked to long-term outcomes (τ_{SI} , $S_{I,\text{target}}$, τ_{β} , P_{max} , A_{max} , A_b), expressed as relative percentage changes from baseline. The distributions are shown for both outcomes (T2D and no T2D) within the PA and no PA groups. As an example, τ_{SI} values are generally higher in no T2D individuals compared to those developing T2D. Notably, the PA-T2D subgroup displays even lower τ_{SI} values than the no PA-T2D subgroup, suggesting that individuals who progress to T2D despite physical activity carry stronger adverse predispositions, such as faster baseline disease progression. Comparable patterns are observed for $S_{I,\text{target}}$, P_{max} , and A_{max} , which are related to lower insulin sensitivity, reduced beta-cell production, and increased beta-cell loss in the PA-T2D group. Together, these findings emphasize the capacity of physical activity to provide protection across a wide range of simulated risk profiles, even in presence of adverse a-priori risk predispositions [2].

- *Dietary intake and metabolic balance.*

Table 2.6– Simulation Set DIMB summarizes the fraction of no T2D individuals at the end of the simulation horizon for two baseline risk profiles, high ($\tau_{SI} = 150$) and moderate ($\tau_{SI} = 180$). The results demonstrate that variability in dietary intake and metabolic regulation can markedly affect long-term outcomes, even under identical exercise regimens. This impact is strongly modulated by baseline predisposition: in the high-risk cohort ($\tau_{SI} = 150$), approximately 45% of individuals still develops T2D despite regular exercise. This finding underscores that dietary fluctuations may constrain the benefits of physical activity in individuals with unfavorable disease progression. The distributions of the parameter R_0 across the T2D subjects, reported in Figure 2.7 as percentage deviations from nominal values, further illustrate this effect. In the high-risk cohort, the distribution shifts to lower values and exhibits greater variability, while in the moderate-risk cohort it is narrower and centered at a higher median (+8%). These results suggest a compensatory mechanism between exercise and diet: in moderate-risk individuals, T2D onset occurs primarily among those experiencing substantial dietary and metabolic challenges, whereas in high-risk individuals, even smaller dietary loads are sufficient to outweigh the protective effect of physical activity [2].

- *Individual response to exercise and adherence to intervention.*

Table 2.6 – Simulation Set IRAI-a presents the percentage of simulated individuals remaining non-diabetic in the high-risk ($\tau_{SI} = 150$) and moderate-risk ($\tau_{SI} = 180$) groups. The exercise intervention yields high success rates across both cohorts, with higher efficacy observed in the moderate-risk group, consistent with its more favorable baseline predisposition.

Further insights are provided in Figure 2.8, which shows a 3D scatter plot of deviations in three parameters ($k_s, k_{n,SI}, \zeta_3$) among T2D individuals in the high-risk cohort. The data largely cluster near a surface corresponding to approximately +40% variation in k_s , the parameter regulating the decay of exercise benefits. This trend indicates that faster decay rates markedly increase T2D risk, largely independent of $k_{n,SI}$ and ζ_3 . Notably, some individuals with lower k_s values still progress to T2D when associated with negative deviations in ζ_3 , showing that severely impaired improvements in insulin sensitivity cannot be compensated by exercise alone. These findings emphasize that inter-individual heterogeneity in the persistence and magnitude of exercise effects plays a crucial role in determining intervention outcomes [2].

Table 2.6–Simulation Set IRAI-b explores the impact of intra-individual variability in adherence to the training program. The results demonstrate that behavioral fluctuations strongly modulate the effectiveness of physical activity, especially with higher dropout probabilities. For instance, in the high-risk group, raising the skipping probability from 10% to 15% decreases the success rate from 96% to 41%. In contrast, the moderate-risk group remains robust until dropout levels become substantially higher, but even then, at 35% skipping probability, approximately half of individuals develops T2D [2].

Figure 2.9 supports these observations, showing the distributions of CED deviations relative to baseline for representative scenarios in simulated individuals who develop T2D. As expected, CED decreases linearly with increasing dropout probability. The distributions are relatively narrow, confirming a limited effect of the parameter fluctuations [2].

Simulation Set	Group	no T2D (%)
GEP	PA group	66%
	no PA group	35%
DIMB	$\tau_{SI} = 150$	55%
	$\tau_{SI} = 180$	80%
IRAI-a	$\tau_{SI} = 150$	70%
	$\tau_{SI} = 180$	96%
IRAI-b	$\tau_{SI} = 150, pr = 10\%$	96%
	$\tau_{SI} = 150, pr = 15\%$	41%
	$\tau_{SI} = 180, pr = 10\%$	100%
	$\tau_{SI} = 180, pr = 15\%$	100%
	$\tau_{SI} = 180, pr = 35\%$	50%

Table 2.6: Percentage of simulated individuals in which G remains below the diabetic threshold in the whole time horizon (no T2D) when parameters related to inter-individual variability are varied in the considered simulation sets.

2.5 Discussion

The proposed model (2.4) [1], [2] successfully captures the dynamics of glucose and insulin concentrations in individuals during T2D progression, characterized by a gradual exponential decline in insulin sensitivity, in the presence of regular physical exercise.

To the best of our knowledge, the model bridges a crucial gap in the literature of T2D modeling and control, as it is the first original formulation in the literature capable of accurately describing the cumulative effects of physical activity on both insulin sensitivity and beta-cell function, with these effects mediated by IL-6 signaling released during exercise. The present formulation incorporates physiologically realistic insulin sensitivity dynamics, consistent with clinical observations reported in [23], [31] (see Figure 2.1). Moreover, it captures the influence of different exercise intensities on fasting glucose levels, thus providing a comprehensive and mechanistic description of exercise-induced metabolic adaptations during T2D progression. The model predictions under a range of simulated conditions align well with findings reported in the literature and with evidence from clinical guidelines, as illustrated in Figures 2.4 to 2.5 and Tables 2.4 to 2.5. In particular, the model accurately reproduces the dose-response relationship between physical activity and diabetes risk reduction [17], [23], [32]. The simulations show that, holding all other parameters constant (weekly exercise duration, τ_{SI} , and initial conditions), the risk of developing T2D decreases as exercise intensity increases (see Figure 2.4, Tables 2.4–2.5). More in detail, the long-horizon simulations produced by the proposed model indicate a clear beneficial effect of physical activity intensity on glycaemic outcomes: increasing u leads to progressively lower glycaemic trajectories. Over a five-year simulation horizon, the terminal glucose level decreases monotonically as u is increased from 0% up to 70%, suggesting that, under comparable overall exercise duration, higher-intensity exercise produces a stronger protective metabolic effect in the model. This behavior is in line with the evidence in [17], where, for comparable total exercise duration, increasing intensity is associated with a progressive reduction in diabetes incidence in the studied cohorts. While the clinical endpoints in those studies are typically expressed in terms of incidence/risk rather than simulated glucose concentrations, the direction of the outcome is coherent: higher intensity, at similar duration, corresponds to improved long-term metabolic control. Analogously, the model reproduces the complementary dose-response trend with respect to total exercise duration: for fixed intensity, increasing the cumulative duration of physical activity produces progressively lower glucose levels. This second trend further aligns with the findings in [17], [23], [32].

Furthermore, the model predicts comparable benefits when performing 150 minutes per week of moderate-intensity exercise or 75 minutes per week of vigorous-intensity exercise (Figure 2.5), in agreement with physical activity guidelines issued by the WHO [13], the American Diabetes Association [33], and the American College of Sports Medicine [34]. Finally, the model also captures the saturation effect in the benefits of moderate-intensity exercise at approximately 400 minutes per week, as documented in recent studies [17] (see Figure 2.5).

The proposed model is also capable of reproducing sustained benefits following the discontinuation of the exercise intervention, as shown in Table 2.5. These results are consistent with findings from both FDPS and the CDQDPS. Notably, in the FDPS [31], the average fasting glucose concentration among participants who eventually developed T2D was reported to be 126 mg/dL at the fourth year. This value is closely approximated by the model, which predicts an average glucose concentration $G_{4\text{th-year}}$ of 127 mg/dl under comparable conditions (Table 2.5).

Similar consistency is observed for the CDQDPS training and de-training experiment. The simulation results in Table 2.5 reproduce the range of T2D incidence trajectories and fasting glucose levels reported by Li et al. [15]. In some cases, diabetes develops within six years or over the full 20-year horizon; in others, disease progression is substantially slowed or even reversed.

Interestingly, two clear trends emerge from the model outputs (see Table 2.4 and Table 2.5): (i) higher exercise intensity (u) combined with a slower decline in insulin sensitivity (S_I) is associated with lower predicted glucose levels at both the end of the intervention and at the end of the follow-up period; (ii) when the basal glucose level remains below the diabetic threshold at the end of the intervention, the model predicts that benefits are maintained throughout the follow-up period, up to 14 years post-discontinuation, provided that a moderate-intensity exercise regimen has been performed consistently for

six years.

In addition to the specific studies discussed above (e.g., [14], [15], [17]), numerous other trials have reported similar effects of physical activity on the prevention or delay of T2D (e.g., [16], [35], [36]). These comparisons suggest that the model provides a realistic representation of average glycemic trajectories in a generalized “virtual subject” undergoing lifestyle-based diabetes prevention. The model also demonstrates the protective effects of regular physical activity on diabetes progression when inter-individual variability is introduced through random perturbations of model parameters (see Table 2.6). Simulations incorporating person-to-person variability yield outcomes that align well with major clinical evidence, consistently indicating a 30–60% reduction in diabetes incidence in physically active individuals compared to control groups [14], [16], [26], [35], [37], [38]. Specifically, in simulations targeting variability in genetic and environmental predisposition, the model predictions are highly consistent with findings reported in [39], which observed up to a 74% reduction in diabetes incidence among individuals with elevated genetic risk who engaged in regular physical activity. In addition, the model captures the impact of behavioral variability, including differences in adherence and exercise program structure. Despite these variations, physical activity continues to provide preventive benefits, although the success rate varies across scenarios. This is in agreement with findings from clinical studies such as [16], [40], which report that even irregular or intermittent physical activity patterns can significantly reduce the risk of developing T2D. Despite the promising results, the model shows several limitations that should be addressed in future work. The first limitation is related to the lack of explicit modeling of meals and future developments should be aimed at integrating the effects of the caloric balance between intake and expenditure, as explored in other modeling studies [22], [41], representing a promising direction for model refinement. An additional limitation involves the lack of explicit modeling of demographic and comorbidity-specific contributions, of biomechanical constraints and pharmacological treatments. Another limitation lies in the assumption made in this study regarding the initial condition: a relatively rapid decline in insulin sensitivity was simulated, effectively modeling the onset of high-risk conditions for T2D at time $t = 0$. Future work should investigate the impact of physical activity in healthy individuals at low risk for T2D. This would involve exploring alternative formulations for the decline in insulin sensitivity and incorporating varying levels of baseline fitness to better capture the early preventive role of physical activity [19], [42]. Nevertheless, the most important limitation lies in the fact that the presented results are mainly simulation-based evidence and that the model lacks validation against clinical data. To allow a broader generalizability of the findings, future works will be aimed at overcoming the aforementioned limitations and at assessing the direct validation of the model predictions against data.

2.6 Conclusion

The role of physical activity in delaying or preventing the onset of T2D is well established. Regular exercise has been shown to significantly reduce the burden of this chronic condition, which is projected to pose a major global health challenge in the coming decades [43]. What makes exercise particularly crucial in the context of T2D prevention is its ability to preserve insulin sensitivity and beta-cell function, one of the key physiological factors in preventing disease onset and progression [9], [10]. The model presented in this chapter represents a meaningful step toward a mechanistic understanding of the protective effects of exercise programs on T2D progression, as it represents the first mathematical model suitable to capture the complex physiological machinery that is primarily involved in driving exercise benefits. The results of the model show promising and support the design of targeted recommendations for the effective monitoring and prevention of T2D by means of the design and development of model-based control strategies for personalized intervention planning. This model has already been employed in both theoretical [44] and applicative [45] frameworks. Future developments will enhance

the predictive power of the model and its capacity to support the creation of personalized lifestyle recommendations for T2D prevention. Ultimately, this line of research holds the potential to guide individuals in achieving sustainable health targets, while also offering substantial benefits for both patients and healthcare systems.

References

- [1] De Paola, P. F., Paglialonga, A., Palumbo, P., Keshavjee, K., Dabbene, F., and Borri, A., “The long-term effects of physical activity on blood glucose regulation: A model to unravel diabetes progression,” *IEEE Control Systems Letters*, vol. 7, pp. 2916–2921, 2023. DOI: [10.1109/LCSYS.2023.3290774](https://doi.org/10.1109/LCSYS.2023.3290774).
- [2] De Paola, P. F., Borri, A., Dabbene, F., Keshavjee, K., Palumbo, P., and Paglialonga, A., “Modeling the cumulative benefits of regular physical activity on type 2 diabetes progression,” *Computers in Biology and Medicine*, vol. 198, p. 111 194, 2025. DOI: <https://doi.org/10.1016/j.compbiomed.2025.111194>.
- [3] Palumbo, M. C., Morettini, M., Tieri, P., Diele, F., Sacchetti, M., and Castiglione, F., “Personalizing physical exercise in a computational model of fuel homeostasis,” *PLoS computational biology*, vol. 14, no. 4, e1006073, 2018.
- [4] Kartono, A., “Modified minimal model for effect of physical exercise on insulin sensitivity and glucose effectiveness in type 2 diabetes and healthy human,” *Theory in Biosciences*, vol. 132, pp. 195–206, 2013.
- [5] Al Ahdab, M., Leth, J., Knudsen, T., Vestergaard, P., and Clausen, H. G., “Glucose-insulin mathematical model for the combined effect of medications and life style of type 2 diabetic patients,” *Biochemical Engineering Journal*, vol. 176, p. 108 170, 2021.
- [6] Derouich, M. and Boutayeb, A., “The effect of physical exercise on the dynamics of glucose and insulin,” *Journal of biomechanics*, vol. 35, no. 7, pp. 911–917, 2002.
- [7] Singh, S. and Kumar, D., “A mathematical model on glucose-insulin regulatory system with the impact of physical activities,” *Computational and Mathematical Biology*, vol. 5, no. 1, pp. 5–8, 2016.
- [8] Roy, A. and Parker, R. S., “Dynamic modeling of exercise effects on plasma glucose and insulin levels,” *J. of Diabetes Science and Technology*, vol. 1, no. 3, pp. 338–347, 2007.
- [9] Petersen, A. M. W. and Pedersen, B. K., “The anti-inflammatory effect of exercise,” *Journal of applied physiology*, vol. 98, no. 4, pp. 1154–1162, 2005.
- [10] Curran, M. et al., “The benefits of physical exercise for the health of the pancreatic β -cell: A review of the evidence,” *Experimental Physiology*, vol. 105, no. 4, pp. 579–589, 2020.
- [11] Morettini, M., Palumbo, M. C., Sacchetti, M., Castiglione, F., and Mazza, C., “A system model of the effects of exercise on plasma interleukin-6 dynamics in healthy individuals: Role of skeletal muscle and adipose tissue,” *PloS one*, vol. 12, no. 7, e0181224, 2017.
- [12] Ha, J., Satin, L. S., and Sherman, A. S., “A mathematical model of the pathogenesis, prevention, and reversal of type 2 diabetes,” *Endocrinology*, vol. 157, no. 2, pp. 624–635, 2016.
- [13] Bull, F. C. et al., “World Health Organization 2020 guidelines on physical activity and sedentary behaviour,” *British journal of sports medicine*, vol. 54, no. 24, pp. 1451–1462, 2020.
- [14] Lindström, J. et al., “Sustained reduction in the incidence of type 2 diabetes by lifestyle intervention: Follow-up of the finnish diabetes prevention study,” *The Lancet*, vol. 368, no. 9548, pp. 1673–1679, 2006.

-
- [15] Li, G. et al., “The long-term effect of lifestyle interventions to prevent diabetes in the China Da Qing diabetes prevention study: A 20-year follow-up study,” *The Lancet*, vol. 371, no. 9626, pp. 1783–1789, 2008.
 - [16] Diabetes Prevention Program Research Group, “Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin,” *New England journal of medicine*, vol. 346, no. 6, pp. 393–403, 2002.
 - [17] Boonpor, J. et al., “Dose–response relationship between device-measured physical activity and incident type 2 diabetes: Findings from the UK biobank prospective cohort study,” *BMC medicine*, vol. 21, no. 1, p. 191, 2023.
 - [18] Strain, T. et al., “Quantifying the relationship between physical activity energy expenditure and incident type 2 diabetes: A prospective cohort study of device-measured activity in 90,096 adults,” *Diabetes Care*, vol. 46, no. 6, pp. 1145–1155, 2023.
 - [19] Topp, B., Promislow, K., Devries, G., Miura, R. M., and T Finegood, D., “A model of β -cell mass, insulin, and glucose kinetics: Pathways to diabetes,” *J. of Theoretical Biology*, vol. 206, no. 4, pp. 605–619, 2000.
 - [20] Dâmaso, A. R. et al., “Aerobic plus resistance training was more effective in improving the visceral adiposity, metabolic profile and inflammatory markers than aerobic training in obese adolescents,” *Journal of sports sciences*, vol. 32, no. 15, pp. 1435–1445, 2014.
 - [21] Bird, S. R. and Hawley, J. A., “Update on the effects of physical activity on insulin sensitivity in humans,” *BMJ open sport & exercise medicine*, vol. 2, no. 1, e000143, 2017.
 - [22] De Gaetano, A. and Hardy, T. A., “A novel fast-slow model of diabetes progression: Insights into mechanisms of response to the interventions in the diabetes prevention program,” *PLoS One*, vol. 14, no. 10, e0222833, 2019.
 - [23] Balducci, S. et al., “Sustained decreases in sedentary time and increases in physical activity are associated with preservation of estimated β -cell function in individuals with type 2 diabetes,” *Diabetes Research and Clinical Practice*, vol. 193, p. 110 140, 2022.
 - [24] Mujika, I. and Padilla, S., “Detraining: Loss of training-induced physiological and performance adaptations. part I: Short term insufficient training stimulus,” *Sports medicine*, vol. 30, pp. 79–87, 2000.
 - [25] Steil, G. M., “Algorithms for a closed-loop artificial pancreas: The case for proportional-integral-derivative control,” *Journal of diabetes science and technology*, vol. 7, no. 6, pp. 1621–1631, 2013.
 - [26] Pan, X.-R. et al., “Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance: The Da Qing IGT and diabetes study,” *Diabetes care*, vol. 20, no. 4, pp. 537–544, 1997.
 - [27] De Paola, P. F., Borri, A., Paglialonga, A., Dabbene, F., and Palumbo, P., “Long-term diabetes prevention via physical activity: An output-feedback MPC approach,” *IEEE Control Systems Letters*, 2025. DOI: [10.1109/LCSYS.2025.3578031](https://doi.org/10.1109/LCSYS.2025.3578031).
 - [28] Di Ferdinando, M., Pepe, P., Di Gennaro, S., Borri, A., and Palumbo, P., “Quantized sampled-data static output feedback control of the glucose–insulin system,” *Control Engineering Practice*, vol. 112, p. 104 828, 2021.
 - [29] Borri, A., Pola, G., Pepe, P., Di Benedetto, M. D., and Palumbo, P., “Symbolic control design of an artificial pancreas for type-2 diabetes,” *IEEE Transactions on Control Systems Technology*, vol. 30, no. 5, pp. 2131–2146, 2021.
 - [30] Borri, A., Cacace, F., De Gaetano, A., Pompa, M., and Panunzi, S., “Optimal periodic impulsive strategies in glycemic control,” *IEEE Trans. on Contr. Syst. Technology*, vol. 32, no. 6, pp. 2062–2074, 2024.

- [31] Uusitupa, M. et al., “Long-term improvement in insulin sensitivity by changing lifestyles of people with impaired glucose tolerance: 4-year results from the finnish diabetes prevention study,” *Diabetes*, vol. 52, no. 10, pp. 2532–2538, 2003.
- [32] Gregg, E. W. et al., “Association of an intensive lifestyle intervention with remission of type 2 diabetes,” *Jama*, vol. 308, no. 23, pp. 2489–2496, 2012.
- [33] Colberg, S. R. et al., “Physical activity/exercise and diabetes: A position statement of the american diabetes association,” *Diabetes care*, vol. 39, no. 11, pp. 2065–2079, 2016.
- [34] Haskell, W. L. et al., “Physical activity and public health: Updated recommendation for adults from the american college of sports medicine and the american heart association,” *Circulation*, vol. 116, no. 9, pp. 1081–1093, 2007.
- [35] Ramachandran, A. et al., “The indian diabetes prevention programme shows that lifestyle modification and metformin prevent type 2 diabetes in asian indian subjects with impaired glucose tolerance (IDPP-1),” *Diabetologia*, vol. 49, pp. 289–297, 2006.
- [36] Kosaka, K., Noda, M., and Kuzuya, T., “Prevention of type 2 diabetes by lifestyle intervention: A japanese trial in IGT males,” *Diabetes research and clinical practice*, vol. 67, no. 2, pp. 152–162, 2005.
- [37] Tuomilehto, J. et al., “Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance,” *New England journal of medicine*, vol. 344, no. 18, pp. 1343–1350, 2001.
- [38] Sakane, N. et al., “Prevention of type 2 diabetes in a primary healthcare setting: Three-year results of lifestyle intervention in japanese subjects with impaired glucose tolerance,” *BMC public health*, vol. 11, no. 1, p. 40, 2011.
- [39] Luo, M., Yu, C., Cruz, B. D. P., Chen, L., and Ding, D., “Accelerometer-measured intensity-specific physical activity, genetic risk and incident type 2 diabetes: A prospective cohort study,” *British Journal of Sports Medicine*, vol. 57, no. 19, pp. 1257–1264, 2023.
- [40] Kany, S. et al., “Associations of “weekend warrior” physical activity with incident disease and cardiometabolic health,” *Circulation*, vol. 150, no. 16, pp. 1236–1247, 2024.
- [41] Camerlingo, N., Vettoretti, M., Del Favero, S., Facchinetti, A., and Sparacino, G., “Mathematical models of meal amount and timing variability with implementation in the type-1 diabetes patient decision simulator,” *Journal of Diabetes Science and Technology*, vol. 15, no. 2, pp. 346–359, 2021.
- [42] DeFronzo, R. A. and Abdul-Ghani, M. A., “Preservation of β -cell function: The key to diabetes prevention,” *The Journal of Clinical Endocrinology & Metabolism*, vol. 96, no. 8, pp. 2354–2366, 2011.
- [43] Khan, M. A. B., Hashim, M. J., King, J. K., Govender, R. D., Mustafa, H., and Kaabi, J. A., “Epidemiology of type 2 diabetes – global burden of disease and forecasted trends,” *Journal of Epidemiology and Global Health*, vol. 10, pp. 107–111, 1 2019.
- [44] Multerer, L., De Paola, P. F., Lenatti, M., Paglialonga, A., and Azzimonti, L., “Homogenization of ordinary differential equations for the fast prediction of diabetes progression,” *arXiv preprint arXiv:2412.16261*, 2024.
- [45] Lenatti, M. and Zaffalon, M. and Antonucci, A. and De Paola, P.F. and Multerer, L. and Mongelli, M. and Paglialonga, A. and Azzimonti, L., “Counterfactual inference using ordinary differential equations to assess the effect of physical activity on type 2 diabetes onset,” in *International Conference on Artificial Intelligence in Medicine*, Springer, 2025, pp. 212–221. DOI: [10.1007/978-3-031-95838-0_21](https://doi.org/10.1007/978-3-031-95838-0_21).

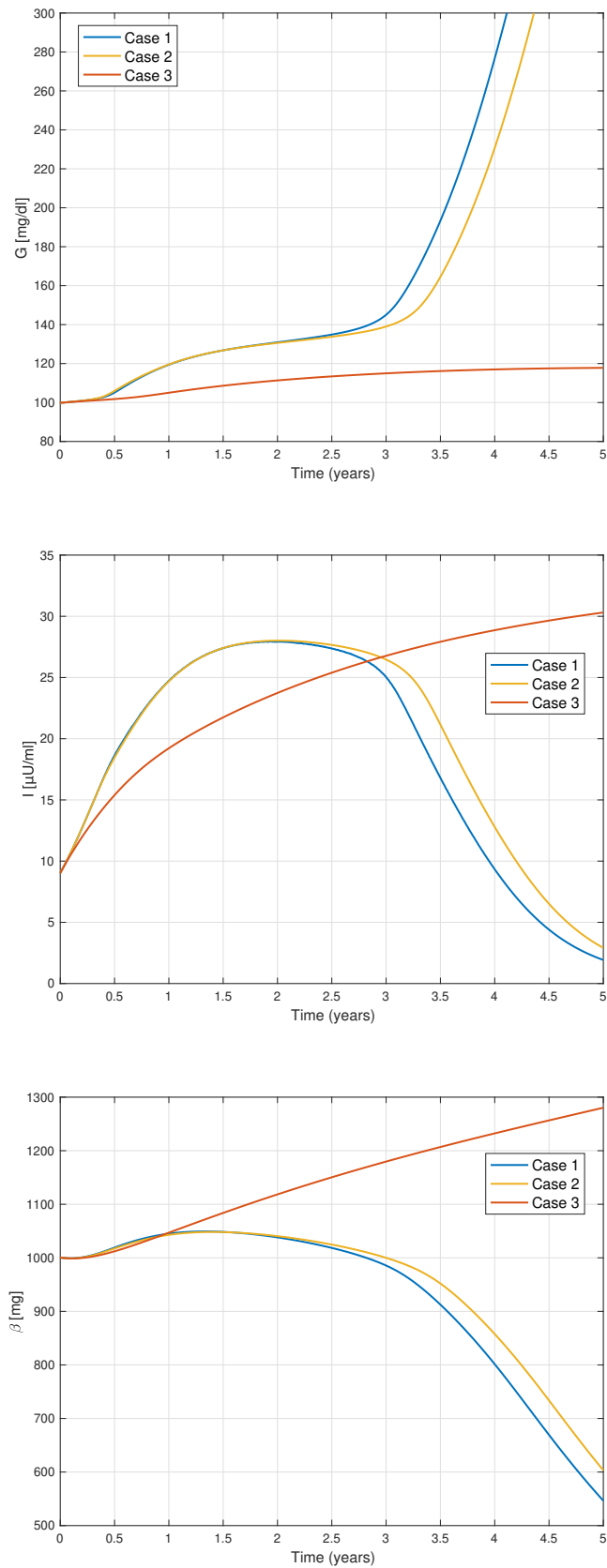


Figure 2.3: Top panel: Plasma basal glucose concentration as a function of time observed in Case 1 (blue), Case 2 (yellow), Case 3 (red). Center panel: Serum basal insulin concentration in the same simulated cases. Bottom panel: beta-cell mass in the same simulated cases.

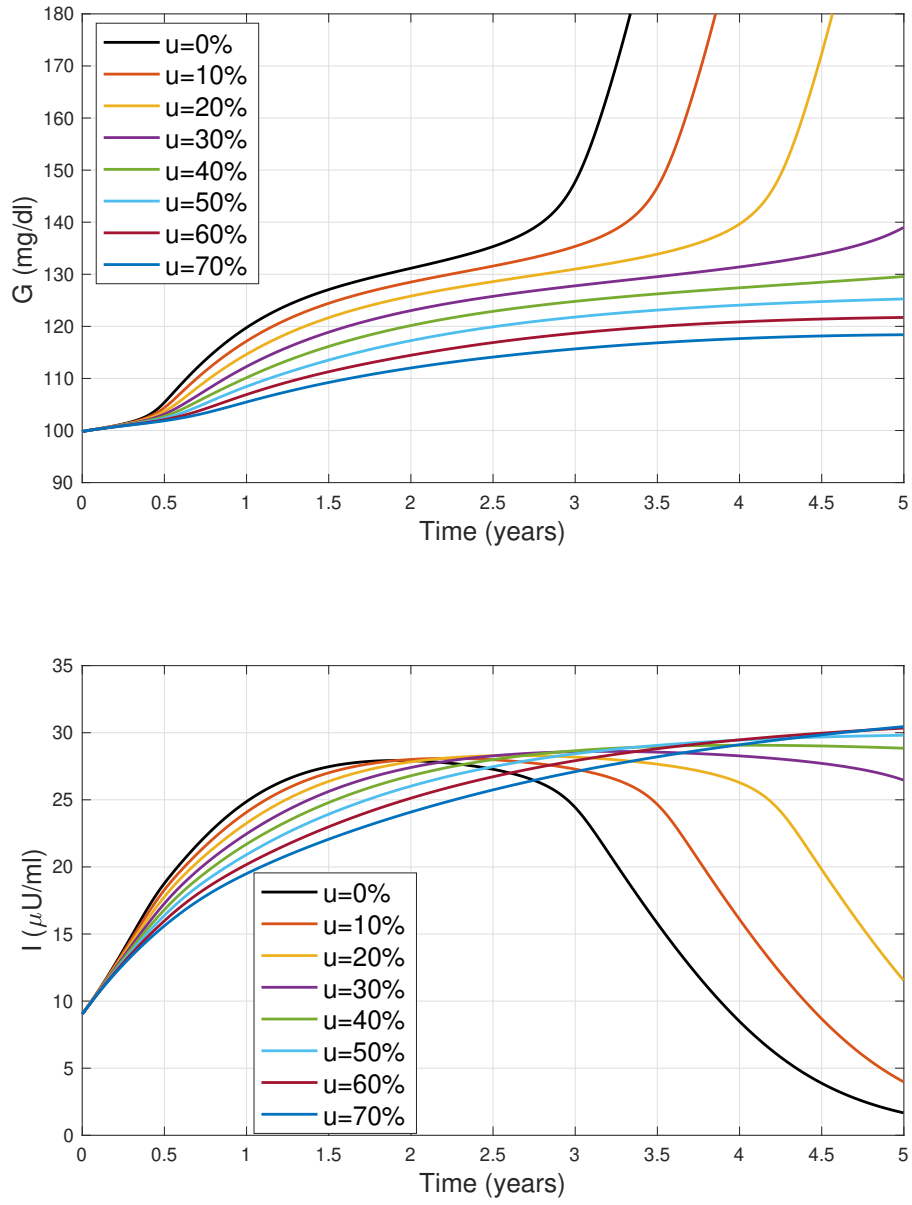


Figure 2.4: Basal glucose concentration (top panel) and insulin concentration (bottom panel) as a function of u over a five-year simulation horizon (simulations: 60 minutes/session, 3 sessions/week, $\tau_{SI} = 150$ days).

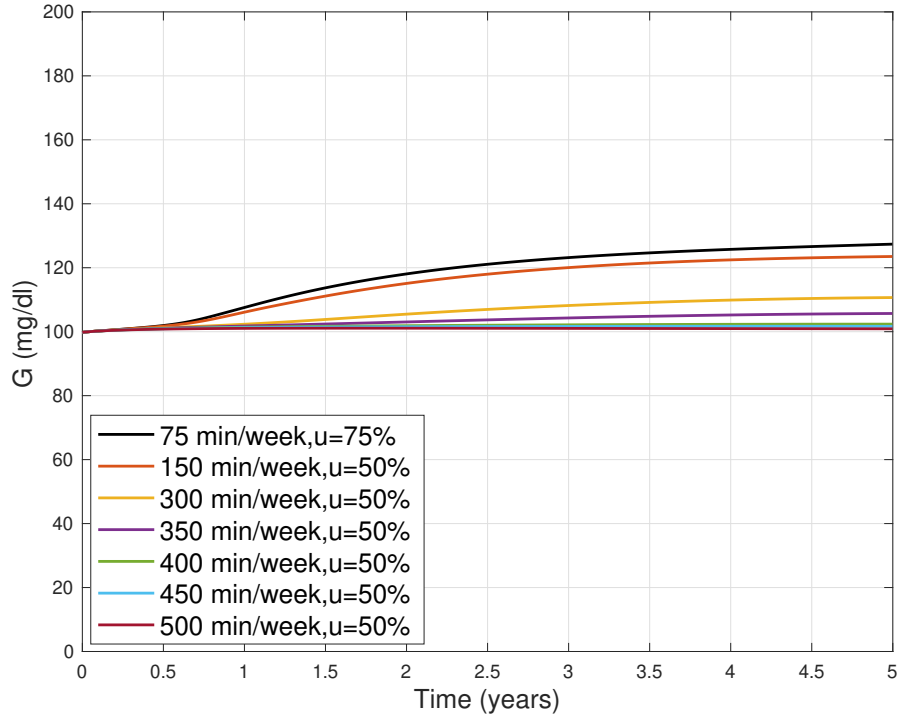


Figure 2.5: Basal glucose concentration with vigorous exercise ($u = 75\%$) for 75 minutes/week and moderate exercise ($u = 50\%$) as a function of exercise duration, from 150 to 500 minutes/week, using $\tau_{SI} = 180$ days.

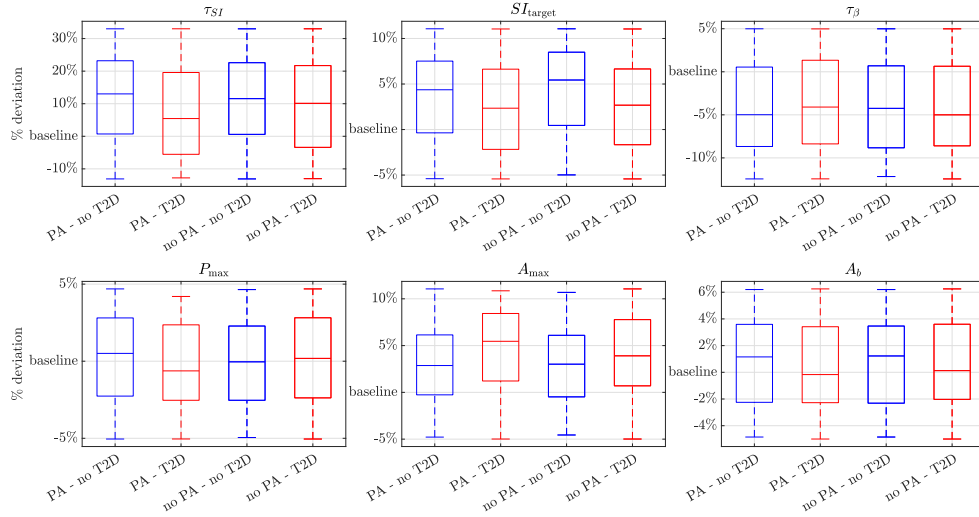


Figure 2.6: Distribution of six out of 18 parameters related to genetic predisposition and environmental effects (i.e., τ_{SI} , SI_{target} , τ_{β} , P_{max} , A_{max} , A_b) in Simulation Set GEP. Variation is expressed as percentage change from the nominal value.

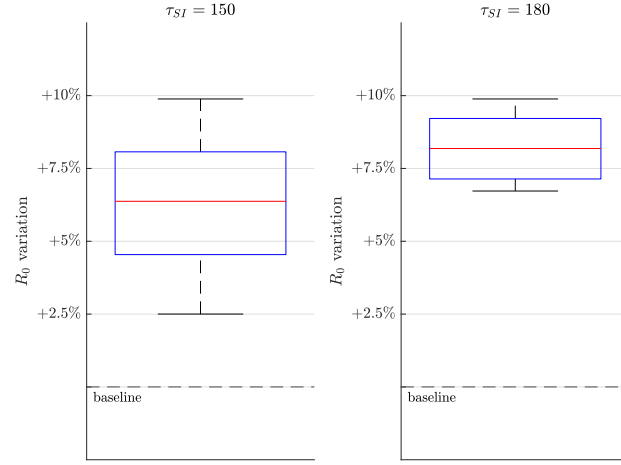


Figure 2.7: Distribution of R_0 variation in T2D subjects with high ($\tau_{SI} = 150$) and moderate ($\tau_{SI} = 180$) baseline risk in Simulation Set DIMB. Variation is expressed as percentage deviation from the nominal value of R_0 .

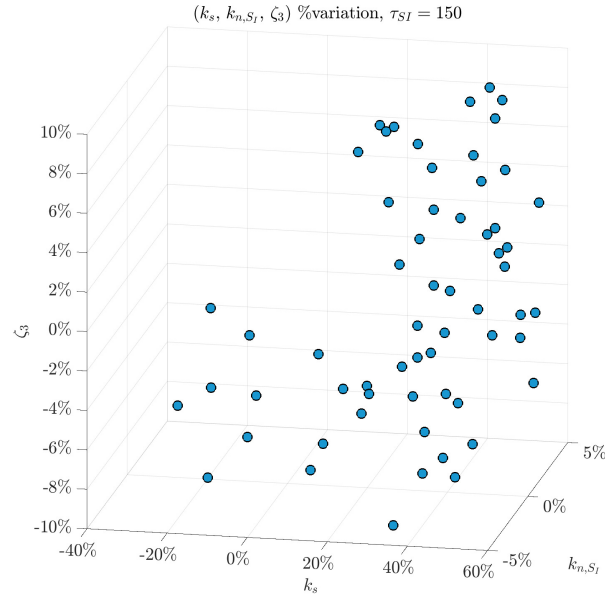


Figure 2.8: Distribution of the variations of $k_s, k_n, \tau_{SI}, \zeta_3$ in simulated individuals with T2D at the end of the simulation window in Simulation Set IRAI-a. Variations are expressed as percentage deviation from the nominal values of the parameters.

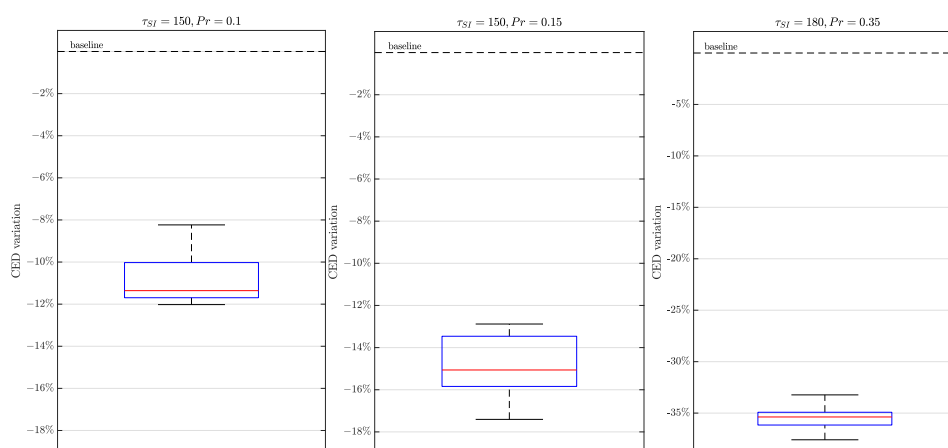


Figure 2.9: Distribution of the CED variation for simulated individuals with T2D at the end of the simulation window in Simulation Set IRAI-b. Variation is expressed as percentage deviation from the dose of the nominal training program, for $\tau_{SI} = 150, Pr = 0.1$ (left panel) and $\tau_{SI} = 150, Pr = 0.15$ (center panel) and $\tau_{SI} = 180, Pr = 0.35$ (right panel).

Chapter 3

On the Qualitative Analysis of the Model

Contents

3.1	Introduction	45
3.2	Formal Analysis of the Model	45
3.3	Properties of the Solutions	55
3.4	Conclusion	58

3.1 Introduction

This chapter is devoted to the formal analysis of the model presented in Chapter 2 [1], [2]. While in Chapter 2 we have shown how the model is suitable of reasonably assessing the benefits of physical activity in a wide range of simulations [1], [2], this chapter will focus on the analysis of its qualitative behavior. Qualitative analysis includes a broad spectrum of techniques and is widely employed in mathematical biology to characterize model stationary behavior by identifying equilibria and by identifying additional structural characteristics related to bifurcation, sensitivity analysis or further properties of the solutions [3], [4], [5], [6], [7], [8], [9], [10], [11]. Specifically, the contribution of this chapter is multi-fold.

First, with the further aim of exploiting the model for control purposes, we show how it is possible to simplify the exercise representation. To do this, we derive a *Quasi-steady State Representation* (QSSA) of the effect of the exercise, encoding in compact way all the related features (intensity, duration and period). Moreover, in this way, we show that the order of the full-state system can be reduced.

Then, we analyze the steady-state behavior of the model and provide a general qualitative analysis. In this regard, results clearly show the existence of two asymptotically stable stationary solutions on the long term: one related to a healthy equilibrium, where hyperglycemia is avoided thanks to the support of insulin produced by pancreatic beta-cells; the other related to a frank diabetic equilibrium where beta-cells are asymptotically cleared out and no more circulating insulin is available.

Finally, we prove key properties of the solutions of the system, that turn out to be crucial for exploiting the model for the design and development of model-based control techniques, such as the existence and uniqueness, positiveness and boundedness.

3.2 Formal Analysis of the Model

This section is devoted to the formal analysis of the model. More in detail, in this section we illustrate the QSSA representation of the exercise and how this can lead to a reduced-order model. Moreover, we provide a comprehensive analysis of the steady-state behavior of the model.

The full list of model equations, both differential and algebraic, with parameters we refer to in the subsequent developments are reported in Appendix A.



3.2.1 Quasi Stationary Approximation

As discussed in Chapter 2, the input u is a PWC input defined as

$$u(t) = \begin{cases} \bar{u} & t \in [kT, kT + D) \\ 0 & t \in [kT + D, (k+1)T), \end{cases} \quad (3.1)$$

where $\bar{u}$ is the exercise intensity, D the duration of the exercise session that occurs with a given period T . The working hypothesis is to assume a fixed shape of physical activity (i.e. constant $\bar{u}$, D and T) [12]. This fact allows to find out an average stationary value for PVO_2 to be exploited in the qualitative behavior analysis on the long term. From (A.1h), according to input ((3.1)) we can write the evolution of PVO_2 with $t \in [kT, (k+1)T]$ as follows:

$$PVO_2(t) = \begin{cases} PVO_2(kT)e^{-0.8(t-kT)} + \int_{kT}^t e^{-0.8(t-\tau)} 0.8\bar{u}d\tau, & t \in [kT, kT + D) \\ \left[PVO_2(kT)e^{-0.8D} + \int_{kT}^{kT+D} e^{-0.8(D+kT-\tau)} 0.8\bar{u}d\tau \right] e^{-0.8(t-kT-D)}, & t \in [kT + D, (k+1)T) \end{cases}$$

that is

$$PVO_2(t) = \begin{cases} PVO_2(kT)e^{-0.8(t-kT)} + \bar{u}(1 - e^{-0.8(t-kT)}), & t \in [kT, kT + D) \\ PVO_2(kT)e^{-0.8(t-kT)} + \bar{u}(1 - e^{-0.8D})e^{-0.8(t-kT-D)}, & t \in [kT + D, (k+1)T). \end{cases}$$

If we compute the average value in a given interval, we obtain:

$$\begin{aligned} \langle PVO_2 \rangle_k &= \frac{1}{T} \int_{kT}^{kT+D} \left(PVO_2(kT)e^{-0.8(t-kT)} + \bar{u}(1 - e^{-0.8(t-kT)}) \right) dt \\ &\quad + \frac{1}{T} \int_{kT+D}^{(k+1)T} PVO_2(kT)e^{-0.8(t-kT)} dt \\ &\quad + \frac{1}{T} \int_{kT+D}^{(k+1)T} \bar{u}(1 - e^{-0.8D})e^{-0.8(t-kT-D)} dt \\ &= \frac{1}{T} \int_{kT}^{(k+1)T} PVO_2(kT)e^{-0.8(t-kT)} dt + \frac{\bar{u}}{T} \int_{kT}^{kT+D} (1 - e^{-0.8(t-kT)}) dt \\ &\quad + \frac{\bar{u}}{T} \int_{kT+D}^{(k+1)T} (e^{-0.8(t-kT-D)} - e^{-0.8(t-kT)}) dt \\ &= \frac{1}{0.8T} PVO_2(kT)(1 - e^{-0.8T}) + \frac{\bar{u}D}{T} - \frac{\bar{u}}{0.8T} (1 - e^{-0.8D}) \\ &\quad + \frac{\bar{u}}{0.8T} (1 - e^{-0.8(T-D)} - e^{-0.8D} + e^{-0.8T}) \\ &= \frac{1}{0.8T} PVO_2(kT)(1 - e^{-0.8T}) + \bar{u}\delta + \frac{\bar{u}}{0.8T} (e^{-0.8T} - e^{-0.8(T-D)}) \\ &= \frac{1}{0.8T} PVO_2(kT)(1 - e^{-0.8T}) + \bar{u}\delta + \frac{\bar{u}}{0.8T} e^{-0.8T} (1 - e^{0.8D}). \end{aligned}$$

Then, in order to obtain a recursive equation for $\langle PVO_2 \rangle_k$, we write:

$$\langle PVO_2 \rangle_{k+1} = \frac{1}{0.8T} PVO_2((k+1)T)(1 - e^{-0.8T}) + \bar{u}\delta + \frac{\bar{u}}{0.8T} e^{-0.8T} (1 - e^{0.8D});$$

that is

$$\langle PVO_2 \rangle_{k+1} - \langle PVO_2 \rangle_k = \frac{1 - e^{-0.8T}}{0.8T} \left(PVO_2((k+1)T) - PVO_2(kT) \right).$$

Since

$$PVO_2((k+1)T) = e^{-0.8T} PVO_2(kT) + \bar{u} \left(e^{-0.8(T-D)} - e^{-0.8T} \right),$$

then

$$\langle PVO_2 \rangle_{k+1} - \langle PVO_2 \rangle_k = \frac{1 - e^{-0.8T}}{0.8T} \left(\bar{u} \left(e^{-0.8(T-D)} - e^{-0.8T} \right) - (1 - e^{-0.8T}) PVO_2(kT) \right).$$

Finally, since

$$PVO_2(kT)(1 - e^{-0.8T}) = 0.8T \langle PVO_2 \rangle_k - 0.8T \bar{u} \delta - \bar{u} e^{-0.8T} (1 - e^{0.8D}),$$

we have

$$\begin{aligned} \langle PVO_2 \rangle_{k+1} - \langle PVO_2 \rangle_k &= \frac{1 - e^{-0.8T}}{0.8T} \left(\bar{u} \left(e^{-0.8(T-D)} - e^{-0.8T} \right) - 0.8T \langle PVO_2 \rangle_k + 0.8T \bar{u} \delta \right. \\ &\quad \left. + \bar{u} e^{-0.8T} (1 - e^{0.8D}) \right) \\ &= \frac{1 - e^{-0.8T}}{0.8T} \left(-0.8T \langle PVO_2 \rangle_k + 0.8T \bar{u} \delta \right), \end{aligned}$$

so that

$$\langle PVO_2 \rangle_{k+1} = e^{-0.8T} \langle PVO_2 \rangle_k + \bar{u} \delta (1 - e^{-0.8T})$$

that is a sequence asymptotically converging to the solution of

$$\langle PVO_2 \rangle_\infty = e^{-0.8T} \langle PVO_2 \rangle_\infty + \bar{u} \delta (1 - e^{-0.8T}) \implies \langle PVO_2 \rangle_\infty = \bar{u} \delta = \bar{u} \cdot \frac{D}{T}.$$

Such a value may be reasonably upper bounded by a $\bar{u} \leq 60$ simulating moderate-to-vigorous exercise [13] and by a duty cycle δ ranging from a minimum of 0 to a maximum around 400 minutes/weeks (that is the threshold at which the benefits of physical activity are supposed to saturate according to clinical evidence [14]), with fixed period T . In summary, we have:

$$\langle PVO_2 \rangle_\infty = \bar{u} \delta = \bar{u} \cdot \frac{D}{T} \leq 2.5.$$

Hence, leveraging the use of a constant, in average, $\langle PVO_2 \rangle$, we have:

$$u_{eq} \simeq \langle PVO_2 \rangle_\infty = \bar{u} \cdot \frac{D}{T} \quad (3.2)$$

and we can readily obtain the stationary solutions for the fast dynamics (A.1i)-(A.1l):

$$G_{\text{prod},ss} = \frac{a_1}{a_2} u_{eq}, \quad (3.3)$$

$$G_{\text{up},ss} = \frac{a_3}{a_4} u_{eq}, \quad (3.4)$$

$$I_{e,ss} = \frac{a_5}{a_6} u_{eq}, \quad (3.5)$$

$$IL_{6,ss} = \frac{SR_{\text{ex}}}{k_m} u_{eq}. \quad (3.6)$$

The IL-6 stationary value in (3.6) allows to obtain the stationary solution for V_l as :

$$V_{l,ss} = \frac{SR_{\text{ex}}}{k_s k_m} u_{eq}, \quad (3.7)$$

leading to the following low dynamics algebraic equations for the rest of stationary values:

$$R_0 - (Eg_0 + S_{I,ss} \cdot I_{ss}) G_{ss} + \frac{W}{Vol_G} \left(\frac{a_1}{a_2} - \frac{a_3}{a_4} \right) u_{eq} = 0, \quad (3.8a)$$

$$\beta_{ss} \frac{ISR(G_{ss}, \gamma_{ss}, \sigma_{ss})}{V} - kI_{ss} - \frac{a_5}{a_6} u_{eq} \cdot \text{Heav}(I_{ss}) = 0, \quad (3.8b)$$

$$(\bar{P}_{ISR}(G_{ss}, \gamma_{ss}, \sigma_{ss}, V_{l,ss}) - \bar{A}_{M_G}(G_{ss}, V_{l,ss})) \beta_{ss} = 0, \quad (3.8c)$$

$$\gamma_{ss} = \gamma_{\infty, G}(G_{ss}), \quad (3.8d)$$

$$\sigma_{ss} = \sigma_{ISR_{M_\infty}}(G_{ss}, \gamma_{ss}, \sigma_{ss}), \quad (3.8e)$$

$$(S_{I,ss} - S_{I,\text{target}}) \cdot \left(1 - \zeta_3 \frac{V_{l,ss}^n}{k_{n,S_I}^n + V_{l,ss}^n} \right) = 0. \quad (3.8f)$$

The last equation (3.8f) can be simplified as

$$S_{I,ss} = S_{I,\text{target}}, \quad (3.9)$$

since there is no reason to suppose the equality constraint

$$\zeta_3 \frac{V_{l,ss}^n}{k_{n,S_I}^n + V_{l,ss}^n} = 1, \quad (3.10)$$

among the model parameters.

As discussed, the model is two-timescale and involves both fast and slow dynamics. The concept of u_{eq} induces a QSSA of the fast subsystem (A.1h)-(A.1l). Hence, when adopting u_{eq} , the fast variables can be practically considered at their stationary values, leading to an equivalent reduced-order model without affecting the overall long-term behavior of the slow dynamics.

3.2.2 Analysis of the Steady State

Equation (3.9) allows to write the stationary insulinemia as a function of stationary glycemia:

$$I_{ss} = \frac{R_0 + \frac{W}{Vol_G} \left(\frac{a_1}{a_2} - \frac{a_3}{a_4} \right) u_{eq}}{G_{ss} S_{I,\text{target}}} - \frac{Eg_0}{S_{I,\text{target}}} := \varphi_1(G_{ss}, u_{eq}). \quad (3.11)$$

Note that a necessary condition to make equation (3.11) meaningful is that

$$R_0 + \frac{W}{Vol_G} \left(\frac{a_1}{a_2} - \frac{a_3}{a_4} \right) u_{eq} > 0, \quad (3.12)$$

otherwise we would have $I_{ss} < 0$ whatever the positive values of the model parameters and of the stationary glycemia.

The following result ensures that this condition is always satisfied.

Proposition 3.2.1

The necessary condition in (3.12) always holds for $u_{eq} < u_{eq,thr} \simeq 83.63$. $\square$

Proof. The proof straightforwardly follows observing that from (3.12) we get

$$u_{eq} < \frac{Vol_G}{W} R_0 \left(\frac{a_2 a_4}{a_2 a_3 - a_1 a_4} \right) = u_{eq,thr},$$

with $u_{eq,thr} \approx 83.63$ in the parameter values, which concludes the proof. $\square$

As previously stated, it is reasonable to suppose that $u_{eq} \leq 2.5$, therefore the necessity condition is well satisfied by a wide margin. It can be easily noted that $\varphi_1(G_{ss}, u_{eq})$ is a monotonic decreasing function and moreover

$$\lim_{G_{ss} \rightarrow 0^+} \varphi_1(G_{ss}, u_{eq}) = +\infty, \quad \lim_{G_{ss} \rightarrow +\infty} \varphi_1(G_{ss}, u_{eq}) = -\frac{Eg_0}{S_{I, \text{target}}} < 0 \quad (3.13)$$

and

$$\varphi_1(G_{ss}, u_{eq}) = 0 \quad \implies \quad G_{ss} = \frac{R_0}{Eg_0} + \frac{W}{Eg_0 Vol_G} \left(\frac{a_1}{a_2} - \frac{a_3}{a_4} \right) u_{eq}. \quad (3.14)$$

It turns out that in the values of the parameters such a stationary glucose threshold to keep positive stationary insulinemia is about 600mg/dl. Figure 3.1 shows $\varphi_1(G_{ss}, u_{eq})$ for different values of u_{eq} .

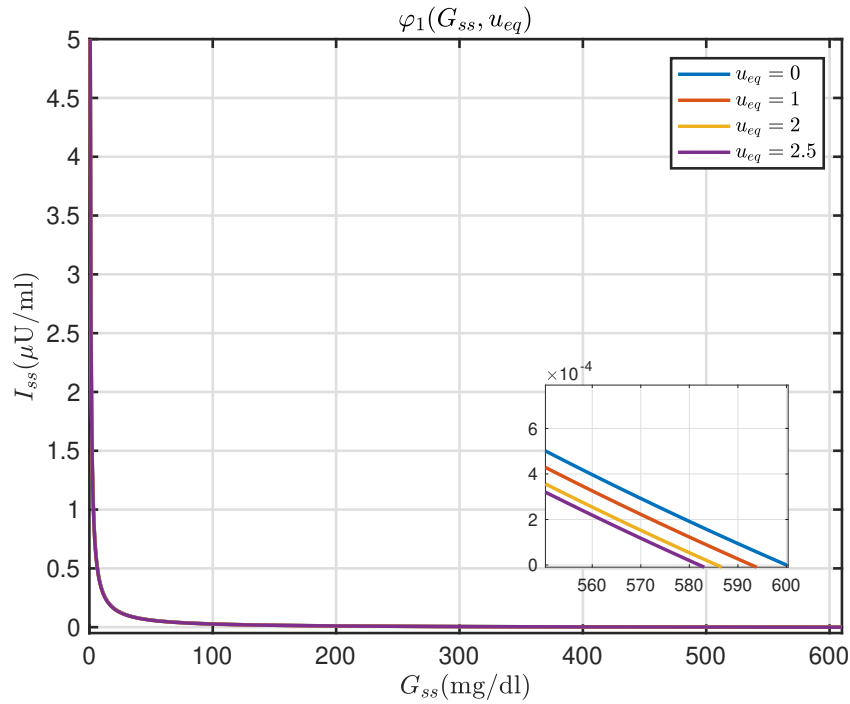


Figure 3.1: Stationary plasma insulinemia as a function of stationary plasma glycemia, at different values of u_{eq} .

For what concerns γ_{ss} , it may be written as a function of the only stationary glycemia, independent of the physical activity:

$$\gamma_{ss} = \gamma_{\infty, G}(G_{ss}) = \frac{\gamma_{\max}}{1 + e^{-\frac{G_{ss} - \gamma_s}{\gamma_n}}} - \gamma_0 := \varphi_2(G_{ss}). \quad (3.15)$$

It can be easily observed that this function is monotonically decreasing with

$$\lim_{G_{ss} \rightarrow 0^+} \varphi_2(G_{ss}) = \frac{\gamma_{\max}}{1 + e^{\frac{\gamma_s}{\gamma_n}}} - \gamma_0 \simeq -0.1, \quad \lim_{G_{ss} \rightarrow +\infty} \varphi_2(G_{ss}) = \gamma_{\max} - \gamma_0 \simeq 0.1. \quad (3.16)$$

Figure 3.2 reports the saturating-function-like behavior of $\gamma_{ss}(G_{ss})$ according to the parameter values.

Let us now consider equation (3.8c), that turns out to be crucial for the computation of the steady state of the model. Indeed, it can be readily noted that equation (3.8c) induces two steady-state conditions for the system, associated respectively with a non-trivial equilibrium point at $\beta \neq 0$ and a trivial equilibrium point at $\beta = 0$, that we analyze separately in the following.

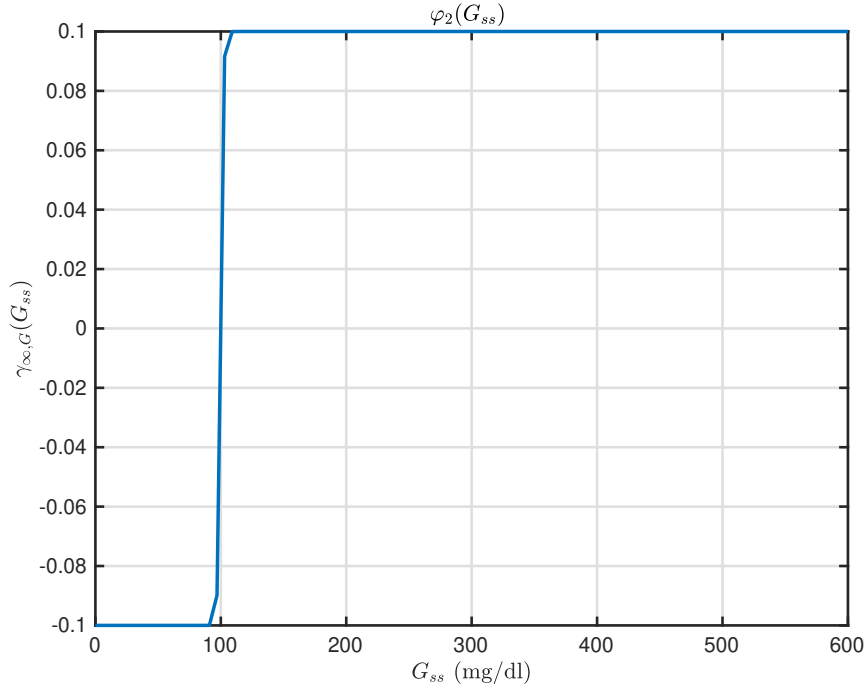


Figure 3.2: Stationary γ as a function of stationary plasma glycemia.

3.2.2.1 The Stationary Case with $\beta_{ss} \neq 0$ and Related Stability Analysis

For the stationary case with $\beta_{ss} \neq 0$, equation (3.8c) becomes:

$$\bar{P}_{ISR}(G_{ss}, \gamma_{ss}, \sigma_{ss}, V_{l,ss}) = \bar{A}_M(G_{ss}, V_{l,ss}). \quad (3.17)$$

By denoting

$$H_1(G_{ss}) := \frac{(M_G(G_{ss}) + \gamma_{ss})^{k_{ISR}}}{(M_G(G_{ss}) + \gamma_{ss})^{k_{ISR}} + \alpha_{ISR}^{k_{ISR}}} = \frac{(M_G(G_{ss}) + \varphi_2(G_{ss}))^{k_{ISR}}}{(M_G(G_{ss}) + \varphi_2(G_{ss}))^{k_{ISR}} + \alpha_{ISR}^{k_{ISR}}}, \quad (3.18)$$

$$H_2(G_{ss}) := A_{\max} \frac{M_G(G_{ss})^{k_A}}{M_G(G_{ss})^{k_A} + \alpha_A^{k_A}} + A_b \quad (3.19)$$

and

$$\Xi_1 := 1 + \frac{\zeta_1 V_{l,ss}^2}{k_{n,P}^2 + V_{l,ss}^2} > 0, \quad \Xi_2 := 1 - \frac{\zeta_2 V_{l,ss}^2}{k_{n,A}^2 + V_{l,ss}^2}, \quad (3.20)$$

it is possible to shortly write

$$P_{\max} \frac{\sigma_{ss}^{k_P} H_1(G_{ss})^{k_P}}{\sigma_{ss}^{k_P} H_1(G_{ss})^{k_P} + \alpha_P^{k_P}} \Xi_1 = H_2(G_{ss}) \Xi_2. \quad (3.21)$$

For what concerns Ξ_2 , it is easy to observe that $\Xi_2 > 0$, being the Hill-function

$$\frac{V_{l,ss}^2}{k_{n,A}^2 + V_{l,ss}^2} < 1 \quad (3.22)$$

by definition whatever the stationary value of V_l in the values of the parameters. This makes the steady-state equation (3.21) meaningful.

If we define

$$\varphi_3(G_{ss}) := \frac{\Xi_2}{\Xi_1 P_{\max}} H_2(G_{ss}), \quad (3.23)$$

from equation (3.21) it is possible to obtain

$$\sigma_{ss}^{k_P} = \frac{\alpha_P^{k_P} \varphi_3(G_{ss})}{H_1(G_{ss})^{k_P} (1 - \varphi_3(G_{ss}))}. \quad (3.24)$$

It is worth noting that $\varphi_3(\cdot)$ formally depends of the physical activity by means of Ξ_1 and Ξ_2 , both depending of u_{eq} by means of $V_{l,ss}$, (3.7). However, according to parameter values, both $\Xi_1 \simeq \Xi_2 \simeq 1$, therefore φ_3 may be thought practically independent of the physical activity. According to equation (3.24), it is possible to state the following.

Lemma 3.2.1

A solution to (3.24) exists provided that

$$G_{ss} < \alpha_M \sqrt[k_M]{\frac{\xi}{1-\xi}}, \quad (3.25)$$

with

$$\xi = \alpha_A \sqrt[k_A]{\frac{\Theta}{1-\Theta}} \in (0, 1), \quad \Theta = \frac{\Xi_1 P_{\max}}{\Xi_2 A_{\max}} - \frac{A_b}{A_{\max}} \in (0, 1) \quad (3.26)$$

□

Proof. According to parameter values, $k_P = 4$, so that solution to (3.24) exists provided that

$$\varphi_3(G_{ss}) < 1, \quad \implies \quad H_2(G_{ss}) < \frac{\Xi_1 P_{\max}}{\Xi_2}. \quad (3.27)$$

Since $H_2(G_{ss}) \in \mathcal{I}$, with $\mathcal{I} = [A_b, A_{\max}] = [0.8, 9.8]$ and $\Xi_1, \Xi_2 \simeq 1$ with $P_{\max} = 4.55$ in the parameters, equality (3.24) holds if

$$A_{\max} \frac{M_G(G_{ss})^{k_A}}{M_G(G_{ss})^{k_A} + \alpha_A^{k_A}} + A_b < \frac{\Xi_1 P_{\max}}{\Xi_2}, \quad (3.28)$$

that is

$$M_G(G_{ss})^{k_A} < \alpha_A^{k_A} \cdot \frac{\Theta}{1-\Theta}, \quad \Theta = \frac{\Xi_1 P_{\max}}{\Xi_2 A_{\max}} - \frac{A_b}{A_{\max}} \in (0, 1). \quad (3.29)$$

In turn, this implies, being $k_A = 6$:

$$M_G(G_{ss}) = \frac{G_{ss}^{k_M}}{G_{ss}^{k_M} + \alpha_M^{k_M}} < \alpha_A \sqrt[k_A]{\frac{\Theta}{1-\Theta}}. \quad (3.30)$$

With $\alpha_A = 0.44$ from the parameter values, this last equation is meaningful, hence we have (for $k_M = 2$):

$$G_{ss} < \alpha_M \sqrt[k_M]{\frac{\xi}{1-\xi}}, \quad \xi = \alpha_A \sqrt[k_A]{\frac{\Theta}{1-\Theta}} \in (0, 1). \quad (3.31)$$

□

Such an upper bound for G_{ss} is roughly equal to 126.6 mg/dl, according to parameter values so that, in summary, we may state that a stationary solution for $\beta_{ss} \neq 0$ exists provided that the stationary value for plasma glycemia is below 126.6 mg/dl, which is a meaningful result since that is a healthy equilibrium.

Moreover, this allows to rewrite σ_{ss} as a function of G_{ss} as follows:

$$\sigma_{ss} := \varphi_4(G_{ss}) = \alpha_P \sqrt[k_P]{\frac{\varphi_3(G_{ss})}{H_1(G_{ss})^{k_P} (1 - \varphi_3(G_{ss}))}}. \quad (3.32)$$

Figure 3.3 reports the evolution of $\varphi_4(\cdot)$ within the meaningful range $G_{ss} \in [0, 126.6]$ mg/dl. Finally, it is possible to state the following main result.

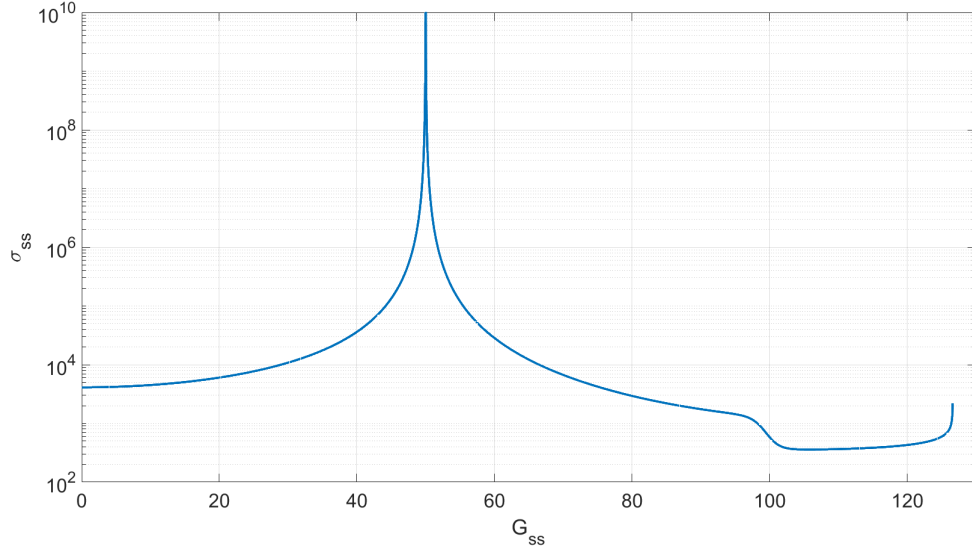


Figure 3.3: Stationary insulin secretion cabability versus stationary plasma glycemia.

Proposition 3.2.2

The model exhibits two positive solutions for the stationary glycemia values G_{ss}

$$G_{1,ss} \simeq 100\text{mg/dl}, \quad G_{2,ss} \simeq 125\text{mg/dl}.$$

□

Proof. Considering the steady-state equation for the insulin secretion capability dynamics (3.8e) and substituting (3.32) in (3.8e) we obtain

$$\varphi_5(G_{ss}) = \varphi_4(G_{ss}) - \sigma_{ISR_\infty}(G_{ss}, \varphi_2(G_{ss}), \varphi_4(G_{ss})) \cdot \sigma_{M_\infty}(G_{ss}) - \sigma_b = 0, \quad (3.33)$$

which uniquely depends on G_{ss} . Consequently, solving numerically

$$\varphi_5(G_{ss}) = 0$$

the result follows. □

It should be noted that $\varphi_5(G_{ss})$ does not practically depend on physical activity: indeed, $\varphi_5(\cdot)$ depends on $\varphi_3(\cdot)$ and $H_1(\cdot)$, with $H_1(\cdot)$ independent of the physical activity and $\varphi_3(\cdot)$ formally depending on the physical activity by means of Ξ_1 and Ξ_2 , both depending of $V_{l,ss}$. However, as previously stated, both Ξ_1 and Ξ_2 are $\simeq 1$, so that the stationary values of G_{ss} may be thought of as independent of the physical activity.

By replacing G_{ss} in $\gamma_{ss} = \varphi_2(G_{ss})$ and in $\sigma_{ss} = \varphi_4(G_{ss}) > 0$ we find two triples of meaningful stationary values for plasma glycemia, shift in glucose response γ and insulin secretion capability σ independent of the physical activity. Finally, according to $I_{ss} = \varphi_1(G_{ss}, u_{eq})$ and to the parameter values, we obtain a pair of positive stationary values for plasma insulinemia, these ones clearly depending of the physical activity, as well as the pair of stationary values for beta cells obtained by inverting equation (3.8b):

$$\beta_{ss} = \frac{V}{ISR(G_{ss}, \gamma_{ss}, \sigma_{ss})} \cdot \left(kI_{ss} + \frac{a_5}{a_6} u_{eq} \right).$$

In summary, we obtain a pair of meaningful stationary solutions for the long-term system.

For what concerns the stability analysis, it is analytically complex, hence it is evaluated numerically. Evaluating the eigenvalues of the Jacobian matrix at the equilibrium points associated with the stationary glycemia values $G_{1,ss} \simeq 100\text{mg/dl}$ and $G_{2,ss} \simeq 125\text{mg/dl}$ it results that the equilibrium points are respectively an asymptotically stable fixed point and a saddle.

3.2.2.2 The Stationary Case with $\beta_{ss} = 0$ and Related Stability Analysis

In the case of $\beta_{ss} = 0$ in (3.8c), we straightforwardly have from (3.8b) that also $I_{ss} = 0$. The new equilibrium we are searching clearly refers to an insulin-less stationary case. In turn, by substituting $I_{ss} = 0$ in (3.8a), we straightforwardly compute

$$G_{ss} = \frac{R_0}{Eg_0} + \frac{W}{Eg_0 Vol_G} \left(\frac{a_1}{a_2} - \frac{a_3}{a_4} \right) u_{eq} \simeq 600 \text{mg/dl}, \quad (3.34)$$

that is positive because of (3.12). The stationary value for γ_{ss} is always constrained to G_{ss} according to (3.8d), so that, according to the hyperglycemic state, we have

$$\gamma_{ss} \simeq \gamma_{max} - \gamma_0 \simeq 0.1. \quad (3.35)$$

The steady-state value for σ is achieved by solving equation (3.8e):

$$\sigma_{ss} = \sigma_{ISR_\infty}(G_{ss}, \gamma_{ss}, \sigma_{ss}) \sigma_{M_\infty}(G_{ss}) + \sigma_b. \quad (3.36)$$

The left-hand side of equation (3.36) is the bisector of the positive orthant. Dealing with the right-hand side as a function of σ_{ss} , it is a monotonically increasing function from a positive value

$$\sigma_{ISR_\infty}(G_{ss}, \gamma_{ss}, \sigma_{ss} = 0) \sigma_{M_\infty}(G_{ss}) + \sigma_b > 0, \quad (3.37)$$

up to the positive asymptotic value

$$\sigma_{ISR_\infty}(G_{ss}, \gamma_{ss}, \sigma_{ss} = +\infty) \sigma_{M_\infty}(G_{ss}) + \sigma_b > 0, \quad (3.38)$$

so that there exists a unique intersection. Such an intersection may be approximately found by exploiting the hyperglycemic state, providing that

$$\sigma_{M_\infty}(G_{ss}) \simeq 0 \quad \implies \quad \sigma_{ss} \simeq \sigma_b. \quad (3.39)$$

In summary, this equilibrium point is:

$$G_{ss} \simeq 600, \quad I_{ss} = 0, \quad \beta_{ss} = 0, \quad \gamma_{ss} \simeq \gamma_{max} - \gamma_0 \quad \sigma_{ss} \simeq \sigma_b \quad (3.40)$$

with $S_{I,ss} = S_{I,target}$.

For what concerns the stability analysis, the following result holds.

Theorem 3.2.1

The equilibrium point associated with $\beta_{ss} = 0$ is asymptotically stable. □

Proof. We observe that the Jacobian matrix can be partitioned as

$$J = \begin{bmatrix} J_{11} & J_{12} \\ O_{6 \times 6} & J_{22} \end{bmatrix} \quad (3.41)$$

where $O_{6 \times 6}$ is a 6×6 block of zero. Hence the stability analysis can be carried out according to J_{11} and J_{22} independently. Regards to J_{22} we have

$$J_{22} = \begin{bmatrix} -k_s & 0 & 0 & 0 & 0 & 1 \\ 0 & -0.8 & 0 & 0 & 0 & 0 \\ 0 & a_1 & -a_2 & 0 & 0 & 0 \\ 0 & a_3 & 0 & -a_2 & 0 & 0 \\ 0 & a_5 & 0 & 0 & -a_6 & 0 \\ 0 & SR_{ex} & 0 & 0 & 0 & -k_m \end{bmatrix} \quad (3.42)$$

whose eigenvalues can be readily obtained from the main diagonal although the Jacobian is not a diagonal matrix. It can be readily checked that they are all negative, i.e. asymptotically stable. Dealing with matrix J_{11} , we have

$$J_{11} = \begin{pmatrix} J_{11}(1,1) & J_{11}(1,2) & 0 & 0 & 0 & -I_{ss}G_{ss} \\ J_{11}(2,1) & -k & \frac{ISR(G_{ss}, \gamma_{ss}, \sigma_{ss})}{V} & J_{11}(2,4) & J_{11}(2,5) & 0 \\ J_{11}(3,1) & 0 & J_{11}(3,3) & J_{11}(3,4) & J_{11}(3,5) & 0 \\ \frac{\varphi'_2(G_{ss})}{\tau_\gamma} & 0 & 0 & -\frac{1}{\tau_\gamma} & 0 & 0 \\ J_{11}(5,1) & 0 & 0 & J_{11}(5,4) & J_{11}(5,5) & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{\Xi_3(V_{l,ss})}{\tau_{SI}} \end{pmatrix} \quad (3.43)$$

where

$$J_{11}(1,1) = -(Eg_0 + S_{I,target}I_{ss}), \quad J_{11}(1,2) = -S_{I,target}G_{ss} \quad (3.44)$$

$$J_{11}(2,1) = \frac{\beta_{ss}}{V} \frac{\partial ISR}{\partial G} \Big|_{G_{ss}, \gamma_{ss}, \sigma_{ss}}, \quad J_{11}(2,4) = \frac{\beta_{ss}}{V} \frac{\partial ISR}{\partial \gamma} \Big|_{G_{ss}, \gamma_{ss}, \sigma_{ss}} \quad (3.45)$$

$$J_{11}(2,5) = \frac{\beta_{ss}}{V} \frac{\partial ISR}{\partial \sigma} \Big|_{G_{ss}, \gamma_{ss}, \sigma_{ss}}, \quad J_{11}(3,3) = \frac{\bar{P}_{ISR} - \bar{A}_{MG}}{\tau_\beta} \Big|_{G_{ss}, \gamma_{ss}, \sigma_{ss}, V_{l,ss}} \quad (3.46)$$

$$J_{11}(3,1) = \frac{\beta_{ss}}{\tau_\beta} \left(\frac{\partial \bar{P}_{ISR}}{\partial G} - \frac{\partial \bar{A}_{MG}}{\partial G} \right) \Big|_{G_{ss}, \gamma_{ss}, \sigma_{ss}, V_{l,ss}} \quad (3.47)$$

$$J_{11}(3,4) = \frac{\beta_{ss}}{\tau_\beta} \frac{\partial \bar{P}_{ISR}}{\partial \gamma} \Big|_{G_{ss}, \gamma_{ss}, \sigma_{ss}, V_{l,ss}}, \quad J_{11}(3,5) = \frac{\beta_{ss}}{\tau_\beta} \frac{\partial \bar{P}_{ISR}}{\partial \sigma} \Big|_{G_{ss}, \gamma_{ss}, \sigma_{ss}, V_{l,ss}} \quad (3.48)$$

$$J_{11}(5,1) = \frac{1}{\tau_\sigma} \frac{\partial \sigma_{ISRM\infty}}{\partial G} \Big|_{G_{ss}, \gamma_{ss}, \sigma_{ss}}, \quad J_{11}(5,4) = \frac{1}{\tau_\sigma} \frac{\partial \sigma_{ISRM\infty}}{\partial \gamma} \Big|_{G_{ss}, \gamma_{ss}, \sigma_{ss}} \quad (3.49)$$

$$J_{11}(5,5) = \frac{1}{\tau_\sigma} \left(\frac{\partial \sigma_{ISRM\infty}}{\partial \sigma} \Big|_{G_{ss}, \gamma_{ss}, \sigma_{ss}} - 1 \right), \quad (3.50)$$

$$\Xi_3(V_{l,ss}) = 1 - \zeta_3 \frac{V_{l,ss}^n}{k_{n,SI}^n + V_{l,ss}^n}. \quad (3.51)$$

From computations, the Jacobian matrix J_{11} simplifies into:

$$J_{11} = \begin{pmatrix} -Eg_0 & J_{11}(1,2) & 0 & 0 & 0 & 0 \\ 0 & -k & \frac{ISR(G_{ss}, \gamma_{ss}, \sigma_{ss})}{V} & 0 & 0 & 0 \\ 0 & 0 & J_{11}(3,3) & 0 & 0 & 0 \\ \frac{\varphi'_2(G_{ss})}{\tau_\gamma} & 0 & 0 & -\frac{1}{\tau_\gamma} & 0 & 0 \\ J_{11}(5,1) & 0 & 0 & J_{11}(5,4) & J_{11}(5,5) & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{\Xi_3(V_{l,ss})}{\tau_{SI}} \end{pmatrix} \quad (3.52)$$

In light of the block triangular form of J_{11} , it readily comes that the eigenvalues come out from the diagonal. Hence to conclude about the asymptotic stability, we need to investigate the sign of $J_{11}(3,3)$ and $J_{11}(5,5)$. For what concerns to $J_{11}(3,3) < 0$, we have:

$$\begin{aligned} P_{\max} \frac{ISR(G_{ss}, \gamma_{ss}, \sigma_b)^{k_P}}{ISR(G_{ss}, \gamma_{ss}, \sigma_b)^{k_P} + \alpha_P^{k_P}} \Xi_1(V_{l,ss}) \\ < \left(A_{\max} \frac{M_G(G_{ss})^{k_A}}{M_G(G_{ss})^{k_A} + \alpha_A^{k_A}} + A_b \right) \Xi_2(V_{l,ss}) \end{aligned} \quad (3.53)$$

that simplifies into

$$P_{\max} \frac{ISR(G_{ss}, \gamma_{ss}, \sigma_b)^{k_P}}{ISR(G_{ss}, \gamma_{ss}, \sigma_b)^{k_P} + \alpha_P^{k_P}} < \left(A_{\max} \frac{M_G(G_{ss})^{k_A}}{M_G(G_{ss})^{k_A} + \alpha_A^{k_A}} + A_b \right) \quad (3.54)$$

Inequality (3.54) is widely verified for meaningful values of physical activity.

Regarding $J_{11}(5, 5) < 0$, we have:

$$\left. \frac{\partial \sigma_{ISR_{M_\infty}}}{\partial \sigma} \right|_{G_{ss}, \gamma_{ss}, \sigma_b} < 1 \quad (3.55)$$

that is

$$\left. \frac{\partial \sigma_{ISR_\infty}}{\partial \sigma} \right|_{G_{ss}, \gamma_{ss}, \sigma_b} \cdot \sigma_{M_\infty}(G_{ss}) < 1 \quad (3.56)$$

that is broadly achieved as well because of the hyperglycemic state of G_{ss} making $\sigma_{M_\infty}(G_{ss}) \simeq 10^{-17}$. $\square$

In summary, the system has two asymptotically stable equilibrium points. One refers to a healthy state where diabetes progression is somehow limited to a stationary plasma glycemia of about 100 mg/dl. Such a stationary state is not influenced by the physical activity, that impacts on the transient and, possibly, may impact on the basin of attraction. The other equilibrium point is a frank T2D position, providing a hyperglycemic state of about 600mg/dl with zero beta cells and, straightforwardly, zero endogenous insulin. In this case the physical activity may impact the position of the hyperglycemic state.

3.3 Properties of the Solutions

This section is devoted to illustrating some key properties of the model, such as positiveness, boundedness, existence and uniqueness of the solution. For what concerns positiveness and boundedness, we rigorously prove these properties for the reduced-order system. When the full-order system is considered, for the remaining subset of variables associated with the linear equations (A.1h)-(A.1l) these properties straightforwardly hold.

3.3.1 Positiveness of the Solution

To prove positiveness, we rely on the following result.

Lemma 3.3.1 (Positiveness)

The solutions $G(t), I(t), \beta(t), \sigma(t), S_I(t), V_I(t)$ are nonnegative for any nonnegative initial condition. $\square$

Proof. To prove that the solutions of the system are nonnegative for any nonnegative initial conditions, we examine each differential equation showing that the nonnegative orthant is invariant under the dynamics of the system. Consider equation (A.1a). It can be observed that at the boundary $G = 0$ it holds

$$\left. \frac{dG}{dt} \right|_{G=0} = R_0 + \frac{W}{Vol_G} (G_{\text{prod}} - G_{\text{up}}),$$

that is always positive according to Proposition 3.2.1. Hence, $\forall t \geq 0, \forall G(0) \geq 0, G(t) \geq 0$. Consider equation (A.1c). At $\beta = 0$

$$\left. \frac{d\beta}{dt} \right|_{\beta=0} = 0,$$

thus, being $\beta(0) \geq 0$ by assumption, $\beta(t) \geq 0 \forall t \geq 0$
 Consider equation (A.1b). At the boundary $I = 0$ it holds

$$\left. \frac{dI}{dt} \right|_{I=0} = \beta \frac{ISR}{V} \geq 0,$$

as $\beta \geq 0$, $ISR \geq 0$ being a Hill function, and $V > 0$ in the parameters values. Hence $I(t) \geq 0$, $\forall I(0) \geq 0$, $\forall t \geq 0$.

Consider equation (A.1e). At the boundary $\sigma = 0$, it holds

$$\left. \frac{d\sigma}{dt} \right|_{\sigma=0} = \frac{\sigma_{ISR_M \infty}}{\tau_\sigma} > 0,$$

in the parameters, as discussed. It follows that $\sigma(t) \geq 0$, $\forall \sigma(0) \geq 0$, $\forall t \geq 0$. The positivity of $S_I(t)$ and $V_I(t)$ is proved trivially. $\square$

3.3.2 Boundedness of the Solution

To prove the boundedness of the solutions we rely on the following result.

Lemma 3.3.2 (Boundedness)

The solutions $G(t), I(t), \beta(t), \gamma(t), \sigma(t), S_I(t), V_I(t)$ are bounded. $\square$

Proof. To prove the statement we need to prove separately that each solution is bounded, that is, if we define $\kappa(t) = (G(t), I(t), \beta(t), \gamma(t), \sigma(t), S_I(t), V_I(t))^T$, we need to show for each component κ_i that

$$\limsup_{t \rightarrow \infty} \kappa_i(t) = \kappa_M < \infty, \quad i = 1, \dots, 7.$$

Consider $\kappa_1(t) = G(t)$ and let's assume $\limsup G(t) \rightarrow \infty$. Then, there should exists a sequence

$$\{t_n\} : t_n \rightarrow \infty \implies G(t_n) \rightarrow \infty, \quad \dot{G}(t_n) \geq 0.$$

Consequently, we get

$$\dot{G}(t_n) = R_0 - (Eg_0 + S_I(t_n) \cdot I(t_n)) G(t_n) + \frac{W}{Vol_G} (G_{\text{prod}} - G_{\text{up}}).$$

Being R_0 and $\frac{W}{Vol_G} (G_{\text{prod}} - G_{\text{up}})$ bounded, as discussed, with $G(t_n) \rightarrow \infty$, $\dot{G}(t_n)$ would become negative, which contradicts the assumption. Moreover, $G(t)$ is nonnegative, as discussed. Hence, $G(t)$ is bounded.

Consider $\kappa_3(t) = \beta(t)$ and let's assume $\limsup \beta(t) \rightarrow \infty$. Then, there should exist a sequence

$$\{t_n\} : t_n \rightarrow \infty \implies \beta(t_n) \rightarrow \infty, \quad \dot{\beta}(t_n) \geq 0.$$

Consequently, we get

$$\dot{\beta}(t_n) = \frac{(\bar{P}_{ISR} - \bar{A}_{MG})\beta(t_n)}{\tau_\beta}$$

As $\beta(t_n) \rightarrow \infty$, the term $(\bar{P}_{ISR} - \bar{A}_{MG}) \rightarrow (\bar{P}_{\text{lim}} - \bar{A}_{\text{lim}}) < 0$ in the parameters. Consequently, with $\beta(t_n) \rightarrow \infty$, $\dot{\beta}(t_n)$ would become negative, which contradicts the assumption. Moreover, $\beta(t)$ is nonnegative, as discussed. Hence, $\beta(t)$ is bounded.

Consider $\kappa_2(t) = I(t)$ and let's assume $\limsup I(t) \rightarrow \infty$. Then, there should exists a sequence

$$\{t_n\} : t_n \rightarrow \infty \implies I(t_n) \rightarrow \infty, \quad \dot{I}(t_n) \geq 0$$

Consequently, we get

$$\dot{I}(t_n) = \beta(t_n) \frac{ISR}{V} - kI(t_n) - I_e \text{Heav}(I(t_n)).$$

Being $\beta(t_n)$ and ISR bounded, as discussed, with $I(t_n) \rightarrow \infty$, $\dot{I}(t_n)$ would become negative, which contradicts the assumption. Moreover, $I(t)$ is nonnegative, as discussed. Hence, $I(t)$ is bounded.

Consider $\kappa_5(t) = \sigma(t)$ and let's assume $\limsup \sigma(t) \rightarrow \infty$. Then, there should exists a sequence

$$\{t_n\} : t_n \rightarrow \infty \implies \sigma(t_n) \rightarrow \infty, \dot{\sigma}(t_n) \geq 0.$$

Consequently, we get

$$\dot{\sigma}(t_n) = -\frac{(\sigma(t_n) - \sigma_{ISR_{M\infty}})}{\tau_\sigma}$$

Being $\sigma_{ISR_{M\infty}}$ bounded, with $\sigma(t_n) \rightarrow \infty$, $\dot{\sigma}(t_n)$ would become negative which contradicts the assumption. Moreover, $\sigma(t)$ is nonnegative, as discussed. Hence, $\sigma(t)$ is bounded.

Consider $\kappa_6(t) = S_I(t)$ and let's assume $\limsup S_I(t) \rightarrow \infty$. Then, there should exists a sequence

$$\{t_n\} : t_n \rightarrow \infty \implies S_I(t_n) \rightarrow \infty, \dot{S}_I(t_n) \geq 0$$

Consequently, we get

$$\dot{S}_I(t_n) = -\frac{(S_I(t_n) - S_{I,\text{target}}) \left(1 - \zeta_3 \frac{V_I(t_n)^n}{k_{n,S_I}^n + V_I(t_n)^n}\right)}{\tau_{SI}}$$

The term $\left(1 - \zeta_3 \frac{V_I(t_n)^n}{k_{n,S_I}^n + V_I(t_n)^n}\right)$ is bounded, as discussed, hence, with $S_I(t_n) \rightarrow \infty$ $\dot{S}_I(t_n)$ would become negative, which contradicts the assumption. Moreover, $S_I(t)$ is nonnegative, as discussed. Hence, $S_I(t)$ is bounded.

To prove the boundedness of $V_I(t)$ it is sufficient to consider the closed form solution of the linear equation (A.1g) with u_{eq} :

$$V_I(t) = \frac{SR_{ex}u_{eq}}{k_m}k_s + \left(V_I(0) - \frac{SR_{ex}u_{eq}}{k_m}k_s\right)e^{-k_s t}.$$

It is straightforward to observe that, being u_{eq} bounded, $V_I(t)$ is bounded.

The boundedness of $\gamma(t)$ directly follows from the boundedness of G . $\square$

3.3.3 Lipschitzity Property

In this section we prove the Lipschitzity property of the system. Relying on the Lipschitzity property, it is possible to straightforwardly derive the existence and uniqueness of the solutions. To do this, we state the following result.

Theorem 3.3.1 (Lipschitz continuity)

Let x denote the vector of state and $\dot{x} = f(x, u)$ denote in compact form the system in (A.1). Then, the vector field f defining system (A.1) is Lipschitz continuous on any compact subset K of the closed invariant nonnegative orthant, $K \subset \{x \geq 0\}$. $\square$

Proof. Every right-hand side in (A.1) is $\mathcal{C}^1$ in x on $\{x \geq 0\}$. Let $J(x) = \partial f / \partial x$ denote the Jacobian; then J is continuous on $\{x \geq 0\}$. By compactness of K and continuity of J , the Extreme Value Theorem gives $L_K := \sup_{z \in K} \|J(z)\|_2 < \infty$. For any $x, y \in K$, the mean-value (integral) inequality along the line segment from y to x yields

$$\|f(x) - f(y)\|_2 \leq \left(\sup_{z \in K} \|J(z)\|_2\right) \|x - y\|_2 = L_K \|x - y\|_2,$$

so f is Lipschitz on K . If desired, one may bound $\|J(z)\|_2$ by $\sqrt{\|J(z)\|_1 \|J(z)\|_\infty}$ to obtain an explicit L_K . $\square$

Corollary 3.3.1 (Existence and uniqueness)

Fix a bounded measurable input $u(\cdot)$. For any initial condition $x_0 \in \{x \geq 0\}$ there exist $T > 0$ and a unique Carathéodory solution $x : [0, T] \rightarrow \{x \geq 0\}$ of system (A.1) with $x(0) = x_0$. $\square$

Proof. By Theorem 3.3.1, $f(\cdot, u)$ is (locally) Lipschitz in x on the closed invariant nonnegative orthant. Since $u(\cdot)$ is bounded and measurable, the map $t \mapsto f(x, u(t))$ is measurable for each fixed x and locally bounded on compact sets. Hence $f(x, u(t))$ satisfies the Carathéodory conditions (measurable in t , locally Lipschitz in x). Therefore, by the Carathéodory/Picard–Lindelöf existence–uniqueness theorem, there exist $T > 0$ and a unique solution $x(\cdot)$ on $[0, T]$ with $x(0) = x_0$. Forward invariance of $\{x \geq 0\}$ ensures $x(t) \in \{x \geq 0\}$ on $[0, T]$. $\square$

3.4 Conclusion

This chapter has illustrated some distinctive properties of the multi-scale model [1], [2] and has provided a comprehensive qualitative analysis of the model.

Interestingly, we have detailed how it is possible to encode the whole information about the training program in a very compact form, by means of a QSSA representation of the effects of the exercise and this allows to describe the steady-state operating behavior of the model and to identify the equilibrium points of the model.

Consistently with the literature evidence, suggesting the physical activity seems to act on the transient behavior of the disease progression [14], [15], [16], [17], [18], [19], [20], the model exhibits a bistability conditions as most of the diabetes progression models [3], [4], [21], associated with a normoglycemic and an hyperglycemic steady-state condition.

In the subsequent developments, we will investigate how it is possible to leverage most of these qualitative properties in the design of model-based control techniques exploiting physical activity as control input.

References

- [1] De Paola, P. F., Paglialonga, A., Palumbo, P., Keshavjee, K., Dabbene, F., and Borri, A., “The long-term effects of physical activity on blood glucose regulation: A model to unravel diabetes progression,” *IEEE Control Systems Letters*, vol. 7, pp. 2916–2921, 2023. DOI: [10.1109/LCSYS.2023.3290774](https://doi.org/10.1109/LCSYS.2023.3290774).
- [2] De Paola, P. F., Borri, A., Dabbene, F., Keshavjee, K., Palumbo, P., and Paglialonga, A., “Modeling the cumulative benefits of regular physical activity on type 2 diabetes progression,” *Computers in Biology and Medicine*, vol. 198, p. 111 194, 2025. DOI: <https://doi.org/10.1016/j.compbiomed.2025.111194>.
- [3] Topp, B., Promislow, K., Devries, G., Miura, R. M., and T Finegood, D., “A model of β -cell mass, insulin, and glucose kinetics: Pathways to diabetes,” *J. of Theoretical Biology*, vol. 206, no. 4, pp. 605–619, 2000.
- [4] Ha, J., Satin, L. S., and Sherman, A. S., “A mathematical model of the pathogenesis, prevention, and reversal of type 2 diabetes,” *Endocrinology*, vol. 157, no. 2, pp. 624–635, 2016.
- [5] Palumbo, P., Panunzi, S., De Gaetano, A., et al., “Qualitative behavior of a family of delay-differential models of the glucose-insulin system,” *Discrete and Continuous Dynamical Systems Series B*, vol. 7, no. 2, p. 399, 2007.
- [6] Mavelli, G., Borri, A., Palumbo, P., and Bersani, A. M., “On the qualitative behaviour of oscillating biochemical systems: The stochastic approach,” in *Problems in Mathematical Biophysics: A Volume in Memory of Alberto Gandolfi*, Springer, 2024, pp. 183–205.

-
- [7] Palumbo, P., Papa, F., Vanoni, M., Alberghina, L., et al., “Qualitative behavior of a coarse-grain growth model,” in *INES 2019-IEEE 23rd International Conference on Intelligent Engineering Systems, Proceedings April 2019*, IEEE, 2019, pp. 41–46.
 - [8] De Gaetano, A. et al., “Mathematical models of diabetes progression,” *American Journal of Physiology-Endocrinology and Metabolism*, vol. 295, no. 6, E1462–E1479, 2008.
 - [9] Bennett, D. and Gourley, S., “Global stability in a model of the glucose-insulin interaction with time delay,” *European Journal of Applied Mathematics*, vol. 15, no. 2, pp. 203–221, 2004.
 - [10] Saber, S. and Alalyani, A., “Stability analysis and numerical simulations of ivgtt glucose-insulin interaction models with two time delays,” *Mathematical Modelling and Analysis*, vol. 27, no. 3, pp. 383–407, 2022.
 - [11] Li, J. and Kuang, Y., “Analysis of a model of the glucose-insulin regulatory system with two delays,” *SIAM Journal on Applied Mathematics*, vol. 67, no. 3, pp. 757–776, 2007.
 - [12] De Paola, P. F., Borri, A., Paglialonga, A., Dabbene, F., and Palumbo, P., “Long-term diabetes prevention via physical activity: An output-feedback MPC approach,” *IEEE Control Systems Letters*, 2025. DOI: [10.1109/LCSYS.2025.3578031](https://doi.org/10.1109/LCSYS.2025.3578031).
 - [13] Roy, A. and Parker, R. S., “Dynamic modeling of exercise effects on plasma glucose and insulin levels,” *J. of Diabetes Science and Technology*, vol. 1, no. 3, pp. 338–347, 2007.
 - [14] Boonpor, J. et al., “Dose–response relationship between device-measured physical activity and incident type 2 diabetes: Findings from the UK biobank prospective cohort study,” *BMC medicine*, vol. 21, no. 1, p. 191, 2023.
 - [15] Diabetes Prevention Program Research Group, “Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin,” *New England journal of medicine*, vol. 346, no. 6, pp. 393–403, 2002.
 - [16] Li, G. et al., “The long-term effect of lifestyle interventions to prevent diabetes in the China Da Qing diabetes prevention study: A 20-year follow-up study,” *The Lancet*, vol. 371, no. 9626, pp. 1783–1789, 2008.
 - [17] Strain, T. et al., “Quantifying the relationship between physical activity energy expenditure and incident type 2 diabetes: A prospective cohort study of device-measured activity in 90,096 adults,” *Diabetes Care*, vol. 46, no. 6, pp. 1145–1155, 2023.
 - [18] Gregg, E. W. et al., “Association of an intensive lifestyle intervention with remission of type 2 diabetes,” *Jama*, vol. 308, no. 23, pp. 2489–2496, 2012.
 - [19] Ramachandran, A. et al., “The indian diabetes prevention programme shows that lifestyle modification and metformin prevent type 2 diabetes in asian indian subjects with impaired glucose tolerance (IDPP-1),” *Diabetologia*, vol. 49, pp. 289–297, 2006.
 - [20] Kosaka, K., Noda, M., and Kuzuya, T., “Prevention of type 2 diabetes by lifestyle intervention: A japanese trial in IGT males,” *Diabetes research and clinical practice*, vol. 67, no. 2, pp. 152–162, 2005.
 - [21] De Gaetano, A. and Hardy, T. A., “A novel fast-slow model of diabetes progression: Insights into mechanisms of response to the interventions in the diabetes prevention program,” *PLoS One*, vol. 14, no. 10, e0222833, 2019.

Chapter 4

Model Predictive Control Approaches for Long-term Glucose Control via Physical Activity

Contents

4.1	Introduction	60
4.2	Preliminary Control Approach	61
4.3	Comprehensive Control Approach	65
4.4	Conclusion	71

4.1 Introduction

This chapter¹ is devoted to show how the model presented in Chapter 2 [3], [4] can be leveraged for the design and development of model-based control techniques for long-term T2D control via physical activity management. In this regard, the lack of mathematical models suitably describing the physiological mechanisms involving the effect of physical activity on diabetes course led to an important gap in the research line involving glucose control, with the results that no model-based control techniques so far have leveraged physical activity management for glucose control and diabetes prevention in the long term [5], [6], [7], [8], [9], [10], [11]. The study presented in this chapter bridges this critical gap and proposes novel strategies for controlling T2D course via physical activity. The control framework accounts for both a state- and output-feedback design, involves a widespread compact model, merely employed for the design of the feedback law, and the model presented in Chapter 2 [3], [4], also referred to as *extended model* in the subsequent developments, serving as model for the validation of the control law. More in detail, this chapter proposes two approaches:

- 1) a preliminary state-feedback MPC approach [1], exploiting the model by Topp et al. [12], suitably modified to account for the effect of exercise based on the original formulation introduced in [3], [4]; in the following we will refer to this preliminary approach as Preliminary Control Approach.
- 2) a more structured, output-feedback MPC approach [2], designed on a modified version of the minimal model in [12], exploiting a nonlinear observer to provide a real-time estimation of the current state. The effectiveness of this approach is assessed, validating the designed control law through an extensive simulation campaign involving the closed-loop control of a population of virtual subjects simulated by means of the extended model introduced in Chapter 2 [3], [4], also accounting for additional non-idealities; in the following we will refer to this approach as Comprehensive Control Approach.

While Preliminary Control Approach serves as a proof of concept showing how physical activity can be leveraged to design control strategies [1], Comprehensive Control Approach definitely sets the basis for the development of a general, quantitatively assessed protocol

¹The contents in this chapter are based on the results published in the works by the author in [1] and [2].

suitable to be used for diabetes control via physical activity management [2].

4.2 Preliminary Control Approach

This control approach [1] represents a first conceptual attempt to design a control law leveraging exercise. This approach does not involve directly the extended model [3], [4] but it inherits its formulation of the exercise benefits to exploit a more compact model.

4.2.1 The Control Model

The model here exploited to design the control algorithm [1] is a modification to the model by [12], and describes the evolution of glucose/insulin regulation mechanism in the long period (months/years):

$$\dot{G} = R_0 - (E_{g0} + S_I I) G, \quad (4.1a)$$

$$\dot{I} = \beta \sigma \frac{G^2}{\alpha + G^2} - k I, \quad (4.1b)$$

$$\dot{\beta} = (\bar{P} - \bar{A}) \beta, \quad (4.1c)$$

$$\dot{S}_I = -c(S_I - S_{I,target}) \cdot (1 - \zeta_{si} \frac{V_l}{k_{n,si} + V_l}), \quad (4.1d)$$

$$\dot{V}_l = \frac{SR}{K_{IL6}} \cdot u - k_s V_l, \quad (4.1e)$$

where we omit time dependencies, with the state variables defined as follows:

- G is the plasma glucose concentration (mg/dl);
- I is the serum insulin concentration (μ U/ml);
- β is the beta cell mass (mg);
- S_I is the insulin sensitivity (ml/ μ U/d);
- V_l accounts for the integral effect of IL-6 released during exercise sessions ((pg/ml)·min).

It should be observed that the state variable V_l is inherited from the extended model [3],[4], and reports the effects of the physical activity on insulin sensitivity and beta-cells dynamics. In particular, we have

$$\bar{P} = P(G) \cdot \psi_1(V_l), \quad \psi_1(V_l) = \left(1 + \frac{\zeta_p \cdot V_l^2}{k_p^2 + V_l^2}\right), \quad (4.2a)$$

$$\bar{A} = A(G) \cdot \psi_2(V_l), \quad \psi_2(V_l) = \left(1 - \frac{\zeta_a \cdot V_l^2}{k_a^2 + V_l^2}\right), \quad (4.2b)$$

$$P(G) = (r_{1r} \cdot G - r_{2r} \cdot G^2), \quad (4.2c)$$

$$A(G) = (d_0 - r_{1a} \cdot G + r_{2a} \cdot G^2). \quad (4.2d)$$

Glucose/insulin, beta-cells, and insulin sensitivity dynamics are inherited from the model by Topp et al. [12]. In more detail, the beta-cell dynamics is modified with respect to the original work of [12] where ψ_1 and ψ_2 were formally set to 1; we reformulated this equation to highlight the separate contribution of physical activity on beta-cell proliferation $P(G)$ and apoptosis $A(G)$ by means of Hill functions of the V_l state variable, similarly to the formulation in the extended model [3],[4]. Besides, with respect to [12], we introduce a general positive $S_{I,target}$ in S_I dynamics, as in [3], [4], to avoid the unrealistic increase in beta-cell mass that characterizes the original model in [12]. Finally, we model the additional effect of exercise on insulin sensitivity, widely known in the literature [13], according to the formulation in [4]. Parameters of the model are listed in Table 4.1.

Parameters	Value	Unit
R_0	864	mg/dl/d
E_{g0}	1.44	1/d
σ	43.2	$\mu\text{U/ml/d}$
α	20000	mg^2/dl^2
k	432	1/d
d_0	0.06	1/d
c	0.05	1/d
r_{1r}	$0.42 \cdot 10^{-3}$	dl/mg/d
r_{2r}	$0.12 \cdot 10^{-5}$	dl/mg/d
r_{1a}	$0.42 \cdot 10^{-3}$	dl/mg/d
r_{2a}	$0.12 \cdot 10^{-5}$	dl/mg/d
ζ_p	10^{-4}	/
k_p	10^6	((pg/ml)min)
ζ_a	10^{-4}	/
k_a	10^6	((pg/ml)min)
$S_{I,target}$	0.028	ml/ $\mu\text{U/d}$
ζ_{si}	1.4	/
$k_{n,si}$	$5 \cdot 10^6$	(pg/ml)min
SR	0.045	pg/ml/min
K_{IL6}	0.004	1/min
k_s	$-\frac{\log(0.8)}{(80640)}$	1/min

Table 4.1: Parameters of the control model with values and units for Preliminary Control Approach.

The model (4.1) can be readily restated in nonlinear state-space form

$$\dot{x} = f(x, u), \quad (4.3)$$

with

$$\begin{aligned} x &= [x_1 \ x_2 \ x_3 \ x_4 \ x_5]^T \\ &= [G \ I \ \beta \ S_I \ V_l]^T \in \mathbb{R}_{\geq 0}^5. \end{aligned}$$

For what concerns the control input, we rely on the QSSA of the effect of the exercise exploiting the concept of equivalent control input u_{eq} introduced in Chapter 3. That means, instead of the PWC input:

$$u(t) = \begin{cases} 0 & t \in [kT, (k+1)T - D_k) \\ \bar{u}_k & t \in [(k+1)T - D_k, (k+1)T) \end{cases} \quad k \in \mathbb{N}, \quad (4.4)$$

where D_k is the duration of the k -th exercise session, T is the period of the training and $\bar{u}_k$ is the activity level of the k -th exercise session, exploiting the result (3.2) in Chapter 3 we represent physical activity in a compact form through u_{eq} [1]

$$u(t) = u_{eq,k} = \frac{\bar{u}_k \cdot D_k}{T}, \quad t \in [0, T] \quad (4.5)$$

which basically embeds the time average of the exercise intensity over the period T . The concept of equivalent input u_{eq} allows to represent the contribution of physical activity in a compact form, encoding all the features related to the exercise program and accounting for all the information about the amount of exercise to be performed.

In this regard, we remark that the V_l dynamics in (4.1) comes from the equations of the extended model [3], [4] when the control input u is supposed fixed to its average value u_{eq} and the fast dynamics of extended model [3], [4], describing how physical activity effects propagate on the glucose/insulin dynamics, are at their *quasi-stationary* values. Moreover, we remark that in (4.1) we assume that the dynamics of V_l is known and correctly specified. This assumption should be interpreted as an idealized, proof-of-concept modeling hypothesis adopted to enable the subsequent formulation of an ideal state-feedback MPC controller and to assess its potential performance under full-state information. In practice, the dynamics of V_l is difficult to characterize accurately, due

to inter-individual variability, limited observability, and uncertainty in the underlying biochemical pathways.

With the aim of designing a feedback law by means of an MPC controller, the role of u_{eq} provides a great advantage, as it allows for a continuous formulation of the control with respect to the original impulsive representation encoded through repeated exercise sessions as in [3],[4]. The general information of the control $u_{eq,k}$ computed at each optimization step by the optimizer, can then be translated into a duration control, by keeping fixed the exercise intensity $\bar{u}$ and the period T of the exercise sessions, by means of the inverse map

$$D_k = \frac{u_{eq,k} \cdot T}{\bar{u}}. \quad (4.6)$$

In this way, the general exercise program encoded in u_{eq} is converted into precise recommendations in terms of duration of the exercise to be performed [1].

4.2.2 Design of the Control Law

The control law is designed by means of a state-feedback MPC strategy. We define the cost function [1]:

$$J_1(x, u_{eq}) \Big|_{[t_1, t_2]} \triangleq \|x_1(\cdot)\|_{L^2([t_1, t_2])}^2 + \lambda(t_2 - t_1)|u_{eq}|^2, \quad (4.7)$$

where $[t_1, t_2]$ is the time interval of interest, x_1 is the current glycemia value, u_{eq} is the equivalent control input predicted by the MPC, assumed constant in the period of interest. The cost function in (4.7) allows to achieve the best trade-off between the reduction of current glucose levels and the feasibility of the amount of the exercise to be undertaken. In order to allow for a flexible tuning of the balance between performance and control effort, the cost function introduces a positive real penalty coefficient λ weighing the control input. When λ tends to 0, the control effort is scarcely accounted for in the optimization process, resulting in the trivial solution that the maximum allowed amount of physical activity level is recommended: $u_{eq} = u_{eq}^{\max}$, which is not realistic. On the other hand, for sufficiently large values of λ , the achieved glycemia marginally affects the cost function, which will be optimized in absence of physical exercise: $u_{eq} = 0$.

On top of the continuous-time formulation of the cost function, due to the continuous-time nature of the system (4.1) that we aim to control, we assume a sampled-data implementation of the feedback strategy, i.e. the equivalent input $u(t)$ is implemented as a PWC signal, with hold interval equal to the period of the physical activity:

$$u(t) = u(kT), \quad t \in [kT, (k+1)T), \quad k \in \mathbb{N}.$$

The motivation for the sampled-data, discrete-time, implementation is twofold: from a computational viewpoint, the optimization is performed on discrete time instants, which is much less demanding; on the other hand, since the equivalent input will be translated into a train of pulses, whose duration is modulated once per period, in agreement with the inverse map (4.6), it would be pointless to obtain a continuous-time evaluation of the equivalent input, which we may want to update (at most) at every training session.

The optimizer computes the optimal control sequence $\{u_{eq,k}^*\}_{k \in \mathbb{N}}$ such that

$$\begin{aligned} u_{eq,k}^* &= \arg \min J_1(x, u_{eq}) \Big|_{[kT, (k+N)T]} \\ \text{s.t.} \quad & (4.1), \\ & u_{eq} \in [0, 2.5], \\ & u(t) = u_{eq,k}^*, \quad t \in [kT, (k+1)T), \quad k \in \mathbb{N}, \end{aligned} \quad (4.8)$$

where T is the training period and N represents the integer length of the prediction window in terms of number of training periods considered in the optimization horizon.

The constraints account for the dynamics of the system and for the bounds on u_{eq} . More in detail, we set $u_{eq}^{\max} = 2.5$ as, when this maximal regime is translated into a duration recommendation via the inverse map (4.6) (with moderate-to-vigorous exercise, i.e. $\bar{u} = 60\%$, $T = 2$ days), it results in an overall duration of exercise around 400 minutes/week, which seems to be a saturating threshold for the effects of exercise on diabetes progression [14].

4.2.3 Results

Figure 4.1 illustrates the basal glucose concentration over time in both the open-loop (dashed line, no exercise) and closed-loop (solid line) scenarios. In the open-loop case, the system exhibits a severe progression of T2D, with glucose levels rising and the disease trajectory becoming effectively irreversible due to the failure of beta-cell compensation, primarily driven by glucotoxic effects [12]. As a consequence, the system evolves toward the hyperglycemic steady state [1], conventionally defined as $G = 600$ mg/dl in the model by Topp et al. [12]. Conversely, the closed-loop simulation shows the effectiveness of the proposed MPC strategy. When the MPC controller predicts the equivalent exercise input u_{eq} , the system is able to mitigate disease progression, delay the onset of critical glucose elevation, and, over the long term, to reverse T2D course [1]. This ultimately leads to a restoration of normoglycemic conditions, with the system stabilizing around the physiological steady state at $G = 100$ mg/dl, as defined in the Topp model [12].

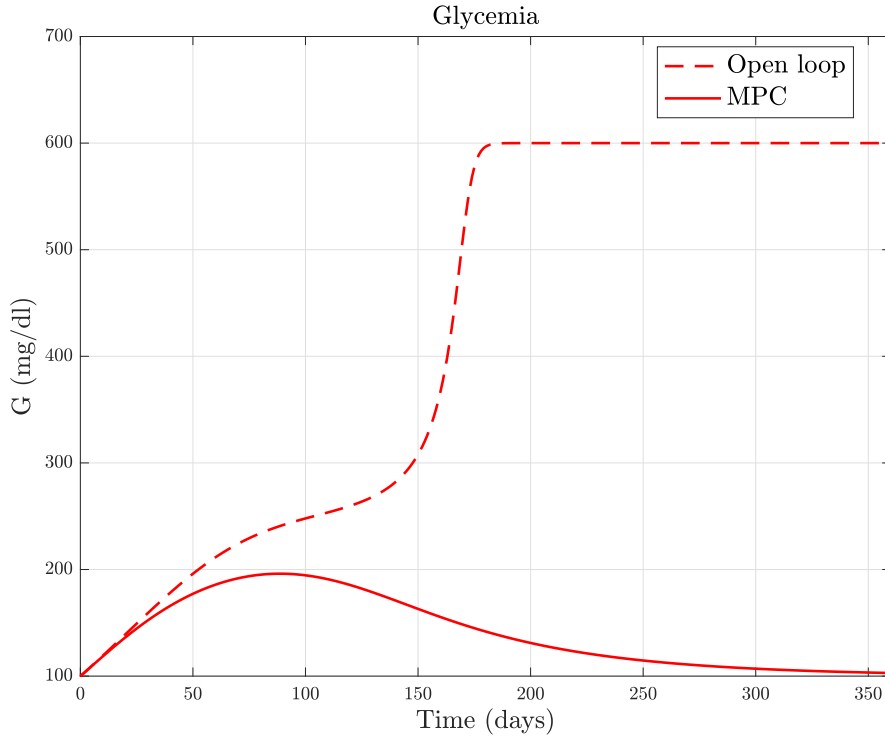


Figure 4.1: Trajectory of basal glucose concentration over time: open-loop (dashed line) versus controlled case (solid line).

Figure 4.2 illustrates how the computed control law, expressed in terms of the equivalent input u_{eq} , encoding the information about the overall exercise program, can be translated into a precise, time-varying recommendation for physical activity to be undertaken. In this regard, by fixing the exercise intensity (e.g., $\bar{u} = 50\%$, corresponding to moderate-intensity activity) and the exercise period (e.g., $T = 2$ days), the optimal control input computed via the MPC strategy is mapped, through the inverse transformation (4.6), into a specific

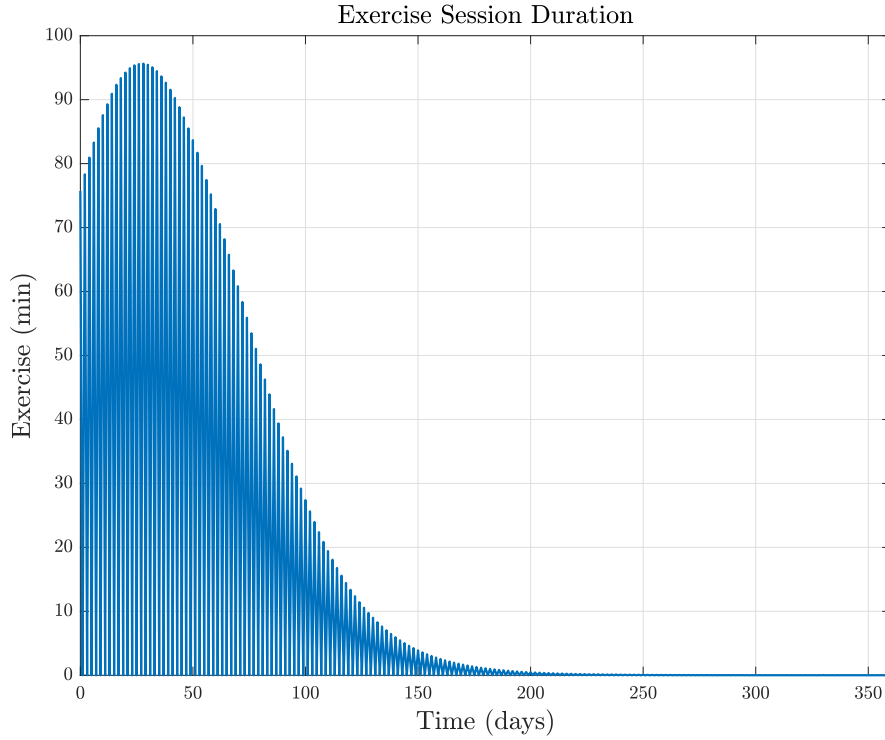


Figure 4.2: Evolution of the recommended single-session exercise duration over time, as computed through the inverse map in the MPC-controlled case ($\bar{u} = 50\%$, $T = 2$ days).

recommendation for the duration of each exercise session [1]. This formulation enables a direct and interpretable translation of abstract control inputs into actionable lifestyle guidance. As observable in the plot, during the early stages of T2D progression, the recommended total exercise duration is approximately 300 minutes/week. Over time, as glucose levels decrease under the effect of the intervention, the MPC dynamically adjusts the duration of exercise sessions, progressively reducing the required effort in accordance with the improving glycemic profile [1].

4.3 Comprehensive Control Approach

Preliminary Control Approach constitutes a first attempt to design a closed-loop control law for long-term T2D prevention and mainly serves as proof of concept. In this section we describe Comprehensive Control Approach [2], a more structured approach that overcomes the drawbacks associated with Preliminary Control Approach, mainly related to the unfeasible complete knowledge of the model, including real-time measurements of the state variables, and to the relatively short time range involved, one year, compared to the longer, more realistic time horizon involved in T2D progression.

To overcome these drawbacks, following a well-assessed technique broadly employed in glucose control [6], [8], [11], [15], the approach here proposed relies on two steps [2]:

- 1) we perform the design of the control law on a minimal model of diabetes progression;
- 2) the loop is closed validating the control law on the extended model [3], [4], accounting for both the case of *average virtual* patient (i.e., simulating the extended model with nominal parameters) and the case of a population of virtual subjects simulated by means of parameter perturbations as detailed in the followings.

The minimal model employed here for the design of the control law is different from that previously leveraged in Preliminary Control Approach, as clearly illustrated in the

followings. Moreover, Comprehensive Control approach directly involves the extended model for the validation of the control law, whereas in Preliminary Control Approach the loop was closed on the same model in (4.1).

4.3.1 Control Model

For the control design [2], we rely on the widespread model of diabetes progression [12], which, as discussed, shares with the extended model [3], [4] some basic features as basal glycemia and insulinemia, that we suitably modify here to include the effect of the exercise in a very minimal way, adopting an even simpler representation than that employed in Preliminary Control Approach. For this model the state vector is $\chi(t) = [G(t), I(t), \beta(t), S_I(t)]^T$, whose dynamics is given by

$$\dot{\chi}_1 = R_0 - (E_{g0} + \chi_4 \chi_2) \chi_1, \quad (4.9a)$$

$$\dot{\chi}_2 = \chi_3 \sigma \frac{\chi_1^2}{\alpha + \chi_1^2} - k \chi_2, \quad (4.9b)$$

$$\dot{\chi}_3 = (-d_0 + r_1 \chi_1 - r_2 \chi_1^2) \chi_3, \quad (4.9c)$$

$$\dot{\chi}_4 = -c(\chi_4 - S_{I,target}) \cdot (1 - u_v), \quad (4.9d)$$

with measured output

$$y = \Gamma \chi, \quad \Gamma = [1 \ 0 \ 0 \ 0]. \quad (4.10)$$

Differently from the modified version (4.1) of the model in [12] proposed in Preliminary Control Approach, in (4.9) we account for the effect of exercise in a compact form by means of the (normalized) virtual input function $u_v : \mathbb{R}_{\geq 0} \rightarrow [0, 1]$, which embeds the overall effect of exercise in terms of relative reduction of the insulin sensitivity decay rate, thus further simplifying the representation of the exercise effect on insulin sensitivity and omitting the additional contribution of physical activity on beta cells. This allows also the subsequent development of a control scheme that does not rely on the explicit modeling of V_l in (4.1), thereby relaxing the assumption of the exact knowledge and correctness of the V_l submodel and moving toward a more realistic control architecture. Coherently with the extended model [3], [4] and (4.1), we consider the positive steady state value $S_{I,target}$ for S_I to avoid the unrealistic unbounded increase in beta-cell mass that the original model [12] may exhibit. The parameters of the minimal model (4.9) are inherited from [12], with $S_{I,target} = 0.028$. It is worth noticing that such a value refers to a very low insulin sensitivity [12], [16], corresponding to the long-term asymptotic impairment provided by an uncontrolled diabetes progression. The role of the virtual control u_v is to delay as far as possible such a progression, since it is exerted on the rate of the insulin sensitivity dynamics [2].

4.3.2 MPC Design

In our framework, optimal physical activity strategies are decided based on the state, which is essentially unknown in a real case scenario, except for the variable G that is measured. To prove the validity of our approach, in the following we compare the case of state-feedback and output-feedback control law [2]. For the sake of clearness, we remark that the state-feedback MPC considered in this case is different with respect to that proposed in Preliminary Control Approach, as it is designed on a different, even simpler, model, and, as it will be discussed below, it involves a different design and validation setting of the control law.

4.3.2.1 State-feedback MPC

The MPC solves an optimization problem involving the cost functional [2]:

$$J_2(\chi, u_v) \Big|_{[t_1, t_2]} \triangleq \|\chi_1(\cdot) - G_d\|_{L^2([t_1, t_2])}^2 + \lambda(t_2 - t_1)|u_v - u_{v,b}|^2, \quad (4.11)$$

where $[t_1, t_2]$ is the time interval of interest, G_d is the reference glycemia value and u_v is the virtual control input predicted by the MPC, assumed constant in the period of interest, whereas $u_{v,b} \in [0, 1]$ is a feedforward virtual baseline control. The cost function $J_2(\chi, u_v)$ is designed to balance the best trade-off between glucose levels and the practical feasibility of the amount of the required exercise. We remark that the design choice of penalizing the deviation of the control u_v from a baseline level $u_{v,b}$ is motivated by both control considerations and clinical rationale. Indeed, penalizing the total exercise could lead, as in Preliminary Control Approach, the optimizer to drive the control toward minimal exercise, potentially very close to zero minutes/week as a cost-minimizing solution. Such behavior would be physiologically unrealistic and clinically undesirable [17]. Conversely, penalizing the deviation from a baseline level $u_{v,b}$ ensures that the optimizer drives the control to track the baseline amount of physical exercise, while allowing for dynamical adjustments. The optimizer computes the optimal virtual control sequence $\{u_{v,k}^*\}_{k \in \mathbb{N}}$ such that

$$\begin{aligned} u_{v,k}^* &= \arg \min J_2(\chi, u_v) \Big|_{[kT, (k+N)T]} \\ \text{s.t. } & (4.9), \\ & u_v \in [0, 1], \\ & u(t) = u_{v,k}^*, \quad t \in [kT, (k+1)T), \quad k \in \mathbb{N}, \end{aligned} \quad (4.12)$$

where T is the training period and N represents the integer length of the prediction window in terms of number of training periods considered in the optimization horizon.

4.3.2.2 Observer-based Output-feedback MPC

The optimization problem in (4.12) assumes an unrealistic complete knowledge of the full state χ and serves in this framework as ideal benchmark controller. For subsequent analysis [2], we recast it in a compact form as

$$\dot{\chi}(t) = f(\chi(t)) + g(\chi(t))u_v(t). \quad (4.13)$$

As only plasma glycemia can be directly measured, the other components of the state χ are estimated in real time through an observer. To properly handle the discrete-time character of the sampled data, these measurements are modeled as continuous outputs with a time-varying observation delay expressed as

$$y(kT) = \Gamma \chi(t - \varrho(t)), \quad (4.14)$$

with $\varrho(t) = t - kT$, $t \in [kT, (k+1)T)$, $k \in \mathbb{N}$.

In this context, we consider the observer scheme presented in [18], [19]. Its design follows the dynamics below, making use only of the sampled measurements $y(kT)$, $k = 1, 2, \dots$:

$$\begin{aligned} \dot{\hat{\chi}} &= f(\hat{\chi}(t)) + g(\hat{\chi}(t))u_v(t) \\ &\quad + e^{-\eta \varrho(t)} \cdot Q^{-1}(\hat{\chi}(t))\Lambda(y(kT) - \Gamma \hat{\chi}(kT)). \end{aligned} \quad (4.15)$$

Here, $\eta > 0$ is a design parameter that progressively reduces the weight assigned to more recent measurements, while Λ denotes the observer gain and determines the eigenvalues of the error matrix $A_b - \Lambda \Gamma$, with A_b being the Brunowski matrix,

$$A_b = \begin{bmatrix} 0_{3 \times 1} & I_3 \\ 0 & 0_{1 \times 3} \end{bmatrix}.$$

The development of the observer requires the *drift-observability* map

$$\Psi(\chi) \triangleq [\chi_1 \quad L_f \chi_1 \quad L_f^2 \chi_1 \quad L_f^3 \chi_1]^T,$$

where $L_f^k \chi_1$ denotes the k -th order Lie derivative for $k = 1, 2, 3$, to be a diffeomorphism on $\mathbb{R}_{\geq 0}^4$.

This condition implies that its Jacobian

$$Q(x) = \frac{\partial \Psi(x)}{\partial x}$$

which is exploited in (4.15), must be invertible. This property has been verified locally by finding that $Q(\cdot)$ remains invertible within the region of the state space explored by simulation. Convergence has been investigated through simulation, as shown in Fig. 4.3, that illustrates the trend of the state estimation of the observer (yellow line), the true state (plant, blue line) and the predictions of the model when simulated with wrong initial conditions (model prediction, red line). Sufficient conditions for global or semiglobal convergence of the observer are provided in [18], [19]. However, in the absence of a separation principle allowing for the independent treatment of control and state estimation, such conditions do not extend to guaranteeing the same properties for the closed-loop observer.

In the spirit of the separation principle, we reformulate the MPC problem (4.12) in an output-feedback framework by replacing the true, and generally unknown, state with its real-time estimate (4.15). Within this framework, the optimal virtual control sequence is expressed as $\{u_{v,k}^*\}_{k \in \mathbb{N}}$ such that [2]

$$\begin{aligned} \hat{u}_{v,k}^* &= \arg \min J(\hat{\chi}, u_v) \Big|_{[kT, (k+N)T]} \\ \text{s.t. } & (4.15), \\ & u_v \in [0, 1], \\ & u(t) = \hat{u}_{v,k}^*, \quad t \in [kT, (k+1)T), \quad k \in \mathbb{N}. \end{aligned} \tag{4.16}$$

The proposed formulation relies on a continuous-time dynamics along with a piecewise-constant MPC and provides some advantages that appear to perfectly align with the nature of the system under consideration and with the overall control framework, with system variables evolving in continuous time and measurements and control updates occurring in discrete time. The continuous-time formulation allows to capture and regulate the intersample behavior, while leveraging the broader and more mature continuous-time nonlinear control design methodologies and preserving the theoretical background of the observation design techniques.

In light of its formulation, the proposed MPC optimization problem is inherently both initially and recursively feasible. This is the consequence of the fact that the only constraints imposed are the system dynamics and a bounded control input $u_v \in [0, 1]$. Additionally, the absence of hard state constraints or terminal constraints prevents the possibility of feasibility loss due to unreachable state targets. Consequently, feasibility is ensured at the initial time and is preserved recursively over system evolution.

4.3.2.3 Validation of the Control Law and Simulation Setup

This section is devoted to illustrate how the loop is closed [2] validating the control law on the extended model [3], [4]. The study considers both a nominal virtual patient and a virtual population generated through random parameter perturbations, while accounting for non-idealities such as quantized measurements and disturbances.

Prior to closing the loop on the extended model [3], [4], the optimal output-feedback MPC formulation must be translated into a PWC control as in (4.4). We achieve this by representing physical activity as pulses of optimal duration D_k^* , constrained to integer multiples of $\mu_D = 5$ minutes to ensure practical feasibility. The exercise intensity is fixed at $\bar{u}_k = \bar{u} = 60\%$ for all intervals, consistent with moderate-to-vigorous activity [3], [4], [20]. The resulting optimal MPC control is interpreted as the fraction of the maximal

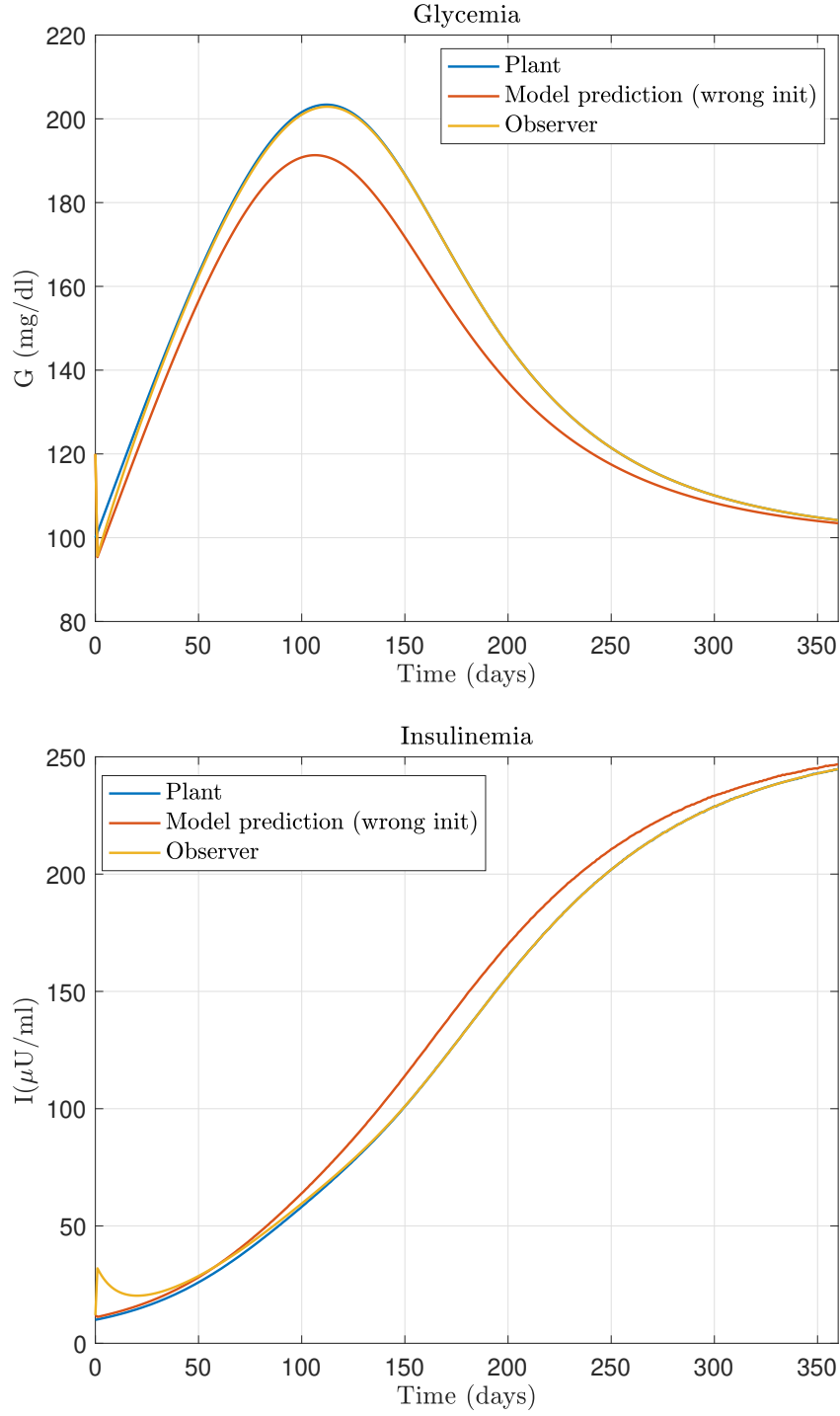


Figure 4.3: Top panel: Basal glucose concentration in the comparison between the true state (plant, blue line), the model predictions with wrong initial conditions (red line) and the observer estimation (yellow line). Bottom panel: Insulin concentration in the comparison between the true state (plant, blue line), the model predictions with wrong initial conditions (red line) and the observer estimation (yellow line).

equivalent input $\bar{u}_{eq,k}$, leading, via the mapping from the result (3.2), to

$$\frac{D_k^*}{T} \bar{u} = \hat{u}_{v,k}^* \cdot u_{eq,\max}. \quad (4.17)$$

By incorporating quantization, the pulse duration becomes

$$D_k^* = \left\lceil \frac{u_{eq,\max} T}{\bar{u}} \cdot \hat{u}_{v,k}^* \right\rceil_{\mu_D}, \quad k \in \mathbb{N}, \quad (4.18)$$

where $[\cdot]_\mu : \mathbb{R}_{\geq 0} \rightarrow \mu\mathbb{N}$ denotes the quantization operator defined by

$$[x]_\mu = \mu \left\lceil \frac{x}{\mu} \right\rceil$$

and $[\cdot]$ denotes rounding to the nearest natural number.

In the simulations, we set $T = 2$ days (physical activity period, coinciding with sampling and actuation), and $N = 10$ (MPC prediction window length), yielding a prediction horizon of 20 days. The MPC weight parameter in (4.11) is chosen as $\lambda = 150$, while $u_{eq,\max} = 2.5$, corresponding via (4.18) to a maximal exercise duration per session around $D_{\max} = 120$ minutes. We further set $G_d = 100$ mg/dl and $u_{v,b} = 0.36$. Through (4.18), the virtual control $u_{v,k}^*$ encodes the effect of physical activity as an exercise duration D_k^* ranging from 0 up to approximately 400 minutes/week, that is the level at which the benefits of physical activity are reported to saturate in clinical studies [14]. This choice reflects a deliberate design rationale, as the threshold of 400 minutes/week aligns with clinical evidence. Similarly, the baseline virtual input $u_{v,b}$ is calibrated to represent about 150 minutes/week of exercise, consistent with WHO recommendations [17].

We set the initial condition of the extended model as $x(0) = [99.7604 \ 9.025 \ 1000.423 \ -0.00666 \ 536.67 \ 0.8 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]^T$ and initialize the observer (4.15) at $\hat{x}(0) = [100 \ 4 \ 990 \ 0.72]^T$. To further account for non-idealities beyond the theoretical framework, glucose measurements are modeled as quantized, i.e. $y(kT) = [G(kT)]$, where glycemia is rounded to the nearest integer. The simulations are performed over a 20-year horizon, in line with clinical evidence and prevention studies on diabetes progression, assessing that diabetes course develops on a time horizon ranging from five years onward [14], [21], [22].

4.3.3 Results

4.3.3.1 Numerical Results on the Average Patient Model

In Fig. 4.4 (top panel), we compare the simulation outcomes of the extended model under four configurations: open loop (blue dashed line), baseline feedforward control aligned with WHO recommendations of 150 min/week (red dashed line), MPC with state feedback (yellow dotted line), and MPC with output feedback (purple solid line). The plot highlights a marked difference across the scenarios. In the open-loop case, the subject becomes diabetic from the third year of simulation, whereas under baseline WHO feedforward control (150 min/week), diabetes onset is delayed until the seventh year. In both cases, the disease trajectory becomes irreversible, and the system converges to the hyperglycemic equilibrium at $G = 600$ mg/dl over the 20-year horizon. Conversely, when MPC is applied, both state- and output-feedback, the controller is able to delay and even reverse the disease course, thereby restoring normoglycemia. Under these conditions, basal glucose concentrations remain consistently below the diabetic threshold of 126 mg/dl [2].

The bottom panel of Fig. 4.4 depicts the evolution of the control input, expressed as weekly exercise time in minutes. It can be observed that the MPC with state feedback (yellow dotted line) and the MPC with output feedback (purple solid line) result in comparable exercise prescriptions throughout the 20-year simulation. The state-feedback strategy, however, produces a marginally higher average exercise level, which translates into lower glucose concentrations with respect to the output-feedback case, as shown in the top panel of Fig. 4.4 [2]. The MPC state-feedback strategy yields a lower overall cost than the output-feedback strategy ($J_{sf} = 1.47 \cdot 10^6$ vs. $J_{of} = 1.66 \cdot 10^6$), as reasonably expected.

Importantly, these results are consistent with the WHO guidelines, which recommend

a minimum of 150 minutes of moderate physical activity per week for the prevention of chronic diseases and explicitly highlight that additional benefits can be achieved by exceeding this threshold [17].

As observable from Fig. 4.5, the output-feedback controller is able to guarantee that glucose concentration remains within safe and acceptable bounds, even under random perturbations simulating circadian variability of basal glycemia of up to $\pm 10\%$, in line with the studies in [23], [24].

4.3.3.2 Numerical Results with Validation on a Population of Virtual Subjects

To further assess the performance of the output-feedback controller [2], we close the loop on a population of 400 virtual patients generated by perturbing the nominal parameters of the extended model [3], [4]. The objective is to assess, over a 20-year horizon, how many individuals experience a reversal of diabetes progression. Specifically, 200 virtual patients are simulated using independent log-normal perturbations with a standard deviation of 5% relative to nominal values, while the remaining 200 are generated using independent uniform perturbations within $\pm 5\%$ of the nominal values, as in [25]. Similar ranges of parameter variability have been considered in previous works [9], [11]. A summary of the results across the full population is reported in Table 4.2, where the columns represent the percentage of subjects reversing diabetes, their average basal glycemia at year 20, and the average weekly physical activity prescribed by the controller. The outcomes are comparable across both parameter distributions and indicate a promising success rate.

On average, the recommended exercise duration aligns with established guidelines for

	% of success	G (mg/dl)	Exercise (min/week)
U	75	106 (4.3)	201 (39.3)
Log	71	104 (3.7)	199 (38.5)

Table 4.2: Validation results for the cohort of 400 virtual subjects.

diabetes prevention. Adjusting the MPC penalty parameter λ in (4.11) may increase the proportion of successful cases; however, this comes at the cost of prescribing greater amounts of weekly physical activity. Fig. 4.6 reports scatter plots of basal glucose and insulin concentrations obtained from the 200 virtual subjects with log-normal parameter perturbations. These plots provide insights into the dynamics of disease progression, indicating that the course of diabetes, either progression or reversal through sustained physical activity, appears to be definitely set after the first 10 years.

4.4 Conclusion

In this chapter we have proposed a novel control approach for long-term diabetes progression via physical activity. Results are consistent with the literature. The philosophy of this approach is to complement public health guidelines, such as those from the WHO, by providing a quantitative assessment, that suggests the **minimal level** of physical activity that could be sufficient to prevent or slow the progression of T2D within the modeled physiological dynamics. Specifically, our control approach is aimed at providing a personalized lower bound on the required physical activity level, aimed at identifying the minimum amount of exercise necessary to prevent disease progression in a given individual under a given physiological profile. Indeed, while performing less than this level (e.g., according to the WHO suggested exercise) may be insufficient to prevent diabetes according to the individual conditions, performing more could be often beneficial (“more is better”) but may be unfeasible for some class of patients. Thus, the purpose is to ensure that the minimal dose of exercise sufficient to prevent diabetes is delivered, *one that is both clinically necessary and individually achievable*. These personalized recommendations should be viewed as a dynamic translation of public health guidance into individualized, clinically grounded advice. By enabling a personalized and adaptive recommendation,

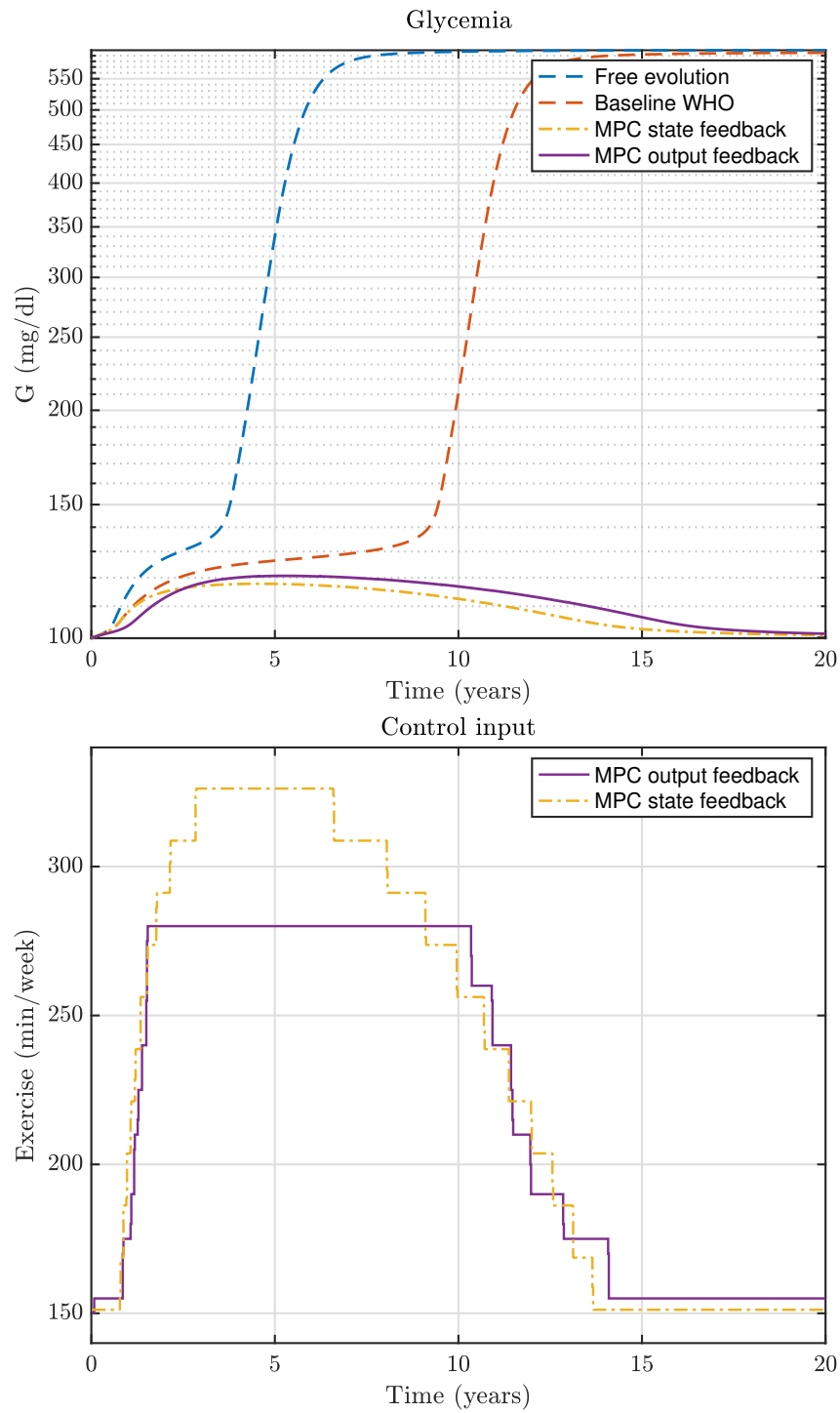


Figure 4.4: Top panel: trend of the basal glucose concentration over the 20-year horizon. Bottom panel: trend of the control input in terms of recommended exercise per week.

the proposed method offers a way to translate general guidelines into actionable and individualized, quantitatively assessed, targets, particularly in the context of digital health tools and patient self-management.

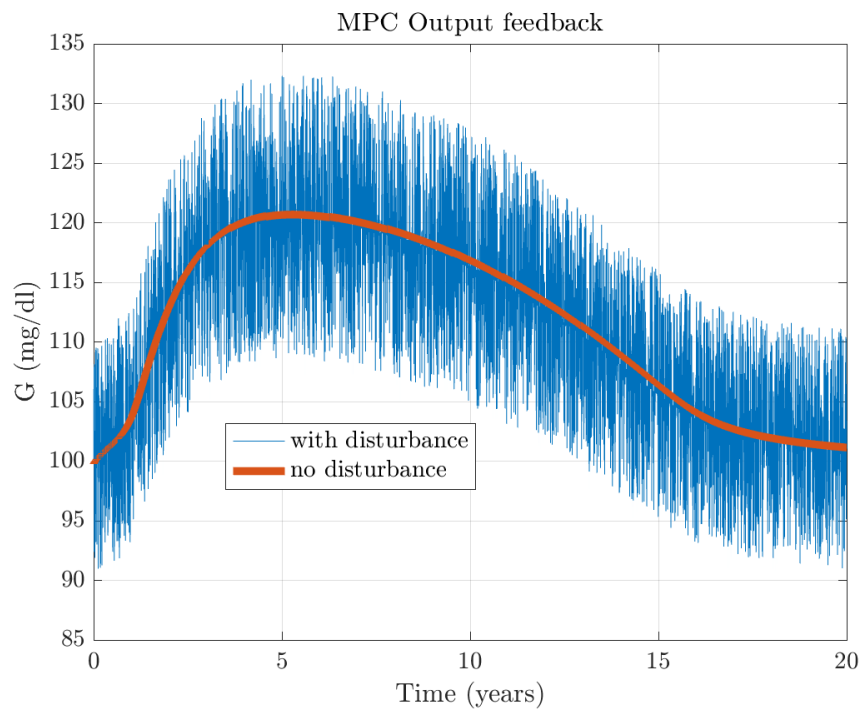


Figure 4.5: Comparison between the basal glucose concentration in the undisturbed case (red line) and when a random disturbance is introduced.

References

- [1] De Paola, P. F., Borri, A., Paglialonga, A., Palumbo, P., and Dabbene, F., “A model-based approach for glucose control via physical activity,” in *Studies in health technology and informatics*, vol. 327, 2025, pp. 27–31. DOI: [10.3233/SHTI250267](https://doi.org/10.3233/SHTI250267).
- [2] De Paola, P. F., Borri, A., Paglialonga, A., Dabbene, F., and Palumbo, P., “Long-term diabetes prevention via physical activity: An output-feedback MPC approach,” *IEEE Control Systems Letters*, 2025. DOI: [10.1109/LCSYS.2025.3578031](https://doi.org/10.1109/LCSYS.2025.3578031).
- [3] De Paola, P. F., Paglialonga, A., Palumbo, P., Keshavjee, K., Dabbene, F., and Borri, A., “The long-term effects of physical activity on blood glucose regulation: A model to unravel diabetes progression,” *IEEE Control Systems Letters*, vol. 7, pp. 2916–2921, 2023. DOI: [10.1109/LCSYS.2023.3290774](https://doi.org/10.1109/LCSYS.2023.3290774).
- [4] De Paola, P. F., Borri, A., Dabbene, F., Keshavjee, K., Palumbo, P., and Paglialonga, A., “Modeling the cumulative benefits of regular physical activity on type 2 diabetes progression,” *Computers in Biology and Medicine*, vol. 198, p. 111 194, 2025. DOI: <https://doi.org/10.1016/j.compbiomed.2025.111194>.
- [5] Incremona, G. P., Messori, M., Toffanin, C., Cobelli, C., and Magni, L., “Model predictive control with integral action for artificial pancreas,” *Control engineering practice*, vol. 77, pp. 86–94, 2018.
- [6] Messori, M., Incremona, G. P., Cobelli, C., and Magni, L., “Individualized model predictive control for the artificial pancreas: In silico evaluation of closed-loop glucose control,” *IEEE Control Systems Magazine*, vol. 38, no. 1, pp. 86–104, 2018.
- [7] Borri, A., Palumbo, P., Manes, C., Panunzi, S., De Gaetano, A., et al., “Sampled-data observer-based glucose control for the artificial pancreas,” *Acta Polytechnica Hungarica*, vol. 14, no. 1, pp. 79–94, 2017.

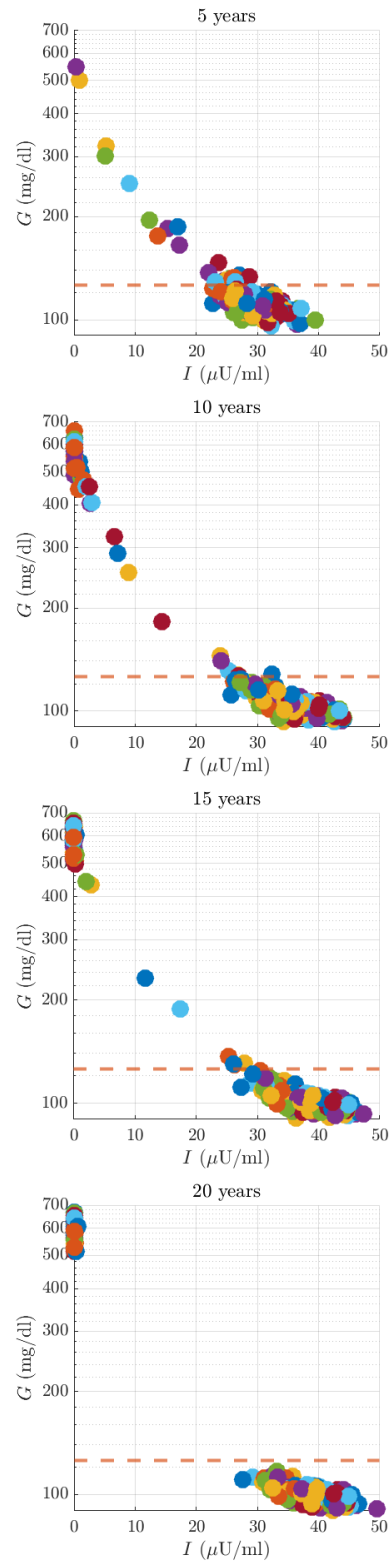


Figure 4.6: Scatter plots of G and I for the 200 virtual subjects simulated by means of log-normal parameter perturbations at 5,10,15,20 years. The red dashed line represents the diabetic threshold (126 mg/dl).

-
- [8] Di Ferdinando, M., Pepe, P., Palumbo, P., Panunzi, S., and De Gaetano, A., “Semiglobal sampled-data dynamic output feedback controller for the glucose–insulin system,” *IEEE Transactions on Control Systems Technology*, vol. 28, no. 1, pp. 16–32, 2018.
 - [9] Di Ferdinando, M., Pepe, P., Di Gennaro, S., Borri, A., and Palumbo, P., “Quantized sampled-data static output feedback control of the glucose–insulin system,” *Control Engineering Practice*, vol. 112, p. 104828, 2021.
 - [10] Kovács, L., Szalay, P., Almássy, Z., and Barkai, L., “Applicability results of a nonlinear model-based robust blood glucose control algorithm,” *Journal of diabetes science and technology*, vol. 7, no. 3, pp. 708–716, 2013.
 - [11] Borri, A., Pola, G., Pepe, P., Di Benedetto, M. D., and Palumbo, P., “Symbolic control design of an artificial pancreas for type-2 diabetes,” *IEEE Transactions on Control Systems Technology*, vol. 30, no. 5, pp. 2131–2146, 2021.
 - [12] Topp, B., Promislow, K., Devries, G., Miura, R. M., and T Finegood, D., “A model of β -cell mass, insulin, and glucose kinetics: Pathways to diabetes,” *J. of Theoretical Biology*, vol. 206, no. 4, pp. 605–619, 2000.
 - [13] Bird, S. R. and Hawley, J. A., “Update on the effects of physical activity on insulin sensitivity in humans,” *BMJ open sport & exercise medicine*, vol. 2, no. 1, e000143, 2017.
 - [14] Boonpor, J. et al., “Dose–response relationship between device-measured physical activity and incident type 2 diabetes: Findings from the UK biobank prospective cohort study,” *BMC medicine*, vol. 21, no. 1, p. 191, 2023.
 - [15] Palumbo, P., Pizzichelli, G., Panunzi, S., Pepe, P., and De Gaetano, A., “Model-based control of plasma glycemia: Tests on populations of virtual patients,” *Mathematical biosciences*, vol. 257, pp. 2–10, 2014.
 - [16] Ha, J., Satin, L. S., and Sherman, A. S., “A mathematical model of the pathogenesis, prevention, and reversal of type 2 diabetes,” *Endocrinology*, vol. 157, no. 2, pp. 624–635, 2016.
 - [17] Bull, F. C. et al., “World Health Organization 2020 guidelines on physical activity and sedentary behaviour,” *British journal of sports medicine*, vol. 54, no. 24, pp. 1451–1462, 2020.
 - [18] Cacace, F., Germani, A., and Manes, C., “An observer for a class of nonlinear systems with time varying observation delay,” *Systems & Control Letters*, vol. 59, no. 5, pp. 305–312, 2010.
 - [19] Cacace, F., Germani, A., and Manes, C., “A chain observer for nonlinear systems with multiple time-varying measurement delays,” *SIAM J. Control Optim.*, vol. 52, no. 3, pp. 1862–1885, 2014.
 - [20] Roy, A. and Parker, R. S., “Dynamic modeling of exercise effects on plasma glucose and insulin levels,” *J. of Diabetes Science and Technology*, vol. 1, no. 3, pp. 338–347, 2007.
 - [21] Diabetes Prevention Program Research Group, “Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin,” *New England journal of medicine*, vol. 346, no. 6, pp. 393–403, 2002.
 - [22] Li, G. et al., “The long-term effect of lifestyle interventions to prevent diabetes in the China Da Qing diabetes prevention study: A 20-year follow-up study,” *The Lancet*, vol. 371, no. 9626, pp. 1783–1789, 2008.
 - [23] De Gaetano, A., Sakulrang, S., Borri, A., Pitocco, D., Sungnul, S., and Moore, E. J., “Modeling continuous glucose monitoring with fractional differential equations subject to shocks,” *J. of Theoretical Biology*, vol. 526, p. 110 776, 2021.

- [24] Panunzi, S., Borri, A., D'Orsi, L., and De Gaetano, A., "Order estimation for a fractional brownian motion model of glucose control," *Communications in Nonlinear Science and Numerical Simulation*, vol. 127, p. 107554, 2023.
- [25] Borri, A., Cacace, F., De Gaetano, A., Pompa, M., and Panunzi, S., "Optimal periodic impulsive strategies in glycemic control," *IEEE Trans. on Contr. Syst. Technology*, vol. 32, no. 6, pp. 2062–2074, 2024.

Part 2:

Counterfactuals for Dynamical Systems

Chapter 5

A General Overview About Counterfactuals

Contents

5.1	General Context	78
5.2	Beyond the Philosophical Framework	78
5.3	A Control System Framework for Counterfactuals	80

5.1 General Context

Counterfactuals are concepts inherently grounded in the field of philosophy and logic. They represent hypothetical “what-if” statements that consider alternatives to reality. For example, conditionals like: “what if Martin Luther King had died in 1958?”, or, “what if the Americas had never been colonized?” [1] are examples of counterfactual thinking. In their inherent essence, counterfactuals are endowed with the distinctive ideas of *causation* and of *closeness* or *similarity* between possible situations [2]. From a theoretical perspective, *causation* is associated with the strict causality relation embedded in a given conditional statement. For instance, considering the sentence

- if it had been the case of event “*c*”, than it would been the case of event “*e*”

the conditional assigns the cause-effect principle, in the sense that in case if “*c*” had not occurred, “*c*” would not have occurred [3]. The “*c*” condition is often referred to as *antecedent*.

Indeed, the question of *closeness* or *similarity* is more challenging. Considering the previous sentence and considering the *actual world* to be the *no-c world*, that is the world in which event “*c*” does not occur, the *closest* counterfactual is associated with the world that *minimally* differs from the *actual world*. That is, among all the possible “*c-worlds*”, it is exactly like the *actual world* except for the fact that “*c*” occurs [2], [3]. Despite their hypothetical framework, counterfactuals play a foundational role across many disciplines, such as philosophy and psychology to history, AI, and beyond.

Indeed, while the main theoretical background of counterfactuals derives from philosophy, the practice of counterfactual reasoning proves to be notable also in engineering and scientific domains. Engineers often engage in implicit counterfactual thinking through “what-if analysis” and scenario planning, for instance, during design and processes they are involved in reasoning as: “What if this component fails? What if the environmental conditions were extreme?”. Such questions are essentially counterfactual-driven questions, envisioning alternative scenarios to test system resilience or account for worst-case scenarios. Counterfactual reasoning provides a valuable theoretical setting to explore alternate outcomes and infer causal relationships by comparing what did happen with what could have happened under different conditions [1].

5.2 Beyond the Philosophical Framework

Beyond the aforementioned practical and often implicit use of counterfactual reasoning, the field of science and engineering where counterfactuals have been most rigorously

formalized within a methodological framework is AI [4]. In recent years, counterfactuals have emerged as a central tool for interpretable and trustworthy AI [4], particularly in the context of ML models whose internal mechanisms are opaque to human users. Indeed, one of the most important challenge for AI is interpretability, inherently related to the fact that many AI systems are based on black-box models that overlook the capability of the human of interpreting the decision.

The adoption of counterfactuals in the context of AI is aimed at providing new perspectives to the interpretability of ML models [4]. In this regard, *counterfactual explanations* suggest what should be different in the input variables to vary the output of the ML model [4], [5], [6]. We report here the formal definition of *counterfactual explanation* provided in [4] and [7].

Definition 5.2.1 (Counterfactual Explanation)

Given a classifier b that outputs the decision $y = b(x)$ for an instance x , a counterfactual explanation consists of an instance x' such that the decision for b on x' is different from y , i.e., $b(x') \neq y$, and such that the distance between x and x' is minimal. $\square$

As usual in practice, we will often refer to the original instance (the reference, actual, one) x as factual. The analogy with *causality* and *closeness* of the classical counterfactual concept is straightforward. Different methods have been proposed for counterfactual retrieval in the AI context. The most common methods available in the literature are outlined in the followings, each reflecting different assumptions about the underlying data and models:

- 1) *Optimization-based approaches* pose the search for counterfactuals as a constrained optimization problem, minimizing predefined cost function between the factual input and a perturbed version that changes the model prediction [5].
- 2) *Heuristic search methods* rely on iterative procedures to explore the input space, guided by feasibility criteria or proximity metrics, which make them computationally lighter though sometimes less precise [8].
- 3) *Instance-based techniques* exploit the availability of training data, for example by identifying nearest neighbors that yield a different outcome, or by adapting real examples to serve as plausible counterfactuals [9].
- 4) *Model-specific methods*, those based on decision trees or rule lists, generate counterfactuals by exploiting the interpretable structure of the predictive model, which allows tracing paths leading to alternative predictions [4].

For the sake of an example, a graphical representation of possible factual-counterfactual pairs in a classification problem in the AI framework is provided in Fig. 5.1, where A, B, C, D represent possible counterfactual instances defined with respect to different metrics. Counterfactuals have been widely applied to classification problems in various applications, for example in health prediction and prevention where different classes are typically associated with the risk of developing a given disease, or in safety-critical applications, with classes representing safe and unsafe sets. In this regard, we refer to the works in [10], [11], [12], where mainly heuristic search techniques are exploited to generate counterfactuals on data from a dataset with records of individuals at risk of developing T2D. It is important to emphasize that, in the AI context, counterfactuals are inherently *data-based*. ML models operate on datasets, and counterfactual reasoning is therefore framed in terms of input-output relations encoded by the model. Consequently, the search for counterfactuals consists in finding alternative data points (often close in some metric sense to the factual one) that would lead the model to yield a different prediction. This data-centric perspective distinguishes counterfactual reasoning in AI from more abstract philosophical treatments, grounding it firmly in the novel methods for assessing interpretability in ML models. AI-driven counterfactual explanations provide significant advantages, as they can leverage the vast amounts of data available nowadays and exploit the powerful modern ML models which allows the identification of subtle

and complex patterns that would be inaccessible by analytical modeling alone. This allows to generate counterfactuals even in highly nonlinear and high-dimensional settings where classical approaches would fail. Moreover, their ability to provide *local, instance-specific explanations* makes them attractive for applications where interpretability and user interaction are key.

Nevertheless, counterfactuals explanations exhibits a *crucial* limitation. Being inherently data-driven these approaches exploit associations between features and remain agnostic to the underlying mechanisms, thus neglecting the governing dynamics of the system under consideration [2], [4], [13]. As a result, the counterfactuals usually proposed may be essentially *conceptual* in nature, suggesting input modifications that are consistent with the data and the ML model, yet they may not be physically plausible or practically actionable in the real world. This gap between data-driven counterfactuals and the physics of dynamical systems highlights the need for complementary approaches, grounded in system and control theory, that can ensure both interpretability and feasibility of the generated counterfactuals.

5.3 A Control System Framework for Counterfactuals

To address the limitations of the data-driven approaches, we propose a control system theoretical framework for counterfactuals, with the aim of assessing a physics-informed approach suitable to account for the underlying mechanisms driving the change of class. The design of the theoretical setting is performed in analogy with the spirit of the classical interpretation of counterfactuals, while grounded in the context of dynamical systems theory.

Specifically, we address this problem by means of two optimization-based approaches, aimed at providing two alternative control system approaches for counterfactual definition:

- *optimal control based formulation*: here, the concept of counterfactual in the control system framework is introduced by means of an optimal control problem, aimed at computing the minimum control law steering a given initial condition from a given subset of the state space (e.g. the unsafe set) to another one (e.g. the safe set). Thus, the what-if analogy is straightforward: “If the system had been subject to a *minimal* optimal control input, the final state would have been safe.” This proposal will be discussed in Chapter 6.
- *optimal transport based formulation*: this formulation is more challenging from a theoretical perspective and defines the counterfactual by solving an optimal transport problem, aimed at identifying the *closest* initial condition (with respect to a reference unsafe one) that lets the system evolve safely. The counterfactual reasoning is faithfully represented: “If the system had started from a slightly different set of initial conditions, the resulting system evolution would have been safe.” This proposal will be discussed in Chapter 7.

Indeed, the control-related literature proposes very few approaches of counterfactual reasoning frameworks, mainly in the context of safety verification and fault diagnosis, typically building on explanatory or diagnostic purposes [14], [15], [16]. The methods proposed in these works [14], [15], [16] mainly operate retrospectively, evaluating slight *post-hoc* modifications to observed behaviors or control policies under fixed dynamics to assess whether a failure could have been avoided. While providing useful insights into causal attribution and policy refinement, they do not formulate counterfactuals based on full-fledged system-theoretic, physics constrained, optimization problems. Rather than perturbing known trajectories or behaviors *post-hoc*, in our framework counterfactuals are defined constructively as solutions to *forward-looking*, synthesis-oriented, optimization problems.

It should be emphasized that both the optimal control based and optimal transport based formulations proposed in our setting encode the distinctive features of the classical

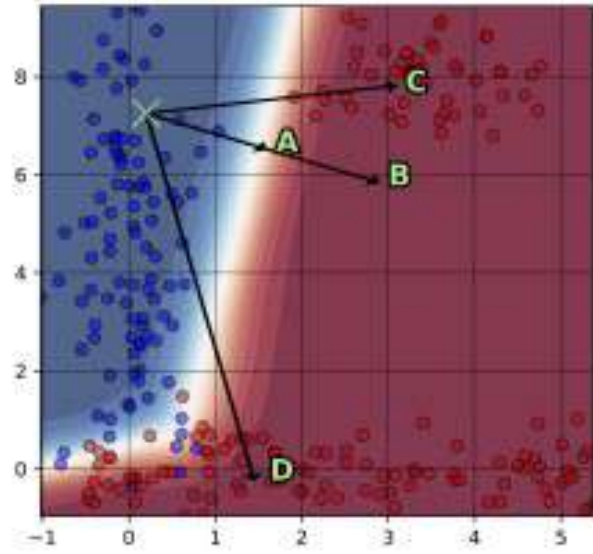


Figure 5.1: Graphical representation of possible factual-counterfactual pairs in a classification problem. The picture is from [8].

counterfactual concept in the sense of both *causality* and *closeness*. Indeed, in our control-theoretic setting, *causality* is enforced by the system dynamics representing the underlying mechanisms, while *closeness* is ensured by the optimization-based formulation, which allows to formalize the philosophical notion of “nearest world” by quantifying and minimizing deviation from the factual (the reference, actual, scenario).

For optimization purposes, both approaches rely on the tools of occupation measures and *Moment - Sum of Squares* (moment-SOS) hierarchy, which show undisputable advantages, most notably their applicability to general nonlinear systems and their inherent capability of handling both deterministic and stochastic settings. These tools allow to cast the original optimization program as an infinite-dimensional problem in the space of Borel measures and subsequently solve them by means of the moment-SOS hierarchy through a sequence of convex relations in the space of the moments [17], [18], [19], [20]. Moreover, we highlight that, while the optimal control problem formulation could, in principle, be solved using other classical techniques such as NMPC [21] or Pontryagin’s Maximum Principle [22], these tools are not suitable to address the optimal transport based formulation due to its structural complexity. Conversely, the occupation measure framework allows us to treat optimal control and optimal transport within a common measure-theoretic, convex-based setting, making occupation measures and moment-SOS relaxations a powerful and versatile tool for both problems. The mathematical background about the measure setting, occupation measures and moment-SOS hierarchy is provided in Appendix B.

Within the context of biomedical systems and life science applications, introducing the concept of counterfactuals in a control system framework paves the way to the integration of control and AI, with the aim of feeding ML models with additional system knowledge acquired by means of the proposed control methodology, with possible applications, for the sake of an example, to disease risk control.

References

- [1] Byrne, R. M., *The rational imagination*, 2005.
- [2] Kment, B., “Counterfactuals and explanation,” *Mind*, vol. 115, no. 458, pp. 261–310, 2006.
- [3] Lewis, D., *Counterfactuals*. John Wiley & Sons, 2013.

- [4] Guidotti, R., “Counterfactual explanations and how to find them: Literature review and benchmarking,” *Data Mining and Knowledge Discovery*, vol. 38, no. 5, pp. 2770–2824, 2024.
- [5] Wachter, S., Mittelstadt, B., and Russell, C., “Counterfactual explanations without opening the black box: Automated decisions and the GDPR,” *Harv. JL & Tech.*, vol. 31, p. 841, 2017.
- [6] Lucic, A., Ter Hoeve, M. A., Tolomei, G., De Rijke, M., and Silvestri, F., “Cf-gnnexplainer: Counterfactual explanations for graph neural networks,” in *International Conference on Artificial Intelligence and Statistics*, PMLR, 2022, pp. 4499–4511.
- [7] Molnar, C., *Interpretable machine learning*. Lulu. com, 2020.
- [8] Poyiadzi, R., Sokol, K., Santos-Rodriguez, R., De Bie, T., and Flach, P., “Face: Feasible and actionable counterfactual explanations,” in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 2020, pp. 344–350.
- [9] Keane, M. T. and Smyth, B., “Good counterfactuals and where to find them: A case-based technique for generating counterfactuals for explainable AI (XAI),” in *International Conference on Case-Based Reasoning*, Springer, 2020, pp. 163–178.
- [10] Lenatti, M., Carlevaro, A., Guergachi, A., Keshavjee, K., Mongelli, M., and Paglialonga, A., “A novel method to derive personalized minimum viable recommendations for type 2 diabetes prevention based on counterfactual explanations,” *Plos one*, vol. 17, no. 11, e0272825, 2022.
- [11] Carlevaro, A., Lenatti, M., Paglialonga, A., and Mongelli, M., “Counterfactual building and evaluation via explainable support vector data description,” *IEEE Access*, vol. 10, pp. 60 849–60 861, 2022.
- [12] Carlevaro, A., Lenatti, M., Paglialonga, A., and Mongelli, M., “Multi-class counterfactual explanations using support vector data description,” *IEEE Transactions on Artificial Intelligence*, 2023.
- [13] Stepin, I., Alonso, J. M., Catala, A., and Pereira-Fariña, M., “A survey of contrastive and counterfactual explanation generation methods for explainable artificial intelligence,” *IEEE Access*, vol. 9, pp. 11 974–12 001, 2021.
- [14] Gjørsum, V. B., Strümke, I., Lekkas, A. M., and Miller, T., “Real-time counterfactual explanations for robotic systems with multiple continuous outputs,” *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 7–12, 2023.
- [15] Parreaux, J., Piribauer, J., and Baier, C., “Counterfactual causality for reachability and safety based on distance functions,” *arXiv preprint arXiv:2308.11385*, 2023.
- [16] Han, H., Taheri, M., Chung, S.-J., and Hadaegh, F. Y., “A counterfactual reasoning framework for fault diagnosis in robot perception systems,” *arXiv preprint arXiv:2509.18460*, 2025.
- [17] Henrion, D., Lasserre, J. B., and Savorgnan, C., “Nonlinear optimal control synthesis via occupation measures,” in *2008 47th IEEE Conference on Decision and Control*, IEEE, 2008, pp. 4749–4754.
- [18] Lasserre, J. B., Henrion, D., Prieur, C., and Trélat, E., “Nonlinear optimal control via occupation measures and lmi-relaxations,” *SIAM journal on control and optimization*, vol. 47, no. 4, pp. 1643–1666, 2008.
- [19] Henrion, D., Korda, M., and Lasserre, J. B., *Moment-sos Hierarchy, The: Lectures In Probability, Statistics, Computational Geometry, Control And Nonlinear Pdes*. World Scientific, 2020, vol. 4.
- [20] Lasserre, J. B., *Moments, positive polynomials and their applications*. World Scientific, 2009, vol. 1.
- [21] Allgöwer, F. and Zheng, A., *Nonlinear model predictive control*. Birkhäuser, 2012, vol. 26.

- [22] Pontryagin, L. S., *Mathematical theory of optimal processes*. Routledge, 2018.

Chapter 6

An Optimal Control Based Formulation for Counterfactuals via Occupation Measures

Contents

6.1	Introduction	84
6.2	Fully Parametrized Systems	84
6.3	Uncertain Systems	94
6.4	Conclusion	99

6.1 Introduction

As discussed in the previous chapter, this chapter¹ is devoted to the introduction of the *optimal control based* formulation for counterfactuals definition in the context of dynamical systems.

The scenario we address can be framed as a safety problem: starting from an initial condition assumed to be an unsafe condition (the factual), we ask whether it would be possible to identify a point in the safe set (the counterfactual) reachable with *minimum control effort*, with free terminal time. This problem is naturally cast within the framework of optimal control.

The analysis we propose unfolds across two levels of abstraction. First, we define and solve the optimization problem for the simple case of fully parametrized systems, that is, assuming a perfectly known underlying dynamics.

Then, we generalize the discussion and show also how the proposed approach can be extended to account for uncertainties on the knowledge of the parameters of the system under consideration. Biomedical systems and, in general, systems in life science applications are highly nonlinear and prone to uncertainties that can turn out to be critical in many real-world applications. Hence, defining a general “robust” framework for control-driven counterfactuals is crucial for the reliability of the proposed approach. In both cases, the optimization problem is then here cast as an infinite-dimensional problem in the space of measures and subsequently solved by means of the moment-SOS hierarchy through a sequence of convex relaxations in the space of the moments [3], [4], [5], [6], thus allowing to design a general methodology tailored on nonlinear systems applications. This proposal establishes a first bridge between the philosophical notion of counterfactual reasoning [7], [8], [9] and the mathematical framework of control theory.

6.2 Fully Parametrized Systems

6.2.1 Problem Statement

Consider a general optimal control problem with free terminal time in the form:

$$P^* = \min_{u, \tau \in [0, T]} J(x, u) = \Phi(x(\tau)) + \int_0^\tau \phi(x(t), u(t)) dt \quad (6.1a)$$

¹The contents in this chapter are based on an extended version of the results published in the work by the author in [1], [2]

$$\text{s.t. } \dot{x}(t) = f(x(t), u(t)), \quad t \in [0, \tau], \quad (6.1b)$$

$$x(0) \in X_0, \quad x(t) \in X, \quad t \in [0, \tau], \quad (6.1c)$$

$$x(\tau) \in X_T, \quad (6.1d)$$

$$u(t) \in U, \quad t \in [0, \tau], \quad (6.1e)$$

where $x(t) \in \mathbb{R}^n$ represents the state of the system at time t , $u(t) \in \mathbb{R}^m$ represents the control input, and $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ is the vector field describing the system dynamics, τ is the free terminal time, $\Phi(x(\tau))$ is the terminal cost, a function of the final state $x(\tau)$, $\phi(x(t), u(t))$ is the running cost, a function of the state $x(t)$ and of the control $u(t)$. X_0 , X , X_T represent respectively the set of initial conditions, the set of the trajectories and the set of terminal states, T is the time horizon.

Within the framework of the optimal control problem outlined above, in this study we define the concepts of factual and counterfactual (OC-factual and OC-counterfactual) as follows [2].

Definition 6.2.1 (OC-factual)

Consider a general optimal control problem of the form (6.1a)-(6.1e). We define the OC-factual as

$$x_f := x(0) \in X_0.$$

□

Definition 6.2.2 (OC-counterfactual)

Consider a general optimal control problem (6.1a)-(6.1e) and the solution $u^*(\cdot) = \arg \min_{u(\cdot), \tau} J(x(t), u(t))$. Given a OC-factual $x_f \in X_0$, we define the OC-counterfactual x_{cf} associated with the factual $x_f \in X_0$ as

$$x_{cf} := x(\tau \mid x_f, u^*), \quad x_{cf} \in X_T,$$

where $x(t \mid x_0, u)$ denotes the state reached at time t from x_0 with input function u . □

It can be observed that, by definition, the OC-counterfactual encodes the concept of “minimum effort” for a vector field to access the terminal set X_T . For the purpose of this study we will consider the following assumption.

Assumption 6.2.1

The set $X_0 \subseteq X_U$ and the set $X_T \subseteq X_S$ where X_U and X_S are the unsafe and the safe set for the system, respectively. □

From now on the terms “factual” and “counterfactual”, when clear from the context, will specifically refer to the OC-factual and OC-counterfactual. The following part of the work illustrates how it is possible to extract counterfactuals by leveraging occupation measures and moment-SOS hierarchy.

6.2.2 Problem Solution in Measure Space

Optimal control problems of the form (6.1a)-(6.1e) can be bounded by infinite-dimensional Linear Program (LP) in measures as the following [3], [4], [5]:

$$p^* = \inf_{\mu, \mu_\tau} \langle \phi, \mu \rangle + \langle \Phi, \mu_\tau \rangle \quad (6.2a)$$

$$\text{s.t. } \delta_0 \otimes \mu_0 + \mathcal{L}^* \mu = \mu_\tau \quad (6.2b)$$

$$\mu \in \mathcal{M}_+([0, \tau] \times X \times U), \quad (6.2c)$$

$$\mu_\tau \in \mathcal{M}_+(X_T), \quad (6.2d)$$

where $\mu_0 = \delta_{x_0}$ is the initial measure with support on the initial condition x_0 , μ is the occupation measure of the system trajectories, whereas $\mu_\tau = \delta_\tau \otimes \delta_{x_\tau}$ generalizes the concept of terminal measure with free terminal time. It holds that $\langle 1, \mu \rangle = \tau$ (see Appendix B).

Although the original control problem is generally nonlinear in the state and in the control, in light of the system dynamics and cost functional, the bounding program of the form ((6.2a))–((6.2d)) is linear in the space of the measures. The linearity follows from the fact that decision variables are measures μ, μ_τ and, in this setting, the objective functionals involve integral of functions against measures

$$\langle \phi, \mu \rangle = \int \phi(x, u) d\mu, \quad \langle \Phi, \mu_\tau \rangle = \int \Phi d\mu_\tau,$$

which are linear in μ and μ_τ . Moreover, the Liouville's equation constraint in (6.2b) is linear as the adjoint operator $\mathcal{L}^*$ acts linearly on measures (see Appendix B).

The dual problem in the space of the continuous functions reads as [5], [10]:

$$q^* = \sup_{v \in \mathbb{R}} v \quad (6.3a)$$

$$\text{s.t. } v \geq v(x), \quad \forall x \in X_0 \quad (6.3b)$$

$$\mathcal{L}_f v(x) \geq -\phi(x, u), \quad \forall (x, u) \in X \times U \quad (6.3c)$$

$$v(x) \leq \Phi(x), \quad \forall x \in X_T \quad (6.3d)$$

$$v \in C^1(X). \quad (6.3e)$$

The polynomial $v(x(t))$ that solves (6.3a)–(6.3e) provides a polynomial subsolution of the Hamilton-Jacobi-Bellman equation which approximates the value function along all the optimal trajectories of the system.

The following assumptions are made in program (6.1a)–(6.1e) for the development of this work [2].

Assumption 6.2.2

X_0, X, X_T are semialgebraic sets, $X_0 = \{x \mid w_{0k}(x) \geq 0, k = 1, \dots, N_0\}$, $X = \{x \mid w_k(x) \geq 0, k = 1, \dots, N\}$, $X_T = \{x \mid w_{Tk}(x) \geq 0, k = 1, \dots, N_T\}$. Moreover, X_0, X, X_T are compact. □

Assumption 6.2.3

The vector field $\dot{x} = f(x(t), u(t))$ is Lipschitz in each argument in the compact set $[0, T] \times X \times U$ and is polynomial, thus $f(x(t), u(t)) \in \mathbb{R}^n[x, u]$. Moreover, the vector field is affine in the control input, $\dot{x} = h(x(t)) + g(x(t))u$. □

Assumption 6.2.4

The effort of the vector field is considered to be the L^2 -norm of the control input u , $\|u(t)\|_{L^2([0, T])} = \left(\int_0^T \|u\|^2 dt \right)^{1/2}$; no terminal cost is considered in the functional (6.1a). □

In light of Assumption 6.2.4, the cost functional in (6.2a) simplifies naturally, with $\phi = \|u\|^2$. Hence the cost functional reduces to

$$\langle \phi, \mu \rangle = \int \|u\|^2 d\mu = \langle \|u\|^2, \mu \rangle.$$

As a consequence, for the development of the proposed methodology, the general formulation of (6.2a)–(6.2d) takes the form [2]:

$$p^* = \inf_{\mu, \mu_\tau} \langle \|u\|^2, \mu \rangle \quad (6.4a)$$

$$\text{s.t. } \delta_0 \otimes \mu_0 + \mathcal{L}^* \mu = \mu_\tau \quad (6.4b)$$

$$\mu \in \mathcal{M}_+([0, \tau] \times X \times U), \quad (6.4c)$$

$$\mu_\tau \in \mathcal{M}_+(X_T). \quad (6.4d)$$

6.2.3 Linear Matrix Inequalities Formulation

The problem in (6.4a)-(6.4d) is solved in the space of moments with relaxation order d with the following *Linear Matrix Inequalities* (LMI) formulation [4], [5] :

$$p_d^* = \min_{y, y_\tau} \quad y_{02} \quad (6.5a)$$

$$\text{s.t. } \quad Liou_\alpha(y_0, y, y_\tau) \quad \forall \alpha = 1, \dots, 2d \quad (6.5b)$$

$$\mathbb{M}_d(y) \succeq 0, \quad \mathbb{M}_d(y_\tau) \succeq 0 \quad (6.5c)$$

$$\mathbb{M}_{d-d_k}(w_k y) \succeq 0, \quad \forall k = 1 \dots, N \quad (6.5d)$$

$$\mathbb{M}_{d-d_k}(w_{Tk} y_\tau) \succeq 0, \quad \forall k = 1 \dots, N_T, \quad (6.5e)$$

where y_0, y, y_τ represent the moment sequence respectively of measures μ_0, μ, μ_τ , y_{02} is the cost in (6.4a) expressed as 2^{nd} - order moment of the control input u with respect to the occupation measure μ . As discussed in Appendix B, constraint (6.5b) represents the Liouville equation in (6.4b) in the moment space and Moment and Localizing matrices in constraints (6.5c)-(6.5e) require moments up to degree $2d$ and enforce the measure support constraints in (6.4c)-(6.4d).

Remark 6.2.1

In (6.1) we assume a time invariant vector field. The generality of the proposed approach is preserved for time varying systems solving the optimization problem by means of time dependent auxiliary functions $\varphi(x, t)$ and extending the support measure constraints. In this case, Liouville's equation in moment space results in

$$\langle x^\alpha t^\beta, \delta_0 \otimes \mu_0 \rangle + \langle \mathcal{L}_f(x^\alpha t^\beta), \mu \rangle = \langle x^\alpha t^\beta, \mu_\tau \rangle$$

for all test functions $\varphi(x, t) = x^\alpha t^\beta$. □

6.2.4 Counterfactual Extraction

In this section we illustrate two algorithms that can be exploited independently to extract counterfactuals by means of the methodology illustrated above [2].

Proposition 6.2.1 (Counterfactual Extraction via Moments)

Consider the problem in (6.4a)-(6.4d) and its d -degree solution and consider a factual $x_f = \text{spt}(\mu_0 = \delta_{x_f})$. Then, the counterfactual $x_{cf} = \text{spt}(\mu_\tau = \delta_{x_\tau})$ and it holds that $x_{cf} = y_{1,\tau}$, where $y_{1,\tau}$ denotes the 1st-order moment of the final measure μ_τ . □

Proof. The Liouville constraint (6.4b) expresses the balance of mass between the initial, occupation, and terminal measures:

$$\delta_0 \otimes \mu_0 + \mathcal{L}^* \mu = \mu_\tau,$$

where $\mathcal{L}^*$ is the adjoint operator associated with the system dynamics. To recover conservation of total mass, we integrate both sides against a smooth test function $\varphi(x)$. By the definition of the adjoint operator (Appendix B), it holds

$$\langle \varphi, \mathcal{L}^* \mu \rangle = \langle \mathcal{L} \varphi, \mu \rangle,$$

with $\mathcal{L}\varphi = \frac{\partial \varphi}{\partial t} + \nabla_x \varphi \cdot f(x, u)$. Choosing $\varphi(x) \equiv 1$ gives $\mathcal{L}\varphi = 0$, since both derivatives vanish. Substituting into the weak form of the Liouville equation yields

$$\langle 1, \mu_\tau \rangle = \langle 1, \mu_0 \rangle.$$

Because the initial measure is $\mu_0 = \delta_{x_f}$, we have $\langle 1, \mu_0 \rangle = 1$, and therefore $\langle 1, \mu_\tau \rangle = 1$. This means that the terminal measure μ_τ also has unit mass. Under deterministic dynamics with a unique solution, the pushforward (Appendix B) of the Dirac initial measure by the optimal flow at time τ is a Dirac measure:

$$\mu_\tau = \delta_{x_\tau},$$

supported on the terminal state $x_{cf} = x_\tau$ reached by the optimal trajectory. The vector of first-order moments of μ_τ is then

$$y_{1,\tau} = \int x d\mu_\tau(x) = \int x d\delta_{x_\tau}(x) = x_\tau.$$

Hence we have that $x_{cf} = y_{1,\tau}$ which concludes the proof. $\square$

Hence, the counterfactual x_{cf} can be extracted from the Moment Matrix of the final measure $\mathbb{M}_d(y_\tau)$. Algorithm 6.1 summarizes this procedure.

Remark 6.2.2

If (6.4) admits r optimal solutions, then the final measure μ_τ is r -atomic (i.e. $\mu_\tau = \sum_{i=1}^r \beta_i \cdot \delta_{x_{\tau_i}}$). Hence, the counterfactual $x_{cf} = \text{spt}(\mu_\tau = \delta_{x_\tau}) = \sum_{i=1}^r \beta_i \cdot x_{\tau_i}$. $\square$

Moreover, considering the dual formulation (6.3), the following alternative method holds.

Proposition 6.2.2 (Counterfactual Extraction via Trajectory Recovery)

Consider the problem in (6.4a)-(6.4d) and its dual formulation in the form (6.3) with a given d -degree solution that yields a subsolution $v_d(x)$ of the value function. Let $u_d^*(x) = \arg \min_{u \in U} \{\mathcal{L}_f v_d(x) + \phi(x, u)\}$ and let $\tau_d = \text{mass}(\mu)$. Then, for a given $x_f \in X_0$, $x_{cf} = x(\tau_d | x_f, u_d^*(x))$. $\square$

Proof. Let v_d be any dual-feasible d -degree subsolution for (6.3), so that

$$\mathcal{L}_f v_d(x) + \phi(x, u) \geq 0 \quad \forall (x, u) \in X \times U, \quad v_d(x) \leq \Phi(x) \quad \forall x \in X_T.$$

Define the residual

$$r_d(x, u) := \mathcal{L}_f v_d(x) + \phi(x, u) \geq 0.$$

Let $u_d^*(x) = \arg \min_{u \in U} r_d(x, u)$, and let $x(\cdot)$ be the solution of $\dot{x} = f(x, u_d^*(x))$ with $x(0) = x_f$. Denote by τ_d the hitting time of X_T under this feedback; equivalently, τ_d is the mass of the occupation measure induced by $(x(t), u_d^*(x(t)))$ on $[0, \tau_d]$.

Along the closed loop,

$$\frac{d}{dt} v_d(x(t)) = \mathcal{L}_f v_d(x(t)) = -\phi(x(t), u_d^*(x(t))) + r_d(x(t), u_d^*(x(t))).$$

Integrating on $[0, \tau_d]$ gives

$$v_d(x(\tau_d)) - v_d(x_f) = - \int_0^{\tau_d} \phi(x(t), u_d^*(x(t))) dt + \int_0^{\tau_d} r_d(x(t), u_d^*(x(t))) dt.$$

Rearranging,

$$\int_0^{\tau_d} \phi(x(t), u_d^*(x(t))) dt = v_d(x_f) - v_d(x(\tau_d)) + \int_0^{\tau_d} r_d(x(t), u_d^*(x(t))) dt.$$

Since $\Phi \equiv 0$,

$$\int_0^{\tau_d} \phi(x(t), u_d^*(x(t))) dt = v_d(x_f) + \int_0^{\tau_d} r_d(x(t), u_d^*(x(t))) dt + (\Phi(x(\tau_d)) - v_d(x(\tau_d))).$$

By feasibility of v_d , we have $r_d \geq 0$ and $v_d \leq \Phi$ on X_T , hence both correction terms are nonnegative. Therefore

$$J(u_d^*; x_f) := \int_0^{\tau_d} \phi(x(t), u_d^*(x(t))) dt \geq v_d(x_f).$$

By strong duality for (6.4)–(6.3), $p^* \geq v_d(x_f)$, so

$$0 \leq J(u_d^*; x_f) - p^* \leq \underbrace{\int_0^{\tau_d} r_d(x(t), u_d^*(x(t))) dt}_{\text{integrated residual}} + \underbrace{(\Phi(x(\tau_d)) - v_d(x(\tau_d)))}_{\text{terminal slack}}.$$

Hence, for any finite relaxation degree d , the control law u_d^* produces an approximately optimal trajectory with suboptimality gap bounded by the integrated residual and terminal slack. The corresponding terminal measure is induced by the closed-loop trajectory,

$$\mu_\tau = \delta_{x(\tau_d) | x_f, u_d^*},$$

whose support is the counterfactual

$$x_{\text{cf}} = x(\tau_d | x_f, u_d^*(x)).$$

□

Algorithm 6.2 summarizes the procedure. Proposition 6.2.1 and Proposition 6.2.2 follow from the developments in [3], [4].

For input-affine control system and running cost of the form $\phi(x, u) = \phi_x(x) + u^T u$, the control law $u^*(x)$ can be derived from the first order optimality condition [4][5]:

$$\frac{\partial(\phi(x, u) + \nabla v(x) \cdot (h(x) + g(x)u))}{\partial u} = 2u + \nabla v(x)u = 0 \quad (6.6)$$

$$\Rightarrow u^*(x) = -\frac{1}{2} \nabla v^*(x) \cdot g(x). \quad (6.7)$$

Algorithm 6.1 Counterfactual Extraction via Moments

Require: Sets X_0, X, X_T , dynamics f , cost functional, initial degree d

Ensure: Factual-Counterfactual pair generation

- 1: Define μ_0 (factual $x_f = \text{spt}(\mu_0 = \delta_{x_f})$)
 - 2: Solve (6.4) at degree d
 - 3: Extract $y_{1,\tau}$ from $\mathbb{M}_d(y_\tau)$
 - 4: $x_{\text{cf}} = y_{1,\tau}$
-

As discussed in Appendix B, solving (6.4a)–(6.4d) and (6.3a)–(6.3e) for a given relaxation order d yields lower bound $p_d^* = q_d^* \leq p^* = q^*$ and the convergence is guaranteed as $d \rightarrow \infty$. Moreover, in the next theorem we show a formal result that ensures the boundedness of the trajectory leading to the counterfactual under the optimal control.

Algorithm 6.2 Counterfactual Extraction via Trajectory Recovery**Require:** Sets X_0, X, X_T , dynamics f , cost functional, initial degree d **Ensure:** Factual-Counterfactual pair generation

- 1: Define μ_0 (factual $x_f = \text{spt}(\mu_0 = \delta_{x_f})$)
- 2: Solve (6.4) at degree d
- 3: Extract $\tau_d = \text{mass}(\mu)$
- 4: Retrieve $v_d(x)$ from the dual program (6.3)
- 5: Retrieve $u_d^*(x) = \arg \min_{u \in U} \{\mathcal{L}_f v_d(x) + \phi(x, u)\}$
- 6: Simulate the system in closed-loop $\dot{x} = f(x(t), u_d^*(x(t)))$
- 7: $x_{cf} = x(\tau | x_f, u_d^*(x))$

Theorem 6.2.1 (Boundedness of the Trajectory)

Consider the problem (6.4a)-(6.4d) with dual in the form (6.3a)-(6.3e) and its solution for a given relaxation order d . Then, the retrieved trajectory $(x(t) | x_f, u_d^*(x))$ remains bounded on $[0, \tau_d]$. Consequently, the family of trajectories $\{x(\cdot | x_f, u_d^*)\}_d$ is uniformly bounded as $d \rightarrow \infty$. $\square$

Proof. By feasibility, the closed-loop trajectory $(x(t), u_d^*(x(t)))$ satisfies $x(t) \in X$ for all $t \in [0, \tau_d]$. Since X is compact, there exists $M > 0$ such that $\|x\| \leq M$ for all $x \in X$. Hence $\|x(t)\| \leq M$ for all $t \in [0, \tau_d]$, and the trajectory $(x(t) | x_f, u_d^*(x))$ is bounded. The bound M does not depend on d , so the trajectories are uniformly bounded as $d \rightarrow \infty$. $\square$

As additional theoretical result, we will prove the following theorem, ensuring asymptotic stability.

Theorem 6.2.2 (Asymptotic Stability)

Consider the problem (6.4a)-(6.4d) with dual in the form (6.3a)-(6.3e) and its solution for a given relaxation order d . As $d \rightarrow \infty$, the terminal set X_T containing the counterfactual x_{cf} is asymptotically stable. $\square$

Proof. By the assumptions of (6.3a)–(6.3e), $v_d \in C^1(X)$ and, by [4], as $d \rightarrow \infty$ one has $v_d \rightarrow \mathcal{V}$ uniformly on compact sets, where $\mathcal{V}$ is the value function. Moreover, with $\Phi \equiv 0$ and $\phi(x, u) = \|u\|^2 \geq 0$,

$$\mathcal{V}(x) = 0 \text{ on } X_T, \quad \mathcal{V}(x) > 0 \text{ on } X \setminus X_T,$$

so $\mathcal{V}$ is positive definite with respect to the set X_T . From (6.3), with $d \rightarrow \infty$, we have

$$\min_{u \in U} \{\mathcal{L}_f \mathcal{V}(x) + \phi(x, u)\} = 0.$$

and, evaluating at the minimizing control

$$u^*(x) = \arg \min_{u \in U} \{\mathcal{L}_f \mathcal{V}(x) + \phi(x, u)\},$$

along the closed-loop system $\dot{x} = f(x, u^*(x))$, it follows that

$$\dot{\mathcal{V}}(x) = \mathcal{L}_f \mathcal{V}(x, u^*(x)) = -\|u^*(x)\|^2 \leq 0.$$

Hence $\mathcal{V}$ decreases along system trajectories and remains constant only when $\|u^*(x)\| = 0$. Hence, LaSalle's invariance principle implies that all trajectories converge to X_T . Therefore, the terminal set X_T is asymptotically stable for the closed-loop system as $d \rightarrow \infty$. $\square$

The boundedness of the closed-loop trajectories ensures that the dynamics reconstructed under $u_d^*(x)$ remain well-posed and enforces that the support constraints imposed in the optimization correspond to an invariant and physically meaningful region for the actual system evolution. Moreover, the asymptotic stability of X_T implies that the system converges to the desired terminal configuration, corresponding to the counterfactual x_{cf} , under the optimal control. In this sense, X_T acts as an attractor for all admissible trajectories starting from the factual initial conditions x_f . Moreover, we prove the following result.

Theorem 6.2.3 (Bounded Distance)

Consider the problem (6.4a)–(6.4d) and its dual (6.3a)–(6.3d). Let us assume it is feasible and that a solution exists for all suitable relaxation orders d . Let denote with $\eta : S \times S \rightarrow \mathbb{R}$ a general continuous metric on the euclidean space $\mathbb{R}^n$ ($S \subseteq \mathbb{R}^n$). As $d \rightarrow \infty$, the distance $\eta(x_{cf,d}, x_{cf,\infty})$ is bounded. □

Proof. By assumption, X_T is compact in $\mathbb{R}^n$. Hence $X_T \times X_T$ is compact. Moreover, η is continuous on $X_T \times X_T$. By the extreme value theorem, the continuous real-valued map $(x, y) \mapsto \eta(x, y)$ attains a finite maximum on $X_T \times X_T$. Since $x_{cf,d}, x_{cf,\infty} \in X_T$, it follows that

$$\eta(x_{cf,d}, x_{cf,\infty}) \leq \max_{(x,y) \in X_T \times X_T} \eta(x, y) < \infty,$$

which proves boundedness. □

This result ensures that the sequence of counterfactual states $x_{cf,d}$ remains bounded with respect to the asymptotic counterfactual $x_{cf,\infty}$. This implies that, even for low relaxation orders, the counterfactuals reconstructed from the moment-SOS hierarchy cannot deviate arbitrarily from the limiting solution. In other words, the optimization framework produces counterfactuals that are dynamically consistent and physically meaningful even at low orders, with a guaranteed upper bound on their deviation from the true asymptotic counterfactual.

6.2.5 Results

The proposed methodology for counterfactual extraction via occupation measures is applied [2] to a general glucose-insulin regulation model, the well acknowledged model by Bergman et al.[11], described by the following system of differential equations:

$$\dot{x}_1 = -p_1 x_1 - x_2 x_1 + p_1 G_b \quad (6.8a)$$

$$\dot{x}_2 = -p_2 x_2 + p_3 x_3 \quad (6.8b)$$

$$\dot{x}_3 = -n x_3 + p_4 u \quad (6.8c)$$

where:

- x_1 is the blood glucose concentration, also denoted as G in the following (mg/dl);
- x_2 is the remote insulin concentration ($\mu\text{U}/\text{ml}$);
- x_3 is the serum insulin concentration, also denoted as I in the following ($\mu\text{U}/\text{ml}$);
- u represents exogenous insulin administration ($\mu\text{U}/\text{ml}/\text{min}$).

For the values of the parameters reference is made to [11]. To improve the numerical behavior of the *Semidefinite Programming* (SDP) solvers, variables should be normalized, i.e. scaled within boxes $[0,1]$ (see e.g. [3], [4], [5]). Hence, by defining the scaled variables

$\hat{x}_1 = \frac{x_1}{x_{1max}}, \hat{x}_2 = \frac{x_2}{x_{2max}}, \hat{x}_3 = \frac{x_3}{x_{3max}}$ the system results as follows:

$$\dot{\hat{x}}_1 = -p_1\hat{x}_1 - \hat{x}_2\hat{x}_1x_{2max} + \frac{p_1G_b}{x_{1max}} \quad (6.9a)$$

$$\dot{\hat{x}}_2 = -p_2\hat{x}_2 + p_3\hat{x}_3 \frac{x_{3max}}{x_{2max}} \quad (6.9b)$$

$$\dot{\hat{x}}_3 = -n\hat{x}_3 + \frac{p_4u}{x_{3max}}, \quad (6.9c)$$

where $x_{1max} = 600$ mg/dl, $x_{2max} = 1$ μ U/ml, $x_{3max} = 100$ μ U/ml represent reasonable maximum values respectively for the variables x_1, x_2, x_3 [11]. Moreover, time is scaled by multiplying the system (6.9a)-(6.9c) by the time horizon T , here assumed to be $T = 80 \cdot 60$ (sec), which is a time range consistent with the model by Bergman et al.[11]. The terminal set is considered to be $X_T = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid 80 \leq x_1 < 126\}$ [12], whereas the set of initial conditions is assumed to be $X_0 = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid (126 \leq x_1 \leq 260), (0 \leq x_3 \leq 30)\}$, which represent respectively reasonable safe and unsafe set when dealing with glucose-insulin dynamics.

Fig. 6.1 illustrates factual-counterfactual pairs as obtained when randomly sampling 20 factuals over X_0 , solving the optimization problem for $d = 4$ and extracting counterfactuals via Algorithm 6.1. The plot identifies the association between each factual and the related counterfactual in the phase plane (G, I) . Counterfactuals lie in the safe set in the vicinity of the boundary (i.e., the diabetic threshold) separating the safe set from the unsafe set. Fig. 6.2 shows the results obtained applying Algorithm 6.1 when factuals (red dots) are distributed as 150 random initial conditions over the set X_0 whereas counterfactuals (blue dots) are extracted solving the optimization problem for two different relaxation orders, $d = 4$ (top panel) and $d = 6$ (bottom panel). Both plots show a remarkable feature of counterfactuals, related to the fact that most of the counterfactuals cluster in a dense region within an interval of values of I between 20 and 25 μ U/ml. This is a notable analogy with respect to the AI-driven framework, that similarly denotes presence of dense regions of counterfactuals obtained by means of the classification algorithms, as discussed in [8], [13].

Fig. 6.3 shows an example of the actual trajectory leading from a given factual to its related counterfactual. The trajectory can be retrieved by applying Algorithm 6.2. More in detail, for the factual in figure $x_f = (250, 0, 0)$ (thus we simulate a diabetic virtual subject, with no endogenous insulin production), by solving the optimization problem for $d = 2$, we get

$$v_{d=2}^*(\hat{x}) = \zeta_1 + \zeta_2\hat{x}_1 + \zeta_3\hat{x}_2 + \zeta_4\hat{x}_3 + \zeta_5\hat{x}_1^2 + \zeta_6\hat{x}_1\hat{x}_2 \quad (6.10)$$

$$+ \zeta_7\hat{x}_1\hat{x}_3 + \zeta_8\hat{x}_2^2 + \zeta_9\hat{x}_2\hat{x}_3 + \zeta_{10}\hat{x}_3^2 \quad (6.11)$$

Coefficient	Value
ζ_1	-2.7308
ζ_2	18.7159
ζ_3	-353.7267
ζ_4	0.6326
ζ_5	-27.9368
ζ_6	82.6978
ζ_7	-2.0188
ζ_8	-755356.4616
ζ_9	-318.4366
ζ_{10}	-0.48977
γ_1	-1.4879
γ_2	4.7483
γ_3	748.9628
γ_4	2.3039

Table 6.1: Coefficients for $v_{d=2}^*(x)$ and $u_{d=2}^*(x)$

From (6.7), it is possible to derive

$$u_{d=2}^*(\hat{x}) = -\frac{1}{2}\nabla v_{d=2}^*(\hat{x}) \cdot g(\hat{x}) \quad (6.12)$$

$$= \gamma_1 + \gamma_2 \hat{x}_1 + \gamma_3 \hat{x}_2 + \gamma_4 \hat{x}_3. \quad (6.13)$$

As described in Algorithm 6.2, the trajectory and the counterfactual x_{cf} are retrieved by simulating in closed-loop the system $\dot{\hat{x}} = \hat{f}(\hat{x}) + \hat{g}(\hat{x})u_{d=2}^*(\hat{x})$. As observable from Fig. 6.3, the retrieved trajectory is able to steer the system to the safe set and the reached counterfactual lies in the vicinity of the boundary, as expected. Conversely, in open loop, the system evolution remains unsafe.

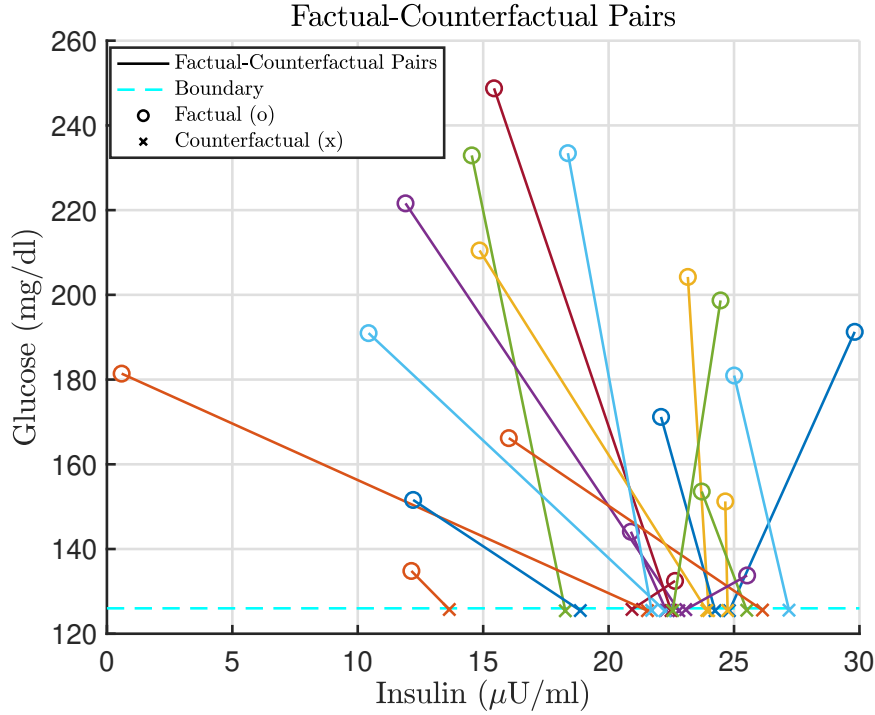


Figure 6.1: Factual-counterfactual pairs as obtained by solving the optimization problem in the moment space with relaxation order $d = 4$.

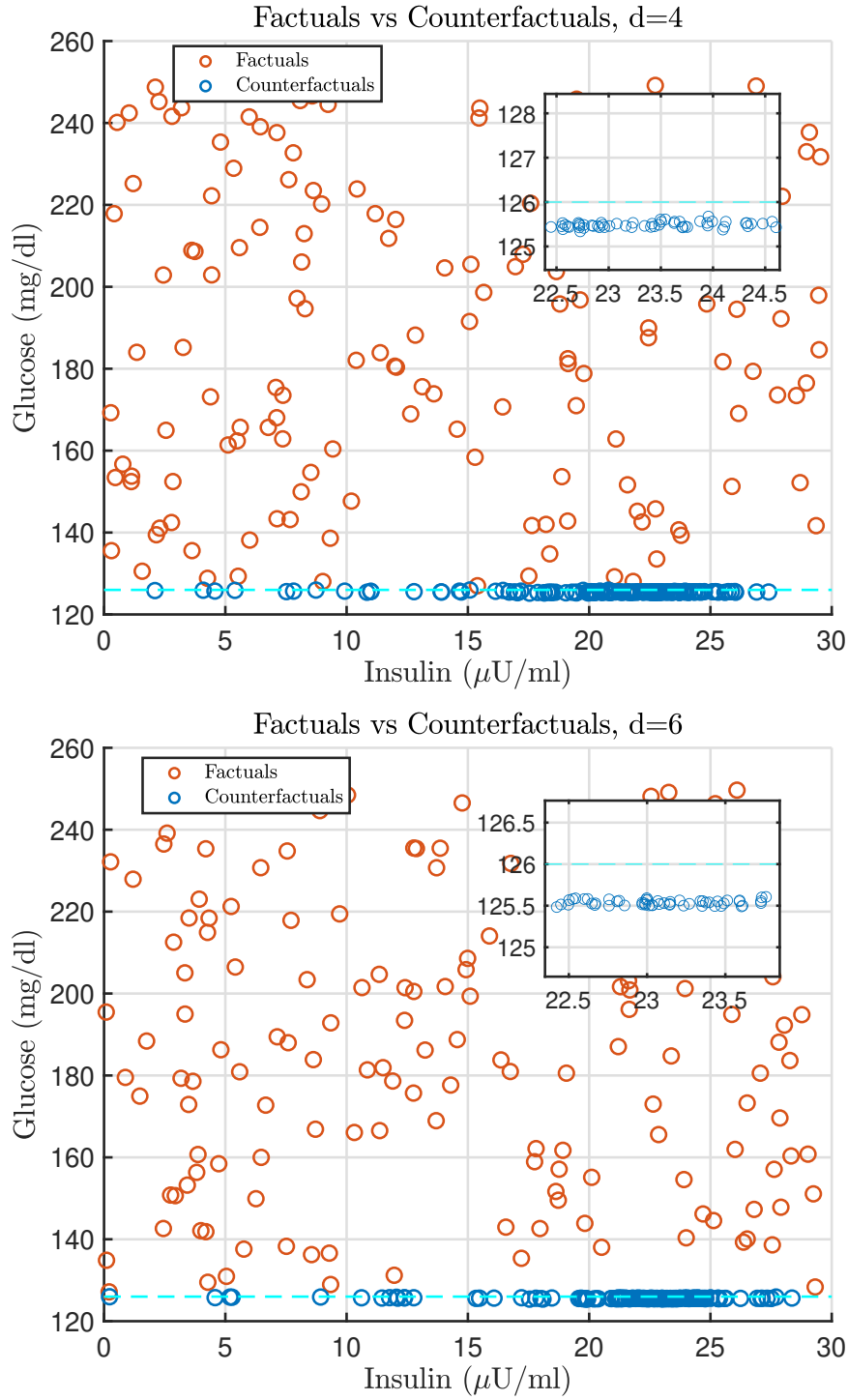


Figure 6.2: Factuals (red dots) and counterfactuals (blue dots) obtained by solving the optimization problem in the moment space with relaxation order $d = 4$ (top panel) and $d = 6$ (bottom panel).

6.3 Uncertain Systems

This section is devoted to illustrate how the methodology previously discussed for fully parametrized systems can be generalized in a robust framework to the case of uncertain systems [1].

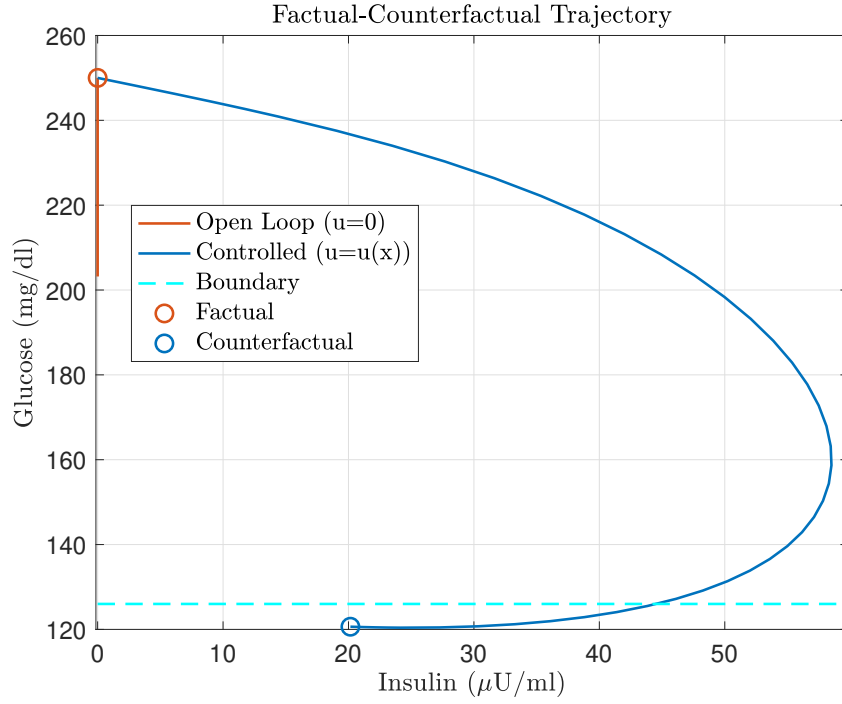


Figure 6.3: Example of counterfactual trajectory as retrieved for the factual (250,0,0) by applying Algorithm 6.2.

6.3.1 Problem Definition

In this framework [1] we consider a general time invariant differential vector field of the form:

$$\dot{x} = f(x, u, \xi) \quad (6.14)$$

where $x \in \mathbb{R}^n$ is the state of the system, $u \in \mathbb{R}^m$ is the control and $\xi \in \mathbb{R}^{n_\xi}$ are the system parameters and $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ is the vector field of the system dynamics. We assume uncertainties on the knowledge of the parameters of the system as follows.

Assumption 6.3.1

A subset $\zeta \subseteq \xi$ of the parameters is unknown and takes values in a set $\Sigma \subset \mathbb{R}^{n_\zeta}$. These values are considered to be time-independent and constant along trajectories. $\square$

Then, for the system (6.14), the general free terminal time optimal control problem can be readily restated as follows:

$$J = \inf_{u(\cdot), t^* \in [0, T]} \Psi(x(t^*)) + \int_0^{t^*} \psi(x(t), u(t), \xi) dt \quad (6.15a)$$

$$\text{s.t. } \dot{x}(t) = f(x(t), u(t), \xi), \quad (6.15b)$$

$$x(0) \in X_0, x(t) \in X, \quad (6.15c)$$

$$x(t^*) \in X_T, \quad (6.15d)$$

$$u(t) \in U, \quad t \in [0, t^*], \quad (6.15e)$$

where t^* is the free terminal time, $\Psi(x(t^*))$ is the terminal cost, $\psi(x(t), u(t), \xi)$ is the stage cost.

X_0 , X , X_T constitute respectively the set of initial states, the set of the trajectories and the set of terminal states, T is the time horizon. We will consider in general $X_0 \subset X$, $X_T \subseteq X$.

Assumptions 6.2.1, 6.2.2, 6.2.3, 6.2.4 previously made for the case of fully parametrized system continue to hold also for the case of uncertain systems.

To leverage the optimal control problem in the measures framework, following the developments in [4], [14], we will include the unknown parameters extending the state space representation of the system as follows:

$$\kappa := [x \quad \zeta_j]^T, \dot{\zeta}_j = 0 \quad (6.16)$$

for $j = 1, \dots, n_\zeta$.

In analogy with the classical definition of counterfactual and to what discussed above we will consider the *factual* to be the initial condition of the general optimal control problem (6.15a)-(6.15e) applied to the system (6.16), $\kappa_f := \kappa(0)$. Similarly, we will indicate as *counterfactual* the terminal state reached by the system in closed-loop with $u^*(\cdot) = \arg \min_{u(\cdot), t^*} J(\kappa(t), u(t))$, $\kappa_{cf} = \kappa(t^* \mid \kappa_f, u^*)$. We will denote with χ_0, χ, χ_T respectively the sets of initial conditions, of the state trajectories and of terminal states of the extended system (6.16).

Remark 6.3.1

Sets χ_0, χ, χ_T inherit the algebraic properties of X_0, X, X_T . □

6.3.2 Problem Solution in Measures

According to the developments in [3], [5], [14] the LP measure program bounding the problem in (6.15a)-(6.15e) reads as [1]

$$p^* = \inf_{\mu_0, \mu, \mu_{t^*}} \langle \psi, \mu \rangle + \langle \Psi, \mu_{t^*} \rangle \quad (6.17a)$$

$$\text{s.t.} \quad \delta_0 \otimes \mu_0 + \mathcal{L}^* \mu = \mu_{t^*} \quad (6.17b)$$

$$\langle 1, \mu_0 \rangle = 1, \quad (6.17c)$$

$$\mu \in \mathcal{M}_+([0, t^*] \times \chi \times U), \quad (6.17d)$$

$$\mu_{t^*} \in \mathcal{M}_+(\chi_T), \quad (6.17e)$$

where $\mu_0 = \delta_{\kappa_f}$ is the initial measure supported on the initial condition κ_f , μ is the occupation measure, $\mu_{t^*} = \delta_{t^*} \otimes \delta_{\kappa_{cf}}$ generalizes the concept of terminal measure with free terminal time. The infinite-dimensional optimization problem in the dual space of the continuous functions is [4], [15]:

$$q^* = \sup_{\gamma \in \mathbb{R}} \gamma \quad (6.18a)$$

$$\text{s.t.} \quad \gamma \geq v(\kappa), \quad \forall \kappa \in \chi_0 \quad (6.18b)$$

$$\mathcal{L}_f v(\kappa) \geq -\psi(\kappa, u), \quad \forall (\kappa, u) \in \chi \times U$$

$$v(\kappa) \leq \Psi(\kappa), \quad \forall \kappa \in \chi_T \quad (6.18c)$$

$$v \in C^1(\chi). \quad (6.18d)$$

The polynomial $v(\kappa(t))$ approximates the value function along all the optimal trajectories of the system.

As design choice for the development of our approach, we consider zero terminal cost ($\Psi = 0$) and we consider the *effort* (i.e., ψ) of the vector field to be the L^2 -norm of the control u . Consequently, the problem (6.17a)-(6.17e) results in [1]

$$p^* = \inf_{\mu_0, \mu, \mu_{t^*}} \langle \|u\|^2, \mu \rangle \quad (6.19a)$$

$$\text{s.t.} \quad \delta_0 \otimes \mu_0 + \mathcal{L}^* \mu = \mu_{t^*} \quad (6.19b)$$

$$\langle 1, \mu_0 \rangle = 1, \quad (6.19c)$$

$$\mu \in \mathcal{M}_+([0, t^*] \times \chi \times U), \quad (6.19d)$$

$$\mu_{t^*} \in \mathcal{M}_+(\chi_T) \quad (6.19e)$$

Constraint (6.19c) induces an atomic representation for μ_0 . From (6.17b) by choosing $\varphi = 1$ as test function it directly follows that also the final measure μ_{t^*} is atomic, being

$$\langle 1, \mu_0 \rangle + \langle \mathcal{L}_f 1, \mu \rangle = \langle 1, \mu_{t^*} \rangle.$$

More in detail, whenever (6.19) admits s optimal solutions, then the final measure μ_{t^*} is s -atomic, i.e., it is a linear combination of atomic measures supported on s points.

Remark 6.3.2

The case of completely known model (no uncertainty on model parameters) discussed above is a particular case of the more general one discussed here. In that case, the optimal control problem (and the related program in measure space) refers to the original state space of the system. The uncertainties on parameters here result in a different optimization program in measures (6.19a)-(6.19e) with different decision variables and constraints.

□

6.3.3 LMI Formulation

The problem (6.19a)-(6.19e) can be formulated in the truncated moment space and solved as follows [4], [5]:

$$p_d^* = \min_{y_0, y, y_{t^*}} \quad y_{02} \quad (6.20a)$$

$$\text{s.t.} \quad Liou_\alpha(y_0, y, y_{t^*}) \quad \forall \alpha = 1, \dots, 2d \quad (6.20b)$$

$$y_{0,0} = 1 \quad (6.20c)$$

$$\mathbb{M}_d(y_0), \mathbb{M}_d(y), \mathbb{M}_d(y_{t^*}) \succeq 0 \quad (6.20d)$$

$$\mathbb{M}_{d-d_k}(g_{0k}y) \succeq 0, \quad \forall k = 1 \dots, N_0 \quad (6.20e)$$

$$\mathbb{M}_{d-d_k}(g_k y) \succeq 0, \quad \forall k = 1 \dots, N \quad (6.20f)$$

$$\mathbb{M}_{d-d_k}(g_{Tk} y_{t^*}) \succeq 0, \quad \forall k = 1 \dots, N_T \quad (6.20g)$$

Measures μ_0, μ, μ_{t^*} have representing moment sequences respectively y_0, y, y_{t^*} guaranteed by the algebraic properties of the support sets. The cost is expressed as 2^{nd} -order moment of the control u with respect to the occupation measure μ (y_{02}). $Liou_\alpha(y_0, y, y_{t^*})$ in (6.20b) denotes here Liouville's equation expressed in the moment space and ensures the relation

$$\langle \kappa^\alpha, \delta_0 \otimes \mu_0 \rangle + \langle \mathcal{L}_f \kappa^\alpha, \mu \rangle = \langle \kappa^\alpha, \mu_{t^*} \rangle$$

holds for all test functions $\varphi(\kappa) = \kappa^\alpha$. Moment and Localizing matrices in constraints (6.20d)-(6.20g) represent the support set constraints (6.19d)-(6.19e) in the truncated moments space (see Appendix B).

6.3.4 Counterfactual Extraction

This section shows how Proposition 6.2.1 previously proposed can be extended to the generalized robust framework to allow the extraction of counterfactuals. We extend Proposition 6.2.1 as follows [1].

Proposition 6.3.1 (Robust Counterfactual Extraction via Moments)

Consider the problem in (6.19a)-(6.19e) and its d -degree solution. Then, the counterfactual $\kappa_{cf} = \text{spt}(\mu_{t^*} = \delta_{\kappa_{t^*}})$ associated with the factual $\kappa_f = \text{spt}(\mu_0 = \delta_{\kappa_f})$ can be retrieved from the moment matrix and specifically it holds that $\kappa_{cf} = y_{1,t^*}$, where y_{1,t^*} is the 1st-order moment of the final measure μ_{t^*} . $\square$

Algorithm 6.3, which extends the previously proposed Algorithm 6.1, summarizes the procedure. The formulation we design in this study inherently yields a safety problem. The formal results expressed by Theorem 6.2.1, Theorem 6.2.2, Theorem 6.2.3 hold also for the case here studied of uncertainties on the model for the extended system (6.16).

Algorithm 6.3 Robust Counterfactual Extraction via Moments

Require: Sets χ_0, χ, χ_T , vector field f , cost functional, degree d

Ensure: Robust Counterfactual extraction

- 1: Define the known support components of μ_0 , $\text{spt}(\mu_0) = \kappa_f$
- 2: Solve (6.19) at degree d
- 3: Extract y_{1,t^*} from $\mathbb{M}_d(y_{t^*})$
- 4: $\kappa_{cf} = y_{1,t^*}$

Remark 6.3.3

Solving the dual program (6.18a)-(6.18d), counterfactuals can be equivalently extracted retrieving $u^*(\kappa) = \arg \min_{u \in U} \{\mathcal{L}_f v(\kappa) + \psi(\kappa, u)\}$ and simulating the closed-loop system via Algorithm 6.2. $\square$

6.3.5 Results

We apply [1] the proposed approach to the system (6.8) scaled as (6.9), where we assume that parameters (p_2, p_3) are unknown, $p_2 \in [0.049, 0.051]$ and $p_3 \in [2.7 \cdot 10^{-5}, 3 \cdot 10^{-5}]$. Indeed this is a reasonable choice, as parameters p_2 and p_3 are related to the dynamics of the remote insulin (x_2) hence identifying their values is not trivial in a real-world scenario. We assume the safe set to be $X_T = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid 80 \leq x_1 < 126\}$ whereas initial conditions are sampled within the set $X_0 = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid (126 \leq x_1 \leq 260), (0 \leq x_3 \leq 30)\}$ [16].

In the following we will compare the results obtained when the system is assumed to be known, as discussed above, and the case when the system is assumed to be uncertain. Robust counterfactuals for the uncertain system are extracted via Algorithm 6.3 leveraging the proposed methodology. Fig. 6.4 shows pairs of factuals and counterfactuals when 10 factuals are randomly sampled over the unsafe set and the optimization problem is solved for $d = 4$ (top panel) and $d = 6$ (bottom panel) both for the known system and the uncertain one [1].

For a given factual (initial, unsafe condition for the system, red dot), the plot highlights in yellow and blue respectively the association between the counterfactual extracted in case of known system and the counterfactual robust with respect to the model uncertainties. As it can be observed from the plot, when the system is assumed known, counterfactuals tend to lie on the boundary (dashed line) separating the two sets. Conversely, in the robust scenario, counterfactuals lie in the safe set in points at lower values of G . This result appears to be reasonable, as the robust optimal control law ensures that the system keeps safe despite the uncertainties on its dynamics and this results in an increased average distance from the boundary.

It should be remarked that the generated counterfactuals are points reached by the system dynamics with lower bound with respect to the cost achieved as $d \rightarrow \infty$. Indeed Fig. 6.5 shows the results obtained when robust counterfactuals are extracted solving the optimization problem for $d = 4$ (top panel) and $d = 6$ (bottom panel) and factuals are 50 randomized points over the space of the initial conditions. As it can be observed, the plot

highlights a specific property of counterfactuals, as they tend to cluster in dense regions of the state space, independently of the relaxation order d that solves that optimization problem as observed also previously for the case of fully parametrized systems. It can be easily observed that in Fig. 6.5 the dense region lies within an interval for the values of I between 0 and $10 \mu\text{U/ml}$ and this dense region is different with respect to that denoted in Fig. 6.2 when the system is assumed to be known. It would be of interest to investigate to what extent the presence of dense regions is in general associated with specific properties of the system under consideration, such as separatrix, basins of attractions and multiple equilibria.

6.4 Conclusion

The work presented in this chapter represents a novel proposal aimed at introducing the concept of counterfactuals in the field of control and its applications, leveraging the powerful tools of occupation measures and moment-SOS hierarchy. Control-driven counterfactuals are inherently informed about the system dynamics, thus embedding the information about how the system can be steered to the safe set and which terminal states can be reached at a given terminal time. This distinctive feature represents a conceptual step forward with respect to the AI-framework, where most counterfactuals overlook the knowledge of the system dynamics and may represent points not reachable by the system in a real-world scenario. Further developments will investigate whether some of the formal properties described in [8] and [17] can be extended to the case of control-driven counterfactuals. Moreover, future work will deal with a stochastic oriented setting, leveraging different initial distributions and final distributions of counterfactuals. Within the context of life science applications, this study set the basis for future works leveraging the proposed approach of control-driven counterfactuals to integrate control and AI with the aim of feeding machine learning algorithms with the physics knowledge acquired via the system framework, with specific reference to applications in the field of T2D control [18], [19], [20].

References

- [1] De Paola, P. F., Miller, J., Borri, A., Paglialonga, A., and Dabbene, F., “Robust control-driven counterfactual generation for uncertain systems,” in *11th IFAC Symposium on Robust Control Design ROCOND 2025*, vol. 59, 2025, pp. 116–121. DOI: <https://doi.org/10.1016/j.ifacol.2025.10.089>.
- [2] De Paola, P. F., Miller, J., Borri, A., Paglialonga, A., and Dabbene, F., “A control system framework for counterfactuals: An optimization based approach,” *arXiv preprint arXiv:2501.12914*, 2025.
- [3] Henrion, D., Lasserre, J. B., and Savorgnan, C., “Nonlinear optimal control synthesis via occupation measures,” in *2008 47th IEEE Conference on Decision and Control*, IEEE, 2008, pp. 4749–4754.
- [4] Lasserre, J. B., Henrion, D., Prieur, C., and Trélat, E., “Nonlinear optimal control via occupation measures and lmi-relaxations,” *SIAM journal on control and optimization*, vol. 47, no. 4, pp. 1643–1666, 2008.
- [5] Henrion, D., Korda, M., and Lasserre, J. B., *Moment-sos Hierarchy, The: Lectures In Probability, Statistics, Computational Geometry, Control And Nonlinear Pdes*. World Scientific, 2020, vol. 4.
- [6] Lasserre, J. B., *Moments, positive polynomials and their applications*. World Scientific, 2009, vol. 1.
- [7] Lewis, D., *Counterfactuals*. John Wiley & Sons, 2013.

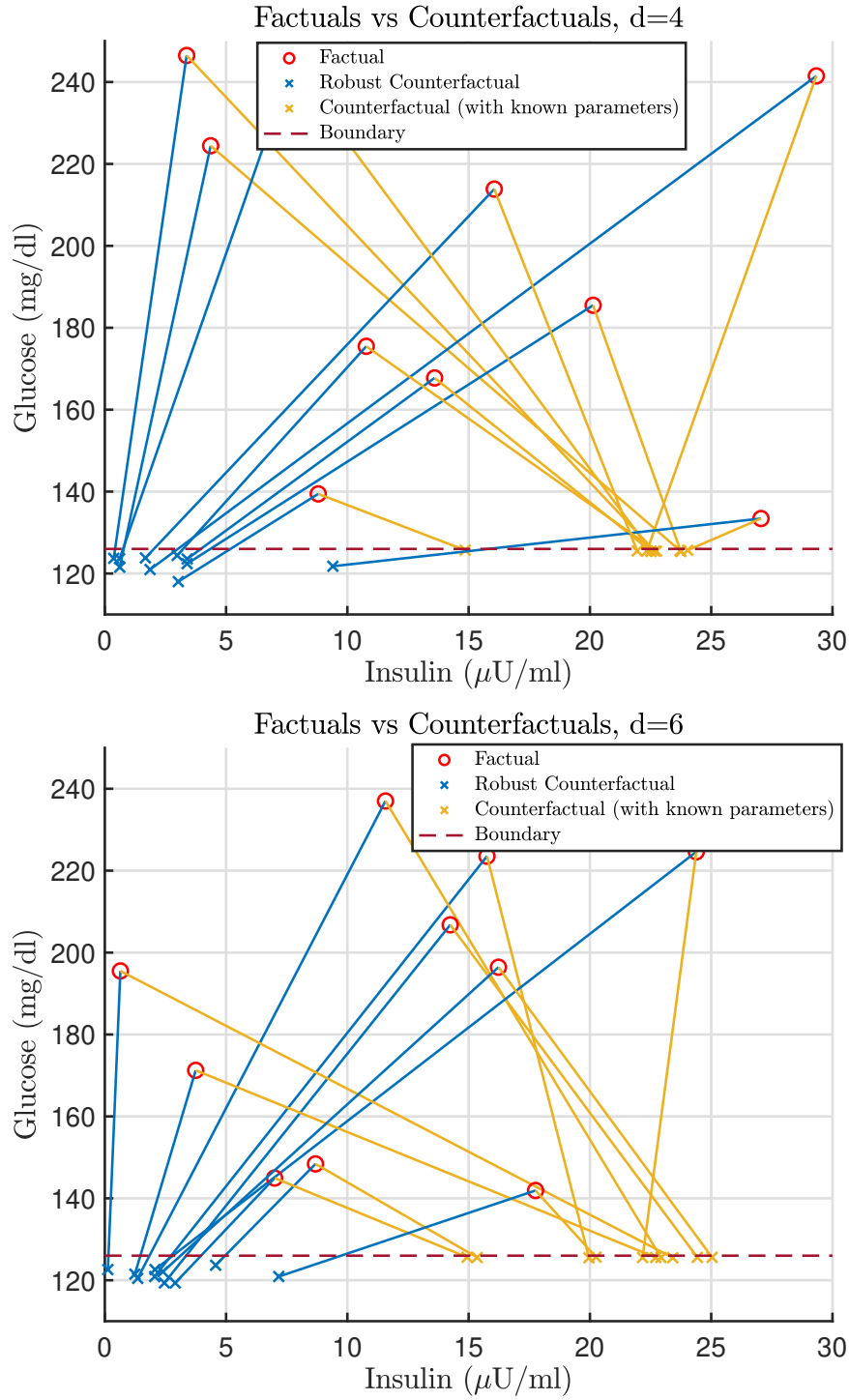


Figure 6.4: Factual-counterfactual pairs as obtained by solving the optimization problem in the moment space with relaxation order $d = 4$ (top panel) and $d = 6$ (bottom panel). Robust counterfactuals lie at lower values of G with increased average distance from the boundary.

- [8] Guidotti, R., “Counterfactual explanations and how to find them: Literature review and benchmarking,” *Data Mining and Knowledge Discovery*, vol. 38, no. 5, pp. 2770–2824, 2024.

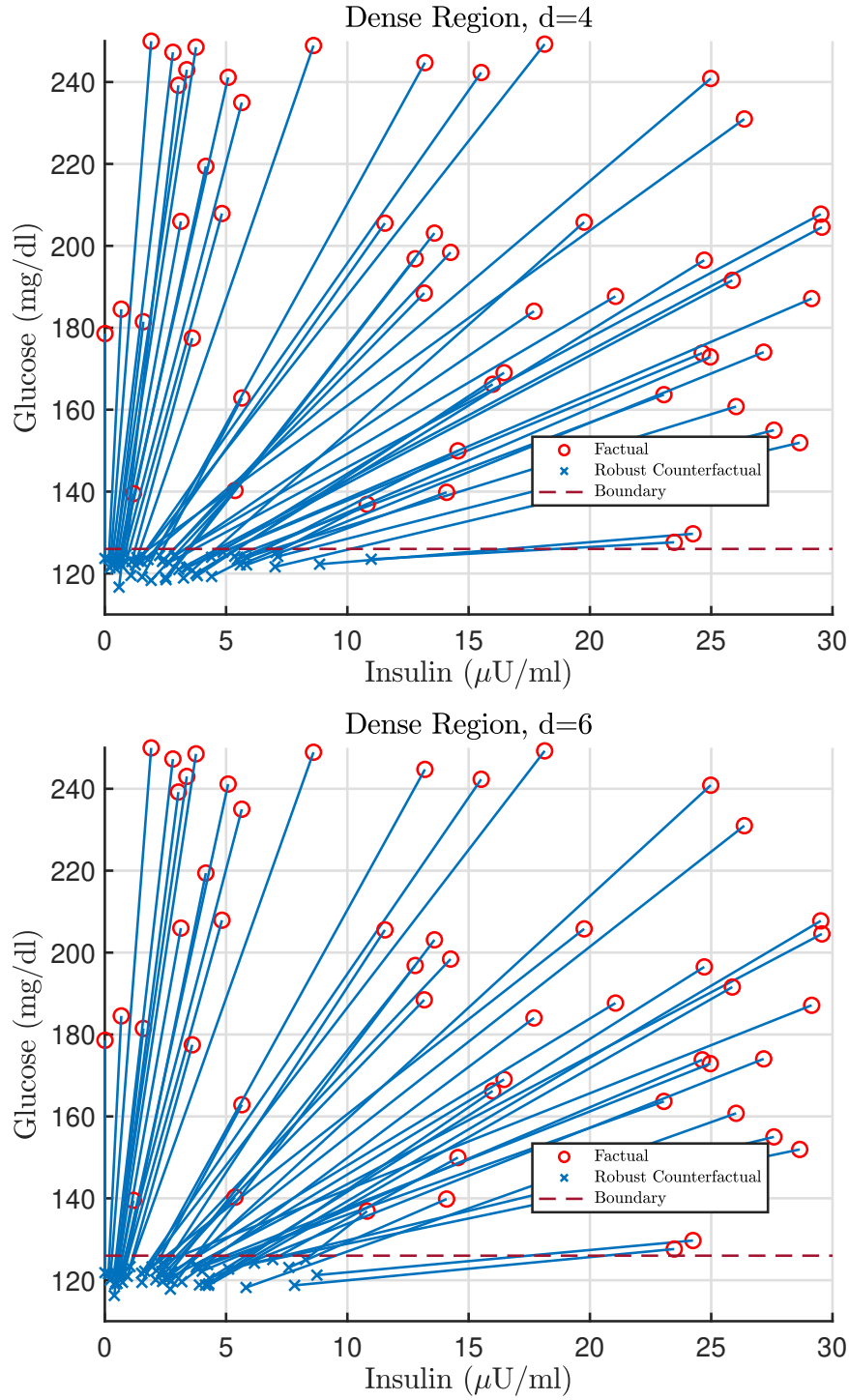


Figure 6.5: Factuals vs Counterfactuals obtained for $d = 4$ (top panel) and $d = 6$ (bottom panel). Both the plots highlight the same dense region of counterfactuals within the interval of values of I between 0 and 10.

- [9] Wachter, S., Mittelstadt, B., and Russell, C., “Counterfactual explanations without opening the black box: Automated decisions and the GDPR,” *Harv. JL & Tech.*, vol. 31, p. 841, 2017.
- [10] Fattorini, H. O., *Infinite dimensional optimization and control theory*. Cambridge University Press, 1999, vol. 54.

- [11] Bergman, R. N., Phillips, L. S., Cobelli, C., et al., “Physiologic evaluation of factors controlling glucose tolerance in man: Measurement of insulin sensitivity and beta-cell glucose sensitivity from the response to intravenous glucose,” *The Journal of clinical investigation*, vol. 68, no. 6, pp. 1456–1467, 1981.
- [12] Whelan, M. and Bell, L., “The English national health service diabetes prevention programme (NHS DPP): A scoping review of existing evidence,” *Diabetic Medicine*, vol. 39, no. 7, e14855, 2022.
- [13] Poyiadzi, R., Sokol, K., Santos-Rodriguez, R., De Bie, T., and Flach, P., “Face: Feasible and actionable counterfactual explanations,” in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 2020, pp. 344–350.
- [14] Henrion, D., Ganet-Schoeller, M., and Bennani, S., “Measures and lmi for space launcher robust control validation,” *IFAC Proceedings Volumes*, vol. 45, no. 13, pp. 236–241, 2012.
- [15] Fantuzzi, G. and Goluskin, D., “Bounding extreme events in nonlinear dynamics using convex optimization,” *SIAM journal on applied dynamical systems*, vol. 19, no. 3, pp. 1823–1864, 2020.
- [16] Bergman, R. N., Ider, Y. Z., Bowden, C. R., and Cobelli, C., “Quantitative estimation of insulin sensitivity,” *American Journal of Physiology-Endocrinology And Metabolism*, vol. 236, no. 6, E667, 1979.
- [17] Carlevaro, A., Lenatti, M., Paglialonga, A., and Mongelli, M., “Multi-class counterfactual explanations using support vector data description,” *IEEE Transactions on Artificial Intelligence*, 2023.
- [18] Lenatti, M., Carlevaro, A., Guergachi, A., Keshavjee, K., Mongelli, M., and Paglialonga, A., “A novel method to derive personalized minimum viable recommendations for type 2 diabetes prevention based on counterfactual explanations,” *Plos one*, vol. 17, no. 11, e0272825, 2022.
- [19] De Paola, P. F., Paglialonga, A., Palumbo, P., Keshavjee, K., Dabbene, F., and Borri, A., “The long-term effects of physical activity on blood glucose regulation: A model to unravel diabetes progression,” *IEEE Control Systems Letters*, vol. 7, pp. 2916–2921, 2023. DOI: [10.1109/LCSYS.2023.3290774](https://doi.org/10.1109/LCSYS.2023.3290774).
- [20] De Paola, P. F., Borri, A., Dabbene, F., Keshavjee, K., Palumbo, P., and Paglialonga, A., “Modeling the cumulative benefits of regular physical activity on type 2 diabetes progression,” *Computers in Biology and Medicine*, vol. 198, p. 111 194, 2025. DOI: <https://doi.org/10.1016/j.compbiomed.2025.111194>.

Chapter 7

An Optimal Transport Based Formulation for Counterfactuals via Occupation Measures

Contents

7.1	Introduction	103
7.2	Problem Statement	104
7.3	Counterfactual Extraction	111
7.4	Results	112
7.5	Conclusion	113
7.6	Background and Additional Contents	115

7.1 Introduction

This chapter is devoted to the introduction of a novel proposal for counterfactuals definition in the context of dynamical systems. In the previous chapter we have proposed a framework for counterfactuals based on an optimal control problem formulation, in which basically the counterfactual coincides with the final state reached by the dynamical system under a control law designed to balance the best “effort” to reach a given safe set. The approach provided here is focused on a different proposal, indeed closer to the philosophical “what-if” spirit encoded in the classical concept of counterfactual [1], [2], [3]. Motivated by a scenario in which a given set of initial states (the factuals) of the system evolve toward an undesirable or unsafe set, the central question we address here is whether it is possible to identify alternative initial conditions, which we call here counterfactuals, that are as close as possible in the sense of a given metric to the factuals, but such that they steer the system towards a safe set. This viewpoint captures the essence of counterfactual reasoning in the context of dynamical systems: “what would have happened if the system had started in slightly different way?”

To formalize this question, we adopt the framework of optimal transport [4], [5], [6]. Here, factual and counterfactual initial conditions are treated either as individual points or as probability distributions, and the aim is to minimize a suitable notion of distance between them, subject to the dynamical constraints of the system. In this way, the optimal transport formulation naturally quantifies how “close” a counterfactual is to the factual reference, while ensuring that trajectories reach safe terminal conditions. The proposed analysis develops through three layers of abstraction:

- we first define a pointwise formulation, in which factual and counterfactual initial conditions are represented by single states, and the problem reduces to a constrained projection onto the set of admissible safe initial conditions;
- we then generalize to the measure-theoretic setting, where both factual and counterfactual initial conditions are described by probability measures. In this case, the quadratic Wasserstein distance provides a natural extension of the Euclidean distance between points [6]. We also show how it is possible to treat the pointwise problem as particular case of this distributional level setting;

- finally, we reformulate the problem in terms of occupation measures [7], [8], [9], [10], which encode the evolution of state distributions along trajectories and allow us to cast the counterfactual problem as a LP over measures.

The occupation-measure viewpoint is particularly powerful, as it converts nonlinear dynamics into possible infinite dimensional linear constraints, laying the ground to systematic convex relaxations and to the solution of the problem by means of the moment-SOS hierarchy. The methodology here discussed is able to translate the intuitive notion of factual and counterfactual scenarios into an optimal transport problem, thereby providing a mathematically precise yet conceptually faithful realization of the classical philosophical notion of factual and counterfactual reasoning into the field of dynamical systems and control theory. While the primary focus of this Thesis is on modeling and control of T2D, the methodologies developed throughout this chapter and Chapter 6 are not limited to this application, as they constitute a more general framework for the analysis of nonlinear systems. To highlight this broader relevance, the methodology proposed in this final chapter will be illustrated through an example drawn from a general biomedical system. In doing so, we emphasize the methodological generality of the proposed approach and its potential to provide insight into a wide range of biomedical applications beyond diabetes control.

7.2 Problem Statement

7.2.1 Pointwise Formulation

Consider a general autonomous nonlinear system on a state space $X \subset \mathbb{R}^n$

$$\dot{x}(t) = f(x(t)), \quad (7.1)$$

with initial condition $x(0) = x_0 \in X_0 \subseteq X$.

For the system (7.1) we denote $X_{0,u}$ the set of initial states that let the system evolve toward an unsafe set $\mathcal{U}$ at a given terminal time $T > 0$,

$$X_{0,u} := \{x_0 \in X : x(T|x_0) \in \mathcal{U}\}, \quad X_{0,u} \subseteq X_0.$$

Analogously, we denote as $X_{0,s}$ the set of initial states that steer the system to a safe set $\mathcal{S}$ at the terminal time $T > 0$,

$$X_{0,s} := \{x_0 \in X : x(T|x_0) \in \mathcal{S}\}, \quad X_{0,s} \subseteq X_0.$$

The following assumptions are made for system (7.1) for the subsequent development of this chapter.

Assumption 7.2.1

The vector field $\dot{x} = f(x(t))$ is polynomial, thus $f(x(t)) \in \mathbb{R}^n[x]$. X , $X_{0,u}$, $X_{0,s}$ are semialgebraic sets, $X = \{x \mid w_k(x) \geq 0, k=1, \dots, N_X\}$, $X_{0,u} = \{x \mid w_{0u,k}(x) \geq 0, k=1, \dots, N_0\}$, $X_{0,s} = \{x \mid w_{0s,k}(x) \geq 0, k=1, \dots, N_0\}$, $\deg(w_k(x)) = \deg(w_{0u,k}(x)) = \deg(w_{0s,k}(x)) = d_k$ for all k . Moreover, we assume that X , $X_{0,u}$, $X_{0,s}$ are compact.

□

The polynomial assumption implies the (local) Lipschitzity of the vector field f on X , thus guaranteeing that there exists a unique Carathéodory solution on some maximal interval of interest for all $x_0 \in X_0$. Moreover, the polynomial assumption is a necessary condition for leveraging moment-SOS hierarchy. For the system (7.1) we consider the following optimization program, whose solution provides a first pointwise formulation for counterfactuals.

Problem 7.2.1 (Optimization Problem 1 (Pointwise Formulation))

Given $x_{0,u} \in X_{0,u}$, solve :

$$\min_{x_{0,s}} \quad \frac{1}{2} \|x_{0,u} - x_{0,s}\|^2 \quad (7.2a)$$

$$s.t. \quad \dot{x}(t) = f(x(t)) \quad \text{for a.e. } t \in [0, T], \quad (7.2b)$$

$$x(0) = x_{0,s}, \quad (7.2c)$$

$$x_{0,s} \in X_{0,s} \quad (7.2d)$$

$$x \in X \quad (7.2e)$$

□

Within the framework of this chapter we define the *factual* and the *counterfactual* as follows.

Definition 7.2.1 (Factual)

Consider the general optimization problem of the form (7.2). We define the *factual* as

$$x_f := x_{0,u} \in X_{0,u}.$$

□

Definition 7.2.2 (Counterfactual)

Consider the general optimization problem of the form (7.2). Given a *factual* x_f , we define the *counterfactual* x_{cf} associated with the *factual* x_f as

$$x_{cf} := x_{0,s}^* = \arg \min_{x_{0,s} \in X_{0,s}} \frac{1}{2} \|x_{0,u} - x_{0,s}\|^2$$

□

In other words, we define the *factual* as the initial condition, defined as a point, which induces an unsafe trajectory for the system, ending up into the unsafe set $\mathcal{U}$. Conversely, the *counterfactual* is defined to be the point, closest to the *factual* in the sense of Euclidean distance, capable of steering the system to the safe set $\mathcal{S}$ over time evolution. Hence, identifying the *counterfactual* by means of Problem 7.2.1 results in finding the closest variation in the initial conditions able to change the evolution in terms of safety of the system.

7.2.2 Distributional Extension (Measure Level)

If we consider the initial conditions to be distributions of states instead of points it is possible to provide a measure-level representation of Problem 7.2.1. For the preliminary background related to the general measure setting reference is made to Appendix B.

Let $\mathcal{M}_+(X)$ denote the set of Borel measures on X , let $\Psi(X)$ denote the set of probability measures on X and let $\mu_{0,u}$ and $\mu_{0,s}$ be respectively the initial unsafe and safe distributions and let $\Phi_t : X \rightarrow X$ denote the flow map associated with the vector field f at time $t \in [0, T]$. It is possible to generalize the optimization problem in the form (7.2) at the distributional level as follows.

Problem 7.2.2 (Optimization Problem 2 (Measure Level))

Given an initial unsafe distribution $\mu_{0,u} \in \Psi(X_{0,u})$, solve:

$$\mathcal{P}^* = \inf_{\mu_{0,s} \in \Psi(X_0)} \quad \frac{1}{2} W_2^2(\mu_{0,u}, \mu_{0,s}) \quad (7.3a)$$

$$s.t. \quad \text{spt}((\Phi_T)_\# \mu_{0,s}) \subseteq \mathcal{S}, \quad (7.3b)$$

where $(\Phi_T)_\# \mu_{0,s}$ denotes the pushforward measure (Appendix B) of $\mu_{0,s}$ through the

$\text{map}(\Phi)_t$ and the cost functional (7.3a) represents the quadratic Wasserstein distance

$$W_2^2(\mu_{0,u}, \mu_{0,s}) := \min_{\pi \in \Pi(\mu_{0,u}, \mu_{0,s})} \int_{X \times X} \|x - y\|^2 d\pi(x, y), \quad (7.4)$$

where $\Pi(\mu_{0,u}, \mu_{0,s})$ denotes the set of transport plans between $\mu_{0,u}$ and $\mu_{0,s}$. $\square$

Consequently, the concept of factual and counterfactual can be readily restated as follows.

Definition 7.2.3 (μ -Factual)

Consider the general optimization problem of the form (7.3). We define the μ -factual $\mathcal{X}_f$ as

$$\mathcal{X}_f := \text{spt}(\mu_{0,u}).$$

$\square$

Definition 7.2.4 (μ -Counterfactual)

Consider the general optimization problem of the form (7.3). Given a μ -factual $\mathcal{X}_f$, we define the μ -counterfactual $\mathcal{X}_{cf}$ associated with the μ -factual $\mathcal{X}_f$ as

$$\mathcal{X}_{cf} := \text{spt}(\mu_{0,s}^*) = \text{spt}\left(\arg \min_{\mu_{0,s} \in \Psi(X_0)} \left\{ \frac{1}{2} W_2^2(\mu_{0,u}, \mu_{0,s}) : \text{spt}((\Phi_T)_\# \mu_{0,s}) \subseteq \mathcal{S} \right\}\right).$$

$\square$

In simple terms, Problem 7.2.2 generalizes the pointwise Problem 7.2.1 to the case of distributions of states. Specifically, the factual distribution $\mu_{0,u}$ describes the initial measure of states that, under the system dynamics, evolve into the unsafe region. The counterfactual distribution $\mu_{0,s}$, by contrast, represents an alternative choice of initial states that are guaranteed to reach the safe region at time T . Indeed, the condition involving the pushforward measure $(\Phi_T)_\# \mu_{0,s}$ (see Appendix B) enforces the condition that, when the counterfactual distribution is transferred forward through the flow of the system, its mass must end up inside the safe set $\mathcal{S}$. Then, the optimizer selects among all those distributions, the one that is closest to the unsafe factual distribution, with closeness measured in the quadratic Wasserstein sense, that generalizes to measures the Euclidean distance in problem in the form (7.2).

This problem is naturally framed within the framework of optimal transport. Figure 7.1 provides a graphical overview of the problem: the Wasserstein distance quantifies the minimal effort required to transport mass from a measure $\mu_{0,u}$ supported in $X_{0,u}$ to a measure $\mu_{0,s}$ supported in $X_{0,s}$ and these sets are associated respectively with unsafe and safe evolution for the system under consideration.

7.2.3 Lifting the Problem into Occupation Measures

As mentioned above, the formulation in (7.3) inherently yields an optimal transport problem with dynamical constraints. Indeed, classical optimal transport are infinite-dimensional LPs over plans $\pi \in \Pi(\mu, \nu)$. Although tractable in some cases (e.g. discrete settings), the general problem over continuous spaces and nonlinear dynamics is nontrivial to solve [4], [5], [6]. In our framework, to address the complexity of the problem (namely related to the infinite-dimensional formulation, to the nonlinear dynamics and to the induced nonconvex constraints) and to allow a manageable solution via tractable approximations in the space of the moments, we rely on occupation measures. In this way, we recast the original problem into a convex formulation solvable via discretization and SDP relaxations through moment-SOS hierarchy [7], [8], [9], [10]. To recast the problem into an occupation measure formulation we exploit the following result.

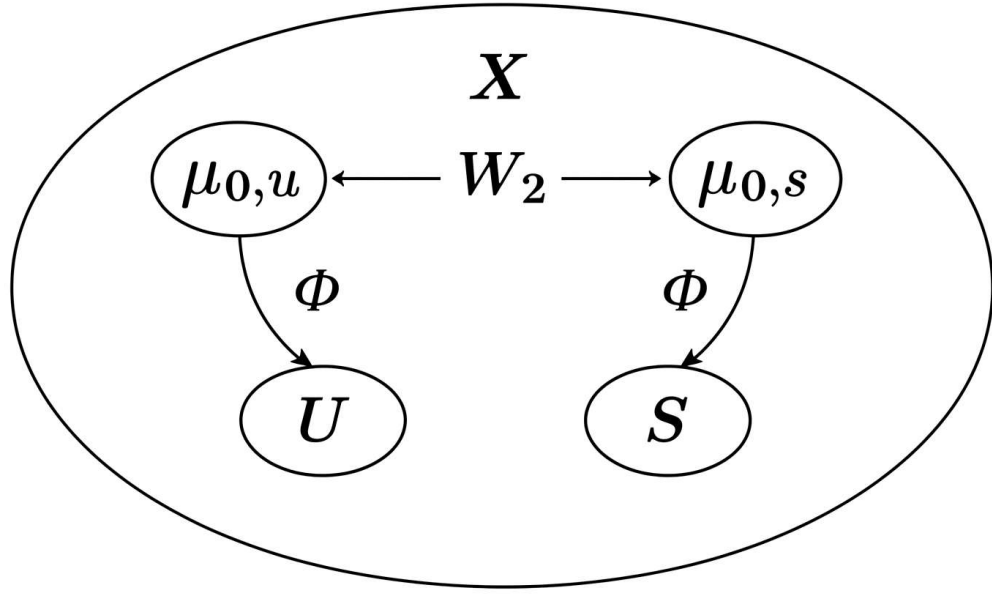


Figure 7.1: Graphical representation of the counterfactual problem in the measure framework.

Theorem 7.2.1 (Convex Reformulation via Occupation Measures)

Consider the optimization problem in the form (7.3). The problem is equivalent to the following optimization program over occupation measures:

$$\mathcal{P}_\mu^* = \inf_{\mu_{0,s}, \eta, \mu_s} \frac{1}{2} \langle \|v\|^2, \eta \rangle \quad (7.5a)$$

$$\text{s.t.} \quad \frac{\partial \eta}{\partial \tau} + \nabla_x \cdot (v\eta) = 0, \quad (7.5b)$$

$$\frac{\partial \mu_s}{\partial t} + \nabla_x \cdot (f\mu_s) = 0, \quad (7.5c)$$

$$\eta(\tau = 0) = \mu_{0,u}, \quad (7.5d)$$

$$\eta(\tau = 1) = \mu_{0,s}, \quad (7.5e)$$

$$\mu_s(t = 0) = \mu_{0,s}, \quad (7.5f)$$

$$\text{spt}(\mu_s(t = T)) \subseteq \mathcal{S}, \quad (7.5g)$$

$$\mu_{0,s} \in \Psi(X_0), \quad (7.5h)$$

$$\eta \in \mathcal{M}_+(X \times [0, 1] \times \mathcal{V}), \quad (7.5i)$$

$$\mu_s \in \mathcal{M}_+(X \times [0, T]). \quad (7.5j)$$

Moreover, it holds that $\mathcal{P}^* = \mathcal{P}_\mu^*$. □

Proof. Consider any feasible $\mu_{0,s} \in \Psi(X_0)$ for Problem 7.2.2, such that $\text{spt}((\Phi_T)_\# \mu_{0,s}) \subseteq \mathcal{S}$. By the Benamou-Brenier formula, there exist a narrowly continuous curve of measures $\rho(\cdot, \tau)$ and a Borel vector field $w(\cdot, \tau)$ such that

$$\frac{\partial \rho}{\partial \tau} + \nabla \cdot (\rho w) = 0, \quad \rho(\cdot, 0) = \mu_{0,u}, \quad \rho(\cdot, 1) = \mu_{0,s}$$

and

$$\int_0^1 \int_X \|w\|^2 \rho \, dx \, d\tau = W_2^2(\mu_{0,u}, \mu_{0,s}).$$

Define the *transport occupation measure* $\eta \in \mathcal{M}_+(X \times [0, 1] \times \mathcal{V})$

$$d\eta(x, \tau) := \rho(x, \tau) \, dx \, d\tau \quad (7.6)$$

and set $v := w$. Testing the strong continuity equation with $\varphi \in C_c^1(X \times [0, 1])$ and integrating by parts gives the weak continuity equation (see section *Background and Additional Contents (7.6), Eq. (7.12)*) with boundary data $\eta(\tau = 0) = \mu_{0,u}$ and $\eta(\tau = 1) = \mu_{0,s}$. Therefore the continuity constraint and its boundary conditions in (7.5) hold.

Moreover,

$$\frac{1}{2} \langle \|v\|^2, \eta \rangle = \frac{1}{2} \int_0^1 \int_X \|w\|^2 \rho \, dx \, d\tau = \frac{1}{2} W_2^2(\mu_{0,u}, \mu_{0,s}),$$

so the objective in (7.5) equals the Wasserstein cost.

Let μ_s be the state occupation measure on $X \times [0, T]$ induced by the flow Φ_t with initial boundary $\mu_s(t = 0) = \mu_{0,s}$. By the superposition principle, μ_s satisfies the weak Liouville equation (*Background and Additional Contents (7.6), Eq. (7.15)*) with Lipschitz vector field f , and its terminal marginal is $\mu_s(t = T) = (\Phi_T)_\# \mu_{0,s}$. Since $\text{spt}((\Phi_T)_\# \mu_{0,s}) \subseteq \mathcal{S}$, the terminal support constraint in (7.5) holds. Thus $(\mu_{0,s}, \eta, \mu_s)$ is feasible for (7.5) and achieves cost $\frac{1}{2} W_2^2(\mu_{0,u}, \mu_{0,s})$.

Taking the infimum over all feasible $\mu_{0,s}$ in Problem 7.2.2 gives $\mathcal{P}_\mu^* \leq \mathcal{P}^*$.

Conversely, let $(\mu_{0,s}, \eta, \mu_s)$ be feasible for (7.5). By the weak continuity equation (*Background and Additional Contents (7.6), Eq. (7.12)*) with boundary $\eta(\tau = 0) = \mu_{0,u}$ and $\eta(\tau = 1) = \mu_{0,s}$, the pair (η, v) is an admissible transport interpolation from $\mu_{0,u}$ to $\mu_{0,s}$ with kinetic action $\frac{1}{2} \langle \|v\|^2, \eta \rangle$. Hence, by the variational characterization of W_2 ,

$$\frac{1}{2} W_2^2(\mu_{0,u}, \mu_{0,s}) \leq \frac{1}{2} \langle \|v\|^2, \eta \rangle.$$

Furthermore, by the weak Liouville equation (*Background and Additional Contents (7.6), Eq. (7.15)*) with boundary $\mu_s(t = 0) = \mu_{0,s}$, we have $\mu_s(t = T) = (\Phi_T)_\# \mu_{0,s}$; the terminal support constraint in (7.5) then implies $\text{spt}((\Phi_T)_\# \mu_{0,s}) \subseteq \mathcal{S}$. Thus $\mu_{0,s}$ is feasible for Problem 7.2.2 and $\mathcal{P}^* \leq \mathcal{P}_\mu^*$ follows by taking infima. $\square$

Problem (7.5) can be readily interpreted as follows. The cost (7.5a) expresses the Benamou-Brenier kinetic energy as squared L^2 -norm of the velocity field v . Constraint (7.5b) is the continuity equation, recording the transfer of mass according to the transport occupation measure η . The boundary conditions (7.5d)–(7.5e) enforce that this transport starts from the factual distribution $\mu_{0,u}$ at $\tau = 0$ and ends at the counterfactual distribution $\mu_{0,s}$ at $\tau = 1$. Constraint (7.5c) is the Liouville equation in the physical time variable t , which guarantees that the counterfactual initial distribution $\mu_{0,s}$ evolves according to the system dynamics f . Its boundary condition (7.5g) requires that the resulting terminal states at time T lie entirely in the safe set $\mathcal{S}$. The program comprehensively searches for (i) a safe initial distribution $\mu_{0,s}$, (ii) a transport occupation measure η describing how to morph the unsafe distribution into this safe one with minimal Benamou-Brenier cost, and (iii) an occupation measure μ_s ensuring that this safe distribution indeed flows to the safe region.

Remark 7.2.1

The pointwise problem in the form (7.2) is a particular case of the more general distributional-level problem. Indeed, in the case of initial conditions distributed as points, it holds

$$\mu_{0,u} = \delta_{x_{0,u}}, \quad \mu_{0,s} = \delta_{x_{0,s}}.$$

Moreover, by (7.4), we have

$$W_2^2(\delta_{x_{0,u}}, \delta_{x_{0,s}}) = \inf_{\pi \in \Pi(\delta_{x_{0,u}}, \delta_{x_{0,s}})} \int_{X \times X} \|x - y\|^2 d\pi(x, y).$$

whose right-hand side can be readily computed as the set $\Pi(\delta_{x_{0,u}}, \delta_{x_{0,s}})$ only contains one element, specifically $\pi = \delta_{x_{0,u}} \otimes \delta_{x_{0,s}}$. Substituting we get

$$W_2^2(\delta_{x_{0,u}}, \delta_{x_{0,s}}) = \int_{X \times X} \|x - y\|^2 \delta_{x_{0,u}} \delta_{x_{0,s}} dx dy = \|x_{0,u} - x_{0,s}\|^2.$$

□

In light of this, for the subsequent developments we will refer only to the distributional-level case and the pointwise case will be treated as a particular case when measures are indeed distributed as Dirac measures.

7.2.4 General Extension with Unknown Factual Distribution

In this section we provide an extension of Problem 7.2.2 in the case of unknown factual distribution.

Problem 7.2.3 (Generalized Measure-Level Problem)

The measure level problem in the form (7.3) with unknown factual distribution reads as:

$$\mathcal{P}^* = \inf_{\mu_{0,u}, \mu_{0,s} \in \Psi(X_0)} \frac{1}{2} W_2^2(\mu_{0,u}, \mu_{0,s}) \quad (7.7a)$$

$$\text{s.t.} \quad \text{spt}((\Phi_T)_\# \mu_{0,u}) \subseteq \mathcal{U} \quad (7.7b)$$

$$\text{spt}((\Phi_T)_\# \mu_{0,s}) \subseteq \mathcal{S}, \quad (7.7c)$$

with cost (7.7a) defined as in (7.4). □

The interpretation of problem in the form (7.7) is straightforward from problem (7.3). In this general extension, the factual distribution $\mu_{0,u} \in \Psi(X_0)$ is now a decision variable. The additional constraint involving the pushforward measure $(\Phi_T)_\# \mu_{0,u}$ enforces the condition that the initial unsafe measure $\mu_{0,u}$ is transferred by the system flow to the unsafe set $\mathcal{U}$.

While problem (7.7) proposes a natural generalization of the proposed measure-level framework to the case of unknown factual distributions, this generalized setting is not the primary focus of our analysis. For completeness, the theoretical result ensuring the equivalent occupation-measure formulation and its moment-SOS relaxation are presented in the Background and Additional Contents section (7.6).

7.2.5 LMI Formulation

In this section we illustrate the LMI formulation associated with the d -degree relaxation in the moment space for the optimization problem in (7.5).

Problem 7.2.4 (LMI Formulation for the Measure-Level Problem)

The problem in (7.5) can be formulated in the truncated moment space and solved at a given relaxation order d as follows.

$$p_d^* = \min_{y_{0,s}, y_s, y_{T,s}, y_\eta} \frac{1}{2} y_{002,\eta} \quad (7.8a)$$

$$\text{s.t.} \quad Liou_{\alpha,\beta}(y_\eta, y_{0,s}, y_{0,u}) \quad \forall (\alpha, \beta) : |\alpha| + \beta \leq 2d, \quad (7.8b)$$

$$Liou_{\alpha,\beta}(y_s, y_{0,s}, y_{T,s}) \quad \forall (\alpha, \beta) : |\alpha| + \beta \leq 2d, \quad (7.8c)$$

$$\mathbb{M}_d(y_{0,s}) \succeq 0, \quad (7.8d)$$

$$\mathbb{M}_{d-d_{0s,k}}(w_{0s,k} y_{0,s}) \succeq 0, \quad k = 1, \dots, N_0, \quad (7.8e)$$

$$\mathbb{M}_d(y_s) \succeq 0, \quad (7.8f)$$

$$\mathbb{M}_{d-d_{X,i}}(w_i y_s) \succeq 0, \quad i = 1, \dots, N_X, \quad (7.8g)$$

$$\mathbb{M}_{d-1}(t(T-t) y_s) \succeq 0, \quad (7.8h)$$

$$\mathbb{M}_d(y_{T,s}) \succeq 0, \quad (7.8i)$$

$$\mathbb{M}_{d-d_{S,\ell}}(s_\ell y_{T,s}) \succeq 0, \quad \ell = 1, \dots, N_S, \quad (7.8j)$$

$$\mathbb{M}_d(y_\eta) \succeq 0, \quad (7.8k)$$

$$\mathbb{M}_{d-d_{X,i}}(g_i y_\eta) \succeq 0, \quad i = 1, \dots, N_X, \quad (7.8l)$$

$$\mathbb{M}_{d-1}(\tau(1-\tau) y_\eta) \succeq 0, \quad (7.8m)$$

$$\mathbb{M}_{d-d_{V,j}}(h_j y_\eta) \succeq 0, \quad j = 1, \dots, N_V. \quad (7.8n)$$

□

The cost is expressed as 2^{nd} -order moment of the vector field v with respect to the transport occupation measure η . As discussed in Appendix B, measures $\mu_{0,s}, \eta, \mu_s$ are represented by moment sequences respectively $y_{0,s}, y_\eta, y_s$, guaranteed by the algebraic properties of the support sets. Constraints (7.8b) and (7.8c) translate the continuity equations in (7.5b) and (7.5c) in the moment space while also accounting for the marginal (Appendix B) constraints in (7.5d), (7.5e), (7.5f) by means of the moment sequences $y_{0,u}$ and $y_{T,s}$. Moment and Localizing matrices in (7.8d)-(7.8n) enforce measure support constraints in (7.5g)-(7.5j) in moment space.

We state now the following result, that ensures the convergence of the hierarchy as the relaxation order d increases.

Theorem 7.2.2 (Convergence of the Hierarchy)

Consider the problem in the form (7.5) and its d -degree relaxation in (7.8). Then, as $d \rightarrow \infty$, it holds

$$\lim_{d \rightarrow \infty} p_d^* = \mathcal{P}_\mu^*.$$

□

Proof. The result follows from standard convergence arguments in the moment-SOS hierarchy. Under Assumption 7.2.1, for each relaxation order d , the truncated hierarchy provides a lower bound on the true optimum:

$$p_d^* \leq \mathcal{P}_\mu^*, \quad \forall d,$$

since every feasible measure induces feasible moment sequences up to order $2d$. By Putinar's Positivstellensatz (Appendix B), any polynomial that is nonnegative on a compact semialgebraic set with Archimedean quadratic modules (Appendix B) can be represented (up to arbitrary precision) as a sum of squares weighted by the defining constraints of the set. Consequently, the feasible sets of the truncated moment problems converge, in the weak-star sense, to the feasible set of the infinite-dimensional measure problem.

Therefore, the sequence of optimal values $\{p_d^*\}_{d \in \mathbb{N}}$ is nondecreasing and bounded above, hence convergent, satisfying

$$\lim_{d \rightarrow \infty} p_d^* = \mathcal{P}_\mu^*.$$

Moreover, there exists a sequence of feasible truncated moment tuples whose corresponding measures converge weakly to an optimal measure triplet $(\mu_{0,s}^*, \eta^*, \mu_s^*)$ of (7.5). □

7.3 Counterfactual Extraction

This section is devoted to the illustration of algorithms that can be employed to generate counterfactuals by means of the proposed methodology. We first illustrate the pointwise case (i.e., with measures distributed as Dirac measures) and then we extend the analysis to the case of general measures.

7.3.1 Counterfactual Extraction (Pointwise Problem)

For counterfactual extraction in the pointwise case, we will rely on the following result.

Proposition 7.3.1 (Counterfactual Extraction via Moment Matrix)

Consider the problem in (7.5) and its d -degree solution and consider a μ -factual $x_f = \text{spt}(\mu_{0,u} = \delta_{x_{0,u}})$. Then, the μ -counterfactual $x_{cf} = \text{spt}(\mu_{0,s}^* = \delta_{x_{0,s}})$ can be retrieved from the moments of $\mu_{0,s}^*$ and specifically it holds

$$x_{cf} = y_{1,\mu_{0,s}^*}$$

where $y_{1,\mu_{0,s}^*}$ denotes the first order moment of the measure $\mu_{0,s}^*$. □

Proof. By hypothesis $\mu_{0,s}^* = \delta_{x_{0,s}}$. For any Borel integrable function φ , the pushforward property of the Dirac measure gives $\int_X \varphi(x) d\delta_{x_{0,s}}(x) = \varphi(x_{0,s})$. In particular, for the coordinate functions $\varphi(x) = x_i$, $i = 1, \dots, n$, we obtain

$$(y_{1,\mu_{0,s}^*})_i = \int_X x_i d\mu_{0,s}^*(x) = \int_X x_i d\delta_{x_{0,s}}(x) = (x_{0,s})_i.$$

Hence $y_{1,\mu_{0,s}^*} = x_{0,s}$. Since $x_{cf} = \text{spt}(\mu_{0,s}^*) = \{x_{0,s}\}$, we conclude $x_{cf} = y_{1,\mu_{0,s}^*}$. □

Proposition 7.3.1 allows counterfactual extraction via moment matrix. Algorithm 7.1 summarizes this procedure.

Algorithm 7.1 Counterfactual Extraction via Moment Matrix

Require: Sets X_0 , X , $\mathcal{S}$, dynamics f , cost functional, initial degree d

Ensure: Factual-Counterfactual pair generation

- 1: Define $\mu_{0,u} = \delta_{x_f}$ (factual $x_f = \text{spt}(\mu_{0,u})$)
- 2: Solve the d -degree relaxation (7.8) of (7.5)
- 3: Extract $y_{1,\mu_{0,s}^*}$ from $\mathbb{M}_d(y_{0,s})$
- 4: $x_{cf} = y_{1,\mu_{0,s}^*}$

Remark 7.3.1

If the d -degree relaxation is flat (Appendix B), i.e. $\text{rank } \mathbb{M}_d(y_{0,s}) = \text{rank } \mathbb{M}_{d-1}(y_{0,s}) = 1$, then $\mu_{0,s}^*$ is a Dirac and the support point is recovered as $x_{cf} = y_{1,\mu_{0,s}^*}$. If the rank is $r > 1$, the problem admits r optimal solutions, then measure $\mu_{0,s}^*$ is r -atomic (i.e. $\mu_{0,s}^* = \sum_{i=1}^r \beta_i \cdot \delta_{x_{\mu_{0,s}^*,i}}$). Hence, the counterfactual $x_{cf} = \sum_{i=1}^r \beta_i \cdot x_{\mu_{0,s}^*,i}$. □

7.3.2 Counterfactual Extraction (Measure-Level Problem)

Counterfactual extraction for the measure-level problem is more challenging and involves the retrieval of the support of the involved measure. To this end, we will rely on the following result.

Proposition 7.3.2 (Counterfactual Extraction via Christoffel-Darboux Polynomial)

Consider the problem in (7.5) and the related μ -factual $\mathcal{X}_f$. Consider the d -degree

solution $\mu_{0,s}^*$, the associated Moment Matrix $\mathbb{M}_d(y_{0,s})$ and define the Christoffel-Darboux polynomial as

$$p_{\epsilon_d}(x) := b(x)^T (\mathbb{M}_d(y_{0,s}) + \epsilon_d I)^{-1} b(x),$$

where $\epsilon_d > 0$ is a design small parameter and $b(x)$ is the vector of monomials. Then

$$\mathcal{X}_{cf} = \text{spt}(\mu_{0,s}^*) \approx \{x \in X : \frac{1}{p_{\epsilon_d}(x)} \geq \alpha_d\}$$

with $\alpha_d > 0$. □

Proposition 7.3.2 allows counterfactual extraction by the identification of the support of the measure, which can be approximated by suitable superlevel sets of the inverse of the Christoffel-Darboux polynomial. The approximation becomes tighter as the relaxation order d increases. In this way, Proposition 7.3.2 allows to highlight the regions where the measure places most of its mass, thus indicating where the counterfactual is most likely to lie. Proposition 7.3.2 follows almost directly from the results in [11], [12]. Algorithm 7.2 summarizes this procedure.

Algorithm 7.2 Counterfactual Extraction via Christoffel-Darboux Polynomial

Require: Sets $X_0, X, \mathcal{S}$, dynamics f , cost functional, initial degree d , $\epsilon_d > 0$, $\alpha_d > 0$

Ensure: Factual-Counterfactual pair generation

- 1: Define $\mu_{0,u}$
 - 2: Solve the d -degree relaxation (7.8) of (7.5)
 - 3: Extract $\mathbb{M}_d(y_{0,s})$
 - 4: Build the Christoffel-Darboux polynomial $p_{\epsilon_d}(x) = b(x)^T (\mathbb{M}_d(y_{0,s}) + \epsilon_d I)^{-1} b(x)$
 - 5: $\mathcal{X}_{cf} = \text{spt}(\mu_{0,s}^*) \approx \{x \in X : \frac{1}{p_{\epsilon_d}(x)} \geq \alpha_d\}$
-

The Christoffel-Darboux polynomial $p_{\epsilon_d}(x) = b(x)^T (\mathbb{M}_d(y_{0,s}) + \epsilon_d I)^{-1} b(x)$ is computed with a small Tikhonov regularization parameter $\epsilon_d > 0$ to stabilize the condition number of $(\mathbb{M}_d(y_{0,s}) + \epsilon_d I)$. In practice, ϵ_d is chosen such that $\kappa(\mathbb{M}_d(y_{0,s}) + \epsilon_d I)$ remains moderate and $\alpha_d > 0$ is calibrated to a desired coverage. Under compactness assumption, there exist level sequences for which these sets converge to the true support as $d \rightarrow \infty$ and $\epsilon_d \rightarrow 0$ [11].

7.4 Results

We apply the proposed methodology to the following system, which mimics the tumor growth dynamics in a prey-predator model fashion [13]:

$$\dot{x}_1 = T\left(-a(x_1 - p_1)(x_1 - p_2)(x_1 - p_3) - b\left(x_2 - (x_1^2 - c x_1 + (x_1 - p_1)(x_1 - p_3) + C_0)\right)\right), \quad (7.9)$$

$$\dot{x}_2 = T\left((x_1 - p_1)(x_1 - p_3) - b\left(x_2 - (x_1^2 - c x_1 + (x_1 - p_1)(x_1 - p_3) + C_0)\right)\right). \quad (7.10)$$

The state variable $x_1 \in [0, 1]$ represents a normalized tumor size, with $x_1 = 0$ corresponding to absence of tumor and $x_1 = 1$ to a maximal reference size. The auxiliary state $x_2 \in [0, 1]$ captures a immune activity response, also normalized between 0 and 1. Both variables are dimensionless and represent fractions of maximal reference values. This normalization renders the system scale-free and emphasizes the qualitative dynamics rather than absolute magnitudes.

The system exhibits a bistable behavior characterized by two stable equilibria separated by an intermediate unstable saddle point. Specifically, there exists a stable safe equilibrium $E_1 = (0.1, 0.20)$, corresponding to a low tumor burden state, and a stable unsafe

Parameter	Value
T	10
a	10
b	1.0
c	0.05
p_1	0.1
p_2	0.5
p_3	0.9
C_0	0.2

Table 7.1: Parameter value

equilibrium $E_2 = (0.9, 0.96)$, representing a high tumor burden state. These two regimes are separated by an unstable saddle point $E_3 = (0.4, 0.19)$. For the simulation setting, we consider the safe set to be $\mathcal{S} = \{x \in X : 0 \leq x_1 \leq 0.3\}$, reflecting a region of the state space where the tumors burden remains minimal.

The simulation scenarios involve the simulation of the pointwise problem, with factuals distributed as Dirac measures with support randomized over points in the set $X_{0,u} = \{x \in X : 0.7 \leq x_1 \leq 1, 0.6 \leq x_2 \leq 0.8\}$, and the simulation of the measure-level problem, with the initial measure $\mu_{0,u}$ distributed as a uniform measure supported over the set $X_{0,u} = \{x \in X : 0.8 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}$. The choice of the sets of the unsafe initial conditions, both in the pointwise and measure-level cases, reflects advanced or unfavorable tumor conditions, including scenarios where the immune system remains active but insufficient to suppress tumor progression. Fig. 7.2-top panel illustrates the factual-counterfactual pairs when considering the pointwise problem with initial factuals distributed as points over the initial unsafe set. Counterfactuals are extracted via Algorithm 7.1 when solving the optimization program with relaxation order $d = 6$. The plot highlights the link between each factual and the associated counterfactual.

Fig. 7.2-bottom panel illustrates the trajectories associated with the same factual and counterfactual pairs of Fig. 7.2-top panel. As clearly observable from Fig. 7.2-bottom panel, counterfactuals steer the system, at the end of the simulation horizon, to the safe equilibrium condition, whereas factuals let the system evolve toward the unsafe equilibrium: this provides a perfect graphical synthesis of the factual-counterfactual concept, contrasting unsafe factual outcomes with safe counterfactual ones, as also represented in the time evolution plots in Fig. 7.3.

Fig. 7.4 shows the results obtained when solving the measure-level problem for relaxation order $d = 6$ and $d = 8$. Specifically, the plots illustrate the support of the measure $\mu_{0,s}^*$ which solves the optimization problem, retrieved via Algorithm 7.2 by means of the Christoffel-Darboux Polynomial associated with the measure $\mu_{0,s}^*$.

As it can be readily observed, the measure $\mu_{0,s}^*$ appears to be concentrated over a subset of the plane in the neighborhood of $x_1 \approx 0.7, x_2 \approx 1$. The retrieved support shrinks over narrower set as the relaxation order d increases, reflecting a sharper approximation of the true support of the measure associated with the increasing d , as reasonably expected.

7.5 Conclusion

Optimal transport theory, originally developed to study the most efficient way of morphing distributions, has recently found growing relevance in control and systems theory, where it enables the characterization of state transitions, uncertainty propagation and stochastic control, filtering and estimation, path planning [14], [15], [16], [17], [18], [19].

In this chapter we have leveraged the framework of optimal transport to provide an original rigorous formulation of the factual-counterfactual paradigm in the setting of dynamical systems.

Starting from the pointwise formulation, we extended the setting to distributions and reformulated the resulting optimal transport problem in terms of occupation measures,

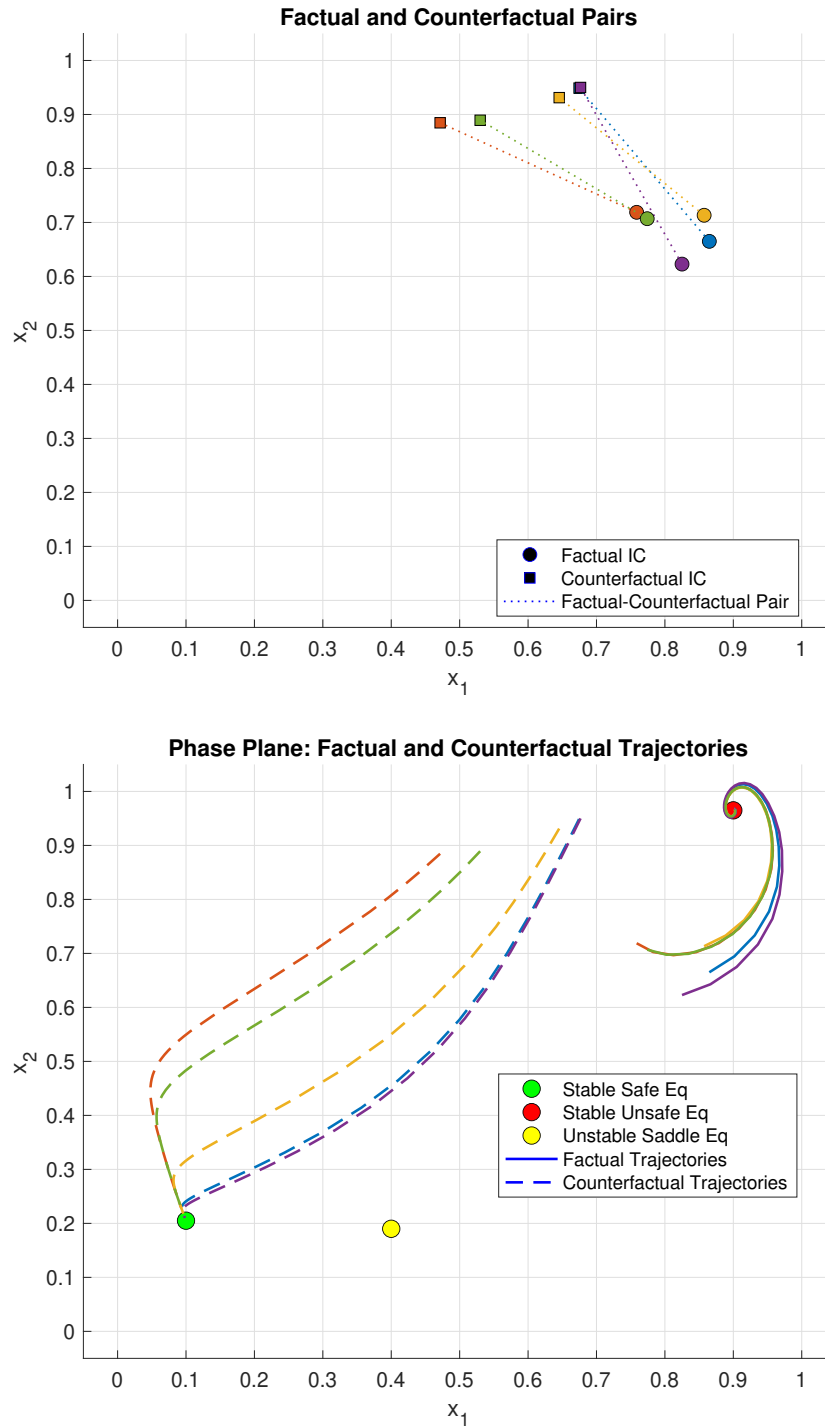
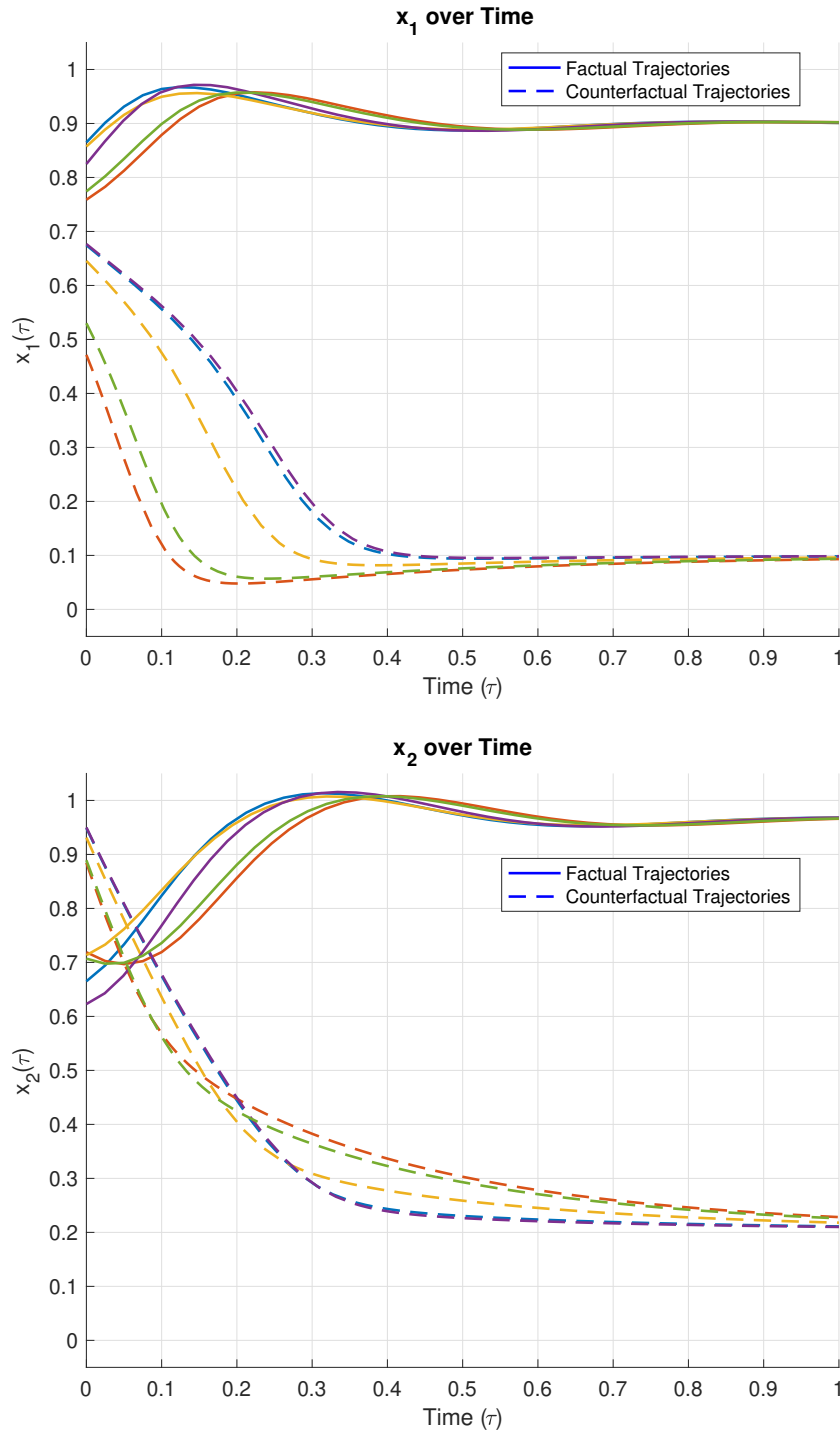


Figure 7.2: Factuals vs Counterfactuals obtained for $d = 6$

through convex program encoding both transport and system dynamics. Beyond its theoretical appeal, this framework straightforwardly offers important implications. Potential applications range from safety verification and robust control design to biomedical modeling for disease control [20], [21], [22], where interpretable counterfactual reasoning can support decision-making in complex systems.

Figure 7.3: Factuals vs Counterfactuals obtained for $d = 6$

7.6 Background and Additional Contents

7.6.1 Preliminaries

7.6.1.1 Preliminaries on Optimal Transport

The theory of optimal transport [5], [6] provides a mathematical framework to quantify the most efficient way of rearranging one distribution of mass into another. Originally

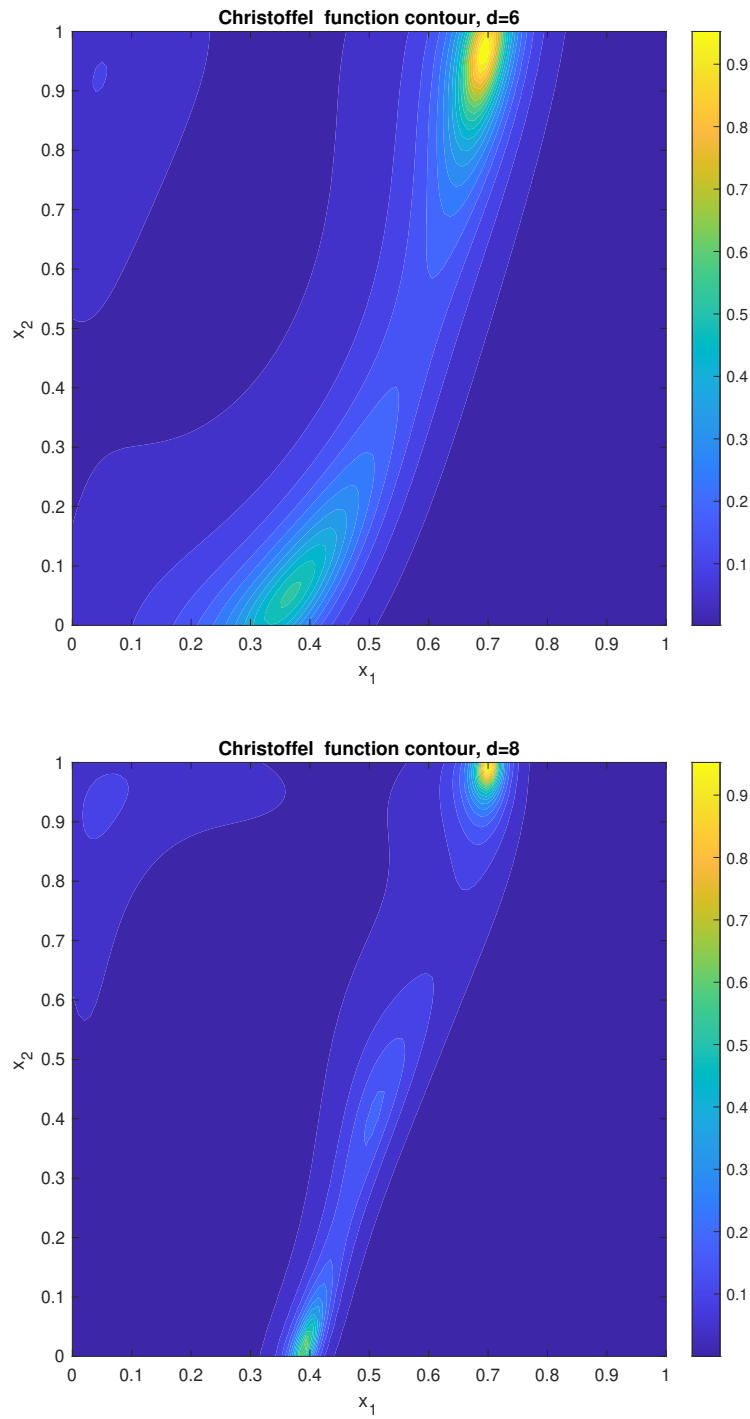


Figure 7.4: Factuals vs Counterfactuals obtained for $d = 6$ and $d = 8$

posed by Gaspard Monge in 1781 in the context of moving soil, the problem has since evolved into a rich area of analysis with profound connections to probability, geometry, and partial differential equations. We briefly recall the classical formulations and their interconnections.

7.6.1.2 The Monge Problem

Let μ_0 and μ_1 be two probability measures on $\mathbb{R}^d$, representing the initial and target mass distributions, respectively. Given a cost function $c : \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, +\infty)$, Monge formulated the problem of finding a measurable map $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that

$$T_{\#}\mu_0 = \mu_1,$$

where $T_{\#}$ denotes the pushforward operator, i.e. $(T_{\#}\mu_0)(B) = \mu_0(T^{-1}(B))$ for Borel sets B . The objective is then

$$\inf_{T: T_{\#}\mu_0 = \mu_1} \int_{\mathbb{R}^d} c(x, T(x)) d\mu_0(x).$$

This formulation requires that each unit of mass located at x is transported entirely to a single location $T(x)$. However, such a map may not exist in general, rendering Monge's problem ill-posed in certain cases.

7.6.1.3 The Kantorovich Relaxation

In 1942, Kantorovich introduced a relaxation that allows splitting of mass. Instead of a deterministic transport map, one considers a transport plan π , i.e. a probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ whose marginals coincide with μ_0 and μ_1 :

$$\pi(A \times \mathbb{R}^d) = \mu_0(A), \quad \pi(\mathbb{R}^d \times B) = \mu_1(B).$$

The set of admissible couplings is denoted $\Pi(\mu_0, \mu_1)$. The Kantorovich problem is then

$$\inf_{\pi \in \Pi(\mu_0, \mu_1)} \int_{\mathbb{R}^d \times \mathbb{R}^d} c(x, y) d\pi(x, y).$$

This relaxation guarantees existence of minimizers under mild conditions. Moreover, if a Monge-type map T is admissible, the induced plan $\pi_T = (\text{Id} \times T)_{\#}\mu_0$ is feasible in the Kantorovich problem, which shows that the Monge formulation is a special case of Kantorovich's.

7.6.1.4 The Benamou–Brenier Dynamic Formulation

When the cost is quadratic, the optimal transport cost admits a dynamic representation due to Benamou and Brenier. Let $\rho(x, t)$ denote a time-dependent density and $v(x, t)$ a velocity field. The evolution must satisfy the continuity equation

$$\partial_t \rho(x, t) + \nabla \cdot (\rho(x, t)v(x, t)) = 0, \quad \rho(\cdot, 0) = \mu_0, \quad \rho(\cdot, 1) = \mu_1.$$

The squared 2-Wasserstein distance between μ_0 and μ_1 can then be expressed as

$$\frac{1}{2} W_2^2(\mu_0, \mu_1) = \inf_{\rho, v} \int_0^1 \int_{\mathbb{R}^d} \frac{1}{2} \|v(x, t)\|^2 \rho(x, t) dx dt,$$

subject to the continuity constraint. This formulation highlights the geometric interpretation of W_2 as the minimal kinetic energy required to continuously deform μ_0 into μ_1 . It also establishes Wasserstein spaces as geodesic metric spaces, with geodesics given by displacement interpolations of probability measures.

7.6.1.5 Continuity and Liouville Equations

Let $X \subseteq \mathbb{R}^n$ and consider a time horizon $[0, T]$. Given a (possibly time-dependent) velocity field $v : X \times [0, T] \rightarrow \mathbb{R}^n$ and a family of probability measures $(\mu_t)_{t \in [0, T]}$ on X , the continuity equation expresses conservation of mass:

$$\partial_t \mu_t + \nabla_x \cdot (v \mu_t) = 0. \tag{7.11}$$

In general, μ_t need not admit a smooth density, so (7.11) is understood in the following weak sense: for all test functions $\varphi \in C_c^1(X \times [0, T])$,

$$\int_{X \times [0, T]} \left(\partial_t \varphi(x, t) + \nabla_x \varphi(x, t) \cdot v(x, t) \right) d\mu(x, t) = \int_X \varphi(x, T) d\mu_T(x) - \int_X \varphi(x, 0) d\mu_0(x), \quad (7.12)$$

where $\mu \in \mathcal{M}_+(X \times [0, T])$ is the occupation measure defined by $d\mu(x, t) = d\mu_t(x) dt$. Equation (7.12) is the weak form of (7.11) and ensures mass conservation without requiring smooth densities.

If the dynamics of the state $x(t)$ are governed by

$$\dot{x}(t) = f(x(t), u(t)), \quad u(t) \in U, \quad (7.13)$$

then the continuity equation reduces to the *Liouville equation*,

$$\partial_t \mu_t + \nabla_x \cdot (f(x, u) \mu_t) = 0, \quad (7.14)$$

expressing that the distribution of states evolves deterministically along the flow induced by f . In weak form, this reads: for all $\psi \in C_c^1(X \times [0, T])$,

$$\int_{X \times [0, T]} \left(\partial_t \psi(x, t) + \nabla_x \psi(x, t) \cdot f(x, u) \right) d\mu(x, t) = \int_X \psi(x, T) d\mu_T(x) - \int_X \psi(x, 0) d\mu_0(x). \quad (7.15)$$

These weak formulations are the general backbone for all measure-based descriptions of dynamical systems and optimal transport problems. Specializations (e.g. the Benamou–Brenier transport PDE or autonomous Liouville equations) are obtained as particular cases.

7.6.2 Occupation Measure and LMI Formulation for the Generalized Problem

By means of the following result it is possible to recast Problem 7.2.3 via occupation measures.

Theorem 7.6.1

Consider the optimization problem in the form (7.7). The problem is equivalent to the following optimization program over occupation measures:

$$\mathcal{P}_\mu^\star = \min_{\mu_{0,u}, \mu_u, \mu_{0,s}, \mu_s} \frac{1}{2} \langle \|v\|^2, \eta \rangle \quad (7.16a)$$

$$\text{s.t.} \quad \frac{\partial \eta}{\partial \tau} + \nabla_x \cdot (v\eta) = 0 \quad (7.16b)$$

$$\frac{\partial \mu_u}{\partial t} + \nabla_x \cdot (f\mu_u) = 0 \quad (7.16c)$$

$$\frac{\partial \mu_s}{\partial t} + \nabla_x \cdot (f\mu_s) = 0 \quad (7.16d)$$

$$\eta(\tau = 0) = \mu_{0,u} \quad (7.16e)$$

$$\eta(\tau = 1) = \mu_{0,s} \quad (7.16f)$$

$$\mu_u(t = 0) = \mu_{0,u} \quad (7.16g)$$

$$\mu_s(t = 0) = \mu_{0,s} \quad (7.16h)$$

$$\text{spt}(\mu_u(t = T)) \subseteq \mathcal{U} \quad (7.16i)$$

$$\text{spt}(\mu_s(t = T)) \subseteq \mathcal{S} \quad (7.16j)$$

$$\mu_{0,u} \in \Psi(X_0), \quad (7.16k)$$

$$\mu_{u,s} \in \mathcal{M}_+(X \times [0, T]), \quad (7.16l)$$

$$\mu_{0,s} \in \Psi(X_0), \quad (7.16m)$$

$$\eta \in \mathcal{M}_+(X \times [0, 1] \times \mathcal{V}), \quad (7.16n)$$

$$\mu_s \in \mathcal{M}_+(X \times [0, T]). \quad (7.16o)$$

Moreover, it holds that $\mathcal{P}^* = \mathcal{P}_\mu^*$. $\square$

Proof. The argument follows the same structure as the proof of Theorem 7.2.1, with the only difference that both the factual and the counterfactual initial laws are now decision variables.

Consider any feasible pair $(\mu_{0,u}, \mu_{0,s})$ for Problem 7.2.3. By the Benamou–Brenier formula, there exist a narrowly continuous curve $\rho(\cdot, \tau)$ and Borel velocity $w(\cdot, \tau)$ connecting $\mu_{0,u}$ to $\mu_{0,s}$ and satisfying

$$\int_0^1 \int_X \|w\|^2 \rho \, dx \, d\tau = W_2^2(\mu_{0,u}, \mu_{0,s}).$$

Define the transport occupation measure η as in (7.6) and set $v = w$. Then the weak continuity equation (Eq. (7.12)) holds with boundaries $\eta(\tau = 0) = \mu_{0,u}$, $\eta(\tau = 1) = \mu_{0,s}$, so constraints (7.16b)–(7.16f) are satisfied.

Then define μ_u as the state occupation measure induced by the autonomous flow Φ_t with initial law $\mu_u(t = 0) = \mu_{0,u}$. By the superposition principle, μ_u satisfies the weak Liouville equation (Eq. (7.15)) with boundary $\mu_u(t = 0) = \mu_{0,u}$ and terminal marginal $\mu_u(t = T) = (\Phi_T)_\# \mu_{0,u}$. The constraint (7.16i) then follows from (7.7b). Analogously, define μ_s from the flow with initial law $\mu_{0,s}$; it satisfies the Liouville equation with $\mu_s(t = T) = (\Phi_T)_\# \mu_{0,s}$, and the terminal support condition (7.16j) follows from (7.7c).

Thus $(\mu_{0,u}, \mu_u, \mu_{0,s}, \mu_s, \eta)$ is feasible for (7.16) with cost $\frac{1}{2} W_2^2(\mu_{0,u}, \mu_{0,s})$. Taking the infimum over feasible pairs $(\mu_{0,u}, \mu_{0,s})$ shows $\mathcal{P}_\mu^* \leq \mathcal{P}^*$.

Conversely, let $(\mu_{0,u}, \mu_u, \mu_{0,s}, \mu_s, \eta)$ be feasible for (7.16). By the weak continuity equation with boundary data $\eta(\tau = 0) = \mu_{0,u}$, $\eta(\tau = 1) = \mu_{0,s}$, the pair (η, v) is an admissible interpolation between $\mu_{0,u}$ and $\mu_{0,s}$ with kinetic action $\frac{1}{2} \langle \|v\|^2, \eta \rangle$. By the variational characterization of W_2 ,

$$\frac{1}{2} W_2^2(\mu_{0,u}, \mu_{0,s}) \leq \frac{1}{2} \langle \|v\|^2, \eta \rangle.$$

By the Liouville equations, the marginals satisfy $\mu_u(t = T) = (\Phi_T)_\# \mu_{0,u}$ and $\mu_s(t = T) = (\Phi_T)_\# \mu_{0,s}$, and by (7.16i)–(7.16j) their supports are included in $\mathcal{U}$ and $\mathcal{S}$, respectively. Hence $(\mu_{0,u}, \mu_{0,s})$ is feasible for Problem (7.7), and taking infima gives $\mathcal{P}^* \leq \mathcal{P}_\mu^*$. $\square$

Problem in the form (7.16) accounts for additional constraints in the occupation measure formulation related with the additional decision variables $\mu_{0,u}, \mu_u$. Specifically, constraints (7.16c) induces Liouville equation for the measure μ_u , whereas constraints (7.16g) and (7.16i) ensures that the mass of measure $\mu_{0,u}$ is entirely transferred into a measure on the unsafe set $\mathcal{U}$ by means of the occupation measure μ_u .

The LMI formulation for the generalized measure-level problem with unknown initial unsafe distribution is provided below.

Problem 7.6.1 (LMI Formulation for the Generalized Measure-level Problem)

The problem in (7.16) can be formulated in the truncated moment space and solved as follows.

$$p_d^* = \min_{y_{0,u}, y_u, y_{T,u}, y_{0,s}, y_s, y_{T,s}, y_\eta} \frac{1}{2} y_{002,\eta} \quad (7.17a)$$

$$s.t. \quad Liou_{\alpha,\beta}(y_u, y_{0,u}, y_{T,u}) \quad \forall (\alpha, \beta) : |\alpha| + \beta \leq 2d, \quad (7.17b)$$

$$Liou_{\alpha,\beta}(y_\eta, y_{0,s}, y_{0,u}) \forall(\alpha, \beta) : |\alpha| + \beta \leq 2d, \quad (7.17c)$$

$$Liou_{\alpha,\beta}(y_s, y_{0,s}, y_{T,s}) \forall(\alpha, \beta) : |\alpha| + \beta \leq 2d, \quad (7.17d)$$

$$\mathbb{M}_d(y_{0,u}) \succeq 0, \quad (7.17e)$$

$$\mathbb{M}_{d-d_{0u,\kappa}}(w_{0u,\kappa} y_{0,u}) \succeq 0, \quad \kappa = 1, \dots, N_0, \quad (7.17f)$$

$$\mathbb{M}_d(y_{0,s}) \succeq 0, \quad (7.17g)$$

$$\mathbb{M}_{d-d_{0s,k}}(w_{0s,k} y_{0,s}) \succeq 0, \quad k = 1, \dots, N_0, \quad (7.17h)$$

$$\mathbb{M}_d(y_u) \succeq 0, \quad (7.17i)$$

$$\mathbb{M}_{d-d_{X,i}}(w_i y_u) \succeq 0, \quad i = 1, \dots, N_X, \quad (7.17j)$$

$$\mathbb{M}_{d-1}(t(T-t) y_u) \succeq 0, \quad (7.17k)$$

$$\mathbb{M}_d(y_{T,u}) \succeq 0, \quad (7.17l)$$

$$\mathbb{M}_{d-d_{U,\ell}}(s_\ell y_{T,u}) \succeq 0, \quad \ell = 1, \dots, N_U, \quad (7.17m)$$

$$\mathbb{M}_d(y_s) \succeq 0, \quad (7.17n)$$

$$\mathbb{M}_{d-d_{X,i}}(w_i y_s) \succeq 0, \quad i = 1, \dots, N_X, \quad (7.17o)$$

$$\mathbb{M}_{d-1}(t(T-t) y_s) \succeq 0, \quad (7.17p)$$

$$\mathbb{M}_d(y_{T,s}) \succeq 0, \quad (7.17q)$$

$$\mathbb{M}_{d-d_{S,\ell}}(s_\ell y_{T,s}) \succeq 0, \quad \ell = 1, \dots, N_S, \quad (7.17r)$$

$$\mathbb{M}_d(y_\eta) \succeq 0, \quad (7.17s)$$

$$\mathbb{M}_{d-d_{X,i}}(g_i y_\eta) \succeq 0, \quad i = 1, \dots, N_X, \quad (7.17t)$$

$$\mathbb{M}_{d-1}(\tau(1-\tau) y_\eta) \succeq 0, \quad (7.17u)$$

$$\mathbb{M}_{d-d_{V,j}}(h_j y_\eta) \succeq 0, \quad j = 1, \dots, N_V. \quad (7.17v)$$

□

The problem interpretation is similar to the previous discussed case (7.8), with additional moment decision variables and constraints accordingly with the extended formulation in (7.16). Specifically, the additional moment sequences $y_{0,u}, y_u$ represent measures $\mu_{0,u}, \mu_u$ in the moment space whereas constraint (7.17b) represent the continuity equation in (7.16c) while also encoding the marginal constraints. Moment and Localizing matrix constraints in (7.17e)-(7.17v) enforce measure support constraints in (7.16i)-(7.16o) in moment space.

References

- [1] Guidotti, R., “Counterfactual explanations and how to find them: Literature review and benchmarking,” *Data Mining and Knowledge Discovery*, vol. 38, no. 5, pp. 2770–2824, 2024.
- [2] Wachter, S., Mittelstadt, B., and Russell, C., “Counterfactual explanations without opening the black box: Automated decisions and the GDPR,” *Harv. JL & Tech.*, vol. 31, p. 841, 2017.
- [3] Carlevaro, A., Lenatti, M., Paglialonga, A., and Mongelli, M., “Counterfactual building and evaluation via explainable support vector data description,” *IEEE Access*, vol. 10, pp. 60 849–60 861, 2022.
- [4] Villani, C. et al., *Optimal transport: old and new*. Springer, 2008, vol. 338.
- [5] Santambrogio, F., “Optimal transport for applied mathematicians,” 2015.
- [6] Frieesecke, G., *Optimal Transport: A Comprehensive Introduction to Modeling, Analysis, Simulation, Applications*. SIAM, 2024.
- [7] Henrion, D., Korda, M., and Lasserre, J. B., *Moment-sos Hierarchy, The: Lectures In Probability, Statistics, Computational Geometry, Control And Nonlinear Pdes*. World Scientific, 2020, vol. 4.

- [8] Henrion, D., Lasserre, J. B., and Savorgnan, C., “Nonlinear optimal control synthesis via occupation measures,” in *2008 47th IEEE Conference on Decision and Control*, IEEE, 2008, pp. 4749–4754.
- [9] Lasserre, J. B., Henrion, D., Prieur, C., and Trélat, E., “Nonlinear optimal control via occupation measures and lmi-relaxations,” *SIAM journal on control and optimization*, vol. 47, no. 4, pp. 1643–1666, 2008.
- [10] Lasserre, J. B., *Moments, positive polynomials and their applications*. World Scientific, 2009, vol. 1.
- [11] Lasserre, J. B., Pauwels, E., and Putinar, M., *The Christoffel–Darboux kernel for data analysis*. Cambridge University Press, 2022, vol. 38.
- [12] Marx, S., Pauwels, E., Weisser, T., Henrion, D., and Lasserre, J. B., “Semi-algebraic approximation using christoffel–darboux kernel,” *Constructive Approximation*, vol. 54, no. 3, pp. 391–429, 2021.
- [13] Murray, J. D., *Mathematical biology: I. An introduction*. Springer Science & Business Media, 2007, vol. 17.
- [14] Chen, Y., Georgiou, T. T., and Pavon, M., “Optimal transport in systems and control,” *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 4, no. 1, pp. 89–113, 2021.
- [15] Taghvaei, A. and Mehta, P. G., “An optimal transport formulation of the ensemble kalman filter,” *IEEE Transactions on Automatic Control*, vol. 66, no. 7, pp. 3052–3067, 2020.
- [16] Elvander, F., Haasler, I., Jakobsson, A., and Karlsson, J., “Multi-marginal optimal transport using partial information with applications in robust localization and sensor fusion,” *Signal Processing*, vol. 171, p. 107474, 2020.
- [17] Haasler, I., Ringh, A., Chen, Y., and Karlsson, J., “Multimarginal optimal transport with a tree-structured cost and the schrodinger bridge problem,” *SIAM Journal on Control and Optimization*, vol. 59, no. 4, pp. 2428–2453, 2021.
- [18] Zhai, H., Egerstedt, M., and Zhou, H., “Path planning in unknown environments using optimal transport theory,” *arXiv preprint arXiv:1909.11235*, 2019.
- [19] Krishnan, V. and Martínez, S., “Distributed optimal transport for the deployment of swarms,” in *2018 IEEE Conference on Decision and Control (CDC)*, IEEE, 2018, pp. 4583–4588.
- [20] Bergman, R. N., Ider, Y. Z., Bowden, C. R., and Cobelli, C., “Quantitative estimation of insulin sensitivity,” *American Journal of Physiology-Endocrinology And Metabolism*, vol. 236, no. 6, E667, 1979.
- [21] De Paola, P. F., Paglialonga, A., Palumbo, P., Keshavjee, K., Dabbene, F., and Borri, A., “The long-term effects of physical activity on blood glucose regulation: A model to unravel diabetes progression,” *IEEE Control Systems Letters*, vol. 7, pp. 2916–2921, 2023. DOI: [10.1109/LCSYS.2023.3290774](https://doi.org/10.1109/LCSYS.2023.3290774).
- [22] De Paola, P. F., Borri, A., Dabbene, F., Keshavjee, K., Palumbo, P., and Paglialonga, A., “Modeling the cumulative benefits of regular physical activity on type 2 diabetes progression,” *Computers in Biology and Medicine*, vol. 198, p. 111194, 2025. DOI: <https://doi.org/10.1016/j.compbiomed.2025.111194>.

Conclusions

Chapter 8

Conclusions

The work presented in this Thesis contributes to the field through two entwined research directions. The first one, discussed in Part I, proposed a novel mathematical model suitable to capture the cumulative long-term benefits promoted by regular exercise on T2D progression. Building upon this model, an original framework for model-based control is then proposed, aiming to regulate disease progression through exercise interventions. Specifically:

- Chapter 2 introduced the proposed model, describing how it is built in the two-timescale structure and defining its state-space representation. Moreover, Chapter 2 has shown that modeling frameworks omitting the explicit representation of IL-6 in the regulation of beta-cell dynamics and insulin sensitivity are inadequate to reproduce the impact of exercise on the progression of the disease, even when the fast, short-term dynamics of exercise are accounted for in the model structure. Furthermore, Chapter 2 has assessed the model validity by means of a corpus of simulations aimed at reproducing a broad range of scenarios.
- Chapter 3 has developed the formal analysis of the proposed model and investigated its qualitative behavior, discussing its steady-state conditions and some theoretical properties.
- Chapter 4 has leveraged the proposed model to design and develop MPC-based control techniques. Both state- and output- feedback control approaches have been investigated yielding comparable results. The goal was to lay the ground for the development of personalized intervention protocols, in which exercise duration (or, equivalently, the intensity or the period of the training) are optimally tailored to the individual's physiological characteristics and disease course to effectively regulate glucose homeostasis and slow disease progression.

The second research direction, discussed in Part II, has been focused on the proposal of a control system framework for counterfactuals. More in detail, two approaches have been investigated:

- Chapter 6 has developed the first approach leveraging tools from optimal control and occupation measure theory to derive counterfactual trajectories that remain dynamically consistent with the system model while satisfying feasibility and safety constraints.
- Chapter 7 involved a second approach that reformulates the problem within the optimal transport paradigm, interpreting counterfactual reasoning as the minimal redistribution of probability mass between initial distributions associated with unsafe and safe evolution.

The proposed frameworks represent initial steps toward a deeper integration of mathematical modeling and control theory aimed at advancing personalized and explainable approaches to healthcare. Indeed, toward a comprehensive translational impact, some limitations should be addressed in future developments. In this regard, the model presented in Part I has not yet been validated on clinical or population data. This limits its immediate translational applicability and highlights the need for parameter calibration and model refinement based on empirical evidence.

Moreover, although the model includes the anti-inflammatory effects of exercise, this is represented through a compact mechanistic structure. In reality, low-grade systemic inflammation involves a multifaceted network of metabolic interactions that evolve over

different time scales and vary across individuals. The adopted modeling approach aims to preserve model tractability, particularly for control design, but inherently neglects some layers of biological complexity. Additionally, the model does not explicitly account for meal dynamics or the energy intake–expenditure balance. As for the model-based control, further validation on more comprehensive physiological simulators, such as the UVA/Padova simulator, would offer stronger evidence of clinical plausibility and robustness to unmodeled dynamics, inter-individual variability, and behavioral uncertainty.

Whereas, for what concerns the second research direction presented in Part II, the proposed counterfactual reasoning framework, grounded in optimal control and optimal transport formulations, offers a rigorous system-theoretic foundation via occupation measures and moment-SOS hierarchy. While moment-SOS methods are powerful tools and enable the convex reformulation of nonlinear, nonconvex control problems, they are prone to computational challenges, mainly due to the limits in the scalability to high-dimensional systems or to higher-order relaxations that may be required for tighter approximations. Moreover, moment-SOS methods typically require all problem data, such as dynamics, cost functions, and constraints, to be polynomial, or at least approximated by polynomials over compact sets.

Future research should be focused on overcoming the discussed limitations and should pursue the refinement of the developed model through tighter coupling with experimental and clinical data, the extension of control strategies to account for additional source of inter-individual variability and uncertainty and the validation of the control law on higher-dimensional models. Furthermore, by integrating the control-based counterfactual reasoning framework with artificial intelligence methodologies, future research could allow to inform data-driven algorithms with principled knowledge of system dynamics and control, thereby enhancing their interpretability, reliability, and capacity to generate physiologically plausible predictions and interventions.

Of course, the results collected in this Thesis only scratch the surface of the many scientific and technological challenges that lie ahead in the modeling, control, and counterfactual analysis of complex physiological systems. Together, the techniques proposed in this Thesis establish a unified and rigorous framework that connects modeling, control theory and causal reasoning, paving the way toward integrated, explainable, and control-aware decision support tools for personalized and preventive medicine. By fostering a synergy between theory and data, these directions hold the potential to transform the predictive and prescriptive understanding of chronic diseases such as T2D, ultimately contributing to the development of precision medicine strategies grounded in physiological knowledge and control theory principles.

Appendices

Appendix A

Preliminaries

A.1 Model description

Model differential equations:

$$\frac{dG}{dt} = R_0 - (Eg_0 + S_I \cdot I)G + \frac{W}{Vol_G}(G_{\text{prod}} - G_{\text{up}}), \quad (\text{A.1a})$$

$$\frac{dI}{dt} = \beta \frac{ISR}{V} - kI - I_e \text{Heav}(I), \quad (\text{A.1b})$$

$$\frac{d\beta}{dt} = \frac{(\bar{P}_{ISR} - \bar{A}_{MG})\beta}{\tau_\beta}, \quad (\text{A.1c})$$

$$\frac{d\gamma}{dt} = -\frac{(\gamma - \gamma_{\infty, G})}{\tau_\gamma}, \quad (\text{A.1d})$$

$$\frac{d\sigma}{dt} = -\frac{(\sigma - \sigma_{ISR_{M\infty}})}{\tau_\sigma}, \quad (\text{A.1e})$$

$$\frac{dS_I}{dt} = -\frac{(S_I - S_{I, \text{target}}) \left(1 - \zeta_3 \frac{V_l^n}{k_{n, S_I}^n + V_l^n}\right)}{\tau_{SI}}, \quad (\text{A.1f})$$

$$\frac{dV_l}{dt} = IL_6 - k_s V_l, \quad (\text{A.1g})$$

$$\frac{dPVO_2}{dt} = -0.8PVO_2 + 0.8u, \quad (\text{A.1h})$$

$$\frac{dG_{\text{prod}}}{dt} = a_1 PVO_2 - a_2 G_{\text{prod}}, \quad (\text{A.1i})$$

$$\frac{dG_{\text{up}}}{dt} = a_3 PVO_2 - a_4 G_{\text{up}} \quad (\text{A.1j})$$

$$\frac{dI_e}{dt} = a_5 PVO_2 - a_6 I_e, \quad (\text{A.1k})$$

$$\frac{dIL_6}{dt} = SR_{\text{ex}} PVO_2 - k_m IL_6, \quad (\text{A.1l})$$

State-dependent functions:

$$\text{Heav}(I) = \begin{cases} 0, & I \leq 0, \\ 1, & I > 0 \end{cases} \quad (\text{A.2a})$$

$$M_G = \frac{G^{k_M}}{G^{k_M} + \alpha_M^{k_M}} \quad (\text{A.2b})$$

$$ISR = \sigma \frac{(M_G + \gamma)^{k_{\text{ISR}}}}{(M_G + \gamma)^{k_{\text{ISR}}} + \alpha_{\text{ISR}}^{k_{\text{ISR}}}} \quad (\text{A.2c})$$

$$\bar{P}_{ISR} = P_{\text{max}} \frac{ISR^{k_P}}{ISR^{k_P} + \alpha_P^{k_P}} \left(1 + \frac{\zeta_1 V_l^2}{k_{n, P}^2 + V_l^2}\right) \quad (\text{A.2d})$$

$$\bar{A}_{MG} = (A_{\text{max}} \frac{M_G^{k_A}}{M_G^{k_A} + \alpha_A^{k_A}} + A_b) \left(1 - \frac{\zeta_2 V_l^2}{k_{n, A}^2 + V_l^2}\right) \quad (\text{A.2e})$$

Symbol	Description	Unit
G	Plasma glucose concentration	[mg/dl]
I	Serum insulin concentration	[μ U/ml]
β	Beta cell mass	[mg]
γ	Shift in glucose response	(dimensionless)
σ	Insulin secretion capability	[μ U/ μ g/d]
S_I	Insulin sensitivity	[ml/ μ U/d]
PVO_2	Sovrabasal oxygen consumption	(dimensionless)
G_{prod}	Rate of incremental hepatic glucose production promoted by exercise	[mg/kg/min]
G_{up}	Rate of incremental glucose uptake by working tissues promoted by exercise	[mg/kg/min]
I_e	Rate of incremental insulin removal induced by exercise	[μ U/ml/min]
IL_6	Concentration of IL-6 in the muscle compartment	[pg/ml]
V_l	Integral effect of physical activity representing the long-term effect of exercise	[(pg/ml)*min]

Table A.1: Description of the model state variables

Symbol	Description	Unit
$\text{Heav}(I)$	Heaviside function of I	(dimensionless)
M_G	Metabolic function	(dimensionless)
ISR	Insulin secretion rate	[μ U/ μ g/d]
P_{ISR}	Beta cell proliferation	(dimensionless)
A_{M_G}	Beta cell apoptosis	(dimensionless)
$\gamma_{\infty,G}$	Steady-state for dose response shift	(dimensionless)
$\sigma_{\text{ISR}_{M_\infty}}$	Steady-state for σ	[μ U/ μ g/d]
$\sigma_{\text{ISR}_\infty}$	ISR component of $\sigma_{\text{ISR}_{M_\infty}}$	[μ U/ μ g/d]
σ_{M_∞}	M component of $\sigma_{\text{ISR}_{M_\infty}}$	[μ U/ μ g/d]

Table A.2: Description of the state - depending functions

$$\gamma_{\infty,G} = \frac{\gamma_{\text{max}}}{1 + \exp\left(-\frac{(G-\gamma_s)}{\gamma_n}\right)} - \gamma_0 \quad (\text{A.2f})$$

$$M_\sigma = \frac{(G - G_{\sigma_s})^{k_M}}{\alpha_M^{k_M} + (G - G_{\sigma_s})^{k_M}} \quad (\text{A.2g})$$

$$ISR_\sigma = \sigma \frac{(M_\sigma + \gamma)^{k_{\text{ISR}}}}{(M_\sigma + \gamma)^{k_{\text{ISR}}} + \alpha_{\text{ISR}}^{k_{\text{ISR}}}} \quad (\text{A.2h})$$

$$\sigma_{\text{ISR}_\infty} = \frac{\sigma_{\text{ISR}_{\text{max}}}}{1 + \sigma_{\text{ISR}k} \exp\left(-\frac{(ISR_\sigma - \sigma_{\text{ISR}_s})}{\sigma_{\text{ISR}n}}\right)} \quad (\text{A.2i})$$

$$\sigma_{M_\infty} = 1 - \frac{\sigma_{M_{\text{max}}}}{1 + \sigma_{Mk} \exp\left(-\frac{(M_\sigma - \sigma_{Ms})}{\sigma_{Mn}}\right)} \quad (\text{A.2j})$$

$$\sigma_{\text{ISR}_{M_\infty}} = \sigma_{\text{ISR}_\infty} \sigma_{M_\infty} + \sigma_b \quad (\text{A.2k})$$

Parameter	Value	Unit
$\bar{u}$	variable in [0,92] %	dimensionless
ks	$\frac{-\log(0.8)}{8 \times 7 \times 1440}$	1/min
W	70	kg
$S_{I,target}$	0.18	[ml/ μ U/d]
SR_{ex}	0.045	[pg/ml/min]
k_m	0.004	[1/min]
Vol_G	117	[dl]
a_1	0.00158	[mg/kg/min ²]
a_2	0.056	[1/min]
a_3	0.00195	[mg/kg/min ²]
a_4	0.0485	[1/min]
a_5	0.00125	[mg/kg/min ²]
a_6	0.075	[1/min]
τ_{SI}	150	[d]
G_{σ_s}	75	[mg/dl]
R_0	864	[mg/dl/d]
Eg_0	1.44	[1/d]
V	5	[L]
k	700	[1/d]
τ_γ	2.14	[d]
τ_σ	249.9	[d]
τ_β	8570	[d]
k_M	2	dimensionless
α_M	150	[mg/dl]
k_{ISR}	2	dimensionless
α_{ISR}	1.2	dimensionless
P_{max}	4.55	dimensionless
k_P	4	dimensionless
α_P	41.77	[μ U/ μ g/d]
A_{max}	9	dimensionless
k_A	6	dimensionless
α_A	0.44	dimensionless
A_b	0.8	dimensionless
γ_{max}	0.2	dimensionless
γ_s	99.9	dimensionless
γ_n	1	dimensionless
γ_0	0.1	dimensionless
$\sigma_{ISR_{max}}$	600	[μ U/ μ g/d]
σ_{ISR_s}	0.1	dimensionless
σ_{ISR_n}	0.1	dimensionless
σ_{ISR_k}	1	dimensionless
σ_b	3	[μ U/ μ g/d]
$\sigma_{M_{max}}$	1	dimensionless
σ_{M_s}	0.2	dimensionless
σ_{M_n}	0.02	dimensionless
σ_{M_k}	0.2	dimensionless
ζ_1	10^{-4}	dimensionless
ζ_2	10^{-4}	dimensionless
ζ_3	1.4*	dimensionless
$k_{n,P}$	10^6	[(pg/ml)*min]
$k_{n,A}$	10^6	[(pg/ml)*min]
k_{n,S_I}	$5 \cdot 10^6$ *	[(pg/ml)*min]
n	1*	dimensionless

Table A.3: Parameter values used in the model.

*NOTE: Model simulations are performed with parameters as in Table A.3. Under this parameter configuration, the model may exhibit an unstable behavior of S_I dynamics for very high, unrealistic, level of exercise (equivalent to $u_{eq} \gg 2.5$). To solve this issue it is possible to set $\zeta_3 = 1$, $k_{n,S_I} = 2.7 \cdot 10^6$ and $n = 1.28$, to avoid this behavior and

preserve the same simulation results.

Appendix B

Preliminaries on Occupation Measures and Moment-SOS Hierarchy

B.0.1 Mathematical background

$\mathbb{R}^n$ denotes the n -dimensional real euclidean space and $\mathbb{R}[x]$ the ring of polynomials in the variable x . $C(X)$ is the vector space of continuous functions and $\mathcal{M}(X)$ constitutes its dual space of Borel measures on X , with X subset of $\mathbb{R}^n$.

A Borel measure on $\mathbb{R}^n$ is a measure defined on the σ -algebra generated by the open subsets of $\mathbb{R}^n$.

$C(X)$ and $\mathcal{M}(X)$ are paired via the continuous linear functional $Tf = \int_X f(x)d\mu(x)$. $\mathcal{M}_+(X)$ and $C_+(X)$ denotes respectively the nonnegative cones of $\mathcal{M}(X)$ and $C(X)$. $C^1(X) \subset C(X)$ is the set of continuous functions with continuous first derivatives. A set $X \subseteq \mathbb{R}^n$ of the form $X = \{x \mid s_k(x) \geq 0\}$ is said to be basic semialgebraic. The quadratic module $Q[s]$ defined by the constraints induced by a given basic semialgebraic set is the set of polynomials [1]

$$Q[s] = \{\sigma_0(x) + \sum_{i=1}^{N_c} \sigma_i(s)s_i(x)\},$$

with components σ being SOS:

$$\sigma_i(x) \in \Sigma[x] \quad \forall i = 0, \dots, N_c.$$

A given basic semialgebraic set K is compact if there exists a constant $0 \leq R \leq \infty$ such that the set is contained in the ball $R \leq \|x\|_2^2$. The set K is Archimedean if the polynomial $R - \|x\|_2^2$ is included in $Q[s]$. Any compact set can be reduced to Archimedean adding a redundant ball constraint $R - \|x\|_2^2 \geq 0$ into the defining constraints.

The Putinar's Positivstellensatz states that every polynomial $p(x)$ positive on an Archimedean set admits a representation [1]:

$$p(x) = \sigma_0(x) + \sum_{i=1} \sigma_i(s)s_i(x)$$

$$\sigma_0(x) \in \Sigma[x], \quad \sigma_i(x) \in \Sigma[x].$$

A polynomial $\pi(x)$ can be expressed in the monomial basis as $\sum_{\alpha} \pi_{\alpha} x^{\alpha}$ and the degree of $\pi(x)$ ($\deg(\pi(x))$) is considered to be the maximum $|\alpha| = \sum_i |\alpha_i|$.

The indicator function of set $A \subseteq X$ is a function $\mathbb{I}_A : X \rightarrow \{0, 1\}$ such that $\mathbb{I}_A(x) = 1$ if $x \in A$ and $\mathbb{I}_A(x) = 0$ otherwise and the measure of a set A with respect to $\mu \in \mathcal{M}(X)$ is defined as $\mu(A) = \langle \mathbb{I}_A(x), \mu \rangle$.

The α -order moment of a measure μ is the scalar quantity $y_{\alpha} = \langle x^{\alpha}, \mu \rangle$. For $\alpha = 1$ we get the mass of the measure, i.e. μ is $\langle 1, \mu \rangle$. The measure μ is a probability measure if it holds that $\langle 1, \mu \rangle = 1$.

The support of a measure μ is the set of all points $x \in X$ such that every open neighborhood N_x of x has $\mu(N_x) > 0$. We denote the support of a measure μ as $\text{spt}(\mu)$.

A Dirac delta measure or *atomic* measure $\delta_{x=x'} \in \mathcal{M}(X)$ (also referred to as $\delta_{x'}$ in the following) is a measure supported on a single point (i.e., *atom*) $x = x'$.

X

A r -atomic measure is a measure μ such that there exist scalars $\beta_i > 0$ and points $x_i \in X, \forall i = 1, \dots, r$ such that $\mu = \sum_{i=1}^r \beta_i \delta_{x_i}$. The atoms of the measure μ are given by the support points $\{x_i\}_{i=1}^r$.

The product $\mu_1 \otimes \mu_2$ denotes the product of the measures μ_1 and μ_2 that satisfies $(\mu_1 \otimes \mu_2)(A \times B) = \mu_1(A)\mu_2(B)$.

Let X and Y be measurable spaces, μ be a measure on X , and $T : X \rightarrow Y$ be a measurable map on metric spaces and let μ be a probability measure on X . The *pushforward* of μ through T , denoted $T_{\#}\mu$, is the measure on Y defined by

$$(T_{\#}\mu)(B) := \mu(T^{-1}(B)),$$

for all measurable $B \subseteq Y$.

Let $(X, \mathcal{B}(X))$ and $(Y, \mathcal{B}(Y))$ be measurable spaces, and let μ be a measure on the product space $(X \times Y, \mathcal{B}(X \times Y))$. The *marginal measures* of μ are the measures μ_X on X and μ_Y on Y defined by

$$\mu_X(A) := \mu(A \times Y), \quad \forall A \in \mathcal{B}(X),$$

$$\mu_Y(B) := \mu(X \times B), \quad \forall B \in \mathcal{B}(Y).$$

A given linear operator $H : X \rightarrow X$ is endowed with with a unique adjoint operator H^* such that $H^* = H$, that is $\langle Hv, \mu \rangle = \langle v, H^*\mu \rangle, \forall v(t, x) \in C^1([0, T] \times X) \in X, \forall \mu \in \mathcal{M}(X)$. $\mathcal{L}_f$ will denote throughout the text the Lie derivative of a general function with respect to the vector field f .

B.0.2 Occupation Measures

Occupation measures are great tools for solving nonlinear optimization problem for dynamical systems[2], [3], [4]. Let assume that $X \subseteq \mathbb{R}^n$ is the state space of a dynamical system, and let $x(t | x_0)$ denote the trajectory of the system starting from initial condition x_0 over the time interval. The occupation measure μ of a set $A \times B$ with $A \subseteq [0, T]$ and $B \subseteq X$ referred to the trajectory $x(t | x_0)$ is defined as

$$\mu(A \times B | x_0) := \int_0^T \mathbb{I}_{A \times B}(t, x(t | x_0)) dt$$

and can be interpreted as the time spent by the trajectory in the set $A \times B$.

Given a distribution of initial conditions $\mu_0 \in \mathcal{M}(X_0)$ the occupation measure μ is averaged over the initial distribution μ_0 and yields the μ_0 -averaged amount of time that the trajectories dwell in $A \times B$

$$\mu(A \times B) := \int_{X_0} \mu(A \times B | x_0) d\mu_0(x_0).$$

Analogously, the final measure μ_T represents the distribution of the final states as they are transferred by the dynamics according to the initial distribution μ_0 ,

$$\mu_T(B) := \int_{X_0} \mathbb{I}_B(x(T|x_0)) d\mu_0(x_0).$$

Figure B.1 illustrates a graphical representation of the concept of occupation measure. The plotted curves $x(t)$ represent an example of state trajectories of a generic system evolving over time. The blue shaded region intersected by the trajectories corresponds to a subset of the time–state space where the occupation measure is nonzero ($\mu > 0$), as the system spends a positive amount of time in that region. Conversely, the upper

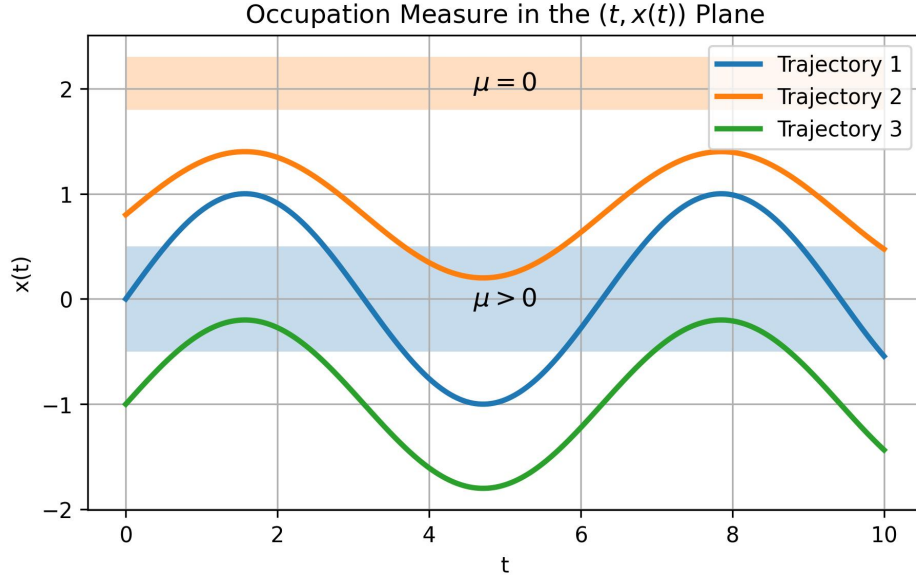


Figure B.1: Intuitive representation of the concept of occupation measure

yellow shaded band is never visited by any trajectory and therefore has zero occupation measure ($\mu = 0$). This graphical representation highlights that the occupation measure assigns mass only to regions of the space-time domain effectively explored by the system dynamics. It results that the mass of the occupation measure coincides with the time spent by the trajectories in the considered regions of space time domain. The three measures μ_0, μ, μ_T are linked together by the Liouville's equation

$$\langle v(T, x), \mu_T \rangle = \langle v(0, x), \mu_0 \rangle + \langle \mathcal{L}_f v(t, x), \mu \rangle$$

that admits also a short-hand formulation independent from the general test function $v \in C^1([0, T] \times X)$ by means of the linear adjoint operator $\mathcal{L}^*$,

$$\delta_T \otimes \mu_T = \delta_0 \otimes \mu_0 + \mathcal{L}^* \mu.$$

In measure-based formulations, test functions play a central role in linking the dynamics of the system to the associated occupation measures. A test function is a sufficiently smooth function against which measures are integrated, allowing the behavior of trajectories to be characterized in a weak, integral sense rather than pointwise. In the occupation measures framework, integrating a test function along system trajectories and over time corresponds to integrating the same function with respect to the occupation measure. This equivalence enables the replacement of trajectory-dependent expressions by linear functionals of measures. Within the moment-SOS framework, test functions are typically chosen as polynomials. This choice allows the resulting integrals to be expressed in terms of moments of the occupation measure, leading to linear constraints on the moment sequence and ultimately to a semidefinite programming formulation.

Liouville's equation describes the evolution of the occupation measure associated with a dynamical system and formalizes the conservation of mass induced by the system flow. In particular, it expresses the fact that the distribution of trajectories in the state (or time-state) space evolves consistently with the underlying dynamics, without creation or loss of mass. From a physical and probabilistic viewpoint, the equation enforces that the mass initially assigned to a set of states is transported along system trajectories as time evolves, and can only be redistributed according to the vector field defining the dynamics. Liouville's equation provides a conservation law that governs the evolution of the occupation measure and explicitly encodes the transfer of mass induced by the system dynamics. In this framework, the occupation measure records how mass originating from

the initial distribution is continuously transported through the state space over time along system trajectories. The equation enforces that mass is neither created nor destroyed, but only redistributed according to the vector field, thereby linking the initial measure, the occupation measure, and the terminal measure through a single balance relation.

Every trajectory $x(t|x_0)$ with $x_0 \in X_0$ and $t \in [0, T]$ induces a measure formulation with $\mu_0 = \delta_{x_0}$ and $\mu_T = \delta_{x_T|x_0}$.

B.0.3 Moment-SOS hierarchy

A standard formulation for an infinite-dimensional LP with measure variables $\mu \in \mathcal{M}_+(X)$ is the following [5]

$$\begin{aligned} p^* &= \inf_{\mu \in \mathcal{M}_+(X)} \langle p(x), \mu \rangle \\ \text{s.t.} \quad &\langle a_i(x), \mu \rangle = b_i, \quad \forall i = 1, \dots, I_{\max} \end{aligned} \quad (\text{B.1})$$

Here $p(x) \in C(X)$ is the cost functional and $\langle a_i(x), \mu \rangle = b_i$ are a set of affine constraints, with $a_i \in C(X)$. A dual program can be formulated over the space of nonnegative continuous functions in the form [6]:

$$\begin{aligned} q^* &= \sup_{v \in \mathbb{R}^m} \sum_i b_i v_i \\ \text{s.t.} \quad &p(x) - \sum_i a_i(x) v_i \geq 0, \\ &\forall i = 1, \dots, I_{\max}, x \in X \end{aligned} \quad (\text{B.2})$$

Under the assumption of $p(x), a_i(x)$ polynomials and X semialgebraic, the program (B.1) can be discretized for tractable computation and restated as an infinite dimensional LP in the space of the moments of the measure y_α and in turn truncated in finite dimensional relaxations in the space of moments up to a given relaxation order d , such that $2d \geq \max(\deg(f(x)), \deg(s_i(x)))$ as follows

$$p_d^* = \min_y \sum_{\alpha} p_{\alpha} y_{\alpha} \quad (\text{B.3a})$$

$$\text{s.t.} \quad \sum_{\alpha} a_{i\alpha} y_{\alpha} = b_i, \quad \forall i = 1, \dots, I_{\max} \quad (\text{B.3b})$$

$$\mathbb{M}_d(y) \succeq 0, \quad (\text{B.3c})$$

$$\mathbb{M}_{d-d_k}(s_k y) \succeq 0, \quad \forall k = 1, \dots, N \quad (\text{B.3d})$$

$\mathbb{M}_d(y)$, $\mathbb{M}_{d-d_k}(s_k y)$ (Moment and the Localizing Matrix) enforce constraints (B.3c)-(B.3d) in the space of truncated moments and guarantee that y are sequences of moments representing measure.

The moment matrix $\mathbb{M}_d(y)$ is a square matrix of size

$$\binom{n+d}{d} \times \binom{n+d}{d}$$

where n is the state space dimension and d is the finite relaxation order. Its entries involve moments up to the degree $2d$. As observable, the size of the matrix increases combinatorially with n and d and this entails a rapid increase in the computational complexity of the moment-SOS relaxations as either the dimension d or the relaxation order d increases. The condition

$$\mathbb{M}_d(y) \succeq 0$$

represents a necessary condition for y to be the moment sequence of a positive Borel measure.

The sequence y is flat (or equivalently has a flat extension) if

$$\text{rank}(\mathbb{M}_d(y)) = \text{rank}(\mathbb{M}_{d-1}(y)).$$

The flatness condition implies that there exists an atomic representing measure for the sequence y with $\text{rank}(\mathbb{M}_d(y))$ atoms. In the context of polynomial optimization, flat extensions enable the certification of global optimality of solutions obtained from moment-SOS relaxations [1].

The primal problem in moments admits a dual SOS formulation in the form:

$$q_d^* = \max_{\lambda} \sum_i b_i \lambda_i \quad (\text{B.4a})$$

$$\text{s.t.} \quad p(x) - \sum_i \lambda_i a_i(x) = \sigma_0(x) + \sum_k \sigma_k(x) s_k(x), \quad (\text{B.4b})$$

$$\sigma_0(x) \in \Sigma[x]_{\leq 2d}, \quad (\text{B.4c})$$

$$\sigma_k(x) \in \Sigma[x]_{\leq 2d - [\deg s_i/2]}, \quad k = 1, \dots, N. \quad (\text{B.4d})$$

The dual variables λ_i are Lagrange multipliers associated with the linear moment constraints $\sum_{\alpha} a_{i\alpha} y_{\alpha} = b_i$. The SOS polynomials σ_0, σ_k certify that the polynomial

$$p(x) - \sum_i \lambda_i a_i(x)$$

admits a Putinar-type decomposition on the semialgebraic set $K = \{x : s_k(x) \geq 0\}$, i.e., it is represented as a sum of squares plus weighted sums of squares of the defining polynomials s_k . Consequently, the constraint enforces $p(x) - \sum_i \lambda_i a_i(x) \geq 0$ on K , so the dual maximizes the largest lower bound (compatible with the equality constraints) that can be certified by SOS certificates of degree at most $2d$.

The moment-SOS hierarchy for d -degree solutions to problem yields sequence of increasing lower bounds to the optimum p^* , $p_d^* \leq p_{d+1}^* \leq p_{d+2}^* \dots$, with convergence guaranteed as $d \rightarrow \infty$ under suitable compactness properties of the support set X . Finite convergence occurs when flat extension conditions are satisfied. In this situation, the moment matrix associated with the optimal solution attains a stabilized rank across consecutive relaxation orders, which implies that the truncated moment sequence admits a finitely atomic representing measure. As a consequence, the global optimizer of the original polynomial optimization problem can be exactly recovered from the finite-order relaxation. However, such rank conditions are not generically satisfied and depend on both the problem structure and the geometry of the feasible set; therefore, finite convergence cannot be guaranteed in general and the hierarchy must be interpreted as providing asymptotically exact approximations.

Strong duality holds between the primal moment relaxation (B.3) and its SOS dual formulation. In particular, for any finite relaxation order d , the optimal values of the primal and dual problems coincide, i.e.,

$$p_d^* = q_d^*.$$

This equality follows from the fact that both problems are convex semidefinite programs and that the primal problem admits a strictly feasible point, ensuring Slater's condition. As a consequence, there is no duality gap between the moment and SOS formulations, and both provide equivalent finite-dimensional representations of the same relaxation of the original infinite-dimensional polynomial optimization problem.

In general, when referring to the moment space, we refer to the Liouville's equation in the moment space as $Liou_{\alpha,\beta}(y_{\alpha}, y_{\beta})$, with α, β multi-indices denoting the degree of the monomials involved in the general moment sequences y_{α}, y_{β} .

References

- [1] Miller, J. F., "Safety quantification for nonlinear and time-delay systems using occupation measures," Ph.D. dissertation, Northeastern University, 2023.

-
- [2] Henrion, D., Lasserre, J. B., and Savorgnan, C., “Nonlinear optimal control synthesis via occupation measures,” in *2008 47th IEEE Conference on Decision and Control*, IEEE, 2008, pp. 4749–4754.
 - [3] Henrion, D., Korda, M., and Lasserre, J. B., *Moment-sos Hierarchy, The: Lectures In Probability, Statistics, Computational Geometry, Control And Nonlinear Pdes*. World Scientific, 2020, vol. 4.
 - [4] Miller, J., Henrion, D., and Sznaiier, M., “Peak estimation recovery and safety analysis,” *IEEE Control Systems Letters*, vol. 5, no. 6, pp. 1982–1987, 2021.
 - [5] Lasserre, J. B., *Moments, positive polynomials and their applications*. World Scientific, 2009, vol. 1.
 - [6] Nash, P. and Anderson, E. J., “Linear programming in infinite-dimensional spaces: Theory and applications,” (*No Title*), 1987.

Appendix C

Additional Contribution

C.1 Stability Analysis of the model by Topp

Following the developments in [1], we provide here¹ a characterization of the regions of attraction (ROAs) the model by Topp *et al.* [2], that we have employed for the analysis in Chapter 4. ROAs are defined as the set of initial conditions of a model leading to a specific target set.

The problem of computing ROAs associated with specific target sets is of central importance in control theory. Determining the ROA provides a deeper understanding of the system dynamics and clarifies how control inputs should be designed to steer trajectories toward the desired steady-state condition. Moreover, it serves as a valuable tool for safety analysis, as it defines the boundaries within which the system can be guaranteed to converge to the target set without leaving admissible operating regions. From a practical perspective, knowledge of the ROA enables the design of controllers that not only stabilize the system but also maximize the domain of safe and effective operation, thereby improving both reliability and performance. As most of the biological systems, the model by Topp *et al.* [2] is inherently nonlinear, hence computing the ROAs of this system is not a trivial problem. Here, we are aimed at estimating the ROAs of its equilibrium points by combining two complementary approaches [1]. First, a Monte Carlo simulation is carried out over the space of admissible initial conditions to obtain an empirical estimate of the ROA. While this method provides useful intuition, it only offers an approximate characterization that depends on the sampling density and coverage of the initial condition space. To complement and validate these results, we adopt the method proposed in [3], which formulates the computation of outer approximations of a ROA as an infinite-dimensional convex linear programming problem defined over the space of measures. This convex relaxation can then be systematically solved by means of the moment-SOS (sum-of-squares) hierarchy, which yields a sequence of increasingly tight approximations. In this way, the Monte Carlo exploration provides an intuitive, simulation-based picture, while the measure-theoretic convex optimization framework provides rigorous guarantees and convergence properties [1].

The model by Topp *et al.* [2] is described by the following system of equations:

$$\dot{G} = R_0 + (E_{g0} + S_I I) G, \quad (\text{C.1a})$$

$$\dot{I} = \beta \sigma \frac{G^2}{\alpha + G^2} - s I, \quad (\text{C.1b})$$

$$\dot{\beta} = (-\bar{d} + r_1 \cdot G - r_2 \cdot G^2) \beta, \quad (\text{C.1c})$$

$$\dot{S}_I = -c(S_I - S_{I,target}) \quad (\text{C.1d})$$

where:

- G is the plasma glucose concentration [mg/dl];
- I is the serum insulin concentration [μ U/ml];
- β is the β -cell mass [mg];
- S_I is the insulin sensitivity [ml/ μ U/d];

¹The contents presented here are based on the results published in the work by the author in [1]

For the values of the model parameter reference is made to [1]. As can be observed from the model equations, the system features a non-polynomial nonlinearity captured by the term

$$h(G) = \frac{G^2}{\alpha + G^2},$$

which corresponds to a second-order Hill function of the glucose concentration. This Hill function governs the insulin response to glucose dynamics and exhibits a characteristic sigmoidal shape, introducing a saturation effect at high glucose levels. In order to apply the methods introduced by [3], we propose a polynomial approximation of the function $h(G)$. This approximation is obtained through polynomial fitting of data sampled over the interval $[0, 600]$ for G . The choice of this interval is motivated by a precise rationale: its upper bound coincides with the hyperglycemic equilibrium identified in the original model of Topp et al. [2]. The analytical expression of the approximating six-degree polynomial in the monomial basis for $h(G)$ is provided in [1].

C.1.1 Methods

This section outlines the derivation of an outer approximation of the ROA of an equilibrium point of the model in (C.1) from the solution of a convex optimization problem [1], and explains how this problem can be addressed using the moment-SOS hierarchy, following the approach of [3]. For a vector field described by the system of differential equations $\dot{x} = f(x(t))$, we formally define the ROA χ_0 as

$$\chi_0 = \{x_0 \in X \mid \exists T > 0 \text{ s.t. } x(T|x_0) \in X_T\},$$

where $X := \{x \in \mathbb{R}^n : g_i^X(x) \geq 0\}$ and $X_T := \{x \in \mathbb{R}^n : g_i^{X_T}(x) \geq 0\} \subset X$, with $g_i^X(x), g_i^{X_T}(x) \in \mathbb{R}[x], i = 1, \dots, n$. In the following of the discussion the vector field $\dot{x} = f(x(t))$ is assumed to be polynomial, $f_i \in \mathbb{R}[x], i = 1, \dots, n$. Sets X, X_T are assumed to be semialgebraic sets.

Following the results in [3], the computation of the ROA χ_0 is formulated as the solution of an optimization problem. In particular, it is obtained as the super-level set of one of the functions that solve the optimization problem [3], as detailed below.

The primal optimization problem is stated as follows [1], [3]:

$$p^* = \sup \quad \mu_0(X) \tag{C.2a}$$

$$\text{s.t.} \quad \delta_T \otimes \mu_T = \delta_0 \otimes \mu_0 + \mathcal{L}^* \mu \tag{C.2b}$$

$$\mu_0 + \hat{\mu}_0 = \lambda \tag{C.2c}$$

$$\mu \geq 0, \mu_0 \geq 0, \mu_T \geq 0, \hat{\mu}_0 \geq 0 \tag{C.2d}$$

$$\text{spt}(\mu) \subset [0, T] \times X, \text{spt}(\mu_0) \subset X \tag{C.2e}$$

$$\text{spt}(\hat{\mu}_0) \subset X, \text{spt}(\mu_T) \subset X_T. \tag{C.2f}$$

The supremum is taken over the triplet (μ, μ_0, μ_T) and aims at maximizing the mass of the initial measure μ_0 , that is, to determine the maximum volume of the distribution of initial conditions [3]. Constraint (C.2b) corresponds to Liouville's equation, which encodes the vector field f through the adjoint operator $\mathcal{L}^* \mu$ in the space of measures. The support constraint (C.2e) on the occupation measure μ ensures that the trajectories remain within the state constraints, while the support constraint on the terminal measure μ_T guarantees that trajectories terminate in the target set [3]. Finally, in constraint (C.2c), the slack variable $\hat{\mu}_0$ ensures that the optimal solution coincides with the restriction of the Lebesgue measure λ to the ROA. The dual problem in the space of the continuous functions reads as [3]:

$$q^* = \inf \int_X w(x) d\lambda(x) \quad (\text{C.3a})$$

$$\text{s.t.} \quad \mathcal{L}_f(v, t) \leq 0, \quad \forall (t, x) \in [0, T] \times X \quad (\text{C.3b})$$

$$w(x) \geq v(0, x) + 1, \quad \forall x \in X \quad (\text{C.3c})$$

$$v(T, x) \geq 0, \quad \forall x \in X_T \quad (\text{C.3d})$$

$$w(x) \geq 0, \quad \forall x \in X. \quad (\text{C.3e})$$

The infimum is taken over $(v, w) \in C^1([0, T] \times X) \times C(X)$. In the dual formulation, the continuous function $v(t, x)$ is required to be non-increasing along the system trajectories (C.3b) and non-negative (C.3b)–(C.3d). Consequently, it follows that $w(x) \geq 1$ on X_0 [3]. It can be shown that strong duality holds between the primal problem in the measure space and the dual problem in the space of continuous functions, implying the absence of a duality gap and ensuring $p^* = q^*$ [3]. An *outer* approximation for the ROA is given by the super-level set:

$$\chi_0 \subset \{x : v(0, x) \geq 0\} = \{x : w(x) - 1 \geq 0\}.$$

To solve the optimization problem, it is approximated through a sequence of finite-dimensional LMI problems in the space of moments, up to a given relaxation order d . The solutions of the sequence of LMI problems converge to the optimum of the infinite-dimensional LP in the measures space as the relaxation order d tends to infinity. The LMI formulation of the considered problem is provided in [1], [3]. We apply the described method and we compare the results with a ROA estimation via Monte Carlo Simulation [1].

C.1.2 Results

C.1.2.1 ROA via Monte Carlo simulation

To compute the ROA, we first carried out a Monte Carlo simulation over the space of initial conditions[1]. Specifically, the model was simulated by randomizing the initial states, sampling $N = 10000$ points in the three-dimensional space $x = [G \ I \ \beta]^T$, with ranges $[0, 600]$ mg/dl for G , $[0, 150]$ $\mu\text{U/ml}$ for I , and $[0, 3000]$ mg for β . The initial condition for $x_{4_0} = S_{I_0}$ was fixed at $0.72 \text{ ml}/\mu\text{U/d}$, corresponding to the maximum value in the model, under the assumption that the simulated subject begins the degradation of S_I (the condition predisposing to diabetes) exactly at $t = 0$ [2].

Fig.C.1 illustrates the ROA computed for the normoglycemic equilibrium point at $G = 100$ mg/dl and for the hyperglycemic equilibrium point at $G = 600$ mg/dl [1]. As shown, the three-dimensional space of initial conditions $x_0 = [G_0 \ I_0 \ \beta_0]^T$ is separated by a surface located around $\beta = 300$ mg. The results indicate that both ROAs are (at least approximately) box-shaped, which is physiologically reasonable: β -cells are responsible for insulin secretion, so larger β -cell mass supports greater insulin production and, consequently, lower glucose concentrations. Thus, under the same decline in insulin sensitivity (i.e., the same disease trajectory), larger initial β -cell mass (above the threshold of $\beta = 300$ mg) prevents β -cell failure and drives the system toward the normoglycemic equilibrium at $G = 100$ mg/dl. Conversely, insufficient initial β -cell mass (below $\beta = 300$ mg) leads to irreversible disease progression toward the hyperglycemic equilibrium at $G = 600$ mg/dl [2]. Notably, this behavior appears to hold regardless of the initial plasma glucose concentration G or serum insulin concentration I [1]. It is important to note that, by construction, the Monte Carlo approximation yields only a *lower bound* of each ROA. Consequently, there is no strict guarantee that all initial conditions within the ‘red’ region in Fig. C.1 necessarily converge to the equilibrium at $G = 600$ mg/dl. This limitation motivates the adoption of the more rigorous approach presented in the following section [1].

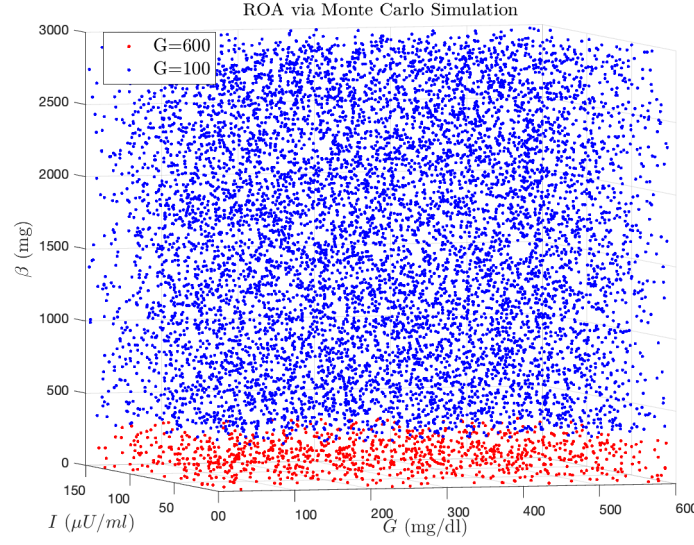


Figure C.1: The blue dots denote the set of initial conditions in the three-dimensional space $x_0 = [G_0 \ I_0 \ \beta_0]^T$ that converge to the equilibrium at $G = 100$ mg/dl (ROA of $G = 100$), whereas the red dots correspond to the set of initial conditions that converge to the equilibrium at $G = 600$ mg/dl (ROA of $G = 600$). The ROAs are box-shaped and separated by a surface nearby $\beta=300$ mg.

C.1.2.2 ROA via convex optimization

As previously discussed, we apply to the model in (C.1) the method in [3]. Fig.C.2 illustrates the ROA of the normoglycemic equilibrium point ($G = 100$ mg/dl), computed for two relaxation orders of the moment problem, $d = 5$ and $d = 11$. In both plots, the colored region corresponds to an *outer* approximation of the true ROA of the normoglycemic equilibrium point, while the area outside the box can be interpreted as an *inner* approximation of the ROA of the hyperglycemic attractor at $G = 600$ mg/dl. As d increases, the volume of the colored region progressively shrinks, and the separating boundary between the two regions moves closer to the surface around $\beta = 300$ mg, consistently with the Monte Carlo results shown in Fig. C.1. It is therefore reasonable to expect increasingly accurate approximations of this surface for larger values of d . Nevertheless, as discussed, convergence to the exact ROA cannot be guaranteed for finite d , since the true ROA is obtained as $d \rightarrow \infty$. Overall, the agreement between the two approaches supports the consistency of the computed ROAs and highlights the presence of a separating surface between the normoglycemic and hyperglycemic attractors at a β -cell mass of approximately 300 mg.

References

- [1] De Paola, P.F. and Borri, A. and Paglialonga, A. and Palumbo, P. and Dabbene, F., “Polynomial approximation of regions of attraction via occupation measures: An application to a biological autonomous system,” in *Proceedings of the IEEE 20th International Conference on Automation Science and Engineering (CASE)*, 2024. DOI: [10.1109/CASE59546.2024.10711611](https://doi.org/10.1109/CASE59546.2024.10711611).
- [2] Topp, B., Promislow, K., Devries, G., Miura, R. M., and T Finegood, D., “A model of β -cell mass, insulin, and glucose kinetics: Pathways to diabetes,” *J. of Theoretical Biology*, vol. 206, no. 4, pp. 605–619, 2000.
- [3] Henrion, D. and Korda, M., “Convex computation of the region of attraction of polynomial control systems,” *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 297–312, 2013.

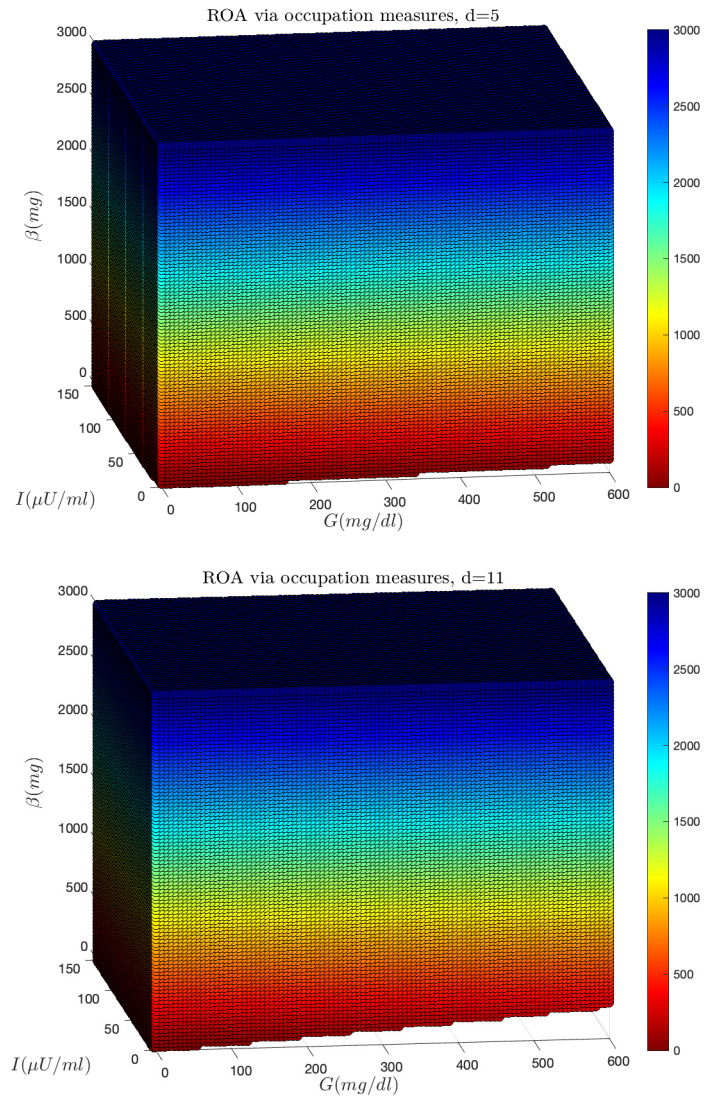


Figure C.2: Outer approximation of the ROA for $G = 100$ mg/dl obtained via occupation measures for $d = 5$ and $d = 11$.

