Constrained optimization under uncertainty of the driver's command for energy saving of high-speed trains using computational stochastic nonlinear dynamics and statistics

Julien Nespoulous

Université Gustave Eiffel

MSME LABORATOIRE MODÉLISATION ET SIMULATION MULTI ÉCHELLE
Sous la co-tutelle de :
CNRS
UNIVERSITÉ GUSTAVE EIFFEL
UPEC - UNIVERSITÉ PARIS-EST CRÉTEIL

SNCF

École doctorale Science, Ingénierie et Environnement (SIE)

Thèse de Doctorat
Spécialité : Sciences de l'ingénieur

présentée par

Julien NESPOULOUS

Constrained optimization under uncertainty of the driver's command for energy saving of high-speed trains using computational stochastic nonlinear dynamics and statistics

Pietro Marco CONGEDO	Directeur de recherche,	INRIA	Examinateur
Christine FUNFSCHILLING	Docteur,	SNCF	Examinatrice
Roger Georges GHANEM	Professeur,	USC	Rapporteur
Mathilde MOUGEOT	Professeure,	ENSIIE	Présidente
Guillaume PERRIN	Chercheur,	UGE	Examinateur
Rodolphe Le RICHE	Directeur de recherche,	CNRS/EMSE	Rapporteur
Christian SOIZE	Professeur Emérite,	UGE	Directeur de thèse

23 Novembre 2022

Abstract

The railway world is undergoing major changes. The advent of new technologies allows us to rethink the train system, but also face new challenges. The autonomous train is a significant advance in the field, but one must not forget all the ecological constraints that are now accentuated by the increase in energy costs. These issues raise a question: how can one make a train run autonomously while reducing its energy consumption? Several ideas can be explored to answer this question. This thesis work focuses on the optimization of the driver's command to limit the energy consumption of the trains. This problem is difficult to solve mainly because of the complexity of the railway system, the large amplitude of uncertainties attributed to the different model parameters, likewise, the importance of the constraints in the optimization problem. This thesis work is based on these three axes.

Firstly, the train is a complex system whose dynamic behavior can be difficult to predict. The construction of a rigid body model allows for describing the elements constituting the train and all their interactions, but it is expensive to solve for long journeys. For this reason, the longitudinal dynamics is often preferred when it is sufficient. The energy consumed by the train must carefully be estimated as it constitutes a key element of this work.

The second point focuses on the identification of the model parameters. This covers both models describing the dynamics and the energy consumption. But all trains do not behave in the same way. Therefore, the use of the probabilistic framework allows us to depict all these behaviors as much as is possible. The use of the Bayesian inference on a set of measurements performed on commercial trains brings the model closer to physical reality.

Finally, the optimization problem is complex to solve. The optimization variables and the search domain must be carefully defined with respect to the physical framework. A set of constraints ensures safety, punctuality, as well as passenger comfort. The cost function must be close to the industrial objective. However, all these quantities are random variables. For this reason, a robust strategy has been set up to be able to take into account all the uncertainty related to the train system. The optimal solutions obtained are compared with measurements from commercial trains.

Résumé

Le monde ferroviaire est en pleine mutation. L'avènement de nouvelles technologies permet de repenser le système du train mais aussi de faire face à de nouveaux enjeux. Le train autonome est une avancée notable dans le domaine mais elle ne doit pas oublier l'ensemble des contraintes écologiques qui sont aujourd'hui accentuées par l'augmentation des coûts de l'énergie. Ces problématiques soulèvent une question: comment faire rouler un train de façon autonome tout en réduisant sa consommation énergétique ? Plusieurs pistes peuvent être explorées pour répondre à cette interrogation. Ce travail de thèse se penche sur l'optimisation de la commande du conducteur pour économiser l'énergie consommée par les trains. Ce problème est difficile à résoudre à cause de la complexité du système ferroviaire, de la grande amplitude d'incertitudes attribuée aux différentes grandeurs du modèle, ou encore de l'importance des contraintes dans le problème d'optimisation. Ce travail de thèse s'articule autour de ces trois axes.

Dans un premier temps, le train est un système complexe dont le comportement dynamique peut s'avérer difficile à prévoir. La construction d'un modèle de corps rigides permet de représenter les éléments constituant le train et l'ensemble de leurs interactions mais celui-ci est coûteux à résoudre pour des trajets de grandes distances. Pour cette raison, la dynamique longitudinale est souvent privilégiée lorsque celle-ci est suffisante. L'énergie consommée par le train doit être estimée avec attention comme elle constitue un élément clé de cette recherche.

Le deuxième point se focalise sur l'identification des paramètres du modèle. Ceux-ci couvrent à la fois des grandeurs décrivant la dynamique, mais aussi la consommation énergétique. Mais les trains ne se comportent pas tous de la même façon. Aussi, l'utilisation du cadre probabiliste permet de représenter autant que possible l'ensemble de ces comportements. L'utilisation de l'inférence Bayésienne sur un ensemble de mesures réalisées sur des trains commerciaux rapproche le modèle de la réalité physique.

Enfin, le problème d'optimisation est complexe à résoudre. Les variables d'optimisation ainsi que le domaine de recherche doivent être définis avec attention pour respecter le cadre physique. Un ensemble de contraintes assure la sécurité, la ponctualité et le confort des passagers. La fonction coût doit s'approcher de l'objectif industriel. Cependant, toutes ces grandeurs sont des variables aléatoires. Pour cette raison, une stratégie robuste a été mise en place pour être capable de tenir compte de l'ensemble des incertitudes liées au système. Les solutions optimales obtenues sont comparées avec des mesures de trains commerciaux.

Acknowledgments

First, I would like to sincerely thank my thesis supervisor, Prof. Christian Soize for his invaluable help throughout my three-year thesis. His advice has always been precise and allowed me to advance in my work. Its scientific stature and his kind pedagogy improved my understanding and assimilation of complex notions in some fields of applied mathematics such as uncertainty quantification and optimization. I particularly appreciate his availability to unblock complex situations and all the time he gave me to help advancing.

I would like to thank Christine Funfshilling and Guillaume Perrin for the touch they could bring to my work. The richness of this work would not have been the same without the complementarity of their supervision.

Christine's expertise in dynamics allowed me to understand the railway field as well as all its complexity. She was often able to guide me to solve technical problems and showed me a great confidence during the realization of this thesis, but also by offering to assist her in giving lessons from the first year of thesis.

Guillaume's scientific knowledge allowed me to acquire a rigorous and a scientific curiosity essential to the smooth running of a scientific research. His advice often helped me to clarify situations or to think to an appropriate solution. The accuracy of his comments have not ceased to improve the quality of my work and my contributions.

I also thank my family, my friends and especially my darling for their foolproof support throughout this period, especially during the writing of the manuscript or during the preparation of the various conferences in which I had the opportunity to participate.

Finally, I would like to thank the entire department called "Direction Technologies Innovations et Projets Groupe" (DTIPG) for its warm welcome, the SNCF doctoral students' association for the organization of small events rich in learning, and more generally the whole engineers, doctoral students, service providers, interns and work-study interns with whom I have been able to collaborate during these three-years.

Remerciements

Dans un premier temps, je souhaite remercier sincèrement mon directeur de thèse, Prof. Christian Soize pour son aide précieuse tout au long de mes trois années de thèse. Ses conseils ont toujours été précis et m'ont permis d'avancer activement dans mes recherches. Son envergure scientifique et sa pédagogie m'ont permis de comprendre et d'assimiler des notions complexes dans le domaine des mathématiques appliquées comme la quantification des incertitudes et l'optimisation. J'apprécie notamment sa disponibilité et tout le temps qu'il m'a accordé pour m'aider à débloquer des situations complexes.

Je tiens à remercier Christine Funfschilling et Guillaume Perrin pour la touche qu'ils ont pu apporter à mon travail. La richesse de ce travail n'aurait pas été la même sans la complémentarité de leur encadrement.

L'expertise de Christine en dynamique m'a permis de comprendre le monde du ferroviaire ainsi que toute sa complexité. Elle a souvent su m'aiguiller pour répondre à des problématiques techniques et m'a témoigné une grande confiance au cours de la réalisation de ce travail de thèse mais aussi en me proposant de l'assister pour donner des cours dès la première année de thèse.

Les connaissances scientifiques de Guillaume m'ont permis de d'acquérir une rigueur et une curiosité scientifique essentielle au bon déroulement d'une recherche scientifique. Ses conseils m'ont souvent aidé à éclaircir des situations ou à réfléchir à une solution adaptée. La justesse de ses commentaires n'ont pas cessé d'améliorer la qualité de mon travail et de mes contributions scientifiques.

Je remercie également ma famille, mes amis et particulièrement ma chérie pour leur soutien à toute épreuve pendant toute cette période, notamment durant la rédaction du manuscrit ou lors de la préparation des diverses conférences auxquelles j'ai eu l'opportunité de participer.

Je remercie enfin l'ensemble de la Direction Technologies Innovations et Projets Groupe (DTIPG) pour son accueil chaleureux, l'association des doctorants de la SNCF pour l'organisation de petits évènements riches en apprentissage, et plus généralement l'ensemble des ingénieurs, doctorants, prestataires, stagiaires et alternants avec qui j'ai pu collaborer durant ces trois années.

Contents

Glossary

Deterministic variable	Lowercase letter: a.
Deterministic vector	Lowercase bold letter: $\boldsymbol{a}$.
Deterministic matrix	Uppercase calligraphic letter between brackets: $[\mathcal{A}]$.
Deterministic functional	Lowercase letter between curly bracket: $\{a\}$.
Random variable	Uppercase letter: A.
Random vector	Uppercase bold letter: $\boldsymbol{A}$.
Max and min value	Exponent max and min: a^{max} and a_{min}.
Optimal solution	Exponent $*$: a^*.
Time derivative	Dot: $\dot{a}$.
Variable perturbed by modeling error	Exponent mod: a^{mod}.

$\mathbb{E}$	Mathematical expectation.
$\mathbb{1}_{\mathcal{D}}$	Indicator function over a set $\mathcal{D}$.
$\det$	Determinant of a matrix.
Γ	Gamma function.
$\arg\max$ and $\arg\min$	Argument of the maxima and minima.
$\propto$	Proportionality relation.
d_1	Distance associated with norm-1.
d_2	Distance associated with norm-2.
$\mathcal{T}$	General characteristics of the running environment.
$\mathcal{V}$	General description of the vehicle.
r	Curve radius.
θ	Declivity of the track.
z^{alt}	Altitude of the track.
v^{max}	Speed limitation on the track.
x_P and y_P	Relative longitude and latitude of P in the spatial discretization.
$\bar{\boldsymbol{v}}_P^w$	Mean part of the wind speed velocity vector at point P.
$\boldsymbol{V}_P^{turb}$	Turbulence part of the wind speed velocity vector at point P.
$\boldsymbol{V}_P^w$	Wind speed velocity vector at the point P.
α_P	Orientation of the wind at point P.
v^w	Wind speed projected on the longitudinal to the track axis.

a	Davis coefficients (static friction).
b	Davis coefficients (dynamic friction).
c	Davis coefficients (aerodynamic friction).
m	Mass of the train.
$\bar{m}$	Deterministic estimation of the train mass.
Δm	Correction of the mass estimation.
p^a	Auxiliary power.
Δp^a	Variation of auxiliary power.
a_η	Traction efficiency coefficient (linear term).
b_η	Traction efficiency coefficient (constant term).
c_η	Braking efficiency coefficient (linear term).
d_η	Braking efficiency coefficient (constant term).
i^{mes}	Measured current intensity.
u^{mes}	Measured voltage.
ϕ^{mes}	Measured phase shift.
f^{mes}	Measured energy.
δt^{mes}	Discretization time step for measurements.
N^{mes}	Number of measured time sampling points.
$\boldsymbol{X}$	Uncertain parameters.
$\boldsymbol{X_D}$	Uncertain dynamic parameters.
$\boldsymbol{X_E}$	Uncertain energy consumption parameters.
$\underline{\boldsymbol{x}}$	Mean value of the uncertain parameters.
$\mathcal{U}$	Admissible set for the driver's command $\{u\}$.
$\{u\}$	Functional driver's command.
$\{u_{\leq t}\}$	Functional driver's command truncated at time t.
$\boldsymbol{u}$	Discretized driver's command.
$\boldsymbol{u}^{mes}$	Experimental driver's command (determined in inverse).
$\underline{\boldsymbol{u}}$	Initial driver's command respecting the constraints.
$\boldsymbol{u}^*_{\boldsymbol{x}_k}$	Optimal driver's command in the specific configuration $\boldsymbol{x}_k$.
$\boldsymbol{u}_{\Delta,\kappa}$	One driver's command vector kept for the PCA.
e_κ	Energy consumption of the train following driver's command $\boldsymbol{u}_{\Delta,\kappa}$.
$\widehat{\boldsymbol{u}}^\kappa$	Normalized driver's command.
y	Train curvilinear position.
$\dot{y}$	Train curvilinear speed.
$\ddot{y}$	Train curvilinear acceleration.
$\mathcal{S}$	Dynamic solution.

f^R	Longitudinal friction force (Davis approximation).
f^C	Corrective term of the longitudinal friction force (in curve).
f^W	Projection of the weight on the longitudinal to the track axis.
f^T	Longitudinal traction force.
f^B	Longitudinal braking force.
f^E	Longitudinal electrodynamic braking force.
p^T	Electrical power during traction phases.
p^B	Electrical power during braking phases.
p^E	Total electrical power.
η^T	Traction efficiency.
η^B	Recovery braking efficiency.
$\mathcal{F}$	Energy consumption of the train along the journey.
$\Delta\mathcal{F}$	Variation of energy consumption.
t	Current time.
s	Curvilinear position on the track.
t_s and t_f	Starting and final time.
s_s and s_f	Starting and final position.
v_s and v_f	Starting and final speed.
t_j^{mes}	Measured time sampling points.
t_i	Simulated time sampling points.
Δt	Simulated discretization time step.
g	Gravitational acceleration.
k^C	Empiric coefficient for the corrective force applied in curve.
k^{rot}	Corrective factor of the inertial term.
N	Number of small discretization intervals.
N^Δ	Number of large discretization intervals.
$N^\mathcal{T}$	Number of different running environments.
N^U	Number of driver's commands selected for the PCA method.
$N^{\boldsymbol{X}}$	Number of realizations used to estimate the expectation value $\mathcal{F}_{\mathbb{E}_{\boldsymbol{X}}}$.
N^{id}	Number of candidates $\bar{\boldsymbol{x}}_k$.
N^{step}	Number of steps of the deterministic identification process.
N^{MH}	Number of MH iteration in one MwG iteration.
N^{MwG}	Number of MwG iterations.
N^{pop}	Number of points in one realization of CMA-ES.
L	Size of the moving average filter.
w_i	Weighting of the moving average filter.
c_T and c_B	Traction and braking factors used during the transformation $\tilde{\boldsymbol{T}}$.
$\boldsymbol{T}_{MA}$	Moving average filter.
$\tilde{\boldsymbol{T}}$	Transformation assuring the respect of the punctuality constraints.
$\boldsymbol{T}$	Complete transformation of the driver's command.

$\widehat{\boldsymbol{x}}$	Central value for the random draw of candidate $\bar{\boldsymbol{x}}_k$.
$\underline{\boldsymbol{x}}_k$	Candidate drawn for the deterministic identification process.
$\mathcal{F}_k$	Energy consumption associated with the candidate $\bar{\boldsymbol{x}}_k$.
Z	General random variable.
f_Z	Probability density function of Z.
$\mathcal{S}_Z$	Support of the probability distribution of Z.
$Z_1 \vert Z_2$	Conditional random variable Z_1 given Z_2.
$\underline{z}$ and σ_Z^2	Mode and variance of Z.
k_Z and θ_Z	Shape and scale hyperparameters of the Gamma random variable Z.
$\left[C^Z\right]$	Covariance matrix of Z.
$\boldsymbol{\varepsilon}^{mod}$	Random variable quantifying the modeling error.
$\boldsymbol{\varepsilon}^F$	Random variable quantifying the dynamic modeling error.
$\tilde{\boldsymbol{\varepsilon}}^F$	Propagation of $\boldsymbol{\varepsilon}^{\boldsymbol{F}}$ in the energy consumption model.
$\boldsymbol{\varepsilon}^P$	Random variable quantifying the energy modeling error.
σ^F	Modeling error hyperparameter associated with $\boldsymbol{\varepsilon}^F$.
$\tilde{\sigma}^F$	Modeling error hyperparameter associated with $\tilde{\boldsymbol{\varepsilon}^{\boldsymbol{F}}}$.
σ^P	Modeling error hyperparameter associated with $\boldsymbol{\varepsilon}^{\boldsymbol{P}}$.
$\boldsymbol{\sigma}$	Vector containing the values of σ^F and σ^P.
$\mathcal{L}$	Likelihood function.
$f_{\boldsymbol{X}}^{prior}$ and $f_{\boldsymbol{X}}^{post}$	Prior and posterior distributions of $\boldsymbol{X}$.
$[\hat{u}]$	Normalized driver's commands matrix.
$[C_{\hat{u}}]$	Covariance matrix of $[\hat{u}]$.
$[\lambda]$ and $[\Phi]$	Eigenvalue and eigenvector matrices of $[C_{\hat{u}}]$.
$[q]$	Orthogonal matrix used by the PCA method.
(m)	Dimension reduction.
$\left[\hat{u}^{(m)}\right]$	Reduced matrices of $[\hat{u}]$.
$\left[q^{(m)}\right]$, $\left[\lambda^{(m)}\right]$, $\left[\Phi^{(m)}\right]$	Reduced matrices of $[q]$, $[\lambda]$, and $[\Phi]$.
f_o	Overtaking function (speed limitation constraint).
c	General constraint function.
$\boldsymbol{c}$	Constraint vector gather c_1, c_2, and c_3.
c_1, c_2, and c_3	Speed limitation, final position, and final speed constraints.
p_1, p_2, and p_3	Probability tolerance thresholds.
ε_1 and ε_2	Acceptability thresholds.
$\mathcal{F}_{\mathbb{E}_{\boldsymbol{X}}}$	Expectation value of the energy consumption over $\boldsymbol{X}$.

Acronyms

CMA-ES	Covariance Matrix Adaptation - Evolution Strategy
GS	Gibbs Sampling
LGV	Ligne à Grande Vitesse (French high-speed track)
MCMC	Monte Carlo Markov Chain
MH	Metropolis-Hastings
MwG	Metropolis-within-Gibbs
PCA	Principal Component Analysis
PDF	Probability Density Function
PK	Point kilométrique (kilometer point: curvilinear abscissa on the track)
RMSE	Root Mean Square Error
SIMPACK®	Multibody simulation software
SNCF	Société Nationale des Chemins de fer Français (French railway company)
TER	Train Express Régional (French regional train)
TGV	Train à Grande Vitesse (French high-speed train)
VAMPIRE®	Vehicle dynAmic Modeling Package In a Railway Environment

Chapter I

Introduction

I.1 Industrial context and motivations

In recent decades, the world is facing new challenges. Among many, the environmental issue is one of the most critical stakes. The transportation sector, as one of the largest energy consumers, is particularly affected by these challenges. The demographic growth encourages companies to modify their way of thinking concerning matters that involve transportation. Momentum has further emerged as a credo. However, for this precept, it is difficult to comply with the ecological challenges.

Railway transport is often associated with the top of the class due to its low-carbon emissions. Still, the high-speed trains consume a large amount of electric energy compared with the other types of trains. The recent skyrocket of the electric energy costs encourages the railway companies to reduce their consumption. In this work, we focus on the running consumption of trains. In order to limit it, three different levers can be activated.

First, the running environment of the train may be modified. Playing on the rail material, protecting the vehicle from the wind, reducing the aerodynamic drag by creating an artificial vacuum environment are few insights that could reduce the energy consumption. Nevertheless, modifying the infrastructure on thousands kilometers does not appear to be a panacea due to the important costs of constructing a new infrastructure.

Second, the vehicle itself can be redesigned. For instance, the shape of the train is slightly transformed to reach better performance in terms of aerodynamics or traction chain efficiency for instance. These modifications are not negligible but often limited due to the different operational constraints.

Finally, the last levers is the driver's behavior. This aspect seems to be relatively negligible in appearance, but it has an impact on the energy consumption of the train as well. Indeed, a previous work has shown that on the same high-speed line, energy consumption between the different circulations could vary by up to 20 or 30%. Climatic conditions (wind, humidity) can partly explain these differences in consumption, however, driving probably also plays an important role. Thus, looking for the best driver's command would be of great concern to reduce the energy consumption of the trains.

The industrial motivations are twofold. First, minimizing the energy consumption is an ecological stake. Second, it leads to important costs reduction involving economical motivations.

I.2 Scientific objectives

As explained in the industrial context, the aim of this work is to build speed profiles minimizing energy consumption. However, the driver's command only plays a role on the traction and braking forces applied to the train system. Consequently, the speed cannot be directly imposed. Therefore, in contrast to other research, this work seeks to directly identify the driver's command for a given environment $\mathcal{T}$ and a specific vehicle $\mathcal{V}$. The problem that will be solved in this work is thus written as follows,

$$\{u^*\} = \underset{\substack{\{u\}\in\mathcal{U} \\ c(\{u\},\mathcal{T},\mathcal{V})=0}}{\arg\min} \ \mathcal{F}\left(\{u\},\mathcal{T},\mathcal{V}\right), \tag{I.1}$$

with $\{u\}$ the driver's command, $\{\mathcal{F}\}$ the objective functional (linked to the energy consumption), $\mathcal{U}$ the admissible set, $c(\{u\},\mathcal{T},\mathcal{V})$ the constraint for a given driver's command $\{u\}$, and $\{u^*\}$ the optimal driver's command.

Solving this problem presents several major scientific challenges. Firstly, the relation between speed and the driver's command that can be imposed on the dynamic system is complex. The construction of a longitudinal dynamic model for high-speed trains has been developed to work in concert with an energetic model. These models enable to calculate the train dynamics and its energy consumption from the three entries:

- the environmental parameters, $\mathcal{T}$, are deterministic and describe all the exterior factors that have an impact on the train system.
- the model parameters, $\mathcal{V}$, are associated with the vehicle dynamics and its energy consumption.
- function $\{u\}$ describes the driver's command.

The models are summarized in Figure I.1.

Secondly, many sources of variability have an important impact on the train system (wind velocity, temperature, humidity, number of passengers, wear state, ...). In particular, the model parameters $\mathcal{V}$ are subject to uncertainty. An identification of the model parameters is necessary for updating the simulation models.

Thirdly, the sought driver's command are functional. Therefore, the optimization problem to be solved is very high dimensional (several hundreds). To this extent, numerical methods are introduced to reduce the dimension of the problem and to solve it.

Finally, punctuality and speed limits are nonlinear constraints that are very sensitive to the model parameters, which are subject to uncertainties. In summary, the problem to be solved is a probabilistic optimization problem under nonlinear constraints and is in very high dimension. A robust strategy is proposed to overcome these difficulties.

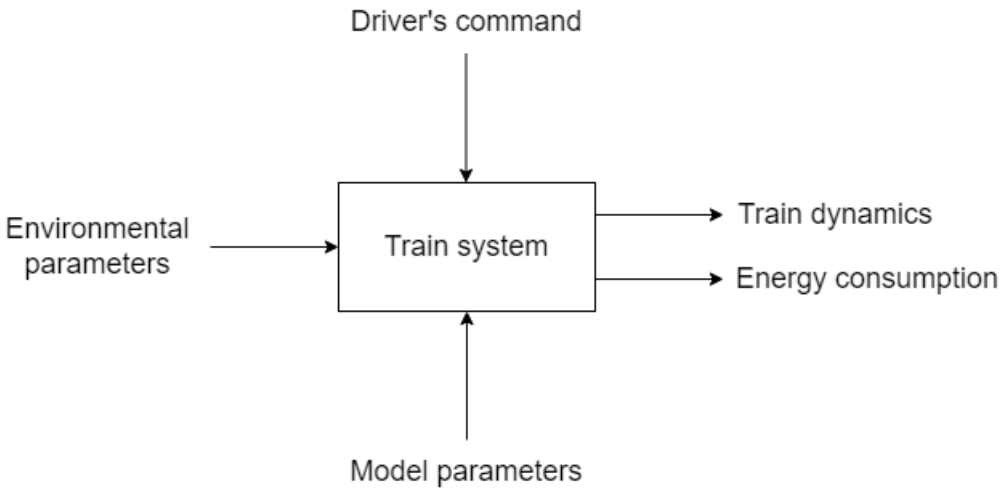

Figure I.1: High-speed train system: inputs, parameters, and outputs of the models.

I.3 State of the art

In the last decades, the optimization of the driver's command has been studied from different points of view. From the train-behavior modeling to the optimization of the driver's command, including the calibration of the models, numerous works have explored these railway fields.

Before trying to optimize the driver's command, the train system has to be modeled properly. Figure I.1 presents the train system. As it can be observed, the entries (inputs and parameters) play a crucial role in the modeling, as listed below.

- The train is very sensitive to the environment. The track and the wind are essential features of the high-speed train dynamic modeling. The first one is described from its geometric design (larger scale), and its irregularities (smaller scale). P. Aknin *et al* describe this formulation in [1]. The second one is decomposed into a mean part and a fluctuation part for describing the turbulence. Both of them have different influences on the train system depending on factors that want to be highlighted. Depending on the type of works performed, the turbulence fluctuations are or not taken into account. In the context of railway field, a wind model is proposed by H. Liu in [2].

- The second entry that cannot be neglected is the description of the vehicle. Few aspects are discussed below, but it only represents a small proportion of all the complexity of the train system. The design of the wheels is closely linked to the contact force. This force is difficult to precisely measure on-track, that is why novel research focuses on their estimation of by the use of force with neural networks [3] or random walk methods [4]. Each of the two levels of suspensions have to be modeled precisely to monitor their behavior [5]. In case of wear, this mechanical behavior is modified, but the wear has to remain acceptable. The traction chain is composed of many different components that is difficult to individually model. When focusing on the energy consumption of the train, this aspect must be considered. In general, in the context of the energy consumption studies,

subjects such as the pantograph/catenary interaction or the acoustics are not taken into account (we will use this assumption in our work).

- Finally, the driver's command directly impacts the train dynamics. Two point of views can be considered: the speed profile and the driver's command. Both of them are connected to the dynamic model. Estimating the speed profile from the driver's command can be achieved solving the train dynamics. On the contrary, determining the driver's command from the speed profile in inverse is much more difficult.

Once the entries are well defined, the train system can be modeled. Once again different approaches can be highlighted. In railway domain, the multibody framework [6] allows us to have a complete description of each component of the train. VAMPIRE® or SIMPACK® are two examples of software, which are frequently used in railway dynamics. But this solution might be relatively numerical time consuming as it results in solving six equations for each rigid body. Another approach consists in observing the train system in the longitudinal to the track axis. A complete overview is proposed by C. Cole in [7].

Obtaining measurements is often expensive and these data are sometimes replaced by simulations. Consequently, the identification of the model parameters is generally necessary. It allows for having an efficient representation of the real system. Two types of uncertainties can exist in the models: the model parameters uncertainties and the model uncertainties induced by modeling errors. For this reason, the uncertainty quantification must be taken into account (see for instance, [8, 9, 10, 11, 12, 13]). In the framework of railway systems, the Bayesian inference has been used in [14, 15] to identify some parameters of prestressed concrete catenary poles and of the suspension components. A. Dib *et al* proposed to use the Bayesian method for predictive maintenance in [16].

Much works have studied the optimization of the driver's command in railway field. From K. Ichikawa [17], one of the first paper dedicated to this domain, different points of view have been considered. For instance, several trains have been considered in [18], the uncertainties are included in the optimization problem in [19], [20], and [21], multi-objectives are taken into account in [22], [23], [24], and [25]. They aim to reduce the energy consumption of the train, to assure the punctuality, to maximize the passenger's comfort, or the rail and rolling stock wear for example. All these points of view modify the optimization problem depending on the objective. These optimization problems can be solved with many different methods like evolutionary algorithms [26], dynamic programming [27], pseudo-spectral methods [28], or the maximum principle [29] between others. The uncertainty also plays an important role in the optimization under uncertainties and the use of general methods can be mandatory (see for instance [30, 31, 32, 33, 34, 35, 36, 37]).

I.4 Positioning and novelty of the thesis work

The present work is innovative in several ways with respect to the state-of-the-art.

(i) First, the models are constructed to be as close as possible to the train system under consideration. In particular, the use of the measured energy consumption has been favored

instead of the electrical power. Indeed, it shows to be more performing due to the fact that the models are constructed so as to estimate the energy and not the power. Many measurements are involved to identify the model parameters of the inputs (track and wind) and of the vehicle. These experimental data are processed to be usable by the models. For instance, an efficient model of the traction and braking forces are constructed in order to well represent the real capacities of the train. In particular, the energy efficiency model that allows for evaluating the yield of the traction chain, has been constructed with the help of measurements. The capacity of the train to recover a part of the consumed energy is taken into account in an algorithm.

(ii) Second, the introduction of model parameter uncertainties is proposed for this railway context. Two random vectors are introduced for representing model uncertainties induced by modeling error. Their probability distributions are constructed by using the maximum entropy principle from Information Theory. Their mean vectors and covariance matrices are calibrated by using experiments.

(iii) Then, the model parameters are calibrated using a methodology in two steps: first, deterministic identification of the model parameters is performed, followed by a Bayesian inference for estimating their posterior distributions. These two steps also identify the driver's command in inverse from the measured train speed profile. Note that the prior distributions are constructed using an available information given by the train constructor. Note also that the Metropolis-within-Gibbs algorithm has been preferred due to the unequal influence of the uncertain parameters on the train system.

(iv) The optimization method is also very different from the one encountered in previous published works. The driver's command has been privileged to the speed of the train as the optimization variable. Moreover, the constraints are related to the imposed time of the train arrival, given position in space, zero speed at arrival, and the curvilinear speed limitation on the track. A complete time discretization of the driver's command is performed from the initial to the final time and is represented by a driver's command vector. Since the optimization problem is in high dimension, a reduced-order representation is constructed using a Principal Component Analysis. Integrating the uncertainty in the optimization problem introduces important difficulties. A given deterministic driver's command vector cannot respect the constraints for all realizations of the uncertain model parameters. Then, a method has thus been developed to deal with this difficulty. It consists in looking for the optimal driver's command vector that is easy to transform in another driver's command vector respecting the constraints. The transformed driver's command vectors minimize the statistical mean of the energy consumption with respect to the model parameters uncertainties. Roughly speaking, this transformation allows for managing the final time, arrival position, and arrival train speed. The use of the Covariance Matrix Adaptation - Evolution Strategy (CMA-ES) algorithm is also innovative to solve this kind of problem. Finally, all the optimal solutions obtained are compared with measurements carried out on commercial trains. This allows for validating the models and the identification methods, but also to verify the good quality of the optimal solution.

I.5 Outline of the thesis

This thesis is divided into three parts. Each of them answers to a specific objective that will be used in the following sections. Overall, the chapters aim to construct a framework that can be used to solve the issue.

First of all, Chapter II focuses on the modeling of trains. A brief description of the environment, the vehicle, and the driver's command is given in Section II.1. The modeling of the high-speed train dynamics is presented in Section II.2. Particular attention is paid to the modeling framework. Section II.3 describes the energy consumption model. A sensitivity analysis is carried out in Section II.4 and conclusions are drawn in Section II.5. A discussion on the modeling choices is also given.

The third chapter deals with the identification of the model parameters. To this extent, two different approaches introduced in Section I.4 are developed. The selection of the uncertain parameters is performed in Section III.1. The identification of the experimental driver's command is realized in Section III.2. Sections III.3 and III.4 present the two steps of the identification method evoked in Section I.4. The first one lies on the deterministic identification of the mean value of each uncertain parameters. The second one estimates their posterior distribution using a Bayesian inference method. Section III.5 concludes on this chapter.

Chapter IV gives a robust answer to what this thesis aims to address, that is to say the optimization of the driver's command to limit the energy consumption of the train under constraints and uncertainties. Its high dimension imposes the use of reduction methods (as explained in Section I.4) that are presented in Section IV.1. The impact of the uncertainties are described in Section IV.2. The optimization problem is formulated in Section IV.3 before presenting the methodology and the developed algorithms for solving it numerically in Section IV.4. The optimal solution is illustrated in Section IV.5. Section IV.6 gives a conclusion on the optimization problem.

Conclusions, perspectives, publications and communications related to this work are presented in Chapter V followed by the references in the Bibliography section.

The Appendix presents A the multibody approach for the train dynamic modeling, B the CMA-ES optimization algorithm, and C another case that we have studied concerning the optimization of the driver's command for autonomous trains in a deterministic framework.

Chapter II

Physics modeling of high-speed trains

This chapter presents the mathematical/physical models of high-speed trains, which have been constructed in order to represent the train behavior in a specific rolling environment and for a given driver's command. Effectively, the choice of optimizing the driver's command instead of the train speed profile requires to know the impact of the driver's command on the train system. In other words, this chapter gives an answer to the following question: how does the driver's command interacts with the train behavior? Section II.1 describes the environmental inputs, denoted by $\mathcal{T}$, of the train system to have a better understanding on its whole complexity (green box of Equation (II.1)). Then, the modeling has been separated into two phases. The first phase details the scientific framework to construct the dynamic model in Section II.2. The second part proposes an energy consumption model specific for high-speed trains in Section II.3. These two models are useful for characterizing the constraints and the cost function (blue boxes of Equation (II.1)). A sensitivity analysis is carried out on a set of model parameters in Section II.4. Finally, Section II.5 concludes and discusses the modeling choices. For a fixed environment $\mathcal{T}$ and a given vehicle $\mathcal{V}$, the general optimization problem is given by:

$$\{u^*\} = \underset{\substack{\{u\}\in\mathcal{U} \\ c(\{u\},\mathcal{T},\mathcal{V})=0}}{\arg\min} \mathcal{F}\left(\{u\},\mathcal{T},\mathcal{V}\right). \tag{II.1}$$

II.1 Description of the system characteristics

In the railway field, each "journey" is defined by a vehicle, moving in a specific environment, with a determined driver's command. Obviously, each type of vehicle has its own response to an excitation track input. Effectively, a high-speed train will not behave as a regional or a fret train. Moreover, each train of a given type has a proper mechanical behavior, which depends on the wear state of the vehicle, the number of passengers, etc. The rolling environment also plays an important role as it directly impacts the train dynamics. By rolling environment, we can gather every element, which is exterior to the train system such as the wind or the track. Without any doubt, the wind modifies the train dynamics especially for high-speed trains, for which the aerodynamic forces play an important role. In the same way, the train behavior is different if it moves on a high-speed track (designed with smooth curvature) compared to a subway line. Finally, the driver's command is chosen by the driver and defines if the motor or the brakes are used. These three elements fully determine the entries of the train system.

In the following sections, we describe each of these entries individually and the mathematical representation we have proposed to model them. We particularly focus on the impact these elements have on the train dynamics and on the associated energy consumption as these physical quantities are crucial in this work.

II.1.1 Characterization of the rolling environment

The environment of the train plays an important role on its dynamic behavior. Clearly, the wind has an impact on the aerodynamic forces applied to the train and the track modifies the wheel/rail contact forces. For example, a strong headwind generates a higher resistant force and thus increases the energy consumption of the train. In the same way, moving forward in a positive track slope (or declivity) will impose to apply a higher traction force to compensate the effect of gravity and consequently, the train will consume more energy. The environment of the train also includes all the electric infrastructures. Indeed, the supply in electric energy may vary from a track to another one depending on the position of the electric substations or simply on the type of electric current provided (continuous or alternative). The three following paragraphs present how each of these factors impacts the train system and how they are formally described. The description of the whole rolling environment of the train is denoted by $\mathcal{T}$.

Description of the track

First of all, the train dynamics is strongly related to the track geometry and irregularities. As a matter of fact, the wheels are in direct contact with the rail and every default modifies the train behavior. In order to be able to precisely characterize the contact, the wheel, the rail, and the track have to be modeled with attention. Each track is described by its geometry and the defaults (irregularities). The geometry contains the information characterizing the position of the mean line and the cant (also called superelevation). The position of the mean line is defined by the vertical curvature (linked to the declivity) and the horizontal curvature. The track cant describes the elevation of a rail over the other. This degree of freedom aims to compensate a part of the centrifugal acceleration by elevating the exterior rail in curve. With these three quantities, we are able to describe every possible mean line. The geometry defaults are of four types. A vertical or lateral alignment irregularity corresponds to a vertical or lateral displacement of the two rails compared with the mean line. The gauge default is a modification in the distance between the two rails and the cant deficiency defines a variation of the elevation of one rail compared with the mean line superelevation. These four quantities entirely describe all the possible defaults of the track. They are represented in Figure II.1. The defaults are correlated as long as one irregularity often leads to others. They can be described by random variables as it is proposed by R.H. Fries *et al* in [38]. Realistic tracks can even be modeled by random fields as it has been developed by G. Perrin *et al* in [39]. The three quantities defining the mean line and the four quantities characterizing the defaults are sufficient enough to perfectly describe the track and its defaults. The defaults have an important impact on the train dynamics, as it has been identified by G. Perrin *et al* in [40]. Moreover, they may vary in time due to many factors. They have to be monitored with high precision as the dynamic response of the train is very sensitive to their evolution [41], [42].

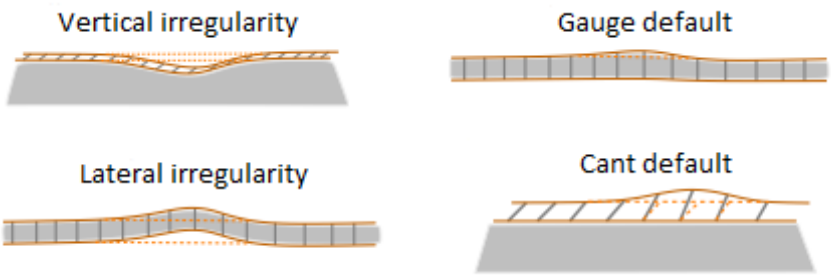

Figure II.1: Representation of the four types of track defaults (irregularities).

The defaults are frequently controlled thanks to specific vehicles, which estimate all the defaults owing to mechanical or optical measurements. The fine description of the defaults is very important for safety and comfort reasons, but they have a relatively low influence on the energy consumption of the train and on the quantities of interest that will be analyzed in this work. For this reason, we choose to neglect the defaults and only consider the mean part of the track. Denoting by s the curvilinear position on the track, the track is composed of the curvature, defined by the curve radius $r = \{r(s), s \in [s_s, s_f]\}$, and the declivity, characterized by the slope $\theta = \{\theta(s), s \in [s_s, s_f]\}$ (quantifying the elevation distance of both rails per meter). A curvilinear speed limitation is also defined on each track to assure the security of the journey. It is written as a function $v^{max} = \{v^{max}(s), s \in [s_s, s_f]\}$ in the remaining part of the document. This speed limitation, which depends on the curvilinear position on the track, is represented in Figure II.2 for the French LGV Rhin-Rhone as an example.

We can observe that there are relatively few parts of the journey with a specific speed limitation apart from the starting position of the line. The curvature, which is the inverse of the curve radius, and the declivity (represented here by the altitude for more clarity) are much more perturbed. For a track characterized by $\mathcal{T}$, the altitude $z^{alt} = \{z^{alt}(s), s \in [s_s, s_f]\}$ can be calculated from the slope θ thanks to the following equation,

$$z^{alt}(s) = \int_{s_s}^{s} \theta(\tilde{s}, \mathcal{T})\, d\tilde{s} + z^{alt}(s_s)\,, \tag{II.2}$$

in which $z^{alt}(s_s) = 0$ is the altitude at the starting point. They are discretized on 1-meter-long intervals. This choice allows a good representation of the track geometry for the need of this work.

Wind characterization

The second factor of the rolling environment that has an important influence on the train system, and especially on the energy consumption, is the wind. It can be decomposed into two quantities: the mean velocity vector $\bar{\boldsymbol{v}}_P^w$ at point P (direction and amplitude) and the fluctuation part [43] of the turbulence $\boldsymbol{V}_P^{turb}$ at point P. Consequently, the total wind speed $\boldsymbol{V}_P^w$ is written, at time t, as

$$\boldsymbol{V}_P^w(t) = \bar{\boldsymbol{v}}_P^w(t) + \boldsymbol{V}_P^{turb}(t)\,. \tag{II.3}$$

The mean part of the wind can be measured with accuracy and strongly impact the energy consumption. It is specific at each position and time. Consequently, it has to properly be

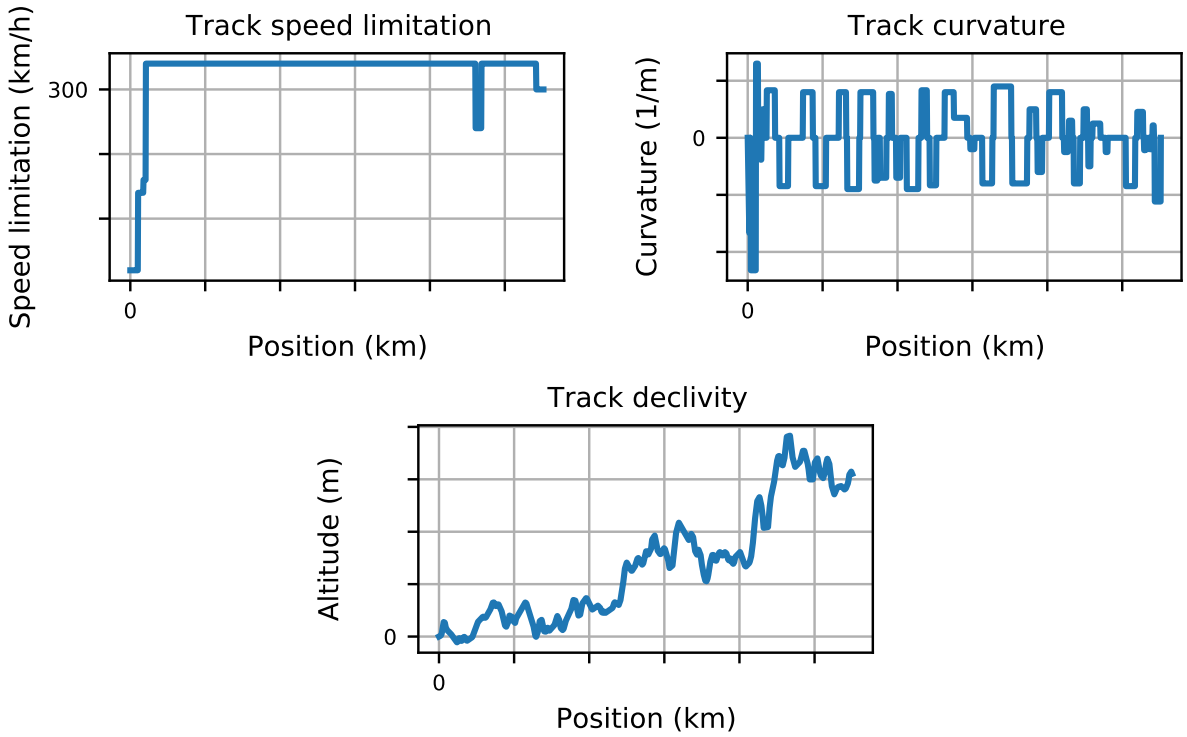

Figure II.2: Speed limitation, curvature, and declivity of the LGV Rhin-Rhone depending on the position. In this thesis, most of the axes of the figures shown are hidden for confidentiality purposes.

represented at each position on the track and at any time during the journey. As mentioned in Chapter I, the turbulence part has no significant effect on the energy consumption, and consequently its model is not necessary.

In this work, we have at our disposal predictions of the mean wind velocity in the region concentrated around the French LGV Rhin-Rhone track. The wind speed profile is then defined as the values of the mean wind velocity projected on the longitudinal track axis. These forecasts are estimated by the French meteorologic Meteo-France company. They are performed once every hour on a grid of 40 points for each longitude and latitude degree. For the studied area (around the LGV Rhin-Rhone), it represents around $2\,000$ points. This grid is interpolated at each needed track position as follows: at each point of the track P, four closest prediction points are extracted. We call them P_{LL}, P_{LR}, P_{UL}, and P_{UR}, the first letter standing for the vertical or latitude position (L for lower and U for upper) and the second for the horizontal or longitude position (L for left and R for right). We write x_P and y_P the normalized relative longitude and latitude of P. Both quantities take their values in $[0, 1]$. For example, $\{x_P = 0, y_P = 1\}$ corresponds to the point P_{UL}. Schema II.3 represents the four selected prediction points and the associated normalized coordinates x_P and y_P. With these notations, we have proposed to interpolate spatially the mean wind at a current time t thanks to the following equation:

$$\begin{aligned}\bar{\boldsymbol{v}}^w_P(t) = (1-x_P)(1-y_P)\,\bar{\boldsymbol{v}}^w_{P_{LL}}(t) + x_P(1-y_P)\,\bar{\boldsymbol{v}}^w_{P_{LR}}(t) + (1-x_P)\,y_P\bar{\boldsymbol{v}}^w_{P_{UL}}(t) \\ + x_P y_P \bar{\boldsymbol{v}}^w_{P_{UR}}(t)\ .\end{aligned} \tag{II.4}$$

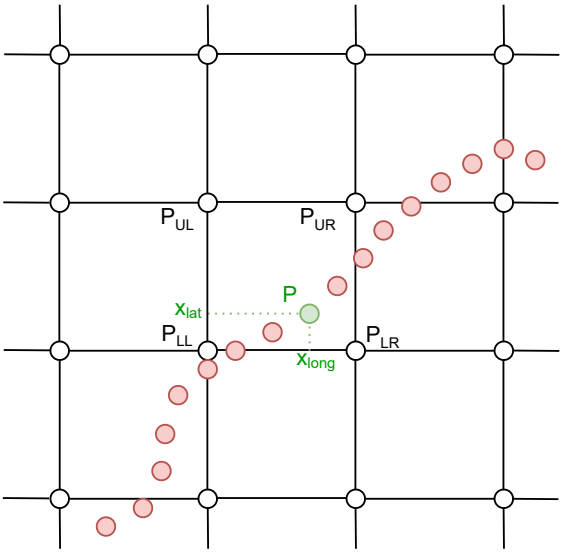

Figure II.3: Scheme describing the interpolation - The red dots represent the track, and the black dots represent the prediction points provided by Meteo France.

Equation (II.4) gives a linear approximation of the wind velocity along the track from the prediction grid. The aerodynamic forces depend on the relative wind velocity. Consequently, we have to project the wind velocity on the longitudinal track axis. To do so, we simply calculate the rotation angle between the track referential and the Cartesian referential. As we only have access to the longitudinal/latitude coordinates of the track, we propose to evaluate this rotation angle α_P at each point P on the track from the longitudinal/latitude of the previous and following points $P-1$ and $P+1$. It is equal to 0 by default when the track is vertical through the positive latitude. We write x_{P+1} and y_{P+1} the longitudinal and latitude coordinates of the point $P+1$. x_{P-1} and y_{P-1} are the longitudinal and latitude coordinates of the point $P-1$. If the track is perfectly horizontal, that is to say $x_{P+1} = x_{P-1}$, two cases exist: if $y_{P+1} > y_{P-1}$, we define $\alpha_P = 0$, otherwise, $\alpha_P = \pi$. When the track is not horizontal, with $x_{P+1} > x_{P-1}$, the angle α_P is given by:

$$\alpha_P = \arctan\left(\frac{y_{P+1} - y_{P-1}}{x_{P+1} - x_{P-1}}\right) + \frac{\pi}{2}\,. \tag{II.5}$$

In the other case, that is to say for $x_{P+1} < x_{P-1}$, the angle α_P is given by:

$$\alpha_P = \arctan\left(\frac{y_{P+1} - y_{P-1}}{x_{P+1} - x_{P-1}}\right) + \frac{3\pi}{2}\,. \tag{II.6}$$

Equations (II.5) and (II.6) give the angle α_P for every point P on the track. It is included in the interval $[0, 2\pi]$. Consequently, the wind speed at point P projected on the longitudinal to the track axis v_P^w can be written

$$\boxed{v_P^w(t) = \|\overrightarrow{\boldsymbol{v}}_P^w(t)\| \sin(\alpha_P)\,.} \tag{II.7}$$

With these notations, $v_P^w(t)$ is a deterministic scalar variable attached to a specific point P and it depends on time. It is positive for a tailwind and negative for a headwind. The lateral wind

can also be estimated with the same method, but it will not be used in this work as it does not impact the energy consumption in comparison to the longitudinal wind. Figure II.4 represents the cartography of one prediction grid associated with a given time t provided by Meteo France (in colors) and the projection on the track $v_P^w(t)$ (in black and white).

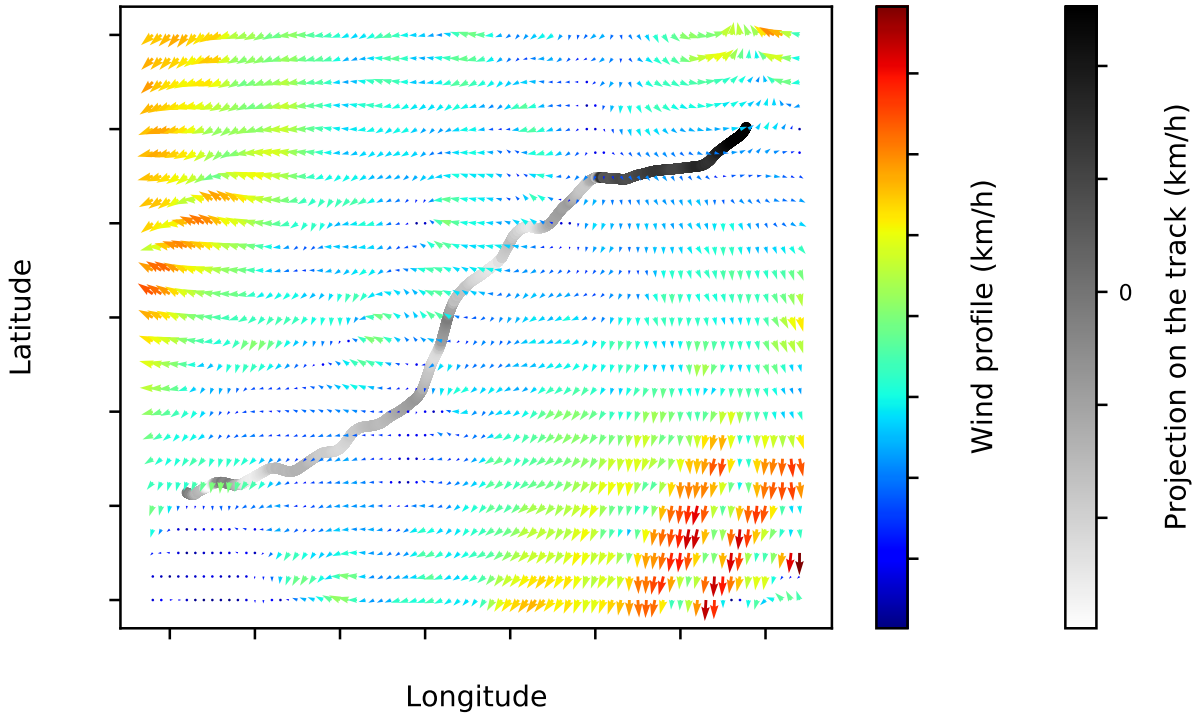

Figure II.4: Wind cartography and projection on the French LGV Rhin-Rhone track of its amplitude.

We can observe that the method presented before is efficient as the projection is positive when the wind "pushes" the train and is negative when it acts as a resistant force. It is around zero when the wind is perpendicular to the track. We also notice that the sign of the projection can vary depending on the curvilinear position. In Figure II.4, it can be seen that it is mainly positive at the beginning of the segment and is negative on the final segment.

Regarding the time discretization, we only have one prediction each hour. As the number of time for which the forecast is available is very small, we choose not to interpolate in time as we did spatially, because it would probably yield a nonrepresentative result. Consequently, we prefer to select the closest time forecast and to consider it as constant during the journey. In practice, with more forecasts, it is possible to give a better time description of the wind. Therefore, the wind speed projected on the longitudinal track axis only depends on the position on the track and not on time anymore. In the following sections, we prefer to denote the point P in the rolling environment $\mathcal{T}$ by its curvilinear abscissa s, and thus v_P^w is rewritten $v^w(s, \mathcal{T})$ along the track. This quantity can be used to estimate the aerodynamic forces applied to the train, as it has been done in Section II.2.2.

Electric energy supply

Between the electric transport network and the pantograph of the train, the electric energy passes through a substation, which provides electricity to the catenary. The substation positions are specific to every tracks. They have an impact on the available electric energy. Indeed, if the train is far from the closest substations, less energy is available, which can result in the limitation of the usable traction power. This constraint could have been implemented but, in this work, we assume that the maximum traction power available is only limited by the motor capacities.

In France, two types of electric energy supply are possible. On the one hand, the high-speed tracks are all supplied by alternative current (voltage $25kV$ and frequency $50Hz$). The same electrification type is present in most of the "classic" (regional) tracks of the south of France. On the other hand, many tracks of the north part of France are electrified with continuous current (voltage $1\,500V$). This organization results from the historic construction of French railway tracks. Figure II.5 represents the organization of electrification of the French railway network. In this work, we propose to study the TGV Dasye on the LGV Rhin-Rhone, which is only supplied by alternative current.

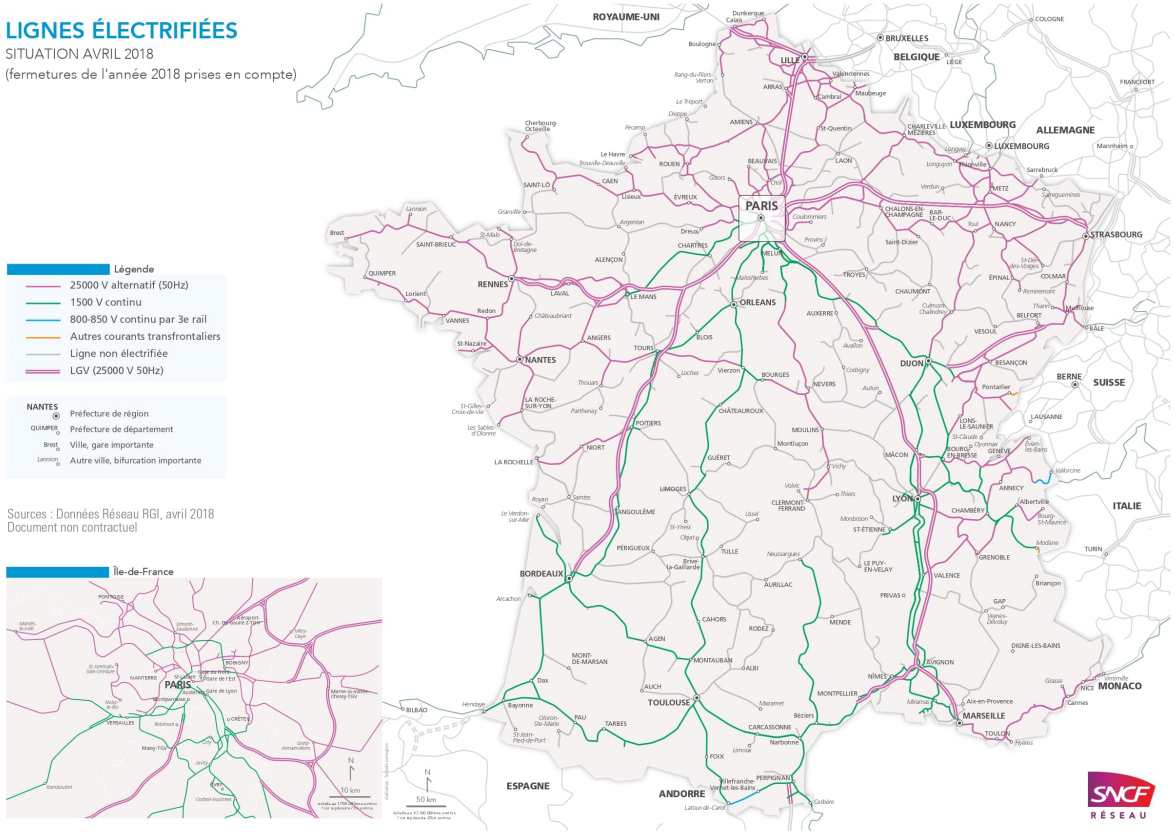

Figure II.5: Map of the types of electrification of the French railway network.

These aspects (positions of the substations and type of electrification) are discussed by R. Bosquet in [44] but will not be considered for reasons of simplicity.

II.1.2 Description of the vehicle

Another entry of the train system is the vehicle. All vehicles do not have the same reaction to a specific environment. It depends of course on the composition of the train but also on its wear state. In fact, the train is a complex nonlinear system, and the slightest change can modify its entire dynamic behavior. Most of commercial trains, and especially the high-speed train TGV Dasye, are composed of three different components: the wheelsets, the bogies, and the cars. All these components are linked to the others by nonlinear suspensions (spring, dampers, bushes, bump stops, etc.). The suspension design plays a specific role to assure comfort and security of the passengers. We can also dissociate the motor car from the other cars of the train. In practice, a French TGV Dasye is composed of two motor cars (situated at each extremity of the train), six passenger cars, four motor bogies (two for each motor cars), nine trailer bogies (some of them being shared by two cars), eight motor wheelsets, and eighteen trailer wheelsets. All these components and the suspensions have their own reactions to a given environment, which is difficult for even the most precise prediction.

The dissociation between motor or trailer components is also important to describe the traction, but also for the braking, as long as French high-speed trains have two different types of brakes. The first one, called the mechanic (or pneumatic) brake, relies on the friction of brake shoes on the braking discs present on the wheel axles. The second type of brake, the electrodynamic (also called dynamic) brake, corresponds to the electric inversion of the motors, which transforms the mechanic energy in electric energy (instead of the opposite in traction mode). Consequently, it is able to recover a part of the energy lost during braking, which will be reinjected in the catenary in order to be used by other trains. The maximum motor or braking capacities of each wheelset depend on the train speed. They can be expressed thanks to a torque applied to the wheelset, but in practice, we directly convert it in a longitudinal force, positive for traction and negative for braking. This conversion is possible if we treat the entire train as one element supposing that the forces are equally distributed to each type of wheelsets. Figure II.6 represents the maximum longitudinal traction or braking forces that are available for a given train speed.

We can observe in Figure II.6 that the motor and trailer wheelsets have their proper traction or braking capacities. These maximum capacities are nonlinearly dependent on the train speed and are specific to each type of train. For instance, the pneumatic braking of motor wheelsets is equal to zero at high speed. These specific capacities should be taken into account in the high-speed train dynamic model so as to have an accurate representation of the traction chain and of the brakes. In practice, the traction or braking forces are equally distributed between all types of wheelsets in order to limit the efforts between cars.

The train dynamics is moreover strongly linked to the wheel-rail contact. This contact is very complex to model because of the shape of the two surfaces in contact. In France, the most popular shape for the rail is the Vignole rail and it is inclined to $1/20$. The wheel is not cylindrical, as cars, but is slightly conical. This shape allows a good curve inscription and the recentering of the wheelsets in the track.

Finally, the shape of the train has an influence on the aerodynamic forces and is thus designed to reduce the drag force. The shape is however constrained by the gauge, different in each country.

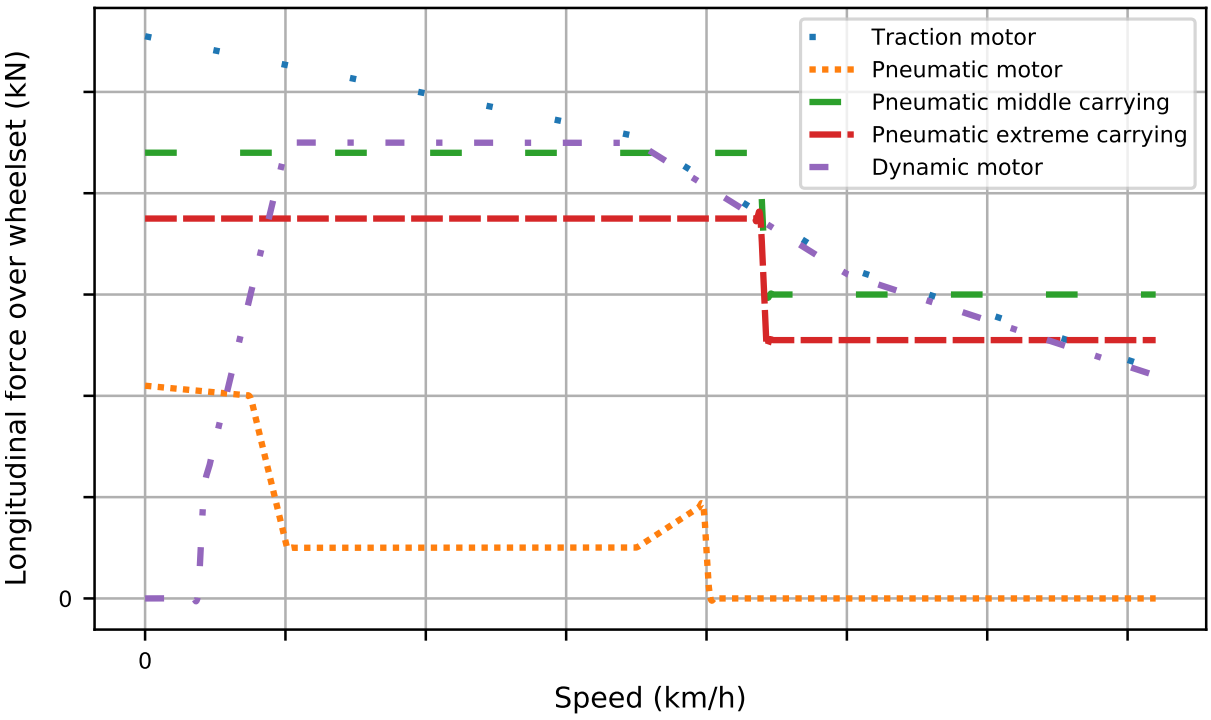

Figure II.6: Longitudinal forces available for the motors and the brakes in function of the train speed for the French TGV Dasye.

This section was dedicated to the description of the main characteristics of the rolling stock. They are not going to be defined in detail as they are not directly linked to the main matter of the work, meanwhile understanding the complexity of the train system is unavoidable. In the following sections, we write $\mathcal{V}$ all the parameters that describe the vehicle. Other parameters (like the mass of the train) depend on $\mathcal{V}$, but to simplify the notations, we propose to include them as well in $\mathcal{V}$. This is the case of the mass of the train, the auxiliary power, the three Davis coefficients, and the four efficiency parameters that are going to be described in the next sections. These parameters may be submitted to uncertainty, as we often do not know the exact wear state of the train or the exact number of passengers for instance. Consequently, we can consider that $\mathcal{V}$ gathers random variables to take into account these uncertainties.

II.1.3 Representation of the driver's command

The traction and braking applied to the train system are chosen by the driver through the driver's command. It is a manipulator, which can take any value between the maximum traction or maximum braking torque. As these values depend on the speed of the train (as shown in Figure II.6), we propose to normalize them and to write the driver's command as a time-dependent variable taking its value in the interval $[-1, 1]$. Consequently, -1 corresponds to the position of the manipulator, which represents the use of maximal braking torque. On the opposite side, 1 stands for the position of the manipulator corresponding to the use of maximum traction torque. The driver's command is written $\{u\}$ and is defined by:

$$\begin{array}{rccc} \{u\} & : & [t_s, t_f] & \rightarrow & [-1, 1] \\ & & t & \mapsto & u(t) \end{array}, \tag{II.8}$$

where t_s and t_f are the starting and final time. If $u(t) < 0$, the train is braking at current time t. On the other hand, a positive value of $u(t)$ is equivalent to the use of the motors for the traction at current time t.

The train dynamics at a given instant t is dependent on the actual value of the driver's command $u(t)$, but also on all the previous values $\{u_{\leq t}\}$ as the state of the system is strongly dependent on the previous states. For instance, braking when the train is already stopped does not have the same effect that braking when the train is at full speed. The notation $\{u_{\leq t}\}$ refers to the values of the function $\{u\}$ before time t. In other words, $\{u_{\leq t}\} = \{u(\tau); t_s \leq \tau \leq t\}$.

As explained in Section II.1.2, we prefer to convert the traction and braking torques in traction and braking longitudinal forces. To do so, we suppose the wheel radius constant. But the wheels are conic and might be damaged. This simplification gives a good estimation if we consider the mean over all the wheelsets. Consequently, $u(t)$ corresponds to the proportion of the maximum of longitudinal traction or braking force injected in the train system at time t. Function $\{u\}$ belongs to the space of continuously differentiable functions, since the physical quantities associated with the driver's command are also continuously differentiable.

II.2 Construction of the train dynamic model

Once the entries are properly represented, that is to say, when all the environmental factors are well defined, when the vehicle is precisely described, and when the driver's command has been defined, the train has a deterministic dynamic behavior. This dynamic behavior is complex, and the use of simplifications might be useful to model the train system. These simplifications introduce a model error, which should be monitored in order to preserve the realistic character of the representation.

The developed computational model has to be sufficiently precise and with a small numerical cost, which allows the quantities of interest to be extracted. In this work, we take into account the longitudinal traction and braking forces and the train speed in order to calculate the energy consumption of the train (as it is the quantity we are going to minimize). Consequently, the next sections propose a dynamic computational model to describe the train mechanical behavior. In particular, Section II.2.1 focuses on the choice of modeling framework, as several points of view can be adopted. The external forces are then presented in Section II.2.2. Finally, Section II.2.3 proposes an expression for the longitudinal dynamic equation that is going to be used in the rest of the work. The train speed and its position are very important to estimate the constraints. For instance, the punctuality constraints (defined in Section IV.2.1) verify if the train arrives at the exact position, at a given final time, and with a given speed (that will be zero).

II.2.1 Choice of the modeling framework

The first step of the modeling is to choose an appropriate framework. Without any doubt, we could model the entire train, take all the components apart, and choose complex models to represent all the interactions between them. An overview of many dynamic phenomena is proposed by G. Boschetti in [45]. Nevertheless, it is not always the best choice as it may lead to an important calculation time. Essentially, calculating with a high precision the deformation

of one wheelset may be useless if we focus on the energy consumed by the train. To this extent, two approaches are frequently used in the literature.

For the first approach, the rigid body framework proposes to consider each component of the train (cars, bogies, and wheelsets) as nondeformable bodies, and attention is only paid to the internal interactions (suspensions) and to the external interactions (with the environment). An example of this approach is given by N. Bosso *et al* in [46] to represent the whole train dynamics. This point of view presents the advantage of neglecting the deformation of the car bodies, the bogies, and the wheelsets (which are of little importance in this work), while preserving a good precision by modeling the nonlinear suspensions. The dynamic equations can directly be expressed from the Lagrangian equations. Some industrial software has already been developed to solve these equations and are adapted to the railway system such as VAMPIRE®. Nevertheless, the calculation time is around one minute for long journeys, which is relatively important for solving the driver's command optimization problem.

The second approach is to consider the train as one element. The internal interactions are considered constant and the whole train behaves as one rigid body. Consequently, the problem is simplified as only the external interactions impact the train dynamics. This method is presented by Q. Wu *et al* in [47]. This approach is linked to the previous one as it is demonstrated in Appendix A. Moreover, understanding the connection between the two allows us to have a better idea of the hypotheses that are necessary to proceed from one approach to the other. This method considerably reduces the calculation costs as the dynamic behavior is summarized in one longitudinal equation.

In this work we choose to use this second approach because it accelerates the calculation without significantly affecting the estimation of the quantities of interest: energy consumption, speed profile, and position of the train.

II.2.2 Description of the external forces

We take into consideration a train running in a rolling environment $\mathcal{T}$, with a well described set of parameters $\mathcal{V}$, and a given driver's command $\{u\}$. This train goes from a starting point s_s to a final point s_f (position of the train stations). The departure and arrival times are denoted by t_s and t_f (scheduled by the railway operator except if we consider the case of delayed journeys), and the initial and final speeds are v_s and v_f (zero as the train arrives at a station).

We highlight here that, in this work, a duality exists between time and position as we mainly work on the speed profile of the train. Some figures represent the speed profile of the train and its energy consumption depending on time to verify the departure and arrival time, but others represent the same solution plotted against position to compare the speed profile together with the speed limitation. The two quantities are directly linked.

During the journey, the train moves along the track with a specific longitudinal position $y(\{u\}, \mathcal{T}, \mathcal{V})$, speed $\dot{y}(\{u\}, \mathcal{T}, \mathcal{V})$, and acceleration $\ddot{y}(\{u\}, \mathcal{T}, \mathcal{V})$. These time dependent functions depend not only on the actual value of the driver's command, but also on the previous instants. At a given time t, the longitudinal position and speed are thus written as

$$y(t) := y(t; \{u_{\leq t}\}, \mathcal{T}, \mathcal{V}) \quad , \quad \dot{y}(t) := \dot{y}(t; \{u_{\leq t}\}, \mathcal{T}, \mathcal{V}) \ . \tag{II.9}$$

The complete dynamic solution is written as

$$\mathcal{S}(\{u\}, \mathcal{T}, \mathcal{V}) = \{y(\{u\}, \mathcal{T}, \mathcal{V}), \dot{y}(\{u\}, \mathcal{T}, \mathcal{V}), \ddot{y}(\{u\}, \mathcal{T}, \mathcal{V})\} \ .$$

This dynamic solution is specific to the driver's command, the rolling environment, and the vehicle. With these notations, modeling the train longitudinal dynamics requires to estimate all the forces applied to the train system. Some interactions are difficult to model, such as the friction forces. The traction and braking forces are very important as they directly come from the driver's command. Finally, the weight of the train also has an important impact.

Friction forces

First, the aerodynamic forces are quite difficult to represent as the shape of the train is very specific, especially for high-speed trains, for which the prediction with computational fluid dynamics is not so easy. Some works focus on the wind effects such as X. Quost in [48], with the intention of modeling the rollover of a train caused by strong wind.

In the same manner, the wheel-rail contact forces are nonlinear and difficult to model. The Hertzian contact, proposed by H. Hertz in [49], combined with Kalker's theory [50], estimates the lateral contact forces, and is often used in the railway domain. These contact models need the exact geometry of the rail as well as the wheel of the train, but also many other parameters such as the air humidity and the friction coefficient between others. These models allows us to have a precise idea of the wheel-rail tridimensional contact and of the associated efforts. An example of the application of these theory applied to the railway system is given by B. Pecile in [51].

As we have previously explained, we assume that only the mean speed of the wind has an influence on the consumed energy for a complete journey. The average wind on the track is estimated in Section II.1.1. Therefore, we do not model the tridimensional aerodynamic and contact forces, but only the longitudinal drag and contact forces, as it fits with the chosen approach (longitudinal dynamics for the whole train). The Davis forces, first proposed by William J. Davis in [52], are a well-known simplification often used in railway literature, which approximate the longitudinal friction forces applied to the train with a relatively simple expression. The longitudinal friction force f^R applied to the train at a given time t is estimated by:

$$f^R(\dot{y}(t), v^w(y(t), \mathcal{T}), \mathcal{V}) = a(\mathcal{V}) + b(\mathcal{V})\,\dot{y}(t) + c(\mathcal{V})\,(\dot{y}(t) - v^w(y(t), \mathcal{T}))\left|\dot{y}(t) - v^w(y(t), \mathcal{T})\right| . \tag{II.10}$$

In this expression, v^w is the mean wind speed, the coefficient c refers to an aerodynamic coefficient, the coefficient b is associated with the dynamic friction, and the coefficient a refers to the static friction. Thus, a and b correspond to the longitudinal contact force and c to the longitudinal aerodynamic force applied to the train. These three coefficients are provided by the train manufacturer for each type of vehicle. The friction force depends thus on $\mathcal{V}$.

The term $\dot{y}(t) - v^w(y(t), \mathcal{T})$ represents the longitudinal relative train speed with respect to the wind speed. When the wind "pushes" the train at a position $y(t)$, the associated average wind speed $v^w(y(t), \mathcal{T})$ is positive (as explained in Section II.1.1). Hence, the square term of Equation (II.10) is lower, that is to say the resistance force is lower. Finally, the aerodynamic part of Equation (II.10) is written with only one absolute value of the relative speed, so that this term can be negative in case of a train speed $\dot{y}(t)$ smaller than the wind speed $v^w(y(t), \mathcal{T})$.

This approximation of the longitudinal friction force can be discussed as it is a very simplified model, but yet allows us to estimate this force without going through the use of complex models. In practice, it has been observed that the friction force is higher in curve, where the wheel-rail friction is higher. Consequently, we propose to use a corrective term also frequently met in the railway literature. This corrective term depends on the mass of the train, as the friction force in the curve is not the same for heavy trains than for light ones, but it also depends on the curve radius of the track r, as the friction is higher in tight curves. For this reason the corrective term f^C is conveyed depending on the specific curve radius at the train position. f^C is generally written

$$f^C(y(t), \mathcal{T}, \mathcal{V}) = m(\mathcal{V})\, g \frac{k^C(\mathcal{V})}{r(y(t), \mathcal{T})}, \tag{II.11}$$

where $m(\mathcal{V})$ is the mass of the train. It can be considered as a random variable because the exact mass of the train is unknown due to the number of passengers, which is not precisely known. The coefficient k^C is estimated by the train manufacturer for each specific vehicle $\mathcal{V}$, g is the gravity acceleration, and r is the curve radius of the track.

Equations (II.10) and (II.11) give an estimation of the longitudinal friction force applied to the train. Nevertheless, the friction is nonlinearly dependent on the train speed because of the aerodynamic term. The Davis approximation and the choice of the values for the coefficient a, b, and c is detailed by G. Boschetti *et al* in [53].

Traction and braking forces

The second group of forces, which has a great influence on the train dynamics, is the traction and braking forces. These forces are applied by the motors and the brakes directly to the motor or carrying wheelsets as a torque whose amplitude is defined by the driver. Here, we suppose that all the motor torques are transmitted to the contact. Consequently, it can be converted in a global longitudinal traction or braking force applied to the whole train system.

For a given train speed, the maximum traction or braking force available is specific to the vehicle $\mathcal{V}$. The forces for the French TGV Dasye are shown in Figure II.6. We write the maximum traction (resp. braking) force available $f^{T,max}(\dot{y}(t), \mathcal{V})$ (resp. $f^{B,max}(\dot{y}(t), \mathcal{V})$) and it simply corresponds to the sum of the maximum traction (resp. braking force) available over each motor (resp. brake). In this case, we need to call to mind that some trains have several types of brakes. The TGV Dasye combines both mechanical and electrodynamic brakes. Consequently, the maximum braking available is the sum of all the kind of brakes that can be found in the train,

$$f^{B,max}(\dot{y}(t), \mathcal{V}) = f^{M,max}(\dot{y}(t), \mathcal{V}) + f^{E,max}(\dot{y}(t), \mathcal{V}), \tag{II.12}$$

where $f^{M,max}$ and $f^{E,max}$ denote the maximum braking force available from mechanical and electrodynamic brakes. As it is explained in Section II.1.2, the traction and braking forces are equally distributed over each type of wheelsets to limit the efforts between cars. Thus, the maximum traction, mechanical and electrodynamic braking forces are equivalent to the unitary capacity available on one wheelset (which can be read in Figure II.6) multiplied by the number of wheelsets having the ability to tract or to brake by mechanical or electrodynamic equipment.

The driver's command is defined as the proportion of the maximum traction or braking force used. At time t, if $u(t)$ is positive, the braking force is equal to zero and only traction is used. On the opposite, if it is negative, only braking is used, and the traction force is equal to zero. With these notations, the longitudinal traction and braking forces f^T and f^B are given (due to the normalization of $u(t)$ in $[-1, 1]$) by:

$$f^T(\dot{y}(t), u(t), \mathcal{V}) = \begin{cases} u(t) f^{T,max}(\dot{y}(t), \mathcal{V}) & \text{if } u(t) > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{II.13}$$

$$f^B(\dot{y}(t), u(t), \mathcal{V}) = \begin{cases} 0 & \text{if } u(t) > 0 \\ u(t) f^{B,max}(\dot{y}(t), \mathcal{V}) & \text{otherwise.} \end{cases} \tag{II.14}$$

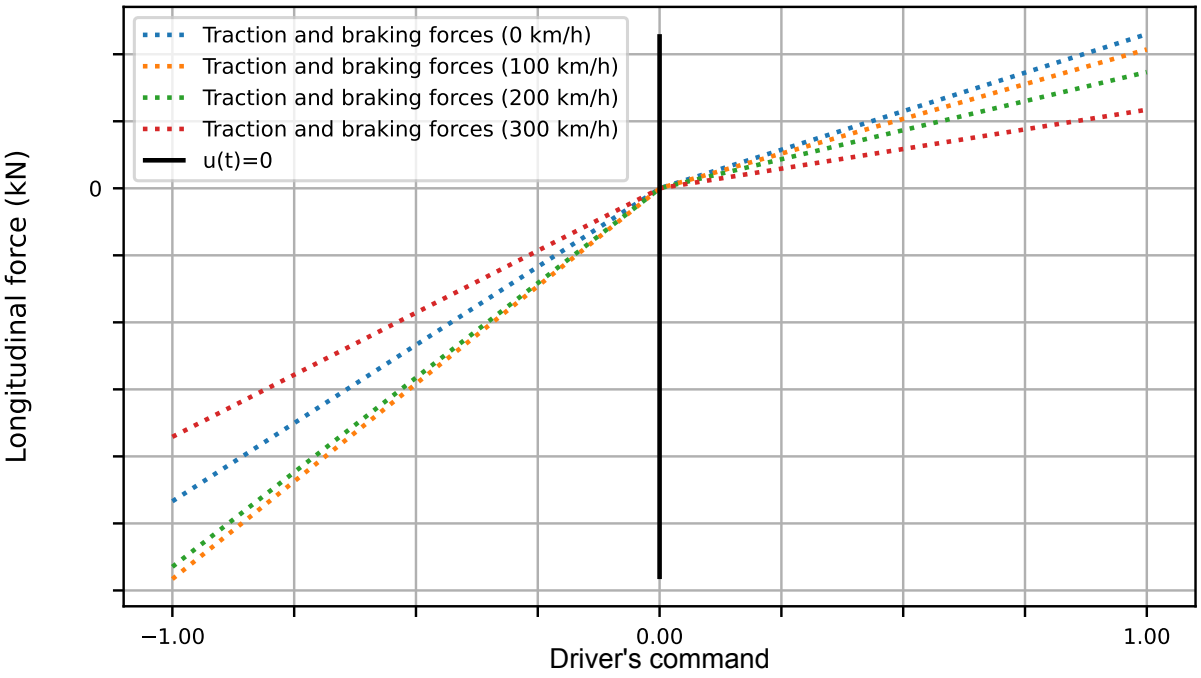

Figure II.7: Longitudinal traction and braking forces applied to the train depending on the driver's command for several train speeds.

Figure II.7 represents the longitudinal traction or braking force applied to the train at a given time t depending on the driver's command $u(t)$ for several specific train speeds (0, 100, 200, and $300 km/h$). We can observe that the traction or braking force applied to the train is not differentiable at $u(t) = 0$ for a constant train speed, because this driver's command is the

interface between the traction and the braking models. When the driver's command is positive, it quantifies the proportion of maximum traction used, whereas it describes the proportion of maximum braking when negative. Moreover, if the speed of the train is not constant, the maximum traction or braking available also changes, despite a constant driver's command. For $u(t) = -1$ or $u(t) = 1$, the equivalent longitudinal force is the sum of all the braking or traction forces available at each wheelsets (represented in Figure II.6). The traction force available decreases with the train speed, which is logical according to Figure II.6. The same analysis with the braking force is not that simple. Indeed, some wheelsets are only equipped with pneumatic braking and others with dynamic braking.

All these reasons highlight the fact that even if the mathematical representation of the driver's command is relatively simple (linear proportion of the maximum force available), the resulting longitudinal force is quite complex. It essentially depends on the train speed and the force is not differentiable.

Weight

With respect to the impact of the track declivity on the energy consumption, the weight of the train has no impact on a flat track because it is perpendicular to the train movement. In case of a ramp (positive declivity) or a slope (negative declivity), its projection on the longitudinal axis is not equal to zero and has to be taken into account. In case of a ramp, the weight has a negative resulting force once projected on the longitudinal to the track axis: it pulls the train backwards. For a negative declivity (slope), the projection of the force is positive, and the weight helps the train to move forwards. The projection of the weight on the longitudinal to the track axis f^W is given by:

$$f^W\left(y\left(t\right), \mathcal{T}, \mathcal{V}\right) = -m\left(\mathcal{V}\right) g \sin\left(\arctan\left(\theta\left(y\left(t\right), \mathcal{T}\right)\right)\right) . \tag{II.15}$$

For instance, if the track declivity is equal to 1 at a given curvilinear abscissa, that is to say a ramp of 1 meter per meter advanced, the associated track angle has to be equal to $\arctan\left(1\right) = \frac{\pi}{4}$. For such a declivity, the projection of the weight pulls the train backwards. Thus, the projection of the weight has to be negative, (negative sign in Equation (II.15)). This equation also shows that this force is dependent on the train position.

II.2.3 Longitudinal dynamics for high-speed trains

As we explained in the previous section, we only consider the train longitudinal dynamics of one train. Details on the hypotheses and the developments used to extract the longitudinal dynamics from the multi-body approach are provided in Appendix A. In this section, we only highlight the resulting dynamic equation, which will be used in the next chapters, and which is written as

$$\begin{aligned} m\left(\mathcal{V}\right) k^{rot} \ddot{y}\left(t\right) = {} & f^T(\dot{y}\left(t\right), u(t), \mathcal{V}) - f^B(\dot{y}\left(t\right), u(t), \mathcal{V}) \\ & - f^R\left(\dot{y}\left(t\right), v^w\left(y\left(t\right), \mathcal{T}\right), \mathcal{V}\right) - f^C\left(y\left(t\right), \mathcal{T}, \mathcal{V}\right) + f^W\left(y\left(t\right), \mathcal{T}, \mathcal{V}\right) , \end{aligned} \tag{II.16}$$

for $t \in [t_s, t_f]$ and with initial Cauchy conditions. It is assumed that for given u, $\mathcal{T}$, and $\mathcal{V}$, and without constraints, this nonlinear differential equation has a unique solution. The left-hand

side of Equation (II.16) represents the inertial term, including the rotation of the wheels thanks to the corrective factor k^{rot}. In the literature, we often find $k^{rot} = 1.04$ (the explanation of this value is given in the development made in Appendix A). The right-hand side stands for the external forces applied to the train. We recognize the forces previously listed, that is to say the traction and braking forces (which depend on the driver's command), the resistant force (estimated thanks to the Davis equation (II.10)), the corrective term applied in curves, and the weight of the train projected on the longitudinal track axis.

From the entries and driver's command presented in Section II.1, Equation (II.16) allows us to simulate the complete longitudinal dynamics of the train. In practice, we are mostly interested in the train speed and the train position, as well as the traction and braking forces. Surely, these quantities are the key to compute the energy consumed by the train, while verifying all the constraints linked to the train system (which will be described in Section IV.2.1).

II.3 Definition of the energy consumption model for the train

Once the longitudinal dynamics of the train is defined, we have to construct an energy model to compute the energy consumption of the train. It depends on the driver's command and is directly linked to the train dynamics. Therefore, the train speed is required to estimate the energy consumption. In the next subsections, we propose to present the energy consumption model. We begin by describing the auxiliary power in Section II.3.1, before presenting the energy-efficiency models in Section II.3.2. Section II.3.3 focuses on the whole energy consumption model.

II.3.1 Description of the auxiliary power

The electrical power transmitted to the train system has several distinct uses. Obviously, a major part is used by the traction chain to assure the train motion, but a second part is transmitted to the auxiliary equipment of the train. These equipment assure the security of the train by the use of different air-cooling systems that guarantee the proper functioning of the traction chain. They also assure the comfort of the passengers thanks to the combination of the air conditioning, the provision of electrical outlets, the lights, etc. The electrical power transmitted to the auxiliary equipment is called the auxiliary power and is denoted by p^a. It depends on $\mathcal{V}$. The electrical power transmitted by the catenary is denoted by p^E and is such that:

$$p^E\left(\dot{y}\left(t\right), u(t), \mathcal{V}\right) = p^T\left(\dot{y}\left(t\right), u(t), \mathcal{V}\right) - p^B\left(\dot{y}\left(t\right), u(t), \mathcal{V}\right) + p^a\left(\mathcal{V}\right), \tag{II.17}$$

where p^T is the electrical power estimated during traction phases, and p^B is the electrical power recovered by the braking (this explains the minus sign). The recovered energy returns to the electrical network by the catenary. Consequently, we assume that the energy recovered by the braking is fully used, that is why we simply remove it from the electrical power (Equation (II.17)). During traction phases, the recovered energy is equal to zero and $p^B = 0$ and during braking phases, the traction chain does not consume energy and $p^T = 0$.

The term p^a is not a constant, as it depends on the outside temperature (as it is linked to the use of air conditioning), on the number of passengers (as it is linked to the use of the electrical outlets),

and on the driver's command (as it is related to the use of the air-cooling systems). Nevertheless, in this work, we consider p^a as a constant (variations sufficiently small). A first estimation of the value of the auxiliary power is performed by estimating the energy consumption of the train when it is stopped. This estimation does not include the electrical energy consumed by the air-cooling systems. For this reason, we consider that p^a is not perfectly known, and its uncertainty will be taken into account in Chapter III.

II.3.2 Construction of models representing the energy efficiency

All the electrical power transmitted to the traction chain is not converted in mechanical power. A part of it is lost due to the friction effects and the heat dissipation of the traction chain components. In the same way, all the mechanical power lost by the electrodynamic brakes is not recovered. More precisely, the traction efficiency, η^T, is defined by the proportion of electrical power that is effectively converted into mechanical power during traction phases. On the opposite side, the braking efficiency, η^B, describes the proportion of mechanical power that is injected in the catenary. An example of a 2D-plot of traction efficiency of a French regional train is given in Figure II.8.

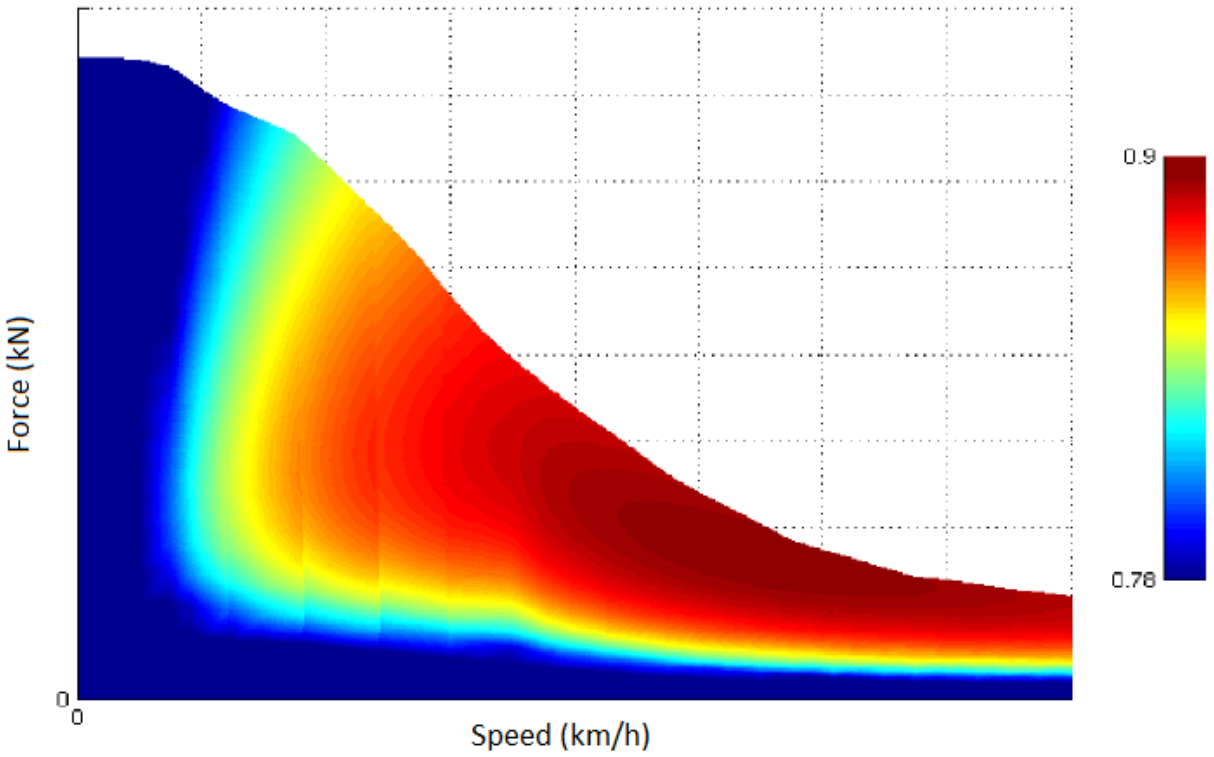

Figure II.8: 2D-plot of the efficiency of the traction chain of a French regional train, according to M. Debruyne [54].

We can notice that the efficiency is not constant. It depends on the train speed, but also on the driver's command itself. As we do not have access to the equivalent 2D-plot for high-speed

trains and we do not have information on the braking efficiency, we propose to construct two models, which is inspired from Figure II.8. Even if the traction chain is responsible for the traction and braking efficiency as it converts electrical power into mechanical power (or vice versa during braking), there is no reason for the two efficiencies to be equal. Nevertheless, we consider that they can be approximated with the same model, but with different values for the parameters. Regarding Figure II.8, we propose an affine function of p^T and p^B respectively for η^T and η^B,

$$\eta^T\left(\dot{y}\left(t\right), u(t), \mathcal{V}\right) = a_\eta\left(\mathcal{V}\right) p^T\left(\dot{y}\left(t\right), u(t), \mathcal{V}\right) + b_\eta\left(\mathcal{V}\right) \ , \tag{II.18}$$

$$\eta^B\left(\dot{y}\left(t\right), u(t), \mathcal{V}\right) = c_\eta\left(\mathcal{V}\right) p^B\left(\dot{y}\left(t\right), u(t), \mathcal{V}\right) + d_\eta\left(\mathcal{V}\right) \ , \tag{II.19}$$

where a_η, b_η, c_η, and d_η are four parameters that have to be indirectly identified using information on the electric energy, except that the global efficiencies are included in $[0, 1]$. As a first approximation, we propose to define the values for these parameters by imposing the boundary values of the 2D-plot in Figure II.8. With these values, the simulated 2D-plot is represented in Figure II.9.

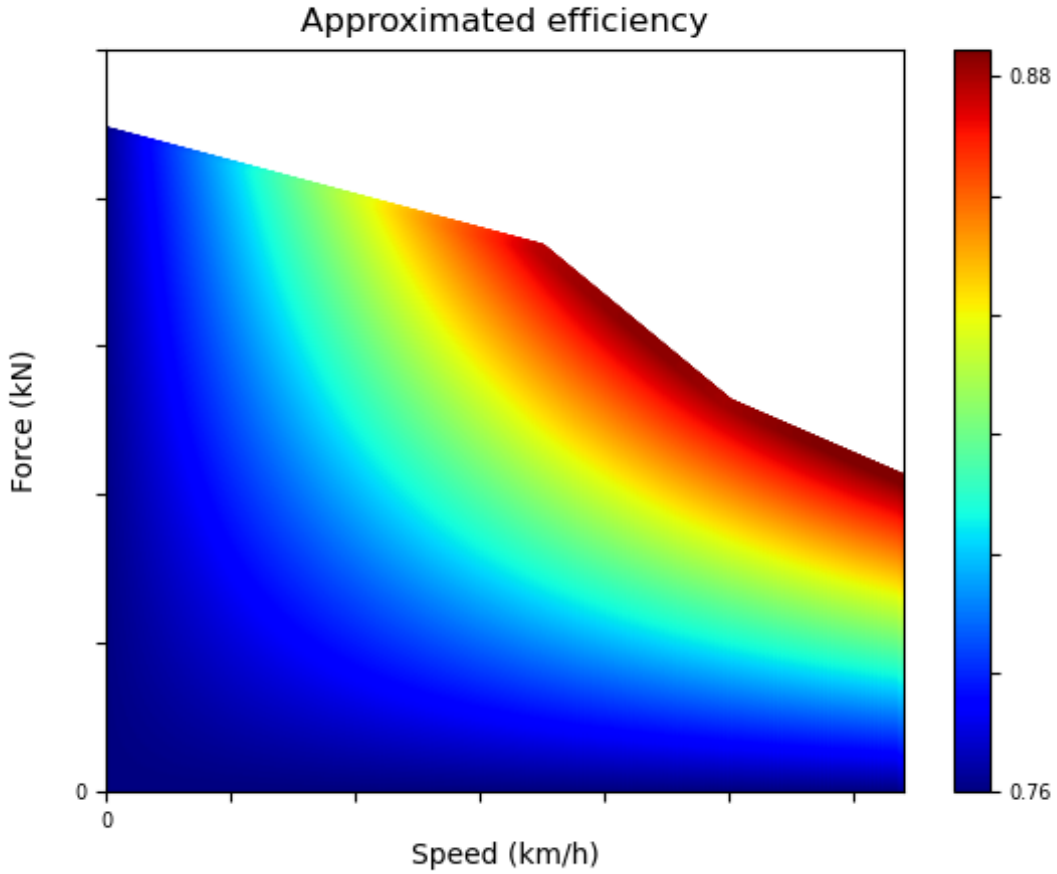

Figure II.9: 2D-plot of the efficiency of the traction chain simulated.

The 2D-plot boundary values are similar in the two figures. Even if the 2D-plot are not exactly the same, we consider these models sufficient as a first approach. However, uncertainties will be

implemented for these four parameters and will be identified in Chapter III. With these models, we can estimate the electrical power p^T and p^B during traction and braking phases, which are such that:

$$\eta^T\left(\dot{y}\left(t\right), u(t), \mathcal{V}\right) p^T\left(\dot{y}\left(t\right), u(t), \mathcal{V}\right) = f^T(\dot{y}\left(t\right), u(t), \mathcal{V})\, \dot{y}\left(t\right) \,, \tag{II.20}$$

$$p^B\left(\dot{y}\left(t\right), u(t), \mathcal{V}\right) = \eta^B\left(\dot{y}\left(t\right), u(t), \mathcal{V}\right) f^E(\dot{y}\left(t\right), u(t), \mathcal{V})\, \dot{y}\left(t\right) \,, \tag{II.21}$$

where f^E is the longitudinal electrodynamic braking force. Note here that the traction and braking efficiencies are defined by two different points of view. On the one hand, the traction efficiency describes the quantity of electrical traction power converted into mechanical power. On the other hand, the braking efficiency is defined in inverse, that is to say the quantity of mechanical power that is injected in the catenary. For this reason, the inverse of the traction efficiency is used in Equation (II.20), and the braking efficiency is directly used as it is in Equation (II.21). This choice of notation is justified so that the efficiencies are always included in $[0, 1]$. The electrical power is estimated multiplying the longitudinal force (traction f^T or braking f^B) by the train speed $\dot{y}\left(t\right)$ and applying the traction or braking efficiency presented in Equations (II.18) and (II.19).

II.3.3 Energy consumption for high-speed trains

For a given driver's command, we can solve the train dynamic equation, and estimate the associated train speed. From these quantities (train speed $\dot{y}$ and driver's command $\{u\}$), we can estimate the traction and braking efficiencies (Equations (II.18) and (II.19)), which allow us to calculate the traction and braking powers from Equations (II.20) and (II.21). Then, Equation (II.17) can be used to estimate the electrical power. The energy consumed by the train, noted $\mathcal{F}$ in the following, is simply the integral over time of the electrical power. Consequently, it is given by:

$$\boxed{\mathcal{F}\left(\dot{y}, \{u\}, \mathcal{V}\right) = \int_{t_s}^{t_f} p^E\left(\dot{y}\left(t\right), u(t), \mathcal{V}\right) dt\,.} \tag{II.22}$$

As the train speed also depends on driver's command $\{u\}$, rolling environment $\mathcal{T}$, and vehicle $\mathcal{V}$, we propose to rewrite the energy consumption as a function of $\{u\}$ and $\mathcal{V}$: $\mathcal{F}\left(\dot{y}, \{u\}, \mathcal{V}\right) = \mathcal{F}\left(\{u\}, \mathcal{T}, \mathcal{V}\right)$. This model allows for estimating the energy consumed by the train for a given driver's command $\{u\}$, a specific rolling environment $\mathcal{T}$, and vehicle $\mathcal{V}$.

II.4 Sensitivity analysis

The dynamic and energy models defined before are very sensitive to driver's command $\{u\}$, rolling environment $\mathcal{T}$, and the train characterized by $\mathcal{V}$. The system scheme of the model is summarized in Figure II.10. We propose to analyze the sensitivity of the two models for a subset of the parameters gathered in $\mathcal{V}$. These parameters include both dynamic parameters, as train mass m, Davis coefficients a, b, and c, but also energy consumption parameters, like auxiliary power p^a, traction efficiency parameters a_η and b_η, and braking efficiency c_η and d_η. Figures II.11 to II.19 represent the impact of a variation of $\pm 20\%$ over each model parameter (taken one-by-one) on the dynamic model (on the left) and the energy consumption model (on the right). Equations (II.16) and (II.22) are numerically solved for a given environment $\mathcal{T}$ (the

French LGV Rhin-Rhone track), a specific vehicle characterized by $\mathcal{V}$ (the French Dasye high-speed train), and a driver's command $\{u\}$ that is identified in inverse from the measurements (see Section III.2). The associated speed profile and energy consumption simulated by the models are plotted in Figures II.11 to II.19. We propose to describe the sensitivity of the models first with respect to the dynamic parameters and then with respect to the energy consumption parameters.

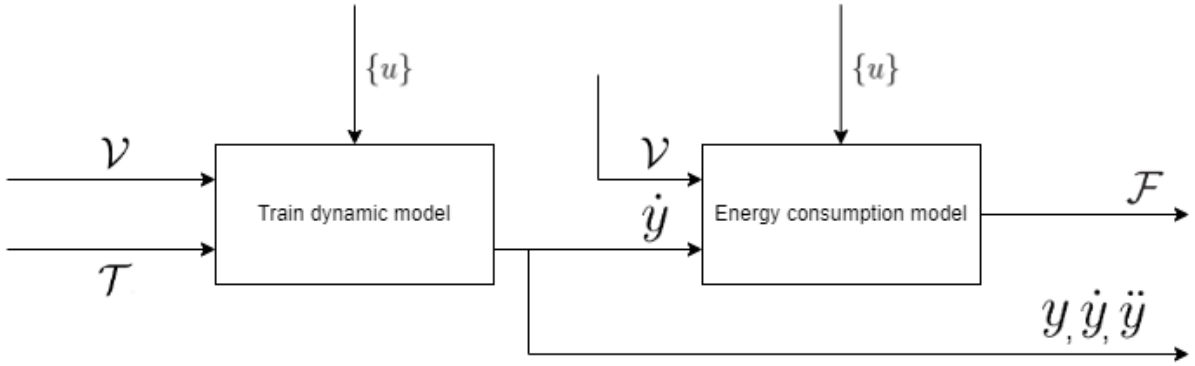

Figure II.10: System scheme of the model: inputs, driver's command, and outputs.

II.4.1 Sensitivity to the dynamic coefficients

Obviously, mass m is an important parameter in the dynamic model described by Equation (II.16). Figure II.11 displays the speed profile and the consumed energy as a function of m. It can be seen the importance of mass m in the model. Effectively, at the final time, the variation of energy consumption is around 14% for a variation of mass of $\pm 20\%$. This is quite logical, as the energy consumption model is closely linked to the dynamic model as it uses the output of the latter to estimate the energy consumed by the train along the journey. In practice, variations of $\pm 20\%$ cannot be encountered in reality, as the mass of the train only varies around $\pm 5\%$ around a mean value.

The second group of model parameters gathers the three Davis parameters, a, b, and c (see Equation (II.10)). A variation of their values has a different influence on the resistant force (and in the same way on the longitudinal dynamics). We can see in Figure II.12 that the speed profile is slightly modified by a variation of coefficient a. The associated energy consumption is almost unchanged (less than 0.25% at the final time when applying $\pm 20\%$ variations). Coefficient a does not have a significant impact on the high-speed train, because for the high speed the square velocity term is dominant, which depends on c.

Figure II.13 represents the impact of a variation of coefficient b. As for coefficient a, coefficient b has an important influence on the resistant force, and thus on the dynamic response (but has higher influence at high speed because b is multiplied by the train speed). This impact can be seen all along the journey but is more visible when the train speed is over $100km/h$ compared with Figure II.12. The influence on the energy consumption model is still small (around 0.5% at the final time when applying $\pm 20\%$ variations) but is superior to the influence of coefficient a.

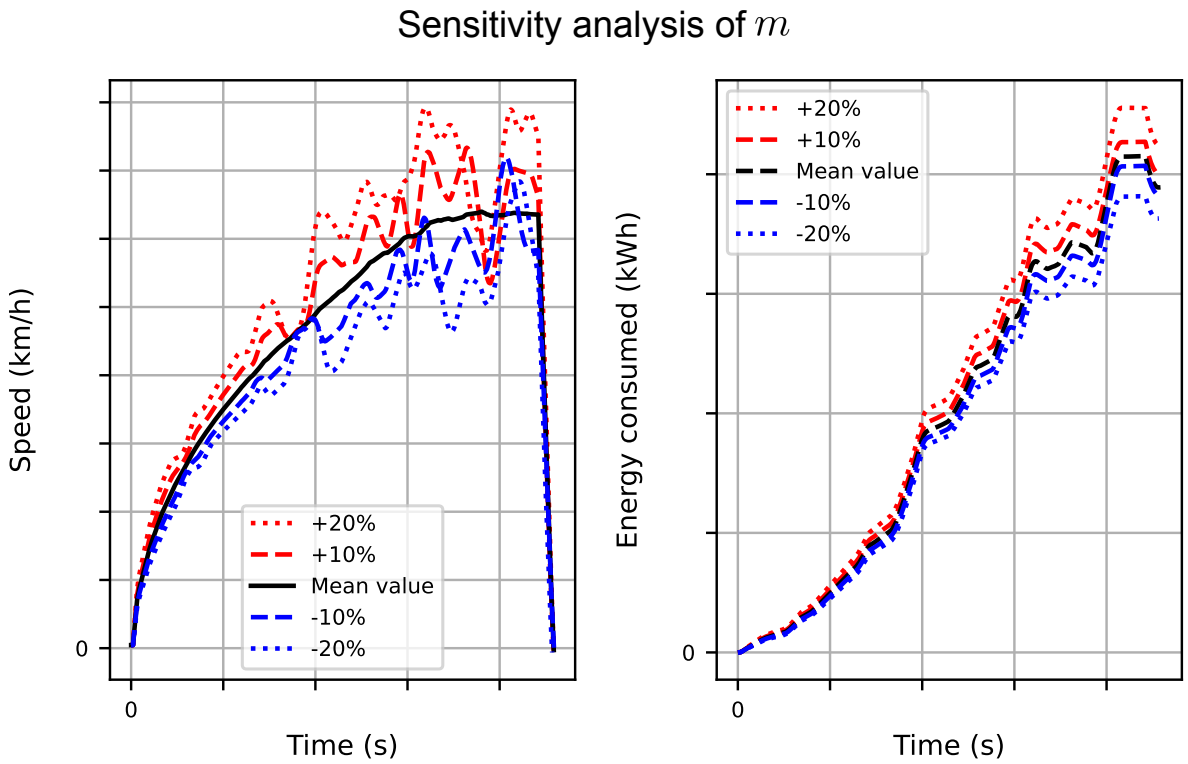

Figure II.11: Sensitivity of the models to mass m.

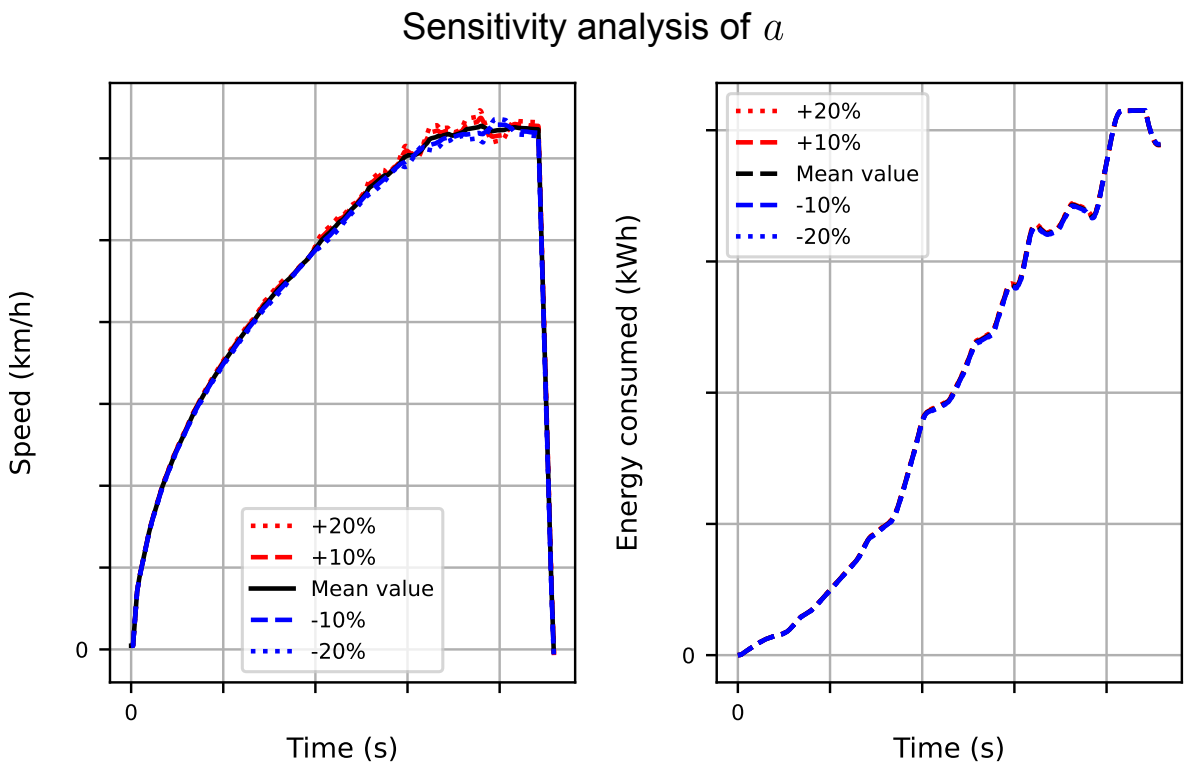

Figure II.12: Sensitivity of the models to Davis coefficient a.

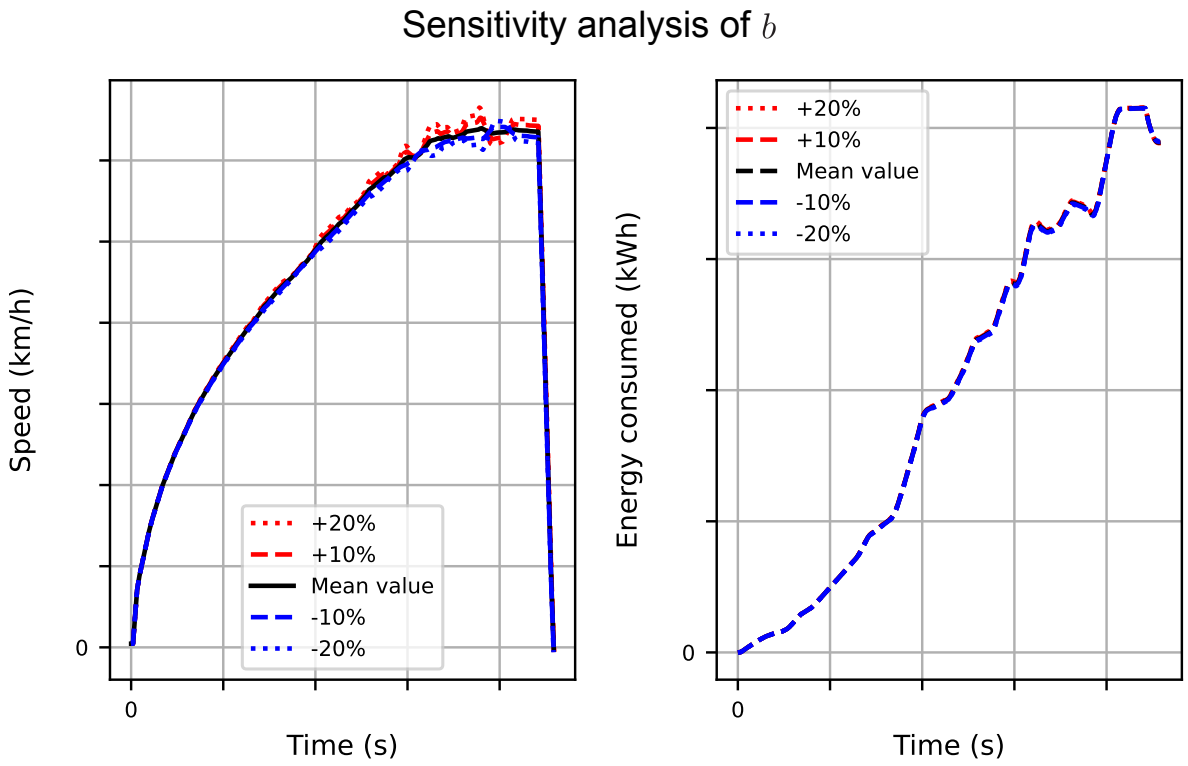

Figure II.13: Sensitivity of the models to Davis coefficient b.

Finally, the influence of coefficient c is represented in Figure II.14. It is associated with the square of the train speed, and thus has an important influence at high speed. We can notice that on the left figure, where the train speed profile is modified a lot, when the train speed is around $150km/h$. The influence of energy consumption is still relatively low (around 0.75% at the final time when applying $\pm 20\%$ variations) and cannot be observed without any zoom. On longer journeys, where higher speed is involved, the influence of these three parameters could be more important.

The sensitivity analysis performed on the dynamic parameters has enhanced their influence on the train dynamics and the energy consumption. We demonstrate that they mainly have an impact on the dynamic model, as they are directly linked to it, but they also have an influence on the energy consumption model, because the two models are directly linked. The sensitivity to the energy consumption parameters is analyzed in the next section.

II.4.2 Sensitivity to the energy consumption parameters

A brief sensitivity analysis is also carried out for the energy consumption parameters. The first one is auxiliary power p^a that we have assumed to be a constant and is represented in Figure II.15. This parameter has no influence on the train dynamics (as it only appears in Equation (II.22)), but we can observe that perturbing it modifies the associated energy consumption. This influence can be determined analytically. Effectively, it is simply added to the traction and braking power to obtain the electrical power (Equation (II.17)) and is then integrated over time to obtain the energy consumption (Equation (II.22)). Consequently, to

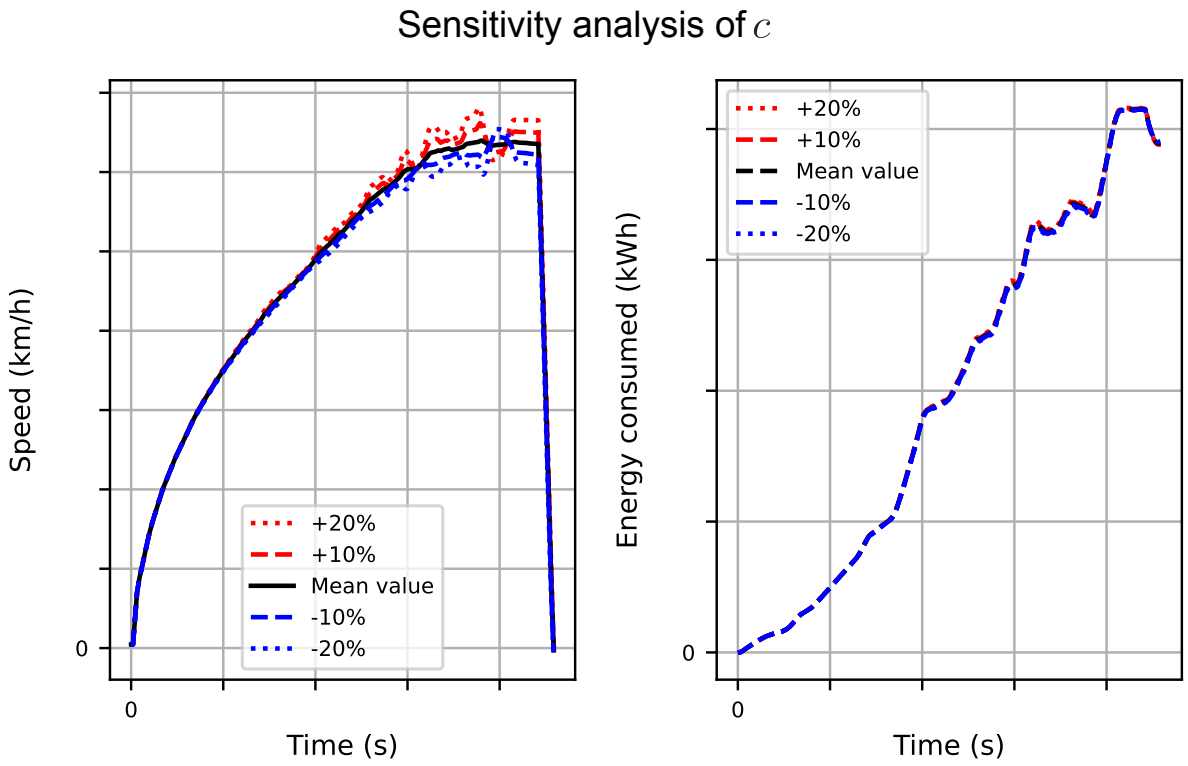

Figure II.14: Sensitivity of the models to Davis coefficient c.

perturb the value of the auxiliary power as $p^a = \bar{p^a} + \Delta p^a$ is equivalent to modify the energy consumption by $\Delta\mathcal{F}(t) = (t - t_s)\,\Delta p^a$. Therefore, $\Delta\mathcal{F}$ is linear in $(t - t_s)$. This can be observed in Figure II.15. Finally, the variation of energy consumption reaches 2.5% at the final time for perturbation of $\pm 20\%$ of the value of parameter p^a.

The other group of parameters contains the traction and braking efficiency parameters. First, coefficients a_η and b_η used in Equation (II.18) only have an influence on the consumed energy (and not on the dynamics) and only during traction phases, as they directly intervene in the estimation of the traction efficiency. This can be viewed in Figure II.17, where the main variation of the energy consumption occurs during the first part of the journey (when the train accelerates to reach its maximum speed).

Coefficient b_η is the constant part for the efficiency in Equation (II.18) (Figure II.17), whereas a_η (Figure II.16) models approximately its dependency on the train speed and the driver's command. The influence of coefficient a_η on the total energy consumption is smaller (less than 2% for $\pm 20\%$ variations) than the influence of b_η (33% for the same variations), because the first one corresponds to a small adjustment of the efficiency model compared to the second one, which has a higher sensitivity on the estimation of the value of the traction efficiency.

In the same way, we can compare the two other braking coefficients c_η and d_η (see Equation (II.19)). As it was aforementioned, those coefficients do not have any influence on the dynamic model, and they only influence the estimation of the energy recovered, while braking. We can observe this phenomenon in Figure II.19, where the consumed energy is slightly

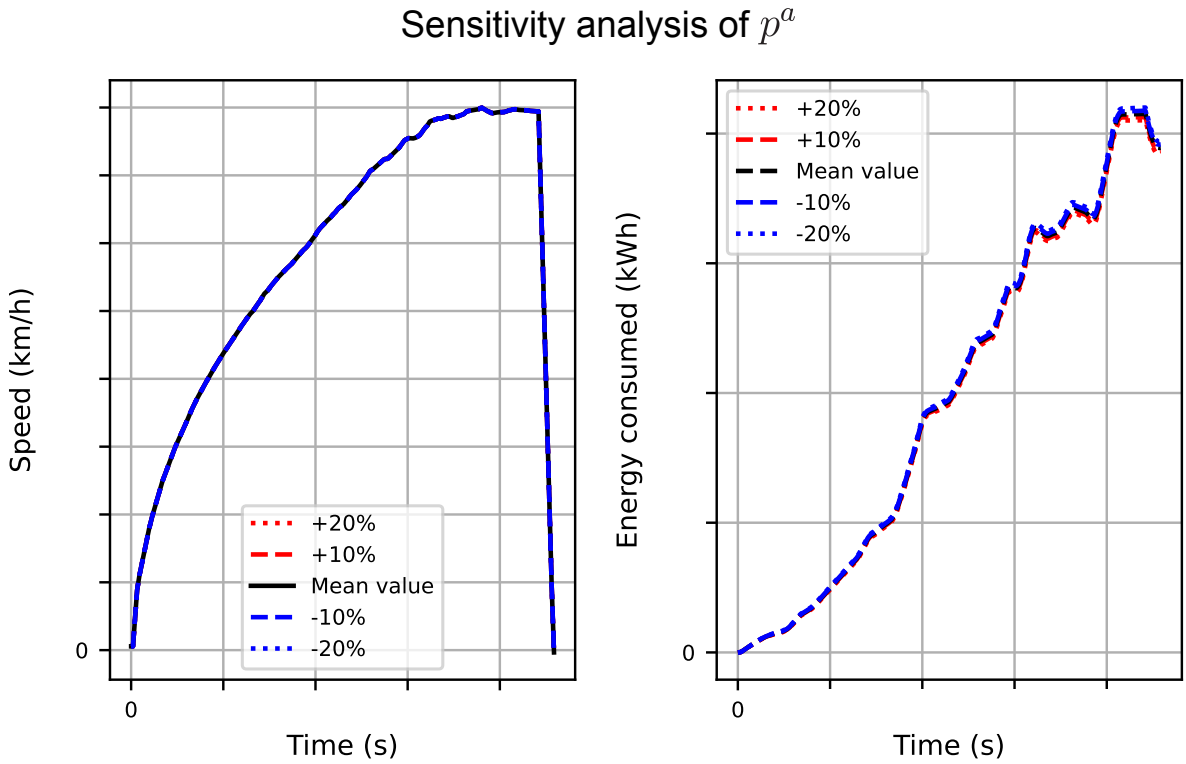

Figure II.15: Sensitivity of the models to auxiliary power p^a.

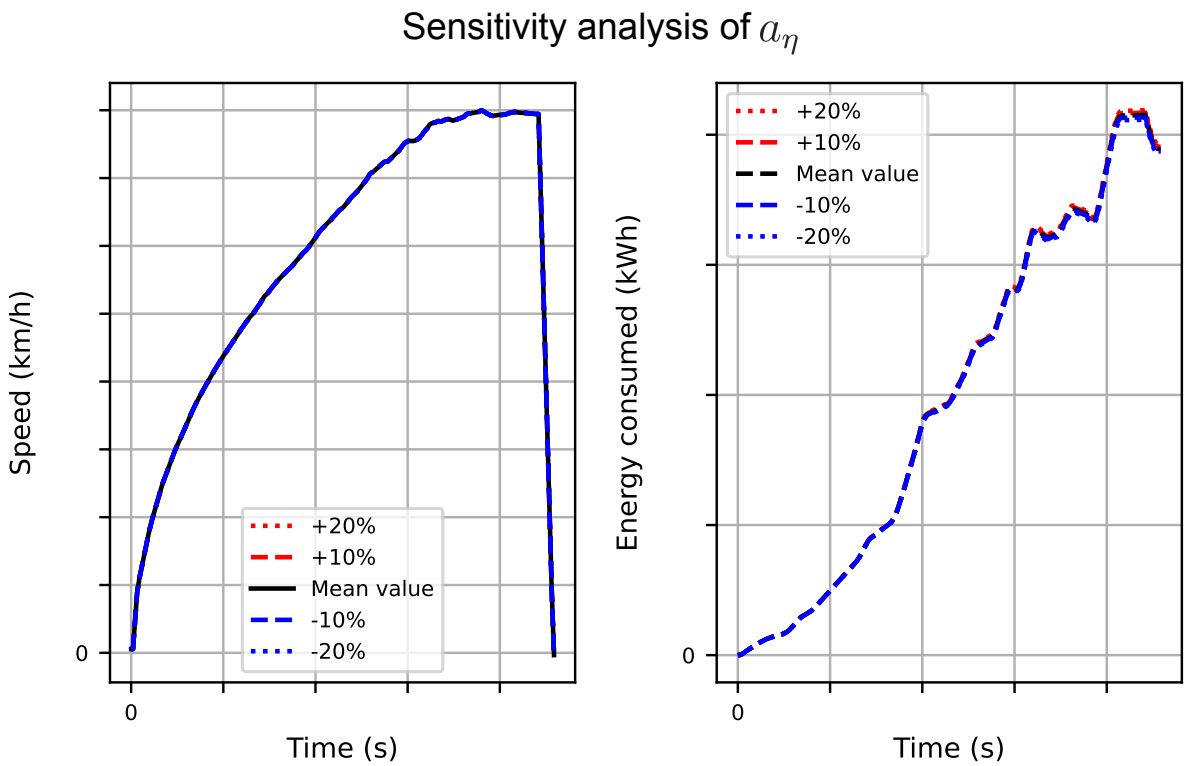

Figure II.16: Sensitivity of the models to traction efficiency coefficient a_η.

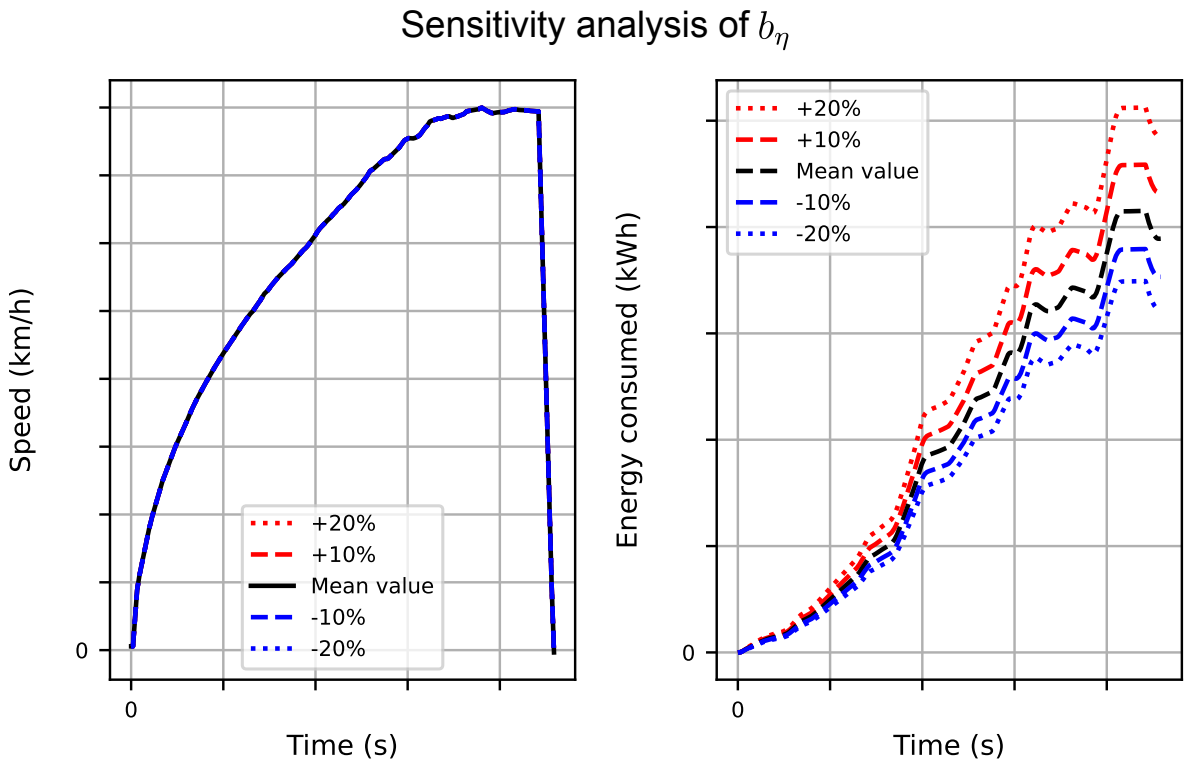

Figure II.17: Sensitivity of the models to traction efficiency coefficient b_η.

modified on the last part of the journey. As it was previously stated, coefficient c_η is a small adjustment of the model and has a small sensitivity on the braking efficiency, compared with coefficient d_η. The variation of energy consumption at the final time is around 0.3% for coefficient c_η and 4.5% for coefficient d_η. The influence of these parameters is relatively small compared with the traction efficiency parameters because this journey is mainly composed of traction phases. The opposite situation would have probably been observed in case of a journey composed of large braking phases and only few traction phases.

The parameters related to the energy consumption model have no influence on the train dynamics, as they do not intervene in the dynamic model. They only modify the energy consumed by the train. Many other parameters could have been studied, like corrective coefficient k^C or wind speed v^w. We have chosen to only focus on parameters m, p^a, a, b, c, a_η, b_η, c_η, and d_η, as they are the most important.

II.5 Conclusion and discussion on the modeling choices

In this chapter, we have constructed a model to represent the train, its environment, and the driver's command. Particular attention has been paid to provide a precise definition of the track and the wind, which have an important influence on the train dynamics. Then, a dynamic model for the train has been developed, derived from a multi-body model, but simplified in order to extract the longitudinal dynamics for the entire train. Additionally, an energy consumption model has been proposed as a function of the auxiliary power, the traction and

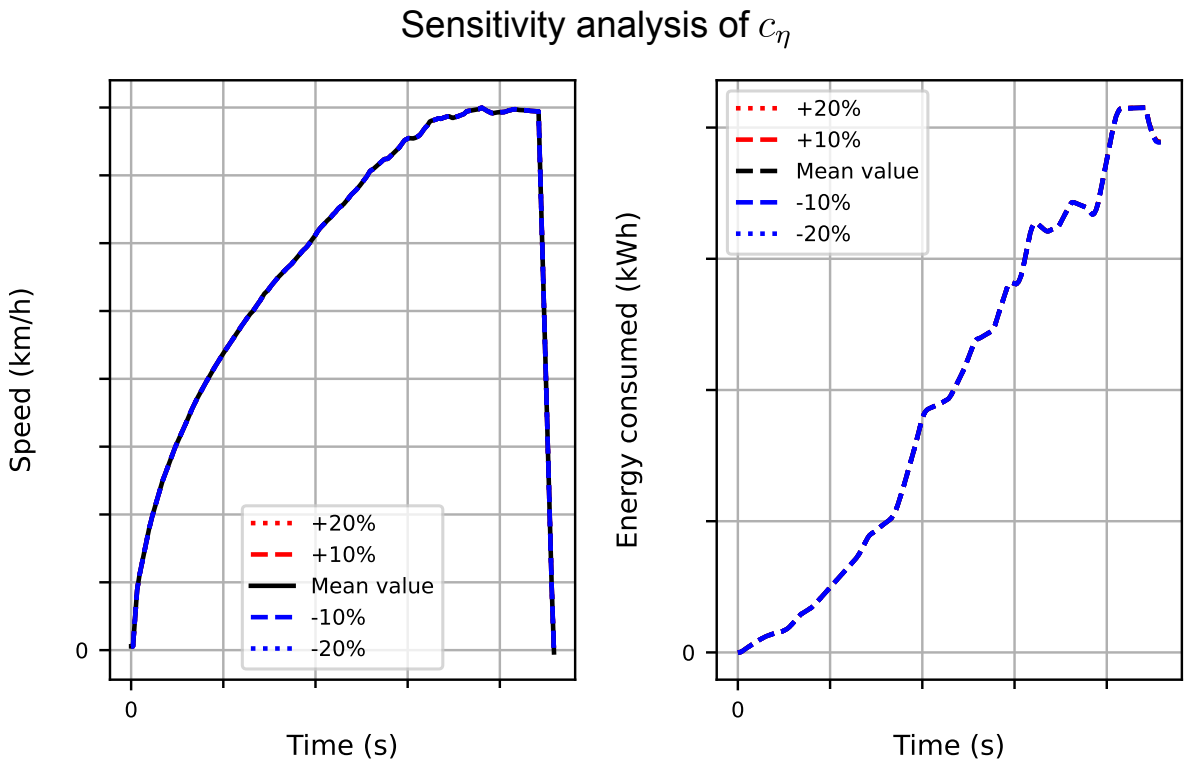

Figure II.18: Sensitivity of the models to braking efficiency coefficient c_η.

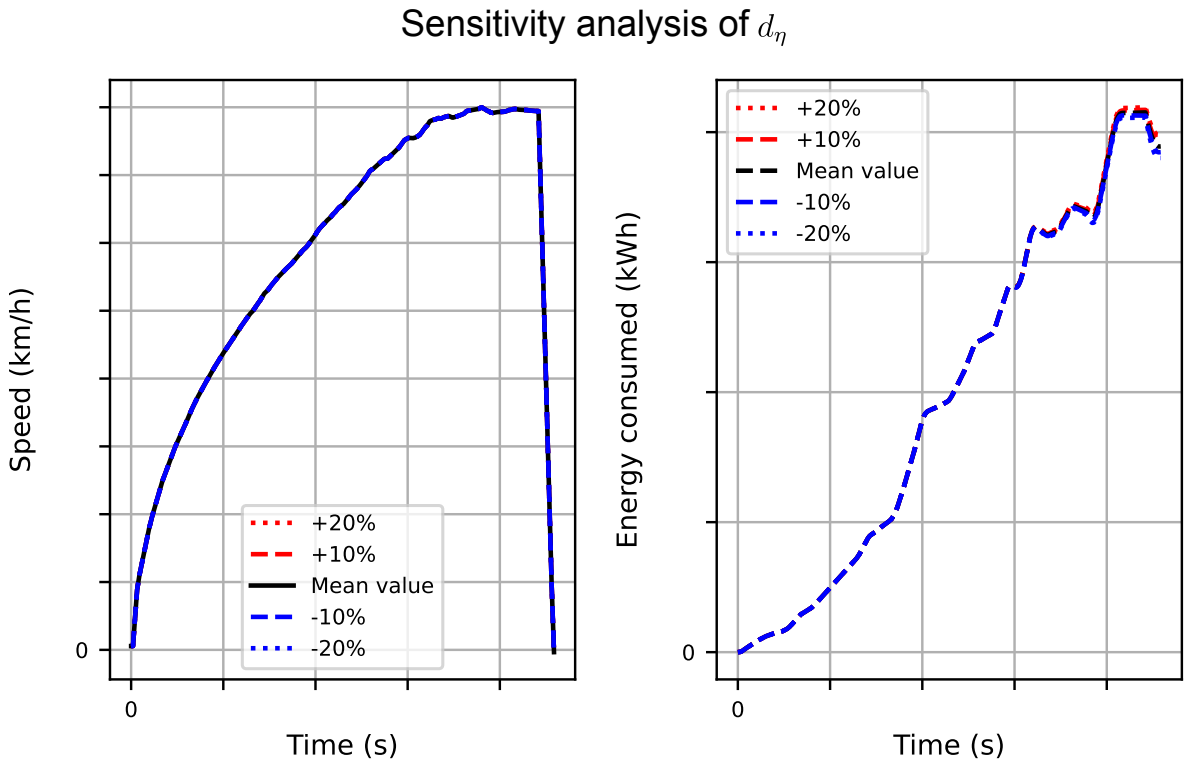

Figure II.19: Sensitivity of the models to braking efficiency coefficient d_η.

braking efficiencies. The chosen hypotheses have been carefully justified. Finally, a sensitivity analysis has been carried out to quantify the influence of the most important model parameters on the dynamic behavior and on the energy consumption of the train.

In Chapter III, the identification of these model parameters is performed. In fact, all trains do not behave in the same way and the identification of parameter values for the train under consideration is mandatory. The introduction of the uncertainty in the model parameters is necessary in order to deal with the variations that exist between trains, and also the lack of knowledge for every train.

Chapter III

Identification of the random model parameters

In order to calculate the train dynamics and its energy consumption, Chapter II have highlighted the importance of precisely defining the entries of the train system and the driver's command. The driver's command is well defined in Section II.1.3. What we have called entries are composed of the rolling environment $\mathcal{T}$ (the definition of the track and wind), but also of the vehicle description $\mathcal{V}$. Even though the environment has well been defined in Section II.1.1, some vehicle parameters are either not perfectly known (lack of knowledge, such as the energy efficiency) or have intrinsic fluctuations (such as the mass depending on the number of passengers in the train). These types of parameters are then considered as uncertain and are modeled by random variables. Thus, this chapter focuses on the identification of the random parameters from measurements carried out on commercial trains, which are grouped in the random vector $\boldsymbol{X}$, while the notation $\mathcal{V}$ gathers all the model parameters related to the vehicle. These model parameters correspond to the red box of the optimization problem (see Equation (III.1))

$$\{u^*\} = \underset{\substack{\{u\}\in\mathcal{U} \\ c(\{u\},\mathcal{T},\boxed{\mathcal{V}})=0}}{\arg\min} \mathcal{F}\left(\{u\},\mathcal{T},\boxed{\mathcal{V}}\right). \tag{III.1}$$

It should be noted that only a few works have been published concerning the identification problem of uncertain parameters in the context of the railway domain, see for instance [55, 15, 56]. In this work, the followed methodology for identifying the uncertain parameters is constituted of two steps. In a first step, the mean value $\underline{\boldsymbol{x}}$ of the random vector $\boldsymbol{X}$ is estimated using the least square method and the measurements of observations of the train system. In a second step, we use a Bayesian inference for estimating the posterior probability distributions of $\boldsymbol{X}$, for which the prior model of $\boldsymbol{X}$ is controlled by $\underline{\boldsymbol{x}}$ and by the variances of the components of $\boldsymbol{X}$ that are fixed to sufficiently large values in order that the support of the posterior probability distributions of $\boldsymbol{X}$ be consistent with the measurements. The likelihood function (conditional probability function) is constructed using the following hypotheses. The model uncertainties induced by the modeling errors are taken into account by the output prediction error method, for which a Gaussian additive noise is introduced on the observations of the train system. The multivariate Gaussian distribution of this noise is centered, and its covariance matrix is estimated by using the maximum likelihood method and the measurements. The sampling of the posterior will be performed using MCMC algorithm. Concerning the

available measurements, only the train speed profile and the energy consumption are measured but the associated experimental driver's command is unknown. Since we need the experimental driver's command for performing the statistical inverse problem, we have to identify it.

To carry out the identification, the available measurements are achieved a specific high-speed line (LGV Rhin-Rhone) for a French high-speed train (TGV Dasye). Several journeys have been used, during which the train longitudinal position s^{mes} and speed v^{mes}, as well as the RMS (Root Mean Square) values of the electrical current intensity i^{mes}, the voltage u^{mes}, and the phase shift, ϕ^{mes}, are measured through the pantograph at each δt^{mes}. In this chapter, the measured time sampling points correspond to an under sampling of all the measured time sampling points with a factor equal to 40. The number N^{mes} of measured time sampling points $\{t_j^{mes}, j = 1, \ldots, N^{mes}\}$ is smaller than the number N of time sampling points $\{t_i, i = 1, \ldots, N\}$ used for the simulations, but each t_j^{mes} coincides with a t_i. The electrical power transmitted (positive and negative) to the train (and consequently its energy consumption) can thus be estimated with a great confidence. The energy consumed by the train, f^{mes}, is written as

$$f^{mes}\left(t, i^{mes}, u^{mes}, \phi^{mes}\right) = \int_{t_s}^{t} i^{mes}\left(\tau\right) u^{mes}\left(\tau\right) \cos\left(\phi^{mes}\left(\tau\right)\right) d\tau \,. \tag{III.2}$$

Section III.1 defines the uncertain parameters that constitute the components of random vector $\boldsymbol{X}$. The identification of the experimental driver's command from the experimental train speed profile and the train model is presented in Section III.2. Section III.3 deals with the identification of $\underline{\boldsymbol{x}}$ using the least square method. In Section III.4, we present the Bayesian inference and the sampling of the posterior probability model using an MCMC algorithm. Finally, conclusions are drawn on the identification method in Section III.5.

III.1 Selection of the uncertain parameters

The accuracy of the prediction of the train model constructed in Chapter II strongly depends on the vehicle parameters. These parameters vary from one train to another one, but differences are also observed when comparing two trains of the same type, as long as they have undergone different wear levels induced by different loadings or different running conditions.

Several parameters of the models defined in Chapter II are not well-known or are defined with very little precision. We have selected nine of them as we observed that they have a strong influence on the dynamic and energy consumption models: the mass m, the Davis coefficients a, b, c, the auxiliary power p^a, the traction and braking efficiencies $a_\eta, b_\eta, c_\eta, d_\eta$.

Measurements made with the same train on several journeys are available for the identification. It is reasonable to assume that these parameters, apart the mass that may change according to the occupancy, vary only slightly from one circulation to another. In order to take into account this particularity related to the mass, the mass is written as

$$m\left(\mathcal{V}\right) = \bar{m}\left(\mathcal{V}\right) + \Delta m \,. \tag{III.3}$$

in which $\bar{m}\left(\mathcal{V}\right)$ is a deterministic nominal value, which is assumed to change from one journey to another. This nominal value is estimated using Algorithm 1, while fixing all the other

parameters to their nominal values. Parameter Δm, called the mass error, is uncertain and will simultaneously be identified with the other parameters in order to integrate the cross influences of the mass with the other uncertain parameters on the train dynamics. Note that we have chosen the representation defined by Equation (III.3) because, in the near future, $\bar{m}$ could directly be estimated by counting the number of passengers entering the train or by analyzing the resulting forces applied to the vertical suspensions of the train between two stations. In such a case, Δm will appear as the error associated with this estimation.

The algorithm is based on the following method. We look for $\bar{m}(\mathcal{V})$ such that the simulated energy consumption is close to the energy consumption measurement. We apply a dichotomy method starting from the possible minimum and maximum values of the mass, m^{min} and m^{max}, until reaching the convergence associated with a given tolerance ($100\,kg$). As the experimental driver's command is not available, and we need a driver's command for estimating $\bar{m}$. This estimation is performed (as explained in Section III.2) using the experimental speed profile and the train model for which all the parameters are fixed to their nominal values. Note that in this algorithm, the nominal value of the mass, $\bar{m}$, changes at each iteration of the dichotomy. This algorithm allowing $\bar{m}$ to be estimated is summarized in Algorithm 1.

Initialization:

- $m_1 \leftarrow m^{min}$, $m_2 \leftarrow m^{max}$
- Compute in inverse, $u_1\left(m_1\right)$ and $u_2\left(m_2\right)$
- Compute the consumed energy $\mathcal{F}_1 := \mathcal{F}\left(\{u_1\}, \mathcal{T}, \mathcal{V}(m_1)\right)$ and $\mathcal{F}_2 := \mathcal{F}\left(\{u_2\}, \mathcal{T}, \mathcal{V}(m_2)\right)$

while $m_2 - m_1 > 100$ **do**

- $m_3 = (m_1 + m_2)/2$
- Compute in inverse $u_3\left(m_3\right)$
- Compute the consumed energy $\mathcal{F}_3 := \mathcal{F}\left(\{u_3\}, \mathcal{T}, \mathcal{V}(m_3)\right)$

if $d_2\left(\mathcal{F}_1, f^{mes}\right) < d_2\left(\mathcal{F}_2, f^{mes}\right)$: **then**

- $m_2 \leftarrow m_3$
- $u_2 \leftarrow u_3$
- $\mathcal{F}_2 \leftarrow \mathcal{F}_3$

else if $d_2\left(\mathcal{F}_2, f^{mes}\right) < d_2\left(\mathcal{F}_1, f^{mes}\right)$: **then**

- $m_1 \leftarrow m_3$
- $u_1 \leftarrow u_3$
- $\mathcal{F}_1 \leftarrow \mathcal{F}_3$

end

Return $(m_1 + m_2)/2$

Algorithm 1: Estimation of $\bar{m}$ with a dichotomy method.

The distance d_2 is defined by

$$d_2\left(f, g\right) = \sqrt{\sum_{j=1}^{N^{mes}} \left(f\left(t_j^{mes}\right) - g\left(t_j^{mes}\right)\right)^2}. \tag{III.4}$$

Finally, the random variables (written with uppercase letters) can be divided into two groups. We define $\boldsymbol{X_D} = (\Delta M, A, B, C)$ the random variables modeling Δm, a, b, c involved in the dynamic model, $\boldsymbol{X_E} = (P^a, A_\eta, B_\eta, C_\eta, D_\eta)$ the random variables modeling p^a, a_η, b_η, c_η, d_η involved in the energy consumption model, and we define $\boldsymbol{X}$ such as $\boldsymbol{X} = (\boldsymbol{X_D}, \boldsymbol{X_E})$ that is the random vector to be identified. Note that all the random variables are normalized by their nominal values before being estimated but we keep the same notations.

Uncertain parameters associated with the dynamic model X_D

By construction, $m(\mathcal{V}) = \bar{m}(\mathcal{V}) + \Delta m$ is included in the interval $[m^{min}, m^{max}]$, where m^{min} is the mass of the empty train, and m^{max} is the mass for a train full of passengers. As we mentioned before, parameter Δm is considered uncertain.

The Davis coefficients are used to estimate the longitudinal friction forces (see Equation (II.10)). Nominal values for these coefficients are given by the train manufacturer for each type of vehicle. Nevertheless, the wear state of the wheels has an influence on the contact force, and thus on coefficients a and b. In the same manner, the aerodynamic force is also very simplified and coefficient c also needs to be properly estimated. Therefore, these three parameters are considered uncertain.

Definition of the uncertain parameters associated with the energy consumption model X_E

Auxiliary power p^a describes the part of the electrical power that is not transmitted to the traction chain. This electrical power can be decomposed in two contributions. A first part is used by the comfort tools of the train, such as the air conditioning or the light, between others. It should be possible to identify this first part when the train is stopped, but it appeared that this value was underestimated, as all the tools are generally not used at stop. A second part is distributed to the systems that ensure the safety of the journey (air-cooling motor systems). This second part can only be measured when the train is in motion. The sum of these two contributions may vary from one journey to another, depending on many variables such as the air temperature, the number of passengers, and so on. Figure III.1 shows the electric power measurement of the train as a red dotted line. It can be compared to the nominal value of the auxiliary power provided by the manufacturer that is plotted in black solid line.

As explained in Chapter II, the traction and braking efficiencies depend on four parameters: a_η and b_η for traction efficiency, c_η and d_η for braking efficiency. The nominal values of parameters a_η and b_η have been chosen to be similar to the 2D plot shown in Figure II.8. The simulated 2D plot (with nominal values) is shown in Figure II.9.

The nominal values of parameters c_η and d_η can also be estimated. We choose to keep the same 2D-plot than the one found for a_η and b_η because the same components of the train are involved in the two phenomena. Nevertheless, the values of the braking efficiency parameters are slightly smaller than the ones describing the traction efficiency. In practice, we know that the braking efficiency is around 0.82 against 0.87 for the traction efficiency. Consequently, we propose to reduce the nominal values of c_η and d_η of around 5% so that this proportion is respected.

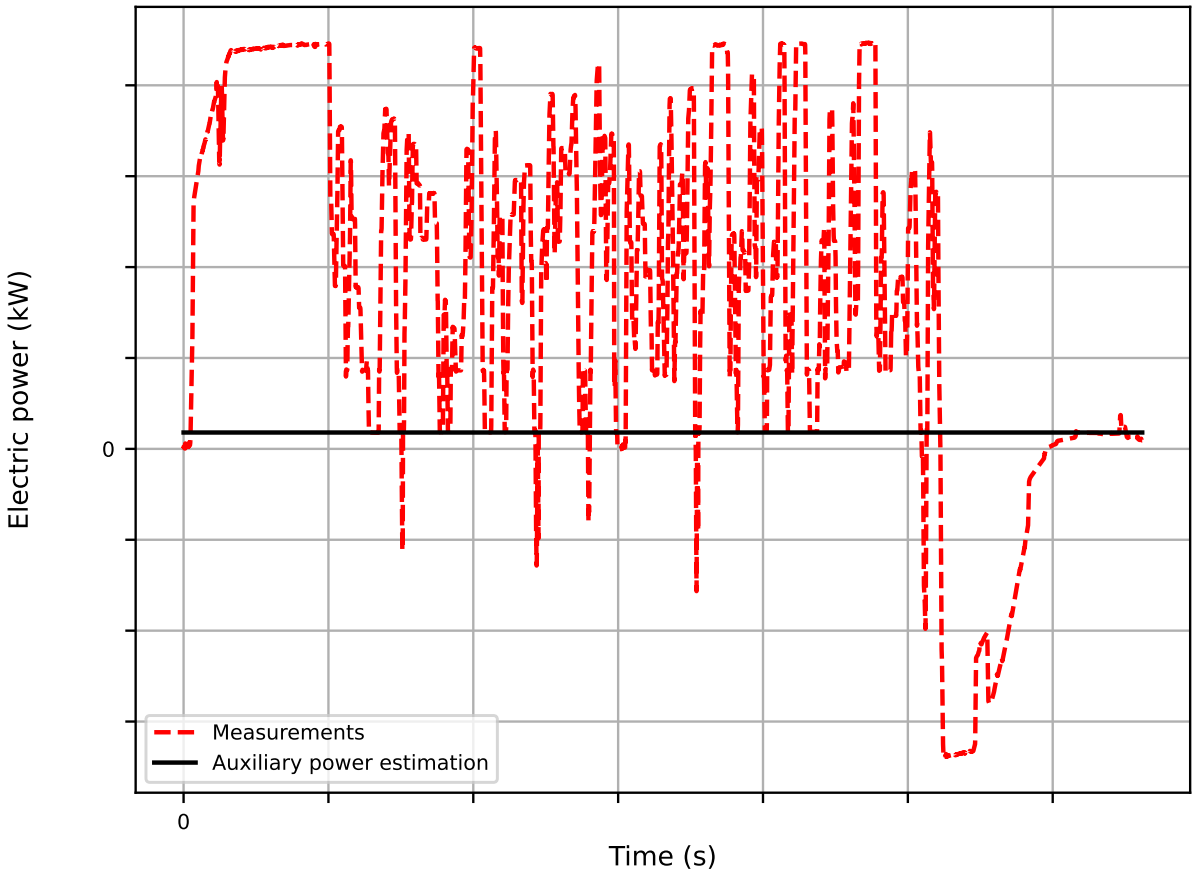

Figure III.1: Measured electrical power consumed by the train (red dotted line) and auxiliary power estimation (black solid line).

III.2 Identification of the experimental driver's command from the experimental train speed profile and the train model

As previously explained, for a given realization $\boldsymbol{x}$ of $\boldsymbol{X}$, the first part of the identification process aims to estimate in inverse the experimental discretized driver's command is denoted as $\boldsymbol{u}^{mes}$, for each journey, so that the simulated speed profile is as close as possible to the measured speed profile (t^{mes}, v^{mes}). As previously explained, the measurements are all discretized on a N^{mes}-dimensional time grid with intervals $\delta t^{mes} = 8\,s$, each of them being associated to a given time t_j^{mes}. The simulated discretized driver's command is denoted by $\boldsymbol{u} = (u_1, ..., u_N)$ with $u_i = u(t_i)$. The same notations are used for the discretized position, speed, and acceleration of the train.

With these notations, Equation (II.16) allows us to estimate the train acceleration at each time t_i using the position and the speed of the train as well as the simulated discretized driver's command,

$$\ddot{y}_i = \frac{\sum\limits_{\alpha} f^{\alpha}(y_i, \dot{y}_i, u_i, \boldsymbol{x}, \mathcal{T})}{m\, k^{rot}} \quad , \quad 1 \leq i \leq N\,. \tag{III.5}$$

Here, f^{α} gathers the traction and braking forces, the resistant force, the correction applied in curves, and the weight of the train. Given the acceleration at time t_i, we deduce the train position and speed at the next time step t_{i+1} thanks to the Euler scheme,

$$\dot{y}_{i+1} = \ddot{y}_i \, \Delta t + \dot{y}_i \quad , \quad y_{i+1} = \dot{y}_i \, \Delta t + y_i \, . \tag{III.6}$$

In practice, the train dynamics is numerically computed according to Equations (III.5) and (III.6). For the computation at point t_{i+1}, the initial speed and position are $\dot{y}_i$ and y_i. In Equation (III.5), the simulated discretized driver's command u_i is generated from u_j^{mes} using repetition. The distance d_1 between the simulated speed and the measured one is then calculated by,

$$d_1 \left(f, g \right) = \sum_{j=1}^{N^{mes}} \left| f \left(t_j^{mes} \right) - g \left(t_j^{mes} \right) \right| . \tag{III.7}$$

To find the experimental discretized driver's command $u_j^{mes} \in [-1, 1]$, we minimize this distance using the Brent's method. It uses the advantages of the bisection method, the secant method, and the inverse quadratic interpolation, switching from one method to another to accelerate the convergence. It has been proposed by R. Brent in [57] from the base of the algorithm developed by T. Dekker in [58]. The estimation process is summarized in Algorithm 2.

Initialization:

- Selection a value of $\underline{\boldsymbol{x}}$

for $j = 1 : N^{mes}$ **do**

- Initialize the position and the speed with y_j^{mes} and $\dot{y}_j^{mes}$
- Iterate Brent's method to find u_j^{mes}

while the convergence condition is not respected **do**

- Brent algorithm selects a candidate value u_j^{mes} using the values of distance d_1 calculated at the previous iteration
- Solve the dynamic using Equations (III.5) and (III.6)
- Calculate distance d_1 between the measured speed and its corresponding simulated speed

end

end

Return $\boldsymbol{u}^{mes}(\underline{\boldsymbol{x}})$

Algorithm 2: Identification of $\boldsymbol{u}^{mes}(\underline{\boldsymbol{x}})$ from experimental speed profile and model.

The quality of this identification is verified in Section III.3.2 (quality assessment).

III.3 Identification of the mean values of the random parameters using the least-square method

III.3.1 Prediction of the energy consumption

Henceforth, the driver's command is estimated, and we can try to evaluate the quality of the energy-consumption prediction for a given value $\underline{\boldsymbol{x}}$ of the parameters. Note that rather than comparing powers, we compare simulated and measured energy consumption because the model presented in Chapter II is constructed to precisely compute the energy. Therefore, we propose to work on the energy consumed by the train on the given times defined by t_j^{mes}. This last point is realized as a result of distance d_2 presented in Equation (III.4).

At that time, we are able to quantify the quality of $\underline{\boldsymbol{x}}$. But the question is now: how can we find the optimal value of $\underline{\boldsymbol{x}}$? Indeed, the dynamic and the energy consumption models both depend on $\underline{\boldsymbol{x}}$ and $\boldsymbol{u}^{mes}(\underline{\boldsymbol{x}})$. This inter-dependency, coupled with the fact that there is no perfect fit because of the model error, makes the problem not necessarily well posed, while being *a priori* nonconvex and nonregular in $\underline{\boldsymbol{x}}$. In fact, a small modification of $\underline{\boldsymbol{x}}$ undoubtedly conducts to a modification of the driver's command $\boldsymbol{u}^{mes}(\underline{\boldsymbol{x}})$, and these two modifications both have an impact on the consumed energy.

To circumvent this difficulty, a sample-based optimization method is used. It relies on the evaluation of a certain number N^{id} of values for $\{\underline{\boldsymbol{x}}_k, k = 1, \ldots, N^{id}\}$, which are drawn uniformly around a chosen value $\widehat{\boldsymbol{x}}$ of the parameters ($\widehat{\boldsymbol{x}}$ will be modified as a function of the algorithm iteration). The support of the uniform distribution for the samples associated with the component $\widehat{x}_l$ of $\widehat{\boldsymbol{x}}$ is chosen as $[0.9\,\widehat{x}_l, 1.1\,\widehat{x}_l]$, as a compromise between the volume of the research domain and the reduction of the calculation time (many iterations is required to explore a large volume). Number N^{id} is chosen depending on the acceptable computational time. For each realization $\underline{\boldsymbol{x}}_k$, $\boldsymbol{u}_k^{mes}(\underline{\boldsymbol{x}}_k)$ is identified with Algorithm 2 in order to calculate the simulated energy consumption before comparing it with the measurements.

Once all the iterations are completed, we extract the best candidate $\underline{\boldsymbol{x}}^*$ from the set of realizations $\{\underline{\boldsymbol{x}}_k, k = 1, \ldots, N^{id}\}$. We iterate this process N^{step} times, replacing $\widehat{\boldsymbol{x}}$ by the best candidate $\underline{\boldsymbol{x}}^*$ in the former procedure. If $\underline{\boldsymbol{x}}^*$ is far from $\widehat{\boldsymbol{x}}$, the same support is kept for the next iteration (because we have yet to explore the new domain). On the opposite, if the best candidate is relatively well centered in the support, this support is reduced for the next iteration to gain in precision. The complete procedure (which is iterated several times) is described in Algorithm 3.

Initialization:

- Choose N^{id} and N^{step}
- Choose $\widehat{\boldsymbol{x}}$

for $p = 1 : N^{\text{step}}$ **do**
  for $k = 1 : N^{id}$ **do**
    - Draw a candidate $\underline{\boldsymbol{x}}_k$ centered on $\widehat{\boldsymbol{x}}$
    - Determine in inverse $\boldsymbol{u}_k^{mes}(\underline{\boldsymbol{x}}_k)$ with Algorithm 2
    - Estimate the consumed energy $\mathcal{F}_k$
    - Calculate distance d_2 between the simulated energy consumption and the measured one
  end
  - Select the best candidate $\underline{\boldsymbol{x}}^*$
  - Update $\widehat{\boldsymbol{x}} \leftarrow \underline{\boldsymbol{x}}^*$
  - Potentially adapt the support
end
Return $\underline{\boldsymbol{x}}^*$.

Algorithm 3: Identification of the value $\underline{\boldsymbol{x}}^*$ of $\underline{\boldsymbol{x}}$.

At convergence, the best candidate $\underline{\boldsymbol{x}}^*$ is considered as a good approximation of the most likely value of the vehicle parameters. This expected value can be used to describe the train vehicle $\mathcal{V}$ in Equations (II.16) and (II.22). In order to make this estimate more robust, the comparisons between the measured and simulated energy consumptions in the global cost function rely on 10 different journeys. This modification only enriches the right-hand side in the definition of d_2.

Figures III.2 and III.3 show the comparison between the measurements and the simulations for the train speed profile and the energy consumption for 2 (serie 1 and serie 2) of the 10 different journeys used in the identification process. In these figures, the simulations have been performed using the estimated measured discretized driver's command.

As a first comment, we can see in the left figures in Figures III.2 and III.3 that $\boldsymbol{u}^{mes}$ has very well been identified. In addition, the simulated speed profile is well predicted comparing to the measurements both for the values of $\underline{\boldsymbol{x}}^*$ and for that of $\underline{\boldsymbol{x}}$. We can observe the weak effect of choosing a large value for δt^{mes} yielding small differences between simulations and measurements.

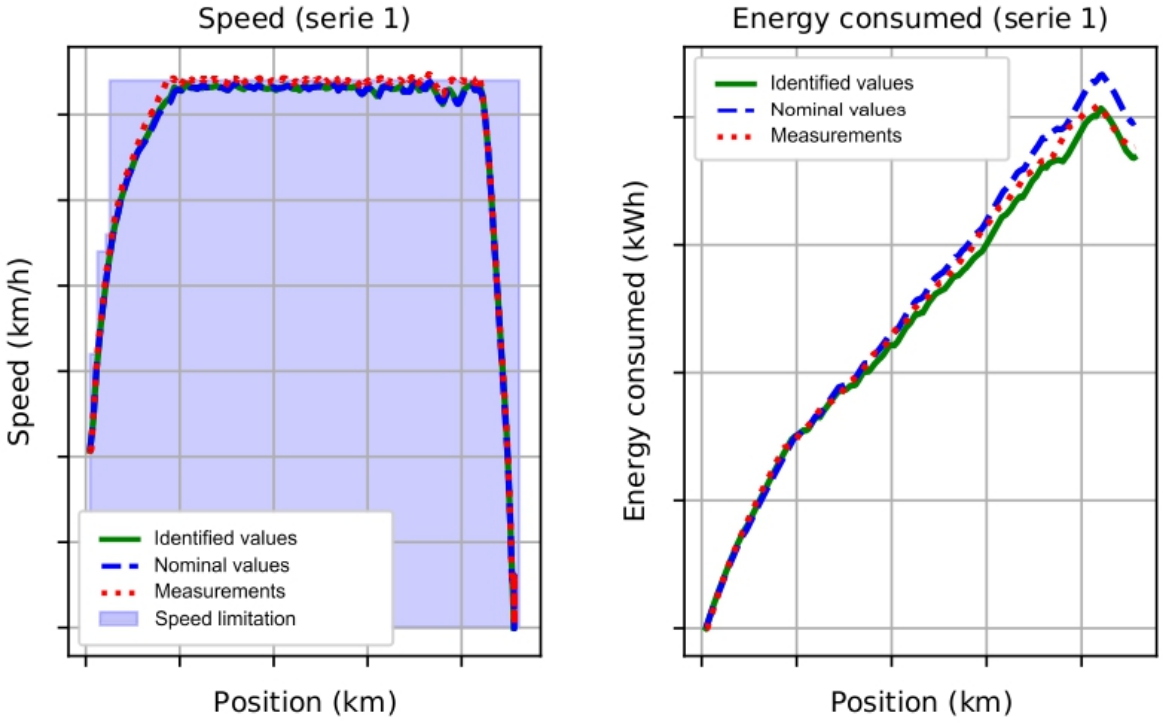

Figure III.2: Speed profile of the train (left) and consumed energy (right) depending on the position (i) for the identified values $\underline{\boldsymbol{x}}^*$ of the parameters (green), (ii) for their nominal values (blue), and (iii) for the corresponding measurements (red) - Serie 1.

The second comment that should be given is that the identification performed with Algorithm 3 is efficient. Indeed, the energy consumed by the train computed with $\underline{\boldsymbol{x}}^*$ is closer to the measurements than the one estimated with $\underline{\boldsymbol{x}}$. This means that, in spite of the difficulty of solving the problem, we are able to find a value $\underline{\boldsymbol{x}}^*$ of $\underline{\boldsymbol{x}}$, which fits the measurements. *A priori*, value $\underline{\boldsymbol{x}}^*$ could not be optimal. Nevertheless, we will see in Section III.3.2 that $\underline{\boldsymbol{x}}^*$ appears to be satisfactory.

Finally, we can calculate the Root Mean Square Error (RMSE) between the simulations and the measurements. With the nominal values, this error reaches 6.4%, and it is around 3.0% with $\underline{\boldsymbol{x}}^*$. Consequently, Algorithm 3 allows for improving the quality of the models by finding more adapted values of $\underline{\boldsymbol{x}}$.

III.3.2 Quality assessment of algorithms for the identification of the mean values of the random parameters

The quality of the presented method is assessed on 10 other journeys that have not been used by the identification algorithms. The value $\underline{\boldsymbol{x}}^*$ of the vehicle parameter is supposed to remain identical for these 10 journeys because the same vehicle has been used to realize all the measurements. The experimental discretized driver's command $\boldsymbol{u}^{mes}$ is again determined in inverse using Algorithm 2.

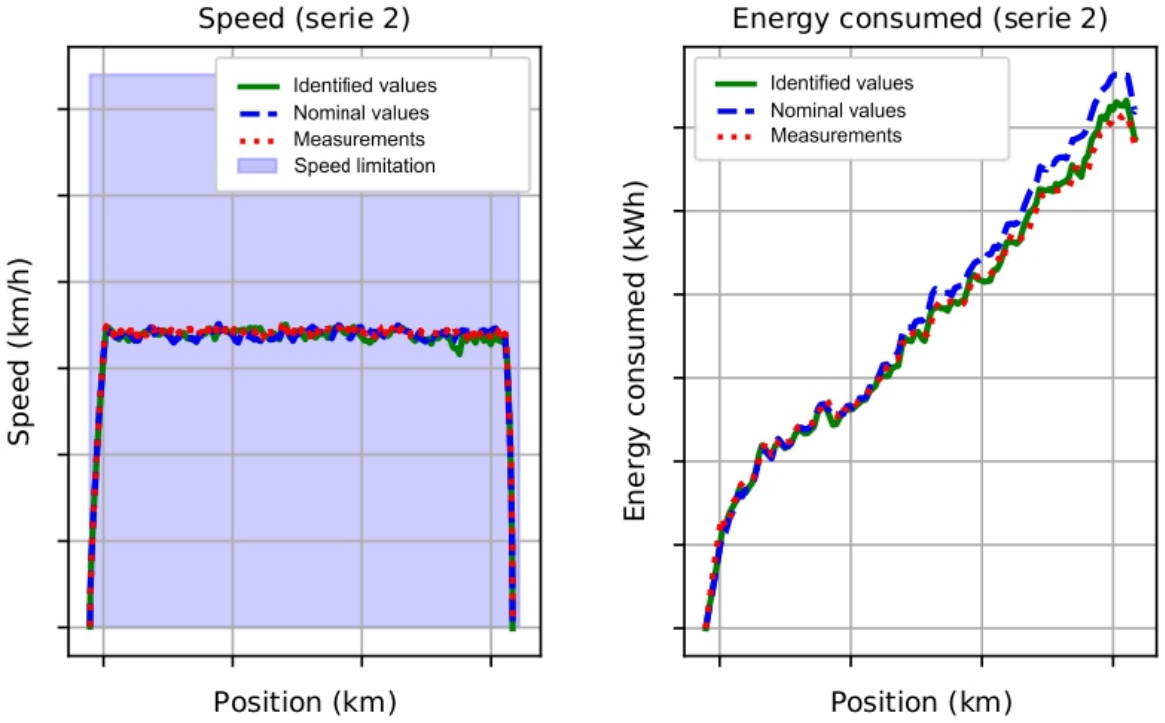

Figure III.3: Speed profile of the train (left) and consumed energy (right) depending on the position (i) for the identified values $\underline{\boldsymbol{x}}^*$ of the parameters (green), (ii) for their nominal values (blue), and (iii) for the corresponding measurements (red) - Serie 2.

Figures III.4 and III.5 show the speed profile and the energy consumed by the train for two journeys (series 12 and 18). Once again, we notice in the left figures that the proposed algorithms yield a very good comparison. Likewise, the measured consumed energy in red is better predicted using $\underline{\boldsymbol{x}}^*$ in green than using $\underline{\boldsymbol{x}}$ in blue (see right figures). Consequently, Algorithm 2 and 3 appear to be very interesting.

In a quantitative point of view, the RMSE over the 10 new journeys reaches 6.9% with $\underline{\boldsymbol{x}}$ and is reduced to 3.1% with $\underline{\boldsymbol{x}}^*$. The error values are slightly higher than the ones for the 10 journeys used by the algorithms, which is coherent.

III.3.3 Discussion about the identification method

This first step of the identification that we have proposed already yields a good prediction of the speed profile and of the energy consumption. However, this identification has been performed without taking into account the uncertainty inherent in the system: parameter uncertainties such as the mass, the auxiliary power, the Davis coefficients, and the energy efficiency and the model uncertainties induced by modeling errors. The next section is devoted to the second step of the identification. As previously explained, we present a probabilistic framework for performing the second step of the parameters identification, for which the first step results will be used to construct the prior probability distributions of the uncertain parameters.

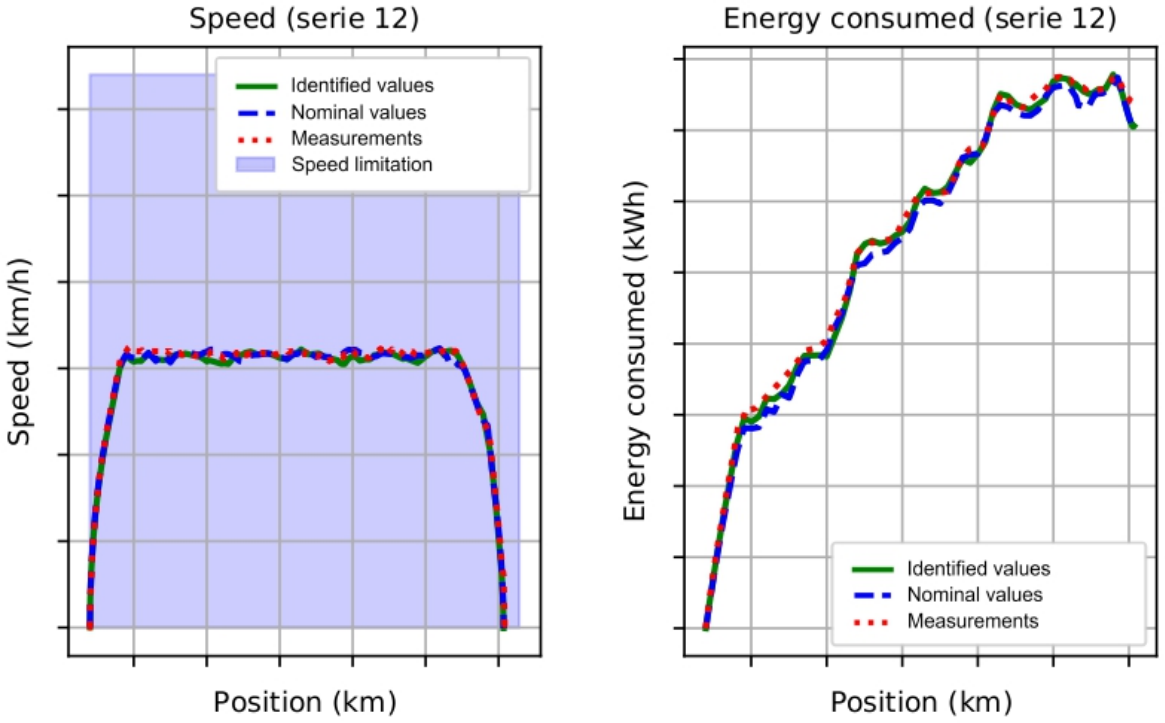

Figure III.4: Speed profile of the train (left) and consumed energy (right) depending on the position (i) for the identified values $\underline{\boldsymbol{x}}^*$ of the parameters (green), (ii) for their nominal values (blue), and (iii) for the corresponding measurements (red) - Serie 12.

III.4 Bayesian inference of the uncertain parameters

The Bayesian inference method that we have developed is based on general statistical tools (see for instance [59, 60, 61, 62, 63, 64] for general aspects relative to Bayesian inference).

The method used has been summarized at the beginning of Chapter III. It is composed of three steps. The first one is the construction of the prior distributions of the model parameters (Section III.4.1), and the choice of a structure for the modeling error (Section III.4.2). The second step deals with the construction of the likelihood function and is detailed in Section III.4.3. The estimation of the posterior distribution constitutes the third step. Section III.4.4 deals with the Markov Chain Monte Carlo (MCMC) method in order to infer this posterior distribution. Finally, in Section III.4.5, we present the uncertainty propagation through the train dynamic system and the energy consumption model. This propagation will be quantified in computing the confidence regions.

III.4.1 Construction of the prior probability distributions

The prior distributions are mainly constructed in the framework of Information Theory and are presented (i) for the Davis coefficients A, B, C, (ii) for the mass error ΔM, the auxiliary power P^a, the efficiency coefficients B_η, D_η, and (iii) for the efficiency coefficients A_η, C_η. Depending on the characteristics of this available information, several constructions are proposed.

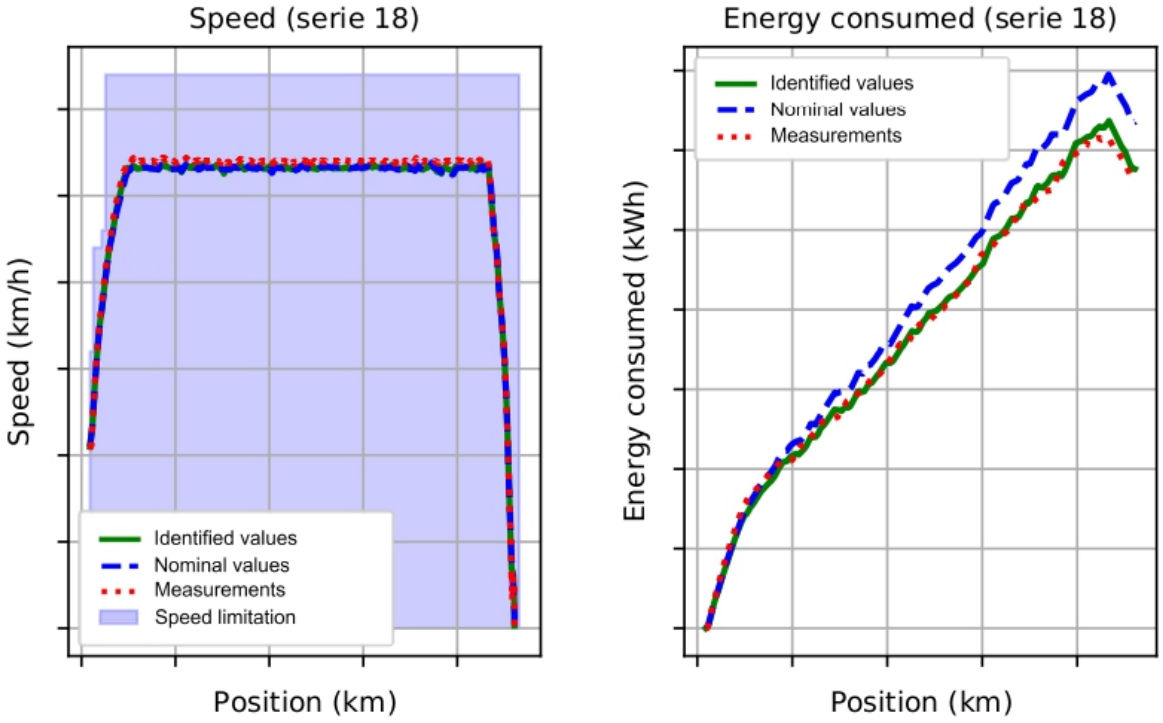

Figure III.5: Speed profile of the train (left) and consumed energy (right) depending on the position (i) for the identified values $\underline{\boldsymbol{x}}^*$ of the parameters (green), (ii) for their nominal values (blue), and (iii) for the corresponding measurements (red) - Serie 18.

Davis coefficients A, B, and C

The prior probability distributions of the random variables A, B, and C are constructed as follows. Let Z be the random variable representing anyone of these three random variables. The available information for Z is positive-valued random variable, mean value and variance given. Consequently, the use of the Maximum Entropy Principle [65, 66, 67, 68] under the constraints defined by this available information, is a Gamma probability distribution. The hyperparameters of such a distribution are the shape parameter k_Z and the scale parameter θ_Z. The mean value is $\underline{z} = k_Z\theta_Z$ and the variance is $\sigma_Z^2 = k_Z\theta_Z^2$. We choose to define the mean value as the value extracted from $\underline{\boldsymbol{x}}^*$ and to fix the value of the variance. The probability density function (PDF) of Z is then written as

$$f_Z(z) = \frac{z^{k_Z-1}e^{-z/\theta_Z}}{\Gamma\left(k_Z\right)\theta_Z^{k_Z}}\mathbb{1}_{\mathbb{R}^+}(z)\,, \tag{III.8}$$

with $k \mapsto \Gamma(k)$ the Gamma function and $x \mapsto \mathbb{1}_{\mathbb{R}^+}(x)$ the indicator function over $\mathbb{R}^+$. The shape and scale hyperparameters are thus written

$$\left\{\begin{array}{rcl} k_Z &=& \underline{z}^2/\sigma_Z^2\,, \\ \theta_Z &=& \sigma_Z^2/\underline{z}\,. \end{array}\right. \tag{III.9}$$

Mass error ΔM, auxiliary power P^a, and efficiency coefficients B_η and D_η

We have previously mentioned in Section III.1 that the total mass of the train is decomposed in an approximated deterministic mass $\bar{m}$ estimated by the devices of the trains, and a mass error, denoted by Δm, and modeled by a random variable ΔM. The available information we have to define the probability distribution of ΔM is that ΔM is with values in $[-1\,000, 1\,000]\,kg$. In addition, we have seen that the random total mass M of the train, is with values in the interval $[m^{min}, m^{max}]$. Consequently, the support of the probability distribution of ΔM is

$$\mathcal{S}_{\Delta M} = [\max(-1\,000, m^{min} - \bar{m})\,,\, \min(1\,000, m^{max} - \bar{m})]\,.$$

With this available information, the use of the Maximum Entropy Principle yields a uniform distribution for ΔM for which the support is $\mathcal{S}_{\Delta M}$.

Regarding the random auxiliary power P^a, the available information is the support of its probability distribution (which is a positive interval) and which is centered on its mean value extracted from $\underline{\boldsymbol{x}}^*$. Consequently, this probability distribution is uniform. We fix a large support in order to well estimate its posterior distribution. Consequently, the support is written

$$\mathcal{S}_{P^a} = [p^{a,min}\,,\, p^{a,max}] \subset \mathbb{R}^+\,,$$

with $p^{a,min} = 350\,kW$ and $p^{a,max} = 650\,kW$.

Concerning the energy efficiency, Equations (II.18) and (II.19) can be rewritten as

$$\eta^T(p^T) = a_\eta p^T + b_\eta \quad , \quad \eta^B(p^B) = c_\eta p^B + d_\eta\,,$$

in which

$$p^T \in [0\,, p^{T,max}] \quad , \quad \eta^T\left(p^T\right) \in [0.1\,, 1]\,, \tag{III.10}$$

$$p^B \in [0\,, p^{B,max}] \quad , \quad \eta^B\left(p^B\right) \in [0.1\,, 1]\,, \tag{III.11}$$

where $p^{T,max}$ and $p^{B,max}$ are the available maximum traction and braking powers reached for a traction and braking forces equal to $f^{T,max}$ or $f^{B,max}$. Traction efficiency $\eta^T(p^T)$ is modeled by a random variable $H^T(p^T)$ and $H^B(p^B)$ by a random variable $H^B(p^B)$ that are written as

$$H^T(p^T) = A_\eta p^T + B_\eta \quad , \quad H^B(p^B) = C_\eta p^B + D_\eta\,. \tag{III.12}$$

It can be deduced that the support of the probability distributions of random variables B_η and D_η are

$$\mathcal{S}_{B_\eta} = [0.1\,, 1] \quad , \quad \mathcal{S}_{D_\eta} = [0.1\,, 1]\,.$$

We only have information on the support of the probability distributions of B_η and D_η. Using the Maximum Entropy Principle yields a uniform distribution with support $\mathcal{S}_{B_\eta}$ and $\mathcal{S}_{D_\eta}$, respectively. Note that the mean values are not imposed.

Efficiency coefficients A_η and C_η

For Equations (III.10) to (III.12), it can be deduced that

$$H^T(p^{T,max}) = A_\eta p^{T,max} + B_\eta \quad , \quad H^B(p^{B,max}) = C_\eta p^{B,max} + D_\eta \,. \tag{III.13}$$

The supports of the probability distributions of $H^T(p^{T,max})$ and $H^B(p^{B,max})$ are, respectively,

$$\mathcal{S}_{H^T(p^{T,max})} = [0.1\,,1] \quad , \quad \mathcal{S}_{H^B(p^{B,max})} = [0.1\,,1]\,. \tag{III.14}$$

The probability distributions of the two last random parameters, A_η and C_η, are constructed as follows. First from Equations (III.13) and (III.14), it can be deduced that, almost surely, we have

$$\begin{aligned} (0.1 - B_\eta)/p^{T,max} \leq & 0 \leq A_\eta \leq (1 - B_\eta)/p^{T,max}\,, \\ (0.1 - D_\eta)/p^{B,max} \leq & 0 \leq C_\eta \leq (1 - D_\eta)/p^{B,max}\,. \end{aligned} \tag{III.15}$$

The constraints defined by Equation (III.15) have to be verified for any realization of B_η and D_η. Let b_η and d_η be realizations of random variables B_η and D_η. We introduce the conditional random variables $A_\eta|\{B_\eta = b_\eta\}$ and $C_\eta|\{D_\eta = d_\eta\}$. The supports of these two random variables are $[0, 1 - b_\eta]$ and $[0, 1 - d_\eta]$. Using the Maximum Entropy Principle under the constraints defined by this available information yields conditional uniform distributions with these supports.

Prior probability density function of random vector $\boldsymbol{X}$

The prior probability density function $f_{\boldsymbol{X}}^{prior}(\boldsymbol{x})$ of the $\mathbb{R}^9$-valued random variable $\boldsymbol{X}$ defined in Section III.1 is directly deduced from the independent properties and the constructed margin probability density functions related to the components.

Generators and graphs of the prior probability density functions of the random parameters

Given the former prior probability distributions and using a random generator of realizations for each one of the prior probability distributions, we have estimated the prior probability density functions using the Gaussian Kernel Density Estimation (see [69] and [70]). The graphs of the prior PDFs of each one of the nine random parameters are shown in Figure III.6.

In these figures, it can be seen that the prior PDFs are consistent with respect to the knowledge we had for these parameters. In particular, the support of the PDFs are well represented and the mean values for the Davis random variables have correctly been adjusted.

III.4.2 Choice of the structure for the modeling error

As explained at the beginning of Chapter III, for the computation of the random energy consumption, the model uncertainties induced by the modeling errors are taken into account by the

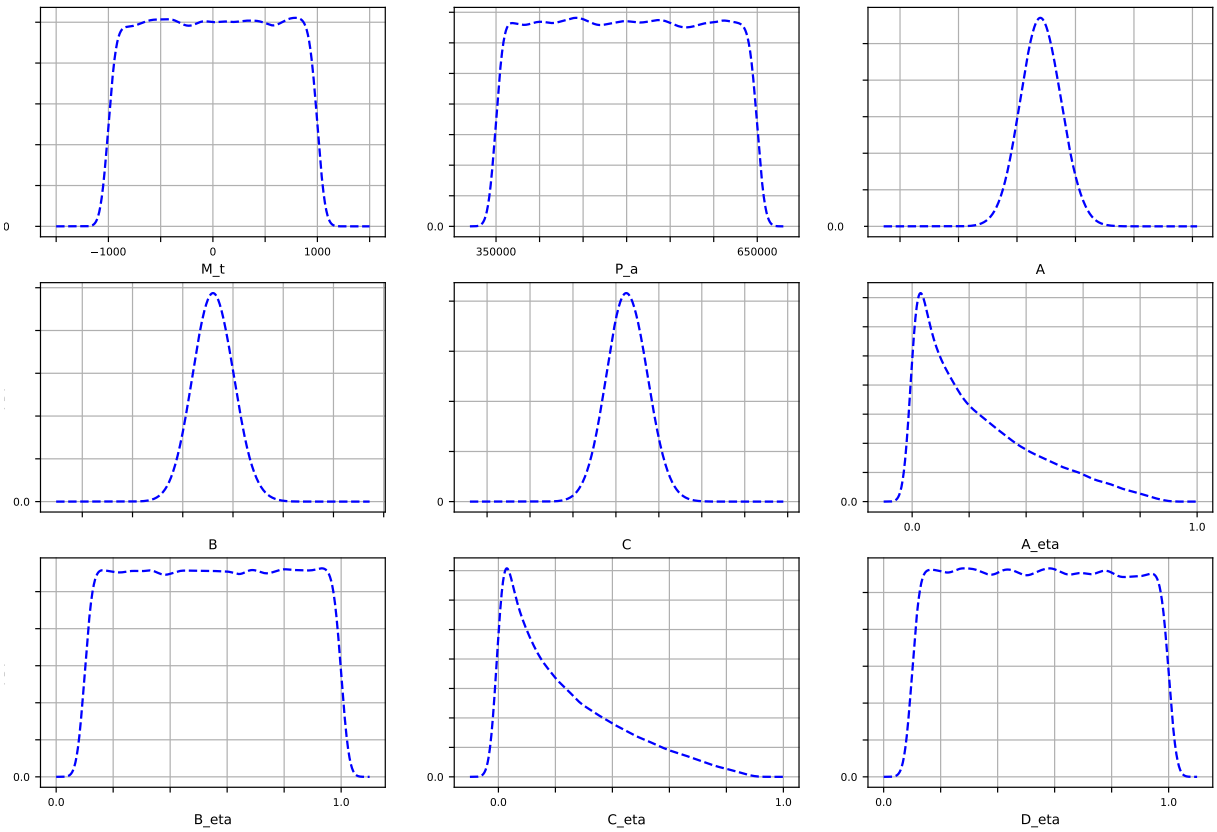

Figure III.6: Prior probability density function of each one of the uncertain parameters. From top left to bottom right (line-by-line): Mass error ΔM, auxiliary power P^a, Davis coefficients A, B, C, and traction and braking efficiency parameters A_η, B_η, C_η, and D_η.

introduction of an additive noise. Using measurements time sampling t_j^{mes}, $j = 1, \ldots, N^{mes}$, we define the additive noise $\boldsymbol{\varepsilon}^{mod}$ by the equality of the $\mathbb{R}^{N^{mes}}$-valued random vectors,

$$\boldsymbol{\mathcal{F}}^{mod}(\boldsymbol{u}, \mathcal{T}, \boldsymbol{X}, \boldsymbol{\sigma}) = \boldsymbol{\mathcal{F}}(\boldsymbol{u}, \mathcal{T}, \boldsymbol{X}) + \boldsymbol{\varepsilon}^{mod}(\boldsymbol{X}, \boldsymbol{\sigma}) \,, \tag{III.16}$$

in which $\boldsymbol{\sigma}$ is a hyperparameter vector that will be presented below. The components of each random vectors correspond to their values at time t_j^{mes}. We define the conditional random vector $\boldsymbol{\varepsilon}^{mod}(\boldsymbol{X}, \boldsymbol{\sigma})|\{\boldsymbol{X} = \boldsymbol{x}\}$, simply denoted as $\boldsymbol{\varepsilon}^{mod}(\boldsymbol{x}, \boldsymbol{\sigma})$. Under adapted hypotheses, we prove below that random vector $\boldsymbol{\varepsilon}^{mod}(\boldsymbol{x}, \boldsymbol{\sigma})$ is a $\mathbb{R}^{N^{mes}}$-valued centered Gaussian random vector with covariance matrix $[C^{mod}(\boldsymbol{x})]$ whose structure will be defined and will be identified using the maximum likelihood on measurements.

The Bayesian inference is going to be applied using measurements $\boldsymbol{f}^{mes}$ associated with $\boldsymbol{\mathcal{F}}^{mod}$. Note that there are two sources of modeling errors: the first one, $\boldsymbol{\varepsilon}^F$, is related to the train-dynamics model and the other one, $\boldsymbol{\varepsilon}^P$, to the energy consumption model.

Error related to the train-dynamics model

The train-dynamics model is responsible for the first source of modeling error. This error results from various simplifications (such as the use of the Davis model for the friction forces), from the use of the longitudinal model instead of the 3D one (see Appendix A), but also from the use of the experimental driver's command presented in Section III.2.

The dynamic equation (Equation (II.16)) is used to compute the traction and the braking forces in order to identify in inverse the experimental driver's command associated with a given speed profile. In this equation, all the quantities may be subjected to a modeling error that is represented by the $\mathbb{R}^{N^{mes}}$-valued centered Gaussian random variable $\boldsymbol{\varepsilon}^F$. For each time t_j^{mes}, we have

$$\begin{cases} F_j^{T,mod}\big(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \sigma^F\big) = f_j^T(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}) + \varepsilon_j^F(\sigma^F) & \text{if } u_j > 0\,, \\ F_j^{B,mod}\big(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \sigma^F\big) = f_j^B(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}) + \varepsilon_j^F(\sigma^F) & \text{otherwise.} \end{cases} \tag{III.17}$$

It is assumed that the components of random vector $\boldsymbol{\varepsilon}^F$ are not correlated, and consequently, are statistically independent. Therefore, the covariance matrix $\big[C^F\big]$ of $\boldsymbol{\varepsilon}^F$ is diagonal. In addition, the train-dynamics modeling errors do not depend on time. Thus, the variance $(\sigma^F)^2$ is chosen independent of time. We then have $\big[C^F\big] = \big(\sigma^F\big)^2\, \mathbb{1}_{N^{mes}}$, where σ^F is a hyperparameter that has to be estimated from the measurements.

Error related to the energy consumption model

Without modeling error, at time sampling t_j^{mes}, Equation (II.17) yields the vectorial random equation,

$$p_j^E\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) = p_j^T\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) - p_j^B\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) + p^a(\boldsymbol{X})\,, \tag{III.18}$$

where p_j^T and p_j^B are given (see Equations (II.20) and (II.21)) by

$$p_j^T\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) \eta_j^T\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) = f_j^T\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) \dot{y}_j\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right)\,, \tag{III.19}$$

$$p_j^B\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) = \eta_j^B\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) f_j^B\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) \dot{y}_j\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right)\,. \tag{III.20}$$

For the modeling error induced by the simplification of the efficiency model, the second source error for the energy consumption model is related to the electrical power and is represented by the $\mathbb{R}^{N^{mes}}$-valued centered Gaussian random variable $\boldsymbol{\varepsilon}^P$ that is assumed to be independent of $\boldsymbol{\varepsilon}^F$. Let $\boldsymbol{\sigma} = (\sigma^F, \sigma^P)$ be the vector in $\mathbb{R}^2$ whose components are σ^F and σ^P. Similarly to Equation (III.18), for each time t_j^{mes}, we have

$$P_j^{E,mod}(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \boldsymbol{\sigma}) = P_j^{T,mod}\big(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \sigma^F\big) - P_j^{B,mod}\big(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \sigma^F\big) + p^a(\boldsymbol{X}) + \varepsilon_j^P(\sigma^P)\,, \tag{III.21}$$

where $P_j^{T,mod}$ and $P_j^{B,mod}$ (inspired of Equations (II.20) and (II.21)) are such that

$$P_j^{T,mod}\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \sigma^F\right) \eta_j^T\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) = F_j^{T,mod}\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \sigma^F\right) \dot{y}_j\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right)\,, \tag{III.22}$$

$$P_j^{B,mod}\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \sigma^F\right) = \eta_j^B\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) F_j^{B,mod}\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \sigma^F\right) \dot{y}_j\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right)\,. \tag{III.23}$$

It should be noted that, in Equations (III.22) and (III.23), we have used $\eta_j^T\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right)$, $\eta_j^B\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right)$, and $\dot{y}_j\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right)$ instead of their corresponding random values that would take into account modeling errors, in order to avoid an implicit complex formulation. This hypothesis is justified by the fact that these three quantities are low sensitive to $\boldsymbol{\varepsilon}^F$.

As for $\boldsymbol{\varepsilon}^F$, the components of random vectors $\boldsymbol{\varepsilon}^P$ are not correlated, its covariance matrix $\left[C^P\right]$ is thus diagonal, and the variance $(\sigma^P)^2$ is chosen independent of time. We then have $\left[C^P\right] = \left(\sigma^P\right)^2 \mathbb{1}_{N^{mes}}$, where σ^P is a hyperparameter that has to be estimated from the measurements.

III.4.3 Definition of the likelihood function for the Bayesian inference

The likelihood function is the conditional probability density function of the observed random variable (vector-valued random energy consumption) given $\boldsymbol{X} = \boldsymbol{x}$. As mentioned, we will prove below that the random variable $\boldsymbol{\varepsilon}^{mod}(\boldsymbol{x}, \boldsymbol{\sigma})$ is Gaussian and consequently, is defined by its mean vector and its covariance matrix depending on $\boldsymbol{x}$. From Equations (III.21) to (III.23), it can be deduced that

$$P_j^{T,mod}\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \sigma^F\right) = \frac{f_j^T\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) + \varepsilon_j^F(\sigma^F)}{\eta_j^T\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right)}\, \dot{y}_j\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right), \tag{III.24}$$

$$P_j^{B,mod}\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \sigma^F\right) = \eta_j^B\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right)\left(f_j^B\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) + \varepsilon_j^F(\sigma^F)\right)\, \dot{y}_j\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right), \tag{III.25}$$

$$P_j^{E,mod}\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \boldsymbol{\sigma}\right) = p_j^E\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) + \tilde{\varepsilon}_j^F\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \sigma^F\right) + \varepsilon_j^P(\sigma^P), \tag{III.26}$$

in which

$$\tilde{\varepsilon}_j^F\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \sigma^F\right) = \left(1/\eta_j^T\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) - \eta_j^B\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right)\right) \dot{y}_j\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}\right) \varepsilon_j^F(\sigma^F). \tag{III.27}$$

Note that the conditional random vector $\tilde{\boldsymbol{\varepsilon}}^F | \{\boldsymbol{X} = \boldsymbol{x}\}$ remains a $\mathbb{R}^{N^{mes}}$-valued centered Gaussian random vector. Using time sampling $\{t_j^{mes}, j = 1, \ldots, N^{mes}\}$ for discretizing the integral in time of the electrical power, the random energy consumption $\mathcal{F}_j^{mod}$ that takes into account modeling error, is written as

$$\mathcal{F}_j^{mod}\left(\boldsymbol{u}_{\leq j}, \mathcal{T}, \boldsymbol{X}, \boldsymbol{\sigma}\right) = \sum_{k=1}^{j}\left(p_k^E\left(\boldsymbol{u}_{\leq k}, \mathcal{T}, \boldsymbol{X}\right) + \tilde{\varepsilon}_k^F\left(\boldsymbol{u}_{\leq k}, \mathcal{T}, \boldsymbol{X}, \sigma^F\right) + \varepsilon_k^P(\sigma^P)\right) \delta t^{mes}. \tag{III.28}$$

Comparing Equations (III.16) and (III.28) yields

$$\varepsilon_j^{mod}(\boldsymbol{X}, \boldsymbol{\sigma}) = \sum_{k=1}^{j}\left(\tilde{\varepsilon}_k^F\left(\boldsymbol{u}_{\leq k}, \mathcal{T}, \boldsymbol{X}, \sigma^F\right) + \varepsilon_k^P(\sigma^P)\right) \delta t^{mes}. \tag{III.29}$$

Since $\tilde{\varepsilon}_k^F(\boldsymbol{X})$ is a linear function of ε_k^F, it can be seen that the conditional random variable $\varepsilon_j^{mod}(\boldsymbol{x}, \boldsymbol{\sigma})$ for $\boldsymbol{X} = \boldsymbol{x}$ is effectively Gaussian and centered.

Since $\tilde{\boldsymbol{\varepsilon}}^F(\boldsymbol{x},\boldsymbol{\sigma})$ and $\boldsymbol{\varepsilon}^P$ are independent and centered, for all j and j' in $\{1,\ldots,N^{mes}\}$, the entry $[C^{mod}(\boldsymbol{x};\boldsymbol{u},\mathcal{T}_n,\boldsymbol{\sigma})]_{jj'}$ of the covariance matrix of the random vector $\boldsymbol{\varepsilon}^{mod}(\boldsymbol{x},\boldsymbol{\sigma})$ is written as

$$[C^{mod}(\boldsymbol{x};\boldsymbol{u},\mathcal{T}_n,\boldsymbol{\sigma})]_{jj'} = (\delta t^{mes})^2 \sum_{k=1}^{\min(j,j')} \left(\left(\tilde{\sigma}_k^F\left(\boldsymbol{u}_{\leq k},\mathcal{T},\boldsymbol{x},\sigma^F\right)\right)^2 + \left(\sigma_k^P\right)^2 \right) . \tag{III.30}$$

In addition, since σ_k^P and σ_k^F are independent of k, simply rewritten as σ^P and σ^F, Equation (III.30) yields

$$[C^{mod}(\boldsymbol{x};\boldsymbol{u},\mathcal{T}_n,\boldsymbol{\sigma})]_{jj'} = (\delta t^{mes})^2 \sum_{k=1}^{\min(j,j')} \left(\left(\tilde{\sigma}_k^F\left(\boldsymbol{u}_{\leq k},\mathcal{T},\boldsymbol{x},\sigma^F\right)\right)^2 + \left(\sigma^P\right)^2 \right) , \tag{III.31}$$

in which

$$\tilde{\sigma}_k^F\left(\boldsymbol{u}_{\leq k},\mathcal{T},\boldsymbol{x},\sigma^F\right) = \left(1/\eta_k^T\left(\boldsymbol{u}_{\leq k},\mathcal{T},\boldsymbol{x}\right) - \eta_k^B\left(\boldsymbol{u}_{\leq k},\mathcal{T},\boldsymbol{x}\right)\right)^2 \left(\dot{y}_k\left(\boldsymbol{u}_{\leq k},\mathcal{T},\boldsymbol{x}\right)\right)^2 (\sigma^F)^2 . \tag{III.32}$$

For a given rolling environment $\mathcal{T}_n$, let $\mathcal{L}_n(\boldsymbol{f}^{mod};\boldsymbol{x},\boldsymbol{u},\mathcal{T}_n,\boldsymbol{\sigma})$ be the likelihood function defined as the conditional probability density function of the $\mathbb{R}^{N^{mes}}$-valued random variable $\boldsymbol{\mathcal{F}}^{mod}(\boldsymbol{u},\mathcal{T}_n,\boldsymbol{x})$ given $\boldsymbol{X}=\boldsymbol{x}$ at any point $\boldsymbol{f}^{mod}$ in $\mathbb{R}^{N^{mes}}$. Since $\boldsymbol{\varepsilon}^{mod}(\boldsymbol{x},\boldsymbol{\sigma})$ is a Gaussian centered random variable with values in $\mathbb{R}^{N^{mes}}$, with covariance matrix $[C^{mod}(\boldsymbol{x};\boldsymbol{u},\mathcal{T}_n,\boldsymbol{\sigma})]$, we have

$$\boxed{\mathcal{L}_n(\boldsymbol{f}^{mod};\boldsymbol{x},\boldsymbol{u},\mathcal{T}_n,\boldsymbol{\sigma}) = g_{\boldsymbol{\varepsilon}^{mod}(\boldsymbol{x},\boldsymbol{\sigma})}(\boldsymbol{f}^{mod}-\boldsymbol{\mathcal{F}}(\boldsymbol{u},\mathcal{T}_n,\boldsymbol{x}))\,,} \tag{III.33}$$

in which $g_{\boldsymbol{\varepsilon}^{mod}(\boldsymbol{x},\boldsymbol{\sigma})}$ is the multivariate Gaussian probability density function with zero mean vector and with covariance matrix $[C^{mod}(\boldsymbol{x};\boldsymbol{u},\mathcal{T}_n,\boldsymbol{\sigma})]$ (note that it can be verified that this covariance matrix is invertible for all $\boldsymbol{x}$ and $\mathcal{T}_n$).

For each rolling environment $\mathcal{T}_n$, there is an associated measurement vector $\boldsymbol{f}^{mes,n}\in\mathbb{R}^{N^{mes}}$. In the application, the given set $\{\mathcal{T}_1,\ldots,\mathcal{T}_{N^{\mathcal{T}}}\}$ of $N^{\mathcal{T}}$ is composed of 30 rolling environments. Concerning the identification of vector-valued hyperparameter $\boldsymbol{\sigma}$, we use the maximum likelihood method, that we write as:

$$\boxed{(\boldsymbol{x}^*,\boldsymbol{\sigma}^*) = \arg\max_{\boldsymbol{x},\boldsymbol{\sigma}} \sum_{n=1}^{N^{\mathcal{T}}} \log\left(\mathcal{L}_n(\boldsymbol{f}^{mes,n};\boldsymbol{x},\boldsymbol{u},\mathcal{T}_n,\boldsymbol{\sigma})\right) .} \tag{III.34}$$

For solving this optimization problem, the Basin-hopping algorithm, conceptualized by David J. Wales and Jonathan Doye in [71], is used due to its robustness and its ability not to be blocked in a local minimum. As a result, we obtain the optimal value $\boldsymbol{\sigma}^*$ of $\boldsymbol{\sigma}$ and the optimal value $\boldsymbol{x}^*$ of $\boldsymbol{x}$, which maximize the likelihood. Optimal value $\boldsymbol{\sigma}^*$ will be injected in the likelihood function (Equation (III.33)) in the formulation of the Bayesian inference and $\boldsymbol{x}^*$ will be used as a starting point for the MCMC algorithm for sampling the posterior probability distribution of $\boldsymbol{X}$.

III.4.4 Bayesian inference for estimating the posterior distribution

Using the prior probability density function defined in Section III.1 and the likelihood function defined in Equation (III.33), the Bayes rule allows us to derive the *a posteriori* probability distribution of the model parameters:

$$f_{\boldsymbol{X}}^{post}(\boldsymbol{x}) \propto \left(\prod_{n=1}^{N^{\mathcal{T}}} \mathcal{L}_n(\boldsymbol{f}^{mes,n}; \boldsymbol{x}, \boldsymbol{u}, \mathcal{T}_n, \boldsymbol{\sigma}^*) \right) f_{\boldsymbol{X}}^{prior}(\boldsymbol{x})\,. \tag{III.35}$$

In order to sampling the posterior probability distribution, we use a Markov Chain Monte Carlo (MCMC) approach. Among these methods, the Metropolis-Hastings (MH) algorithm (first described in [72] before being completed in [73]) relies on the initialization of a random walk, which evolves in the support of PDF $f_{\boldsymbol{X}}^{prior}$ uncertain parameters domain. At each iteration, a candidate is proposed, and it is accepted or rejected according to a specific acceptation rate, which depends on the prior distributions and the likelihood function.

In our case, random vector $\boldsymbol{X}$ has 9 components that have different sensitivity on the posterior (see Section II.4). Because of this, the MH algorithm shows a relatively low efficiency. Effectively, the acceptation of the candidate, which is randomly drawn, is going to be determined almost solely by the most influent random component of $\boldsymbol{X}$. A large number of iterations are likely to be needed to correctly estimate the less influent random component of $\boldsymbol{X}$. For instance, the braking efficiency parameters only have an influence on the braking part of the journey, and the influence of these parameters may be hidden by the influence of the other parameters in the new candidate.

For this reason, the Metropolis-within-Gibbs (MwG) algorithm has been used. It combines the classic Metropolis-Hastings algorithm with the Gibbs Sampling (GS) presented by S. and D. Geman in [74]. The main idea is that at each iteration k, the value x_k of a single component X_k changes according to several iterations of the classic MH algorithm. Hence, all the components of $\boldsymbol{X}$ are considered one-by-one, and their influences do not interfere (more details on the algorithm are given in [75]).

As all the MCMC algorithms, the MwG algorithm needs to be correctly initialized in order to observe a fast convergence and we have chosen the optimal value $\boldsymbol{x}^*$ estimated in Section III.4.3. First of all, we need to choose, at each iteration $k+1$, the new value x_k of component of X_k to be modified. We just propose to select them at random uniformly.

Second it is necessary to choose a transition probability for drawing the candidates of the random walk. As generally done in the literature, we have chosen to use the normal distribution for this task. The variance of this normal distribution is a key parameter for the convergence of the MCMC algorithm. Indeed, a low variance favors the observation of candidates close to the current one. In that case, the exploration, by the random walk, of the parameter domain will be slow. On the contrary, a high value for the variance results in value proposals that can be very different from the current point. A large proportion of them are likely to be rejected and the random walk may also have difficulties to converge.

In this work, the chosen proposal distribution is Gaussian, and we have adapted the variance to each component of $\boldsymbol{X}$ in order to get an acceptation rate close to 50%. We have also chosen $N^{MH} = 5$ meaning that each MwG iteration relies on 5 MH iterations (which seemed to be a good compromise for the considered numerical application).

For each candidate, we need to compute the likelihood function, which requires the simulation of the dynamic and energy models for $N^{\mathcal{T}} = 30$ different journeys. As these calculations are completely independent and are thus performed in parallel computing.

In order to assess the convergence of the method, we estimate $\mathbb{E}\{\|\boldsymbol{X}\|^2\}$. Its value remains relatively stable after N^{MwG} iterations. This means that the random walk has entered in its stationary state. Keeping the iterations following the N^{MwG}-th iteration allows us to gather points that are approximately distributed according to the posterior probability distribution of $\boldsymbol{X}$. Finally, the posterior probability distribution of $\boldsymbol{X}$ can be reconstructed from these samples thanks to a KDE method.

The application to the train system is presented in Figure III.7 and Figure III.8. The number of iterations needed for reaching convergence (abscissa axis) varies between the different graphs in Figure III.7 because the number of MwG iterations is not the same for all the component of $\boldsymbol{X}$.

In Figure III.8, we notice that the posterior probability distributions of random variables A, B and C remain relatively close to their prior ones. This means that the nominal values used for constructing the prior probability distributions were well adapted. This is not the case for the posterior probability distributions of mass error ΔM, auxiliary power P^a, and the four efficiency coefficients A_η, B_η, C_η, and D_η, which are very different from the prior ones, due to the fact that we had only little information for constructing the priors. In that case, the energy measurements allow us for better characterizing the posterior distributions.

III.4.5 Propagation of parameter uncertainties

In this section, we propose to quantify the impact of the residual variability associated with the posterior probability distribution of $\boldsymbol{X}$ on the energy consumption model. To this end, we use realizations of the posterior distributions generated by the MCMC algorithm during the Bayesian inference process. The dynamic equation and the energy consumption equation are solved for these realizations, and we rebuilt the PDFs of the random train speed and the random energy consumption thanks to the use of the Gaussian KDE method. In this section, we propose to only construct the energy-consumption probability distribution, as it involves both dynamic and energy consumption models, and we can directly compare it with measurements.

Figures III.9 and III.10 display (i) some realizations of the random energy consumption generated with the posterior distributions, (ii) its mean value, (iii) the realizations including the modeling error $\boldsymbol{\varepsilon}^{mod}$, and (iv) the envelops of the confidence regions for the 95% quantiles. The PDFs of (i) and (iii) are plotted at three given times. Finally, the energy measurements are plotted.

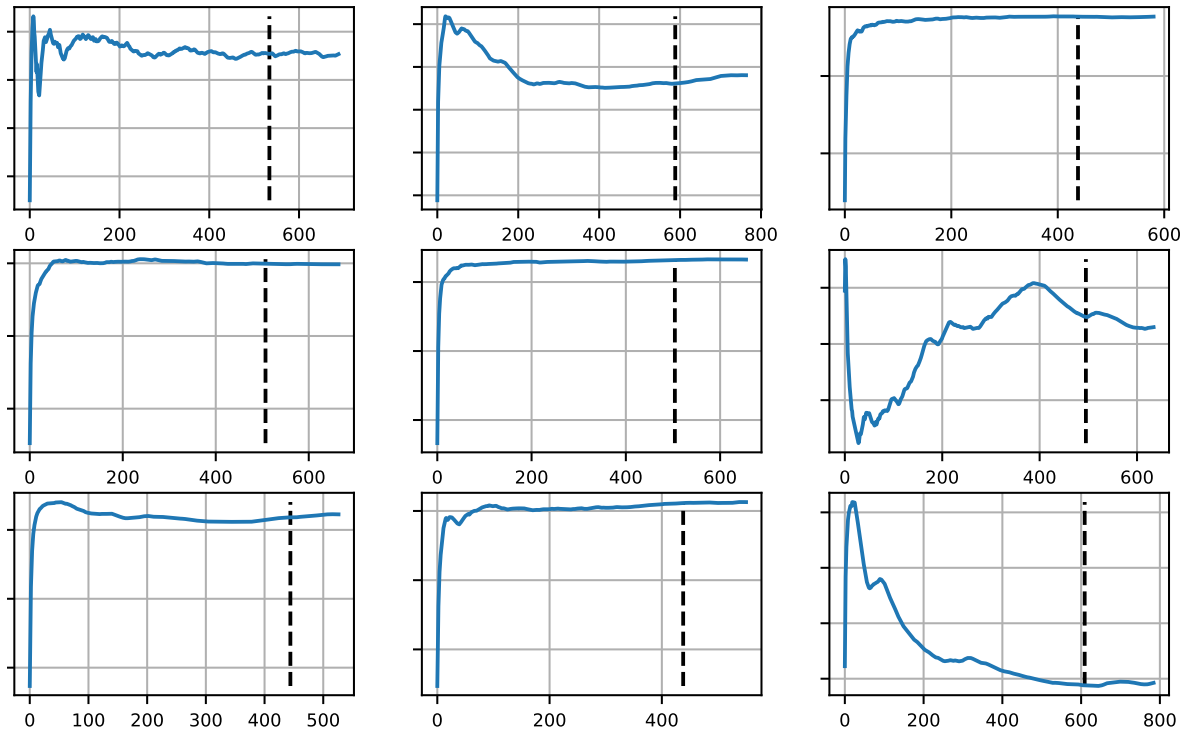

Figure III.7: Second-order moment of component of $\boldsymbol{X}$ as a function of the number of iterations. From top left to bottom right: ΔM, P^a, A, B, C, A_η, B_η, C_η, and D_η. Iteration number N^{MwG} is represented by a vertical dashed line in each figure.

The results displayed in Figure III.9 correspond to one of the 30 journeys used for the Bayesian inference. In this case the measurements are very close to the mean value estimated with the posterior probability distributions. In general, most of the energy consumption of the 30 journeys used in the Bayesian inference are well described by the blue envelop, but some of them are slightly outside (but never outside the green envelop). In other words, the estimated posterior probability distribution correctly characterizes the variability encountered during the 30 measured journeys. Without surprise, the envelop is growing over time as the uncertainty accumulates along the journey.

We also propose to assess the quality of the posterior probability distribution by plotting the statistics of the random energy consumption for journeys that have not been used for carrying out the Bayesian inference. These statistics are shown in Figure III.10 using the same conventions that the one used in Figure III.9.

Once again, the blue envelop is sufficient to describe the measurements, but in a slightly higher number of journeys, we observe measurements that are outside this blue envelope while remaining inside the green envelope. This justifies the introduction of σ^F and σ^P to properly quantify the modeling errors. The posterior probability distributions seem well adjusted because they are

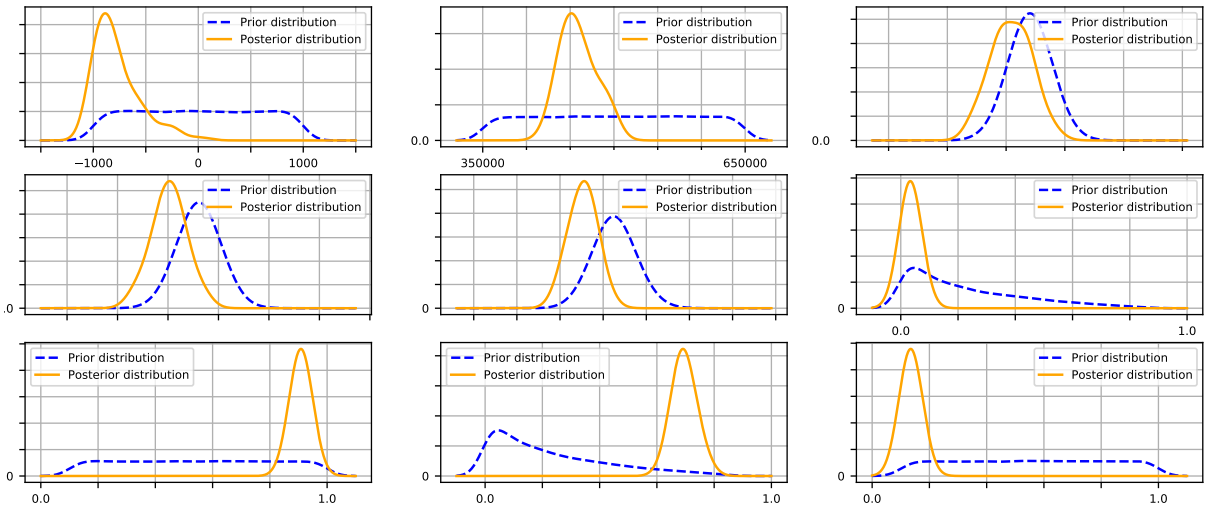

Figure III.8: PDF of the priors (blue) and posteriors (orange) for each uncertain parameter. From top left to bottom right (line-by-line): Mass error ΔM, auxiliary power P^a, Davis coefficients A, B, C, and the traction and braking efficiency parameters A_η, B_η, C_η, and D_η.

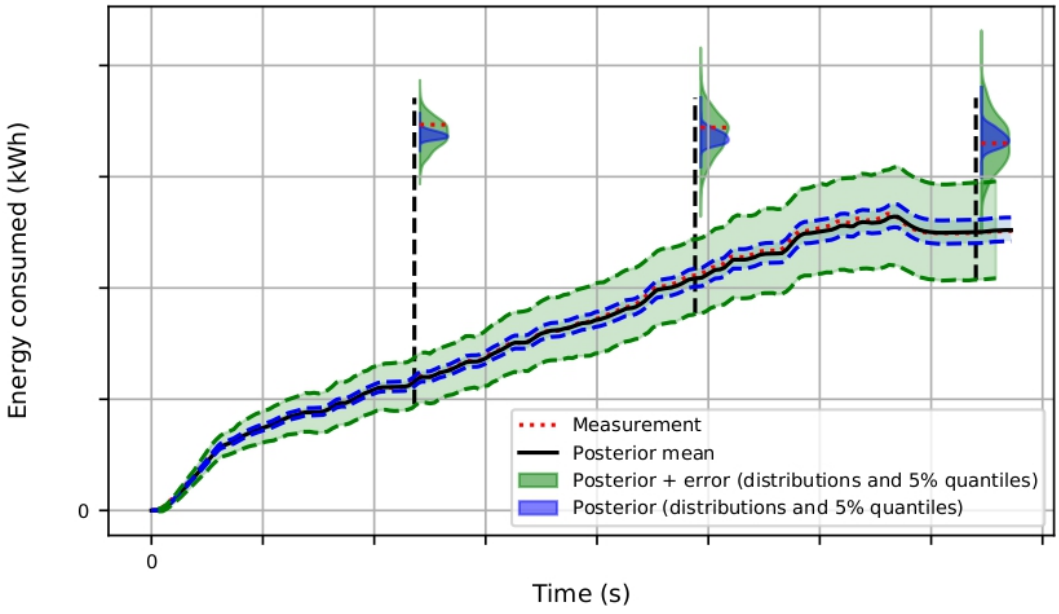

Figure III.9: (i) Realization of the random energy consumption generated with the posterior distributions (in blue), (ii) its mean value (in black), (iii) the realizations including the modeling error ε^{mod} (the green curves), and (iv) the envelops of the confidence regions for the 95% quantiles. The PDFs of (i) and (iii) are plotted at three given times. Finally, the energy measurements are plotted with a red dotted line - Serie 5

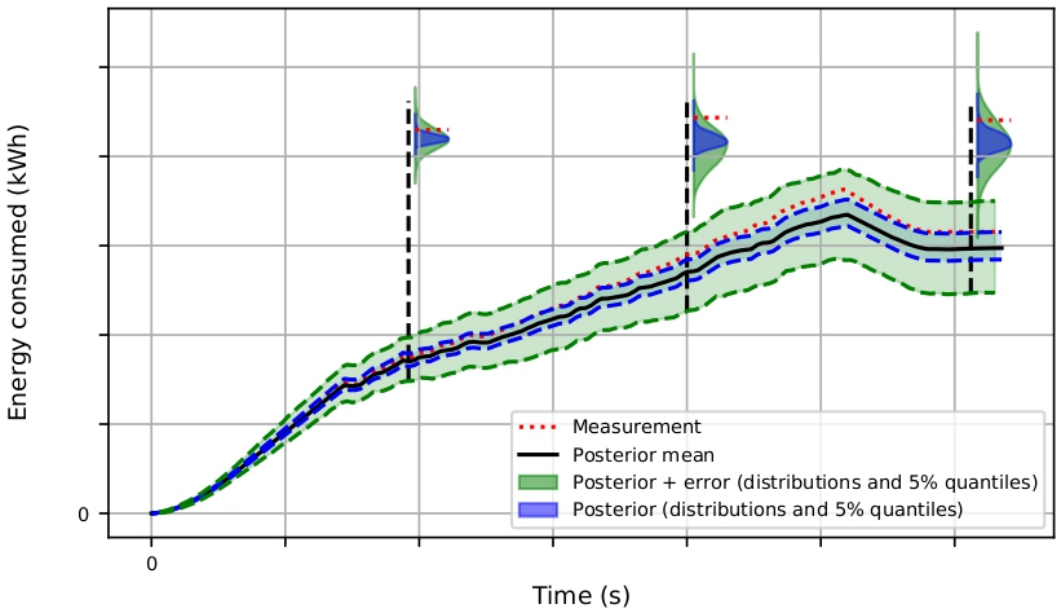

Figure III.10: (i) Realization of the random energy consumption generated with the posterior distributions (in blue), (ii) its mean value (in black), (iii) the realizations including the modeling error ε^{mod} (the green curves), and (iv) the envelops of the confidence regions for the 95% quantiles. The PDFs of (i) and (iii) are plotted at three given times. Finally, the energy measurements are plotted with a red dotted line - Serie 10

able to capture the observed variability among the trains responses, even for rolling environment $\mathcal{T}$ that have not been used during the Bayesian inference method.

III.5 Conclusion on the identification of the model parameters

Two different steps for the identification of the model parameters have been proposed in this chapter. The first step estimates punctual values for these model parameters. We look for experimental driver's commands associated with the measured speed of the train. Then, we search the train parameters that make the measured and simulated energy consumptions be as close as possible.

A Bayesian inference method has been proposed as a second identification step. After introducing the modeling errors, the posterior probability distributions of the model parameters have been sampled using a Metropolis-within-Gibbs algorithm. This algorithm has been selected to accelerate the convergence of the parameters having a lower influence on the model.

Using the developments presented in Chapter II and in this chapter, we can now perform the optimization of the driver's command in order to minimize the energy consumption of the train, while respecting constraints. This optimization is presented in Chapter IV.

Chapter IV

Optimization of the driver's command under constraints and uncertainty

The entries of the train system and the models are defined (in Chapter II) and the model parameters are identified (in Chapter III). This chapter focuses on the optimization of driver's command $\{u\}$, in order to minimize energy consumption $\mathcal{F}$, while ensuring punctuality (final position, final time, and final train speed) and traffic safety (curvilinear speed limitation on the track), all these constraints being gathered in Equation $c(\{u\}, \mathcal{T}, \mathcal{V}) = 0$. The optimization problem is defined by Equation (IV.1), but we have to keep in mind that model parameters $\mathcal{V}$ may be subjected to uncertainties (see Chapter III),

$$\{u^*\} = \underset{\substack{\{u\}\in\mathcal{U} \\ c(\{u\},\mathcal{T},\mathcal{V})=0}}{\arg\min} \; \mathcal{F}(\{u\}, \mathcal{T}, \mathcal{V}) \; . \tag{IV.1}$$

Solving this problem raises several difficulties, which are at the core of this work. We highlight below the main difficulties, and we present the methodology and the algorithms that we have implemented for solving this optimization problem.

- The **admissible space** $\mathcal{U}$ is in infinite dimension due to the definition of the driver's commands as functional. Then, its discretization will be in high dimension. The calculation costs become rapidly very high, that is why a reduced representation of the discretized driver's command has been implemented using a Principal Component Analysis. Concerning the discretization of the driver's command, we use larger intervals than the one used for solving the dynamic equations in order to reduce the numerical cost while the quality of the approximation is preserved. Then, the discretized driver's command are searched around an initial driver's command, which complies with the constraints.

- Due to the use of a probabilistic model of uncertainties, the optimization problem is formulated **probabilistic framework**. An appropriate methodology is introduced to integrate the constraints and to define the objective function. The MCMC is used for generating samples of random quantities for which probability distributions are given, allowing for estimating different mathematical expectations. The constraints are sensitive to uncertainties, and a robust strategy with respect to uncertainties is proposed.

- The optimization problem is very difficult due to its *a priori* **nonconvexity**. In particular, the constraints and cost function are **nonlinear**. The dynamic and energy consumption models are treated as black boxes.

- The fact that the dynamics verifies a chronology, we have to impose constraints that depend on the past and on the future inside the time interval $[t_s, t_f]$, which is the **major difficulty** and strongly impacts the complexity of the algorithm that will be proposed.

To overcome these difficulties, the resolution algorithm has been chosen carefully. To solve the dynamic problem, the work is decomposed in several steps. Section IV.1 details the choice of the optimization variable and the admissible set, emphasizing the methods that has been implemented to reduce the dimension of the search space that is originally very large. Section IV.2 describes the impact of the uncertainty on the optimization problem and proposes robust solution to take them into account. The optimization problem is then formulated in Section IV.3 and it is solved numerically in Section IV.4. The optimal solution is shown in Section IV.5. Finally, Section IV.6 draws conclusions of the chapter.

IV.1 Dimension reduction of the admissible set

As mentioned before, the admissible set of the optimization problem presented in Equation (IV.1) is in infinite dimension due to the definition of the driver's command as a functional. This first section aims reducing this dimension with two methods: a finite approximation followed by the use of a Principal Component Analysis (PCA).

IV.1.1 Finite approximation of the driver's command

The driver's command $\{u\}$ is a function that is assumed to belong to the space $\mathcal{C}^1([t_s, t_f])$ of the continuously differentiable functions defined in $[t_s, t_f]$ with values in $[-1, 1] \subset \mathbb{R}$ (see Chapter II). The problem is then to construct a finite approximation $\{u^{apx}\} \in \mathcal{C}^1([t_s, t_f])$ of $\{u\}$. Taking into account the available information, there are several possibilities for constructing such a finite approximation. The one retained consists in approaching the continuous function by a step function, and then to regularize the step function (piecewise constant function) by filtering to obtain a regularized function in $\mathcal{C}^1([t_s, t_f])$.

Discretization of the driver's command

As explained, in a first step, we approximate the driver's command with a piecewise constant function. The driver's command becomes a vector, $\boldsymbol{u} \in \mathbb{R}^N$, with a finite number of components, made up of the values attached to each discretized interval. The discretization step size Δt is chosen constant. Note that a small size implies a large number of discretization intervals, and consequently a large dimension N of vector $\boldsymbol{u}$ that has to be optimized. On the other hand, a large step size means that the driver's command is discretized with a small value of N. The precision would then be rough, and we might miss a best solution of the optimization problem, which could not be represented by such a discretization (see Chapter III).

Regularization by filtering of the discretized driver's command

After the discretization step, driver's command function $\{u\}$ is represented by a finite vector $\boldsymbol{u} \in \mathbb{R}^N$. From $\boldsymbol{u}$, we have now to rebuild $\boldsymbol{u}^{apx}$, which is the sampling of length N associated with a continuously differentiable function $\{u^{apx}\}$. We have used the following discrete linear filter that we directly present in the discrete form without given the underlying continuous linear one. Several types of discrete filters could be constructed. We have chosen the moving average discrete filter, denoted by T_{MA} and such that:

$$T_{MA}\left(\boldsymbol{u}, i\right) = \frac{1}{L+1} \sum_{n=-L/2}^{L/2} w_n u_{i+n} \qquad \text{for } n = L/2, ..., N - L/2\,, \tag{IV.2}$$

with $L+1$ the size of the moving average window (L is chosen as an even integer). A window covering 5 discretization intervals, which is equivalent to $L+1 = 201$ points, is a good compromise between smoothing the data, and preserving information. We complete the missing values with 1 at the beginning (as the train has to accelerate) and -1 at the end (as the train needs to stop). The weight coefficients w_n have to verify the condition $\sum_{n=-L/2}^{L/2} w_n = 1$. They are chosen to favor the central points. A parabolic profile is chosen, and the coefficients are such that:

$$w_n = \frac{-6}{L^3}\left(n - L/2\right)\left(n + L/2\right), \qquad n \in \{-L/2, \ldots, 0, \ldots, L/2\}\,. \tag{IV.3}$$

The weight coefficients are shown in Figure IV.1. An example of the discretized driver's command and its regularization is given in Figure IV.2. The regularization driver's command is well in $\mathcal{C}^1$ without losing information.

IV.1.2 Reduction of the admissible search using a PCA

While the finite approximation $\boldsymbol{u}^{apx}$ (regularization of the discretization) of the driver's command $\{u\}$ will be used in the dynamic and energy consumption models, the PCA is performed using the discretized driver's command $\boldsymbol{u}$. The covariance matrix for this PCA is estimated with a finite family of vector-valued solutions $(\boldsymbol{u}^*_{\Delta,k})_k$ of a family of optimization problems, which are constructed in this section.

Construction of the family of optimization problems and their solutions

For given train parameters (drawn from the posterior distributions of $\boldsymbol{X}$), the optimization problem is written as

$$\boxed{\boldsymbol{u}^*_{\Delta,k} = \underset{\substack{\boldsymbol{u}_\Delta \in [-1,1]^{N_\Delta} \\ \boldsymbol{c}(\boldsymbol{T}(\boldsymbol{u}_\Delta, \boldsymbol{x}_k), \mathcal{T}, \boldsymbol{x}_k) = \boldsymbol{0}}}{\arg\min} \mathcal{F}\left(\boldsymbol{T}\left(\boldsymbol{u}_\Delta, \boldsymbol{x}_k\right), \mathcal{T}, \boldsymbol{x}_k\right),} \tag{IV.4}$$

where $\boldsymbol{T}$ is derived from $\boldsymbol{T}_{MA}(\boldsymbol{u}) = (1, \ldots, T_{MA}(\boldsymbol{u}, L/2), \ldots, T_{MA}(\boldsymbol{u}, N - L/2), \ldots, -1)$ for which the constraints are taken into account (see Section IV.2.1) and where N_Δ is the

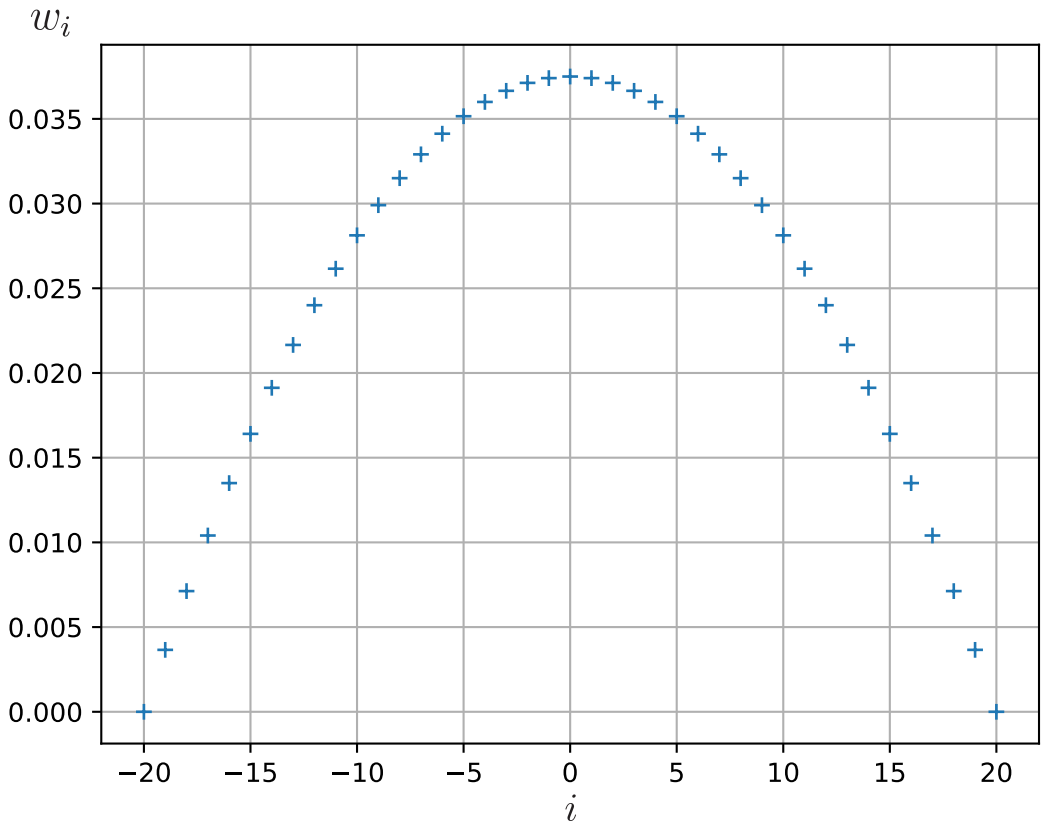

Figure IV.1: Weight coefficients of the moving average filter.

number of components of $\boldsymbol{u}$ without repetition, the reduced vector obtained being denoted by $\boldsymbol{u}_\Delta$. This optimization problem is deterministic, as the uncertain vector is no longer involved (but only its realizations). The constraints and the objective function are deterministic as well. In this paragraph, we show the different steps that allowed the construction of the solutions of the studied problem. First, since the problem is very sensitive to the constraints, we propose to start the optimization process with a value $\underline{\boldsymbol{u}}$ of discretized driver's command $\boldsymbol{u}$, which respects the constraints. Second, the constraints and objective function can directly be defined from the deterministic models (see Equations (IV.23), (IV.24), and (IV.25) for the constraints, and Equation (IV.36) for the cost function). Finally, this constrained optimization problem is difficult to solve because the constraints are a nonlinear mapping of $\boldsymbol{u}$. For that, we first transform this constrained optimization problem in an unconstrained optimization problem thanks to the use of Lagrange multipliers. Second, the resulting unconstrained optimization problem is solved using the CMA-ES algorithm [76] and presented in Appendix B.

For each realization $\boldsymbol{x}_k$, there is an associated optimization problem defined by Equation (IV.4). The optimal solutions are noted $\boldsymbol{u}^*_{\Delta,k}$ and are gathered in what we call the family of solution $\left(\boldsymbol{u}^*_{\Delta,k}\right)_k$. The solutions $\boldsymbol{u}^*_{\Delta,k}$ are optimal in specific configurations $\boldsymbol{x}_k$, which are representative of the whole variability of trains (if enough realizations are drawn). About 100 different configurations have been drawn and the associated optimization problems have been

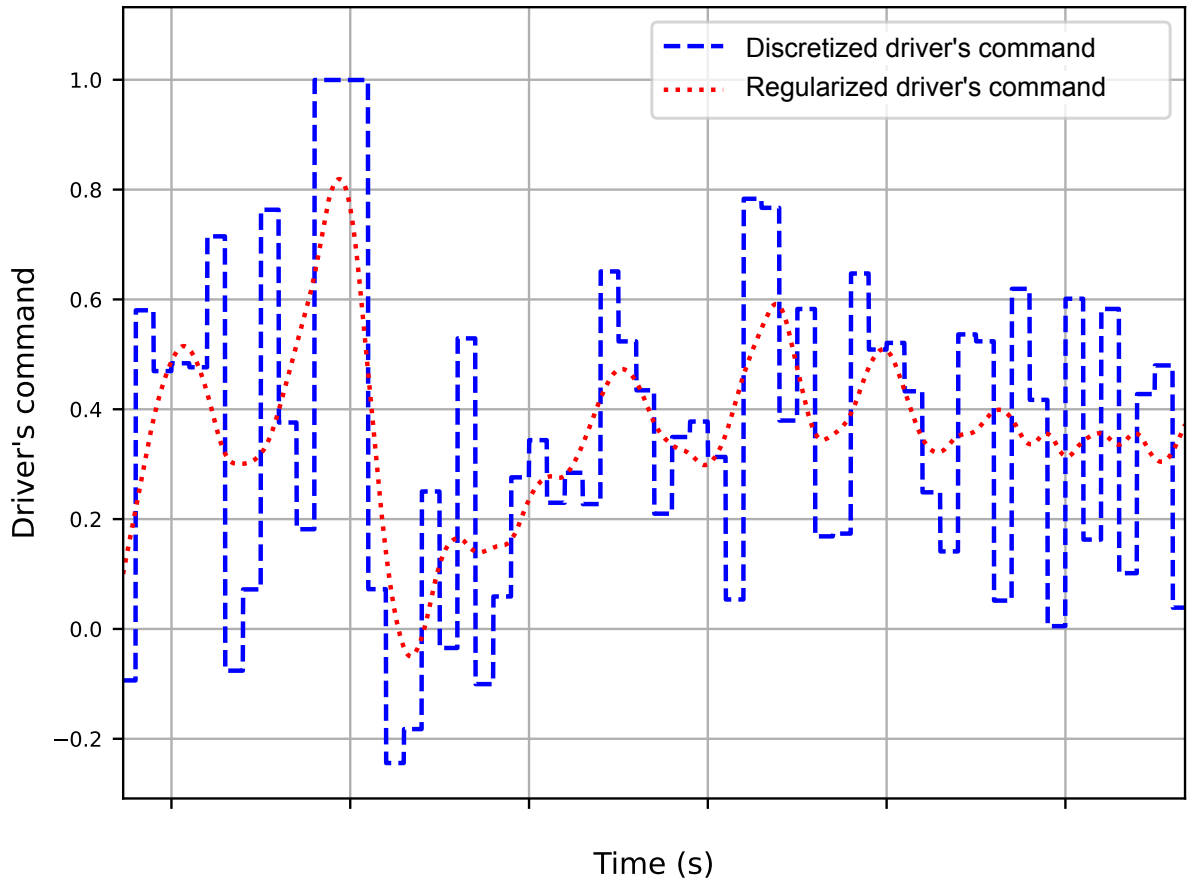

Figure IV.2: Discretized driver's command (in blue) and regularized driver's command (in red) as a function of time.

solved. The optimal speed profiles of the different solutions are shown in Figure IV.3. The measurements are shown in red, and the initial speed profile and energy consumption associated with the initial point of CMA-ES algorithm is plotted in green. The speed limitation is in blue.

We can notice that the energy consumption of the train with the optimal driver's command is 28% lower than the measured trajectory. The solution highlights that reducing the maximal speed of the train also reduces the resistant force applied to the train (especially the aerodynamic force at high-speed). Moreover, the maximum speed of the optimal trajectory is adapted to the track layout. Indeed, it is not efficient to use a lot of traction when the declivity is positive as a higher part of the energy will be injected in the track. As a consequence, the optimal result saves energy, but it also implies delays. As a consequence, the optimal speed profile compensates for the delay accumulated in the middle part of the journey by braking later, in order to remain on time. With this process, all the constraints are verified by the optimal solution. Figure IV.4 presents the same quantities as Figure IV.3 but plotted as a function of time to have a better idea of the time arrival constraint. We also add the driver's command in this figure.

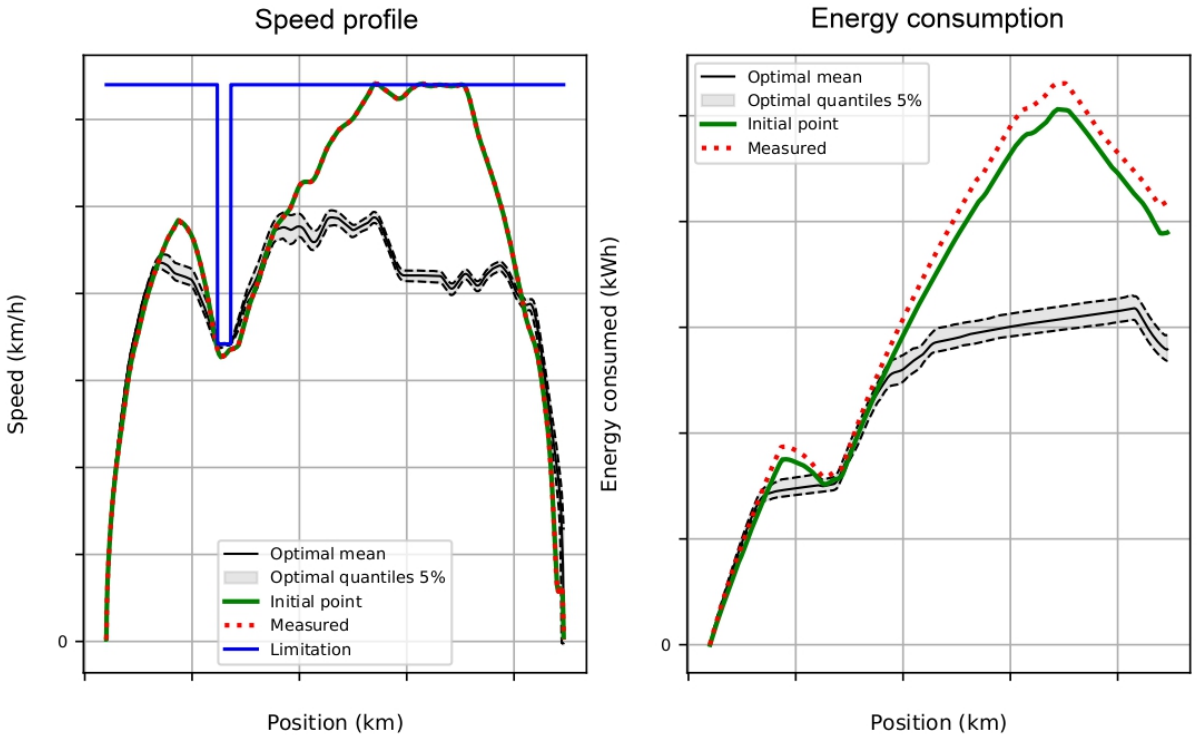

Figure IV.3: Speed profiles (left) and energy consumption (right) associated with the different optimal driver's command, $\boldsymbol{u}^*_{\Delta,k}$, for the configuration $\boldsymbol{x}_k$ depending on the position. The measurements are shown in red, and the initial speed profile and energy consumption associated with the initial point of CMA-ES algorithm is plotted in green. The speed limitation is in blue.

Configuration $\boldsymbol{x}_k$ under consideration has a significant influence on the dynamic solution and its energy consumption. Effectively, between the best and the worst configuration, the variation of the consumed energy reaches 8% during the complete journey. The quantile envelope has not the same size all along the journey. In some parts of the journey (close to the speed-limitation modification, at the starting and final time of the journey), the optimal trajectories have the same speed. This is due to the fact that these parts have an important impact on the energy consumption of the train. Modifying the speed at these positions may drastically reduce the quality of the solution in terms of energy saving. On the contrary, some parts of the journey allow more flexibility depending on configuration $\boldsymbol{x}_k$. In first part of the journey (until the modification of speed limitation), the optimal solution get advance (around 20 seconds) compared to the measurements even if it does not directly reduce the consumed energy. This advance allows for modifying the second part of the journey, which has a higher impact on the energy consumption. All the energy saved by the optimal solution is on the second part of the journey, but it also results in the choices made in the first part.

We can also observe in the center graph of Figure IV.4 that limiting the use of braking allows for saving energy. The optimal driver's command $\boldsymbol{u}^*_{\Delta,k}$ has an important variability on some specific parts of the journey. This means that on some parts of the journey, the optimal driver's

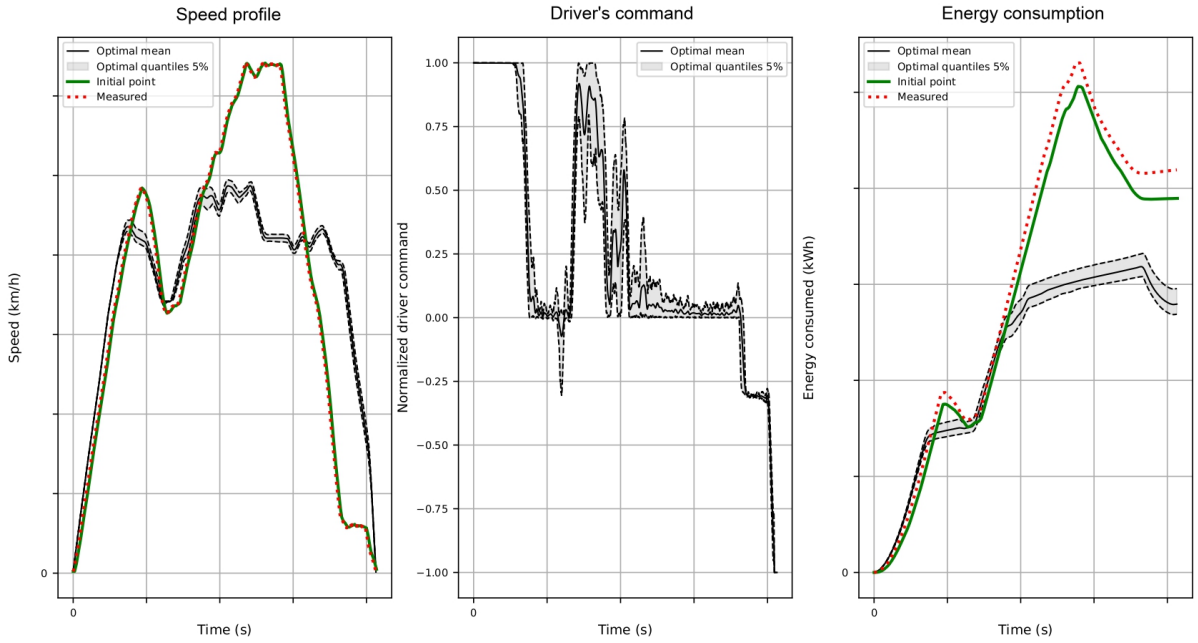

Figure IV.4: Speed profiles (left), driver's command (middle), and energy consumption (right) associated with the different optimal driver's command, $\boldsymbol{u}^*_{\Delta,k}$, for the configuration $\boldsymbol{x}_k$ depending on time. The measurements are shown in red, and the initial speed profile and energy consumption associated with the initial point of CMA-ES algorithm is plotted in green.

command is similar regardless of the train configuration. The PCA will use this property, as it may be useless to consider these parts in the optimization because the optimal solutions are identical whatever the configuration $\boldsymbol{x}_k$.

The driver cannot know which one of the optimal solution in the family of solutions should be used to minimize the consumed energy due to the fact that the configuration $\boldsymbol{x}_k$ of the train is unknown at the beginning of the journey. Therefore, we are going to construct one deterministic driver's command that the driver will be able to use.

Principal Component Analysis as a tool for constructing a reduced representation of the discretized driver's command

The use of the terminology Principal Component Analysis is, in this case, equivalent to the terminology "data compression". The family of solutions computed in the previous paragraph is going to be used in the PCA. The optimal driver's commands $\boldsymbol{u}^*_{\Delta,k}$ correspond to the final iterations of the CMA-ES algorithm that generates driver's commands $\boldsymbol{u}_{\Delta,\kappa}$ for which the constraints are not always satisfied. However, many of them are very similar to each other (as the algorithm has a small amplitude of exploration when it is close to be converged). For this reason, we only keep the driver's command $\boldsymbol{u}_{\Delta,\kappa}$ generated just before the convergence. For instance, a population of 30 points is kept, each 50 CMA-ES iteration (so that the driver's command are not too similar), for the last 500 iterations (so that the constraints are almost respected), for each

of the 100 optimization calculations run in the family of optimal solutions. The subset of kept driver's commands are denoted by:

$$\boldsymbol{u}_{\Delta,\kappa}, \qquad \kappa = 1, \ldots, N_u \,. \tag{IV.5}$$

In order to favor the driver's commands associated with a small energy consumption, we propose to divide the kept driver's commands $\boldsymbol{u}_{\Delta,\kappa}$ by their associated energy consumption e_κ. This operation is possible because $\boldsymbol{u}_{\Delta,\kappa}$ is never associated with an energy consumption equal to zero. We then introduce the following first normalization,

$$\widetilde{\boldsymbol{u}}^\kappa = \frac{\boldsymbol{u}_{\Delta,\kappa}}{e_\kappa} \,. \tag{IV.6}$$

The empirical mean vector and the empirical variance of each component are computed by:

$$\underline{\widetilde{\boldsymbol{u}}} = \frac{1}{N_u} \sum_{\kappa=1}^{N_u} \widetilde{\boldsymbol{u}}^\kappa \quad , \quad {\sigma_j}^2 = \frac{1}{N_u - 1} \sum_{\kappa=1}^{N_u} \left(\widetilde{u}_j^\kappa - \underline{\widetilde{u}}_j \right)^2 \,. \tag{IV.7}$$

Then, we introduce a second normalization denoted by $\widetilde{\boldsymbol{u}}^\kappa$ such as its components are written as

$$\hat{u}_j^\kappa = \frac{\widetilde{u}_j^\kappa - \underline{\widetilde{u}}_j}{\sigma_j} \,. \tag{IV.8}$$

All the normalized driver's commands $\hat{\boldsymbol{u}}^\kappa$ are gathered in the normalized driver's command matrix $[\hat{u}]$,

$$[\hat{u}] = \left[\hat{\boldsymbol{u}}^1, \ldots, \hat{\boldsymbol{u}}^{N_u} \right] \,. \tag{IV.9}$$

This matrix is composed of $N_u = 30\,000$ normalized driver's commands, each of them being a vector of a size close to 200. The PCA is then applied to matrix $[C_{\hat{u}}] = (1/N_u)\,[\hat{u}]\,[\hat{u}]^T$. The first step is solving the eigenvalue problem for matrix $[C_{\hat{u}}]$,

$$[C_{\hat{u}}]\,\boldsymbol{\Phi}^i = \lambda_i \boldsymbol{\Phi}^i \,, \tag{IV.10}$$

where λ_i and $\boldsymbol{\Phi}^i$ are respectively the eigenvalues and eigenvectors. Let $[\lambda]$ be the diagonal matrix of the eigenvalues sorted in descending order and let $[\Phi]$ be the matrix of the eigenvectors sorted in the same order than the eigenvalues. The normalized driver's command matrix $[\hat{u}]$ is such that:

$$[\hat{u}] = [\Phi]\,[\lambda]^{1/2}\,[q] \,, \tag{IV.11}$$

where $[q]$ is a rectangle matrix with 200 rows and N_u columns. The second step of the method is to keep the dominant eigen subspace of dimension $(m) < 200$ in order to obtain a reduced representation of $[\hat{u}]$ (or its compression). The dimension of the optimization problem is then reduced to (m) instead of 200. The value of (m) is chosen regarding the relative truncation error that is written as

$$\frac{\| [\hat{u}] - \left[\hat{u}^{(m)} \right] \|_F}{\| [\hat{u}] \|_F} \leq \varepsilon \,, \tag{IV.12}$$

where ε is small enough and $\left[\hat{u}^{(m)} \right]$ is given by:

$$\boxed{\left[\hat{u}^{(m)} \right] = \left[\Phi^{(m)} \right] \left[\lambda^{(m)} \right]^{1/2} \left[q^{(m)} \right] \,.} \tag{IV.13}$$

The matrix $[q^{(m)}]$ can then be computed as the following projection of $[\hat{u}]$ as

$$\left[q^{(m)}\right] = \left[\lambda^{(m)}\right]^{-1/2} \left[\Phi^{(m)}\right]^T \left[\hat{u}^{(m)}\right] . \tag{IV.14}$$

In Figure IV.5, we have represented $1 - \varepsilon$. We have chosen $(m) = 60$ corresponding to a truncated error of 1%.

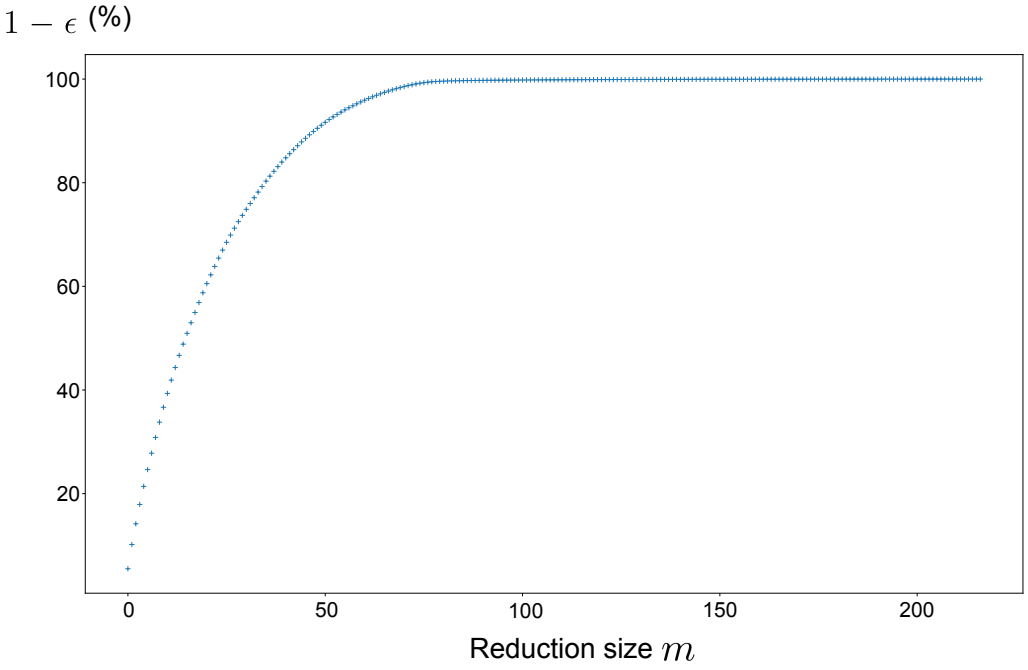

Figure IV.5: Cumulative explained variance depending on the number of principal axes conserved.

Consequently, in the optimization algorithm based on the use of the CMA-ES algorithm, instead of drawing a value of matrix $[\hat{u}]$, we draw a value of matrix $[q^{(m)}]$ and then we compute the associated value of $[\hat{u}] \simeq [\hat{u}^{(m)}]$ using the Equation (IV.13) and then processing the back normalization defined by Equation (IV.8) and (IV.6) in order to compute $\boldsymbol{u}_{\Delta,\kappa}$ for $\kappa = 1, \ldots, N_u$. We verify that the chosen ε is such that the values of the components of $\boldsymbol{u}_{\Delta,\kappa}$ belongs to interval $[-1, 1]$. Finally, the discrete filter is applied.

IV.2 Formulation of the optimization problem under uncertainties

In Chapter III, we have identified the posterior distributions of random vector $\boldsymbol{X}$ in order to be able to take into account uncertainties. The integration of uncertainties in the model is performed in the probabilistic framework. Consequently, we have to adapt the formulation of the constraints to this framework. As a matter of fact, the constraints and the cost function introduced in Equation (IV.1) have to be adapted and are presented in next two subsections.

IV.2.1 Deterministic aspects of the constraints

We begin presenting the deterministic aspects associated with the constraints before adapting them to the probabilistic framework. As we already know, the optimal solution needs to verify a set of constraints. These constraints assure the security, the train punctuality, and the passenger comfort. Note that proposing an optimal solution that would not verify the security criteria would completely be inadmissible. On the opposite, proposing a solution for which the train arrives few seconds earlier in the train station or which slightly overtake the passenger comfort criteria during a short part of the journey is a second priority. This is the reason why, in order to help the convergence of the optimization algorithm, a small tolerance will be added concerning the punctuality. Consequently, we choose to decompose the set of constraints into two groups (two subsets). The security constraints that are a first priority. The comfort and punctuality constraints are a second priority. In particular, the comfort constraint is not restrictive in the longitudinal track axis. It would be more important in the lateral axis, this is why the comfort constraints are always considered as verified (the motor capacity already limits the longitudinal acceleration and jerk).

Searching for a solution, which does not verify the set of constraints is an absurdity. For example, "if we are looking for a solution that minimizes the consumed energy during the journey, a naive solution would be to brake and do not advance, as it does not consume energy at all. Nevertheless, this solution is not interesting, as it does not bring the passengers to their destination". In this problem, the optimal solution is mainly driven by the constraints.

Deterministic formulation of the constraints

In this paragraph, we formulate the constraints for a given realization $\boldsymbol{x}$ of the random vector $\boldsymbol{X}$ (subscript k has been removed). It is a deterministic formulation in the sense that we give only one realization of the train system. We begin describing the security constraints before highlighting the punctuality constraints.

In longitudinal dynamics, the security constraints are assumed to be verified as long as the train speed is smaller than the curvilinear speed limitation v^{max} all along the journey. This curvilinear speed limitation is fixed to the commercial running speed but can decrease on specific part of the journey due to a high-curve radius, a tunnel, a switch, and a well identified state of the track deterioration. The curvilinear speed limitation $\{v^{max}(s), s \in [s_s, s_f]\}$ is attached to a track and depends on its curvilinear abscissa s. Overtaking this value yields the automatic stop of the train, which has an important impact on the consumed energy, but also on the dynamics of the train that follows. The speed limitation constraints can be expressed by:

$$\text{for } t_s \leq t \leq t_f, \qquad \dot{y}\left(t; \{u_{\leq t}\}, \mathcal{T}, \boldsymbol{x}\right) \leq v^{max}\left(y\left(t; \{u_{\leq t}\}, \mathcal{T}, \boldsymbol{x}\right)\right) . \tag{IV.15}$$

This inequality imposes that, at any time t, the train speed $\dot{y}\left(t; \{u_{\leq t}\}, \mathcal{T}, \boldsymbol{x}\right)$ does not exceed the speed limitation $v^{max}\left(y\left(t; \{u_{\leq t}\}, \mathcal{T}, \boldsymbol{x}\right)\right)$, which is defined at the actual train position $y\left(t; \{u_{\leq t}\}, \mathcal{T}, \boldsymbol{x}\right)$. As mentioned before, the driver's command is discretized and regularized thanks to the use of $\boldsymbol{T}$ (see Section IV.1.2) but adapted to the use of $\boldsymbol{q}^{(m)}$. Using the time-sampling notation introduced in Equation (IV.2), Equation (IV.15) becomes, for $n = 1, ..., N$,

$$\dot{y}\left(t_i; \boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right)_{\leq t_i}, \mathcal{T}, \boldsymbol{x}\right) \leq v^{max}\left(y\left(t_i; \boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right)_{\leq t_i}, \mathcal{T}, \boldsymbol{x}\right)\right) , \tag{IV.16}$$

where $\boldsymbol{q}^{(m)}$ is the candidate driver's command proposed by the CMA-ES algorithm. It can be seen that the security gathers N inequality constraints. This high number N of constraints may slow down the convergence of the optimization algorithm. We propose to combine all the inequalities in only one equality, which is written as

$$\sum_{i=1}^{N} f_o\left(t_i; \boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right)_{\leq t_i}, \mathcal{T}, \boldsymbol{x}\right) = 0\,, \tag{IV.17}$$

where f_o is "an overtake function" that, for a given time t_i, yields the exceeding of the train speed,

$$f_o\left(t_i; \boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right)_{\leq t_i}, \mathcal{T}, \boldsymbol{x}\right) = \begin{cases} 0 & \text{if } \dot{y}(t_i) \leq v^{max}\left(y(t_i)\right)\,, \\ \dot{y}(t_i) - v^{max}\left(y(t_i)\right) & \text{otherwise,} \end{cases} \tag{IV.18}$$

where the notation $y(t_i)$ is used instead of $y\left(t_i; \boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right)_{\leq t_i}, \mathcal{T}, \boldsymbol{x}\right)$ and $\dot{y}(t_i)$ is used instead of $\dot{y}\left(t_i; \boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right)_{\leq t_i}, \mathcal{T}, \boldsymbol{x}\right)$.

Note that, replacing the N inequalities by only one equality constraint, which is satisfied when each term is zero, can induce a larger number of iterations than the case for which N inequalities would be imposed. Nevertheless, the numerical experiments that we have performed show that this formulation is faster.

Concerning the punctuality constraints, neglecting the final speed constraint would yield an optimal solution that would not be associated with a driver's command imposing to brake on the final part of the journey for two reasons. On the one hand, braking is dissipating energy (even if a part is recovered). On the other hand, if we do not reduce the speed to stop at the end of the journey, we may arrive at the final position in advance (as the speed is higher). In this situation, the optimal solution is going to reduce the maximum speed to arrive on time (and respect the final position constraint) while reducing the energy consumption.

These constraints impose a specific arrival destination s_f for the train, with an appropriate speed v_f, and a given time t_f. The destination is often chosen as the train station and the final speed as 0, but any value may be chosen. The final time could be the exact arrival time, but it could also be smaller to simulate a departure delay for the journey. The two punctuality constraints (one on the final position, one on the final speed) can be expressed by:

$$\int_{t_s}^{t_f} \dot{y}\left(t, \{u_{\leq t}\}, \mathcal{T}, \boldsymbol{x}\right) dt = s_f - s_s\,, \tag{IV.19}$$

$$\int_{t_s}^{t_f} \ddot{y}\left(t, \{u_{\leq t}\}, \mathcal{T}, \boldsymbol{x}\right) dt = v_f - v_s\,. \tag{IV.20}$$

Once again, we discretize the driver's command and apply the regularization $\boldsymbol{T}$ described in Section IV.1.2. The two constraints are simplified by approaching the integrals by sums,

$$\sum_{i=1}^{N} \dot{y}\left(t_i; \boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right)_{\leq t_i}, \mathcal{T}, \boldsymbol{x}\right) \Delta t \simeq s_f - s_s\,, \tag{IV.21}$$

$$\sum_{i=1}^{N} \ddot{y}\left(t_i; \boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right)_{\leq t_i}, \mathcal{T}, \boldsymbol{x}\right) \Delta t \simeq v_f - v_s \,. \tag{IV.22}$$

Thus, only two equalities are necessary to verify the punctuality constraints. For more clarity, we note c_1, c_2, and c_3 the speed-limitation function (see Equation (IV.17)), the final-speed function (see Equation (IV.21)), and the final-position function (see Equation (IV.22)), respectively. These are function of $\boldsymbol{q}^{(m)}$ and depend on configuration $\boldsymbol{x}$ and rolling environment $\mathcal{T}$ because the speed and acceleration of the train also depend on these quantities,

$$c_1\left(\boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right), \mathcal{T}, \boldsymbol{x}\right) := \sum_{i=1}^{N} f_o\left(t_i, \boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right)_{\leq t_i}, \mathcal{T}, \boldsymbol{x}\right) \,, \tag{IV.23}$$

$$c_2\left(\boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right), \mathcal{T}, \boldsymbol{x}\right) := \sum_{i=1}^{N} \dot{y}\left(t_i; \boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right)_{\leq t_i}, \mathcal{T}, \boldsymbol{x}\right) \Delta t - s_f + s_s \,, \tag{IV.24}$$

$$c_3\left(\boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right), \mathcal{T}, \boldsymbol{x}\right) := \sum_{i=1}^{N} \ddot{y}\left(t_i; \boldsymbol{T}\left(\boldsymbol{q}^{(m)}, \boldsymbol{x}\right)_{\leq t_i}, \mathcal{T}, \boldsymbol{x}\right) \Delta t - v_f + v_s \,. \tag{IV.25}$$

These three functions completely describe the constraints of the train system. For a fixed value of $\boldsymbol{q}^{(m)}$, the constraints are verified if the values of the three functions c_1, c_2, and c_3 at point $\boldsymbol{q}^{(m)}$ are zero. We note $\boldsymbol{c}$ the function such that:

$$\boldsymbol{c} = (c_1, c_2, c_3) \,. \tag{IV.26}$$

Adaptation of the deterministic formulation of the constraints to the probabilistic framework

Considering uncertainties, the physical quantities (train speed, position) attached to the train dynamics are random variables, and the formulation of the constraints function defined by Equations (IV.26) with Equations (IV.23) to (IV.25) has to be adapted. For the security constraint, Equation (IV.15) is replaced by the following one,

$$\text{for } t_s \leq t \leq t_f, \qquad \mathbb{P}\left(\dot{Y}(t) \leq v^{max}\left(Y(t)\right)\right) \geq p_1 \,, \tag{IV.27}$$

where the notations $Y(t)$ and $\dot{Y}(t)$ are defined in Equation (II.9) that is adapted to the probabilistic case. The constant p_1 is a tolerance threshold. The security constraints being a first priority, we choose for p_1 a value close to 1.

The deterministic formulations of the punctuality constraints (Equations (IV.19) and (IV.20)) are also adapted to the probabilistic framework induced by the uncertainty. Since, taking the mathematical expectation of the random constraints is not sufficient, we formulate it in term of probability inequalities,

$$\mathbb{P}\left(\left|\int_{t_s}^{t_f} \dot{y}\left(t, \{u_{\leq t}\}, \mathcal{T}, \boldsymbol{X}\right) dt - s_f + s_s\right| \leq \varepsilon_1\right) \geq p_2 \,, \tag{IV.28}$$

$$\mathbb{P}\left(\left|\int_{t_s}^{t_f} \ddot{y}\left(t, \{u_{\leq t}\}, \mathcal{T}, \boldsymbol{X}\right) dt - v_f + v_s\right| \leq \varepsilon_2\right) \geq p_3\,, \tag{IV.29}$$

where ε_1 and ε_2 are two acceptability thresholds. For instance, the train can arrive at few meters from the exact final position. In practice, we propose to set $\varepsilon_1 = 5\,m$ and $\varepsilon_2 = 0.3\,m/s$. These values might be different depending on the sensitivity of the train system. Finally, p_2 and p_3 have to be very close to 1. For a given configurations $\boldsymbol{x}$, we cannot tolerate that the train does not arrive within the acceptance interval defined before.

Sensitivity of the constraints to the uncertain parameters

Before continuing to define the different items of the optimization problem, we have to verify the plausibility of the constraints definition. Indeed, we seek for one deterministic optimal driver's command, which is robust to uncertainties. Does this solution exist? The answer is not easy. For example, if the identified mass is responsible for variations in the final position of more than 5 meters, we cannot find one unique deterministic driver's command, which verifies the constraints regarding all the possible configurations. Figure IV.6 displays the probability density functions of $c_j\left(\underline{\boldsymbol{u}}, \mathcal{T}, \boldsymbol{X}\right)$, $j = 1, 2, 3$, which show the influence of the randomness of $\boldsymbol{X}$ on the speed limitation, on the final position, and on the final speed constraints for a given driver's command $\underline{\boldsymbol{u}}$. These PDF are estimated using the KDE method on the simulated realizations.

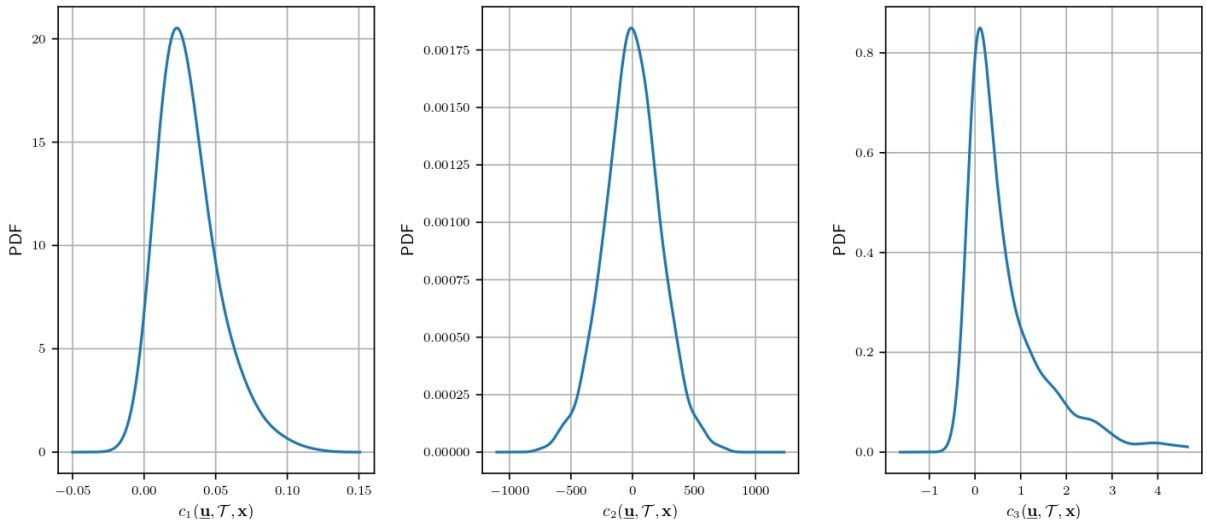

Figure IV.6: Probability density function of $c_j\left(\underline{\boldsymbol{u}}, \mathcal{T}, \boldsymbol{X}\right)$, for $j = 1$ (left), $j = 2$ (middle), and $j = 3$ (right) for a given driver's command $\underline{\boldsymbol{u}}$.

It can be observed on the middle figure that the final position constraint seems to be respected in mean (centered around 0) but the uncertainties are responsible for a wide variability in the final position: this given driver's command yields an arrival position that is located at 500 meters before or after the train station. Most of the realizations $\boldsymbol{x}$ of $\boldsymbol{X}$ satisfy the final speed constraint (right graph), but a significant number of realizations exceed the tolerance value defined by $0.3\,m/s$. Consequently, the problem is not well-posed, and the probabilistic formulation proposed in the previous paragraph cannot be kept and must be improved. Effectively, the ran-

dom vector $\boldsymbol{X}$ has a too high influence on the constraints, and a single deterministic driver's command cannot verify the constraints for all the possible configurations of the train.

Transformation of the driver's command to be robust to uncertainties

As it has been proven in the previous paragraph, the optimization problem defined by Equation (IV.1) is not well-posed in the defined probabilistic framework: it is not possible to find a deterministic driver's command that verifies almost surely the constraints. For this reason, we propose to slightly modify the formulation of the problem. We search for a deterministic driver's command that is easy to transform into a new driver's command, which verifies the constraints, in order to minimize the energy consumption of the train. But what does this transformation refer to? The security and punctuality constraints have two different impacts on the train system. Consequently, the transformation is composed of two major steps each of one being attached to a type of constraints.

First, the security constraints have to be strictly verified. For a given realization $\boldsymbol{x}$ of $\boldsymbol{X}$, the corresponding realization of random mass M being denoted by m and for a given rolling environment $\mathcal{T}$ (we remove $\boldsymbol{x}$ and $\mathcal{T}$ dependency in the notations of the next equations), if driver's command $\boldsymbol{u}$ does not respect the speed limitation constraint at the current time t_{i+1}, we directly modify the driver's command value at that current time. The modified driver's command corresponds to the one allowing for reaching the speed limitation $v^{max}\left(y(t_{i+1})\right)$ without overtaking it. We can find this value using Equation (II.16) as follows. For this realization $\boldsymbol{x}$ of $\boldsymbol{X}$, and for the traction case, the traction force is written, using the notation $y_i := y(t_i)$, $\dot{y}_i := \dot{y}(t_i)$, $\ddot{y}_i := \ddot{y}(t_i)$, and $u_i := u(t_i)$ as

$$f^T(\dot{y}_i, u_i) = \left(f^R\left(\dot{y}_i, v^w\left(y_i\right)\right) + f^C\left(y_i\right) - f^W\left(y_i\right) + mk^{rot}\,\frac{v^{max}\left(y_i\right) - \dot{y}_i}{\Delta t} \right). \tag{IV.30}$$

For the braking case, the braking force is written

$$f^B(\dot{y}_i, u_i) = \left(-f^R\left(\dot{y}_i, v^w\left(y_i\right)\right) - f^C\left(y_i\right) + f^W\left(y_i\right) - mk^{rot}\,\frac{v^{max}\left(y_i\right) - \dot{y}_i}{\Delta t} \right). \tag{IV.31}$$

The current acceleration $\ddot{y}_i$ is replaced by $\left(v^{max}\left(y_i\right) - \dot{y}_i\right)/\Delta t$. This adaptation assures that the speed $\dot{y}_{i+1}$ is equal to $v^{max}\left(y_i\right)$ (it this time, this constraint is saturated). For the traction and the braking cases, using Equations (II.13) and (II.14), we can estimate the driver's command as

$$u_i = \frac{f^T(\dot{y}_i, u_i)}{f^{T,max}\left(\dot{y}_i\right)} \quad , \quad u_i > 0\,, \tag{IV.32}$$

$$u_i = \frac{f^B(\dot{y}_i, u_i)}{f^{B,max}\left(\dot{y}_i\right)} \quad , \quad u_i < 0\,. \tag{IV.33}$$

Sometimes, it is necessary to modify the driver's command by taking into account several time steps t_{i+1}, t_{i+2}, etc. In this case, the algorithm sets the value of the driver's command to -1 at these time steps (equivalent to braking with maximum amplitude) until the train speed be below the speed limitation. Anyway, as braking is equivalent to a loss of energy, these candidates are not selected by the algorithm. The results of this modification on several driver's commands

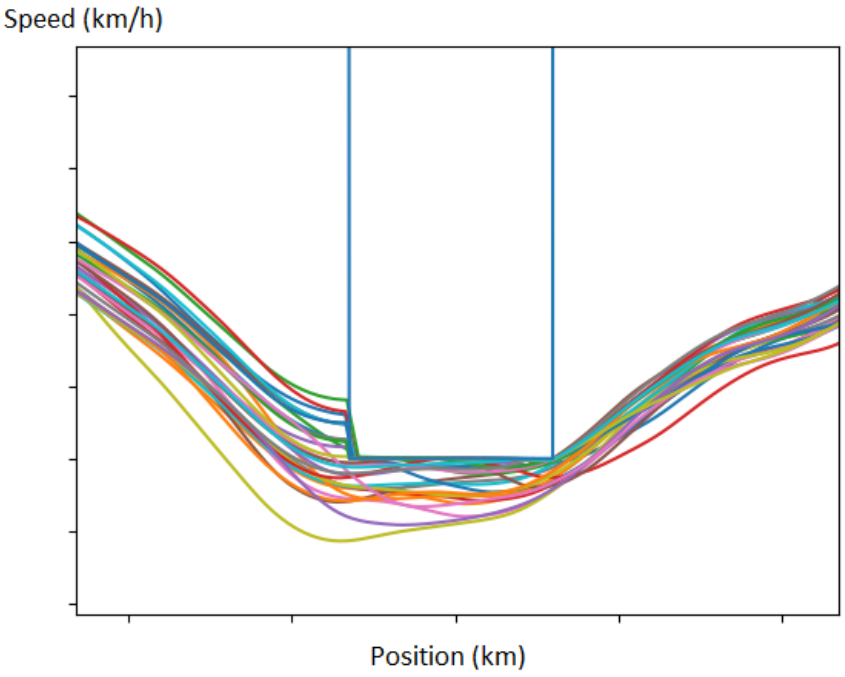

Figure IV.7: Train speed profile of several transformed driver's commands zoomed around a severe speed limitation area (in blue) depending on the position.

are shown in Figure IV.7.

We can see that this transformation exactly reacts as we have designed it. The train speed profiles verify the constraints, and a braking is instantaneously applied as soon as the speed limitation is exceeded. We could add a margin to the speed limitation to assure that the train will have the time to slow down, but once again these solutions should not be selected by the algorithm, as they are not efficient in term of energy saving.

Regarding the punctuality constraints, the second step of the transformation aims to modify very slightly the driver's command by multiply it by two constants. One constant c_T multiplies the driver's command on traction parts of the journey, and the second one c_B multiplies the driver's command on braking parts. To find these constants, we use the two graphs constructed and shown in Figure IV.8, which quantify the influence of c_T and c_B on the punctuality constraints. Note that c_B has only a small influence, as the journey is mainly constituted of traction phases. These graphs are derived from the initial driver's command $\underline{\boldsymbol{u}}$. Driver's command $\underline{\boldsymbol{u}}$ has to carefully be chosen, as the influence of c_T and c_B is dependent on $\underline{\boldsymbol{u}}$.

An example to illustrate this procedure is given below. We solve the dynamic equation for a candidate regularized driver's command and eventually modifying it if this candidate does not verify the security constraints. It turns out that this candidate does not verify the punctuality constraints. If the final position constraint is not verified, we use the value c_T given by the graph in left Figure IV.8, which imposes the train to arrive at the exact position. For instance, if the candidate driver's command makes the journey stopping $5\,km$ after the train station, we

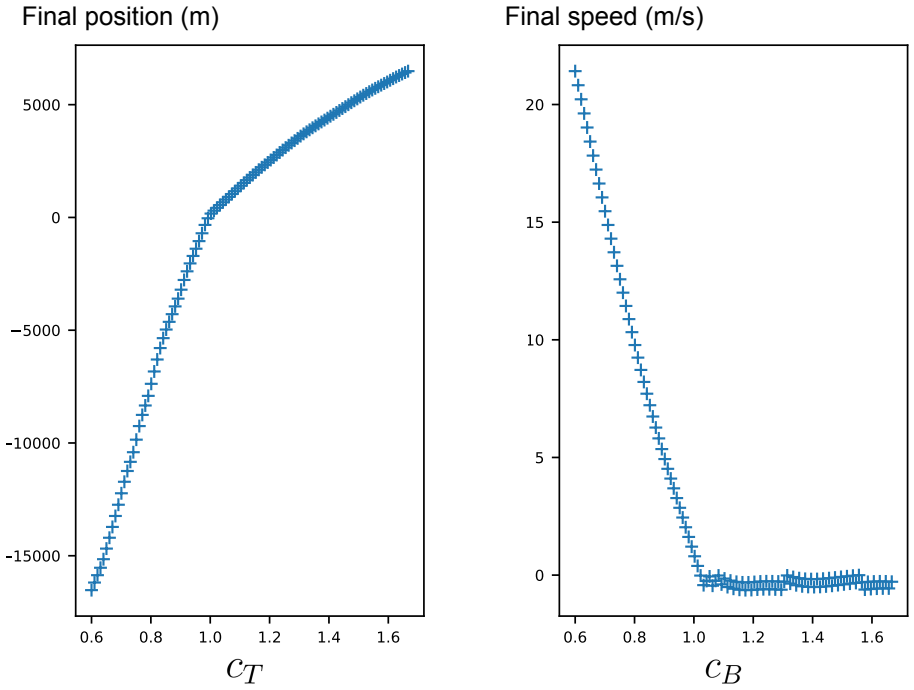

Figure IV.8: Graphs of the influence of the constants c_T and c_B on the final position constraint (left) and on the final speed constraint (right).

modify the driver's command so that the final position is $5\,km$ before. Thus, we can use the value $c_T = 0.85$ given by the graph of the left figure. This value of c_T is tested on the candidate regularized driver's command. The result might be slightly different from what is predicted by this graph, as it is calculated from $\underline{\boldsymbol{u}}$. If necessary, the procedure is iterated until the final position constraint be respected regarding the acceptable tolerance (in general after less than 5 iterations). Secondly, the final speed constraint is analyzed. If it is already verified, we do not modify constant c_B. If not, we increase its value until the constraint be verified. This could have an influence on the final position constraint, but in general, the modification of c_B only has a small influence on the train system, as it only perturbs the braking parts of the journey. This transformation, $\widetilde{\boldsymbol{T}}$, is written as

$$\widetilde{\boldsymbol{T}}\left(\boldsymbol{u}, \boldsymbol{x}\right) = c_T(\boldsymbol{x}) \max\left(\boldsymbol{u}, \boldsymbol{0}\right) - c_B(\boldsymbol{x}) \max\left(-\boldsymbol{u}, \boldsymbol{0}\right) \ , \tag{IV.34}$$

where $\boldsymbol{0}$ is the null vector and $\max$ is applied to each component of the vectors. To further simplify the notation, $\boldsymbol{T}$ is introduced to represent the complete transformation: discretization, regularization, and transformation to satisfy the constraints. The system scheme of this transformation $\boldsymbol{T}$ is shown in Figure IV.9. Note that due to the use of the PCA reduction, $\boldsymbol{u}$ is replaced by the reduced vector $\boldsymbol{q}^{(m)}$.

The probability density functions of $c_j\left(\underline{\boldsymbol{u}}, \mathcal{T}, \boldsymbol{X}\right)$ for $j = 1, 2, 3$ are plotted in Figure IV.10 for the transformed driver's command (the KDE method is used for the PDF estimations). It can be seen that the final position constraint is verified, as whatever the realization, the train arrives close to the station with a deviation of few meters that is lower than the acceptance tolerance of $5\,m$. In the same way, the final speed constraint is verified because it is lower close to 0, and it does not exceed the tolerance for all the tested realizations. The energy consumption remains highly dependent on $\boldsymbol{X}$.

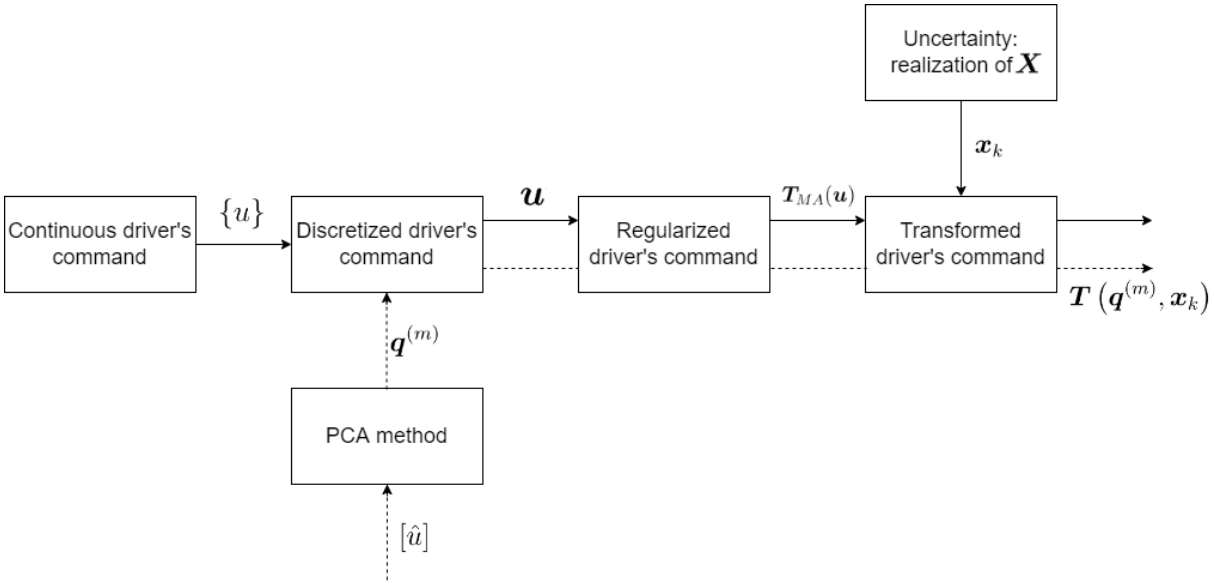

Figure IV.9: Scheme of the different steps of the transformation $\boldsymbol{T}$.

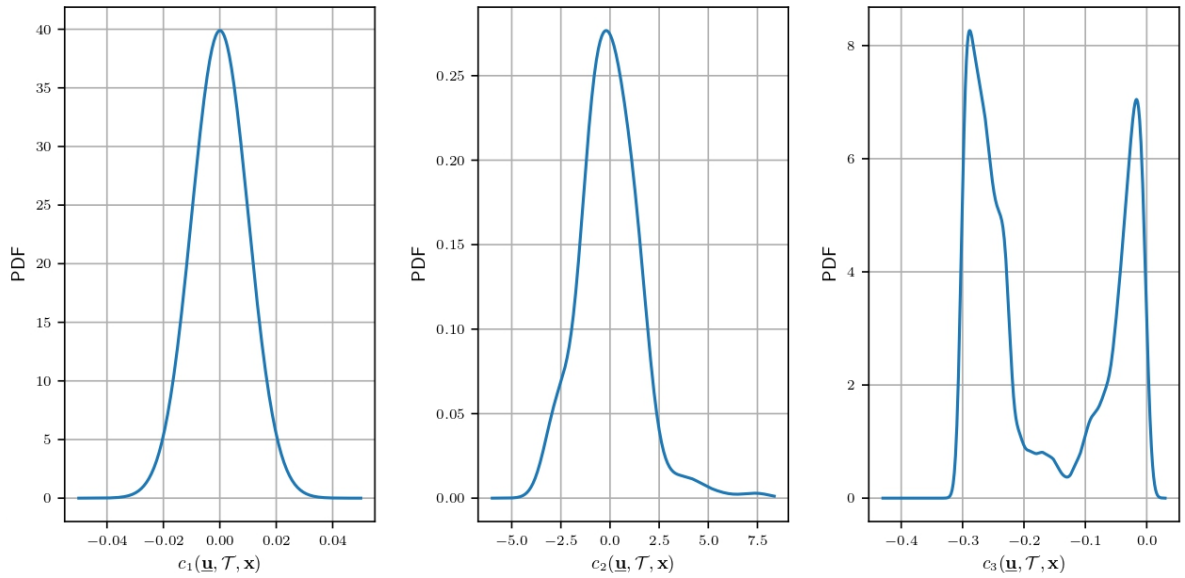

Figure IV.10: Probability density function of $c_j\left(\underline{\boldsymbol{u}}, \mathcal{T}, \boldsymbol{X}\right)$ for $j = 1$ (left), $j = 2$ (middle), and $j = 3$ (right) for the transformed driver's command. The probability density function of the energy consumption is represented on the left.

Finally, in order to assure that transformation $\boldsymbol{T}$ is well used by the algorithm, we have to get sure that it does not impact too much the exploration performed by the algorithm. Otherwise, the algorithm risks to be stuck in a specific subset in which all the candidate driver's commands are always transformed in a same driver's command. In this case, the algorithm cannot continue its exploration. An example of the regularized driver's command and its transformation is shown in Figure IV.11. We can see that both driver's commands are very close to each other (in fact superimposed).

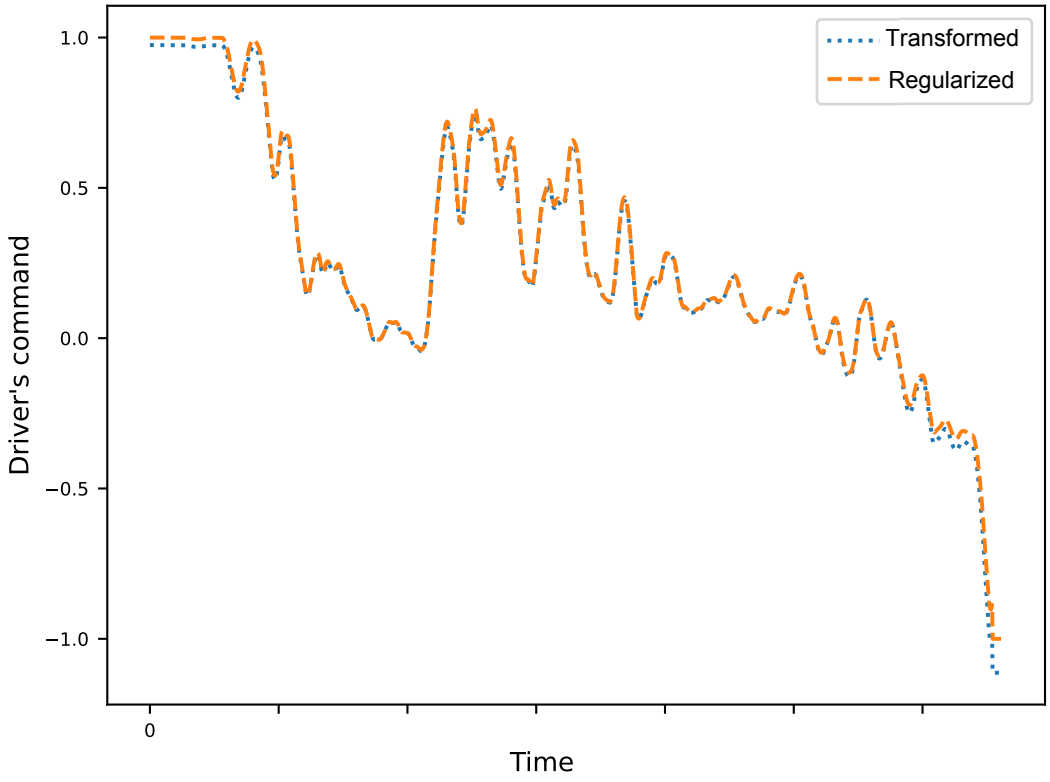

Figure IV.11: Regularized driver's command and its transformation depending on time, which are superimposed.

IV.2.2 Deterministic aspects of the objective function

Deterministic formulation of the objective function

The objective function describes the energy consumption of the train along its journey. For a given realization $\boldsymbol{x}$ of $\boldsymbol{X}$, Equation (II.22) describes the energy consumption of the train,

$$\mathcal{F}\left(\{u\}, \mathcal{T}, \boldsymbol{x}\right) = \int_{t_s}^{t_f} p^E\left(\{u_{\leq \tau}\}, \mathcal{T}, \boldsymbol{x}\right) d\tau \,. \tag{IV.35}$$

Similarly to the formulation performed for the constraints, transformation $\boldsymbol{T}$ is applied to the driver's command and Equation (IV.35) yields

$$\mathcal{F}\left(\boldsymbol{T}\left(\boldsymbol{q}^{(m)},\boldsymbol{x}\right),\mathcal{T},\boldsymbol{x}\right)\simeq\sum_{i=1}^{N}p^{E}\left(\boldsymbol{T}\left(\boldsymbol{q}^{(m)},\boldsymbol{x}\right)_{\leq t_i},\mathcal{T},\boldsymbol{x}\right)\Delta t\,. \tag{IV.36}$$

Adaptation of the deterministic formulation of the objective function to the probabilistic framework

The energy consumption of the train becomes a random variable due to the random vector $\boldsymbol{X}$. For defining the objective function, several choices are possible depending on the strategy we want to set up: do we want to control the extreme values (min or/and max) of the random energy consumption (so controlling the best or the worse), do we want to minimize the mean value of the random energy consumption, etc. We have chosen to construct the objective function as a mathematical expectation of the random energy consumption because it is assumed that its coefficient of variation is not too large, and the PDF is approximately symmetric with respect to its maximum. If we note $\mathcal{F}_{\mathbb{E}_{\boldsymbol{X}}}$ the expected value of the energy consumption of the train, we have

$$\mathcal{F}_{\mathbb{E}_{\boldsymbol{X}}}\left(\{u\},\mathcal{T}\right)=\int_{t_s}^{t_f}\mathbb{E}_{\boldsymbol{X}}\left[P^{E}\left(\{u_{\leq\tau}\},\boldsymbol{X},\mathcal{T}\right)\right]d\tau\,. \tag{IV.37}$$

As aforementioned, the driver's command is transformed by $\boldsymbol{T}$ and Equation (IV.37) becomes

$$\mathcal{F}_{\mathbb{E}_{\boldsymbol{X}}}\left(\boldsymbol{q}^{(m)},\mathcal{T}\right)=\sum_{i=1}^{N}\mathbb{E}_{\boldsymbol{X}}\left[P^{E}\left(\boldsymbol{T}\left(\boldsymbol{q}^{(m)},\boldsymbol{X}\right)_{\leq t_i},\boldsymbol{X},\mathcal{T}\right)\right]\Delta t\,. \tag{IV.38}$$

The stochastic solver used for estimating the objective function as a mathematical expectation of a random quantity is a Monte Carlo method generating random realizations of $\boldsymbol{X}$. Since the posterior probability distribution of $\boldsymbol{X}$ is not represented by a parameterized algebraic expression (See Chapter III), the realizations of $\boldsymbol{X}$ is performed using an MCMC as explained in Chapter III. For each candidate proposed by the CMA-ES algorithm, $N_{\boldsymbol{X}}=30$ new independent realizations of $\boldsymbol{X}$ are used. The objective function is written as

$$\mathcal{F}_{\mathbb{E}_{\boldsymbol{X}}}\left(\boldsymbol{q}^{(m)},\mathcal{T}\right)\simeq\frac{1}{N_{\boldsymbol{X}}}\sum_{k=1}^{N_{\boldsymbol{X}}}\sum_{i=1}^{N}\mathbb{E}_{\boldsymbol{X}}\left[P^{E,mod}\left(\boldsymbol{T}\left(\boldsymbol{q}^{(m)},\boldsymbol{X}\right)_{\leq t_i},\boldsymbol{X},\mathcal{T}\right)\right]\Delta t\,. \tag{IV.39}$$

IV.3 Optimization problem under uncertainty

The optimization problem under uncertainty is different from the one for the deterministic framework. We optimize vector $\boldsymbol{q}^{(m)}$ in order to minimize the objective function defined by Equation (IV.39) for a given rolling environment $\mathcal{T}$. The optimal value $\boldsymbol{q}^{*}\left(\mathcal{T}\right)$ of $\boldsymbol{q}^{(m)}$ is written as

$$\boldsymbol{q}^{*}\left(\mathcal{T}\right)=\underset{\boldsymbol{q}^{(m)}\in\mathcal{Q}}{\arg\min}\,\mathcal{F}_{\mathbb{E}_{\boldsymbol{X}}}\left(\boldsymbol{q}^{(m)},\mathcal{T}\right)\,, \tag{IV.40}$$

where $\mathcal{Q} \subset \mathbb{R}^{(m)}$ is the admissible set for $\boldsymbol{q}^{(m)}$ (see SectionIV.1.2) associated with the generalized coordinates. Note that the applied transformation allows the constraints to be verified, that is why the formulated optimization problem under uncertainties is unconstrained. The Lagrange multipliers are no longer necessary.

IV.4 Numerical solution of the optimization problem

Many methods can be used to solve this *a priori* nonconvex nonlinear optimization problem under uncertainties and under nonlinear probabilistic constraints. In this context of minimization of the energy consumption and without taking into account uncertainties, that is to say for the deterministic optimization problem, some optimization solvers has been used for many authors, the optimization problem being formulated with respect to the driver's command or the train speed profile. We give below a few representative methods used by these authors in this field.

- Pontryagin's maximum principle (L. Cesari [77]) is an analytic optimization method that has been used. It is well suited to find the optimal driver's command of a dynamical system under control or state constraints. I.A. Asnis *et al* proposed the use of this method in the railway field in [78] before being completed by E. Khmelnitsky in [79] or R.R. Liu *et al* in [29] between others. The limitation of the maximum principle method is its difficulty to find a solution in case of nonlinear functions used to describe the traction, braking, and the friction forces.

- The dynamic programming (R. Bellman *et al* [80]) was also widely used for solving this deterministic optimization problem. It relies on the discretization of the speed trajectory as a function of the curvilinear position and on the evaluation of the objective function on the lattice points to find the optimal trajectory for the train. For example, H. Ko *et al* [27] and V. Calderaro *et al* [81] applied this method to solve the optimization problem.

- Another method is the sequential quadratic programming (I.M. Nejdawi *et al* [82]). It is an iterative method, which searches for the optimal direction by looking in the neighborhood of a given point (changing at each iteration). For instance, this method was used by M. Miyatake *et al* [83] and [84].

- The mixed integer linear programming is a variant as it proposes to convert the problem in a complete linear optimization problem. An application for this railway optimization problem was proposed by Y. Wang *et al* [85].

- Evolutionary algorithms are recent alternatives to the previous methods. In particular, swarm particle optimization presented by J. Kennedy *et al* [86] (used by X.H. Zhao *et al* [87]), ant colony optimization proposed by M. Dorigo *et al* [88] (applied in railway field by B.R. Ke *et al* [89] and in [90]), genetic algorithm (employed in [91] by S. Lu *et al*). We were the first to use the Covariance Matrix Adaptation-Evolution Strategy (CMA-ES) [76] in [92]), which belongs to the class of the evolutionary algorithms, for solving this optimization problem in the deterministic framework (that we reuse for solving it in the probabilistic framework in the present work).

- The pseudo-spectral method (D. Dutykh [28]) was used in [93] but for a simplified formulation of the optimization problem.

For the optimization problem under uncertainties that we are considering, all these methods could have advantages and would have drawbacks. They would all be able to integrate the probabilistic nonlinear constraints. The optimization problem under consideration has specific characteristics, for which CMA-ES seems well adapted (and which will prove to be very effective). First, despite efforts to reduce the size of the optimization variable, the search space remains large (dimension will be 60). Second, the problem is strongly nonlinear and *a priori* nonconvex. Note that the use of the CMA-ES algorithm requires many evaluations of the physical model. Hopefully, one evaluation is relatively fast (10 seconds for a journey of 50 kilometers) and in addition, the CMA-ES algorithm can be parallelized.

The CMA-ES algorithm draws a small population of new search points function of an amplitude depending on the iteration and following a multivariate Gaussian distribution, for which its mean vector and its covariance matrix are also re-estimated at each iteration. At each point, the objective function is evaluated (we recall that the optimization problem under uncertainty and constraints has been reformulated in an unconstrained optimization problem). The method is detailed in Appendix B.

The CMA-ES algorithm uses hyperparameters that are related to the amplitude, the mean vector, and the covariance matrix of the multivariate Gaussian distribution. The initial value of the mean vector is chosen as $\underline{\boldsymbol{q}^{(m)}}$ associated with the driver's command $\underline{u}$. The initial value of the amplitude is chosen to 0.1, in order that the algorithm explores a relatively large range of possible solutions. The initial value of the covariance matrix is defined as the identity matrix, not to favor a specific direction. The size of the population of new search points is an important hyperparameter that has to carefully be defined. A large size allows for precisely exploring the neighborhood of the central point defined by the mean vector, but it also induces a more CPU time to evaluate all those points. A small size limits the exploration of the admissible set. In our case, we choose to keep the semi-empirical population size proposed by N. Hansen [76], which is equal to $4 + 3\log((m))$. For $(m) = 60$, the population size reaches $N^{pop} = 16$ points. As they are independently chosen, we can parallelize the evaluation of the cost function.

Finally, a termination criterion has to be defined. The maximum number of iterations could be defined but a good convergence would not be guaranteed. We choose to use two different termination criteria. The first one verifies if the updated amplitude of the new search steps is small (with respect to a given tolerance). In that case, we may consider that all the new search points drawn have similar quality (in terms of the optimization problem), and therefore the optimal solution is assumed to be obtained. The second one is related to the values of the cost function of the population of new search points. If they have approximately the same magnitude (with respect to a given tolerance), the optimal solution is also assumed to be obtained. The tolerances are defined by numerical experiments before running the algorithm. The algorithm is summarized below.

Initialization:

- $i = 0$
- Initiate $\boldsymbol{q}^0$ by $\underline{\boldsymbol{q}^{(m)}}$
- Choose a population size, N^{pop}
- Define the termination criteria

while the termination criteria are not overtaken **do**

- $i = i + 1$
- Draw a population $\boldsymbol{q}^i$ of $\boldsymbol{q}^{(m)}$ with an amplitude σ^{i-1} and from a multivariate Gaussian distribution with mean vector $\boldsymbol{\nu}^{i-1}$ and covariance matrix $[C^{i-1}]$

for $\underline{j = 1 : N^{pop}}$ **do**

- Draw $N_{\boldsymbol{X}} = 30$ independent realizations $\boldsymbol{x}_k$ of $\boldsymbol{X}$
- Estimate the energy consumption for realization $\boldsymbol{x}_k$

end

- Estimate the cost function from Equation (IV.39),
- $\sigma^{i-1} \leftarrow \sigma^i$
- $\boldsymbol{\nu}^{i-1} \leftarrow \boldsymbol{\nu}^i$
- $[C^{i-1}] \leftarrow [C^i]$

end
Return $\boldsymbol{q}^*(\mathcal{T})$.

Algorithm 4: CMA-ES algorithm adapted to optimization under uncertainties.

By starting the iterations close to the optimal solutions found in Section IV.1.2 (family of solutions), we also reduce the number of iterations necessary to reach the convergence. If solving the dynamic equation only takes few seconds, solving it for each of the $N_{\boldsymbol{X}} = 30$ realizations, for all the $N^{pop} = 16$ points, and for all the CMA-ES iterations might be significant. For $5\,000$ iterations to converge, dynamic equations are solved $2\,400\,000$ times. With a 30 cores parallelization and 4 seconds for solving one dynamic equation, the elapsed time is about 3.5 days.

IV.5 Optimal solution under uncertainty

When the CMA-ES algorithm has finished, that is to say when one of the termination criteria is reached, a computed optimal solution under uncertainty is obtained. First of all, we plot the values of the cost function depending on the iteration number (see Figure IV.12). This gives us information on the convergence speed of the algorithm.

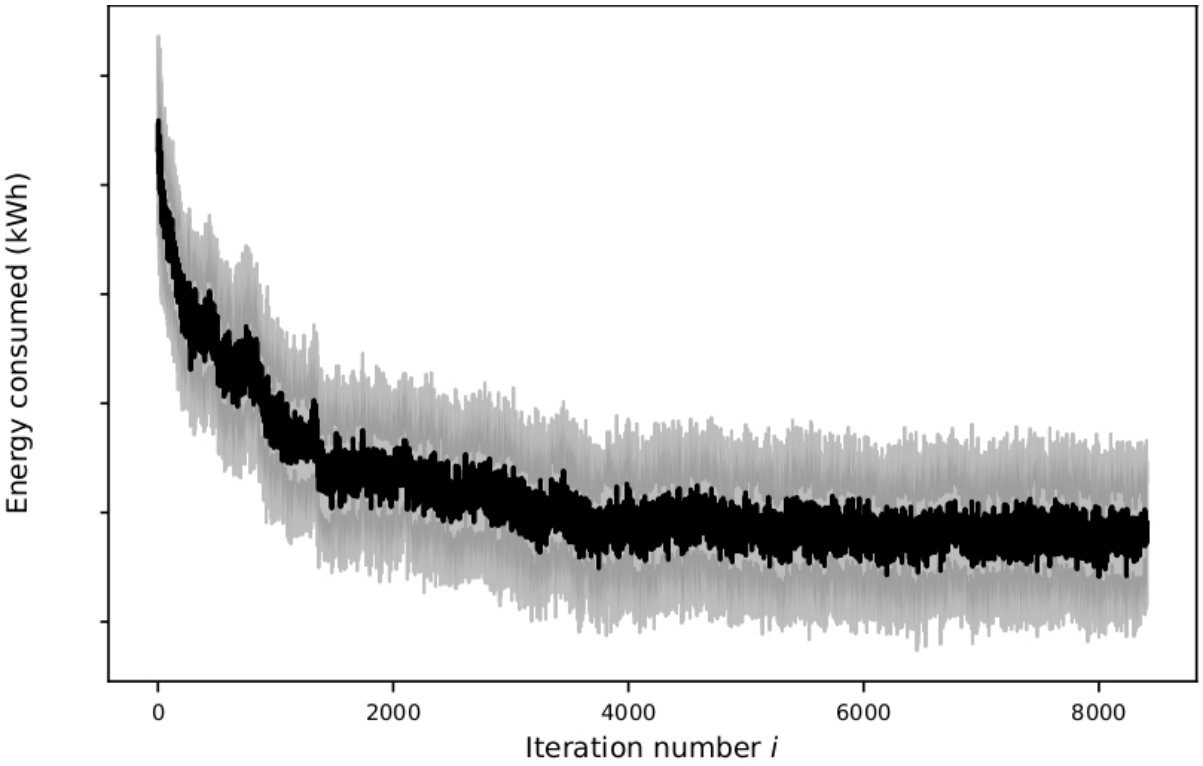

Figure IV.12: Convergence of the cost function evaluated for $\boldsymbol{q}^i$ as a function of iteration number i: mean value over the population (black solid line), standard deviation envelop of the population (grey envelop).

After one thousand iterations, the mean value of the cost function evaluated for $\boldsymbol{q}^i$ decreases rapidly with i. The convergence is stabilized around a value associated with the optimal solution. Note that the width on the region defined by the upper and lower envelop would decrease if $N_{\boldsymbol{X}}$ was increasing, but in increasing the elapsed time. Consequently, we consider that the calculation has converged for $i = 4\,000$, and we propose that the mean of the populations of the last 500 iterations correspond to the optimal value.

This optimal driver's command can be considered as the one computed offline before the train departure. Now, we could imagine analyzing the impact of this optimal driver's command $\boldsymbol{q}^*$ under uncertainty by plotting the confidence regions of the random speed profile and the random energy consumption. However, as explained in Section IV.2.1 the constraints would not be verified. The analysis of the impact of the use of the optimal driver's command is then performed as follows. The optimal driver's command can be adapted to the realization $\boldsymbol{x}_k$ thanks to the use of transformation $\boldsymbol{T}$ (in order to verify the constraints). In practice,

we can assume an online updating of the model parameters along the way. Consequently, the transformed driver's command can be replaced by its online updating $\boldsymbol{T}(\boldsymbol{q}^*, \boldsymbol{x}_k)$ and the constraints remain verified.

Using this online updating, the optimal driver's command $\boldsymbol{q}^*$ is used to compute the optimal speed profile from the dynamic equation and the associated energy consumption. For the optimal driver's command $\boldsymbol{u}^*$, Figure IV.13 displays the quantile of the random speed profile, the speed limitation, and the quantile of the random energy consumption. This optimal solution is compared to the measurements. It can be seen that the energy consumption computed with the online updated optimal driver's command is smaller than the energy consumption associated with the driver's command that was used by the measured train. Thus, it seems that the online updated robust optimization allows for obtaining a significant gain for the analyzed journey under consideration, for which measurements were available.

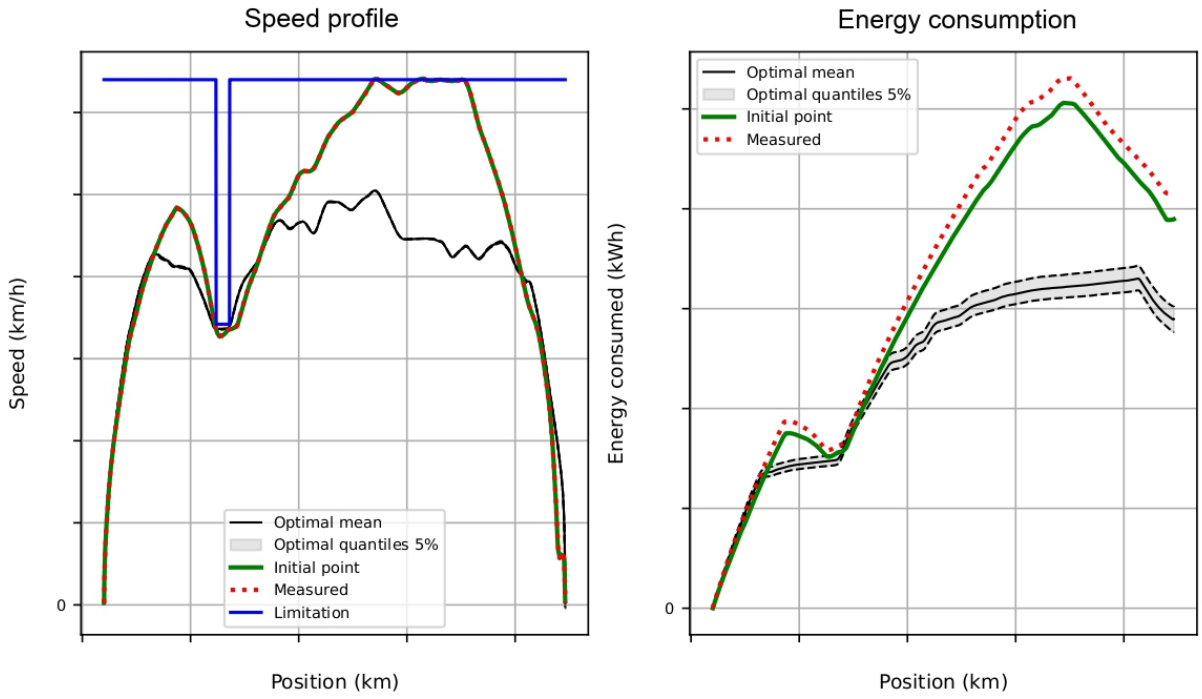

Figure IV.13: Quantile of the random speed profile (left) and quantile of the random energy consumption (right) associated with the online updated optimal driver's command $\boldsymbol{u}^*$. Their mean values are in black, and the grey envelopes stand for the quantile intervals. The initial driver's command $\underline{\boldsymbol{u}}$ is in green. The measurements are plotted in red, and the speed limitation is in blue.

The variability introduced by random vector $\boldsymbol{X}$ has an impact on the transformed optimal driver's command $\boldsymbol{T}(\boldsymbol{q}^*, \boldsymbol{X})$, and thus on the variability of physical quantities (see Figure IV.13). This variability is smaller on the train speed than on the energy consumption.

The most surprising is that in spite of the variability of the train configurations, the optimal speed profile has a variance very close to zero (few km/h), the optimal speed almost seems to be deterministic. An explanation of this small value of the variance is that the transformed

driver's command is adapted to the realization $\boldsymbol{x}_k$ of $\boldsymbol{X}$. Whatever the configuration $\boldsymbol{x}_k$ is drawn, the driver's command is transformed so that the speed of the train remains approximately constant.

We can also observe that the optimal speed profile is very close to the one presented in Figure IV.3, which is logical as they come from two different formulations of optimization problems that come from the same case study. The consumed energy is also very similar, but it appears to be slightly higher for the second approach (Figure IV.13). Indeed, this approach searches for a general solution easy to adapt to the configuration $\boldsymbol{x}_k$, contrary to the first approach, which looks for the optimal solution for random vector $\boldsymbol{X}$. For this reason, the first approach is slightly more efficient in terms of energy saving, but as we explained before, it is more difficult to operate by the drivers as they cannot know which configuration should be considered.

One of the objectives of this work is to help drivers having an optimal behavior in term of energy saving. In that case, providing them the optimal speed profile is interesting as they could adapt the driver's command to the configuration in order to follow the same speed profile, just like we proposed to do in the second approach. The associated energy consumption will vary, as shown in Figure IV.13 but should be minimized in average.

IV.6 Conclusion on the optimization method

To put it in a nutshell, thanks to the dynamic and energy consumption models presented in Chapter II, and to the identification of the parameters performed in Chapter III, we have been able to propose a robust strategy to solve the industrial objective, that is to say to minimize the energy consumed by a train during its journey.

Due to the variability existing between trains, we had to solve a constrained optimization problem under uncertainty. The solution of this problem is not easy, and a first step has consisted in solving a deterministic problem associated with the given train configuration $\boldsymbol{x}_k$. But it should be noted by the family of optimal solutions $\boldsymbol{u}_k^*$ obtained from these deterministic problems are not easy to operate for the drivers, as the drivers cannot know which solution should be used.

For this reason, we have proposed a second step in the methodology, for which the optimization problem is modified. It consists in searching the deterministic optimal driver's command, whose transformation (associated with each realization $\boldsymbol{x}_k$) minimizes in mean the energy consumption of the train. Several tools have been proposed to reduce the dimension of the problem, to take into account uncertainty, and to solve it numerically. The optimal solution shows interesting improvements in terms of energy saving and usability by the drivers.

The transformed driver's command has been replaced by its online updating $\boldsymbol{T}(\boldsymbol{q}^*, \boldsymbol{x}_k)$, which allows the constraints to remain verified.

Chapter V

Conclusion and perspectives

V.1 Conclusion

In this thesis, we have explored several domains linked to the railway dynamics, the stochastic identification, and the driver's command optimization under uncertainty. These topics are often encountered in engineering sciences, as they are closely related to industrial stakes. For reducing the energy consumed by high-speed trains, a methodology has been proposed in this work. Each chapter is devoted to a main part of the development.

The first part (Chapter II) focuses on the construction of a dynamic model and of an energy consumption model. The 3D train dynamics based on a 3D rigid body formulation for the whole train, is projected on the longitudinal track axis in order to obtain the longitudinal dynamic model. The position, speed, acceleration of the train, as well as the traction and braking forces, are estimated solving the dynamic equation. Then, the energy consumption is simulated from the time integral of the electric power, which is itself computed from the train speed, the traction forces, the braking forces, the auxiliary power, and the energy efficiency. The dynamic model and the energy consumption model quantify all the physical quantities related to the optimization problem. Afterwards, the sensitivity analysis of the model parameters has been carried out.

The second part (Chapter III) aims to identify the model parameters taking into account uncertainty. A probabilistic framework has been retained for modeling the uncertainties, which complicates the identification of parameters that are random variables. For this reason, we have proposed two different approaches. The first one only relays on the identification of the mean value of the uncertain parameters. The driver's command is then estimated in inverse from the distance between the simulated speed and the measured speed. Finally, the distance between the measured energy consumption and the simulated one is quantified using the Root Mean Square Error. The mean values of the uncertain parameters are identified as the minimum of this distance. The second ingredient consists in the using the Bayesian inference method to identify the posterior distribution of each parameter. After collecting the available information for constructing the prior distributions, as well as after defining the likelihood function, we have applied the Metropolis-within-Gibbs algorithm. The identified results of both approaches have been validated using data measurements that have not been used for the identification.

The last part (Chapter IV) details the optimization problem, for which we have proposed an innovative formulation and solver. After optimizing several deterministic problems associated with several realizations of the posterior distributions (that correspond to several configurations of the train), we have proposed a robust strategy with respect to uncertainties. The deterministic optimization problems have been replaced by an optimization problem under uncertainty: find the optimal driver's command, which is easy to transform in a new driver's command that verifies the constraints, in order to minimize in mean the energy consumed by all the possible configurations of trains. We then obtain an unconstrained optimization problem under uncertainty. Attention has been paid to develop a methodology that is adapted to the high dimension of the problem and the use of a probabilistic formulation. More specifically, the discretization of the driver's command and the introduction of a PCA of the driver's command have allowed for drastically reducing the dimension of the optimization problem without altering the quality of the prevision. The optimization algorithm has carefully been developed. The online updated optimal driver's command has been compared to the consumption of commercial trains for which measurements were available.

All the developments that have been presented in this work have been implemented in a Software written in Python.

V.2 Perspectives

The constructed methodology proposed for calculating an online-updated optimal driver's command can directly be used by the drivers for minimizing the energy consumption while verifying the constraints (speed limitation, final position, final speed, and final time). However, such an approach requires to get, in quasi real-time, information from the train and its environment along the way. The experimental test on a real train of the developed algorithm is in progress. Some details on that are given in Appendix C.

This work presents the advantage of being adaptable to many real cases and the software developed is flexible. Any kind of trains and any rolling environments can be considered by the method. The only requirements are the knowledge of the wind and track descriptions, and some measurements are necessary to identify the model parameters. The constraints can also be modified to add a starting delay, a new speed limitation in case of a weaken part of the track, or specific crossing points to respect the train timetabling.

Further directions might be explored to extend this work. The first one would consist in a direct continuation of the presented method. The numerical cost generated by solving the dynamic equation could be reduced using a parameterized surrogate model, which is not so easy taking into account the very high dimension of the dynamic model. Such a surrogate model is certainly necessary for considering online updating along the way.

V.3 Publications and communications

Referred journal publications

- Julien Nespoulous, Christian Soize, Christine Funfschilling, Guillaume Perrin. Optimisation of train speed to limit energy consumption. *Vehicle System Dynamics*, (60) 10, 3540-3557,2022.

- Julien Nespoulous, Christian Soize, Christine Funfschilling, Guillaume Perrin. Bayesian inference for high-speed train dynamics and optimization under uncertainty for energy saving using computational stochastic nonlinear dynamics and statistics. *Journal of Computational Physics*, Elsevier, submitted.

Conferences with proceeding

- Julien Nespoulous, Christian Soize, Christine Funfschilling, Guillaume Perrin. Driver's control optimization under uncertainties to reduce energy consumption of high-speed trains. RAILWAYS 2022, 5^{th} International Conference on Railway Technology. 22 - 25 August 2022, Montpellier, France.

- Julien Nespoulous, Christian Soize, Christine Funfschilling, Guillaume Perrin. Bayesian inference for high-speed train dynamics and speed optimization under uncertainty for energy saving. USD 2022, 9^{th} International Conference on Uncertainty in Structural Dynamics. 12 - 14 September 2022, Leuven, Belgium.

Conferences with abstract

- Julien Nespoulous, Christian Soize, Christine Funfschilling, Guillaume Perrin. High-speed train speed optimization for limiting energy consumption. IAVSD 2021, 27^{th} IAVSD Symposium on the Dynamics of Vehicles on Roads and on Tracks. 17 - 19 August 2021, St. Petersburg, Russia.

- Julien Nespoulous, Christian Soize, Christine Funfschilling, Guillaume Perrin. Uncertainty quantification for high-speed train dynamics modeling and optimization under uncertainties to limit energy consumption. WCCM 2022, 15^{th} World Congress on Computational Mechanics & 8^{th} Asian Pacific Congress on Computational Mechanics. 31 July - 5 August 2022, Yokohama, Japan.

Conferences with poster

- Julien Nespoulous, Christian Soize, Christine Funfschilling, Guillaume Perrin. Optimization under uncertainties of high-speed train speed to limit energy consumption. UNCECOMP 2021, 4^{th} International Conference on Uncertainty Quantification in Computational Sciences and Engineering. 28 - 30 June 2021, Athens, Greece.

- Julien Nespoulous, Christian Soize, Christine Funfschilling, Guillaume Perrin. Robust adaptation of the train speed for energy saving under punctuality and security constraints. MASCOT-NUM 2022. 7 - 9 June 2022, Clermont-Ferrand, France.

Bibliography

[1] Patrice Aknin and Hugues Chollet. A new approach for the modelling of track geometry recording vehicles and the deconvolution of versine measurements. In The Dynamics of Vehicles on Roads and on Tracks, pages 59–70. CRC Press, 2021.

[2] Hui Liu. Wind Forecasting in Railway Engineering. Elsevier, 2021.

[3] Pedro Urda, Javier F Aceituno, Sergio Muñoz, and José L Escalona. Artificial neural networks applied to the measurement of lateral wheel-rail contact force: A comparison with a harmonic cancellation method. Mechanism and Machine Theory, 153:103968, 2020.

[4] Vincenzo Niola, Salvatore Strano, and Mario Terzo. A random walk model approach for the wheel-rail contact force estimation. Journal of Dynamic Systems, Measurement, and Control, 140(7):071016, 2018.

[5] Ping Li, Roger Goodall, Paul Weston, Chung Seng Ling, Colin Goodman, and Clive Roberts. Estimation of railway vehicle suspension parameters for condition monitoring. Control Engineering Practice, 15(1):43–55, 2007.

[6] Jens Wittenburg. Dynamics of Systems of Rigid Bodies, volume 33. Springer-Verlag, 2013.

[7] Colin Cole et al. Longitudinal train dynamics. Handbook of Railway Vehicle Dynamics, pages 239–278, 2006.

[8] Roger G Ghanem and Pol D Spanos. Stochastic Finite Elements: a Spectral Approach. Springer-Verlag, 1991.

[9] Christian Soize and Roger Ghanem. Physical systems with random uncertainties: chaos representations with arbitrary probability measure. SIAM Journal on Scientific Computing, 26(2):395–410, 2004.

[10] Gerhart I Schuëller and Hector A Jensen. Computational methods in optimization considering uncertainties–an overview. Computer Methods in Applied Mechanics and Engineering, 198(1):2–13, 2008.

[11] Olivier Le Maître and Omar M Knio. Spectral Methods for Uncertainty Quantification: with Applications to Computational Fluid Dynamics. Springer Science & Business Media, 2010.

[12] Roger Ghanem, David Higdon, Houman Owhadi, et al. Handbook of Uncertainty Quantification, volume 6. Springer, 2017.

[13] Christian Soize. Uncertainty Quantification. Springer, 2017.

[14] F Alkam, I Pereira, and T Lahmer. Qualitatively-improved identified parameters of pre-stressed concrete catenary poles using sensitivity-based bayesian approach. Results in Engineering, 6:100104, 2020.

[15] David Lebel, Christian Soize, Christine Funfschilling, and Guillaume Perrin. Statistical inverse identification for nonlinear train dynamics using a surrogate model in a bayesian framework. Journal of Sound and Vibration, 458:158–176, 2019.

[16] Amir Dib, Charles Truong, Laurent Oudre, Mathilde Mougeot, Nicolas Vayatis, and Heloïse Nonne. Bayesian feature discovery for predictive maintenance. In 2021 29th European Signal Processing Conference (EUSIPCO), pages 1421–1425. IEEE, 2021.

[17] Kunihiko Ichikawa. Application of optimization theory for bounded state variable problems to the operation of train. Bulletin of JSME, 11(47):857–865, 1968.

[18] Pengling Wang and Rob MP Goverde. Two-train trajectory optimization with a green-wave policy. Transportation Research Record, 2546(1):112–120, 2016.

[19] Pengling Wang, Alessio Trivella, Rob MP Goverde, and Francesco Corman. Train trajectory optimization for improved on-time arrival under parametric uncertainty. Transportation Research Part C: Emerging Technologies, 119:102680, 2020.

[20] Li Wang, Lixing Yang, Ziyou Gao, and Yeran Huang. Robust train speed trajectory optimization: A stochastic constrained shortest path approach. Frontiers of Engineering Management, 4(4):408–417, 2017.

[21] Pengling Wang, Alessio Trivella, Rob MP Goverde, and Francesco Corman. Train trajectory optimization for improved on-time arrival under parametric uncertainty. Transportation Research Part C: Emerging Technologies, 119:102680, 2020.

[22] Wei ShangGuan, Xi-Hui Yan, Bai-Gen Cai, and Jian Wang. Multiobjective optimization for train speed trajectory in ctcs high-speed railway with hybrid evolutionary algorithm. IEEE Transactions on Intelligent Transportation Systems, 16(4):2215–2225, 2015.

[23] Cheng-Shang Chang, DY Xu, and HB Quek. Pareto-optimal set based multiobjective tuning of fuzzy automatic train operation for mass transit system. IEE Proceedings-Electric Power Applications, 146(5):577–583, 1999.

[24] Rémy Chevrier, Paola Pellegrini, and Joaquin Rodriguez. Energy saving in railway timetabling: A bi-objective evolutionary approach for computing alternative running times. Transportation Research Part C: Emerging Technologies, 37:20–41, 2013.

[25] Pengling Wang and Rob MP Goverde. Multi-train trajectory optimization for energy efficiency and delay recovery on single-track railway lines. Transportation Research Part B: Methodological, 105:340–361, 2017.

[26] Zbigniew Michalewicz, Dipankar Dasgupta, Rodolphe G Le Riche, and Marc Schoenauer. Evolutionary algorithms for constrained engineering problems. Computers & Industrial Engineering, 30(4):851–870, 1996.

[27] Hideyoshi Ko, T Koseki, and Masafumi Miyatake. Application of dynamic programming to the optimization of the running profile of a train. WIT Transactions on The Built Environment, 74, 2004.

[28] Denys Dutykh. A brief introduction to pseudo-spectral methods: application to diffusion problems. arXiv preprint arXiv:1606.05432, 2016.

[29] Rongfang Rachel Liu and Iakov M Golovitcher. Energy-efficient operation of rail vehicles. Transportation Research Part A: Policy and Practice, 37(10):917–932, 2003.

[30] Donald R Jones, Matthias Schonlau, and William J Welch. Efficient global optimization of expensive black-box functions. Journal of Global Optimization, 13(4):455–492, 1998.

[31] James C Spall. Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control. John Wiley & Sons, 2005.

[32] Kurt Marti et al. Stochastic Optimization Methods, volume 2. Springer, 2005.

[33] Jack PC Kleijnen, Wim Van Beers, and Inneke Van Nieuwenhuyse. Constrained optimization in expensive simulation: Novel approach. European Journal of Operational Research, 202(1):164–174, 2010.

[34] Jing Xie, Peter I Frazier, and Stephen E Chick. Bayesian optimization via simulation with pairwise sampling and correlated prior beliefs. Operations Research, 64(2):542–559, 2016.

[35] Roger Ghanem and Christian Soize. Probabilistic nonconvex constrained optimization with fixed number of function evaluations. International Journal for Numerical Methods in Engineering, 113(4):719–741, 2018.

[36] Pietro Marco Congedo, C Corre, and J-M Martinez. Shape optimization of an airfoil in a BZT flow with multiple-source uncertainties. Computer Methods in Applied Mechanics and Engineering, 200(1-4):216–232, 2011.

[37] Wen Yao, Xiaoqian Chen, Wencai Luo, Michel Van Tooren, and Jian Guo. Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles. Progress in Aerospace Sciences, 47(6):450–479, 2011.

[38] R. H. Fries and B. M. Coffey. A State-Space Approach to the Synthesis of Random Vertical and Crosslevel Rail Irregularities. Journal of Dynamic Systems, Measurement, and Control, 112(1):83–87, 03 1990.

[39] Guillaume Perrin, Christian Soize, Denis Duhamel, and Christine Funfschilling. Track irregularities stochastic modeling. Probabilistic Engineering Mechanics, 34:123–130, 2013.

[40] Guillaume Perrin, Denis Duhamel, Christian Soize, and Christine Funfschilling. Quantification of the influence of the track geometry variability on the train dynamics. Mechanical Systems and Signal Processing, 60:945–957, 2015.

[41] Nicolas Lestoille, Christian Soize, and Christine Funfschilling. Sensitivity of train stochastic dynamics to long-term evolution of track irregularities. Vehicle System Dynamics, 54(5):545–567, 2016.

[42] Nicolas Lestoille, Christian Soize, and Christine Funfschilling. Stochastic prediction of high-speed train dynamics to long-term evolution of track irregularities. Mechanics Research Communications, 75:29–39, 2016.

[43] Theodore De Karman and Leslie Howarth. On the statistical theory of isotropic turbulence. Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences, 164(917):192–215, 1938.

[44] Romain Bosquet. Modélisation énergétique et identification des trains pour l'écoconception des lignes ferroviaires à grande vitesse (in French). PhD thesis, Univesrsité de Nates, 2015.

[45] Vijay Garg. Dynamics of Railway Vehicle Systems. Elsevier, 2012.

[46] Nicola Bosso and Nicolò Zampieri. Long train simulation using a multibody code. Vehicle System Dynamics, 55(4):552–570, 2017.

[47] Qing Wu, Maksym Spiryagin, and Colin Cole. Longitudinal train dynamics: an overview. Vehicle System Dynamics, 54(12):1688–1714, 2016.

[48] Xavier Quost. Modélisation de l'effet du vent sur les trains à grande vitesse.: Une étude dynamique et stochastique appliquée au risque de renversement (in French). PhD thesis, Ecully, Ecole centrale de Lyon, 2005.

[49] Heinrich Hertz. On the contact of elastic solids. Z. Reine Angew. Mathematik, 92:156–171, 1881.

[50] Joost J Kalker. Three-Dimensional Elastic Bodies in Rolling Contact, volume 2. Springer Science & Business Media, 2013.

[51] Bérénice Pecile. Modèle dynamique d'interaction véhicule-voie ferroviaire en présence de défauts géométriques sur les surfaces en contact (in French). PhD thesis, Université de Valenciennes et du Hainaut-Cambresis, 2017.

[52] William J Davis. The Tractive Resistance of Electric Locomotives and Cars, volume 29. General Electric, 1926.

[53] Giorgio Boschetti and Andrea Mariscotti. The parameters of motion mechanical equations as a source of uncertainty for traction systems simulation. In XX IMEKO World Congress, 2012.

[54] Marc Debruyne. La chaîne de traction (in french). Technical report, Estaca - Matériel roulant Propulsion et Guidage, 2011.

[55] Sönke Kraft. Parameter identification for a TGV model. PhD thesis, Châtenay-Malabry, Ecole centrale de Paris, 2012.

[56] David Lebel, Christian Soize, Christine Funfschilling, and Guillaume Perrin. High-speed train suspension health monitoring using computational dynamics and acceleration measurements. Vehicle System Dynamics, 58(6):911–932, 2020.

[57] Richard P. Brent. An algorithm with guaranteed convergence for finding a zero of a function. The Computer Journal, 14(4):422–425, 1971.

[58] Theodorus Jozef Dekker. Finding a zero by means of successive linear interpolation. Constructive Aspects of the Fundamental Theorem of Algebra, pages 37–51, 1969.

[59] Geof H Givens and Jennifer A Hoeting. Computational Statistics, volume 703. John Wiley & Sons, 2012.

[60] Bradley P Carlin and Thomas A Louis. Bayesian Methods for Data Analysis. CRC Press, 2008.

[61] Jari Kaipio and Erkki Somersalo. Statistical and Computational Inverse Problems, volume 160. Springer Science & Business Media, 2006.

[62] Peter Congdon. Bayesian Statistical Modelling. John Wiley & Sons, 2007.

[63] Guillaume Perrin and Christian Soize. Adaptive method for indirect identification of the statistical properties of random fields in a bayesian framework. Computational Statistics, 35(1):111–133, 2020.

[64] Christian Soize, Roger G Ghanem, and Christophe Desceliers. Sampling of bayesian posteriors with a non-gaussian probabilistic learning on manifolds from a small dataset. Statistics and Computing, 30(5):1433–1457, 2020.

[65] Robert M Gray. Entropy and Information Theory. Springer Science & Business Media, 2011.

[66] Jagat Narain Kapur and Hiremaglur K Kesavan. Entropy optimization principles and their applications. In Entropy and Energy Dissipation in Water Resources, pages 3–20. Springer, 1992.

[67] Christian Soize. Maximum entropy approach for modeling random uncertainties in transient elastodynamics. The Journal of the Acoustical Society of America, 109(5):1979–1996, 2001.

[68] Christian Soize. Construction of probability distributions in high dimension using the maximum entropy principle: Applications to stochastic processes, random fields and random matrices. International Journal for Numerical Methods in Engineering, 76(10):1583–1611, 2008.

[69] Murray Rosenblatt. Remarks on some nonparametric estimates of a density function. The Annals of Mathematical Statistics, pages 832–837, 1956.

[70] Emanuel Parzen. On estimation of a probability density function and mode. The Annals of Mathematical Statistics, 33(3):1065–1076, 1962.

[71] David J Wales and Jonathan PK Doye. Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. The Journal of Physical Chemistry A, 101(28):5111–5116, 1997.

[72] Nicholas Metropolis and Stanislaw Ulam. The monte carlo method. Journal of the American Statistical Association, 44(247):335–341, 1949.

[73] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. The Journal of Chemical Physics, 21(6):1087–1092, 1953.

[74] Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-6(6):721–741, 1984.

[75] Don Van Ravenzwaaij, Pete Cassey, and Scott D Brown. A simple introduction to markov chain monte–carlo sampling. Psychonomic Bulletin & Review, 25(1):143–154, 2018.

[76] Nikolaus Hansen. The CMA evolution strategy: A tutorial. arXiv preprint arXiv:1604.00772, 2016.

[77] Lamberto Cesari. Optimization—Theory and Applications: Problems with Ordinary Differential Equations, volume 17. Springer Science & Business Media, 2012.

[78] IA Asnis, AV Dmitruk, and NP Osmolovskii. Solution of the problem of the energetically optimal control of the motion of a train by the maximum principle. USSR Computational Mathematics and Mathematical Physics, 25(6):37–44, 1985.

[79] Eugene Khmelnitsky. On an optimal control problem of train operation. IEEE Transactions on Automatic Control, 45(7):1257–1266, 2000.

[80] Richard Bellman and Robert E Kalaba. Dynamic Programming and Modern Control Theory, volume 81. Citeseer, 1965.

[81] Vito Calderaro, Vincenzo Galdi, Giuseppe Graber, Antonio Piccolo, and Dario Cogliano. An algorithm to optimize speed profiles of the metro vehicles for minimizing energy consumption. In 2014 International Symposium on Power Electronics, Electrical Drives, Automation and Motion, pages 813–819. IEEE, 2014.

[82] Imad M Nejdawi, Kevin A Clements, and Paul W Davis. An efficient interior point method for sequential quadratic programming based optimal power flow. IEEE Transactions on Power Systems, 15(4):1179–1183, 2000.

[83] Masafumi Miyatake and Kunihiko Matsuda. Energy saving speed and charge/discharge control of a railway vehicle with on-board energy storage by means of an optimization model. IEEJ Transactions on Electrical and Electronic Engineering, 4(6):771–778, 2009.

[84] Masafumi Miyatake and Hideyoshi Ko. Optimization of train speed profile for minimum energy consumption. IEEJ Transactions on Electrical and Electronic Engineering, 5(3):263–269, 2010.

[85] Yihui Wang, Bart De Schutter, Bin Ning, Noortje Groot, and Ton JJ Van Den Boom. Optimal trajectory planning for trains using mixed integer linear programming. In 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), pages 1598–1604. IEEE, 2011.

[86] James Kennedy and Russell Eberhart. Particle swarm optimization. In Proceedings of ICNN'95-International Conference on Neural Networks, volume 4, pages 1942–1948. IEEE, 1995.

[87] Xin-Hong Zhao, Bwo-Ren Ke, and Kuo-Lung Lian. Optimization of train speed curve for energy saving using efficient and accurate electric traction models on the mass rapid transit system. IEEE Transactions on Transportation Electrification, 4(4):922–935, 2018.

[88] Marco Dorigo, Vittorio Maniezzo, and Alberto Colorni. Ant system: optimization by a colony of cooperating agents. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 26(1):29–41, 1996.

[89] Bwo-Ren Ke, Meng-Chieh Chen, and Chun-Liang Lin. Block-layout design using max–min ant system for saving energy on mass rapid transit systems. IEEE Transactions on Intelligent Transportation Systems, 10(2):226–235, 2009.

[90] B-R Ke, C-L Lin, and C-C Yang. Optimisation of train energy-efficient operation for mass rapid transit systems. IET Intelligent Transport Systems, 6(1):58–66, 2012.

[91] Shaofeng Lu, Stuart Hillmansen, Tin Kin Ho, and Clive Roberts. Single-train trajectory optimization. IEEE Transactions on Intelligent Transportation Systems, 14(2):743–750, 2013.

[92] Julien Nespoulous, Christian Soize, Christine Funfschilling, and Guillaume Perrin. Optimisation of train speed to limit energy consumption. Vehicle System Dynamics, 60(10):3540–3557, 2022.

[93] Pengling Wang and Rob MP Goverde. Multi-train trajectory optimization for energy efficiency and delay recovery on single-track railway lines. Transportation Research Part B: Methodological, 105:340–361, 2017.

[94] Kouhei Nishida and Youhei Akimoto. Population size adaptation for the cma-es based on the estimation accuracy of the natural gradient. In Proceedings of the Genetic and Evolutionary Computation Conference 2016, pages 237–244, 2016.

[95] U Tan, Olivier Rabaste, Claude Adnet, and J-P Ovarlez. On the eclipsing phenomenon with phase codes. In 2019 International Radar Conference (RADAR), pages 1–5. IEEE, 2019.

Appendix A

Multibody approach

Notations

We consider a train composed of N rigid bodies (N_C cars, N_B bogies, and N_W wheelsets). Let $\mathcal{R}^0$ be the global Cartesian frame of reference $(\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$ and $\mathcal{R}^\alpha$ be the local frame of reference $(\boldsymbol{e}_x^\alpha, \boldsymbol{e}_y^\alpha, \boldsymbol{e}_z^\alpha)$ attached to the rigid body α. The translation and rotation vectors between these two frames of reference are denoted by $\boldsymbol{\zeta}^\alpha$ and $\boldsymbol{\Theta}^\alpha$, respectively. The train system is summarized in Figures A.1 and A.2.

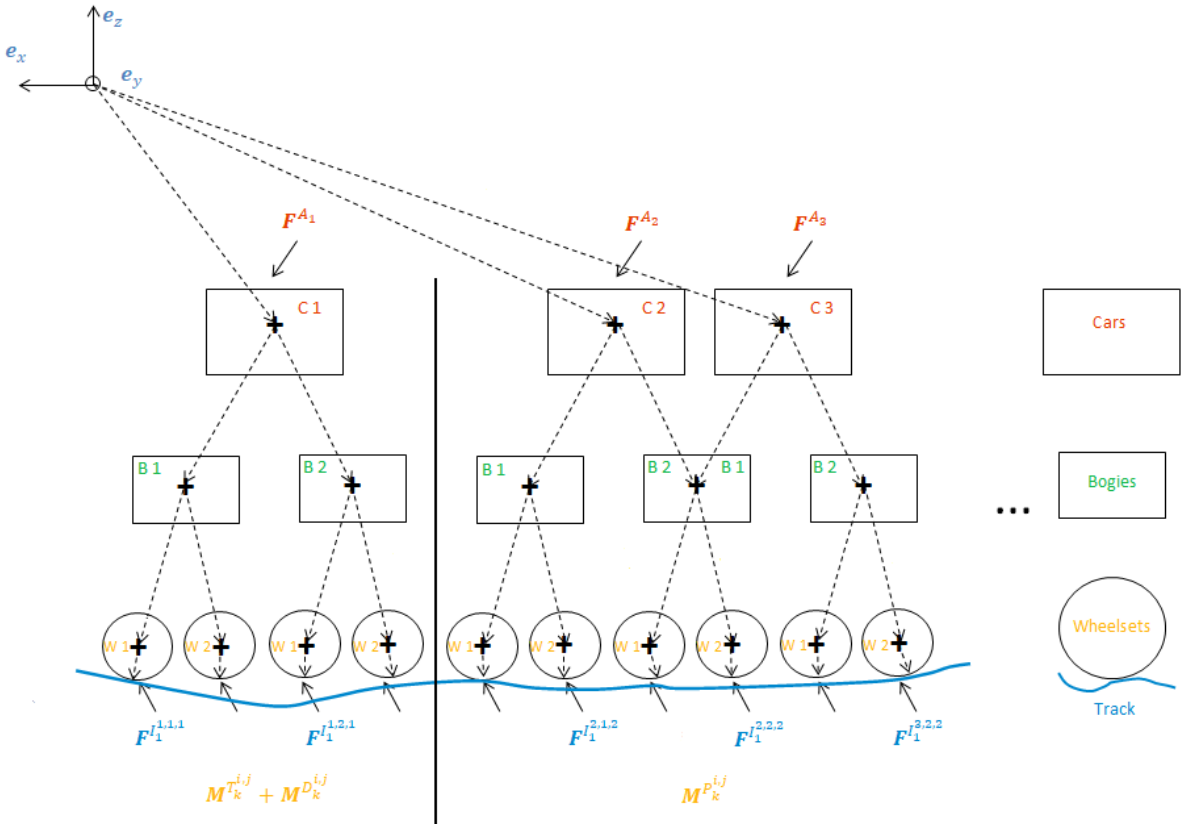

Figure A.1: Scheme of the multibody structure (lateral view).

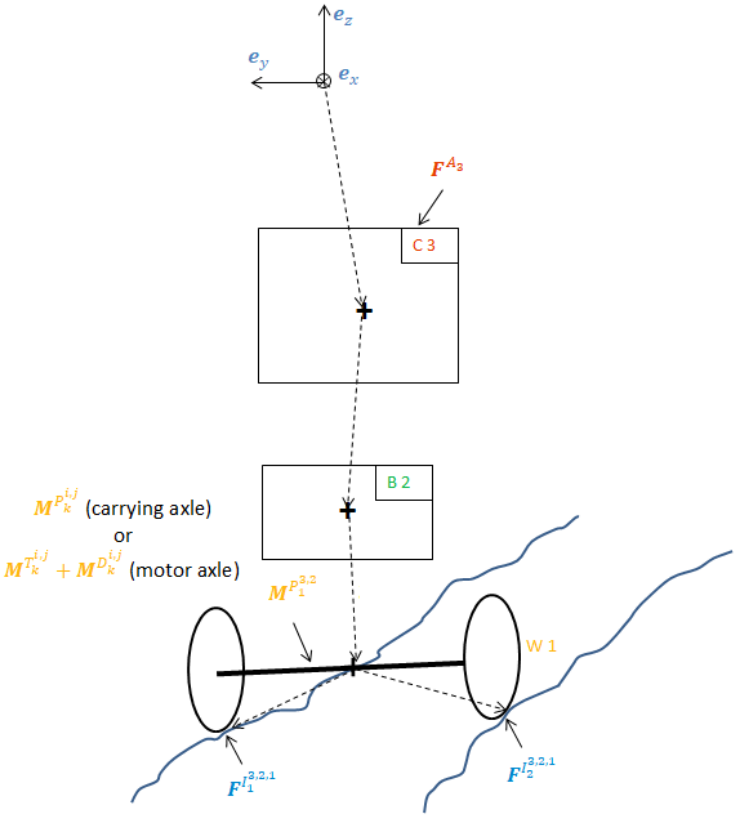

Figure A.2: Scheme of the multibody structure (front view).

The environmental inputs (track curvature, declivity, wind direction and amplitude) is denoted by $\mathcal{T}$, the wheel/rail contact (rail and wheel shape, etc.) is described by $\mathcal{C}$ and $\mathcal{V}$ describes the vehicle. We define the curvilinear position and speed of the train with s and v, respectively, the first car C_1 being the reference,

$$v(t) =< \dot{\boldsymbol{\zeta}}^{C_1}(t), \boldsymbol{e}_x^{Track} > , \tag{A.1}$$

$$s(t) = \int_{t_s}^{t} v(\tau) d\tau \, . \tag{A.2}$$

In order to simplify the comprehension, we use in this appendix a set of colors for denoting each type of rigid bodies (cars, bogies, wheelsets), but also the contact points. A red box is drawn around the simplifications induced by the longitudinal hypotheses.

F	Force applied on cars
F	Force applied on bogies
F	Force applied on wheelsets
F	Force applied at the contact point
$\boxed{F}$	Analytic longitudinal simplifications

The letter C refers to the different cars, B to the bogies, W to the wheelsets, and I to the contact points. If the sum limits are not precised in the following equations, index i browses all cars, index j all the bogies, index k all the wheelsets, and index l all the contact points.

Particular attention is paid to not counting twice the shared bogies.

C_i	i-th car
B_j^i	j-th bogie of the i-th car
$W_k^{i,j}$	k-th wheelset of the j-th bogie of the i-th car
$I_l^{i,j,k}$	l-th contact point of the k-th wheelset, j-th bogie and i-th car

The mass, inertia, stiffness and damping matrices are denoted as follows:

$[M]^\alpha$	mass matrix of the α-th rigid body
$[I]^\alpha$	inertia matrix of the α-th rigid body
$[K]^{\alpha/\beta}$	stiffness matrix between the bodies α and β
$[C]^{\alpha/\beta}$	damping matrix between the bodies α and β

Forces and moments

The forces are of two types. A first group of forces is composed of all the external forces, they are listed below:

- Aerodynamic forces applied on the car bodies (neglected on the other rigid bodies): $\boldsymbol{F}^{A_i}(\dot{\boldsymbol{\zeta}}^{C_i}, \boldsymbol{v}^w(s(t), \mathcal{T}))$
- Aerodynamic moments applied on the car bodies (neglected on the other rigid bodies): $\boldsymbol{M}^{A_i}(\dot{\boldsymbol{\zeta}}^{C_i}, \boldsymbol{v}^w(s(t), \mathcal{T}))$
- Traction moments applied on the motor wheelsets: $\boldsymbol{M}^{T_k^{i,j}}(\boldsymbol{\zeta}^{W_k^{i,j}}, v(t), u(t))$
- Dynamic braking moments applied on the motor wheelsets: $\boldsymbol{M}^{P_k^{i,j}}(\boldsymbol{\zeta}^{W_k^{i,j}}, v(t), u(t))$
- Pneumatic braking moments applied on the passenger wheelsets: $\boldsymbol{M}^{D_k^{i,j}}(\boldsymbol{\zeta}^{W_k^{i,j}}, v(t), u(t))$
- Contact forces applied on the wheels at the contact points: $\boldsymbol{F}^{I_l^{i,j,k}}\left(\boldsymbol{\zeta}^{I_l^{i,j,k}}, \dot{\boldsymbol{\zeta}}^{I_l^{i,j,k}}, \mathcal{T}, \mathcal{C}\right)$
- Contact moments applied on the wheels at the contact points: $\boldsymbol{M}^{I_l^{i,j,k}}\left(\boldsymbol{\zeta}^{I_l^{i,j,k}}, \dot{\boldsymbol{\zeta}}^{I_l^{i,j,k}}, \mathcal{T}, \mathcal{C}\right)$
- Weight applied on all the rigid bodies: $[M]^\alpha \boldsymbol{g}$ (with $\boldsymbol{g} = -9.81\boldsymbol{e}_z$ the gravity acceleration)

The second group of forces is composed of all the internal forces, that is to say all the forces applied by one rigid body on another one. It includes all the suspension elements (primary and secondary) and the interactions between cars.

Lagrangian

The kinetic energy T^α associated with each of the rigid body α is written as follows,

$$T^\alpha(\dot{\boldsymbol{\zeta}}^\alpha, \dot{\boldsymbol{\Theta}}^\alpha) = \frac{1}{2} < [M]^\alpha \dot{\boldsymbol{\zeta}}^\alpha, \dot{\boldsymbol{\zeta}}^\alpha > + \frac{1}{2} < [I]^\alpha \dot{\boldsymbol{\Theta}}^\alpha, \dot{\boldsymbol{\Theta}}^\alpha > . \tag{A.3}$$

Therefore, the total kinetic energy T is the sum of the kinetic energy of all the rigid bodies (Equation (A.3)),

$$\boxed{T(\dot{\boldsymbol{\zeta}}, \dot{\boldsymbol{\Theta}}) = \sum_i T^{C_i}(\dot{\boldsymbol{\zeta}}^{C_i}, \dot{\boldsymbol{\Theta}}^{C_i}) + \sum_{i,j} T^{B^i_j}(\dot{\boldsymbol{\zeta}}^{B^i_j}, \dot{\boldsymbol{\Theta}}^{B^i_j}) + \sum_{i,j,k} T^{W^{i,j}_k}(\dot{\boldsymbol{\zeta}}^{W^{i,j}_k}, \dot{\boldsymbol{\Theta}}^{W^{i,j}_k}) .} \tag{A.4}$$

where $\dot{\boldsymbol{\zeta}}$ and $\dot{\boldsymbol{\Theta}}$ gather the velocities and the rotational velocities of all the rigid bodies. The potential energy of each rigid body is composed of two groups. The first group comes from the gravity. For each rigid body α, the gravity potential energy is given by:

$$V^\alpha(\boldsymbol{\zeta}^\alpha) = < [M]^\alpha \boldsymbol{g}, \boldsymbol{\zeta}^\alpha > . \tag{A.5}$$

Therefore, the total gravity potential energy for all the rigid bodies is expressed by the following expression,

$$V^G(\boldsymbol{\zeta}) = \sum_i V^{C_i}(\boldsymbol{\zeta}^{C_i}) + \sum_{i,j} V^{B^i_j}(\boldsymbol{\zeta}^{B^i_j}) + \sum_{i,j,k} V^{W^{i,j}_k}(\boldsymbol{\zeta}^{W^{i,j}_k}) . \tag{A.6}$$

where $\boldsymbol{\zeta}$ gathers the positions of all the rigid bodies. The second group comes from the stiffness elements of the suspensions. These elements are situated between cars, between the cars and the bogies, and between the bogies and the wheelsets. Between rigid bodies α and β, the stiffness potential energy is such as

$$V^{\alpha/\beta}(\boldsymbol{\zeta}^\alpha, \boldsymbol{\zeta}^\beta) = \frac{1}{2} < [K]^{\alpha/\beta} \left(\boldsymbol{\zeta}^\beta - \boldsymbol{\zeta}^\alpha - \boldsymbol{l}_0^{\alpha/\beta} \right), \boldsymbol{\zeta}^\beta - \boldsymbol{\zeta}^\alpha - \boldsymbol{l}_0^{\alpha/\beta} > . \tag{A.7}$$

where $\boldsymbol{l}_0^{\alpha/\beta}$ is the equilibrium position of β in $\mathcal{R}^\alpha$. The train suspensions are also composed of nonlinear elements (pinlinks, bumpstops, etc.). Their mechanical behavior is considered well known such that the stiffness potential energy is modified as

$$\tilde{V}^{\alpha/\beta}(\boldsymbol{\zeta}^\alpha, \boldsymbol{\zeta}^\beta) = V^{\alpha/\beta}(\boldsymbol{\zeta}^\alpha, \boldsymbol{\zeta}^\beta) + V^{NL,\alpha/\beta}(\boldsymbol{\zeta}^\alpha, \boldsymbol{\zeta}^\beta) . \tag{A.8}$$

The total stiffness potential energy is the sum of the potential energy of all the rigid bodies,

$$V^S(\boldsymbol{\zeta}) = \sum_i \tilde{V}^{C_i/C_{i+1}}(\boldsymbol{\zeta}^{C_i}, \boldsymbol{\zeta}^{C_{i+1}}) + \sum_{i,j} \tilde{V}^{C_i/B^i_j}(\boldsymbol{\zeta}^{C_i}, \boldsymbol{\zeta}^{B^i_j}) + \sum_{i,j,k} \tilde{V}^{B^i_j/W^{i,j}_k}(\boldsymbol{\zeta}^{B^i_j}, \boldsymbol{\zeta}^{W^{i,j}_k}) . \tag{A.9}$$

Finally, the total potential energy is the sum of the gravity potential energy and the stiffness potential energy (Equations (A.6) and (A.9)),

$$\boxed{V(\boldsymbol{\zeta}) = V^S(\boldsymbol{\zeta}) + V^G(\boldsymbol{\zeta}) .} \tag{A.10}$$

The Lagrangian of the train system is written L. It depends on the kinetic energy (Equation (A.4)) and the total potential energy (Equation (A.10)),

$$\boxed{L(\boldsymbol{\zeta}, \dot{\boldsymbol{\zeta}}, \dot{\boldsymbol{\Theta}}) = T(\dot{\boldsymbol{\zeta}}, \dot{\boldsymbol{\Theta}}) - V(\boldsymbol{\zeta}) .} \tag{A.11}$$

Generalized forces

We write q_γ^α one degree of freedom of the train system, with α the rigid body under consideration, and γ the rotation or translation degree of freedom. For example, $q_y^{C_5}$ is the translation degree of freedom of the 5-th car in the y direction. With these notations, we have

$$\boldsymbol{\zeta}^\alpha = q_x^\alpha \boldsymbol{e}_x^\alpha + q_y^\alpha \boldsymbol{e}_y^\alpha + q_z^\alpha \boldsymbol{e}_z^\alpha \,, \tag{A.12}$$

$$\boldsymbol{\Theta}^\alpha = q_\phi^\alpha \boldsymbol{e}_x^\alpha + q_\chi^\alpha \boldsymbol{e}_y^\alpha + q_\psi^\alpha \boldsymbol{e}_z^\alpha \,. \tag{A.13}$$

The total nonpotential force applied on the cars is the sum of the forces between cars, the forces between cars and bogies, and the aerodynamic forces. Note that the first and last cars only interact with one neighboring car,

$$\begin{aligned} F_{\alpha,\gamma}^{totC}(\dot{\boldsymbol{\zeta}}, \boldsymbol{v}^w(s,\mathcal{T})) =& < \boldsymbol{F}^{C_{N_C-1}/C_{N_C}}(\dot{\boldsymbol{\zeta}}^{C_{N_C}}, \dot{\boldsymbol{\zeta}}^{C_{N_C-1}}), \frac{\partial \boldsymbol{\zeta}^{C_{N_C}}}{\partial q_\gamma^\alpha} > - < \boldsymbol{F}^{C_1/C_2}(\dot{\boldsymbol{\zeta}}^{C_1}, \dot{\boldsymbol{\zeta}}^{C_2}), \frac{\partial \boldsymbol{\zeta}^{C_1}}{\partial q_\gamma^\alpha} > \\ &+ \sum_{i=2}^{N_C-1} < \boldsymbol{F}^{C_{i-1}/C_i}(\dot{\boldsymbol{\zeta}}^{C_i}, \dot{\boldsymbol{\zeta}}^{C_{i-1}}) - \boldsymbol{F}^{C_i/C_{i+1}}(\dot{\boldsymbol{\zeta}}^{C_i}, \dot{\boldsymbol{\zeta}}^{C_{i+1}}), \frac{\partial \boldsymbol{\zeta}^{C_i}}{\partial q_\gamma^\alpha} > \\ &- \sum_{i,j} < \boldsymbol{F}^{C_i/B_j^i}(\dot{\boldsymbol{\zeta}}^{C_i}, \dot{\boldsymbol{\zeta}}^{B_j^i}), \frac{\partial \boldsymbol{\zeta}^{C_i}}{\partial q_\gamma^\alpha} > \\ &+ \sum_i < \boldsymbol{F}^{A_i}(\dot{\boldsymbol{\zeta}}^{C_i}, \boldsymbol{v}^w(s,\mathcal{T})), \frac{\partial \boldsymbol{\zeta}^{C_i}}{\partial q_\gamma^\alpha} > + \sum_i < \boldsymbol{M}^{A_i}(\dot{\boldsymbol{\zeta}}^{C_i}, \boldsymbol{v}^w(s,\mathcal{T})), \frac{\partial \boldsymbol{\Theta}^{C_i}}{\partial q_\gamma^\alpha} > \,. \end{aligned} \tag{A.14}$$

The bogies are only submitted to forces between bogies and cars and forces between bogies and wheelsets as follows,

$$\begin{aligned} F_{\alpha,\gamma}^{totB}(\dot{\boldsymbol{\zeta}}) =& \sum_{i,j} < \boldsymbol{F}^{C_i/B_j^i}(\dot{\boldsymbol{\zeta}}^{C_i}, \dot{\boldsymbol{\zeta}}^{B_j^i}), \frac{\partial \boldsymbol{\zeta}^{B_j^i}}{\partial q_\gamma^\alpha} > \\ &- \sum_{i,j,k} < \boldsymbol{F}^{B_j^i/W_k^{i,j}}(\dot{\boldsymbol{\zeta}}^{B_j^i}, \dot{\boldsymbol{\zeta}}^{W_k^{i,j}}), \frac{\partial \boldsymbol{\zeta}^{B_j^i}}{\partial q_\gamma^\alpha} > \,. \end{aligned} \tag{A.15}$$

The forces applied on the wheelsets gather the forces between wheelsets and bogies, the pneumatic braking torques, the dynamic braking torques, and the traction torques. These three torques come from the driver's command and are only applied on a specific type of cars (motor or passenger cars),

$$\begin{aligned} F_{\alpha,\gamma}^{totW}(\dot{\boldsymbol{\zeta}}, v(t), u(t)) =& \sum_{i,j,k} < \boldsymbol{F}^{B_j^i/W_k^{i,j}}(\dot{\boldsymbol{\zeta}}^{B_j^i}, \dot{\boldsymbol{\zeta}}^{W_k^{i,j}}), \frac{\partial \boldsymbol{\zeta}^{W_k^{i,j}}}{\partial q_\gamma^\alpha} > \\ &+ \sum_{i=2}^{N_C-1} \sum_{j,k} < \boldsymbol{M}^{P_k^{i,j}}(\boldsymbol{\zeta}^{W_k^{i,j}}, v(t), u(t)), \frac{\partial \boldsymbol{\Theta}^{W_k^{i,j}}}{\partial q_\gamma^\alpha} > \\ &+ \sum_{i=\{1,N_C\}} \sum_{j,k} < \boldsymbol{M}^{D_k^{i,j}}(\boldsymbol{\zeta}^{W_k^{i,j}}, v(t), u(t)) + \boldsymbol{M}^{T_k^{i,j}}(\boldsymbol{\zeta}^{W_k^{i,j}}, v(t), u(t)), \frac{\partial \boldsymbol{\Theta}^{W_k^{i,j}}}{\partial q_\gamma^\alpha} > \,. \end{aligned} \tag{A.16}$$

Finally, the contact forces are applied on the wheelsets at each contact points,

$$F^{totI}_{\alpha,\gamma}(\boldsymbol{\zeta},\dot{\boldsymbol{\zeta}},\mathcal{T},\mathcal{C}) = \sum_{i,j,k,l} < \boldsymbol{F}^{I^{i,j,k}_l}\left(\boldsymbol{\zeta}^{I^{i,j,k}_l},\dot{\boldsymbol{\zeta}}^{I^{i,j,k}_l},\mathcal{T},\mathcal{C}\right), \frac{\partial \boldsymbol{\zeta}^{I^{i,j,k}_l}}{\partial q^{\alpha}_{\gamma}} > \\ + \sum_{i,j,k,l} < \boldsymbol{M}^{I^{i,j,k}_l}\left(\boldsymbol{\zeta}^{I^{i,j,k}_l},\dot{\boldsymbol{\zeta}}^{I^{i,j,k}_l},\mathcal{T},\mathcal{C}\right), \frac{\partial \boldsymbol{\Theta}^{I^{i,j,k}_l}}{\partial q^{\alpha}_{\gamma}} > . \tag{A.17}$$

The total generalized forces, denoted $Q_{\alpha,\gamma}$, is the sum of all the generalized forces applied on each rigid body. An expression is deduced from Equations (A.14) to (A.17),

$$Q_{\alpha,\gamma}(\boldsymbol{\zeta},\dot{\boldsymbol{\zeta}},v(t),u(t),\mathcal{T},\mathcal{C}) = F^{totC}_{\alpha,\gamma}(\dot{\boldsymbol{\zeta}},\boldsymbol{v}^w(s,\mathcal{T})) + F^{totB}_{\alpha,\gamma}(\dot{\boldsymbol{\zeta}}) + F^{totW}_{\alpha,\gamma}(\dot{\boldsymbol{\zeta}},v(t),u(t)) \\ + F^{totI}_{\alpha,\gamma}(\boldsymbol{\zeta},\dot{\boldsymbol{\zeta}},\mathcal{T},\mathcal{C}) . \tag{A.18}$$

Lagrangian equation

From the Lagrangian (Equation (A.11)) and the generalized forces (Equation (A.18)), the Lagrangian equations are written for all the rigid bodies ($\alpha \in \{C_i, B^i_j, W^{i,j}_k, I^{i,j,k}_l\}$) and for all the degrees of freedom ($\gamma \in \{x, y, z, \phi, \chi, \psi\}$),

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^{\alpha}_{\gamma}}(\boldsymbol{\zeta},\dot{\boldsymbol{\zeta}},\dot{\boldsymbol{\Theta}})\right) - \frac{\partial L}{\partial q^{\alpha}_{\gamma}}(\boldsymbol{\zeta},\dot{\boldsymbol{\zeta}},\dot{\boldsymbol{\Theta}}) = Q_{\alpha,\gamma}(\boldsymbol{\zeta},\dot{\boldsymbol{\zeta}},v(t),u(t),\mathcal{T},\mathcal{C}) . \tag{A.19}$$

By replacing L by the expression of the kinetic energy (Equations (A.3) and (A.4)) and the expression of the potential energy (Equations (A.5) to (A.10)), and $Q_{\alpha,\gamma}$ by the detailed expressions (Equations (A.14) to (A.18)), the dynamic equations for the whole train are obtained from the sum over all the rigid bodies. The formulation for the whole train is given by:

$$
\begin{aligned}
&\frac{d}{dt}\left(\sum_i < [M]^{C_i}\dot{\boldsymbol{\zeta}}^{C_i}, \frac{\partial\dot{\boldsymbol{\zeta}}^{C_i}}{\partial\dot{q}^\alpha_\gamma} > + < [\boldsymbol{I}]^{C_i}\dot{\boldsymbol{\Theta}}^{C_i}, \frac{\partial\dot{\boldsymbol{\Theta}}^{C_i}}{\partial\dot{q}^\alpha_\gamma} > \right.\\
&+\sum_{i,j} < [M]^{B^i_j}\dot{\boldsymbol{\zeta}}^{B^i_j}, \frac{\partial\dot{\boldsymbol{\zeta}}^{B^i_j}}{\partial\dot{q}^\alpha_\gamma} > + < [\boldsymbol{I}]^{B^i_j}\dot{\boldsymbol{\Theta}}^{B^i_j}, \frac{\partial\dot{\boldsymbol{\Theta}}^{B^i_j}}{\partial\dot{q}^\alpha_\gamma} >\\
&\left.+\sum_{i,j,k} < [M]^{W^{i,j}_k}\dot{\boldsymbol{\zeta}}^{W^{i,j}_k}, \frac{\partial\dot{\boldsymbol{\zeta}}^{W^{i,j}_k}}{\partial\dot{q}^\alpha_\gamma} > + < [\boldsymbol{I}]^{W^{i,j}_k}\dot{\boldsymbol{\Theta}}^{W^{i,j}_k}, \frac{\partial\dot{\boldsymbol{\Theta}}^{W^{i,j}_k}}{\partial\dot{q}^\alpha_\gamma} >\right)\\
&+\sum_i < [K]^{C_i/C_{i+1}}\left(\boldsymbol{\zeta}^{C_{i+1}} - \boldsymbol{\zeta}^{C_i} - \boldsymbol{l}_0^{C_i/C_{i+1}}\right), \frac{\partial\left(\boldsymbol{\zeta}^{C_{i+1}} - \boldsymbol{\zeta}^{C_i}\right)}{\partial q^\alpha_\gamma} >\\
&+\sum_{i,j} < [K]^{C_i/B^i_j}\left(\boldsymbol{\zeta}^{B^i_j} - \boldsymbol{\zeta}^{C_i} - \boldsymbol{l}_0{}^{C_i/B^i_j}\right), \frac{\partial\left(\boldsymbol{\zeta}^{B^i_j} - \boldsymbol{\zeta}^{C_i}\right)}{\partial q^\alpha_\gamma} >\\
&+\sum_{i,j,k} < [K]^{B^i_j/W^{i,j}_k}\left(\boldsymbol{\zeta}^{W^{i,j}_k} - \boldsymbol{\zeta}^{B^i_j} - \boldsymbol{l}_0{}^{B^i_j/W^{i,j}_k}\right), \frac{\partial\left(\boldsymbol{\zeta}^{W^{i,j}_k} - \boldsymbol{\zeta}^{B^i_j}\right)}{\partial q^\alpha_\gamma} >\\
&+\sum_i < [M]^{C_i}\boldsymbol{g}, \frac{\partial\boldsymbol{\zeta}^{C_i}}{\partial q^\alpha_\gamma} > + \sum_{i,j} < [M]^{B^i_j}\boldsymbol{g}, \frac{\partial\boldsymbol{\zeta}^{B^i_j}}{\partial q^\alpha_\gamma} > + \sum_{i,j,k} < [M]^{W^{i,j}_k}\boldsymbol{g}, \frac{\partial\boldsymbol{\zeta}^{W^{i,j}_k}}{\partial q^\alpha_\gamma} >\\
&=< \boldsymbol{F}^{C_{N_C-1}/C_{N_C}}(\dot{\boldsymbol{\zeta}}^{C_{N_C}}, \dot{\boldsymbol{\zeta}}^{C_{N_C-1}}), \frac{\partial\boldsymbol{\zeta}^{C_{N_C}}}{\partial q^\alpha_\gamma} > - < \boldsymbol{F}^{C_1/C_2}(\dot{\boldsymbol{\zeta}}^{C_1}, \dot{\boldsymbol{\zeta}}^{C_2}), \frac{\partial\boldsymbol{\zeta}^{C_1}}{\partial q^\alpha_\gamma} >\\
&+\sum_{i=2}^{N_C-1} < \boldsymbol{F}^{C_{i-1}/C_i}(\dot{\boldsymbol{\zeta}}^{C_i}, \dot{\boldsymbol{\zeta}}^{C_{i-1}}) - \boldsymbol{F}^{C_i/C_{i+1}}(\dot{\boldsymbol{\zeta}}^{C_i}, \dot{\boldsymbol{\zeta}}^{C_{i+1}}), \frac{\partial\boldsymbol{\zeta}^{C_i}}{\partial q^\alpha_\gamma} > - \sum_{i,j} < \boldsymbol{F}^{C_i/B^i_j}(\dot{\boldsymbol{\zeta}}^{C_i}, \dot{\boldsymbol{\zeta}}^{B^i_j}), \frac{\partial\boldsymbol{\zeta}^{C_i}}{\partial q^\alpha_\gamma} >\\
&+\sum_i < \boldsymbol{F}^{A_i}(\dot{\boldsymbol{\zeta}}^{C_i}, \boldsymbol{v}^w(s,\mathcal{T})), \frac{\partial\boldsymbol{\zeta}^{C_i}}{\partial q^\alpha_\gamma} > + \sum_i < \boldsymbol{M}^{A_i}(\dot{\boldsymbol{\zeta}}^{C_i}, \boldsymbol{v}^w(s,\mathcal{T})), \frac{\partial\boldsymbol{\Theta}^{C_i}}{\partial q^\alpha_\gamma} >\\
&+\sum_{i,j} < \boldsymbol{F}^{C_i/B^i_j}(\dot{\boldsymbol{\zeta}}^{C_i}, \dot{\boldsymbol{\zeta}}^{B^i_j}), \frac{\partial\boldsymbol{\zeta}^{B^i_j}}{\partial q^\alpha_\gamma} > - \sum_{i,j,k} < \boldsymbol{F}^{B^i_j/W^{i,j}_k}(\dot{\boldsymbol{\zeta}}^{B^i_j}, \dot{\boldsymbol{\zeta}}^{W^{i,j}_k}), \frac{\partial\boldsymbol{\zeta}^{B^i_j}}{\partial q^\alpha_\gamma} >\\
&+\sum_{i,j,k} < \boldsymbol{F}^{B^i_j/W^{i,j}_k}(\dot{\boldsymbol{\zeta}}^{B^i_j}, \dot{\boldsymbol{\zeta}}^{W^{i,j}_k}), \frac{\partial\boldsymbol{\zeta}^{W^{i,j}_k}}{\partial q^\alpha_\gamma} > + \sum_{i=2}^{N_C-1}\sum_{j,k} < \boldsymbol{M}^{P^{i,j}_k}(\boldsymbol{\zeta}^{W^{i,j}_k}, v(t), u(t)), \frac{\partial\boldsymbol{\Theta}^{W^{i,j}_k}}{\partial q^\alpha_\gamma} >\\
&+\sum_{i=\{1,N_C\}}\sum_{j,k} < \boldsymbol{M}^{D^{i,j}_k}(\boldsymbol{\zeta}^{W^{i,j}_k}, v(t), u(t)) + \boldsymbol{M}^{T^{i,j}_k}(\boldsymbol{\zeta}^{W^{i,j}_k}, v(t), u(t)), \frac{\partial\boldsymbol{\Theta}^{W^{i,j}_k}}{\partial q^\alpha_\gamma} >\\
&+\sum_{i,j,k,l} < \boldsymbol{F}^{I^{i,j,k}_l}\left(\boldsymbol{\zeta}^{I^{i,j,k}_l}, \dot{\boldsymbol{\zeta}}^{I^{i,j,k}_l}, \mathcal{T}, \mathcal{C}\right), \frac{\partial\boldsymbol{\zeta}^{I^{i,j,k}_l}}{\partial q^\alpha_\gamma} > + \sum_{i,j,k,l} < \boldsymbol{M}^{I^{i,j,k}_l}\left(\boldsymbol{\zeta}^{I^{i,j,k}_l}, \dot{\boldsymbol{\zeta}}^{I^{i,j,k}_l}, \mathcal{T}, \mathcal{C}\right), \frac{\partial\boldsymbol{\Theta}^{I^{i,j,k}_l}}{\partial q^\alpha_\gamma} >.
\end{aligned}
\tag{A.20}
$$

Longitudinal dynamics

In our case, the system is composed of more than 40 rigid bodies, each of them having 6 degrees of freedom. Consequently, Equations (A.20) gather around 240 dynamic equations. Solving all of them seems far too computational costly, especially when we know that the dynamic equations are solved $2\,400\,000$ by the CMA-ES algorithm. For this reason, we propose to extract from Equation (A.20) the longitudinal dynamics of the train.

(i) We remind that Equation (A.20) is valid for all the degrees of freedom of all the rigid bodies. In our case, we propose to only extract the longitudinal equation, that is to say replace q_γ^α by the longitudinal degree of freedom q_x^α. Only translation q_x^α and rotation q_χ^α (wheels rotation) are involved in longitudinal dynamics. Aerodynamic and contact moments are equal to zero because their effects are negligible in these directions. The green boxes of Equation (A.20) are not kept.

(ii) We assume that the rotation of the cars and the bogies have a small influence on the kinetic energy. For this reason, the associated terms are neglected in Equation (A.20) and the blue boxes are removed.

(iii) We can suppose that the longitudinal speed of all the rigid bodies is approximately equal: $\dot{y}\boldsymbol{e}_x^{Track} \simeq \dot{\boldsymbol{\zeta}}^\alpha$. This signifies that all rigid bodies behave like one. We define q_x the translation degree of freedom common to all rigid bodies. If we assume that all the rigid bodies translate in the longitudinal track axis, we have

$$\frac{\partial \boldsymbol{\zeta}^\alpha}{\partial q_x} = \boldsymbol{e}_x^{Track}\,. \tag{A.21}$$

These hypotheses also implies that the stiffness potential terms (red box of Equation (A.20)) compensate each other. Indeed, for all the rigid bodies α and β, we have $\frac{\partial\left(\boldsymbol{\zeta}^{C_{i+1}} - \boldsymbol{\zeta}^{C_i}\right)}{\partial q_x} = \boldsymbol{e}_x^{Track} - \boldsymbol{e}_x^{Track} = 0$. For the same reason, all the nonpotential forces (yellow boxes of Equation (A.20)) have opposite signs and compensate each other. The red and yellow boxes are also simplified to zero.

With these simplifications, Equation (A.20) is simplified such as

$$\frac{d}{dt}\left(\sum_i < [M]^{C_i}\dot{\boldsymbol{\zeta}}^{C_i}, \boldsymbol{e}_x^{Track} > + \sum_{i,j} < [M]^{B_j^i}\dot{\boldsymbol{\zeta}}^{B_j^i}, \boldsymbol{e}_x^{Track} > + \sum_{i,j,k} < [M]^{W_k^{i,j}}\dot{\boldsymbol{\zeta}}^{W_k^{i,j}}, \boldsymbol{e}_x^{Track} > \right.$$

$$\left. + \sum_{i,j,k} < [\boldsymbol{I}]^{W_k^{i,j}}\dot{\boldsymbol{\Theta}}^{W_k^{i,j}}, \frac{\partial \dot{\boldsymbol{\Theta}}^{W_k^{i,j}}}{\partial \dot{q}_x} > \right)$$

$$+ \sum_i < [M]^{C_i}\boldsymbol{g}, \boldsymbol{e}_x^{Track} > + \sum_{i,j} < [M]^{B_j^i}\boldsymbol{g}, \boldsymbol{e}_x^{Track} > + \sum_{i,j,k} < [M]^{W_k^{i,j}}\boldsymbol{g}, \boldsymbol{e}_x^{Track} >$$

$$= \sum_i < \boldsymbol{F}^{A_i}(\dot{\boldsymbol{\zeta}}^{C_i}, \boldsymbol{v}^w(s,\mathcal{T})), \boldsymbol{e}_x^{Track} > + \sum_{i,j,k,l} < \boldsymbol{F}^{I_l^{i,j,k}}\left(\boldsymbol{\zeta}^{I_l^{i,j,k}}, \dot{\boldsymbol{\zeta}}^{I_l^{i,j,k}}, \mathcal{T}, \mathcal{C}\right), \boldsymbol{e}_x^{Track} >$$

$$+ \sum_{i=\{1,N_C\}} \sum_{j,k} < \boldsymbol{M}^{D_k^{i,j}}(\boldsymbol{\zeta}^{W_k^{i,j}}, v(t), u(t)) + \boldsymbol{M}^{T_k^{i,j}}(\boldsymbol{\zeta}^{W_k^{i,j}}, v(t), u(t)), \frac{\partial \boldsymbol{\zeta}^{I_l^{i,j,k}}}{\partial q_x} >$$

$$+ \sum_{i=2}^{N_C-1} \sum_{j,k} < \boldsymbol{M}^{P_k^{i,j}}(\boldsymbol{\zeta}^{W_k^{i,j}}, v(t), u(t), \frac{\partial \boldsymbol{\zeta}^{I_l^{i,j,k}}}{\partial q_x} > . \tag{A.22}$$

Hypothesis (iii) also allow us to gather all the terms of the blue box of Equation (A.22) as they have the same speed. The blue box is thus equivalent to the total mass of the train M (as we sum the mass over all the rigid bodies) multiplied by the train speed $\dot{y}$.

The red box of Equation (A.22) cannot be omitted as the rotations of the wheelsets are not negligible. If we consider that there is no slip between wheels and rail, we can link the rotation speed of the wheels and the longitudinal train speed. With r_m the mean radius of the wheels, we assume that

$$\dot{\Theta}_y^{W_k^{i,j}} \simeq \frac{\dot{\zeta}_x^{W_k^{i,j}}}{r_m} = \frac{\dot{y}}{r_m} \quad , \quad \frac{\partial \dot{\boldsymbol{\Theta}}^{W_k^{i,j}}}{\partial \dot{q}_x} \simeq \frac{1}{r_m}\frac{\partial \dot{\zeta}_x^{W_k^{i,j}}}{\partial \dot{q}_x} = \frac{1}{r_m}\boldsymbol{e}_x^{Track} . \tag{A.23}$$

Thus, the yellow box is approximated to be equal to the number of wheelsets N_W multiplied by the y diagonal component of the inertia matrix I_y^W (considered equal for all wheelsets) and multiplied to the train speed divided by the mean radius of the wheels. This quantity can be expressed depending on the total mass of the train M,

$$N_W\, I_y^W\, v/(r_m)^2 \simeq 0,04 M\, v\, . \tag{A.24}$$

We recognize the correction factor $k^{rot} = 1,04$ that is defined in order to consider the inertia of wheelsets exposed in Equation (II.16). The terms of the green box can be gathered to obtain the complete weight of the train. The projection of the vertical gravity vector on the longitudinal track axis is equal to the sinus of the declivity angle of the track (Equation (II.15)).

The yellow box represent the friction forces. In train longitudinal dynamics, these forces are approximated with the Davis forces (Equation (II.10)). We often complete this approximation with a corrective term (Equation (II.11)).

Finally, the purple boxes of Equation (A.22) represent the traction and braking torques applied by the driver. These torques can be projected into the longitudinal to the track axis and sum in order to estimate the equivalent traction and braking forces for the whole train. With all these simplifications, the analytic equation becomes

$$\begin{aligned} M\left(\mathcal{V}\right) k^{rot} \ddot{Y}\left(t\right) = & F^{T}\left(\dot{Y}\left(t\right), u(t), \mathcal{V}\right) - F^{B}\left(\dot{Y}\left(t\right), u(t), \mathcal{V}\right) \\ & - F^{R}\left(\dot{Y}\left(t\right), v^{w}\left(Y\left(t\right), \mathcal{T}\right), \mathcal{V}\right) - F^{C}\left(Y\left(t\right), \mathcal{T}, \mathcal{V}\right) + F^{W}\left(Y\left(t\right), \mathcal{V}, \mathcal{T}\right) . \end{aligned} \tag{A.25}$$

Appendix B

CMA-ES optimization algorithm

The CMA-ES method has first been developed by Nikolaus Hansen in [76]. This optimization method uses an iterative random search algorithm that can be applied to nonlinear, nonconvex functions. In our case, CMA-ES is applied to the driver's command $\boldsymbol{q}^{(m)}$. To further simplify the notation, the exponent (m) has been removed. After initializing a starting point $\boldsymbol{\nu}^0$, an initial standard deviation σ^0, a population size N^{pop}, boundary conditions on the optimization variables, an initial direction to explore by defining a specific initial covariance matrix $[C^0]$ (identity matrix by default), termination conditions, and learning rates for example. The number of points N^{pop} inside this population is an hyperparameter chosen by the user but it is recommended to adapt it from the dimension of the optimization problem m. The population size can also be modified at each iteration to be adapted to the current state of the algorithm [94]. In practice, the value of N^{pop} is often chosen constant with $N^{pop} = 4 + \ln(m)$. Then, once the algorithm is initialized properly, the algorithm iterates several steps. We describe bellow the steps of the i-th iteration.

Sampling a population of search points

The first step of the algorithm is to draw a population of search points $(\boldsymbol{q}_1^i, \cdots, \boldsymbol{q}_{N^{pop}}^i)$. This search points are drawn in a direction that has appeared to be interesting during the previous steps. To do so, we use the mean $\boldsymbol{\nu}^{i-1}$, the standard deviation σ^{i-1} and the covariance matrix $[C^{i-1}]$ obtained in the previous iteration,

$$\boldsymbol{q}_k^i = \boldsymbol{\nu}^{i-1} + \sigma^{i-1}\mathcal{N}\left(\mathbf{0}, \left[C^{i-1}\right]\right) \qquad k = 1, \cdots, N^{pop}. \tag{B.1}$$

Selection and Recombination: Update the mean

The second step aims to update the mean value of the search distribution. This mean value is adapted from the search points population. After evaluating the cost function $\mathcal{F}$ (which could be eventually penalized to include constraints), we propose to sort the population depending on the quality of the search points. If the problem aims to minimize $\mathcal{F}$, the sorted search points are such that: $\mathcal{F}(\boldsymbol{q}_1^i) < \cdots < \mathcal{F}(\boldsymbol{q}_{N^{pop}}^i)$. The algorithm only conserves the μ best points regarding the objective (selection step). Finally, a recombination is applied to extract the information of the best selected points to update the mean. In practice, the choice of the $\mu > 1$ conserved points is often defined as $N^{pop}/2$ (half of the points are conserved) and the recombination positive

weights $(w_1, \cdots, w_\mu)$ needs to verify $\sum_{k=1}^{\mu} w_k = 1$. For instance, the weights can be defined by $w_k = \frac{2}{\mu(\mu+1)} (\mu - k + 1), k = 1, \cdots, \mu$ such as

$$\boxed{\boldsymbol{\nu}^i = \boldsymbol{\nu}^{i-1} + c_m \sum_{k=1}^{\mu} w_k \left(\boldsymbol{q}_k^i - \boldsymbol{\nu}^{i-1} \right) ,} \tag{B.2}$$

where $c_m \leq 1$ is an hyperparameter defining the learning rate associated with the mean. It can be chosen by the user but is usually set to 1.

Search an interesting direction: Update the covariance matrix

The covariance matrix also needs to be updated as it gives information on interesting directions to explore. In fact, from previous search population, we want to orientate the exploration through a direction which seems to minimize the cost function (eventually penalized). The update of the covariance matrix can be directly estimated from the previous population of points. Nevertheless, this estimation is relatively imprecise for small populations and two other methods are proposed. The first method is called the rank-μ update. If we call $\widetilde{\boldsymbol{q}}_k^i = (\boldsymbol{q}_k^i - \boldsymbol{\nu}^{i-1}) / \sigma^{i-1}$, the covariance matrix adapted from the rank-μ-method is written

$$\left[C^i\right] = \left(1 - c_\mu \sum_{k=1}^{N^{pop}} w_k\right) \left[C^{i-1}\right] + c_\mu \sum_{k=1}^{N^{pop}} w_k \widetilde{\boldsymbol{q}}_k^i \widetilde{\boldsymbol{q}}_k^{i\,T} , \tag{B.3}$$

with c_μ a learning rate associated with the rank-μ update method. The second method is called the rank-one update and only uses one point at each iteration. We define $\boldsymbol{p}_c^i$ the evolution path, which is initialized at $\mathbf{0}$. At the i-th iteration, it can be determined from

$$\boldsymbol{p}_c^i = (1 - c_c) \, \boldsymbol{p}_c^{i-1} + \left(c_c \left(2 - c_c\right) \mu_{\text{eff}}\right)^{1/2} \frac{\boldsymbol{\nu}^i - \boldsymbol{\nu}^{i-1}}{\sigma^i} , \tag{B.4}$$

where c_c the learning rate associated with the cumulation for the rank-one update and $\mu_{\text{eff}} = \left(\sum_{k=1}^{\mu} {w_k}^2 \right)^{-1}$ the variance effective selection mass for the mean. The covariance estimated with this method is defined by:

$$\left[C^i\right] = (1 - c_1) \left[C^{i-1}\right] + c_1 \boldsymbol{p}_c^i {\boldsymbol{p}_c^i}^T , \tag{B.5}$$

with c_1 a learning rate for the rank-one update method. Finally, the combination of the rank-μ update and the rank-one update methods gives a more precise estimation of the covariance matrix. It is obtained from Equations (B.3) and (B.5),

$$\boxed{\left[C^i\right] = \left(1 - c_1 - c_\mu \sum_{k=1}^{N^{pop}} w_k\right) \left[C^{i-1}\right] + c_1 \boldsymbol{p}_c^i {\boldsymbol{p}_c^i}^T + c_\mu \sum_{k=1}^{N^{pop}} w_k \widetilde{\boldsymbol{q}}_k^i \widetilde{\boldsymbol{q}}_k^{i\,T} .} \tag{B.6}$$

Amplitude of the exploration: Update the standard deviation

The amplitude of exploration of the i-th iteration is given by the standard deviation σ^i. Its estimation depends on an evolution path $\boldsymbol{p}_\sigma^i$, which is initialized at $\mathbf{0}$ such as

$$\boldsymbol{p}_\sigma^i = (1 - c_\sigma)\, \boldsymbol{p}_\sigma^{i-1} + \left(c_\sigma\,(2 - c_\sigma)\,\mu_{\text{eff}}\right)^{1/2} \left([C]^{i-1}\right)^{-1/2} \frac{\boldsymbol{\nu}^i - \boldsymbol{\nu}^{i-1}}{\sigma^{i-1}}\,, \tag{B.7}$$

where c_σ is a learning rate associated with the standard deviation. Then, the updated value for the standard deviation is given by:

$$\sigma^i = \sigma^{i-1} \exp\left(\frac{c_\sigma}{d_\sigma}\left(\frac{\|\boldsymbol{p}_\sigma^i\|}{\mathbb{E}\|\mathcal{N}\,(\mathbf{0}, \mathbb{1})\|} - 1\right)\right)\,, \tag{B.8}$$

with d_σ is a damping parameter assuring the convergence on the standard deviation, $\mathcal{N}\,(\mathbf{0}, \mathbb{1})$ is a multivariate centered normal distribution with unity covariance matrix, $\|.\|$ is the Euclidian norm, and $\mathbb{E}$ designs the expectation operator.

Conclusion of the method

Once the mean, covariance matrix, and standard deviation are updated, the algorithm starts a new iteration until the convergence tolerance is reached. This method proves good results for nonlinear optimization problem with nonconvex, noisy functions. The constraints are implemented with a Lagrangian penalization of the cost function. Moreover, the algorithm can be easily parallelized as the evaluation of the search points are independent to each other. This section has presented an outline of the method, but more details can be found in [76]. The algorithm steps are summarized below.

Initialization: initial point, standard deviation, covariance matrix, hyperparameters (population size, convergence tolerance, learning rates)

for $i = 1 : N^{it}$ **do**

- Generate randomly a population of N^{pop} search points $\boldsymbol{q}_k^i$ (Equation (B.1))
- Select and recombine the search points
- Update mean value $\boldsymbol{\nu}^i$ (Equation (B.2))
- Update covariance matrix $[C^i]$ (Equation (B.6))
- Update standard deviation σ^i (Equation (B.8))

end

Return the computed optimal value.

Algorithm 5: CMA-ES algorithm.

Figure B.1 shows an example of the application of the CMA-ES method. We can see the adaptability of the algorithm and understand how it behaves to estimate an optimal value of the cost function.

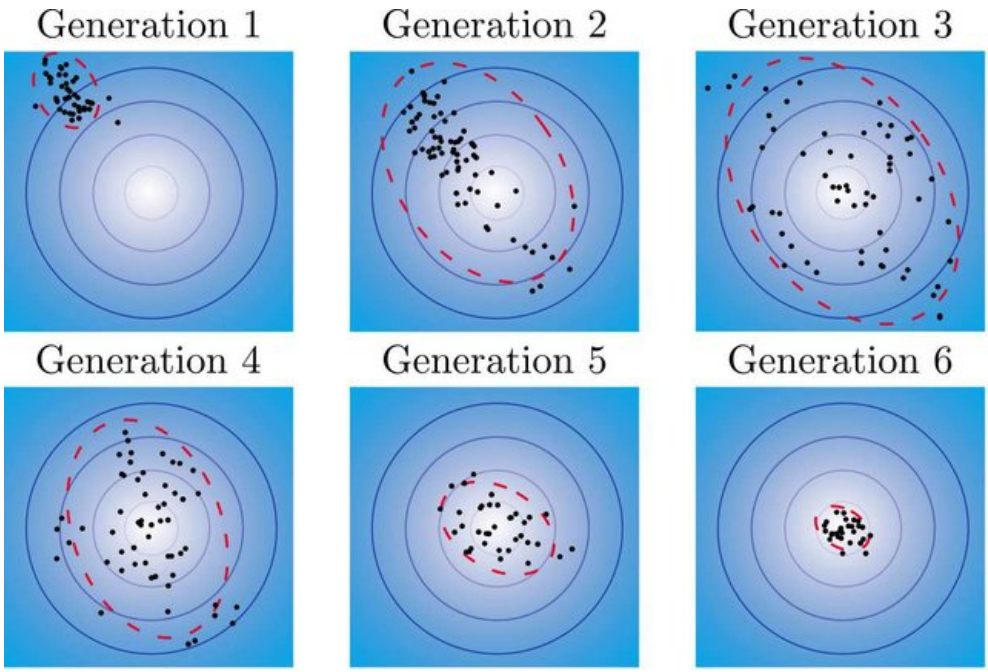

Figure B.1: Illustration of an application of the CMA-ES method on a $2D$ example [95]. Several iterations of population are shown with the associated covariance. The cost function is represented with the background color intensity.

Appendix C

Case study: optimization of the driver's command for autonomous trains

The deterministic method used in Section IV.1.2 has been applied to an industrial case study that aims to test some technologies and advances to prepare the development of the autonomous train. The track considered is referenced as $242,000$. Specifically, the portion of the track between the cities of Busigny and Aulnoye-Aymeries in northeastern France is studied. This journey is composed of four intermediate stops in the stations of Le Cateau, Ors, Landrecies, Hachette before finally stopping in Aulnoye-Aymeries for a total length of $35,km$. The environment is described in Figure C.1, with the speed limitation and the declivity (or the altitude see Equation (II.2)). The track curvature and wind are not considered due to the lack of information in this case study. The departure and arrival times are set to match the actual journey that takes about 30 minutes.

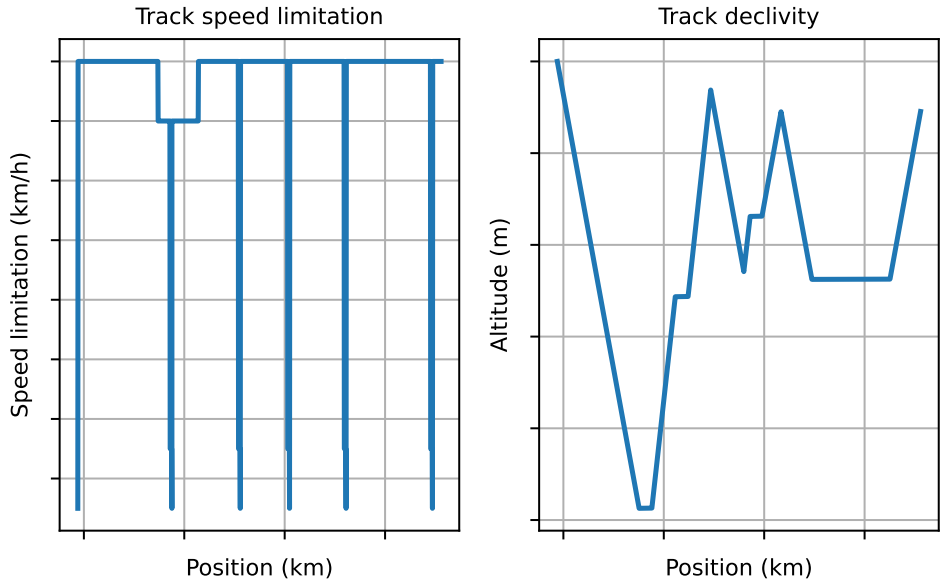

Figure C.1: Rolling environment as a function of the position (speed limitation on the left and altitude on the right).

The measurement vehicle is a well-known French regional train (TER). This train is not equipped with electrodynamic braking, and therefore does not recover energy during braking. The model parameters have been determined relatively accurately for experimental purposes. For this reason, we propose to use the deterministic representation of the model parameters proposed in Section III.3, combined with the deterministic optimization problem presented in Section IV.1.2. The model has been validated from the speed profile measurements performed (see Figure C.2).

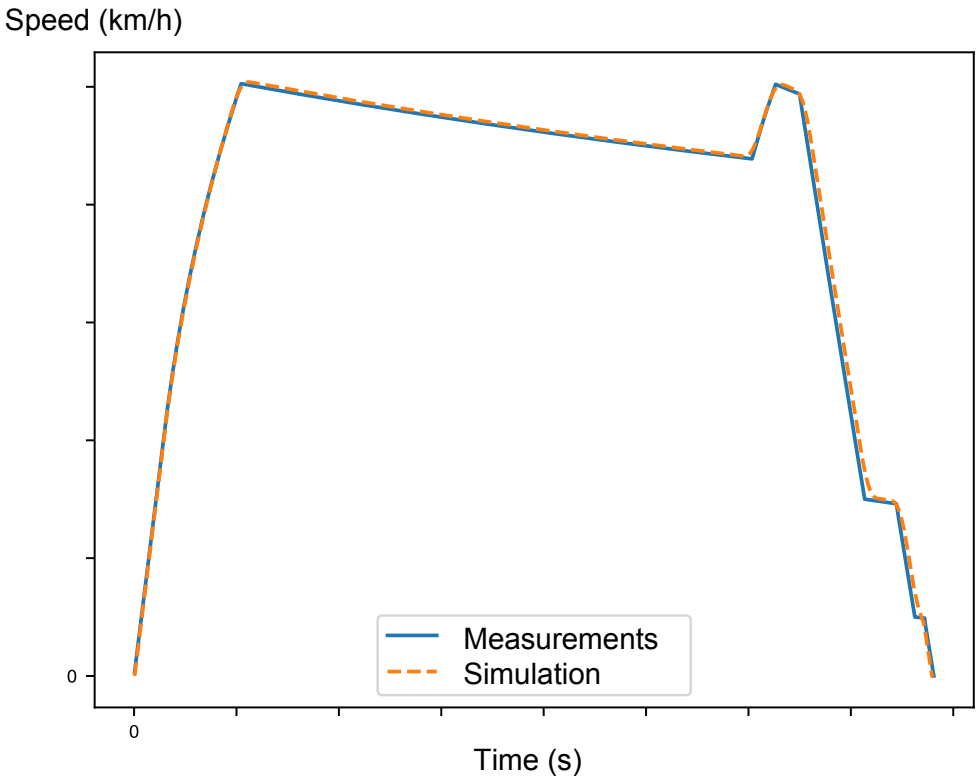

Figure C.2: Validation of the model: Simulated and measured speed profile as a function of time.

The initial driver's command injected in the optimization method is the one realized by the measurement train. This driver's command is determined by a previous coarse optimization algorithm that is used today by SNCF. This coarse algorithm does not consider the wind, and only takes into account the track gradient. Since the journey is relatively short with small variations in gradient, the method does not have many degrees of freedom to reduce energy consumption. Nevertheless, the method presented in this thesis still shows significant improvements. The optimization algorithm is applied to each of the 5 parts of the journey, separately. After several thousand iterations, the algorithm has reached convergence. The journey being short, the computational time is lower than the one of the previous examples in Chapter IV (only one hour is needed). The convergence of the algorithm is shown in Figure C.3 for the first part of the journey.

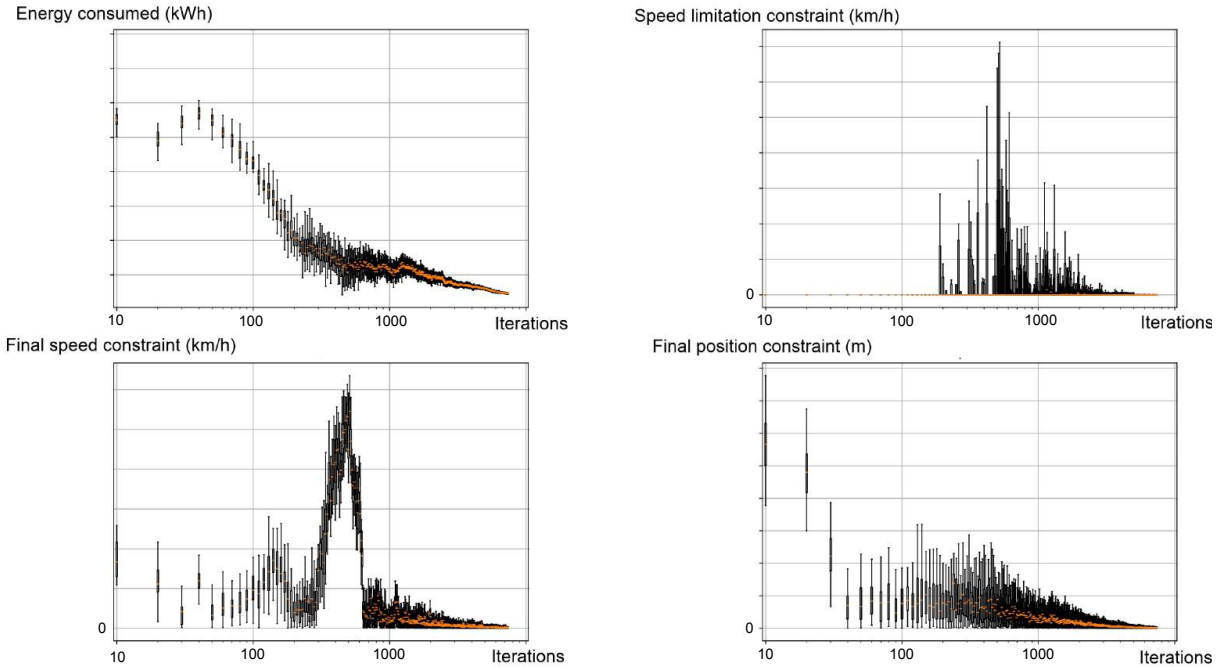

Figure C.3: Convergence of the method. The energy consumption (top left), the speed limitation constraint (top right), the final speed constraint (bottom left), and the final position constraint (bottom right) are plotted against the number of iterations.

After 50 iterations, an interesting direction to explore is discovered, and the energy consumption starts to decrease. The energy consumed by the train starts to stabilize around 500 iterations due to the increasing importance of the constraints. The Lagrange multipliers penalize the cost function, and the algorithm cannot ignore them anymore. After $7\,000$ iterations, the algorithm has reached convergence and the optimal solution respects the constraints. Several iterations have been drawn in Figures C.4 to C.8 to better understand how CMA-ES algorithm works.

The second use of traction just before braking (see Figure C.4) is gradually attenuated because it does not correspond to an efficient behavior in terms of energy saving. It has completely disappeared after 250 iterations. The first and last part of the journey are also heavily modified to verify the constraints. The middle part of the journey varies a lot during the first few iterations, but it seems logical that braking for no reason is equivalent to energy lost. For this reason, the driver's command is close to zero in this part of the journey. The dispersion of the points is very significant at the beginning and decreases with the number of iterations until being very narrow when the algorithm is close to reach convergence. We stop the calculation few hundred meters before the train station due to the fact that the new speed limitation is very difficult to integrate for the algorithm. The last 200 meters are added manually to finish the journey and to be able to compare the optimal solution with the measurements.

When the calculation is performed on each of the 5 parts of the path, the optimal solutions are concatenated. The optimal driver's command is shown in Figure C.9.

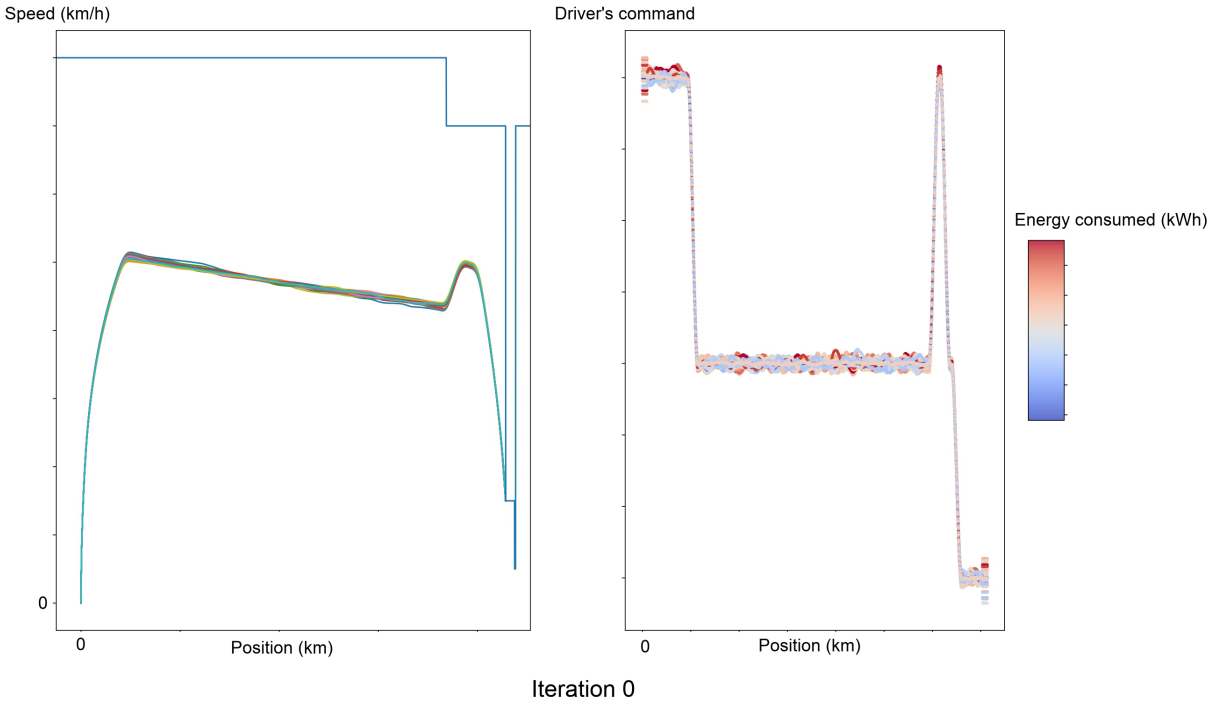

Figure C.4: Speed profile (left) and driver's command (right) depending on the position for the population of points drawn at iteration 0 of CMA-ES algorithm. The consumed energy is represented in color in the right figure.

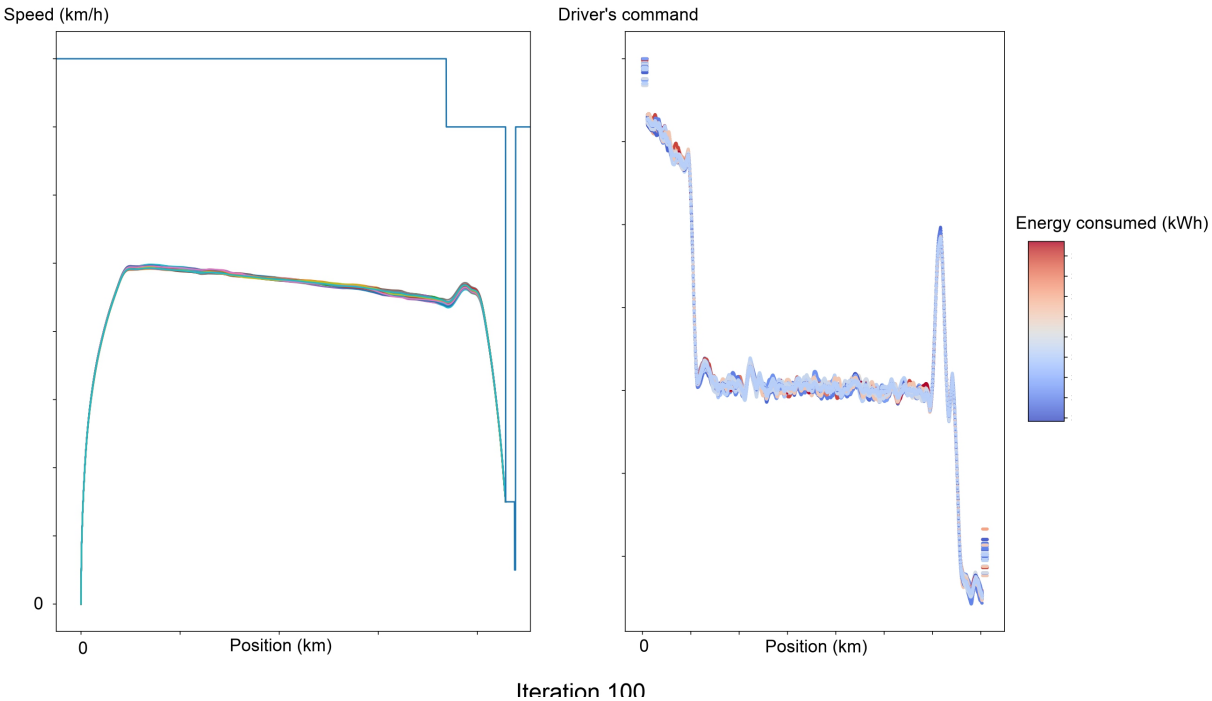

Figure C.5: Speed profile (left) and driver's command (right) depending on the position for the population of points drawn at iteration 100 of CMA-ES algorithm. The consumed energy is represented in color in the right figure.

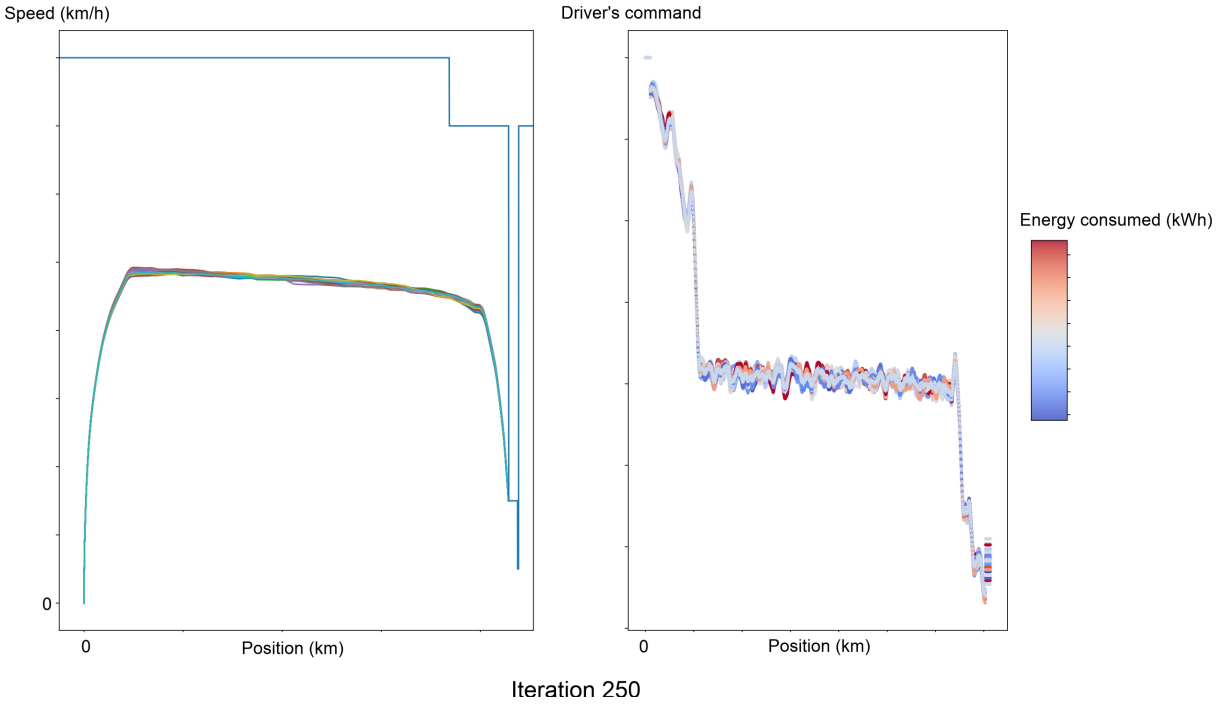

Figure C.6: Speed profile (left) and driver's command (right) depending on the position for the population of points drawn at iteration 250 of CMA-ES algorithm. The consumed energy is represented in color in the right figure.

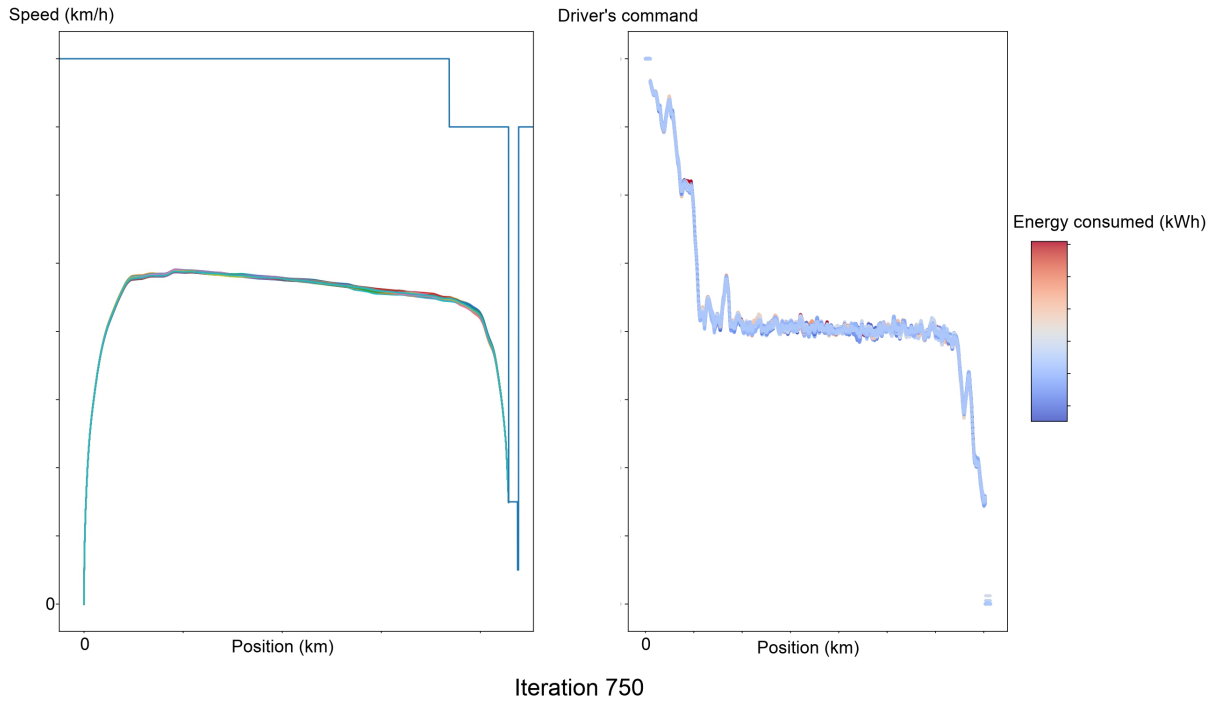

Figure C.7: Speed profile (left) and driver's command (right) depending on the position for the population of points drawn at iteration 750 of CMA-ES algorithm. The consumed energy is represented in color in the right figure.

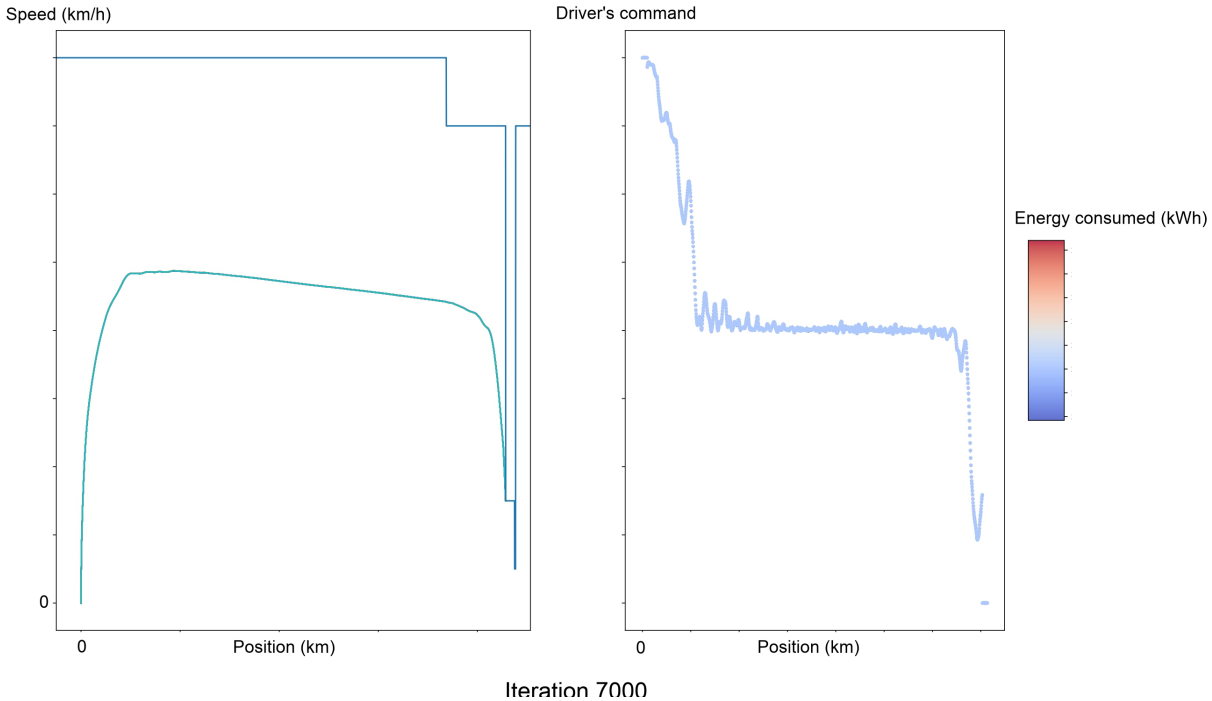

Figure C.8: Speed profile (left) and driver's command (right) depending on the position for the population of points drawn at iteration 7 000 of CMA-ES algorithm. The consumed energy is represented in color in the right figure.

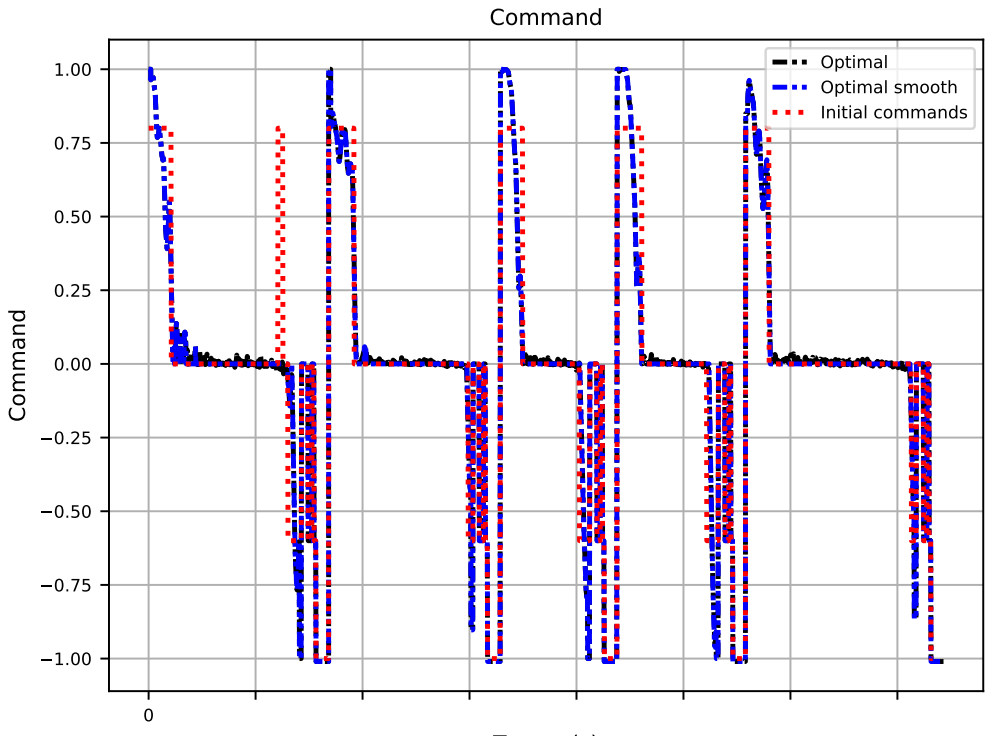

Figure C.9: Optimal driver's command (black), smoothed optimal driver's command (blue), and initial driver's command (red) depending on time.

The optimal solution is plotted in black. The smoothed optimal driver's command (in blue) is directly computed from the optimal driver's command. When the optimal driver's command is close to zero, we set it to zero to limit the effect of noise, without impacting the solution too much. We can observe that the optimal solution is relatively close to the initial solution on the second part of the journey. This means that the algorithm used by the SNCF is already efficient. On the other hand, the optimal solution is very different from the initial solution on the first section. We can observe the impact of these modifications on the train dynamics in Figure C.10.

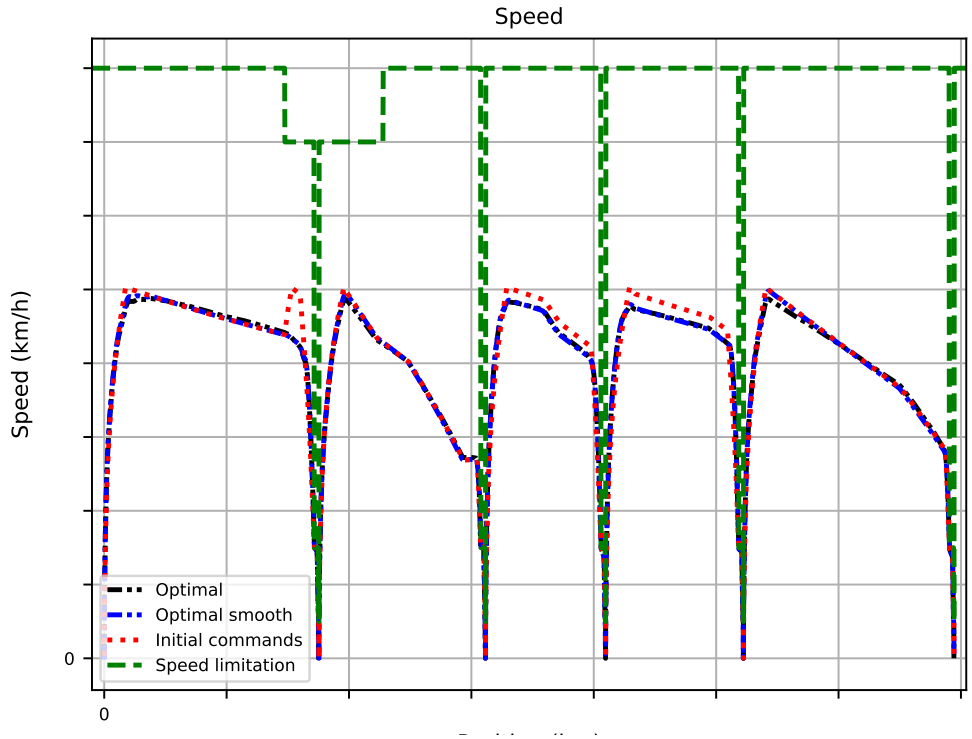

Figure C.10: Speed profile associated with the optimal driver's command (black), the smoothed optimal driver's command (blue), and the initial driver's command (red) depending on the position. The speed limitation is in green.

Undoubtedly, the speed profile associated with the optimal driver's command and with the initial driver's command are very close on the second part of the journey. However, on the first part, the initial driver's command proposes to use motors a second time (just before final brake). This proposal does not seem to be interesting from the energy saving point of view and the optimization algorithm has found a different optimal solution. The maximum speed is slightly reduced on the part three and four of the journey. This is possible due to the fact that the driver's command is higher at the beginning of these sections (see Figure C.9). Finally, the energy consumption is shown in Figure C.11.

We can notice that the optimal solution reduces the energy consumption by about 7% compared to the driver's command provided by the previous algorithm. This reduction reaches about 8% with the smoothed driver's command. The major difference is in the first part of the journey. The energy consumptions are equivalent in the second part of the journey. Despite the small number of degrees of freedom, the algorithm manages to reduce the energy consumption.

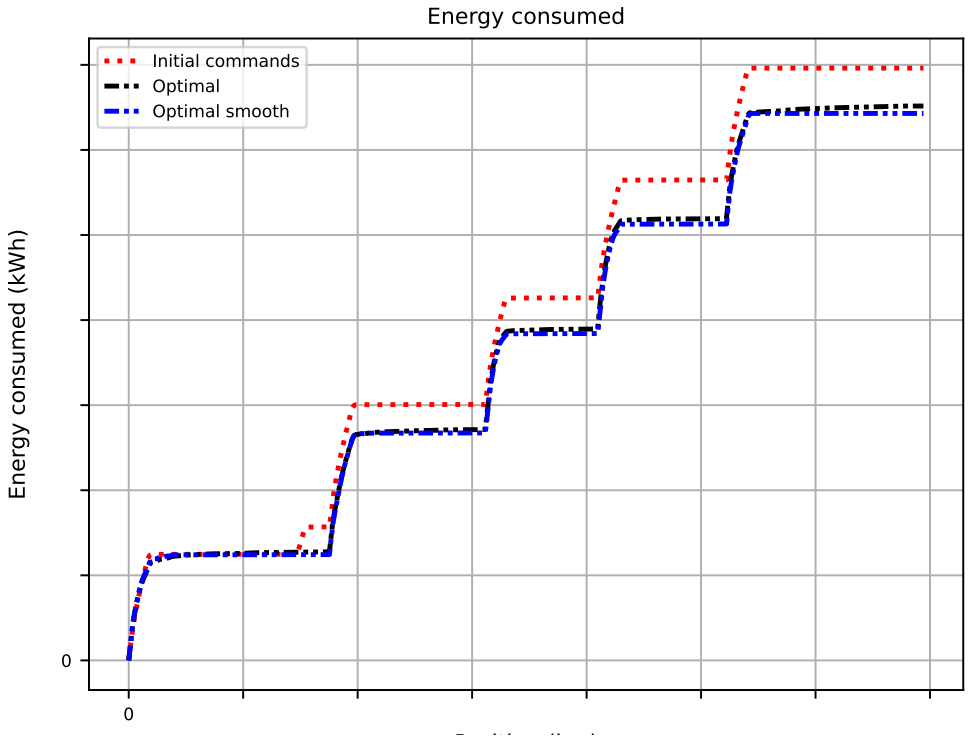

Figure C.11: Energy consumption associated with the optimal driver's command (black), the smoothed optimal driver's command (blue), and the initial driver's command (red) depending on the position.