Universitá degli Studi di Ferrara

Dottorato di Ricerca in

Scienze dell'Ingegneria

Ciclo XIX

Coordinatore: Prof. Stefano Trillo

STOCHASTIC AND DETERMINISTIC SIMULATION TECHNIQUES FOR TRAFFIC AND ECONOMICS

Settore Scientifico Disciplinare MAT/08

Dottorando:
Dott. Piero Foscari W. R.

Tutore:
Prof. Lorenzo Pareschi

Anni 2004/2006

Aknowledgments

Many people have helped me during these years of studies and work.
First of all I would like to express my gratitude to Prof. Lorenzo Pareschi for the stimulating environment he could provide, the interesting topics and ideas suggested, his guidance and support also on all practical issues due to my work that ran in parallel with the PhD studies.

My thanks are surely due to all my colleagues at the Center for Modeling, Computing and Statistics in Ferrara, where a significant part of my investigations was carried out, not only for sharing the office, but also the good times, and for making the everyday work pleasant.

I am grateful to Prof. Michael Mascagni for the research topic he suggested which soon evolved both into practical results and professional value.

I am also indebted to my employers and colleagues, without their support my ongoing parallel research would have not been fulfilled. Firstly at Eurizon Capital I was encouraged in my activity on random number generation. Moreover in particular NEC Labs Europe have significantly supported me in the final phase.

Finally the writing of this dissertation would not have been possible without the continuous support and encouragement of my family, of many friends and in particular of Benedetta.

Contents

Abstract

In this work I present the result of different investigations conducted in the last years in the context of stochastic modeling for decision making in the areas of traffic simulation and economics.

Traffic simulation has seen us from the Center for Modeling, Computing and Statistics involved in a project for the evaluation and planning of two highway stretches in the area around Ferrara. In particular we conducted the modeling and numerical simulation of the highway network, in collaboration with Michael Herty.

Later the study of kinetic analysis and simulation techniques proved useful in another related setting, that is agent based models in economics, a discipline of growing importance in understanding the workings of markets, be they financial or centered on tangible goods.

Due to my job in the asset management industry some of the research activity has been tilted towards practical methods for financial simulations, and in particular that of parallel random number generation is a topic that has been gaining importance during these last years. While at Eurizon Capital I developed a novel fast algorithm for moving over certain widely used random number streams, and at NEC Labs Europe this was further reimplemented as a core block of a professional C++ library for parallel Monte Carlo simulation in finance.

Finally I present a small note on a common numerical artifact arising in Monte Carlo simulations when only a limited number of kinetic particles are used. Already with simple kernels the resulting probability distributions differ significantly from those predicted by theory and obtained with large particle sets.

Abstract

In questo lavoro di tesi sono presentati i risultati ottenuti in una serie di studi condotti negli ultimi anni nell'ambito della modellazione stocastica nei campi delle simulazioni sia per il traffico stradale che per l'economia.

Le simulazioni per il traffico hanno visto noi del Centro per la Modellistica, il Calcolo e la Statistica (CMCS) coinvolti in un progetto per la valutazione e la pianificazione di due tratte autostradali intorno a Ferrara. In particolare, in questo lavoro abbiamo effettuato modellazione e simulazioni numeriche della rete autostradale, in collaborazione con Michael Herty.

Piú tardi, lo studio delle tecniche di simulazione e della teoria cinetica si é dimostrato utile anche in un campo affine, come i modelli agent based in economia, una disciplina la cui importanza é cresciuta per via della necessitá di capire il funzionamento dei mercati, sia finanziari che di beni tangibili.

Per via del mio lavoro nel campo della gestione di patrimoni, una parte della attivitá di ricerca é stata condotta su metodi pratici per simulazioni finanziarie. In particolare, la generazione parallela di numeri casuali é un settore che sta guadagnando rapidamente importanza negli ultimi tempi. Durante gli anni trascorsi in Eurizon Capital abbiamo sviluppato un nuovo algoritmo veloce che consente di muoversi su diverse sequenze di numeri pseudocasuali comunemente utilizzate, mentre alla NEC Labs Europe questo tema é stato ulteriormente sviluppato, reimplementandolo quale funzionalitá di base di una libreria professionale in C++ per la simulazione parallela Monte Carlo in finanza.

Nella parte finale di questo lavoro presentiamo una nota su un comune artefatto numerico che si presenta nelle simulazioni Monte Carlo solo quando si utilizza un numero limitato di particelle. Infatti, anche con semplici kernel, le distribuzioni risultanti differiscono significativamente da quelle predette dalla teoria ed ottenute con un set esteso di particelle.

Chapter 1

Traffic modeling

1.1 Introduction

Research in modeling of traffic flows goes back to 1955 with the pioneering work of Lighthill and Whitham (see [12]). Traffic modeling originates from previous research in different fields especially of physics, and has also seen subsequent reuse of its techniques in other areas such as blood, pedestrian and information network flows.

Nowadays its importance is growing, given the possibility of congestion forecasting, the developements in traffic control and the usefulness in the planning phase for infrastructures.

Traffic data is usually collected by fixed induction loops or highway toll gates. Induction loops are devices coming in couples measuring speed of vehicles traversing them in sequence (therefore standing vehicles will not be detected).

There are at least three different modes for traffic flow (or phases when adopting a statistical mechanics point of view):

1. *Free flow* - Where distance between vehicles is high enough for interactions to be negligible. Therefore each car can proceed at the desired speed, which is often assumed to be tending to the maximum allowed, eventually further modulated by road conditions.

2. *Wide moving jams* - Large slowdowns where a high density is reached and a common velocity is shared.

3. *Stop and go waves* - These are a phenomenon arising in the previous context, they propagate backwards with respect to the traffic direction, and are caused

by the driver behaviour amplifying the braking strength induced by the driver in front. Often more of such waves follow one another, hence the name suggesting the repeater acceleration and braking. This phenomenon occurs more easily in situations of dense traffic and accounts for stops on highways even in apparent lack of a clear cause.

4. *Synchronised flow* - This is also a jammed state, where there is no definite velocity/density relationship anymore: here different speeds arise in nearby areas of similar density. The name comes from the synchronization effect that was observed between lanes, especially on-ramps influencing -and being influenced by- the main flow.

Any satisfactory model should at least be able to reproduce all these penomena when put in the corresponding adequate initial condition.

Modeling of traffic dynamics can be roughly divided into three main categories according to the different level of detail reached:

- *Microscopic modeling.* The most natural approach is direct simulation of each individual car and its response to neighbouring vehicles, this is the realm of *microscopic modeling*, also sometimes called *car following* or *follow the leader* models.

- *Macroscopic modeling.* The opposite view operates at the fluid dynamic level on macroscopic quantities like local densities and average speeds.

- *Kinetic modeling.* An intermediate approach uses Boltzmann-like kinetic equations to reach a good compromise between the computational efficiency of macroscopic models and accuracy of description inherent to car following methods.

Modeling in these three different worlds must be consistent, and many such connections have been worked out, i.e. in the form of hydrodinamic limits linking kinetic models to macroscopic ones.

In mixed modeling these three different approaches naturally lend themselves to treat specific areas where different level of detail is required: macroscopic evolution can be used for highways and freeways, mesoscopic for ordinary roads and microscopic for urban areas.
Models of the car-following and kinetic type are usually also able to handly stocasticity without big efforts.

1.1.1 Characteristics of traffic and the fundamental diagram

Elementary behaviour and constraints for vehicle speeds can be derived from the car density: clearly speed will be maximum in free roads and will tend to decrease at increasing densities, till vanishing at the maximum density determined by a null bumper to bumper distance.

The so called fundamental diagram describes the relation existing between vehicle density $\rho(x,t)$ and flux $f(x,t) = \rho(x,t)\, v(x,t)$. In theory only density should be constrained in $[0, \rho_{\max}]$ while speed could vary (according to the available braking speed), however empirically in certain cases there is a stricter dependence between the two, with actual speed adjusting towards an optimal free flow speed for a certain given density; in fact such speed is influenced by the braking distance available which should match that consented by actual velocity. Consequentely also the flux will be mostly determined by density only, so that it is common practice to describe flux by a fitted relation of the form $q = f(\rho; x, t)$. Furthermore often the space and time homogeneous case is considered so that the corresponding variables can be dropped as in the following.

Such function is usually taken so as to satisfy the following criteria:

1. $f(\rho < f(\sigma))$ for $\rho \in [0, \rho_{\max}]$ σ with $\sigma \in (0, \rho_{\max})$
2. $f'(\rho) \geq 0$ for $\rho \in [0, \sigma)$
3. $f'(\rho) \leq 0$ for $\rho \in (\sigma, \rho_{\max}]$
4. $f(\rho)$ is concave
5. $f(0) = f(\rho_{\max}) = 0$

Clearly the vanishing in (5) is determined by the vanishing of density at 0 and speed at $\rho_{\max}$. The shape determined by the first four points is related to the desired behaviour of velocity; these arise both from the observed empirical distribution and from theoretical considerations in the conservation laws where this relationship is used.

From the shape of $f(\rho)$ one sees that density is a more informative measure than flux, which cannot discriminate between situations of light free flow and heavy congestion. Unfortunately often the flux is the only measured data so that one is forced to integrate this information somehow.

The relation between density and flux breaks down in the case of synchronised traffic, located in the higher part of the density spectrum, so that models making use of a simplifying functional relation like the one above cannot fully capture reality.

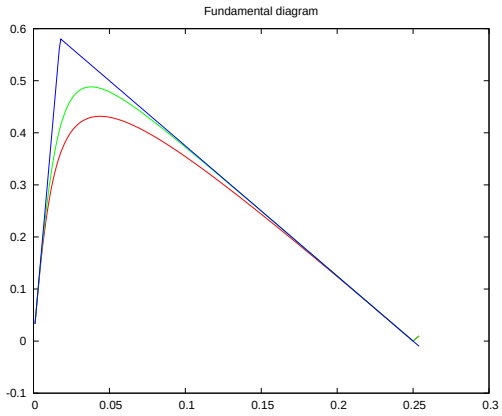


Figure 1.1: *Three different flux/density relations on the fundamental diagram*

1.2 Microscopic modeling

Microscopic modeling is aimed at explicitely simulating the dynamics of each individual vehicle in the system. Here the response of each driver to its predecessor enters directly a second order ordinary differential equation governing motion over time, according to Newton's law.

Therefore atleast the following quantities are tracked for each car numbered j:

$$x_j(t) \in \mathbb{R}, \quad v_j(t) \in [0, v_{max}], \quad t \in \mathbb{R}^+_\nu$$

The acceleration $\dot{v}_j(x_\cdot, v_\cdot, t)$ will thus incorporate all components of driver's behaviour. Often realistic models require a large amount of parameters (even 50) and are therefore of difficult calibration, while simpler models are easy to setup but can only reproduce well global phenomena, not the individual car dynamics.

This driver's perspective represents the so called *Lagrangian* point of view, as opposed to the somewhat "dual" *Eulerian* dynamics: in the former the position and speed are tracked for each given moving particle, while in the latter it is the position being fixed while the corresponding speed and particles (density) are observed.

Note that at variance with gas-dynamics most microscopic traffic models can be applied up to the macroscopic scale (on the whole simulation range) since their computational cost will be high but still affordable (cars on roads are of course fewer than the particles in a gas by several orders of magnitude).

1.2.1 Bando-Hasebe-Nakayama-Shibata-Sugiyama model

In the BHNSS model (presented in 1995 in [2]) the acceleration acts directly adjusting velocity towards a desired value depending on headway distance:

$$\dot{v}_i = a_i \left[V(s_i) - v_i \right]$$

where s_i is the distance to the leading car, and the fixed legal velocity $V(s_i)$ function is monotone and increasing with $V(0) = 0$ and $\lim_{s_i \to \infty} V(s_i) = V^{max}$. A suggested possibility for such a legal velocity function is $V(s_i) = \tanh(s_i - 2) + \tanh 2$.

Clearly to small headway distances must correspond small target velocities, so that the second part of the expression becomes dominant causing braking. On the other hand free space on the road brings target velocities possibly larger than the current one v_i, a difference which will determine a positive acceleration.

It is easy to determine the equilibrium speed at which traffic with a uniform density (and with cars with the same acceleration coefficient $a_i = a$) can advance without changes: $v^e = V(\rho^{-1})$. This immediately gives an explicit solution for uniform free flow:

$$\begin{cases} v_i(t) = v^e \\ x_i(t) = x^0 + i\rho^{-1} + V(\rho^{-1})t \end{cases} \tag{1.1}$$

This is sometimes called a *laminar* state because of its parallel bundle appearance when plotted over time.

Whenever the common initial velocity is different than the equilibrium speed implied by the given density a similar dynamics will develop, but not linear anymore: all cars will advance at a common time varying velocity approaching the implied equilibrium one.

$$\begin{cases} \dot{v}_i(t) = \int a(v^e - v_i(t))dt \\ x_i(t) = x^0 + i\rho^{-1} + \int v_i dt \end{cases} \tag{1.2}$$

Any small deviations from the initial data for such a free flow are amplified with time by the model evolution and lead to the formation of congested areas and more rarified ones with a corresponding faster free flow, with cars entering and exiting from a small interface area between these two regimes. Furthermore such congestions behave "well" by moving backwards with respect to the traffic direction, at a speed of easy calculation: $[V(s_{max})s_{min} - V(s_{min})s_{max}]/(s_{max} - s_{min})$.

In particular the shape and slope of the function $V(s_i)$ determines the the values

s_{min} and s_{max} which behave as "attractors" for the car distance in congested and flowing clusters, respectively. Whenever $V'(s_i) > a/2$ the behaviour is unstable, stable otherwise. For instance, with the given form of $V(s_i)$ three areas emerges: two stables for small and large densities, and and intermediate unstable one.

In this model the fluxes for these two congested and free flow laminar states are equal: the slower speed in jams is exactly compensated by the increased density.

Noteworthy is the fact that no accidents occur, as congestion developes before any of them. Moreover no negative velocities ever appear, even though no direct limitation on speed might appear evident at first: this happens because braking is managed through a decay of velocity, instead of an independent deceleration term.

1.2.2 Treiber-Helbig model

The following Intelligent Driver Model - presented by Treiber and Helbing in [16] - takes a compromise in the number of parameters controlling the driver behaviour, using seven of them for reaching enough flexibility but without risking overfitting of the empirical traffic features. The acceleration term reads:

$$\dot{v}_i = a_i \left[1 - \left(\frac{v_i}{v_i^0} \right)^{\gamma} - \left(\frac{s_i^*}{s_i} \right)^2 \right]$$

where s_i is again the distance to the leading car, $\gamma \geq 1$ (often $\gamma = 4$) and:

$$s_i^* \left(v_i, \Delta v_i \right) = s_i^0 + v_i \max \left(T_i - \frac{\Delta v_i}{\sqrt{4 a_i b_i}}, 0 \right)$$

with v_i^0 as desired speed, Δv_i the relative speed of preceeding vehicle, T_i is the safe time headway in moving congested traffic (that is, the allowed reaction time), a_i maximum acceleration, b_i maximum braking, s_i^0 minimum jam distance.

The acceleration term $a_i \left(1 - \left(\frac{v_i}{v_i^0} \right)^{\gamma} \right)$ represents the acceleration on a free road tending to reach the desired speed v_i^0. Here the exponent γ usually lies between 1 and 5, and is used to modulate between a constant acceleration to the desired speed (for $\gamma \to \infty$) and and exponential acceleration dynamic (when $\gamma = 1$).

The term s_i^* models a "desired minimum gap" from the leading car, so that the deceleration term $- \left(\frac{s_i^*}{s_i} \right)^2$ describes braking intensity, which grows in a Coulomb-repulsion-like fashion as the actual gap s_j shrinks.

In the case of identical drivers and when traffic is in equilibrium (so that $\Delta v_i =$

$\dot{v}_i = 0$) all vehicles assume a corresponding velocity dependent equilibrium gap $s^e(v_i)$ like:

$$s^e(v_i) = s^*(v, 0) \left[1 - \left(\frac{v}{v_0} \right)^\gamma \right]$$

Further for the special subcase of homogeneous jams when $v << v_0$ the equilibrium gap approaches the desired gap: $s^e(v) \approx s_0 + s_1 \sqrt{v/v_0} + vT$. For some parameter values it is possible to derive an analytical value for the corresponding speed.

This model possesses several advantages compared to the rest of the literature; in particular all of its parameters are of practical significance and can be easily measured, while a calibration procedure can fit an empirical fundamental diagram; furthermore there are corresponding mesoscopic and macroscopic version. A very appealing feature is the realistic flow density relation which, while well determined in equilibrium, shows the scattering observed empirically at medium densities, even without the need of mixing vehicles with different characteristics.

Simulation

To evaluate the behaviour of this model we developed a standard C implementation for handling different types of vehicles on a network. A direct solver has been implemented for the two ordinary differential equations involved; the IDM model proved to be quite reliable in practice for not too small time steps.

In figure 1.2 results are shown for the different simulated vehicle dynamics (here being speed and acceleration over time), to investigate the effect of varying parameter values. The red trajectory is the reference one, while the others are for alternative driver behaviours.

In this case the scenario is that of a vehicle initially standing still and having a free road in front, to observe the acceleration phase leading to the desired maximum velocity; this is shown in the left part of the graph. This works as expected with acceleration smoothly decreasing until the desired speed is reached.

On the right part of the graph the driver approaches a still obstacle and reacts accordingly with a smooth braking phase until the car is halted. This is ofcourse the case having the biggest influence on the numerical behaviour of the discretization scheme: a too large timestepping might let the car ”jump” over the obstacle, skipping the singularity in the braking term when distance tends to zero.

Changing the model parameters shows different behaviours: the green trajectory corresponds to a higher value of the acceleration constant a and clearly provokes a faster arrival at the desired speed; however it also has influence on the braking so that

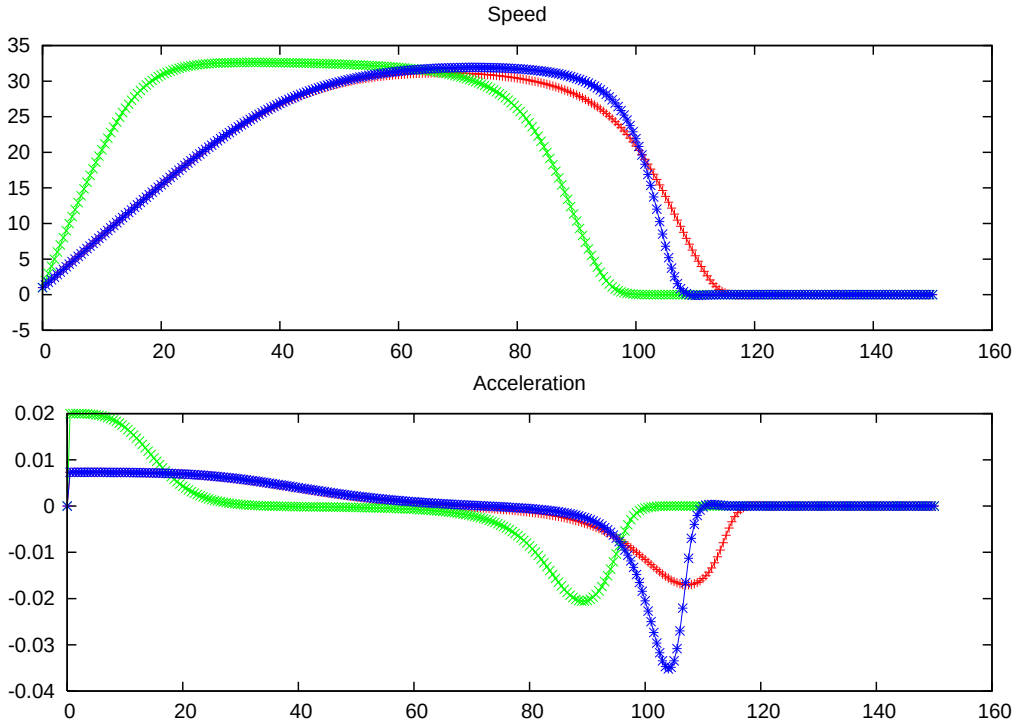


Figure 1.2: *Speed and acceleration for the IDM model*

null speed is reached before. On the other hand the braking parameter b (increased in the blue graph) has a different behaviour: it also makes braking more aggressive, but implicitly allows for it to start later in time, whereas increasing a anticipated it.

Figure 1.3 shows similar profiles in the case of a preceeding car advancing at constant velocity, in place of the previous still obstacle.

1.2.3 Nagel-Schreckenberg model

The Nagel-Schreckenberg model is an example of the so called Cellular Automata. These are update rules operating on a set of connected cells, each having a multi-valued state; usually each cell state is affected just by its neighbours allowing for very efficient implementation, and are also related to Lattice Boltzmann methods in rarified gas theory. Despite their simplicity they can exhibit arbitrarily complex behaviour even for small states (especially famous is the discrete *Rule 101* studied by Wolfram).

This model was introduced and received significant attention because of its high computational efficiency and its ease of implementation. The Swiss road network

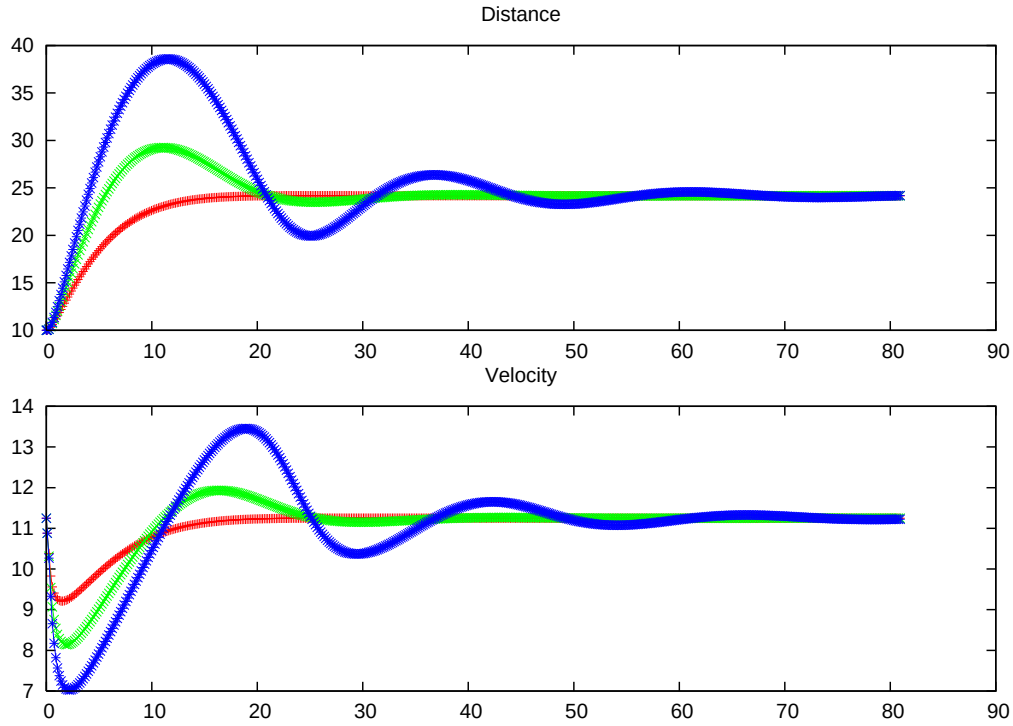


Figure 1.3: *Relaxation to constant velocity for the IDM model*

was one of the first large scale simulations actually performed, and was based on such models.

In the case of the N-S model cells result from a regular discretization of a road, with size corresponding to the front-bumper to front-bumper distance between two cars in the densest jammed state, usually taken as $7,5m$; implicitly this will also induce a discretization in possible velocities once time is discretized too. A cell can just contain one single or no cars; each car will posses an internal state describing velocity v_j in terms of cells traveled per time step. At each time step positions and velocities of all cars are updated in four substeps accounting for acceleration, braking, randomization and advancing:

1. *Acceleration*: $v_j^1(t) = min(v_j(t) + 1, v_{max})$
2. *Braking*: $v_j^2(t) = min(d_j(t) + 1, v_j^1(t))$
3. *Randomize*: $v_j(t+1) = max(v_j^2(t) - 1, 0)$ with probability p
4. *Advance*: car j is moved $v_j(t+1)$ cells forward

At the beginning acceleration accounts for the desire of drivers to reach the maximum possible speed, thus without exceeding the limit.
In the second step deceleration is introduced whenever the distance d_j from the preceeding vehicle is lower than what would be covered in a time step. This guarantees that collisions cannot occur in the following time step.
Later randomization occurs to account for various important phenomena. Asymmetry is introduced to delay acceleration and enhance braking, and the overreaction of drivers to decelerations of the preceeding vehicle is thus captured, thus allowing for stop and go phenomena. Moreover randomized braking accounts for irregularities in driving.
The fourth step finally performs the actual advancing of all the cars.

This set of four rules is minimal for producing a realistic behaviour, and their order is also crucial in obtaining the desired result.

1.3 Mesoscopic modeling

The mesoscopic level of modeling is intermediate between the generating microscopic interaction rules and the macroscopic description of traffic quantities. These models do not track individual vehicles anymore, however they still model the driver behaviour, often probabilistically.

In such a setting a kinetic vehicle density $f(x, v, t)$ is evolved, where x denotes the position on a highway, v a speed in the interval $[0, v_{max}]$ that could be normalised, and t is time. Because of their statistical nature they have been especially advocated for multilane traffic analysis, while macroscopic approaches could be enough for the single lane case. However the case treated in the following is that of a single lane, while the multilane one would be a natural extention without substantial differences.

As usual in kinetic theory it's possible to derive macroscopic quantities from the mesoscopic density moments; in particular the vehicle density:

$$\rho(x,t) := \int_0^{v_{\max}} f(x,v,t)dv$$

and the flux:

$$q(x,t) := \int_0^{v_{\max}} vf(x,v,t)dv$$

and the corresponding average velocity $u(x,t) := u(x,t)/\rho(x,t)$. Of interest are sometimes also higher order moments like average kinetic energy and velocity vari-

ance.

One advantage over the microscopic case is the possibility of achieving analytic results.

The kinetic approach to traffic modeling has been introduced by Prigogine and Andrews in 1960 (see [14]). This was later improved by Paveri and Fontana [13]. Since then a large number of models has been presented.

Herty et al introduced in [4] a convenient framework for the derivation of many kinetic evolution models starting from microscopic interaction models, which will be described in the following.

1.3.1 Kinetic traffic models

In general the kinetic form for traffic flow models is given by a PDE for the evolution of f of the following form:

$$\partial_t f + v\partial_x f = G(f, f^2, x, x + H_G, t, v) - L(f, f^2, x, x + H_L, t, v)$$

Herein increases and decreases in f are given explicitly by the two gain and loss terms G and L, which usually accound for acceleration and braking. When interaction occurs among couples of cars only - as is often the case - the source term depends on $f^2(x, t, H, v, v')$. Additionally for reactions to movements ahead along the road one can add non-local terms $x + H_{L,G}$. One used simplifyng approximation is $f^2 \approx qf(x, v)f(x + H, v')$ where q is a given function.

Such mesoscopic description will be obtained from a generic microscopic model which we now describe. As in the microscopic examples previously described one can directly introduce terms for the different components of driver behaviour, namely:

1. an alignment to a desired maximum velocity v_{max}

2. acceleration coming from two-car interactions, when the second one travels faster

3. braking coming from two-car interactions, when the second one is slower

All such terms can include stochastic components to account for variability in the driver behaviours. Moreover it is also natural to initially consider just the space-homogeneous case. The resulting microscopic interaction rule is then the composition of the following updates, for free flow alignment, acceleration and braking

respectively:

$$v' = v + d(v, v_{max}, \xi)$$
$$v' = v + a(v, w, \xi) \quad v < w$$
$$v' = v - b(v, w, \xi) \quad v > w$$

In this form it is clear how all three rhs terms are related to acceleration of vehicles, performed istantaneously. The uniform random variable ξ accounts for stochasticity. In such a microscopic formulation the time variable is implicit in the iteration procedure so does not appear explicitely as a variable. The update $d(v, \xi)$ does not depend on nearby vehicles and simply adjust speed towards the maximum velocity characteristic of free flow. The remaining two terms account for interaction with the "preceeding" car having velocity w, and thus increase velocity v for a positive difference $w - v$, and brake lowering v in the other case of negative speed difference and thus of decreasing distance.

Most microscopic models can be written in this form, and a few examples will be detailed in the following with the corresponding macroscopic derivations; furthermore this framework also covers the Intelligent Driver Model by Treiber and Helbing described earlier, for which an example of kinetic densities is shown in figure 1.6.

Illner, Klar et al model

This model was studied and presented in [6] and [9]. In this setting no two-car interactions are present and the desired velocity forcing is incorporated into the acceleration term as follows:

1. $d(v, \xi) = 0$
2. $a(v, w, \xi) = (1 - \xi)(v_{max} - v)$
3. $b(v, w, \xi) = \xi v$

Simplified model by Klar et al

This model, presented in [3], operates on the relative velocities among pairs of cars:

1. $d(v, \xi) = \xi v_{max}$
2. $a(v, w, \xi) = (1 - \xi)(w - v)$

3. $b(v, w, \xi) = \xi(v - w)$

Also in [3] closures are found to derive macroscopic equations corresponding to this model.

Other than these two presented cases also a Helbing-like model and other microscopic ones can be restated in a kinetic form.

1.3.2 Derivation of kinetic equations

From the description of driver interactions seen earlier the equations for the evolution of kinetic density can be derived, which in the space homogeneous case reduces to:

$$\partial_t f = G(f, t, v) - L(f, t, v)$$

Let's decompose the gain and loss terms in three components corresponding to each interaction term (while fixing t for shortness):

$$G(f, v) = G_d(f, v) + G_a(f, v) + G_b(f, v)$$

$$L(f, v) = L_d(f, v) + L_a(f, v) + L_b(f, v)$$

The loss subterms are of immediate calculation:

$$L_d(f, v) = f(v) \int \beta^d_{v \mapsto v'} dv' = f(v) \tag{1.3}$$

$$L_a(f, v) = \int_{v<w} \beta^a_{(v,w)\mapsto(v',w)} f(v) f(w) dw \tag{1.4}$$

$$L_b(f, v) = \int_{v>w} \beta^b_{(v,w)\mapsto(v',w)} f(v) f(w) dw \tag{1.5}$$

$$\tag{1.6}$$

Here the three functions β^d, β^a and β^b are composed by the previously introduced correlation function q and further parameters describing the strength of the microscopic interaction.

The gain term is decomposed as follows:

$$
\begin{aligned}
G_d(f,v) &= \int_{N_d(v)} |\det J_d| \beta^d_{v' \mapsto v} f(v') d\xi \\
G_a(f,v) &= \int_0^{v_{\max}} \int_{N_b(v,w)} |\det J_a| \beta^a_{(v',w)\mapsto(v,w)} f(v') f(w) d\xi dw \\
G_b(f,v) &= \int_0^{v_{\max}} \int_{N_b(v,w)} |\det J_b| \beta^b_{(v',w)\mapsto(v,w)} f(v') f(w) d\xi dw
\end{aligned}
$$

where the integrals are defined over the following sets:

$$
\begin{aligned}
N_d(v) &:= \{\xi : 0 \le \xi \le 1, 0 \le v' \le v_{\max}, v = d(v', \xi)\} \\
N_a(v,w) &:= \{\xi : 0 \le \xi \le 1, v' < w, 0 \le v' \le v_{\max}, v = v' + a(v', w, \xi)\} \\
N_b(v,w) &:= \{\xi : 0 \le \xi \le 1, v' > w, 0 \le v' \le v_{\max}, v = v' - b(v', w, \xi)\}
\end{aligned}
$$

and the Jacobians $J^{\cdot}$ are proper of the three respective $(v,w) \mapsto (v',w)$ transformations: $v = d(v', \xi)$, $v = v' + a(\xi, v', w)$ and $v = v' - b(\xi, v', w)$. Here v, w and ξ are the parameters implicitly determining v'. Whenever the interactions d, a and b are given then the previous formulas get simplified, so that all terms can be expressed in terms of v' and w.

For the previously introduced models this procedure results in the two following kinetic equations, with normalized maximum velocity $v_{\max} = 1$.

Illner, Klar et al.

For acceleration, from $G_a(f,v)$ and:

$$a(v,w,\xi) = (1-\xi)(1-v)$$

$$\beta^a(v',w) = |v'-w| q_a(\rho)/\rho$$

where q_a is a correlation function depending solely on macroscopic quantities, one obtains:

$$G_a(f,v) = \frac{1}{\rho} \int_0^1 \int_{v'<w, 0\le v' \le 1} q_a(\rho) |v'-w| f(v') f(w) \frac{1}{1-v'} \chi_{[v',1]}(v) dv' dw.$$

Analogously, for braking:

$$b(v,w,\xi) = \xi v$$

$$\beta^b(v', w) = |v' - w| q_b(\rho)/\rho$$

$$\begin{aligned} G_b(f, v) &= \int_0^1 \int_{M(v,w)} \frac{1}{v'} \beta^b(v', w) f(v') f(w) dv' dw \quad (1.7) \\ &= \int_0^1 \int_{v'>w, 0\le v'\le 1} \frac{1}{v'} \chi_{[0,v']}(v) \beta^b(v', w) f(v') f(w) dv' dw \quad (1.8) \end{aligned}$$

Inserting β^b gives the braking gain term. One can proceed similarly for the loss terms. Assuming further $\forall \rho : q^a(\rho) \neq 0$ and setting:

$$k(\rho) := \frac{q^b(\rho)}{q^a(\rho)}, \quad c(\rho) = q^a(\rho)/\rho$$

one obtains:

$$\begin{aligned} f_t(v,t) &= c(\rho)\Big(k(\rho) \int_0^1 \int_{v'>w} |v' - w| f(v', t) f(w, t) \frac{1}{v'} \chi_{[0,v']}(v) dv' dw \\ &+ \int_0^1 \int_{v'<w} |v' - w| f(v', t) f(w, t) \frac{1}{1 - v'} \chi_{[v',1]}(v) dv' dw \\ &- k(\rho) \int_{v>w} |w - v| f(w, t) f(v, t) dw \\ &- \int_{v<w} |w - v| f(w, t) f(v, t) dw \Big) \\ &=: c\left(k\tilde{G}_B + \tilde{G}_A - k\tilde{L}_B - \tilde{L}_A\right) \quad (1.9) \end{aligned}$$

Klar

Here, analogously to the previous case, for the acceleration term one again calculates $G_a(f, v)$ through:

$$a(v, w, \xi) = (1 - \xi)(w - v)$$

$$\beta^a(v', w) = |v' - w| q_a(\rho)/\rho$$

while for braking:

$$b(v, w, \xi) = \xi(v - w)$$

$$\beta^b(v', w) = |v' - w| q_b(\rho)/\rho$$

Moreover the free acceleration term in this case is:

$$d(v,\xi) = \xi = v'$$

so that $|\det J| = 1$ and $G_d(f,v) = \int f(v')dv'$. The derivation gives the following kinetic equation:

$$\begin{aligned} f_t(v,t) &= c(\rho)\Big(k(\rho)\int_0^1\int_{v'>w}|v'-w|f(v',t)f(w,t)\frac{1}{v'-w}\chi_{[w,v']}(v)dv'dw \\ &+ \int_0^1\int_{v'<w}|v'-w|f(v',t)f(w,t)\frac{1}{w-v'}\chi_{[v',w]}(v)dv'dw \\ &- k(\rho)\int_{v>w}|w-v|f(w,t)f(v,t)dw \\ &- \int_{v<w}|w-v|f(w,t)f(v,t)dw\Big) \\ &=: c\left(k\tilde{G}_b + \tilde{G}_a - k\tilde{L}_b - \tilde{L}_a + 1/c\tilde{G}_d - 1/c\tilde{L}_d\right) \end{aligned} \tag{1.10}$$

1.3.3 A Monte Carlo simulation method

Some simplifying assumptions can be used to obtain a Monte Carlo simulation algorithm in the previously described framework. The two acceleration and braking kernels are then derived from a common one in the following way:

$$\beta_a(v,w) = \begin{cases} \beta(v,w) & v < w \\ 0 & v \geq w \end{cases} \tag{1.11}$$

$$\beta_b(v,w) = \begin{cases} 0 & v \leq w \\ k\beta(v,w) & v > w \end{cases} \tag{1.12}$$

One can split the time dependent equation - rewritten to expose the common kernel so that integrals are carried over the whole velocity range:

$$f_t = kG_b + G_a - kL_b - L_a + G_d - L_d$$

to get:

$$f_t = kG_b + G_a - kL_b - L_a$$

$$f_t = G_d - L_d$$

Then, by rewriting the gain and loss operators:

$$G_a = \int\int \beta^a(v', w) f(v') f(w) d(v', w) \tag{1.13}$$

$$= \int\int \min\{\Sigma, \beta^a(v', w)\} f(v') f(w) d(v', w), \tag{1.14}$$

$$L_a = f(v) \int \beta^a(v, w) f(w) dw \tag{1.15}$$

$$= f(v)\Sigma\rho - f(v) \int \left(\Sigma - \min\{\Sigma, \beta^a(v', w)\}\right) f(w) dw \tag{1.16}$$

$$= \Sigma\rho(f(v) - \tilde{L_a}), \tag{1.17}$$

with a suitable constant $\Sigma > 0$ (which in rarefied gas simulations is usually referred to as *dummy cross section*). The same procedure is then carried for the braking terms.

Finally, applying an explicit Euler method for advancing in time, one obtains:

$$f^{n+1} = f^n(1 - \mu\Delta t) + \mu\Delta t \left(\frac{k}{k+1}(\tilde{G}_b + \tilde{L}_b) + \frac{1}{k+1}(\tilde{G}_a + \tilde{L}_a) \right), \tag{1.18}$$

whith $\mu = \Sigma\rho(k+1)$ and $\tilde{G}_{a,b} = G_{a,b}/(\Sigma\rho)$. The weighting of the different operators can be reinterpreted in probabilistic terms, as usual with kinetic operators:

Probability	Event
$1 - \mu\Delta t$	Speed is mantained
$\frac{\mu\Delta t}{k+1}$	If $v < w$ with probability $\beta(v, w)/\Sigma$ acceleration occurs
$\frac{k\mu\Delta t}{k+1}$	If $v > w$ with probability $\beta(v, w)/\Sigma$ braking occurs

Later the explicit Euler method is applied for the remaining terms:

$$f^{n+1} = f^n(1 - \Delta t) + \Delta t G_D. \tag{1.19}$$

In this case for $\Delta t < 1$ with probability Δt the speed of a car changes is adjusted

towards the desired one.

The corresponding algorithm can be sketched in this way:

- Select N particles from the initial distribution $f_0(v)$ and a timestep Δt
- Divide the particles randomly into two groups, one of $N_1 = N(1 - \Delta t)N$ non interacting particles and $N_c = N - N_1$ interacting ones, and build couples (v_i, v_j) out of the latter group.
- Calculate Σ
- For each interacting pair draw a uniform random number ξ_1 in $[0, 1]$ to choose between acceleration and braking, according to $\xi_1 < 1/(k + 1)$; then calculate $\Sigma_{ij} := \min\{\Sigma, \beta(v_i, v_j)\}$ and draw two further uniforms $\xi_2, 3$

  For acceleration if $v_i < v_j$ and $\xi_2\Sigma < \Sigma_{ij}$ update according to:

  $v_i' = v_i - a(v_i, v_j, \xi_3)$

  $v_j' = v_j$

  For braking, if $v_i \geq v_j$ and $\xi_2\Sigma < \Sigma_{ij}$ update according to:

  $v_i' = v_i - b(v_i, v_j, \xi_3)$

  $v_j' = v_j$
- Perform the free flow update:

  Again subdivide particles among $N_c := \Delta t N$ affected ones and $N - N_c$ unchanged, and perform the updates $v' = d(v, \xi)$ corresponding to G_d
- If stationarity is not reached repeat from the second step

Figure 1.4 shows the stationary solution for the Illner et al. model. for three different values of the parameter k. Similarly $fig.1.5$ depicts results obtained for the simplified Klar model, and $fig.1.6$ for the IDM model by Treiber and Helbing.

1.4 Macroscopic modeling

The macroscopic level of description loses the detailed information available in the kinetic models but on the other hand is often more convenient.

In the case of classical kinetic theory it is possible to obtain the macroscopic PDEs of Euler and Navier-Stokes in important situations of high gas density.

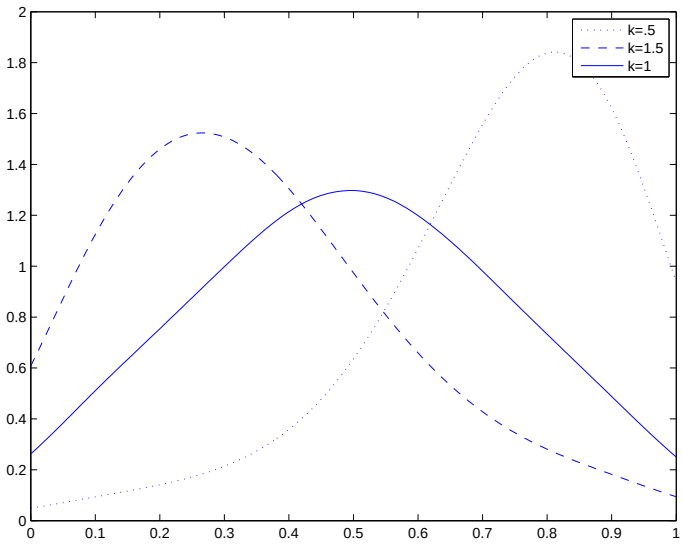


Figure 1.4: *Equilibrium density for the Illner-Klar model at different* k

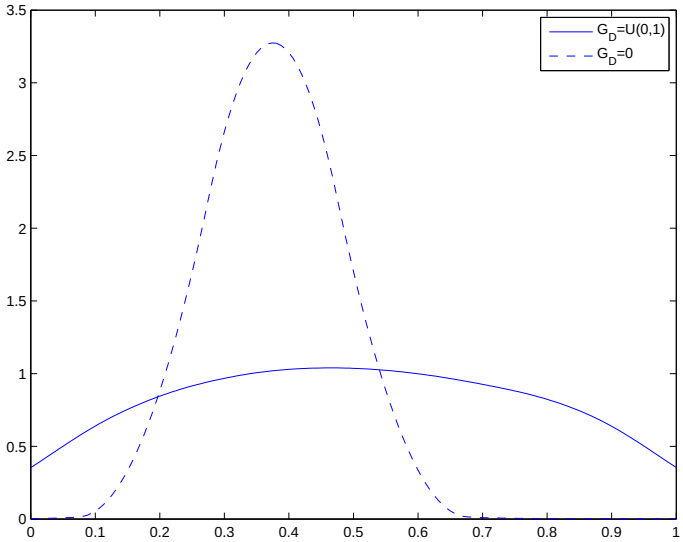


Figure 1.5: *Simplified Klar et al. model*

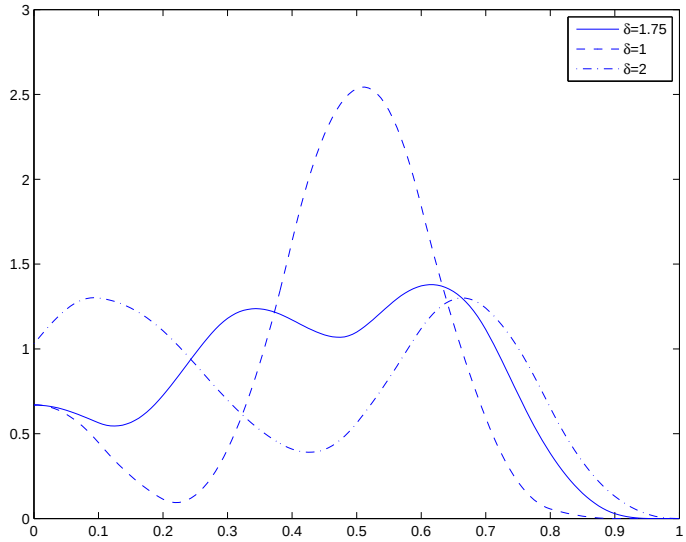


Figure 1.6: *The more complex IDM model by Treiber and Helbing*

A hierarchy of equations involving the different models is created multiplying the kinetic equation by powers of density u^k thus expressing everything in terms of moments. However since each moment depends on the higher ones to solve the hierarchy appropriate closures are needed.

This can be done in the following way: considering that the collision kernel $Q(f, f) = 0$ if and only if f is a Maxwellian, that is a normal distribution. Whenever at kinetic level f is a local Maxwellian it is defined by two variables only (mean and variance, the first two moments). So the third level moment depends on the previous ones, therefore it is possible to close the hierarchy solving for the density.

The hypothesis is valid for $\epsilon \to 0$; ϵ is proportional to the relaxation limit, so if the gas is dense enough the approximation leading to Navier Stokes is reasonable.

The case of traffic is similar but maybe not equally elegant. $Q(f, f) = 0$ $f = f(\rho, \rho u)$. Only one quantity is conserved.

$$\partial_t f + v\partial_x f = Q(f, f)$$

By the multiplication with v^k and integrating over v one obtains:

$$\partial_t m_k + v\partial_x m_{k+1} = \int_0^{v_{\max}} v^k Q(f, f) dv$$

where

$$m_k = \int_0^{v_{\max}} v^k f dv$$

For example, in the case of the first equation, since $m_0 = \rho$ and $m_1 = \rho V$ one obtains the standard continuity equation:

$$\partial_t \rho + \partial_x(\rho V) = 0$$

Here the Lighthill-Whitham equation - described in the following - can be obtained setting $V = V^e(\rho)$.

1.4.1 Lighthill-Whitham-Richards model

In the Lighthill-Whitham-Richards model velocity of vehicles is assumed to be uniquely determined through density by imposing a local equilibrium distribution $V(\rho)$. So one can solve for ρ a hyperbolic equation for the conservation of vehicles:

$$\partial_t \rho(x,t) + \partial_x \left(v(x,t)\rho(x,t)\right) = 0$$

with:

$$v\left(x,t\right) := V\left(\rho\left(x,t\right)\right)$$

This velocity relation $V(\rho)$ must be nonincreasing and nonnegative in the interval $(0, \rho_{max})$. It is often chosen so that the resulting fundamental diagram relation becomes $f\left(\rho\right) = v_{\max}\rho\left(1 - \left(\frac{\rho}{\rho_{\max}}\right)^n\right)$.

Another case sometimes considered is: $f\left(\rho\right) = v_0\left(\rho^{-1} - \frac{1}{\rho_{\max}}\right)$.

The Lighthill-Whitham model is thus described by the following Cauchy problem on the real line for the previous hyperbolic equation:

$$\begin{cases} \partial_t \rho + \partial_x\left(\rho V\left(\rho\right)\right) = 0 & \forall\left(x,t\right) \in (a,b) \times \mathbb{R}^+ \\ \rho\left(x,0\right) = \rho_0\left(x\right) & \forall x \in (a,b) \end{cases} \tag{1.20}$$

This model can describe wave formation, propagation and dissolution both in the forward and in the backward direction: advancing fronts in the former case and queue formation in the latter. It is probably the approach receiving the most attention, under all aspects like the modeling perspective, its analytical properties and the numerical methods for solution. It has also proven easy to extend to networks, multilane and

multiclass driver settings.

However it also possesses various drawbacks; among these are the formation of shocks and the lack of stop and go waves.
For the former issue there are common remedies like the introduction of a dissipative term to prevent the issue. However there seems to be no way to introduce stop and go waves without substantially altering the model. Moreover - as noted previously for microscopic models - imposing a unique relation between density and speed is not in agreement with experimental data, which shows a range of velocities for intermediate densities.

1.4.2 Second order models

The forcing constraint for velocity can be relaxed leading to more general models like that from Payne and Whitham, where treating velocity independently gives an additional equation (hence the term *second order models*, which is not to be confused with any more usual analytical "order"):

$$\begin{cases} \partial_t \rho + \partial_x(\rho v) = 0 \\ \partial_t v + v\partial_x v + \dfrac{p'(\rho)}{\rho}\partial_x \rho = (V^e(\rho) - v)/\tau(\rho) \end{cases} \tag{1.21}$$

Here $p(\rho)$ is the *anticipation coefficient*, a heuristic "density" function slowing down the traffic; while τ is the relaxation time for v to reach the free flow velocity $V^e(\rho)$. However determining in practice such quantities proves to be not so straightforward. Moreover various requirements are violated, namely non-negativity of velocity and speed of information bounded by car velocity. Furthermore it has been argued that all second order models could not account for anisotropy in the information flow for traffic dynamics (drivers react mostly to changes in the situation in front of them, more than behind).

A better approach has been presented in 2000 by Aw and Rascle (see [1], and independently by Zhang [17]) which does not suffer of the same drawbacks. They applied a convective derivative to the pressure term, leading to:

$$\begin{cases} \partial_t \rho + \partial_x(\rho v) = 0 \\ (\partial_t + v\partial_x)(v + p(\rho)) = 0 \end{cases} \tag{1.22}$$

Here $v + p(\rho)$ can be regarded as a preferred velocity, with $p(\rho)$ as before. This function must be smooth, strictly increasing and Lipschitz continuous and satisfy $p(0) = p'(0) = 0$; often it is taken $p(\rho) = \rho^\gamma$ with $\gamma > 0$. This implies that the two

eigenvalues are $\lambda_1 = v - \rho p'(\rho)$ and $\lambda_2 = v$, so that waves do not move faster than traffic itself and no vehicle is influenced by what is happening behind.

It is also possible to add a relaxation term affecting velocity, as seen in the Payne-Whitham case; this models the attempt from drivers to reach an ideal speed and the time lag in the response of each car:

$$\begin{cases} \partial_t \rho + \partial_x(\rho v) = 0 \\ (\partial_t + v\partial_x)(v + p(\rho)) = (V^e(\rho) - v)/\tau(\rho) \end{cases} \tag{1.23}$$

This model is also able to predict instabilities near the vacuum, that is for very light traffic at small values of ρ. The relaxation term is needed on partially empty roads to prevent maximum speed reached by cars from depending on the initial data.

The Aw-Rascle model can be derived from a corresponding microscopic car following rule (see i.e. [5] and references therein).

Bibliography

[1] A. Aw, M. Rascle, *Resurrection of second order models of traffic flow?*, SIAM J. Appl. Math, v. 60, (2000), pp. 916-938

[2] M. Bando, K. Hasebe, A. Nakayama, A. Shibata, Y .Sugiyama, *Dynamic model of traffic congestion and numerical simulation*, Physical Review E, v.51, n.2 (1995), pp. 1035–1042.

[3] M. Günther, A. Klar, T. Materne, R. Wegener, *An explicitly solvable kinetic model for vehicular traffic and associated macroscopic equations*, Math. Comp. Modelling, 35, (2002), p. 591

[4] M. Herty, A. Klar, L. Pareschi, *General kinetic models for vehicular traffic flow and Monte Carlo methods*, Computational Methods in Applied Mathematics, 5, (2005), pp. 154-169.

[5] R. Illner, C. Kirchner, R. Pinnau, *A derivation of the Aw-Rascle traffic models from Fokker-Planck type kinetic models*, Quart. Appl. Math. (2009)

[6] R. Illner, A. Klar, H. Lange, A. Unterreiter and R. Wegener, *A kinetic model for vehicular traffic: Existence of stationary solutions*, J. Math. Anal. Appl., (1999).

[7] B. S. Kerner, *Experimental features and characteristics of traffic jams*, Physical Review E, v. 53, n.2 (1996), pp. R1297-R1300

[8] A. Klar, R. D. Khne, R. Wegener *Mathematical models for vehicular traffic*, Surv. Math. Ind., 6 (1996), pp. 215-239

[9] A. Klar and R. Wegener, *A hierarchy of models for multilane vehicular traffic I: Modeling*, SIAM J. Appl. Math, 59, (1998), pp. 983-1001

[10] A. Klar, R. Wegener, *Kinetic Derivation of Macroscopic Anticipation Models for Vehicular Traffic*, SIAM J. Appl. Math. Volume 60, Issue 5 (2000), pp. 1749-1766

[11] R. J. LeVeque, *Numerical Methods for Conservation Laws*, Birkhauser-Verlag, (1994)

[12] M. Lighthill, J. Whitham, *On kinematic waves*, Proc. Royal Society Edinburgh, A229, (1055), p.281

[13] S. L. Paveri Fontana, *On Boltzmann-like treatments for traffic flow*, Transportation Res. 9, (1975), pp. 225-235.

[14] I. Prigogine and R. Herman, Kinetic Theory of Vehicular Traffic, American Elsevier Publishing Co., (1971), New York.

[15] P. I. Richards, *Shock waves on the highway*, Technical Operations, (1955), pp. 42-51

[16] M. Treiber, A. Hennecke, D. Helbing (2000), *Congested traffic states in empirical observations and microscopic simulations*, Physical Review E 62 (2): 18051824

[17] H. M. Zhang, *A non-equilibrium traffic model devoid of gas-like behavior*, Trans. Res. B,36, (2002), pp. 275-298

Chapter 2

Traffic flow on networks

2.1 Network modeling

Conservation laws on networks are useful in a wide variety of practical fields, and the methods developed for traffic have been also applied to data networks, supply chain management, air traffic management, gas pipelines, irrigation channels etc... (see i.e. [11] for references).

To extend the standard Lighthill-Whitham-Richards conservation law on a single road 1.20 to the case of a network we proceeded in the following way:

Definition 2.1.1 *: A traffic flow network is a connected and directed graph defined by a set of points called* nodes *or* vertices*, connected in couples by* arcs*. Eventually such arcs can extend indefinitely.*

Each arc will represent a corresponding road, and they meet at the nodes representing junctions. To each arc - numbered $j \in \mathcal{J} := \{1, ..., J\}$ - we will then associate an interval $I_j := [a_j, b_j]$, where any of the a_j-s or b_j-s can be possibly infinite.

On each road the traffic will be described by a macroscopic vehicle density function, that is $\rho_j(x,t)$ defined on each interval $x \in [a_j, b_j], \quad \forall j \in \mathcal{J}, t \in \mathbb{R}_+$. The time evolution of each $\rho_j(x,t)$ will proceed independently according to the Lighthill-Whitham-Richards model for a single road, except at the endpoints where one must impose some constraints modeling interactions of different roads at the junctions, that is the so called *coupling conditions* which will guarantee coherence among the partial solutions on each interval.

A general network can have junctions with arbitrary numbers of incoming and outgoing roads; for simplicity we have reworked the graph layout corresponding to

the actual highway network examined, to obtain an equivalent network including junctions with only three one-way roads, as those shown in figure 2.1.

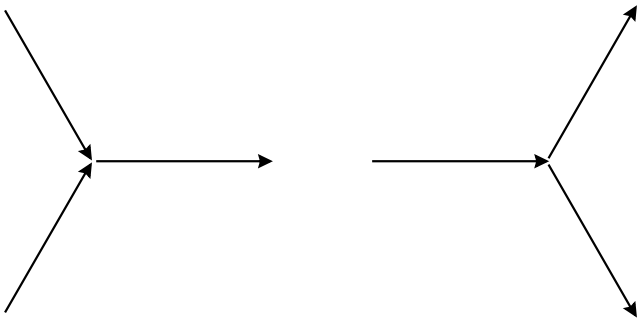


Figure 2.1: *Sketch of three road one-way junctions*

To obtain the desired graph it is necessary to split the higher-level network into basic elements: n-junctions into 3-junctions, two-way junctions into one-way junctions, two-way roads to couples of one way roads.
Even apparently simple configurations can easily lead to non-trivial modeling structures, as can already be seen in a minimal junction among three two-way roads, which gives rise to six basic 3-junctions and six additional arcs, as depicted in figure 2.2. Furthermore any resulting stretches deemed unimportant have been selectively omitted.
However determining a solution at such simple junctions proves already to be not so trivial even for stationary boundary conditions: imposing the conservation of mass in the simple second 3-junction of figure 2.2 is not enough, as one degree of freedom is left open. In the following this issue will be targeted.

2.1.1 Simple junctions

In the following we illustrate some examples of basic junctions to introduce the more general treatment following. The setting is still that of the LWR conservation equation with the corresponding velocity $V(\rho) = v_{max}(\rho_{max} - \rho)$ resulting in the flux $f(\rho) = v_{max}\rho(\rho_{max} - \rho)$.

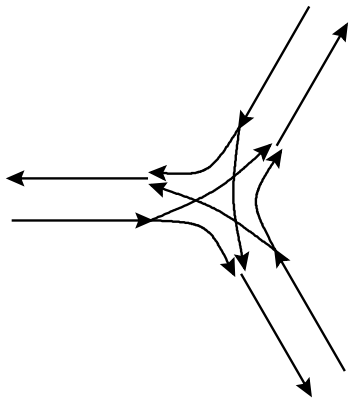


Figure 2.2: *Three road junction split into one-way junctions*

Traffic light

Let's start considering the simplest possible junction, that is a traffic light positioned ad $x = 0$, with the following initial datum:

$$\rho_0(x) = \begin{cases} \rho_{max} & \text{if } x \leq 0 \\ 0 & \text{if } x > 0 \end{cases} \tag{2.1}$$

This describes a situation where the traffic light is red and all cars are queueing before it, so that the road ahead is empty. Let's consider a light turning green at time $t = 0$ when cars start passing through: such dynamics will be well described by the following solution:

$$\rho(x,t) = \begin{cases} \rho_{max} & \text{if } x < f'(\rho_{max})t \\ (f')^{-1}(x/t) & \text{if } f'(\rho_{max})t < x < f'(0)t \\ 0 & \text{if } f'(0)t < x \end{cases} \tag{2.2}$$

In practice this means that on there are two wavefronts exiting from the junction, respectively at speed $f'(\rho_{max})$ and $f'(0)$. Outside of this interval the situation remains the initial datum, while the drivers in the interval $[f'(\rho_{max})t, f'(0)t]$ are moving and

there's a continuous density gradient as vehicles separate from each other, due to the acceleration at the green light.

A similar situation can be observed when a traffic light turns from green to red. Assume that initially the road is used by cars uniformely distributed at density $\rho_0(x) = \sigma$, and that the light turns red at time $t = 0$. The stop can be modeled as a null flux at $x = 0$, and the following solution satisfies the problem:

$$\rho(x,t) = \begin{cases} \sigma & \text{if } x \leq -v_{max}\sigma t \\ \rho_{max} & \text{if } -v_{max}\sigma t < x \leq 0 \\ 0 & \text{if } 0 < x < v_{max}\sigma t \\ \sigma & \text{if } v_{max}\sigma t < x \end{cases} \tag{2.3}$$

This corresponds to cars already ahead of the junction leaving it at speed $v_{max}\sigma$ and leaving an empty stretch back; on the other hand the red light causes the formation of a queue which propagates backwards at the opposite speed $-v_{max}\sigma$, while behind new cars keep arriving at the same density.

Three-junctions

A 3-junction can be more explicative of the problems arising when modeling networks. Let's consider the case of one single incoming road splitting in two, and with the following initial datum:

$$\rho_{0,1}(x) = \rho_{max}, \quad \rho_{0,2}(x) = \rho_{0,3}(x) = 0$$

Assuming that no driver preferences are imposed and road characteristics are homogeneous there are two possible extreme solutions, corresponding to all cars taking each of the two outgoing roads. These are similar to the just described traffic light turning green; the first reads:

$$\rho_1(x,t) = \begin{cases} \rho_{max} & \text{if } x < f'(\rho_{max})t \\ (f')^{-1}(x/t) & \text{if } f'(\rho_{max})t < x < 0 \end{cases}$$

$$\rho_2(x,t) = \begin{cases} (f')^{-1}(x/t) & \text{if } 0 < x < f'(0)t \\ 0 & \text{if } f'(0)t < x \end{cases}$$

$$\rho_3(x,t) = 0$$

The other solution just swaps $\rho_2(x,t)$ and $\rho_3(x,t)$.
Both solutions preserve fluxes: the sum of incoming fluxes is the same as the outgoing ones, but clearly this is not enough for uniqueness. Further conditions will be presented in the following.

2.1.2 Wave front tracking

Wave front tracking is an efficient and flexible numerical method for scalar nonlinear hyperbolic equations that is relatively straightforward to employ on networks. It can be traced back to the work of Dafermos in the early seventies (see [7]), and has since been improved among others by Holden and Risebro (see [10]).

The idea is very simple and is based on the fact that Riemann problems for conservation laws are solved by a time-dependent translations of the Heaviside function[1], as was shown in the simple cases of traffic lights and 3-junctions. In practice when the initial data to the Cauchy problem is piecewise constant, the analytic solution can be decomposed in a set of Riemann problems, each corresponding to the dynamics of each discontinuity or *wave front*. Key to such a decomposition is the finite maximal propagation speed of wavefronts, which keeps the Riemann problems on arcs and junctions independent.

Since the solution of these Riemann problems is much simpler one can then track each wave front until two of them interact together or one interacts with a junction. When such interactions occur the solution will stay piecewise constant and a new Riemann problem is determined, so that the cycle can start again.

For general initial data then one can start with a corresponding sufficiently "near" piecewise approximation $\rho_{j,0}$, where rarefactions are decomposed into many small steps. In the following the Riemann problem for junctions will be covered.

2.1.3 Coupling conditions at a junction

To solve the network problem we first need to be able determining a weak solution of a Riemann problem at the junction. To that aim let's initially introduce weak entropic solutions on arcs and at a junction.

Definition 2.1.2 *A weak entropic solution on an arc I_j is in our case a density function $\rho_j(x,t) : I_j \times [0,+\infty[\to \mathbb{R}$ such that for every smooth test function $\phi(x,t)$:*

[1] Where the Riemann problem is the Cauchy problem with initial data being a Heaviside function, that is a step function $a\chi(\mathbb{R}_+)$.

$I_j \times [0, +\infty[\to [0, +\infty[$ *the following holds:*

$$\int_0^{\infty} \int_{a_j}^{b_j} [\rho_j \partial_t \phi_j + f(\rho_j) \partial_x \phi_j] dx\, dt = 0$$

For an initual datum $\rho_{j,0}$ with bounded total variation the LWR equation admits a unique entropic weak solution, which depends continuously on the initial datum in L^1_{loc}. Furthermore for $\rho_{j,0} \in L^1 \cap L^\infty$ it's possible to achieve Lipschitz continuous dependence in L^1 for $t \mapsto \rho_j(\cdot, t)$.

Definition 2.1.3 *A* weak solution at a junction *for the LWR conservation law (1.20) is a collection of density functions over the roads and intervals* $\{\rho_i(x,t) : I_i \times [0, +\infty[\to \mathbb{R}\}_{i=1,\dots,n+m}$ *such that the following condition is satisfied for every smooth test function on the junction arcs* $\phi = (\phi_1, \dots, \phi_{n+m})$*:*

$$\sum_{j=1}^{n+m} \left[\int_0^{\infty} \int_{a_j}^{b_j} [\rho_j \partial_t \phi_j + f(\rho_j) \partial_x \phi_j] dx\, dt \right] = 0$$

where the test function ϕ *is said smooth across junctions if at each of them it satisfies the following conditions:*

$$\phi_i(b_i) = \phi_j(a_j)$$
$$\partial_x \phi_i(b_i) = \partial_x \phi_j(a_j)$$

with the usual convention of i *denoting incoming roads and* j *outgoing roads.*

Definition 2.1.4 *The* weak solution of a Riemann problem at a junction *is a weak solution for the LWR problem at a junction for an initial datum constant on each road:* $\rho_{j,0}(x) := \rho_{j,0}\ \forall x \in I_j,\ j = 1, \dots, n+m$*, where all arcs* I_j *extend to infinity.*

It is convenient to work with a flux function satisfying the properties given in Chapter 1 (namely f is smooth, strictly concave, $f(0) = f(1) = 0$, $|f'(x)| < C < \infty$ and thus possesses a unique maximum $\sigma \in]0, 1[$).

A related necessary condition to be satisfied is clearly flux conservation, hence the following *Rankine-Hugoniot condition* as a consequence of the previous constraint whenever each $\rho_i(x,t)$ has bounded variation in x:

$$\sum_{i=1}^{n} f_i\left(\rho_i\left(b_i, t\right)\right) = \sum_{j=n+1}^{n+m} f_j\left(\rho_j\left(a_j, t\right)\right)$$

However this is just a single one dimensional constraint so that solution remains still undetermined.
We will also make use of the following:

Definition 2.1.5 *Let $\tau(\rho) : [0,1] \to [0,1],\ \tau(\sigma) = \sigma$ be the map satisfying for any $\rho \neq \sigma$ the following conditions:*

$$f(\tau(\rho)) = f(\rho), \quad \tau(\rho) \neq \rho$$

We can look at the values a solution $\rho_{i|j}(b_i|a_j)$ takes at the junction and denote them in the following way:

$$\overline{\rho}(t) := (\overline{\rho}_1, \dots, \overline{\rho}_{n+m})(t) := (\rho_1(b_1), \dots, \rho_n(b_n), \rho_{n+1}(a_{n+1}), \dots, \rho_{n+m}(a_{n+m}))(t)$$

By imposing appropriate conditions one obtains these values are independent of time, so we can write $\overline{\rho}_{i|j}(t) = \overline{\rho}_{i|j}$. Conversely, setting ρ one can obtain a weak solution by solving the corresponding Riemann problems on each road, respectively for incoming and outgoing:

$$\begin{cases} \rho_t + f(\rho)_x = 0, \quad x \in \mathbb{R}, t > 0 \\ \rho(x,0) = \begin{cases} \rho_{i,0} & if\ x < 0 \\ \overline{\rho}_i & if\ x = 0 \end{cases} \end{cases} \tag{2.4}$$

$$\begin{cases} \rho_t + f(\rho)_x = 0, \quad x \in \mathbb{R}, t > 0 \\ \rho(x,0) = \begin{cases} \overline{\rho}_j & if\ x = 0 \\ \rho_{j,0} & if\ x > 0 \end{cases} \end{cases} \tag{2.5}$$

The desired weak solution at a junction will need to have wavefront speeds on each road moving outwards, that is negative for incoming roads and positive for outgoing. This imposes constraints on the density as follows:

$$\begin{cases} \overline{\rho}_i \in [\sigma, 1] & \rho_{i,0} \geq \sigma \quad i = 1, ..., n \\ \overline{\rho}_i \in \{\rho_{i,0}\} \cup (\tau(\rho_{i,0}), 1] & \rho_{i,0} \leq \sigma \quad i = 1, ..., n \end{cases} \tag{2.6}$$

$$\begin{cases} \overline{\rho}_j \in [0, \sigma] & \rho_{j,0} \geq \sigma \quad j = n+1, ..., n+m \\ \overline{\rho}_j \in [0, \tau(\rho_{j,0})) \cup \{\rho_{j,0}\} & \rho_{j,0} \leq \sigma \quad j = n+1, ..., n+m \end{cases} \tag{2.7}$$

Holden-Risebro conditions

A first example of Riemann solver for traffic networks has been presented in [10]. In this case to define a solution after eq.(2.6),(2.7) an additional constraint is imposed at each junction: the *entropy condition* consisting in the maximization of the following functional:

$$E(\overline{\rho}_1, ..., \overline{\rho}_{n+m}) = \sum_{j=1}^{n+m} g(f(\overline{\rho}_j)/f(\sigma))$$

Here $g(\cdot)$ is a differentiable and strictly concave function. The proof of existence and uniqueness of the solution of such maximization problem is somewhat constructive: it makes use of the front tracking algorithm for increasingly better approximations of the initial data and of the flux function, and obtains the desired result as the limit solution of this procedure.

This method has a drawback in that it is not possible to maximize the flux, as some functions $g(\cdot)$ cannot be treated and $g(f(\overline{\rho}_j)/f(\sigma))$ in particular.

Coclite-Piccoli conditions

Another Riemann solver was later presented in [5] and [6]. The ranges (2.6) and (2.7) for $\overline{\rho}_{i|j}$ are again considered as a constraint for admissible solutions with waves exiting from the junction. Furthermore the following other constraints are imposed:

- We assign a matrix A of turning coefficients describing the fraction α_{ji} of vehicles turning at the junction from road $I_i,\ i \in \{1, \dots, n\}$ to road $I_j,\ j \in \{n+1, \dots, n+m\}$:

$$A := \begin{pmatrix} \alpha_{n+1,1} & \cdots & \alpha_{n+1,n} \\ \vdots & \ddots & \vdots \\ \alpha_{n+m,1} & \cdots & \alpha_{n+m,n} \end{pmatrix}$$

  Such coefficient must satisfy $0 \leq \alpha_{ji} \leq 1$ and $\sum_{j \in Out} \alpha_{ji} = 1$ so that the coupling condition is split into:

$$f_j(\rho_j(a,t)) = \sum_{i=1}^{n} \alpha_{ji} f_i(\rho_i(b,t)) \quad j = n+1, ..., n+m \tag{2.8}$$

  This condition is in agreement and subsumes conservation of cars through the

junction. In fact one can observe the following:

$$\sum_{j=n+1}^{n+m} f_j = \sum_{j=n+1}^{n+m} \sum_{i=1}^{n} \alpha_{j,i} f_i = \sum_{i=1}^{n} \sum_{j=n+1}^{n+m} \alpha_{j,i} f_i = \sum_{i=1}^{n} f_i$$

While the Rankine-Hugoniot condition was not enough to determine a unique solution, one might hope that now the turning coefficient matrix A imposes enough constraints to reach such a goal. Unfortunately this is still not sufficient, in fact a simple counterexample can be shown: let's consider again the three-junction with a single incoming road. The following density is a solution satisfying the additional constraint just considered:

$$\rho_1(x,t) = \rho_{max}, \quad \rho_2(x,t) = \rho_3(x,t) = 0$$

In fact all derivatives vanish on each road, and car conservation through the junction is also satisfied, as all fluxes also vanish, so this is a valid solution. Note that this does not even depend on any particular matrix A; actually in this case equation (2.8) is trivially satisfied and becomes the identity $0 = 0$. Clearly this is not a desirable property of a model, and is due to the fact that the propensity of drivers to cross the junction is not accounted for. Therefore it can be useful to introduce another condition accounting for such desired property.

- To determine a unique solution again a maximization of the cars passing the junction is carried out, that is over the cumulative outgoing flux function:

$$(\overline{\rho}_1, \ldots, \overline{\rho}_{n+m}) = argmax_{(\overline{\rho}_1, \ldots, \overline{\rho}_{n+m})} \left[\sum_{j=n+1}^{n+m} f\left(\overline{\rho}_j\right) \right]$$

  Actually because of car conservation this also corresponds to maximizing incoming fluxes: $\sum_{i=1}^{n} f\left(\overline{\rho}_i\right)$

When there is only one incoming road these conditions are equivalent to maximizing average incoming velocities.
The solution that is found is a time-linear translation of a Heaviside function (as seen already in equations (2.3) and (2.2)), which will allow easy determination of the whole solution on the network by immediate determination of shock interactions.

For junctions where the number of incoming roads is greater than outgoing ones there need to be further conditions. When not all cars can pass the junction it has to be determined which incoming roads will allow cars proceed, that is some right of way parameters are needed to find an unique solution. For example in the case of $n = 2$ and $m = 1$ with equal roads one fixes a parameter q determining the percentage of cars (in terms of flux) coming from I_1 allowed to enter I_3 first, that is having right of way; the remaining quota $(1 - q)$ is left for cars from I_2, however this can relax whenever the actual flux from I_1 is lower than allowed, so that more can come from I_2 instead.

The whole discussion can be further extended to deal with time dependent turning coefficients.

2.1.4 The turning coefficients

This kind of modelling is one of the simplest way to perform traffic flow simulation. The model is justified by the poor information we typically have on large highways. In most circumstances we will have simply daily average traffic measurement obtained by averaging on a period of several months.

We will assume that all junctions are constituted by three incident roads each. This is a usual assumption in traffic flow modelling, but we must note how often an even lower level decomposition is used with one-way roads triplets (we'd need six such blocks to simulate a single two-way roads triplet, see figure 2.2), to obtain directed graphs. By the sole knowledge of the I/O fluxes at a junction we have infinitely many solutions for the flows of cars that take a given path at the junction.

Let us denote by I_i and O_i, $i = 1, 2, 3$ the inflow and outflow of the roads at the junction. Clearly by cars conservation we have $I_1 + I_2 + I_3 = O_1 + O_2 + O_3$. If we denote by α_{ij} the fraction of cars turning from direction i to direction j we obtain readily two values α_{ij}^m and α_{ij}^M such that each $\alpha_{ij} \in [\alpha_{ij}^m, \alpha_{ij}^M]$ is admissible.

These are obtained by solving the linear system

$$\begin{aligned} \alpha_{21} I_2 + \alpha_{31} I_3 &= O_1 \\ \alpha_{12} I_1 + \alpha_{32} I_3 &= O_2 \\ \alpha_{13} I_1 + \alpha_{23} I_2 &= O_3, \end{aligned}$$

with the cars conservation constraint and $\sum_{j \neq i} \alpha_{ij} = 1, \forall i$, $1 \geq \alpha_{ij} \geq 0$.

We will assume these intervals to be characteristic of the traffic in the network (i.e. changing the I_i values should lead to a change of O_i but not of the $\alpha's$). More

precisely we will assume each α_{ij} to be a random variable on the interval $[\alpha_{ij}^m, \alpha_{ij}^M]$.

Similarly the stations can be treated by the knowledge of the inflow/outflow data. Let us denote with I and O the total inflow and outflow at the station from outside. We will denote by I_i and O_i, $i = 1, 2$ the car flows before and after the station in the two directions. Clearly we have by cars conservation $I - O = I_2 - I_1 - O_2 + O_1$. Simple algebra gives the input/otput flows at the station of cars in a given direction.

In fact if we set $I = I_3 + I_4$ and $O = O_3 + O_4$ where I_3 and I_4 stands for cars entering and taking direction 2 and 1 respectively (similarly O_3 and O_4 are exiting form direction 1 and 2 respectively).

The unknown values I_3, I_4, O_3, O_4 can be computed solving the linear system

$$\begin{aligned} I_3 + I_4 &= I \\ O_3 + O_4 &= O \\ I_3 - O_4 &= I_2 - I_1 \\ O_3 - I_4 &= O_2 - O_1. \end{aligned}$$

Once again we have infinite solutions for the fluxes and under the cars conservation constraint and nonnegativity we obtain intervals of existence for the values I_3, I_4, O_3, O_4. Let us set β_{ij}, $i, j = 1, 2, 3, 4$ the fraction of cars that at the station take direction j coming form direction i. Note that $\beta_{14} = \beta_{23} = \beta_{42} = \beta_{31} = 0$. As for the $\alpha's$ the $\beta's$ will be assumed as random variables characteristic of the traffic network.

2.2 The case study

2.2.1 Introduction

The present work describes a study conducted for the local governmental authorities of the Ferrara province to analyse the impact and potential usefulness of two planned new highway stretches. One of them should be connecting the north-to-south A13 highway from Ferrara to the more western A22 stretch. The other would be a highway variant of the already existing E55 motorway heading east to the sea. These have both the potential to lift traffic away from current congested areas opening direct connections to destinations previously needing detours or reachable with

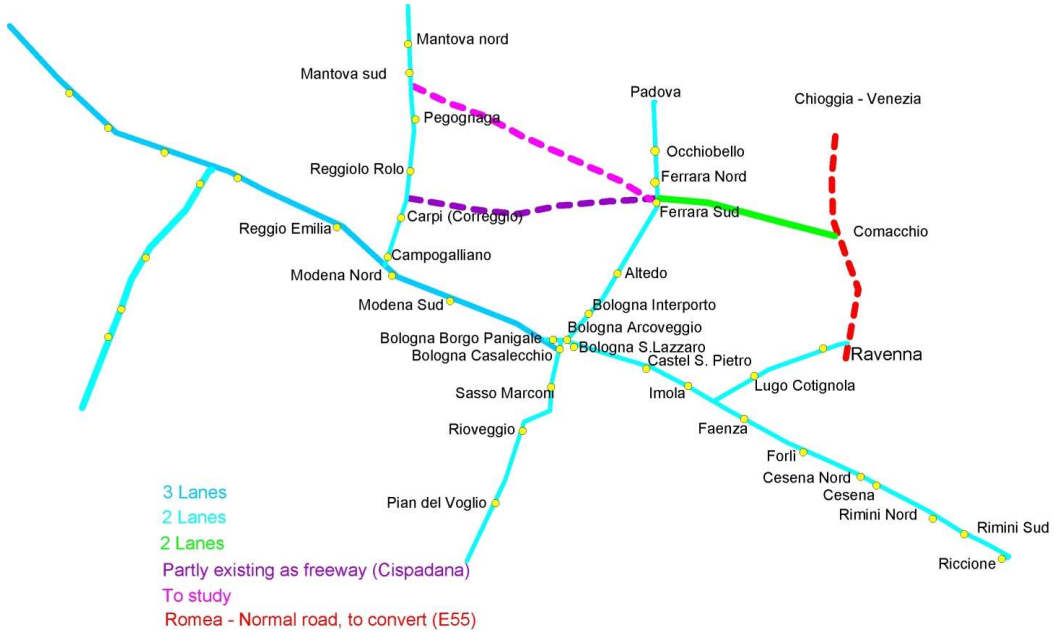


Figure 2.3: *The highway network studied*

ordinary roads only.

The different network scenarios have been evaluated by means of numerical simulation of traffic. We have chosen the Lighthill-Witham-Richards method because of its current wide acceptance, versatility and the readyness of simulation methods on road networks. For evolving the solution in time we used a wave front tracking algorithm, because of its adaptivity and readyness in treating junctions making it the method of choice in the case of traffic networks.

Aside from the implementation of the numerical simulation engine itself, a certain effort has been required in treating the available rough traffic information to obtain more useful data for use as input for initial condition, boundary conditions and turning preferences describing the behaviour of junctions.

2.2.2 Network setting

Our study covered the network visible in figure 2.3, including most of the highway infrastructures available in Emilia Romagna and nearby regions.

In particular the simulation has been conducted with the Lighthill-Witham-Richards (LWR) model together with a front tracking method, updating the vehicle density

profile at successive points in time. The front tracking method is concentratind calculations in particular over discontiuities areas for updating position at each time step. Moreover the LWR model is able to simulate the birth and backward (with respect to the flux) propagation of high density - and thus critical - situations.

We considered different possible scenarios, corresponding to different variants for the stretches Ferrara-A22 and E55. To analyze traffix flows in a clearer way we have conducted analyses corresponding to different directional modes (from north to south, from east to west and the two opposite directions), by selecting subsets of entrances significant for observing how vehicles redistribute over the highway network.

For interpreting the results one needs to highlight how the traffic flux depends from the numerical density of vehicles, thus clearly reaching maximum intensity in an intermediate situation between null and maximum density, which are corresponding to empty and congestioned road. Values shown on graphs are those of traffic flux, so that for a correct reading one needs to keep in mind that low values can point to both light and fluid or congestioned traffic. The relation determining how the flux is governed by numerical density is called *fundamental diagram* and depends on the characteristics of each particular road stretch.

Because simulations are over time we entered constant fluxes into the network until a substantially stable situation was reached, and then graphs were obtained. Because of this it is possible that in a few cases the stationary state has not been fully reached. Simulation graphs correspond to different highway stretches; whenever these enter or exit from the analyzed area we added fixed length padding of 5km, and the label refers to the first toll encountered. The following diagram shows a sketch of the simulated highway network and the symbols adopted for the main nodes. For reading keys one can refer to the initial picture in Appendix 2 and to pictures in the following.

2.3 Simulation scenarios

Analysed scenarios are ordered over groups according to the traffic direction (4 towards South, 2 towards North, 4 towards East and 4 towards West), and correspondingly the entrances and exit points were established. Each of these scenarios is

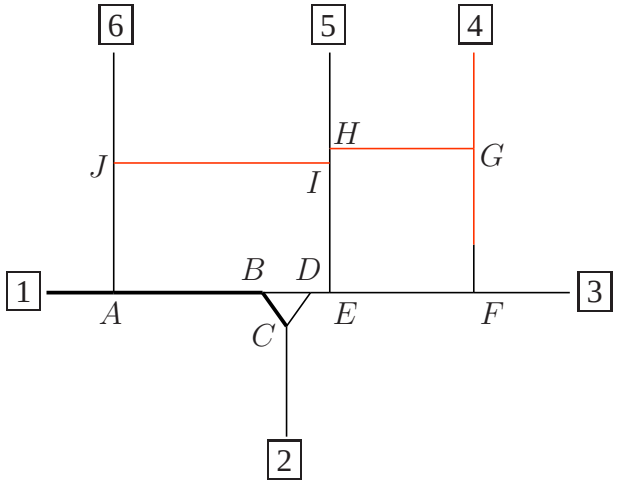


Figure 2.4: *Network and naming*

further comprehending four cases: first come the two variants of the junction A13-A22 according to the entrance on A22 (respectively high near Mantova or lower near Reggiolo Rolo) of the highway stretch from Ferrara on A13 (IJ segment on previous graph) but without the motorway Ferrara-Mare (HG arc in the previous graph). Next both new stretches IK and HG are simultaneously considered.

Scenario	1	2	3	4	5	6	7	8	9	10	11	12	13	14
IJ higher			√	√		√			√	√			√	√
IJ lower	√	√			√		√	√			√	√		
HG		√		√				√		√		√		√
North-South	√	√	√	√										
South-North					√	√								
West-East							√	√	√	√				
East-West											√	√	√	√

It must be stressed how simulations performed can only provide qualitative clues (due to the geometric structure of the network, to the characteristics of toll gates and of current traffic) but no quantitative results; for these it is necessary a further more deep analysis and the solution of various problems (starting from and estimation of induced traffic).

In fact if we consider the stretch from Ferrara to A22 traffic will be made up of both traffic diverted from the current highway network - modeled with all data seen so far - of that moved from the ordinart road network, and of that possible induced traffic which is created by the sole construction of the new infrastructure. Results presented here refer to the diverted highway traffic alone. In this case diverted traffic could be estimated as ad daily flux of 1700 vehicles. A more accurate estimation can clearly only be obtained in a more advanced phase of the project for the highway stretch and its precise position.

In the different scenarios of these simulations we kept constant all coefficients for crossings and junction points already existing; such hypothesis should surely be abandoned when performing a quantitative analysis, which would require an evaluation of influence of different driving paths over the turning coefficients (an illustrative case in this direction is given by the Ravenna junction over A14 which would reach high importance in case of realization of the 4F highway stretch, corresponding to the current Romea). Furthermore an accurate simulation should take into account intrinsic charachteristics of the stretches which influence the behaviour in the so called fundamental diagram which is charachterising traffic in the mathematical model.

2.3.1 Network data

Let's first review major data types that were obtained and analysed during the course of this inquiry:

- distribution of vehicle matriculations over the regional territory
- daily origin/destination matrices for provinces and county seats in Emilia Romagna
- highway traffic fluxes over the highway nat of Emilia Romagna
- entry/leave fluxes at toll gates of such network
- turn coefficients estimated at toll gates and splitting/merging crossings
- hourly course of highway traffic
- fluxes on provincial roads and freeways on the area of concern for E55

The two former points pertained to the first part of the research, while the others have been treated in the second part.

A first facet that has been noticed over data of all tipologies is the large preponderance of short range movements with respect to longer ones, even higher than initially expected. From this follows that the studien infrastructures will be - as detailed below - more useful as local support than capable of diverting long distance movements by the creation of new alternatives.
This partly circular issue is obviously also a limit in the study of new configurations starting from the sole current data, as clearly the E55 Nuova Romea would create a significant change in the geometry of north Italy transport, and the new possibilities would induce the birth of different activities. However it seems reasonable to deem such new activities mostly laying outside of the Ferrara territory, an area suffering mainly from a certain lack of infrastracture which is displayed in the current saturation conditions.

The highway route from Ferrara to A22 could in theory serve different user groups:

- those passing from the Ferrara area towards the Modena-Reggio district, both in local short range moves and in longer ones
- highway heavy traffic from Ferrara towards Reggio Emilia and beyond
- (in the higher scenario) traffic coming from A22 and headed beyond Bologna south-eastwards
- similar transits headed towards the Appennini
- partial rerouting of traffic towards the coastline through the Cispadana - Ferrara Mare
- traffic generated by the sole presence of the infrastructure

Necessarily the a) users typology would constitute the highest volume, and together with b) would be the most affected by the new infrastructure. The component currently making use of ordinary road network can be identified as being of about 1700 vehicles per day.
According to the extrapolations done through the data at toll booths, typology b) of traffic seems to be currently constituting 8.39% of heavy traffic through Ferrara nord, and 4.99% from Ferrara Sud (with roughly exchanged percentages one obtains the

highway component for light traffic of type a)). This does certainly lead to an underestimation of the actual value but it can nevertheless be confirmed in two ways: on the one hand it appears anyway coherent in order of magnitude with the analogous estimates which where obtained with regard to vehicles coming from beyond Verona and to actual values obtained, on the other hand the figures for O/D routes in the regional researches are already very low.

With respect to c) we inferred from the data that these would be very low fluxes, in particular taking in consideration the precision of our sources, and anyway they wouldnt help much in easying the difficult situation on tha A1 between Modena and Bologna. Moreover fluxes of type d) are only slightly higher and while they do lighten the most congestioned stretch of the A1 they also keep gravitating in the Bologna area.
Also the users of e) type routes seem to be of little impact, in particular in the variant without help from E55 where such value can then be neglected.

Up to here the analysis concerned already existing traffic which would be ”diverted” on the new infrastructures, which thus function as traffic lighteners. It’s a whole other matter with regard to traffic induced by the presence of the infrastructures, an additional load that will weigh on both these and the preexisting network, adding to the current traffic with an effect that might even be decisive in nearly critical situations.
However induced traffic f) can hardly be estimated with good precision, also because of its dependance on economical factors. Anyway it is composed of a local part arising from activities growing in the neighborhood and a global part related to all new activities coming from distance (also thanks to a stretch like E55) and insisting on this area.

To evaluate the size off all such flow tipologies one also has to consider in parallel the effect of the Cispadana motorway - currently under construction - which by serving a similar route would also already function for similar flows. In particular such infrastructure is suitable for supporting all small distance movements, which in undergone research also resulted being those most heavyly affecting the highway variant. Therefore this would only be interesting if paired with the E55, for an integration of the respective functionalities.

2.3.2 Simulation setting

The research part concerning the simulation of traffic fluxes has been conducted in collaboration with Kaiserslautern University in Germany.
Performed simulations concerned the highway road network in the two areas divided by the A13 from Ferrara to Bologna as sketched in $fig.2$, thus including:

- the A22 stretch from the junction to A1 (near Campogalliano) up to Mantova
- the A13 highway from Occhiobello to Bologna
- the motorway Ferrara-Mare
- the A1 stretch from Reggio Emilia to Bologna
- the continuation of A1 from Bologna to Sasso Marconi
- A14 from the start in Bologna till Faenza, that is beyond the junction to Ravenna
- the junction to Ravenna from the A14 highway
- the triangle at Bologna knot, comprising the junction between A1 and A4 from Firenze to the Adriatic sea and backwards
- a hypotetical highway stretch from Ferrara to A22, in two different configurations
- the adriatic highway stretch E55 corresponding to Romea

Several among these stretches do currently suffer heavy congestion in different and specific occasions connected to the respective positionings within the network and with respect to the surrounding urban territory being serviced.
Simulations have shown part of the phenomenons known to occur on the current network, as well as suggested possible changes in the way of distributing fluxes arising from the examined configurations.

The fluxes were defined with attention to data on turns which were obtained from current highway flows and from those induced by the creation of new stretches. Fluxes themselves have been set at entrances into the network, as described in the previous relation; they are mean daily fluxes because Autostrade per L'Italia was only able to provide the cumulative yearly data, in this case referring to year 2003. In correspondence to toll gates there are incoming and exit fluxes substantially equal to the daily average, thus the contribution to the respective stretch turns out negligible:

the network behaviour is then mainly due from feedings and sinks within it, and from interactions among these, which also can cause unexpected phenomena.
The considered graphs show together the fluxes of heavy and light traffic, together with the sum. However this is only a hint for the overall situation, since simulations for the different vehicle categories have been conducted separately from each other by changing the respective parameters.

The scenarios we've gone through represent particular situations and aspects of vehicular fluxes, and do not intend to constitute an exhaustive case study as rather an illustrative one. Exhaustiveness and perfect adherence are however hardly reachable, given both the current incompleteness of available data and the arbitraryness of analysis choices which were selected within a very wide spectrum. Therefore we opted for highlighting the most significant scenarios while limiting their number to the minimum necessary for achieving at a time clarity, detail and significativeness.

2.4 Scenario analysis

Performed simulations cover altogether 35 different scenarios (while separately handling heavy and light traffic). In the previous relation only those showing results of some interest were shown, and they're analysed and discussed in more detail in the following.
Given the peculiar geometry of the examined network, it was possible to split the overall traffic according to four different movement modes, corresponding to different destination directions (thereby excluding improbable and in any case neglectable U-shaped routes): in order they are from north to south, from south to north, from west to east and finally from east to west. In the cases of north-south axis the Cispadana has been modeled as bidirectional since it is in practive orthogonal and lacks a preferential directionality. On the other hand the directional selection was used in the cases of traffic on the west-east axis.
These four cases have themselves been further split in more scenarios each, accordint to the realization of the Ferrara-A22 connection in the two variants (near Mantova or near Reggiolo) and further according to the inclusion of Ferrara Mare and E55.

Direction	Scenarios	Numbering	Directions on Cispadana	Inflows
North ⇒ South	4	# 1-4	2	4,5,6
South ⇒ North	2	# 5,6	2	2,3
West ⇒ East	4	# 7-10	1	1,2,6
East ⇒ West	4	# 11-14	1	3,4,5

2.4.1 Results

In the following we describe the results of simulations performed. In appendix for each scenario a set of graphs is shown, each one showing a snapshot of the fluxes on one of the highway stretches involved. Both light and heavy traffic are depicted, as well as the total flux. The time point is usually after or near convergence of the solution, so that mainly one can study the effect of the network structure on fluxes; we're not considering here the specific dynamic details because that would require much more detailed data for model calibration, however the methods applied could in theory be reused in that case. Junctions are visible both with the indication of their position and the changes in flux as traffic is diverted (or joins) to (from) other roads. In the following a detailed description of results for the various scenarios is presented.

North to south direction

Scenarios 1 and 2

The first two scenarios, both involving traffic directed south, are essentially similar in the critical issues highlighted, which therefore turn out to be independent from the position of the junction Ferrara-A22 over this latter. The already intense fluxes coming from Reggio Emilia get increased by the influx of A22, so that a critical situation arises from here to the next junction with A14 in Bologna. On this highway the traffic is lower over the first stretch, but grows significantly on the stretch between the junction with A1 and the entrance of A13; on the other hand beyond this latter highway the traffic movement is eased.
In both cases the traffic over the junction Ferrara-A22 turns out to be higher in the west-east direction.

Scenarios 3 and 4

In these scenarios in comparison with the previous ones two stretches were added: Ferrara Mare and E55; this reflects in the fluxes on A13 which are lowering from 25000 vehicles to 20000. However this proves not enough to improve the congestion situation located at the junction with A14 on the west stretch, on the contrary this turns out slightly worsened; while in east direction the situation is more fluid. This apparently counterintuitive consequence is due to the whole network behaviour, and one must note how fluxes are not to be added or subtracted directly, because - even without changing the amount of vehicles entering in each time unit - the congestion situations cause changes in speed and consequently also flux and density in the interior parts.
The already critical situation already seen on A1 before Bologna does not show any substantial difference in these scenarios.
The two added arterial roads (with respect to the two previous scenarios) would still not carry high fluxes.
Finally, again it can be seen that the two locations for the junction Ferrara-A22 dont provoke any big changes in the global traffic distribution.

South to north direction

Scenarios 5 and 6

For the South-North direction we've considered the case with the highway version of Cispadana only, which shows more traffic in the first scenario. Critical situations arise on the stretch of the Bologna triangle from the entrance of A13 and A1 west, as previously already seen but in the opposite direction
However - differently from what already seen in the opposite direction - traffic on A13 reaches 30000 vehicles per day, which can be consided the danger level that sometimes is already causing slowdowns.

Scenarios 7 and 8

On this direction one starts to notice differences between the two configurations for the junction Ferrara-A22. In particular, besides the high volumes seen already over A1, one finds again the characteristic congestion on the A14 bolognese stretch between the A14-A1 junction and the attachment of A13. The different behaviour between the two geometries shows up in the faster saturation of such stretch which arises in scenario 7.

It is noteworthy how fluxes over A22 seem to be relatively higher than the effect they are inducing over A1, even though this is only considered for a single direction (that is, it is not about a split of the vehicles among the two directions), but especially how the light traffic fluxes become lower (from about 34000 down to roughly 31000). This probably could hint to the change from high traffic still flowing to a slowdown. Therefore the graphs of A1 - even though apparently similar to those seen for North-South moves - are essentially different from these, and this is due to changes in the ratios between heavy and light traffic.

Scenarios 9 and 10

The following two scenarios, where Ferrara Mare and E55 are added, show some effects which were missing in the cases 7 and 8.
First of all the fluxes on A1 after the A22 entrance are even higher (56000 for just the light traffic and 20000 for the heavy one).
With regard to the A14 stretch before the A13 entrance the situation stays unchanged, and the new stretches do not seem to be helpful. Here again however congestions show up faster in the second variant.
On the additional E55 stretch a different behaviour of heavy versus light traffic shows up: the former reach immediately a stationary regime (of about 10000 vehicles), while the latter show more instabilities and the creation of reflexed waves coming from the junction to ravenna from A14.

East to west direction

Scenarios 11 and 12

In these two scenarios the connection A22-Ferrara shows fluxes at a maximum level among those found in this research, of about 21000 vehicles per day. The only relevant effect differentiating the two variants seems to be at Bologna a decrease in light traffic fluxes on the A1 from fork with A14 and the junction.
The situations of highest traffic affect A1 beyond the Bologna triangle towards the Appennino, and on the A14 - beyond Bologna again - while slightly decreasing at the fork to Ravenna.

Scenarios 13 and 14

The addition - with respect to the two previous scenarios - of the stretches beyond Ferrara is determining a drain of E55 traffic on the Ferrara Mare, which seems to enter the A13 without incurring slowdowns (with respect to the previous two scenarios there is in fact a direct link between the flux diverted from E55 by the Ferrara Mare and that added by this to A13) and without the formation of backward moving waves. The connection Ferrara-A22 is subject to a flux decrease of about 25% (and its geometry seems to be only small influence over the A1 on the Bologna triangle); while on the A14 there are increases from 40.000 to 50.000 vehicles per day on the stretch between the merging from Ravenna and the fork with A13.

2.5 Concluding remarks

The undergone study has shown how the construction of even just the highway variant of Cispadana - also because of this very same existing stretch - might have little efefct, while its realization in the context of E55 would be able to significantly help in lowering traffic by offering a different route for various transport needs.
However the critical areas around the Bologna subnetwork are currently already at such a saturation level that the situation there would not get substantially better, even after a lightening of the estimated flows; on the contrary a few cases have emerged where the new network configuration would worsen situations currently almost critical like the Ferrara-Bologna stretch. In this context it could be useful a joint study including also the new infrastructures which are being studied for the Bologna area. These remarks are even more significant when considering the new traffic induced by the added infrastructures, a component which will counterbalance part of the traffic relief obtained through them: here again it will be important that its effect dont get concentrated in the areas now already near to a critical situation.

The applicability of developed software, thanks to its efficiency, still allows the handling of more complex scenarios than those analyzed in this study, be they with or without urban additions. In parallel also the amount of produced data would become more massive so that - to focus on the most interesting ones - it would come handy also the developement of a more user friendly interface and the direct use by those involved in planning.

Among possible variations that could extend the developed system is the addition of hourly fluxes entering into the network, so to be able to study the emergence and vanishing of critical situation in intense traffic moments.

Bibliography

[1] AISCAT, *Flussi autostradali*, (2003)

[2] Autostrade per l'Italia, *Flussi medi di traffico giornalieri*,

[3] ATA Engineering, *Situazione dei trasporti su strada tra SS. 47 Valsugana, A 31 Valdastico, A22 Autobrennero e A4 Serenissima*, (1999)

[4] G. Bretti, R. Natalini, B. Piccoli *Numerical approximations of a traffic flow model on networks* Networks and heterogeneous media v.1 n.1 (2006)

[5] G. M. Coclite, Benedetto Piccoli, *Traffic flow on a road network* preprint (2002)

[6] G. M. Coclite, M. Garavello, and B. Piccoli. *Traffic flow on a road network.* SIAM J. Math. Anal., 36(6):18621886 (electronic), (2005)

[7] Dafermos, *Polygonal approximations of solutions of the initial value problem for a conservation law* J. Math. Anal. Appl. 38 (1972), pp. 33-41.

[8] P. Foscari, M. Herty, L. Pareschi *Macroscopic modeling and simulation of the highway network around Ferrara*, technical report.

[9] M. Herty, *Mathematics of traffic flow networks: modeling, simulation and optimization*, Technische Universitaet Darmstadt, (2004)

[10] H. Holden and N. H. Risebro. *A mathematical model of traffic flow on a network of unidirectional roads.* SIAM J. Math. Anal., 26(4):9991017, (1995)

[11] B. Piccoli *Traffic Flow on Networks: Conservation Laws Models* - 12th International Conference on Hyperbolic Problems: Theory, Numerics, Applications - 2008

[12] B. Piccoli and M. Garavello *Traffic Flow on Networks* - AIMS Applied Mathematics V.1

[13] Provincia di Ferrara, *Rilevazioni sul Ferrarese 1992-2003*,

[14] Provincia di Mantova - Assessorato alle Politiche Ambientali, *Rapporto sullo stato dell'ambiente nel territorio mantovano*, (1999)

[15] Regione Emilia Romagna, *Rilevazioni sul bacino della Romea*, (2003)

[16] Regione Emilia Romagna, *Studio di fattibilitá E55*, (2003)

Chapter 3

Kinetic models for economics

3.1 Introduction

Economic activity is naturally composed of the set of binary exchanges of goods within varying couples of subjects in a society or even between different countries. Even though it would not be feasible to model each of them individually - both for the large numbers involved and the amount of different parameters each trade arguably depends on - it has been noted that some of their aggregated effects can be observed more or less repeating among time and space, and thus lead themselves to modeling. In particular the income and wealth distribution show recurrent inequalities favouring a small minorance of extremely rich people, regardless of the social structure in the situations considered.

This raises the question of whether said invariants can be due to common properties of all market economies once certain base criteria are satisfied. The italian economist Vilfredo Pareto in [22] already noted that such forces seem to be acting similarly in different situations:

It has been noticed long ago that it would be useless to operate a goods subdivision aiming at an evenly distributed wealth. After a while the destroyed unevenness would appear restored. or also in [23]: *La tendenza che ha la popolazione a disporsi secondo una certa forma riguardo alle entrate ha per conseguenza che le modificazioni recate a certe parti della curva delle entrate si ripercuotono sulle altre; onde, in ultimo, la societá riprende l'usata forma* [1].

A kinetic description for market economies seems to be a natural path to follow,

[1] The tendency for population to adjust towards a certain income distribution implies that changes to a part of the income curve are reflected on the remaining ones, so that ultimately society recovers the usual shape.

because it directly mimics the dynamics analysed. By studying the minimal and competing conditions required to reproduce the aimed invariant outcomes it is possible to develope insight into the actual phenomenon and discard unessential prerequisites or overfitting ad-hoc theories.

In this context many techniques initially developed within statistical physics can prove useful. As noted above:

Statistical physicists have determined that physical systems which consist of a large number of interacting particles obey laws that are independent of the microscopic details. This progress was mainly due to the development of scaling theory. Since economic systems also consist of a large number of interacting units, it is plausible that scaling theory can be applied to economics (Stanley et al., 1996).

Many microscopic models have been presented over the last years in particular with the emergence of the so called *econophysics* studies, for some overviews see [25], [14], [20] and [24].

3.1.1 The mesoscopic approach to wealth distribution

In kinetic modeling for economics it is possible to proceed analogously to the case of traffic flow and to the theory of rarified gases. The kinetic density $f(\gamma, w, t)$ describes the distribution of economic agents and is thus always non-negative. Here w represents wealth, t is time as usual, while γ is a third parameter measuring i.e. the trading or investment propensity, or social class. Just like velocity in traffic modeling here wealth is a non negative quantity; moreover it is used as a single unifying measure of all different assets owned and exchenged by each individual.

A mesoscopic description seems to be much better justified for any non-local economy than in the case of traffic flow, due to the significantly larger number of interacting agents actually involved in the real phenomenon to be reproduced: more orders of magnitude separate the few hundred vehicles on a roadway from the hundreds of thausands or more individuals who are partecipating to economic activities.

Sometimes even a simpler kinetic density in the $f(w, t)$ form will suffice, as even basic models are able to fit the empirical data showing Pareto tails for agents with large wealth, that is an inverse power law decay. In particular one has $F_c(w) := Pr[X > x] \sim w^{-\alpha}$, as detailed in the next section.

Data fitting for the tails of such distributions must be handled with care: it has been pointed out (cfr. [19]) how easy it is to fit power laws to relatively small lognormal data sets. This is where kinetic derivations from simple microscopic generative models prove useful by pointing to the effects of basic and widespread underlying

mechanisms and thus providing grounds for confidence in the model itself.

In the following the derivation of kinetic description and its properties will be obtained starting from various models of the microscopic dynamics.

3.1.2 Power laws and scaling

The distribution of income and wealth in various countries and times shares similarities in the higher tail behaviour with many other real world distributions[2].
The common phenomenon of wealth condensation - by which a small minority of individuals controls the mayority of the resources - was studied for income already by the economist Vilfredo Pareto, who proposed as model the following inverse power law, which thereafter took his name:

$$f(x) = \alpha b^{\alpha} x^{-\alpha-1}$$

for $x \geq b$. The corresponding cumulative density function is thus simply $F(x) = 1 - b^{\alpha}x^{-\alpha}$, from which the complementary CDF immediately follows:

$$F_c(x) := Pr[X > x] = b^{\alpha}x^{-\alpha}$$

This representation makes the scale invariance property evident:

$$Pr[X > \beta x] = (b/\beta)^{\alpha}x^{-\alpha} \propto Pr[X > x]$$

Such decay is slower than that of a normal distribution, and is therefore said to be proper of *fat or heavy tails*. The parameter α is often within the range $(0, 2]$. On a log-log scale a power law shows up as a straight line.

A Pareto distribution is only a good fit for large incomes, while the lower majority shows up a lognormal behaviour. This combination of lognormal and power law turns out to be a good model for wealth distribution in most developed countries and across time, from ancient Egypt to Europe, Japan and USA, aswell as in developing nations. Empirically the exponent α for wealth is often measured to be about 2 in western countries, while for financial returns it is around 3 (cfr. [17]).

[2] In particular power laws have been observed in distributions as diverse as that of book sales, stock returns, lunar craters, telephone calls, scientific citations, earthquake severity, sand grain sizes, city sizes, web sites accesses, solar flares, word frequencies etc... In fact it has been stated (Levy-Solomon 1996) that also in multiplicative processes *"power-like systems are expected to arise as naturally as the Boltzmann distribution"*.

3.2 Conservative wealth exchanges

We first review the most fundamental models for wealth exchange and the corresponding resulting distributions, upon wich more elaborate cases are later built. The transaction rules considered here are initially conservative, so that a certain amount of money $\delta_{1,2}$ will change hands passing from an agent with wealth w_1 to another one with wealth w_2:

$$w_1 \to w_1' = w_1 - \delta_{1,2} \qquad w_2 \to w_2' = w_2 + \delta_{1,2} \tag{3.1}$$

Such prototype transaction will be called admissible and thus effectively occur provided that both $w_i \geq 0$. This accounts for prohibiting debt and short selling.

The exchanged amount $\delta_{1,2}$ can be of various type, i.e. of the following form:

$$\delta_{1,2} = \theta C$$

with $C \geq 0$ and θ a random variate uniformly distributed on the interval $(-1, 1)$. As an example one can set:

$$\delta_{1,2} = \theta(w_1 + w_2)$$

Local conservation of money for the couple of agents is always guaranteed, being explicit in equation 3.1 where only a transfer occurs without dissipation or gains, so that:

$$w_1 + w_2 \to w_1' + w_2' = w_1 - \delta_{1,2} + w_2 + \delta_{1,2} = w_1 + w_2 \tag{3.2}$$

Making also use of local wealth conservation in the θ term one can verify reversibility of the interaction in this particular case:

$$\begin{gathered}
(w_1,\, w_2) \\
\Downarrow \\
(w_1 - \theta_1(w_1 + w_2),\, w_2 + \theta_1(w_1 + w_2)) \\
\Downarrow \\
(w_1 - \theta_1(w_1 + w_2) + \theta_2(w_1 + w_2),\, w_2 + \theta_1(w_1 + w_2) - \theta_2(w_1 + w_2)) \\
= \\
(w_1, w_2)
\end{gathered} \tag{3.3}$$

The last equality is always possible with $\theta_2 := -\theta_1$, because the corresponding

"backward" tentative interaction is always defined and admissible by construction.

From this microscopic interaction rule it is possible to introduce a Boltzmann-like kinetic evolution equation for the wealth density f:

$$\partial_t f = \int_{-1}^{1} \int_{0}^{\infty} \beta_{(w,w_*)\leftrightarrow(w',w'_*)} [f(w')f(w'_*) - f(w)f(w_*)]\, dw_*\, d\theta, \tag{3.4}$$

In the case considered the money transfer rate $\beta_{(w,w_*)\leftrightarrow(w',w'_*)}$ is defined as follows, to simply select admissible transactions through the indicator function Ψ:

$$\beta_{(w,w_*)\leftrightarrow(w',w'_*)} = \Psi(w' \geq 0)\Psi(w'_* \geq 0),$$

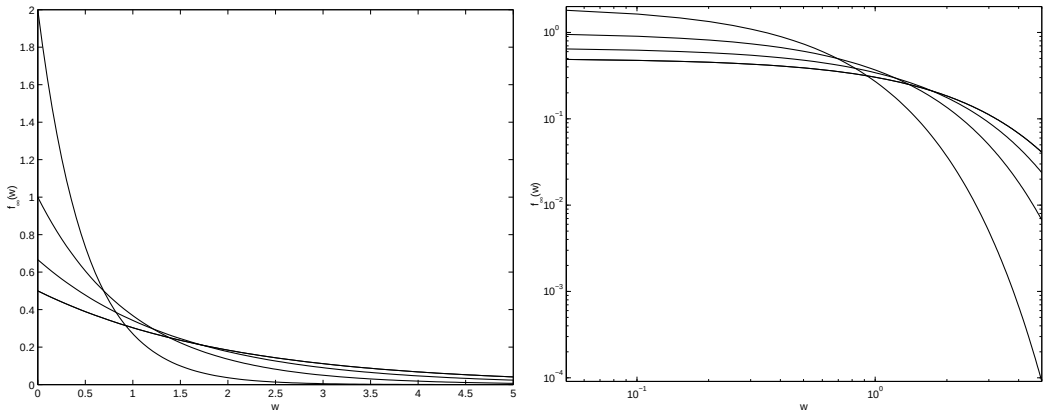


Figure 3.1: *The stationary Boltzmann-Gibbs money distribution for time reversible models in linear (left) and log-log scale (right) with* $\rho = 1$ *and various values of* $\bar{w} = 0.5, 1, 1.5, 2$.

We can define the first two moments of the wealth distribution as follows, the agent density ρ and total money amount $\overline{w}$:

$$\rho = \int_0^\infty f(w)\, dw,$$

$$\overline{w} = \frac{1}{\rho} \int_0^\infty f(w) w\, dw,$$

This model converges to a stable and unique stationary state following a Boltzmann-Gibbs distribution, that is an inverse exponential decay, as shown i.e. in [8]:

$$f_\infty(w) = \frac{\rho}{\bar{w}} e^{-w/\bar{w}}. \tag{3.5}$$

Time reversible interactions

By loosening the reversibility requirement we can allow for more versatile and interesting models. The simplest possibility has been described in [15] and [16], where the following microscopic money transfer is considered:

$$\delta_{1,2} = \begin{cases} \gamma w_1, & \theta < 0 \\ -\gamma w_2, & \theta > 0. \end{cases} \tag{3.6}$$

That is, in practice a constant wealth portion of an agent is transferred to a random other one. This form makes the agent simmetry explicit, but since it is already present in the random choice of agents for interaction it could be enough to set simply $\delta_{1,2} = \gamma w_1$.

A slightly more complex model is the following, presented in [5], and including a marginal saving component:

$$\delta_{1,2} = \gamma[\epsilon w_1 - (1-\epsilon)w_2] \tag{3.7}$$

Here ϵ is an uniform random variate over the interval $(0, 1)$ and controls a convex combination of the two transfers in eq. 3.6, that is a gain with weight ϵ and a loss with the complementary weight. The expected value of the money transfer 3.7 is then comparable to that of eq. (3.6):

$$\mathbb{E}[\delta_{1,2}] = \gamma\{\mathbb{E}[\epsilon w_1] - \mathbb{E}[(1-\epsilon)w_2]\} = (w_1 - w_2)\gamma/2$$

The money transfer can be rewritten as follows:

$$\delta_{1,2} = \gamma w_1 - (1-\epsilon)\gamma(w_1 + w_2) = -\gamma w_2 + \epsilon\gamma(w_1 + w_2) \tag{3.8}$$

corresponding to the interaction:

$$(w_1, w_2) \to ((1-\gamma)w_1 + (1-\epsilon)\gamma(w_1 + w_2), (1-\gamma)w_2 + \epsilon\gamma(w_1 + w_2)) \tag{3.9}$$

This formulation clarifies the roles of γ and ϵ: the former controls the amount of money each agent is willing to invest in the trade, saving the remaining $(1-\gamma)$ quota, while the latter introduces randomness in the trade itself where each agent can gain a portion ϵ of the jointly invested amount. This model has also the advantage that by construction both proposed wealths are positive, so that each interaction is always admissible. For $\gamma = 1$ this *marginal saving propensity* is absent and the

model reduces to that of Angle (see [2], [1], [3], [4]), which asimptotycally leads again to a Boltzmann-Gibbs distribution.

The interaction rule 3.9 leads to the following corresponding kinetic equation:

$$\partial_t f = \int_0^1 \int_0^\infty (\beta_{'\mathbf{w}\to\mathbf{w}} J f('w) f('w_*) - \beta_{\mathbf{w}\to\mathbf{w}'} f(w) f(w_*))\, dw_*\, d\epsilon, \tag{3.10}$$

with $'\mathbf{w} := ('w,' w_*)$ being the agent wealths before the trade which results in the wealths $\mathbf{w} := (w, w_*)$, and clearly with the post trade money couple $\mathbf{w}' := (w', w'_*)$. The Jacobian J for this model is $J = 1/(1-\gamma)$, while the pre-trading wealths are given by:

$$'w = \frac{w - (1-\epsilon)\gamma(w+w_*)}{1-\gamma}, \quad 'w_* = \frac{w_* - \epsilon\gamma(w+w_*)}{1-\gamma}.$$

The transition rates are the following:

$$\beta_{'\mathbf{w}\to\mathbf{w}} = \Psi('w \geq 0)\Psi('w_* \geq 0), \quad \beta_{\mathbf{w}\to\mathbf{w}'} = 1.$$

This kinetic evolution leads in the time limit to a defined stationary state. The first three moments of f can be simply calculated directly: the first two are conserved while the second converges exponentially towards a constant value. Higher order moments also converge to those of the stationary state. Such limit distribution has

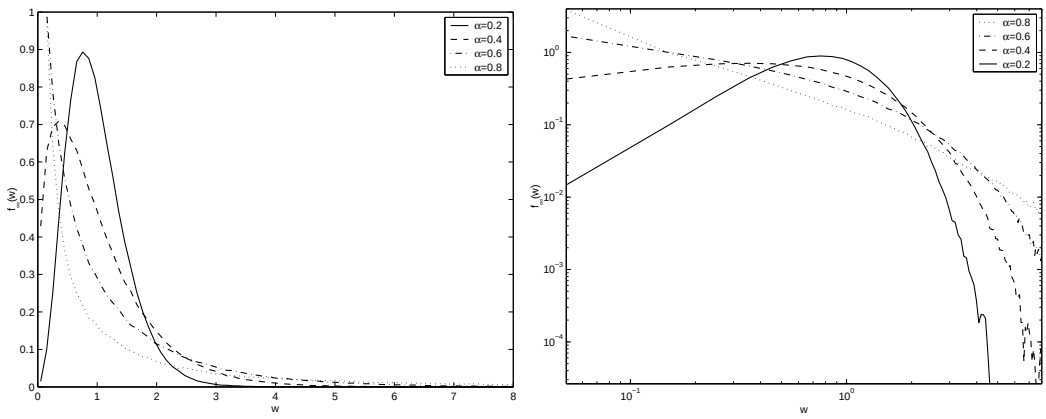


Figure 3.2: *The stationary money distribution for model (3.7) in linear scale (left) and log-log scale (right) with* $\rho = 1$, $\overline{w} = 1$ *and* $\alpha = 0.2, 0.4$. *The large time behavior for* $\alpha = 0.6, 0.8$ *is also shown.*

been empirically well fitted by a gamma distribution, and a numerical evaluation of

its shape is depicted in Figure 3.2. In this case for small values of w the distribution starts as w^λ with $\lambda = -1 - \ln 2/\ln(1-\gamma)$, and the higher follows an exponential decay of Boltzmann-Gibbs type. As anticipated for $\gamma = 1$ the whole distribution reduces to a Boltzmann-Gibbs law. On the other hand for small values of γ the shape of f is similar to a log-normal distribution with the following form:

$$f(w) = \frac{1}{w\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\log(w/\bar{w}))^2}{2\sigma^2}\right), \tag{3.11}$$

with σ^2 being the variance of f.

Inhomogeneous transactions

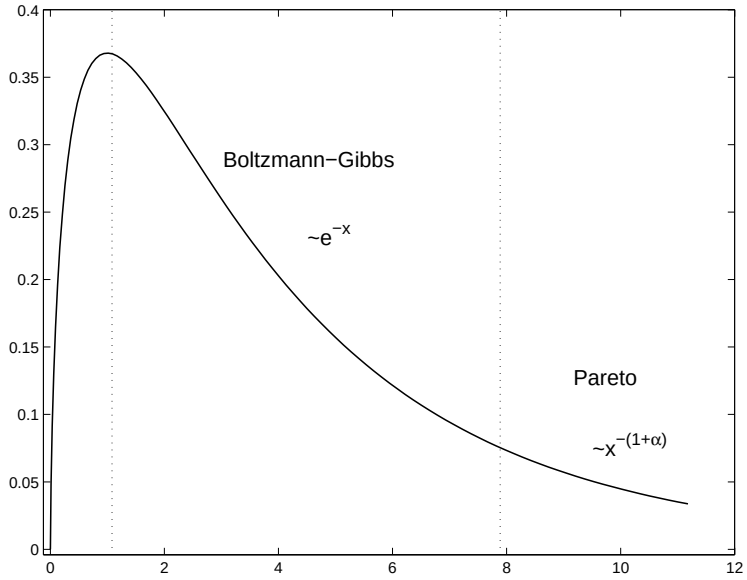


Figure 3.3: *The expected wealth/income distribution behavior. Normal scale (left) and log-log scale (right).*

A natural extension of (3.7) would allow trading propensity γ to vary and be for example agent dependent (while constant in time). The resulting trading rule, analyzed in [5], would then read:

$$\delta_{1,2} = \gamma_1 \epsilon w_1 - \gamma_2(1-\epsilon)w_2 \tag{3.12}$$

The corresponding trading propensities γ_i would in this case be set initially according to a given distribution and kept unchanged; for this the simplest option would be

clearly a uniform distribution on an interval contained in $(0,1)$. Another possible choice is a power law restricted to the same interval $\gamma^\nu : \ \gamma \in (0,1)$. In the case positive values of ν the model keeps the decribed behaviour, with an asymptotic distribution decaying as a Pareto power law w^{-2}, while for a negative ν parameter a Boltzmann-Gibbs like component appears at moderate w values (while higher areas mantain their Pareto decay). This kind of "phase" transition seems to be originated by the relative preponderance of agents with low saving propensity.

The form of the kinetic equation when inhomogeneous γ-s are introduced does not change apart from the Jacobian and pre-trading wealth relation.

3.2.1 A kinetic model

All previous models analysed in their microscopic trading form can be restated within a unique kinetic form. To this aim we first introduce a generalized microscopic interaction rule subsuming them, and then pass to the corresponding Boltzmann equation. As we are considering conservative models we can still concentrate on the amount of money transfered at each trade between agents having wealths w and w_*, this time in a more generic form: $\delta(w, w_*; \gamma, \vartheta)$. The corresponding interaction rule then reads:

$$w' = w - \delta(w, w_*; \gamma, \vartheta), \qquad w'_* = w_* + \delta(w, w_*; \gamma, \vartheta) \tag{3.13}$$

The trade function $\delta(w, w_*; \gamma, \vartheta)$ could also depend on further parameters, as i.e. $\overline{w}$. Here the trade coefficient is restricted to the interval $0 < \gamma(w, w_*) < 1$, while $\vartheta \in \mathbb{R}$ is a parameter controlling the risk of the trade. We recall that the interaction takes place only when both new wealths are non-negative, to avoid the presence of debts. Such a generic trade function comprehends all previously described models as special cases.

Since no particular money unit was fixed, all equations describing the evolution of wealth should then be invariant to homogeneous scalings of individual holdings, which means for the trade function to be scale invariant:

$$\delta_{\cdot,*}(\lambda w, \lambda w_*; \gamma, \vartheta) = \lambda \delta_{\cdot,*}(w, w_*; \gamma, \vartheta).$$

We can state a corresponding Boltzmann model for the dynamics of the wealth density function $f(w,t)$:

$$\partial_t f(w,t) = G(f, w; \gamma) - L(f, w; \gamma) = Q_E(f, f)(w), \tag{3.14}$$

Clearly here $G(f,w;\gamma)$ represents a gain term for money exchanges ending with an agent having money w, while the loss term $L(f,w;\gamma)$ accounts for trades starting from an agent with money w. These two terms can be explicitly stated by the usual parametrization of the remaining variables as follows:

$$G(f,w;\gamma) = \int_{\mathbb{R}}\int_0^{\infty} \beta_{('w,'w_*')\to(w,w_*)} J f('w) f('w_*)\, dw_*\, d\vartheta \tag{3.15}$$

$$L(f,w;\gamma) = \int_{\mathbb{R}}\int_0^{\infty} \beta_{(w,w_*)\to(w',w_*')} f(w) f(w_*)\, dw_*\, d\vartheta \tag{3.16}$$

with $('w,' w_*)$ being the pre-trade wealth pair and J the Jacobian of the transformation from (w, w_*) into (w', w_*'), which explicitely reads:

$$J = \begin{pmatrix} 1 - \dfrac{\partial\delta(w,w_*;\gamma,\vartheta)}{\partial w} & -\dfrac{\partial\delta(w,w_*;\gamma,\vartheta)}{\partial w_*} \\ \dfrac{\partial\delta(w,w_*;\gamma,\vartheta)}{\partial w} & 1 + \dfrac{\partial\delta(w,w_*;\gamma,\vartheta)}{\partial w_*} \end{pmatrix}$$

At this point it's possible and convenient (i.e. to avoid the Jacobian) to study the weak formulation, as in the following

Lemma 3.2.1 *Given an arbitrary test function $\phi(w)$ a weak solution $f(w,t)$ to (3.14) satisfies the following identity*

$$\int_{\mathbb{R}^+} Q_E(f,f)(w)\phi(w)\, dw = \frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}^+}\int_{\mathbb{R}^+} \beta_{(w',w_*')\leftrightarrow(w,w_*)} f(w) f(w_*)\cdot \\ \cdot\left[\phi(w') + \phi(w_*') - \phi(w) - \phi(w_*)\right]\, dw_*\, dw\, d\vartheta. \tag{3.17}$$

Furthermore the following reversibility assumptions on the trade function

$$i) \qquad \delta(w,w_*;\gamma,\vartheta) = \delta(w',w_*';\gamma,\vartheta), \tag{3.18}$$

$$ii) \qquad \frac{\partial\delta(w,w_*;\gamma,\vartheta)}{\partial w} = \frac{\partial\delta(w,w_*;\gamma,\vartheta)}{\partial w_*}, \tag{3.19}$$

lead to the identity

$$\int_{\mathbb{R}^+} Q_E(f,f)(w)\phi(w)\, dw = -\frac{1}{4}\int_{\mathbb{R}}\int_{\mathbb{R}^+}\int_{\mathbb{R}^+} \beta_{(w',w_*')\leftrightarrow(w,w_*)}[f(w')f(w_*') \\ -f(w)f(w_*)]\left[\phi(w') + \phi(w_*') - \phi(w) - \phi(w_*)\right]\, dw_*\, dw\, d\vartheta. \tag{3.20}$$

It turns out that there are two only trade invariant functions, that is for which

$$\int_{\mathbb{R}^+} Q_E(f,f)(w)\phi(w)\,dw = 0,$$

namely $\phi(w) = 1$ and $\phi(w) = w$, corresponding to conservations of total agent number and of wealth respectively.

Asymptotics for reversible models

Whenever the two conditions (3.18) and (3.19) are satisfied, the weak evolution equation (3.20) with test function $\phi(w) = \log f(w)$ gives the following entropy inequality, analogous to the Boltzmann inequality in the theory of dilute gases:

$$\begin{aligned}\int_0^{+\infty} Q_E(f,f)(w)\log f(w)\,dw \;&=\; -\frac{1}{4}\int_{\mathbb{R}}\int_{\mathbb{R}_+}\int_{\mathbb{R}_+} \beta_{(w',w'_*)\leftrightarrow(w,w_*)}[f(w')f(w'_*) \\ &\quad -f(w)f(w_*)] \qquad \log\left(\frac{f(w')f(w'_*)}{f(w)f(w_*)}\right) dw_*\,dw\,d\vartheta \le 0,\end{aligned} \tag{3.21}$$

because due to the monotonicity of logarithm $(x-y)\log(x/y) \ge 0$ always holds, with equality satisfied iff $x = y$. Moreover this equality is satisfied whenever $\log f(w)$ is an invariant, that is if:

$$\log f(w) = c_1 - c_2 w$$

where c_1 and c_2 are nonnegative constants determined by normalization criteria, such that $c_1 = \log(\rho/\overline{w})$ and $c_2 = -1/\overline{w}$. Therefore wealth follows a Boltzmann-Gibbs distribution as in (3.5). This result is robust being independent from both the specific initial wealth distribution and from the trading kernel $\beta_{(w',w'_*)\to(w,w_*)}$.

We can also introduce the entropy function $S(f) = -\int_{\mathbb{R}^+} f(w)\log f(w)dw$, for which the following inequality holds:

$$\frac{dS(f)}{dt} = -\int_0^{\infty} Q_E(f,f)(w)\log f(w)\,dw \ge 0.$$

Therefore entropy also keeps growing independently of the specific wealth distribution, until this reaches the stationary state, in which case the entropy is given by:

$$S(f_\infty) = \rho\left(\log\left(\frac{\rho}{\overline{w}}\right) - 1\right).$$

Non reversible markets and risk

Whenever (3.18) and (3.19) are not satisfied - so that the trade rule is not reversible - it's no longer possible to obtain the previous entropy principle for gaining insight in long time asymptotics. However we can still obtain a better understanding of the trade dynamic in a conservative setting. In [20] the following generalized splitting of the trade rule is suggested:

$$w' = \underbrace{(1-\gamma(w,w_*))w}_{saving\ propensity} + \underbrace{\gamma(w,w_*)w_*}_{potential\ transaction} + \underbrace{\theta\, r(w,w_*)}_{risk}. \tag{3.22}$$

This marginal saving propensity is however different than in (3.18) and [5]. As examples the previously seen trading rules (3.7) and (3.9) result in the following terms:

$$\begin{aligned} w' &= (1-\frac{\gamma}{2})w_i + \frac{\gamma}{2}w_j + \theta\frac{\gamma}{2}(w_i+w_j) \\ w' &= (1-\frac{\gamma_i}{2})w_i + \frac{\gamma_j}{2}w_j + \theta\frac{1}{2}(\gamma_i w_i + \gamma_j w_j) \end{aligned} \tag{3.23}$$

This splitting of microscopic interactions results in relations symmetric with respect to the two individual wealths, with the exception of the risk term which changes sign upon swapping of the agents.

The risk term is controlled by a random variable, namely $\theta \sim \mu(\theta)$, characterized by the first two moments 0 and σ^2; the variance determines an average size for the random transaction component which on the other hand is also linked (and possibly proportional) to the starting wealths involved in the exchange. Examples of risk functions are for example the average money used in [5]:

$$r(w,w_*) = (w+w_*)/2 \tag{3.24}$$

and the minimun money used i.e. in [26]:

$$r(w,w_*) = \min\{w,w_*\}$$

In [26] also a money dependent trade coefficient is considered:

$$\gamma(w,w_*) = \sigma\frac{\min\{w,w_*\}}{w+w_*}$$

which further suggests adopting the following general and convenient form for the

risk term:

$$r(w, w_*) = \gamma(w, w_*)\frac{w + w_*}{\sigma}$$

Here a higher saving propensity ($\gamma(w, w_*) \approx 0$) will correspond to a lower amount of risk, and viceversa. Setting a constant transition rate $\gamma = \tilde{\gamma}/2$ we get again the risk function (3.24) which is used in interaction rule (3.7).

The transition rate in the kinetic formulation will clearly inherit the characteristics of such formulation. It can be expressed as:

$$\beta_{(w,w_*)\to(w',w'_*)} = B(w, w_*, w', w'_*)\,\mu(\theta)$$

Here $B(w, w_*, w', w'_*)$ represents the probability of interaction for a given couple of agents.

Linear trading asymptotycs

Some further analysis can be carried out for trading rules of the following form:

$$w'_1 = (1 - \gamma)w_1 + \gamma w_2 + \theta[\lambda w_1 + (1 - \lambda)w_2] \tag{3.25}$$

with $\lambda \in [0, 1]$ defining the risk. The target wealths produced are always admissible iff $\theta \in [\max\{\gamma/(\lambda - 1), (\gamma - 1)/\lambda\}, \min\{(1 - \gamma)/(1 - \lambda), \gamma/\lambda\},]$ For the case i.e. of $\lambda = 1/2$ and $\gamma \leq 1/2$ this results in $\theta \in [-2\gamma, 2\gamma]$. In the following a simplified transition rate will be considered which includes such positivity constraint but is independent of agent wealths: $\beta_{(w,w_*)\to(w',w'_*)} = \mu_{\gamma,\lambda}(\theta)$.

One can then study the dynamics of the moments for the wealth distribution:

$$m_k(t) = \int_0^\infty f(w, t)w^k\,dw, \quad k \in \mathbb{N}$$

Conservation of the first two moments m_0 and m_1 (number of agents and mean wealth) is guaranteed by construction, so that through a normalization we can write $m_0 = 1$ and $m_1 = \overline{w}$ without losing in generality. The higher moments can be obtained recursively through 3.17 and the derived relations

$$\frac{dm_k}{dt} + A_k m_k = \sum_{j=1}^{k-1} \binom{k}{j} A_{j,k-j} m_j m_{k-j}, \quad k \geq 2$$

with $p = 1 - \gamma + \lambda\theta,\ q = \gamma + \theta(1-\lambda)$ and

$$\begin{aligned} A_k &= \frac{1}{2}\int_{\mathbb{R}} \mu(\theta)(2 - p^k - q^k - (1-p)^k - (1-q)^k)\, d\vartheta, \\ A_{j,k-j} &= \frac{1}{2}\int_{\mathbb{R}} \mu(\theta)(p^j q^{k-j} + (1-p)^j(1-q)^{k-j}) d\theta, \quad j = 1, \dots, k-1. \end{aligned}$$

In the particular case $\mathbb{E}[\theta] = 0$ and $\mathbb{V}[\theta] = \sigma^2$ (the variance) we get for m_2 the following:

$$\begin{aligned} &A_2 = 2\sigma^2\lambda(1-\lambda) - \sigma^2 + 2\gamma(1-\gamma), \\ &A_{11} = \sigma^2\lambda(1-\lambda) + \gamma(1-\gamma), \\ &m_2(t) = e^{-A_2 t} m_2(0) + (1 - e^{-A_2 t}) \left(1 + \frac{\sigma^2}{A_2}\right) \overline{w}^2. \end{aligned} \tag{3.26}$$

Therefore whenever A_2 is positive this moment approaches at an exponential rate a constant value. In the same way all finite higher moments converge exponentially towards a definite value whenever $A_k > 0$, reaching thus a stationary state.

All this depends on the aforementioned values $A_\cdot$, which in turn are determined by λ, γ and $\mu(\theta)$. In particular for $\lambda = 1/2$ and θ uniform over $(-2\gamma, 2\gamma)$ (thus with variance $\sigma^2 = 4\gamma^2/3$), the binary interaction (3.7) is again recovered (with $\gamma/2 \to \gamma$) whose moments converge for:

$$A_k = \frac{1}{2\gamma(k+1)} \left(2\gamma(1+k) + (1-2\gamma)^{k+1} - 1 - (2\gamma)^{k+1}\right) \geq 0,$$

with $\gamma \in [0, 1/2], \quad k \geq 2$. Furthermore, for $\gamma = 1/2$ this reduces to $A_k = (k-1)/(k+1)$ so that the stationary state for the system follows again a Boltzmann-Gibbs distribution.

In the case $\lambda = 1/2$ and $\gamma \leq 1/2$ the requirement $A_2 \geq 0$ implies $\sigma^2 \leq 4\gamma(1-\gamma)$; such condition is luckily satisfied by any probability density over $[-2\gamma, 2\gamma]$ having zero mean. Similar conditions also lead to such convergence for different values of λ.

3.3 Wealth distribution in an open economy

3.3.1 The microscopic interaction

In the following I describe the Cordier-Pareschi-Toscani kinetic model for an open market economy (introduced in 2005 in [6], in short "CPT" from here on) and some of its asymptotics. This is also related to the results of Bouchaud and Mezard in [1].

A very simple rule for binary interactions is already able to account for various aspects of economic activity and wealth distribution.

$$w_1' = (1-\gamma)w_1 + \gamma w_2 + \eta_1 w_1$$
$$w_2' = (1-\gamma)w_2 + \gamma w_1 + \eta_2 w_2$$

Here $0 < \gamma \leq 1/2$ is a fixed constant determining the wealth quota exchanged, while η_i-s are random variables from a common distribution $\Theta(\cdot)$ with null mean and variance σ^2. No debts will be allowed, so that this proposal mapping is accepted if the resulting new wealths w_i' are both positive, otherwise further tentatives are drawn (since this depends on the outcomes of η_i only, for suitable choices of Θ this condition can always be satisfied).

The first two r.h.s. terms tend to redistribute wealth among all agents, leading in the limit to a stationary uniform state: in fact with $\delta := w_2 - w_1$ the microscopic interaction reads $w_i' = w_i \pm \gamma\delta + \eta_i w_i$ so that differences are spread around and thus eliminated, whenever $\gamma < 1$. Moreover the dynamics embodied in these two terms would also be conservative of wealth.

The third term involves a risk and models a speculative factor, where gains and losses are proportional to the starting wealth; therefore the system will be an open economy with exogenous factors. Because of the discarding of trial configurations with negative values the total amount of money will be increasing. This asymmetry is what will cause difficulties in the kinetic analysis.

3.3.2 Kinetic analysis

The kinetic distribution of wealth $f(w,t)$ evolves according to the following Boltzmann-like integro-differential equation:

$$\partial_t f = \int_{\mathbb{R}^2} \int_0^\infty \left[\beta_{'\mathbf{w}\to\mathbf{w}} J^{-1} f('w_1) f('w_2) - \beta_{\mathbf{w}\to\mathbf{w}'} f(w_1) f(w_2)\right] dw_2 d\eta_1 d\eta_2 \quad (3.27)$$

with the couple $'\mathbf{w} = ('w_1,' w_2)$ representing interacting money which leads to the new quantities $\mathbf{w} = (w_1, w_2)$, while $J = (1-\gamma+\eta_1)(1-\gamma+\eta_2)-\gamma^2$ is the Jacobian of the transformation from $\mathbf{w}$ to $\mathbf{w}'$. The Jacobian in the gain term is needed to guarantee conservation of mass (in this case being the number of agents), independently from the choice of β.

In the following the case of a transition rate of the form

$$\beta_{\mathbf{w}\to\mathbf{w}'} = \Theta(\eta_1)\Theta(\eta_1)\Psi(w_1' \geq 0)\Psi(w_2' \geq 0)$$

will be considered, with $\Psi(A)$ being the indicator function of the set A. Thus here the rate function $\beta_{\mathbf{w}\to\mathbf{w}'}$ embodies the effects of the opennes of the modeled economy described by the random variates, and handles the constraint of trading to positive arrival wealths. However in general the trade rule could also include further components, like risks, taxes and subsidies.

In general the rate kernel $\beta_{\mathbf{w}\to\mathbf{w}'}$ depends on $\mathbf{w}'$, but whenever the random variables η_i have density limited to the interval $[-(1-\gamma), (1-\gamma)]$ then $w_i' \geq 0$ will always hold so that $\beta_{\mathbf{w}\to\mathbf{w}'}$ becomes independent from $\mathbf{w}'$. In this case many simplifications are then possible.

Studying the weak solution to the initial value problem corresponding to the previous equation one can see how the total amount of money is increasing exponentially. In particular, if X is a random variable of density $\Theta(\eta)$, taking values on a interval $(-a, a)$, with $a > 1$,

$$\int_{\mathbb{R}} \eta\Theta(\eta)\Psi\left(\eta > 1\right) d\eta = A > 0, \quad (3.28)$$

one can prove the following:

Theorem 3.3.1 *Let the probability density $f_0 \in \mathcal{M}_p$, where $p = 2 + \delta$ for some $\delta > 0$, and let the symmetric random variable Y which characterizes the kernel have a density Θ in $\mathcal{M}_{2+\alpha}$, with $\alpha > \delta$. Then, if Θ has unbounded support, the average*

wealth is increasing with time at least exponentially

$$\int_{\mathbb{R}_+} w f(w,t)\,dw \geq \exp\left\{\frac{A}{4}\,t\right\} \int_{\mathbb{R}_+} w f_0(w)\,dw, \tag{3.29}$$

where the constant A is given by (3.28). Moreover, the average wealth does not increase more than exponentially in time

$$\int_{\mathbb{R}_+} w f(w,t)\,dw \leq \exp\left\{\frac{\sigma^{2+\alpha}}{(1-\gamma)^{1+\alpha}}\,t\right\} \int_{\mathbb{R}_+} w f_0(w)\,dw. \tag{3.30}$$

Similarly, higher order moments does not increase more than exponentially, and the bound

$$\int_{\mathbb{R}_+} w^q f(w,t)\,dw\, e^{\left(\frac{1}{2}p(p-1)c_p A_p(\sigma,\gamma)+2\frac{\sigma^{2+\alpha}}{(1-\gamma)^{2+\alpha}}\right)t} \int_{\mathbb{R}_+} w^q f_0(w)\,dw, \tag{3.31}$$

holds for $q \leq p$.

Further bounds can be derived for the remaining moments, but other results prove difficult to obtain.

3.3.3 A solvable limit case

To gain more information on this model one can take an alternative route and look for asymptotics leading to simplified models, for which steady states are easier to find as in the following case. However any asymptotics cannot be strictly stationary, as we've seen that i.e. average wealth is increasing exponentially; however one might look for a proper scaling factoring out such changes, i.e. by analyzing the behaviour of $\tilde{f}(w,t) := m(t)f(m(t)w,t)$ with $m(t)$ being the average amount of money in the modeled economy. Furthermore particular asymptotics can allow for simpler models (often Fokker-Plank) with easier determination of stationary states.

For ease of calculation it is possible to restrict μ to have support included in $[-(1-\gamma),(1-\gamma)]$, which forces admissibility of the new wealths $w_i' \geq 0$; this allows for simple calculation of the following expected values :

$$\mathbb{E}[w_1' + w_2'] = w_1 + w_2$$
$$\mathbb{E}[w_1' - w_2'] = (1-2\gamma)(w_1 - w_2)$$

The former equation is again the conservation of wealth and in this form shows that the introduction of the stochastic terms $\eta_i w_i$ does not affect such property. The second equation describes a tendency for wealth differences to vanish and holds for any assigned distribution Θ; it corresponds to the energy dissipation found in theory of granular cases. However for general Θ the first equation weakens to an inequality allowing for increase of wealth.

We want to study the limit of small exchanges $\gamma \to 0$ while mantaining the previous two properties at the macroscopic level: mass conservation and the variability in time of mean square difference $A_f(t)$, with a decay in absence of the exogenous factor (as with $\sigma = 0$). The first quantity is given in a kinetic form simply as:

$$\int_{\mathbb{R}^2_+} (w + w_*) f(w) f(w_*) dw dw_* = 2 \int_{\mathbb{R}_+} w f(w) dw = 2m(t) \tag{3.32}$$

and $m(t) = m(0)\ \forall t > 0$ is satisfied as soon as the kernel β is independent of wealth variables. The second kinetic quantity reads:

$$A_f(t) := \int_{\mathbb{R}^2_+} (w - w_*)^2 f(w) f(w_*) dw dw_* \tag{3.33}$$

However the limit of interest is that of continuous trading with vanishing σ, but in this case the dynamics of $A_f(t)$ cannot be studied in general. Fortunately one can use a proper scaling $g(w, \tau) = f(w, t)$ with $\tau = \gamma t$ and equivalently analyze $A_g(t)$, so that through $\lambda = \sigma^2/\gamma$ the following relation is obtained:

$$\frac{dA_g(\tau)}{d\tau} = -(4 - 2\lambda) A_g(\tau) + 2\lambda m^2. \tag{3.34}$$

Here γ and σ can be made disappear, but in a controlled way. It results that for values of $\lambda < 2$ a finite limit value of $\lambda m^2/(2-\lambda$ holds for $A_g(\tau)$, while divergence appears for larger λ values.

In the limit the kinetic equation turns into a much simpler Fokker-Planck equation:

$$\frac{\partial g}{\partial \tau} = \frac{\lambda}{2} \frac{\partial^2}{\partial w^2}(w^2 g) + \frac{\partial}{\partial w}((w - m)g)$$

or equivalently

$$\frac{\partial g}{\partial \tau} = \frac{\partial}{\partial w} \left[\left((1 + \frac{\lambda}{2}) w - m \right) g + \frac{\lambda}{2} w \frac{\partial}{\partial w}(wg) \right]. \tag{3.35}$$

The corresponding steady state can be shown to be:

$$g_\infty(w) = \frac{(\mu-1)^\mu}{\Gamma(\mu)} \frac{e^{-(\mu-1)/w}}{w^{1+\mu}}$$

with $\mu = 1 + 2/\lambda > 1$, therefore the power law decay of such distribution depends on γ and σ^2. This limit is justified with the observation that individual practical economic activity is composed of many transactions of small size relative to personal wealth.

In [12] the CPT model has been recently brought to the hydrodynamic limit with the usual techniques borrowed from kinetic theory of rarified gases, through a closure conducted with the analytic solution of the stationary state.

3.3.4 Numerical simulations

To verify the goodness of the Fokker-Planck model it comes natural to compare the results it provides with those coming from a direct Monte Carlo simulation of the original kinetic model. This latter simulation is performed on a pool of $N = 2000$ agents, initially having all the same amount of money; the binary interaction rule is then applied iteratively to randomly selected couples of agents (provided that the resulting wealth pairs are admissible), until the stationary state is reached. At that point the wealth distribution is saved and averaged with others coming from further interactions to limit the Monte Carlo error. This averaged distribution is then normalized and shown in figures 3.4 and 3.5 for different values of the parameters γ and σ^2: the couple (γ, σ^2) is set to $(0.1, 0.2)$ in the former case, to $(0.01, 0.02)$ in the latter.

Thus in both examples the ratio λ remains 2, and corresponds to an also constant coefficient $\mu = 2$, a value chosen to fit well to empirical income distributions observed in real economies. The numerical results confirm the theoretical analysis, in that keeping a constant parameter ratio λ one can observe a corresponding convergence of the Fokker-Plank model to the ”reference” Boltzmann-like one.

Finally the simulated behaviour of mean wealth is also tracked and compared to the expected theorical dynamics: this can be seen in figure 3.6 corresponding to the same parameters of the previous one. As expected the stochastic simulated increase in average wealth is following an exponential growth.

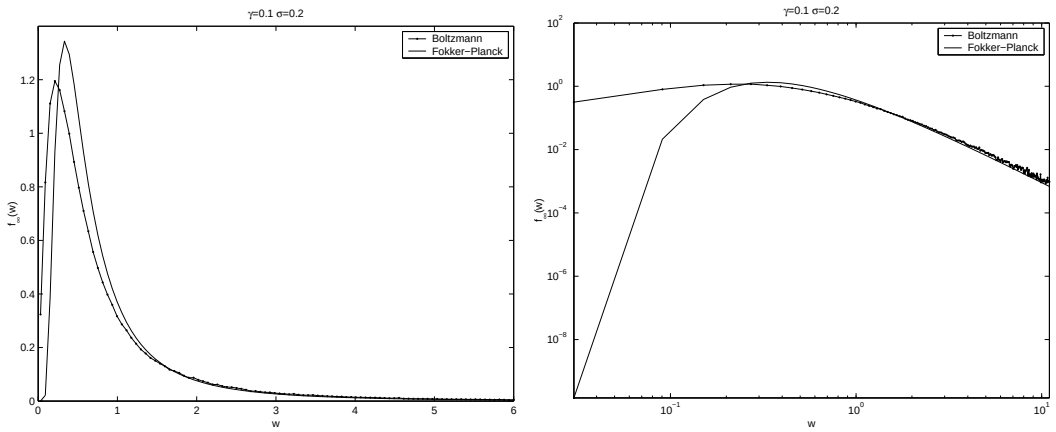


Figure 3.4: *Boltzmann and Fokker-Planck models in the asymptotic limits, for* $\mu = 2.0$, $\gamma = 0.1$ *and* $\sigma = 0.2$

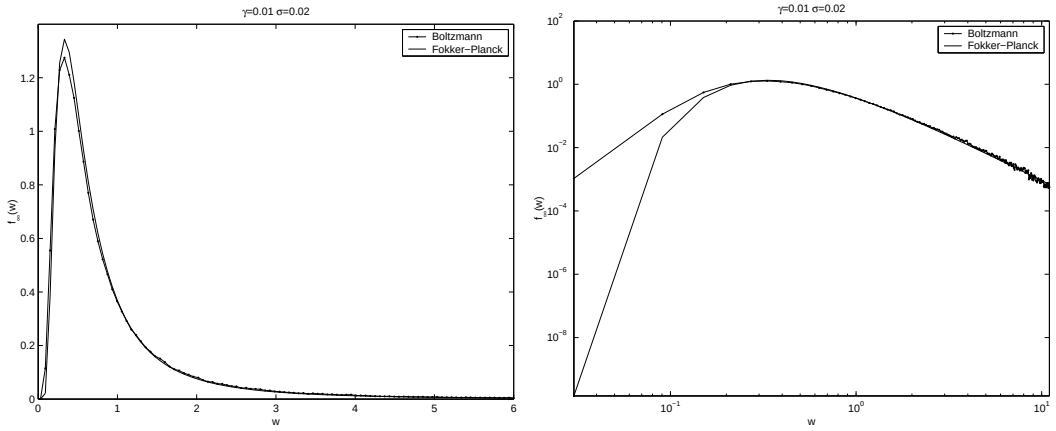


Figure 3.5: *The same as in figure 3.4 with* $\gamma = 0.01$ *and* $\sigma = 0.02$

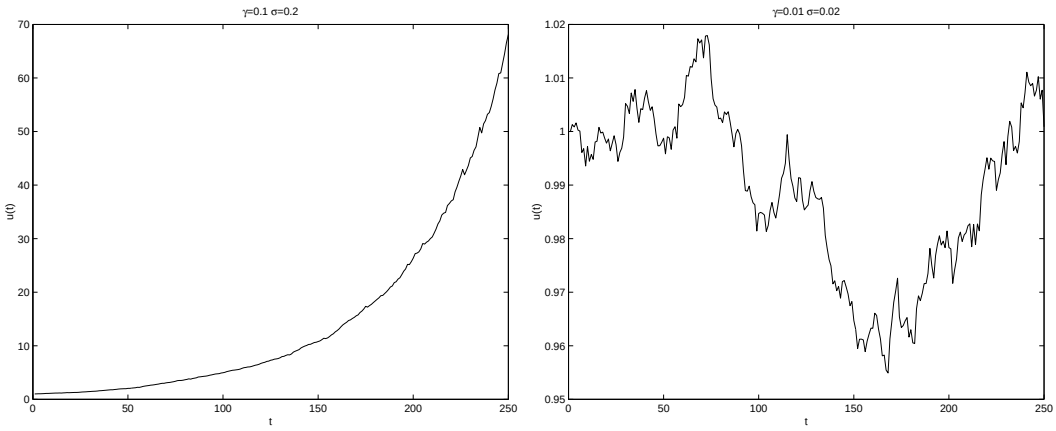


Figure 3.6: *Average amount of money growth as in figure 3.5*

Bibliography

[1] J. Angle, *The surplus theory of social stratification and the size distribution of personal wealth*, Social Forces 65(2), (1986), pp. 293-326.

[2] J. Angle, *The surplus theory of social stratification and the size distribution of personal wealth*, Proceedings of the American Social Statistical Association, Social Statistics Section, Alexandria, VA, (1983) pp. 395.

[3] J. Angle, *The statistical signature of pervasive competition on wage and salary incomes*, I. Math. Sociol. 26 (2002), 217.

[4] J. Angle, *The inequality process as a wealth maximizing process*, Physica A 367 (2006), 388.

[5] A. Chatterjee, B.K. Chakrabarti, S.S. Manna, *Pareto law in a kinetic model of market with random saving propensity*, Phys. A 335 (2004), no. 1-2, 155–163.

[6] S. Cordier, L. Pareschi, G. Toscani, *On a kinetic model for a simple market economy*, Journal of Statistical Physics, 120, (2005), pp. 253-277

[7] S. Cordier, L. Pareschi, C. Piatecki *Mesoscopic modelling of financial markets* (2008)

[8] A. A. Dragulescu and V. M. Yakovenko *Statistical mechanics of money* The European Physical Journal B, v. 17, pp. 723-729 (2000),

[9] A. A. Dragulescu and V. M. Yakovenko *Statistical Mechanics of Money, Income, and Wealth: A Short Survey* Modeling of Complex Systems: Seventh Granada Lectures, AIP Conference Proceedings 661, New York, 2003, pp. 180-183,

[10] B. Düring, D. Matthes, and G. Toscani, *Exponential and algebraic relaxation in kinetic models for wealth distribution*. In: "WASCOM 2007" - Proceedings of

the 14th Conference on Waves and Stability in Continuous Media, N. Manganaro et al. (eds.), pp. 228-238, World Sci. Publ., Hackensack, NJ, 2008.

[11] B. Düring, D. Matthes, and G. Toscani, *Kinetic equations modelling wealth redistribution: a comparison of approaches.* Phys. Rev. E 78(5) (2008), 056103.

[12] B. Düring and G. Toscani, *Hydrodynamics from kinetic models of conservative economies.* Physica A 384(2) (2007), 493-506.

[13] B. Düring and G. Toscani. *International and domestic trading and wealth distribution.* Comm. Math. Sci., 2009.

[14] L. Grilli and F. Ciardiello *Un approccio econofisico alla distribuzione della ricchezza* Econofisica, Metodi per l'Economia, Collana Interdipartimentale di Studi Economici, vol.11, pp. 49-62, (2006), Edizioni Scientifiche Italiane, Napoli.

[15] S. Ispolatov, P. L. Krapivsky, and S. Redner Wealth Distributions in Models of Capital Exchange, Preprint (1997)

[16] S. Ispolatov, P. L. Krapivsky, and S. Redner Wealth distributions in asset exchange models, The European Phisical Journal, 2, pp. 267-276 (1998)

[17] T. Lux, *Financial Power Laws: Empirical Evidence, Models and Mechanism*, Christian-Albrechts-Universität Kiel - Economics Working Paper No 2006-12.

[18] D. Matthes and G. Toscani, *On steady distributions of kinetic models of conservative economies.* J. Stat. Phys. 130(6) (2008), 1087-1117.

[19] M. Mitzenmacher, *A brief history of generative models for power law and lognormal distributions* - Internet Mathematica, v. 1, n. 2 (2004), pp. 226-251.

[20] L. Pareschi, *Microscopic dynamics and mesoscopic modelling in economy*, Working paper

[21] L. Pareschi, G. Toscani, *Self-similarity and power-like tails in nonconservative kinetic models*, Journal of Statistical Physics, 124, (2006), pp. 747-779

[22] V. Pareto, *Cours d'economie politique professé a l'université de Lausanne*, 3 volumes, (1896-1897).

[23] V. Pareto, *Manuale di economia politica con una introduzione alla scienza sociale*, (1906).

[24] M. Patriarca, E. Heinsalu, A. Chakraborti, *The ABCD's of statistical many-agent economy models*, (2008), preprint

[25] E. Samanidou, E. Zschischang, D. Stauffer and T. Lux, *Microscopic models of financial markets* (2006) - Working Paper

[26] N. Scapetta, S. Picozzi and B. J. West, *An out of equilibrium model of the distribution of wealth* Quant. Finance 4 (2004), n.3, 353-364

Chapter 4

Modeling financial markets

4.1 A financial market model

4.1.1 Trading at the microscopic level

Here we analyze a different market model, namely a variation on the Levy-Levy-Solomon first presented in 1994 (see [2], [4] and [5]). This is not based on binary interactions anymore, but the whole set of agents contributes to a process of price formation which in turn influences wealth of each individual.

At each time step the economic agent will have a choice on how to allocate his wealth among two specific different assets: one riskless account and a risky asset, i.e. bonds and company stocks; no liquid cash deposits are allowed. The riskless accout grows at a constant rate r while the risky asset has price S.
Thus one can write $w_i = (1-\gamma_i)w_i + \gamma_i w_i = (1-\gamma_i)w_i + n_i S$ where n_i is the number of shares bought with the risky part of allocation. For simplicity - and unlike in actual markets - the quantity of stocks bought can be a real number, so there is no restriction to integers. However there are restrictions on borrowing and short sales, so that all quantities start non negative and must not become negative after any event; this implies that the allocation tuning parameter is limited to the interval $\gamma_i \in [0,1]$. Bonds are exhogenous and unlimited quantities thereof can be bought, while stock shares are provided to the market in a certain volume n, kept fixed over time for simplicity.

After a time step the updated stock price S' will determine the new wealth of all agents:

$$w_i' = (1-\gamma_i)w_i(1+r) + \gamma_i w_i(1+x)$$

with $x = (S' - S)/S$, where S' is however initially unknown. The only other factor influencing each individual wealth is the constant interest rate, but clearly this is known in advance so all uncertainty is limited to the stock price (because in this case we have not introduced any exhogenous source of stochastic noise yet). To highlight the different contributions of interest rates and stock appreciation one could rewrite the new wealth as follows:

$$w_i' = w_i + (1 - \gamma_i)w_i r + \gamma_i w_i \frac{S' - S}{S} = w_i + (w_i - n_i S)r + n_i(S' - S)$$

The behaviour of each agent will also be affected by an utility function $U(w)$, which can also be used to describe his propensity to risks of different sizes as a function of his starting wealth. Such utility function can be chosen among many different possibilities provided that some basic criteria are satisfied: it is a non decreasing function of w starting at the origin $U(0) = 0$, possibly also convex (that is $U'(w) > 0$ and $U''(w) \leq 0$); a positive first derivative implies that a larger wealth will always be preferred to a smaller one, but less strongly even when when just the *relative* difference gets smaller (because of the decreasing monotony of $U'(w)$). Two common choices are a logarithmic utility $U(w) = \log(w)$ and the von Neumann-Morgenstern utility function:

$$U(w) := w^{1-\alpha}/(1 - \alpha)$$

For any hypotetical price S^h each agent will follow a certain allocation strategy $\gamma_i'(S^h)$ trying to maximize his expected utility $\mathbb{E}[U(w)]$. The strategy will be a monotone and non increasing function such that $\gamma_i'(0) = 1$ and $lim_{s\to\infty}\gamma_i'(s) = 0$. In the original LLS model such strategy was changing with time according to the evolving price history, however in the CPP setting it can be considered given "a priori" and time-homogeneous [1].

The dynamics for the stock will be affected by the strategies of all players in the market: at each iteration the new price is determined by equilibrium of offer and demand. This mechanism is shown in the following, through the total number of stocks traded on the market:

$$n = \sum_i n_i = \sum_i \gamma_i w_i / S$$

[1] In the LLS model the strategy works in the following way: an integer k is fixed, then the future stock return probabilities are taken as k^{-1} for each of the last observed returns, 0 otherwise. In turn this induces an equal probability distribution on the corresponding k possible wealths, which together with the chosen utility function defines an expected utility. An optimization is then run to find the optimal allocation, something that in the case of logarithmic utility can easily be done analytically.

The number of shares owned by each individual is a function of the hypothetical price through both γ_i and w_i; furthermore it is monotone decreasing because of the same monotonicity of both these components and of $(S^h)^{-1}$.

Recalling that the total shares supply on the market is kept constant, at the next time step it must hold the following:

$$n = (S^h)^{-1} \sum_i \gamma_i'(S^h) w_i'(S^h)$$

Thanks to monotonicity of each term n_i' this fixed point equation has one and only one solution $S^h := S'$ (see figure 4.1). This equilibrium price matches offer and demand, so that trading can occur and the whole process advance to the next time step: $w_i'(S')$ will be determined by such value and from the new wealth again a new allocation can follow.

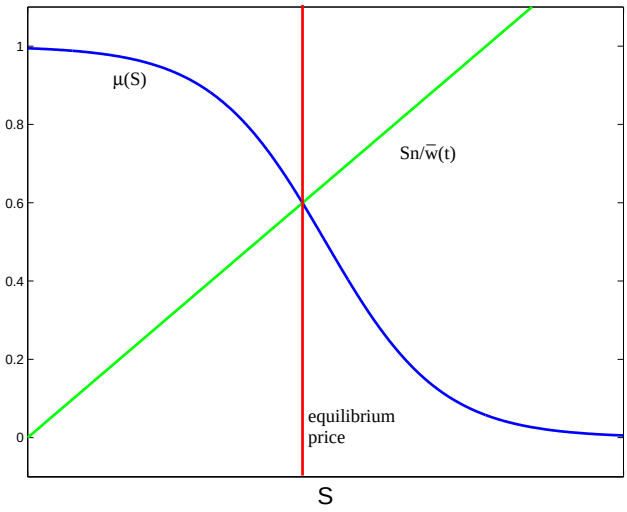


Figure 4.1: *Equilibrium of offer and demand*

All this was in the case of a simple stock return of the form $x := (S' - S)/S$, but the same procedure can be followed for a more realistic return including further components, namely dividends and a stochastic component are often considered, leading to:

$$x(S', \eta) := \frac{S' - S + D + \eta}{S}$$

Here the dividend $D \geq 0$ is a constant amount paid at each iteration, while the random variable $\eta \sim \Theta(\eta)$ is distributed with zero mean and σ^2 variance.

For a more realistic description it's also possible to introduce further stochastic

terms, i.e. accounting for practical deviations of agents from their optimal strategy, or for describing stochastic interest rate dynamics. This will be done in the next section for the mesoscopic approach.

4.1.2 Kinetic formulation

As in the previous section we can describe the time distribution of agents with different wealth $w, t \in \mathbb{R}_+$ through a kinetic density $f(w,t)$.
The invested wealth quota γ will now be of the form $\gamma(t, S, \xi) = \mu(S) + \xi$ with $\mu(S)$ being the optimal choice and ξ a random variable with compact support such that no negative allocation quota can ever occur. Therefore $\gamma \sim \Phi(\mu(S), \xi)$.

Price formation

In analogy with the microscopic case, where $S = n^{-1} \sum_i \gamma_i w_i$, an integral relation holds for the price:

$$S(t) = n^{-1}\mathbb{E}[\gamma w] = n^{-1} \int f(w,t) w \; \Phi(\gamma)\gamma \; d\gamma \; dw$$

However γ and w are independent, so this becomes

$$\overline{w}(t) := \mathbb{E}[w] = \int_0^\infty f(w,t) w \; dw$$

$$\mathbb{E}[\gamma] = \int \Phi(\gamma)\gamma \; d\gamma = \int \Phi(\mu(S), \xi)\xi \; d\xi \equiv \mu(S)$$

$$S(t) = n^{-1}\mathbb{E}[\gamma]\mathbb{E}[w] = n^{-1}\mu(S)\overline{w}$$

For the equilibrium price one can write:

$$S'(t) = n^{-1}\mu(S')\mathbb{E}[w']$$

Since $\mathbb{E}[w'(S', \gamma, \eta)] = \mathbb{E}[w](1 + r) + \mathbb{E}[\gamma w](\mathbb{E}[x(S', \eta)] - r)$ we get:

$$\mathbb{E}[w'(S', \gamma, \eta)] = \overline{w}(1 + r) + \mu(S)\overline{w} \left(\frac{S' - S + D}{S} - r \right)$$

and:

$$S' = n^{-1}\mu(S')\overline{w}(t) \left[(1 + r) + \mu(S) \left(\frac{S' - S + D}{S} - r \right) \right]$$

This can be further manipulated to eliminate dependence on $\overline{w}$ obtaining the following implicit expression for S':

$$\begin{aligned} S'(t) &= \frac{\mu(S')}{\mu(S)}[(1-\mu(S))S(1+r)+\mu(S)(S'+D)] \\ &= \frac{(1-\mu(S))\mu(S')}{(1-\mu(S'))\mu(S)}S(1+r)+\frac{\mu(S')}{1-\mu(S')}D \end{aligned} \tag{4.1}$$

Uniqueness of the solution S' for such an implicit relation can be verified as follows, by writing $g(S) := (1-\mu(S))S/\mu(S)$ the future price satisfies $g(S') = g(S)(1+r)+D$, which gives:

$$\frac{dg(S)}{dS} = -\frac{d\mu(S)}{dS}\frac{S}{\mu(S)^2}+\frac{1-\mu(S)}{\mu(S)} > 0,$$

so that $g(S)$ is monotonically increasing, and thus the solution is unique: $S' = g^{-1}[g(S)(1+r)+D] > S$. Furthermore whenever $r = D = 0$ then the solution is $S' \equiv S$ so that the price does not change over time, independently of the shape of $\mu(\cdot)$.

One can also study the behaviour of stock returns with respect to the bond; through the average stock return:

$$\overline{x}(S') = \mathbb{E}[x(S',\eta)] = \left(\frac{S'-S+D}{S}-r\right)$$

one gets:

$$\overline{x}(S')-r = \frac{(\mu(S')-\mu(S))(1+r)}{(1-\mu(S'))\mu(S)}+\frac{\mu(S')D}{S(1-\mu(S'))}, \tag{4.2}$$

Noting that $\mu(S') \leq \mu(S)$, the sign of the right hand side of such equation - and thus the relative performance of the stock - will depend on the rate of variation of investments being above a constant threshold or not:

$$\frac{\mu(S')-\mu(S)}{\mu(S)\mu(S')}S \geq -\frac{D}{(1+r)}.$$

Kinetic analysis

The linear kinetic equation corresponding to the previously described "microscopic" evolution is:

$$\partial_t f = \int_{-d}^{d} \int_{-z}^{z} \left[\beta_{'w \to w} J^{-1} f('w, t) - \beta_{w \to w'} f(w, t)\right] d\xi d\eta \tag{4.3}$$

where the Jacobian $J(\xi, \eta, t) = 1 + r + \gamma(\xi)(X(S', \eta) - r)$ is needed for conservation of the number of agents. The first term is a gain coming from pre trading wealth $'w = w/J(\xi, \eta, t)$, while the second is a loss term. In this case the interaction kernel is of the form:

$$\beta_{'w \to w} = \Phi(\mu(S), \xi)\Theta(\eta)$$

As usual in kinetic analysis we can take the weak form of the evolution equation (4.3) to gain information on the dynamics of the different moments:

$$\frac{d}{dt} \int_0^\infty f(w, t)\phi(w) dw = \tag{4.4}$$

$$= \int_0^\infty \int_{-D}^{D} \int_{-z}^{z} \Phi(\mu(S), \xi)\Theta(\eta) f(w, t)(\phi(w') - \phi(w)) d\xi \, d\eta \, dw. \tag{4.5}$$

with $\phi(w) = 1$ gives the conservation of the total number of agents, while for $\phi(w) = w$ it returns the average wealth dynamics, which in turn influences price behaviour. Thus one obtains the following bounds:

$$\overline{w}(t) \leq \overline{w}(0) \exp(Mt) \tag{4.6}$$

$$S(t) \leq S(0) \exp(Mt). \tag{4.7}$$

Again one sees that the average wealth increases exponentially with time, and the price is bounded from above by an exponential. For the higher moments the following bound holds:

$$\frac{d}{dt} \int_0^\infty w^p f(w, t) \, dw \leq A_p(S) \int_0^\infty w^p f(w, t) \, dw, \tag{4.8}$$

with

$$A_p(S) = 2C_p(r^p + r^{p-2}\left(1 + \frac{c_2}{S^2}((S'-S)^2 + D^2 + \sigma^2 E(|Y|^2))\right)$$

$$+r^2\left(1 + \frac{c_{p-2}}{S^{p-2}}((S'-S)^{p-2} + D^{p-2} + \sigma^{p-2}E(|Y|^{p-2}))\right)$$

where C_p, c_{p-2} and c_2 are suitable constants.

Fokker-Planck asymptotics

Analogously to the CPT model one can derive a Fokker-Planck limit as stated in the following theorem:

Theorem 4.1.1 *Let the probability density $f_0 \in \mathcal{M}_p$, where $p = 2 + \delta$ for some $\delta > 0$. Then, as $r \to 0$, $\sigma \to 0$, and $D \to 0$ in such a way that $\sigma^2 = \nu r$ and $D = \lambda r$, the weak solution to the Boltzmann equation (4.5) for the scaled density $\tilde{f}_r(w,\tau) = f(v,t)$ with $\tau = rt$ converges, up to extraction of a subsequence, to a probability density $\tilde{f}(w,\tau)$. This density is a weak solution of the following Fokker-Planck equation (4.9).*

$$\frac{\partial}{\partial\tau}\tilde{f} = \frac{\partial}{\partial w}\left[-A(\tau)w\tilde{f} + \frac{1}{2}B(\tau)\frac{\partial}{\partial w}w^2\tilde{f}\right], \tag{4.9}$$

where

$$A(\tau) = 1 + \tilde{\mu}(\tilde{S})\left((\kappa(\tilde{S}) - 1) + \frac{\tilde{\mu}(\tilde{S})(\kappa(\tilde{S}) - 1) + 1}{1 - \tilde{\mu}(\tilde{S})}\frac{\lambda}{\tilde{S}}\right) \tag{4.10}$$

$$B(\tau) = \frac{(\tilde{\mu}(\tilde{S})^2 + \zeta^2)}{\tilde{S}^2}\nu. \tag{4.11}$$

Also in the case of the Fokker-Planck limit the mean wealth increases over time, according to the rate $\dot{\overline{w}}(\tau) = A(\tau)\overline{w}(\tau)$, which gives the following bound:

$$(1 - \tilde{\mu}(\tilde{S}))\overline{w}(\tau) + n\lambda \leq \dot{\overline{w}}(\tau) \leq \overline{w}(\tau) + \frac{n\lambda}{1 - \tilde{\mu}(\tilde{S})}. \tag{4.12}$$

A self similar solution can be found through an adequate scaling; the transformation $\chi = \log(w)$ implies

$$\tilde{f}(w,\tau) = \frac{1}{w}\tilde{g}(\chi,\tau).$$

which can be plugged back into (4.9). It turns then out that the evolution of $\tilde{g}(\chi,\tau)$ itself is governed by a corresponding linear equation of convection-diffusion type:

$$\frac{\partial}{\partial\tau}\tilde{g}(\chi,\tau)=\left(\frac{B(\tau)}{2}-A(\tau)\right)\frac{\partial}{\partial\chi}\tilde{g}(\chi,\tau)+\frac{B(\tau)}{2}\frac{\partial^2}{\partial\chi^2}\tilde{g}(\chi,\tau),$$

which is satisfied by the following solution:

$$\tilde{g}(\chi,\tau)=\frac{1}{(2b(\tau)\pi)^{1/2}}\exp\left(-\frac{(\chi+b(\tau)/2-a(\tau))^2}{2b(\tau)}\right), \tag{4.13}$$

having set

$$a(\tau)=\int_0^\tau A(s)\,ds+C_1,$$
$$b(\tau)=\int_0^\tau B(s)\,ds+C_2.$$

Reexpressing this back in the previous variables w and τ one finally obtains the following lognormal solution in the asymptotic limit:

$$\tilde{f}(w,\tau)=\frac{1}{w(2b(\tau)\pi)^{1/2}}\exp\left(-\frac{(\log(w)+b(\tau)/2-\log(\bar{w}(\tau)))^2}{2b(\tau)}\right). \tag{4.14}$$

with

$$b(\tau)=\log\left(\frac{\bar{e}(\tau)}{(\bar{w}(\tau))^2}\right).$$

and $\bar{e}(\tau)$ being the second order moment, governed by the following evolution:

$$\dot{\bar{e}}(\tau)=(2A(\tau)+B(\tau))\bar{e}(\tau).$$

4.1.3 Numerical examples

In the following numerical simulation results are shown for three different settings. In all cases 1000 agents are participating in a market with 10000 shares. The starting wealth for each investor equals 1000, initially split equally between shares (valued at 50 each) and bonds. The two random variables ξ and η are set being distributed with a truncated normal distribution, to avoid negative wealth values. The first two cases show a comparison between the Monte Carlo simulation of the kinetic model and the direct solution for price. The third example compares the time averaged Monte Carlo asymptotic behaviour to the explicit description given by the Fokker-Planck

evolution.

Test 1

In this test case the specific parameters are set as follows: the interest rate is deterministic and constant $r = 0.01$, the dividend yield is 0.015; furthermore all agents share a common constant investment rule $\mu(\cdot) = C \in (0, 1)$, which is determined by the other parameters so that $C = 0.5$, and the evolutions of both the mean wealth and of the stock price are explicitely known. Results are shown for 400 iterations of the price formation process, with the distributions of the random variates ξ and η having variances of 0.2 and 0.3 respectively.

Figure 4.2 shows the dynamics of the simulated price $S(t)$ (in blue) and the direct solution (in red). Clearly the Monte Carlo behaviour matches the exponential dynamics of the analytical solution. Later the corresponding allocation $\mu(t)$ between assets is also shown over time, oscillating around the optimal reference value in red.

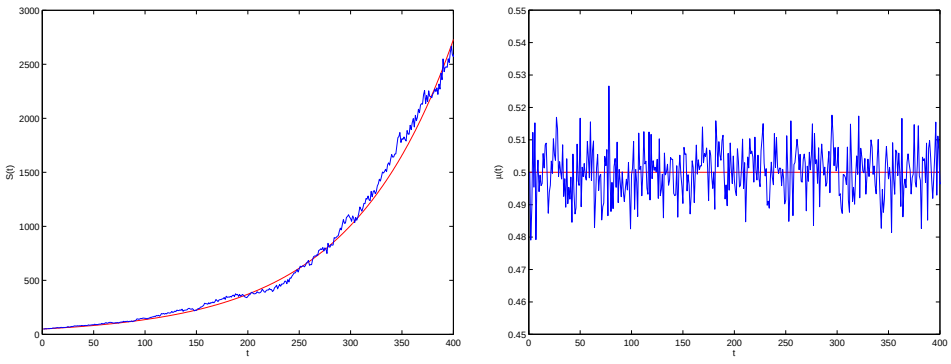


Figure 4.2: *Test 1 - The numerical kinetic simulation shows an exponential growth over time for price (left). Fluctuations for the fraction of investment are depicted on the right.*

Test 2

In this test case the investment quota is varying deterministically according to an exponential decay as stock price grows:

$$\mu(S) = 0.2 + 0.8e^{-C_2 S}$$

where $C_2 = \log(0.8/0.3)/S_0 \approx 0.02$; with these values $S_0 = 50$ satisfies the price equation, and $\mu(S)$ remains in the interval $[0.2, 0.5]$.

Figure 4.3 shows again the exponential price evolution, and the investment behaviour. However in this case the price growth is much slower than in the previous constant investment case.
Next in figure 4.4 the mean wealth growth of both test cases is compared to the risk-

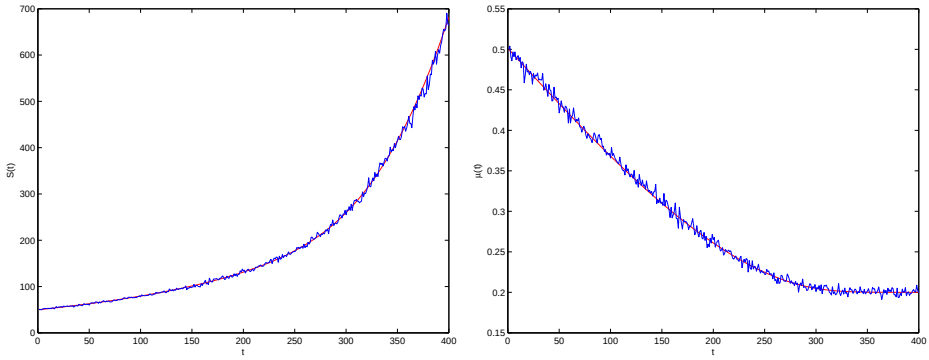


Figure 4.3: *Test 2 - Again the price dynamics and investment fractions, but this time for a non constant investment quota* $\mu(S)$.

less bond investment growing exponentially at the interest rate r, drawn in dashed red. The upper curve corresponding to Test 1 dominates the investment in bonds, while in the other case performance is lower because of the time decay in risky investments.

In the following figure the two averaged final wealth distributions (at $t = 400$) are shown in double logarithmic scale, with a lognormal fit in red for reference. This highlights how even for the Boltzmann model the tails behave lognormally. The same also holds for the number of stocks owned by agents, due to the identity $\gamma_i w_i = n_i S_i$.

Test 3

The last numerical test focuses on the asymptotic limit for the Boltzmann model, comparing it to the explicit solution of the Fokker-Planck model. In this case the kinetic particle simulation is run with parameters $r = 0.001$, $D = 0.0015$, $\mu = 0.5$ and where ξ and $\eta/S(0)$ are distributed with standard deviation 0.05. Figure 4.5 show the resulting wealth distributions at the different time points $t = 50, 200$ and 500; here a good fit of the Fokker-Planck solutions to the Boltzmann results can be seen.

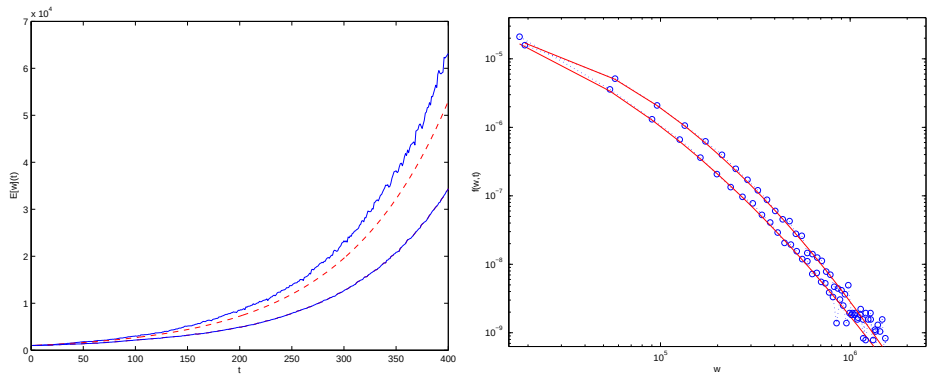


Figure 4.4: *Tests 1 and 2 - The exponential growths of mean wealth in the two cases, versus a bond investment: upper line with constant investment and lower with the decreasing invested quota (left). Also a lognormal fit to the empirical distributions is shown in log-log scale (right).*

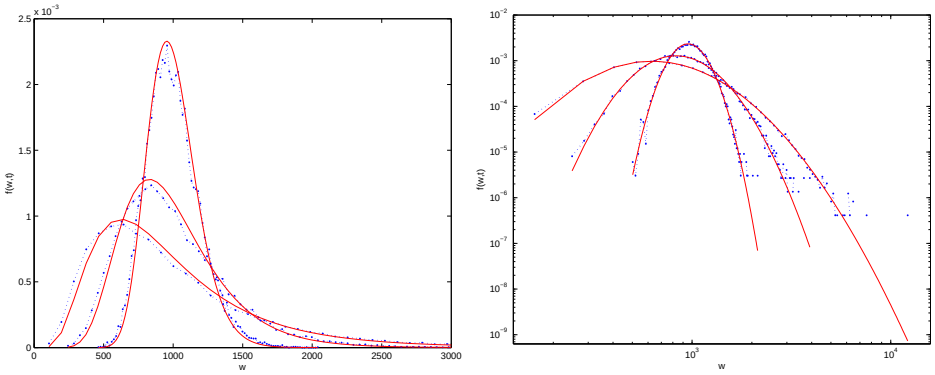


Figure 4.5: *Test 3 - Wealth distribution in the Boltzmann (dotted blue) and Fokker-Planck (continuous red) models*

4.2 A multiclass financial market model

In the following we improve upon the CPP model by allowing for inhomogeneous investment propensity. To more realistically model a market we can introduce a set of kinetic densities $f_k(w,t) \quad k = 1, ..., K$ corresponding to economic agents with different characteristics. For example one could have *fundamentalists* and *chartists* agents interacting on the market, where chartists follow purely quantitative analysis of historical behaviours of assets, while fundamentalists also take into account further information. The following analysis will not include opinion formation, in that after a binary interaction the two traders will remain in the same class, so that the classes

are somewhat separated and evolve in parallel.

Analogously to the homogeneous case we can define the average wealth for each agent class, by implicitly fixing a certan time t:

$$\overline{w}_k = \frac{1}{\rho_k} \int_{\mathbb{R}_+} f_k(w) w \, dw$$

where

$$\rho_k = \int_{\mathbb{R}_+} f_k(w) dw$$

One could further assume $\sum \rho_k = 1$. Moreover we can similarly define a different investor preference function for each agent class: $\mu_k(\cdot)$

The whole procedure of price formation follows closely the one exposed already in the case of homogeneous investing behaviour.

Again we can write an equality on the total wealth invested on the stock market:

$$\begin{aligned} nS &= \sum \mathbb{E}[\gamma_k w_k] \\ &= \sum_{k=1}^{K} \int_{\mathbb{R}_+} \int_{\mathbb{R}} \Phi(\xi) f_k(w) (\mu_k(S) + \xi) w \, d\xi \, dw \\ &= \sum_{k=1}^{K} \mu_k(S) \rho_k \overline{w}_k \end{aligned} \tag{4.15}$$

Here each $\mu_k(S)$ is monotone non increasing, so that the same holds for $\sum_{k=1}^{K} \mu_k(S) \rho_k \overline{w}_k$ and equation (4.15) is satisfied by an unique solution S. At the next time step the individual wealth evolves as:

$$w'_k = (1 - \gamma_k) w_k (1 + r) + \gamma_k w_k (1 + x)$$

coming from the usual trivial identity showing the investment decomposition:

$$w_k = (1 - \gamma_k) w_k + \gamma_k w_k$$

and with $x = (S' - S + D + \eta)/S$. Again x is subject to restrictions guaranteeing $w' \geq 0$ so that for simplicity no debts are allowed; for $\eta \sim \Theta(\eta)$ it follows $\eta \in [-d, d], \quad 0 < d \leq S' + D$. So in the end the future price will satisfy an equation similar to (4.15):

$$\begin{aligned} S' &= \frac{1}{n}\sum_{k=1}^{K}\mathbb{E}[\gamma_k' w_k'] = \frac{1}{n}\sum_{k=1}^{K}\mathbb{E}[\gamma_k']\mathbb{E}[w_k'] \\ &= \frac{1}{n}\sum_{k=1}^{K}\mu_k(S')\int_{\mathbb{R}}\int_{\mathbb{R}}\int_{\mathbb{R}_+}\Theta(\eta)\Phi(\xi)f_k(w) \\ &\qquad\qquad [(1-\gamma_k)w_k(1+r)+\gamma_k w_k(1+x)]\,dw\,d\xi\,d\eta \\ &= \frac{1}{nS}\sum_{k=1}^{K}\mu_k(S')\rho_k\overline{w}_k[(1-\mu_k(S))(1+r)S+\mu_k(S)(S'+D)] \end{aligned} \tag{4.16}$$

We can group the S' terms, thereby obtaining:

$$\begin{aligned} &S'\left(1-\frac{1}{nS}\sum_{k=1}^{K}\mu_k(S')\rho_k\overline{w}_k\mu_k(S)\right) = \\ &\quad = \frac{1}{nS}\sum_{k=1}^{K}\mu_k(S')\rho_k\overline{w}_k[(1-\mu_k(S))(1+r)S+\mu_k(S)D] \end{aligned} \tag{4.17}$$

From here we finally reach an implicit relation for S' through (4.15).

$$\begin{aligned} S' &= \frac{\sum_{k=1}^{K}\mu_k(S')\rho_k\overline{w}_k[(1-\mu_k(S))(1+r)S+\mu_k(S)D]}{nS-\sum_{k=1}^{K}\mu_k(S')\rho_k\overline{w}_k\mu_k(S)} \\ &= \frac{\sum_{k=1}^{K}\mu_k(S')\rho_k\overline{w}_k[(1-\mu_k(S))(1+r)S+\mu_k(S)D]}{\sum_{k=1}^{K}\rho_k\overline{w}_k(1-\mu_k(S'))\mu_k(S)} \end{aligned} \tag{4.18}$$

Again in the Monte Carlo simulation we have used numerical root finding methods to solve for S' the discrete version of this identity.

In the special homogeneous case with $K = 1$ so that $\rho_1 = 1$, equation (4.18) becomes:

$$S' = \frac{\mu_1(S')\overline{w}_1[(1-\mu_1(S))(1+r)S+\mu_1(S)D]}{\overline{w}_1(1-\mu_1(S'))\mu_1(S)} \tag{4.19}$$

This is actually just equation (4.1) as one would expect.

On the other hand, in the scaling limit $r \to 0$ and $D \to 0$ we get

$$S' \sum_{k=1}^{K} \rho_k \overline{w}_k (1 - \mu_k(S')) = S \sum_{k=1}^{K} \mu_k(S') \rho_k \overline{w}_k (1 - \mu_k(S)) \tag{4.20}$$

In the multiclass case the Boltzmann equation (4.3) does not change significantly: the integration extreme d for η does not depend on the class k. On the other hand each investor will be subject to different errors, defined by z_k which are functions of μ_k. This could be unified by using $z = min\{z_k\}$ so that all investors would behave in the same way.

$$\frac{\partial f_k}{\partial t} = \int_{-d}^{d} \int_{-z_k}^{z_k} \left[\beta_{'w \to w} J_k^{-1} f_k('w, t) - \beta_{w \to w'} f_k(w, t) \right] d\xi d\eta \tag{4.21}$$

Here the pre-trading wealths are given through the class-specific Jacobian: $'w = w/J_k(\xi, \eta, t)$ with $J_k(\xi, \eta, t) = 1 + \gamma_k(\xi)(x(S', \eta) - r)$.

4.2.1 Numerical simulation

For the Monte Carlo simulation of the multiclass model we considered the special case of two agent classes with constant investment functions $\mu_1(S) = 0.1$ and $\mu_2(S) = 0.9$. The total wealth was unchanged from previous examples (1000000) while the number of agents in this case is 50000, so that each one initially starts with a wealth of 20. The standard deviations for random variables were set to 0.03 for ξ and 0.22 for $\eta/S(0)$ respectively. We ran a simulation obtaining a well definied bimodal distribution for wealth already after 50 time steps, where each peak corresponds to one of the two investment propensities; this money distribution is depicted in Figure (4.2.1).

This behaviour is to be expected because even though the different $\mu_k(\cdot)$ concur in determining the common share price, each individual dynamic is mostly governed by its own investment propension, which dominates over the influence of share price (in fact this was noted in the homogeneous case already). As a consequence different classes of agents remain somewhat separated.

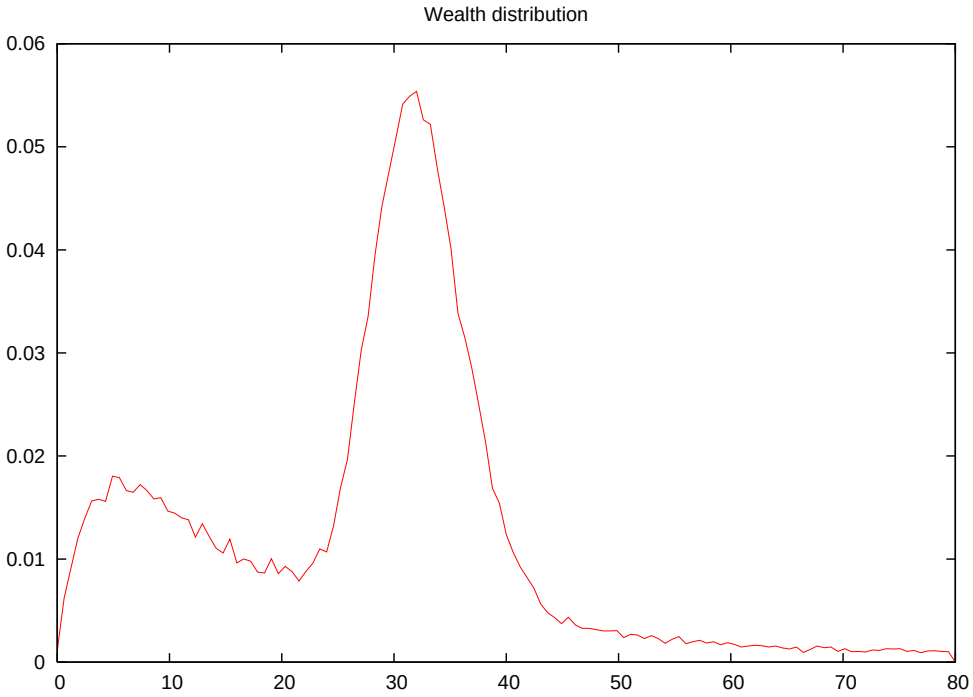


Figure 4.6: Wealth distribution for a multiclass model with two constant investment propensities

Bibliography

[1] J. P. Bouchaud, M. Mézard, *Wealth condensation in a simple model of economy*, Physica A, 282, (2000), pp. 536-545

[2] M. Levy, H. Levy and S. Solomon, *A Microscopic Model of the Stock Market: Cycles, Booms, and Crashes*, Economics Letters, 45, (1994), 103-111

[3] M. Levy, H. Levy and S. Solomon, *Microscopic Simulation of the Stock Market: The Effects of Microscopic Diversity*, Journal de Physique I, 5, (1995), pp. 1087-1107.

[4] H. Levy, S. Solomon and M. Levy, *Microscopic Simulation of Financial Markets: From Investor Behavior to Market Phenomena*, Academic Press, Inc., Orlando, FL, USA (2000).

[5] S. Solomon and M. Levy, *Market Ecology, Pareto Wealth Distribution and Leptokurtic Returns in Microscopic Simulation of the LLS Stock Market Model* in New Directions in Statistical Physics, ed. Luc Wille, Springer-Verlag (2004), p69

Chapter 5

Fast skip ahead for linear recursive pseudorandom generators

Recently linear random number generators based on recurrences modulo 2 have gained wide acceptance in the simulation community thanks to their good statistical properties and computational speed. However sometimes, especially in parallel applications, a skip ahead algorithm is needed to move freely the generator state on the random stream; the usual straightforward matrix multiplication becomes impractical for large state and jump sizes like in the common Mersenne twisters. This work introduces a faster method, practical for the most used generators, exploiting the recursive structure of the iteration matrix.

5.1 Setting

In scientific computing the motivations for jumping ahead within a long sequence of pseudorandom numbers can be quite diverse, but the most common is the need of guaranteeing lack of overlap between subchunks by large separation of their starting points, with the aim of independent initialization.

This is also related to the increasing success of parallel architectures, which is pushing strongly for a wider use of concurrent simulation threads, each needing one or more separate pseudorandom streams.

Other uses for skip ahead include improving the quality of a generator by decimation (see [2]) or just plain investigation of the sequence itself and of its statistical

properties.

There are at least three approaches to providing multiple pseudorandom streams for parallel simulations; one is seen e.g. in the dynamic creation of Mersenne Twisters [6], where different generator parameters are dedicated to each stream (so that the same generator state will often be mapped to different values). Another option is to build PRNGs with a single transition function that directly partitions the state space into multiple disjoint closed orbits: for stream initialization it will suffice being able to pick one (and only one) arbitrary element out of each orbit. Both such methods involve generators that are naturally parallel, but the initialization can be less than trivial and sometimes requires ad-hoc analysis.

The third possibility involves the above mentioned splitting of a single major stream (often of $2^d - 1$ length where d is the state bit size), so in theory parallelization cannot ever extract nonoverlapping substreams, but for large state sizes the starting seeds can nevertheless be chosen "far enough" for practical purposes (as in a real-world setting imposing computational constraints). Such long jumps are the subject of this work.

The focus will be on the class of linear PRNGs over $\mathbb{F}_2^d$ because of their many advantages; among them is the availability of an extensive literature spanning in particular the last 15 years, which provides a well developed theory as opposed to many other good random generatos. They enjoy a fairly easy statistical analysis together with good equidistribution and independence, very large periods, often a simple implementation and high generation speed, because of the light hardware requirements (in fact they only use bit shifts and bitwise logical operation).

Parallelization is also simple, while the only missing ingredient was the availability of fast methods, allowing the practical creation of multistream packages, something that motivated this research.

All these points have made $\mathbb{F}_2^d$-linear PRNGs one of the most popular classes of generators used in the simulation community.

Let then S be the set of PRNG states of dim d over $\mathbb{F}_2$, that is $S \cong \mathbb{F}_2^d$.
Let's consider linear generators whose transition function $f : S \to S$ can be encoded in a $d \times d$ matrix A, driving the succession:

$$\begin{aligned} \boldsymbol{s}_{i+1} &= A\boldsymbol{s}_i \\ \boldsymbol{o}_i &= O\boldsymbol{s}_i \end{aligned}$$

To actually draw random numbers the output transformation (with associated $w \times d$ matrix O) extracts the desired part of the state and might eventually also provide tempering.
This structure encompasses many well studied and used PRNGs: the Linear Feedback Shift Register (LFSR), polynomial LCG, the Generalized FSR (GFSR), the Twisted GFSR (TGFSR), xorshift, the Mersenne Twister (MT), RANROT, WELL, pulmonary MTs, xorgens (see i.e. [6], [8], [9])...

5.1.1 Skipping ahead

The task of skipping ahead to obtain $\boldsymbol{s}_{i+J}$ from $\boldsymbol{s}_i$ is clearly linked to the powers of A, since by induction $\boldsymbol{s}_{i+J} = f^J(\boldsymbol{s}_i) = A^J \boldsymbol{s}_i$. Without loss of generality i will be 0 in the following.

The straightforward approach of iteratively calculating all the explicit matrix powers $\{A^j\}_{1<j\leq J}$ costs $O(d^3 J)$, which can be lowered to $O(d^3 log(J))$ with basic square-and-multiply exponentiation rearrangements ([5]) but still remains impractical for large state sizes. The issue can be tackled by improving either the complexity of the matrix multiplication (at the root of the cubic term) or that of the exponentiation procedure based on it (the logarithmic term); however at the time of writing the former still cannot reach a quadratic growth[1] while the latter techniques can still be applied to the present work.

Precomputations however can spare matrix multiplications for the actual jump, leading to just matrix-vector multiplications of $O(d^2 log(J))$ complexity for arbitrary jumps at the cost of a storage on the same order, or even down to $O(d^2)$ when the jump size is known in advance.
With $d = 19937$ it still means for the popular Mersenne Twister about 47.4 MB, which is not an easy task for the cache memories of most commodity processors to handle. For state sizes of 1024 digits[2] it starts being reasonable (128 KB) but remains still too slow for applications requiring frequent jumps[3]; a major example are

[1] The complexity of matrix multiplication is $O(d^3)$ only in the naive direct metod following the definition. More efficient algorithms have been devised since the seminal work of Strassen in 1969, who lowered the exponent to $log_2 7 \approx 2.807$. Currently the best value is (to the best of my knowledge) ≈ 2.376. However all these ingenious methods come at some cost in either numerical stability, ease of implementation or in the magnitude of the implicit constant, making the actual implementation less viable.

[2] representing a good practical compromise between period size and fast bit mixing

[3] Such size often exceeds already the L1 cache

e.g. certain particle physics simulations requiring continuous spawning of new pseudorandom sequences. For these reasons matrix multiplication is not the best choice since better algorithms have been developed, these will be reviewed and compared after having described the proposed method.

This work exploits the characteristics of a relatively small but important class of RNGs with sparse matrix. The Feedback Shift Register recurrences alter only one (or two) words at a time, while shifting out the remaining ones.
With $T \in M_{w\times d}(\mathbb{F}_2)$, where $d = w \cdot r$ the transition matrix has the following form:

$$A = \begin{bmatrix} \boxed{\begin{matrix} & & T & & \end{matrix}} \\ \begin{matrix} I_w & \cdot & & \cdot & \cdot \\ \cdot & I_w & & \cdot & \cdot \\ & & \ddots & & \\ \cdot & \cdot & & I_w & \cdot \end{matrix} \end{bmatrix}$$

Such generators can also be seen as just pushing at each iteration one new word onto a stack and correspondingly moving over it a window framing the state (as in **Fig**.5.1); this accounts for a much better performance as no data copies need to be carried out anymore[4].

Application of the transition f is fast by construction, however even for a sparse matrix A its higher powers A^j soon become dense so that the straightforward jump by matrix multiplication here bears no advantage compared to a more general transition. Still A^J obviously inherits from A some structure, regardless of the jumps size and even though such structure remains hidden and cannot be exploited directly. The main point of this work is devising a way to expose it, as to be able to take advantage in lowering the algorithm complexity.

5.2 Linear feedback shift register generators

5.2.1 Characteristic basis and a fast skip ahead algorithm

For the special case with $w = 1$ (and thus $r = d$) the PRNG is a standard LFSR, where we use the notation $\boldsymbol{s}_i := (s_{i,d-1}, s_{i,d-2}, ..., s_{i,0})^T$ for the state and its compo-

[4]To limit memory consumption this further prompts the obvious round-robin implementation where the window doesnt move over an infinite stack anymore but over a toroidal fixed height pseudostack having the same size as the state.

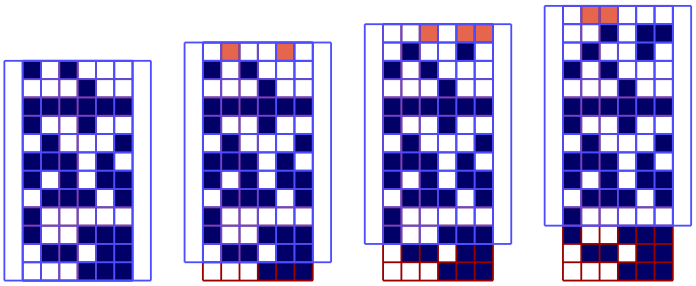


Figure 5.1: *Binary depiction of the transitions for a Feedback Shift Register, showing the sliding state window adding three new words*

nents.

In such setting the main skip-ahead idea is very simple and exploits both linearity of the generators and the feedback shifting structure: linearity implies that any stream starting at $\boldsymbol{s}_i$ can be seen as the componentwise sum of (at most) d streams, whenever the state generating the former is the sum of the states from the latter. On the other hand the FSR structure will allow one to choose a very convenient basis for the generator state, such that jumping on only one stream needs to be performed and that it can be done easily. It will be shown how all this results in a procedure requiring just a polynomial multiplication which actually is implemented as a binary convolution.

Let then $P(t) = \sum_{i=0}^{d} a_i t^{d-i} \in \mathbb{F}_2[t]$ be the characteristic polynomial of the recurrence, and $G(t) = \sum_{i=1}^{\infty} s_{i-1,0} t^{-i} \in \mathbb{F}_2[[t^{-1}]]$

its generating functions[5]: note here that for a LFSR the last bit of the state sequence $s_{i-1,0}$ subsequently receives the shifted values of the preceding ones ($s_{0,i}$ for $i < d$), so that the state already shows the first coefficients of its generating function; such property will be essential for the efficiency of the proposed algorithm.

The notation $\{H(t)\}$ will refer to the "fractional part" of the formal Laurent serie $H(t) \in \mathbb{F}_2[[t^{-1}]]$, and $\{H(t)\}_{,n}$ to the truncation of $\{H(t)\}$ at the n-th term.

Because for a LFSR the first d coefficients of a generating function do *fully* determine it - given $P(t)$ - there is an "identical" bijection between the generator states and the corresponding generating functions, so that it is possible to use each one in

[5] See Appendix A for a brief review of the required terminology

place of the other.

Advancing of the generator state by one iteration is then equivalent to multiplication of the generating function by t and subsequent extraction of the fractional part. In this setting the skip ahead task can be expressed as that of obtaining $\left\{t^J \cdot G(t)\right\}_{,d}$ from $\{G(t)\}_{,d}$.

In a jump it is easy to decouple the initialization from the structure of the recurrence itself and from the start position on the stream:

$$G^J(t) := \left\{t^J \cdot G(t)\right\} = \left\{t^J \cdot (G(t)P(t)) \cdot P(t)^{-1}\right\} =: \left\{g(t) \cdot P(t)^{-1} \cdot t^J\right\}.$$

The term $P(t)^{-1}$ will be central in the jumping procedure: it embodies the LFSR core transform applied to a "basic" element[6], that is a state of the form $(1, 0, ..., 0, 0)^T$. It will be used as a building block for the characteristic basis introduced below.

The coefficients g_j of the term $g(t) = G(t)P(t)$ explicit the components of the projection of $G(t)$ onto the set $\left\{\dfrac{t^j}{P(t)}\right\}_{j\in\mathbb{Z}}$:

$$G(t)P(t) = G(t)/P(t)^{-1}$$
$$\Downarrow$$
$$G(t) = g(t)/P(t) = \left(\sum_{j\in\mathbb{Z}} g_j t^j\right)/P(t) = \sum_{j\in\mathbb{Z}} g_j (t^j/P(t)).$$

In general $g(t) \in \mathbb{F}_2[[t^{-1}]]$ since by construction necessarily $g_j = 0$ for $j \geq d$, but it is known that the primitivity of $P(t)$ also implies $g_j = 0$ for $j < 0$, leading to $g(t) = \sum_{j=0}^{d-1} g_j t^j/P(t)$. Thus it suffices the finite set $\{\pi_j := t^j/P(t)\}_{0\leq j<d}$ to form a basis for the space of generating functions corresponding to $P(t)$, which will be called *characteristic* basis (see **Fig**.5.2 for the equivalent truncated basis $\{\pi_j\}_{,d}$).

Having isolated the three components leading to the target generating function $G^J(t)$ it is now possible to precalculate - for a fixed J - the part that does not depend on the initial generator state: the set $\left\{\overline{\pi}_j := t^{J+j}/P(t)\right\}_{0\leq\ j<d}$.

At this point the actual precalculation process could be easily guessed but it will only be detailed after the jumping procedure, since the latter will be used as the main building block, despite the apparent circularity of such dependencies.

[6]Thus $P(t)^{-1} = t^{-d}+a_1t^{-d-1}+(a_1+a_2)t^{-d-2}+(a_1+a_3)t^{-d-3}+(a_1+a_2+a_1a_2+a_4)t^{-d-4}$....

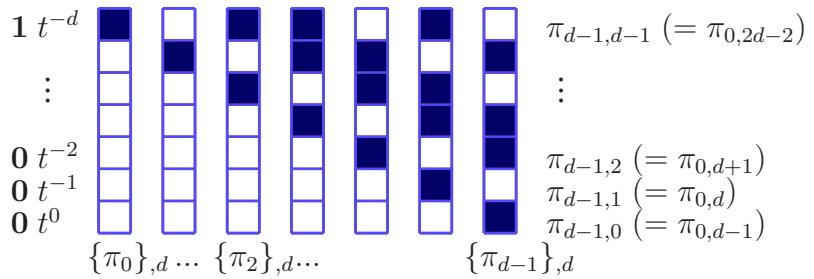


Figure 5.2: *The d coefficient vectors of the (truncated) characteristic basis $\{\pi_j\}_{,d}$ for the space of LFSR generator states*

Summing everything up, thanks to linearity of the generator transform the jump procedure can be expressed compactly as a simple isotropic transformation:

$$\left\{G^J(t)\right\} = \left\{t^J \cdot G(t)\right\} = \left\{t^J g(t)/P(t)\right\} = \left\{t^J \sum_{j=0}^{d-1} g_j \pi_j\right\} = \left\{\sum_{j=0}^{d-1} g_j \overline{\pi}_j\right\}$$

The following part of the section will detail the two steps of decomposition onto $\{\pi_j\}_{0\leq j<d}$ (leading to the coefficients g_j) and reconstruction through the new basis $\{\overline{\pi}_j\}_{0\leq j<d}$.

Handling generating functions

For handling the basis vectors $\{\overline{\pi}_j\}_{j\in\mathcal{J}}$ in practice only the fractional parts will be needed, and more importantly it suffices to obtain and store only the first one, since the others follow immediately by simply skipping its leading j fractional digits.

When implementing this with finite length vectors - representing the generating functions - these will not just require the first d bits of $\left\{t^J/P(t)\right\}$, but the following $d-1$ as well; these are again easily recovered by simple direct iterations through A.

In the end the precalculation needed for skip ahead with this method leads to substantial space (and computational) savings, being of only $3d$ digits, compared to a $d \times d$ bits matrix; $2d-1$ are needed for the arrival basis, $d+1$ for the start basis.

Decomposition

Projecting $G(t)$ onto an arbitrary basis would require solving a dense linear system of d equations in d unknowns, thus a slow $O(d^3)$ task, which luckily lowers to $O(d^2)$ in the considered triangular case. Now let's recall how in the particular case of π_i it can be carried out as a polynomial multiplication over $\mathbb{F}_2$, which is essentially a

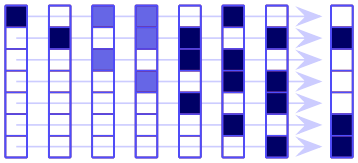


Figure 5.3: *Depiction of a generator state as a linear combination* $\sum_{j\in\mathcal{J}} g_j\pi_j = \mathbf{s}_0$ *on the characteristic basis and with coefficients* $g_j = \{1,1,0,0,1,1,1\}$. *Clearly visible is the triangular structure of the associated linear system for the projection.*

convolution that can be calculated i.e. by FFT in $O(d\ log(d))$ operations[7] ($P(t)$ has $d+1$ terms, while only the first d coefficients of $G(t)$ end up in $g(t)$; these must be padded with zeros to avoid overlapping in the result, so that FFT is applied to two $2d+2$ bit vectors).

This procedure could even be enough for most practical purposes, but it is possible to improve even further bringing complexity down to $O(d)$ as will be shown in Appendix B.

Reconstruction

Just like decomposition was equivalent to polynomial multiplication by the inverse of π_0, in the same way the reconstruction step requires in practice just the multiplication $g(t)\overline{\pi}_0$, where $\overline{\pi}_0$ is considered given. Here again the naive direct approach requires computations on the order of $O(d^2)$, but the use of FFT allows an improvement down to $O(d\ log(d))$.
Despite its lower asymptotic performance the direct reconstruction has the major advantage of being so simple that it can even be implemented on the fly during decomposition, as will be shown later.
In the end the two phases of projection and reconstruction provide together a jump in $O(d\ log(d))$ time, since the latter clearly dominates over decomposition.

Precomputation

The previous steps required the calculation of the target basis which, even when precalculation is admitted, should not be too slow, unlike what happens in the matrix multiplication case. Be it through matrices, through polynomials (as detailed

[7] Actually this is the standard complexity which assumes infinite precision operands; the bit complexity relevant to practical implementations is $O(d\ log(d) log\ log(d))$.

in the following subsection), or by reprojections as in the method just presented, it is possible to obtain longer jumps from smaller ones recursively: regardless of the actual algebraic group used the same addition-chain techniques wellknown in the exponentiation literature (see [5]) can be exploited. Here we'll just just see the simplest method.

The goal of precomputation is obtaining $\overline{\pi}_0$ relative to an arbitrary J and an initial $g(t) = 1$, so let's denote it by $\overline{\pi}_0^J$. Assuming J even and $\overline{\pi}_0^{J/2}$ as given we can obtain $\overline{\pi}_0^J$ by skipping $J/2$ ahead of $\overline{\pi}_0^{J/2}$, that is decomposing $\overline{\pi}_0^{J/2}$ onto the π_i-s and reconstructing with a $J/2$ jump. In this way by induction from $J = 1$ (corresponding to plain application of f) we can precalculate the data needed for all jumps with J being a power of 2.

Moreover through this it is also possible to precalculate everyting needed for an arbitrary jump, through the binary representation of the number J and composition of the jumps corresponding to unit digits. Such procedure requires in the worst case $2\ \log(J)$ jumps, and thus $O(d^2 \log(J)$ or $O(d\ \log(d) \log(j))$ operations depending on the method of choice.

Previous approaches

Recall that we're willing to obtain $\left\{t^J \cdot G(t)\right\}_{,d}$ from $\{G(t)\}_{,d}$. This cannot be calculated explicitly as even just enumerating J terms would be too slow, but since all arithmetic is done on polynomials modulo $P(x)$ one can take some shortcuts. As will become clearer in the following, a simple path for the skip ahead could be just direct polynomial multiplication as in [3]: $\left\{t^J \cdot G(t)\right\}_{,d} = \left\{[t^J\ mod\ P(t)]G(t)\right\}_{,d}$. The factor $[t^J\ mod\ P(t)]$ itself can be precalculated in $O(\log(J))$ steps by repeated squaring, multiply and polynomial modulo reductions. However the CBR method takes a different route.

5.3 The general case - wide word shift registers

For the less restrictive case of a generator making use of word operations with nonunitary word size ($w > 1$) the approach used for LFSR does not hold anymore: apparently there is no state $\boldsymbol{s}_k$ that together with its subsequent iterates $\boldsymbol{s}_{k+1}$, ..., $\boldsymbol{s}_{k+d-1}$ allows for fast decomposition and reconstruction. In particular no triangular iterated basis is necessarily present, since $\boldsymbol{s}_{k+1}$ might always differ from $\boldsymbol{s}_k$ by more than two digits. Here there is a less strong structure than for a LFSR, but depending on r/w

(like for relatively smaller w as happens i.e. in MT19937) it can still be significant. In the following two complementary ways of exploiting it will be presented, focusing on either jump speed or ease of implementation.

In both cases the precalculation stage will follow the very same logic as with LFSR building jumps from shorter ones.

5.3.1 Multiple LFSR

An efficient approach involves splitting the generator stream into w different substreams $\sigma_i = (s_{d-1,i}, ..., s_{1,i}, s_{0,i})^T$ for $(i = 0, ..., w-1)$, which are known to share the common characteristic polynomial $P_A(t)$. Applying A^J to the state is thus equivalent to advancing separately each σ_i as a simpler LFSR through $P_A(t)$ and then composing back the relevant bits into the "full" state generator:

$$s_J = (\overline{\sigma}_{w-1,r-1}, ..., \overline{\sigma}_{0,r-1}, ..., \overline{\sigma}_{w-1,1}, ..., \overline{\sigma}_{0,1}, \overline{\sigma}_{w-1,0}, ..., \overline{\sigma}_{0,0}).$$

The resulting computational complexity can then be traced back to that of LFSR jumps, and is therefore on the order of $wd\,log(d)$.

A more efficient variant on this would be advancing only one among the w substreams, and instead of truncating it to r bits using all d of them to recover the full state S_J; such a transform is guaranteed to exist lowering complexity down to $O(d\,log(d))$.

5.3.2 Pseudocharacteristic basis transform

A slower but much simpler method is perfectly analogous to the LFSR case, using decomposition and reconstruction for the full state; the only issue to be solved is finding the right basis $\{\pi_j\}_{j\in\mathcal{J}}$.

However here no $P(t)$ will be used, and we wont work with proper generating functions. Let's consider instead the $G_i(t)$-s corresponding to the σ_i-s introduced earlier, now the generalized generating function will be $\Sigma := (G_0(t), ..., G_{w-1}(t))^T$.

Where for a LFSR we constructed a basis π_i for the space of generating functions through $d-1$ repeated advancings of π_0, here we will do the same for Σ only $r-1$ times on w different elements such that decomposition is still trivial. Such a set can be in example that corresponding to the w states null everywhere except at one of the

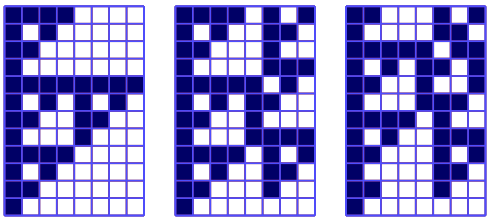


Figure 5.4: *Reconstruction by additive composition of three basis elements*

top w bits, $\{\boldsymbol{e}_1, ..., \boldsymbol{e}_w\}$.

The resulting linear system will still be triangular allowing fast decomposition. However while for a LFSR just a single ”jumped” element was enough for deriving easily the other $d-1$ ones, here we’ll need to keep w different pairs (and obviously also precalculate them).
Reconstruction needs to be carried out directly since convolution is not applicable here. A jump will thus be more memory intensive, up to $3wd$ bits instead of just $3w$, but not slower in the direct method (except for what induced by the larger amount of memory used).

5.4 Generalising further: pulmonary generators

So called *pulmonary* generators ([8],[9]) achieve better efficiency using a more general transition matrix A, in particular a second shorter feedback loop[8] is added to provide continuously changing new inputs to the main one.
The skip ahead for such generators can often be traced back to the case treated in the previous section.

With $T_1 \in M_{w\times d}(\mathbb{F}_2)$, and T_2 aswell, the generator matrix has now the following form:

[8] often referred to as the respiratory circulation, as opposed to the normal blood circulation of a normal FSR; accordingly the state word implementing it is called *lung*

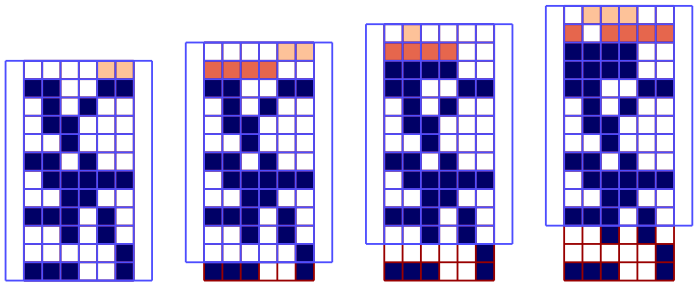


Figure 5.5: *Transitions for a pulmonary FSR*

$$A = \begin{bmatrix} \boxed{T_1} \\ \boxed{T_2} \\ \cdot \; I_w \quad \cdot \quad \cdot \\ \quad \ddots \\ \cdot \quad \cdot \quad I_w \quad \cdot \end{bmatrix}$$

In the practical round-robin implementation two contiguos words are updated in the state instead of just one, see *fig.* 5.5. This means that the top word will be changed twice in subsequent f iterations, so that its first value is only temporary. The ”final” values produced by T_2 grow in a stream for which the previous algorithms still work.

All we need to apply the previous method to such setting is the ability to transform back and forth the top word between its two values. In this setting (A has full rank and is invertible) it must always be possible, however it is not always a fast task, but such is the case i.e. for WELL1024a.
In fact this generator has another good property: transforming the basis ”seeds” $\boldsymbol{e}_1, ..., \boldsymbol{e}_w$ does only add new words without modifying

5.5 Implementation

One major point of characteristic basis skip ahead is the reduction in memory usage (it grows only like dw, that is linearly in the state size), which even in the slower $O(d^2)$ direct algorithm allows in practice for faster execution than matrix multiplication.

A significant advantage is that in the actual implementation there is no need to

obtain explicitely the characteristic polynomial of the generator recurrence, although throughout the analysis it appeared often and at key points. This is possible because all PRNG specific calculations are done directly through the transition function itself, which is given. Other skip ahead methods might instead require resorting to more involved number theoretic operations,

Here is shown the extreme simplicity of the relevant skip ahead code, in the general w case, for a simple $O(d^2)$ jump with direct decomposition and reconstruction:

```
for (i=0; i<r; i++)
 for (j=0; j<w; j++)
  if (src.statew[i] & (1<<j))
   for (k=0; k<r; k++)
   {
    src.statew[k] ^= pi_src[j].statew[r-1+k-i];
    dst.statew[k] ^= pi_dst[j].statew[r-1+k-i];
   }
```

In the inner loop the first binary subtraction solves the triangular system for the source projection by straightforward gaussian elimination, while the following addition updates the target state reconstruction.

The code in the LFSR case would be even simpler due to the collapse of the second loop over j when $w = 1$.

Results The implementation focused on WELL1024a for its good compromise between state/stream size and fast bit mixing leading to quick recovery from almost null states.
The state size is 128 bytes, a start or arrival basis 4 Kb; a full precalculation for all 2^j jumps 16 Mb.
In many consumer processors currently the L1 cache is 16Kb, so that the procedure fits within, without incurring in large penalties.

On an AMD Turion64 1.6Ghz the rate was of about 8000 jumps per second versus 3200 $A\boldsymbol{s}$ matrix multiplications per second, a factor of 2,5.
This translates in a speedup for precalculations of at least $2,5r = 80x$, making it nearly transparent (1-4 secs) on consumer PCs.

It is also possible to get some further speed gains by a constant factor through grouping of calculations in a similar way to [3], however one might consider this an implementational finesse of limited theorical interest, so it won't be treated here.

5.6 Comparisons

As already anticipated this is not the only algorithm presented so far for skip ahead.
Brent recently ([2]) mentioned one making use of generating functions and polynomial multiplication, but the actual reference [1] does not really explain it in much greater detail. Furthermore this would only work for a LFSR, and for a more general setting the computational cost would rise again, expecially when ignoring any FSR structure.

Haramoto et al extended the polynomial method (see [3]) to the more general class of linear PRNGS with arbitrary matrix A; their approach is based on the precalculation of $t^J mod P(t)$ (an $O(d^2 \log(J))$ task) which is then used to independently advance the successions of each state bit and also gives jumps at a cost of $O(d^2)$, with precalculations at only $O(d^2 \log(J)$. On top of such framework they can improve by a constant the algorithm efficiency through careful terms grouping.
Thus the theoric complexity of the jump is not better than those of the approaches presented in this work, but the special structure exploited here allows for much easier implementation, and performance for small state sizes is comparable even in the $O(d)$ case.

Obviously the polynomial and CBR approaches must be equivalent at some level, since they represent different viewpoints for the same calculations. This becomes evident in the simpler LFSR case as both methods amount to conditional summation of consecutive generator states; however while the polynomial method uses precalculation to embody the jump into the coefficients of such summation, CBR keeps it into the summed states already. This latter approach is significantly more intuitive and does not rely explicitely on any underlying algebra of the PRNG other than its linearity, in particular the characteristic polynomial of the recurrence needs not be derived.

5.7 Conclusions

I have shown a simple setting comprising different algorithms for stream movements appropriate for use with many of the most common pseudorandom generators.

Even the more memory intensive versions have a reasonable footprint (much lower than that required by the standard matrix multiplication) that makes them already competitive.

Within the proposed setting it is possible to choose between methods with either simple implementation or high performance:

	Prec. memory	Prec. complexity	Jump complexity
Matrix	d^2	$O(d^3 \log(J))$	$O(d^2)$
Polynomial method	d	$O(d^2 \log(J))$	$O(d^2)$
CBR direct	$3d$	$O(d^2 \log(J))$	$O(d^2)$
CBR FFT	$3d$	$O(d \log(d) \log(J))$	$O(d \log(d))$
$w > 1$ direct	$3wd$	$O(d^2 \log(J))$	$O(d^2)$
$w > 1$ parallel FFT	$3d$	$O(d \log(d) \log(J))$	$O(w\, d \log(d))$

They all work easily as black boxes requiring in input just the start generator state and the function providing the transition: $\boldsymbol{s}_J = F(\boldsymbol{s}_0, f)$, so no generator-specific inner working details are needed.

Although the current direct implementation proved itself already fast enough for most practical needs it will be interesting to test the fast algorithm with an FFT specifically tuned for the $\mathbb{F}_2$ setting. Furhermore a direct empirical comparison with the polynomial method will be useful in assessing the relative merits of the two approaches.

Finally I hope that the characteristic basis will be useful for analysis and developement of new pseudorandom generators, beyond the scope of skip ahead.

Aknowledgements I would like to thank Prof. Matsumoto for the comments at MCQMC 2006 that significantly influenced exposition in this final paper, and Prof. Mascagni for originally suggesting the task as requiring investigation.

5.8 Appendix A - Brief review on finite field arithmetic

A polynomial $P(t) \in \mathbb{F}_2[t]$ over $\mathbb{F}_2 := \mathbb{Z}/2\mathbb{Z}$ having trivial factors only is called *irreducible, reducible* otherwise. An irreducible $P(t)$ of degree d is also *primitive* if $t^i - 1 \neq 0 \bmod P(t) \quad \forall i < 2^d - 1$; in this case the multiplicative group of the field $\mathbb{F}_2[t]/\langle P(t)\rangle$ is cyclic and generated by t, otherwise the order of t is a proper divisor of $2^d - 1$. Thus primality of $2^d - 1$ implies primitivity of $P(t)$.
The characteristic polynomial of A is defined as:
$P_A(t) = det(A - tI) = t^d - a_1 t^{d-1} - ... - a_{d-1}t - a_d$.
Componentwise holds: $s_i = a_1 s_{i-1} + ... + a_d s_{i-d}$.

The set $\mathbb{F}_2[[t^{-1}]]$ of formal Laurent series comprises formal power series having a finite number of terms of positive degree and infinite terms of negative degree. Operations between elements in $\mathbb{F}_2[[t^{-1}]]$ are defined in the natural way but -unlike polynomials- these cannot be evaluated at a specific point t.

It is possible to associate to each output stream (that is to each couple of $P(t)$ and initial state) a formal power series (the generating function)
$G(t) = s_0 t^{-1} + s_1 t^{-2} + s_2 t^{-3} + ... = \sum_{i=1}^{\infty} s_{i-1} t^{-i} \in \mathbb{F}_2[[t^{-1}]]$
It is easy to verify that formally $G(t)P(t) = g(t) := -\sum_{j=0}^{d-1} \sum_{i=0}^{d-1-j} a_{d-i-j-1} s_i t^j \in \mathbb{F}_2[t]$

So we get a correspondence between the series initializations and the $g(t)$-s.
For a generator with state $\boldsymbol{s}_i$ the series spanned by the various digits $s_{i,j}$ are all cycling on the same repeating stream (defined by the common characteristic polynomial $P_A(t)$) but starting from different points. Accordingly it is possible to define the vector $(G_1(t), ..., G_d(t)) = (g_1(t), ..., g_d(t))/P_A(t)$.

5.9 Appendix B - A faster decomposition procedure

Here an $O(d)$ algorithm for projection will be presented in the LFSR case.

Let's first recall the standard triangular system $s_{0,i} = \sum_{j=d-1-i}^{d-1} g_j \pi_{j,i}$ with solution:
$g_{d-1} = s_{0,0}$

$$g_{d-1-i} = s_{0,i} - \sum_{j=d-i}^{d-1} g_j \pi_{j,i}$$

The summations in the second term don't really need to be carried out explicitely each time: thanks again to the structure of the π_j-s (sharing the same ordered coordinates in shifted positions) it is possible to evaluate these expressions in a recursive way. For this we grow aside a stream representing the running sum of the components of $G(t)$ calculated so far.

$g_{d-1} = s_{0,0}$	$\boldsymbol{\gamma}_1 = g_{d-1}\boldsymbol{\pi}_1$
$g_{d-2} = s_{0,1} - \gamma_{1,d-1}$	$\boldsymbol{\gamma}_2 = g_{d-2}\boldsymbol{\pi}_1 + A\boldsymbol{\gamma}_1$
$g_{d-3} = s_{0,2} - \gamma_{2,d-1}$	$\boldsymbol{\gamma}_3 = g_{d-3}\boldsymbol{\pi}_1 + A\boldsymbol{\gamma}_2$
...	...
$g_0 = s_{0,d-1} - \gamma_{d-1,d-1}$	

In the end we have traded a d complexity factor with a bit summation ($g.\boldsymbol{\pi}_1$) plus the application of the transition function, which is assumed built for maximum efficiency and can therefore be considered constant time. This translates in a resulting complexity for the projection of only $O(d)$.

Bibliography

[1] R.P. Brent, *On the Periods of Generalized Fibonacci Recurrences*, Mathematics of Computation archive. v. 63, 207 (1994)

[2] R.P. Brent, *Fast and reliable random number generators for scientific computing* Slides, (2004)

[3] H. Haramoto, M. Matsumoto, T. Nishimura, F. Panneton, P. L'Ecuyer *Efficient Jump ahead for F_2 linear random number generators* INFORMS Journal on computing (2008)

[4] H. Haramoto, M. Matsumoto, P. L'Ecuyer, *A Fast Jump Ahead Algorithm for Linear Recurrences in a Polynomial Space*, Proceedings of SETA 2008

[5] D. Knuth, *The Art of computer programming v. II - Seminumerical algorithms*, Addison Wesley Longman, (1994)

[6] M. Matsumoto, T. Nishimura, *Dynamic creation of pseudorandom number generators*, Monte Carlo and Quasi-Monte Carlo Methods 1998, Springer, (2000), pp 56–69

[7] F. Panneton, *Construction d'ensembles de points basee sur des recurrences lineaires dans un corps fini de caracteristique 2 pour la simulation Monte Carlo et l'integration quasi-Monte Carlo* (2004)

[8] F. Panneton, P. L'Ecuyer, M. Matsumoto, *Improved Long-Period Generators Based on Linear Recurrence Modulo 2: Overview and comparison*, ACM Transactions on Mathematical Software, v.32 n.1 (2006)

[9] M. Saito, H. Haramoto, M. Matsumoto, F. Panneton, T. Nishimura, *Pulmonary LFSR: pseudorandom number generators with multiple feedbacks and reducible transitions*, (2006)

Appendix A

On finite number of particles in Monte Carlo kinetic simulations

Monte Carlo methods are the most popular methods for solving problems in kinetic theory [2, 5]. In this short remark we emphasize some of the side effects due to the use of conservative methods over a finite number of statistical samples (particles) in the simulation. The most relevant aspect is that the steady states of the system are compactly supported and thus they cannot be Maxwellian (or any other non compactly supported statistics) unless the number of particles goes to infinity. These aspects are studied numerically with the help of a simple one-dimensional space homogeneous kinetic model.

A.1 Introduction

The numerical solution of kinetic equations is usually performed through statistical simulation methods such as Monte Carlo [3]. The reason for this is twofold, on the one hand probabilistic techniques provide an efficient toolbox for the simulation due to the reduced computational cost when compared with deterministic schemes, on the other hand the evolution of the statistical samples follows the microscopic binary interaction dynamics thus providing all the relevant physical properties of the system. Traditionally the methods are considered extremely efficient when dealing with stationary problems. In such case, in fact, fluctuations can be eliminated by taking subsequent averages of the solution after then a certain ”stationary time” has been reached. Here we show, with the help of a simple one-dimensional system, that this averaging procedure does not guarantee convergence towards the correct steady state

due to finite number of particles correlations introduced by the microscopic conservation laws. Similar analysis for rarefied gas dynamics have been done in [9, 6].

A.1.1 The model equation

We will consider a simple one–dimensional kinetic model, where the binary interaction between particles obey to the law

$$v' = v\cos\theta - w\sin\theta, \qquad w' = v\sin\theta + w\cos\theta, \tag{A.1}$$

where $\theta \in [-\pi, \pi]$ is a collision parameter. The microscopic energy after the binary interaction rule is conserved

$$(v')^2 + (w')^2 = v^2 + w^2, \tag{A.2}$$

whereas momentum is not.

Let $f(v,t)$ denote the distribution of particles with velocity $v \in \mathbb{R}$ at time $t \geq 0$. The kinetic model can be easily derived by standard methods of kinetic theory, considering that the change in time of $f(v,t)$ depends on a balance between the gain and loss of particles with velocity v due to binary collisions. This leads to the following integro-differential equation of Boltzmann type [4],

$$\frac{\partial f}{\partial t} = \int_{\mathbb{R}}\int_{-\pi}^{\pi} \frac{1}{2\pi}\left(f(v')f(w') - f(v)f(w)\right) d\theta\, dw. \tag{A.3}$$

As a consequence of the binary interaction the second momentum of the solution is conserved in time, whereas the first momentum is preserved only if initially it is equal to zero. For this model one can show that the stationary solution $f_\infty(v)$ is the Maxwell density

$$f_\infty(v) = \frac{1}{\sqrt{2\pi}} e^{-v^2/2}. \tag{A.4}$$

A standard Monte Carlo method for this equation can be easily derived using either Bird's or Nanbu's algorithm for Maxwell molecules [2, 5]. The two algorithms differ mainly in the way the time discretization is treated, but not in the way collisions (sampling from the collision integral operator) are performed. Our results do not differ for the two methods.

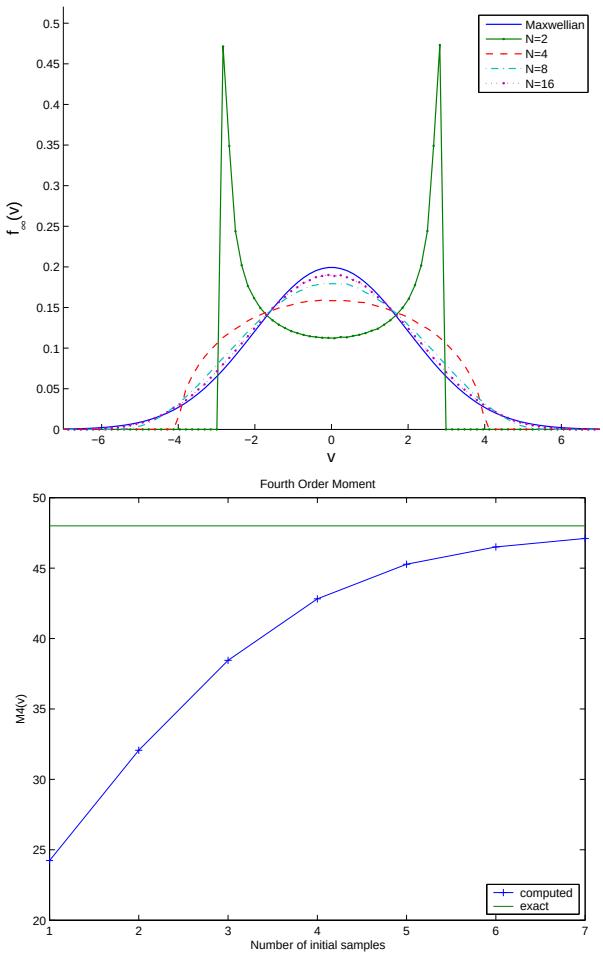


Figure A.1: Equilibrium states for different finite sets of particles vs Maxwellian (left) and equilibrium value of the fourth order moment for the different finite sets of particles.

A.2 Numerical results

The problem we consider here is related to the effect of the finite number of particles in Monte Carlo simulations. Note that given a set of particles $v_1, v_2, \ldots, v_N$ with energy $E = \frac{1}{2}\sum_{i=1}^{N} v_i^2$, we have the inequality

$$|v_i| \leq R_N = \sqrt{2EN}. \tag{A.5}$$

As a consequence of this, any particle dynamic, namely any transformation of the type

$$v_i' = \phi_i(v_1, \ldots, v_N), \quad i = 1, \ldots, N, \tag{A.6}$$

that preserves exactly energy is such that the particle solution remains compactly supported in $[-R_N, R_N]$ at any time. This implies that the distribution of such particles cannot be Maxwellian (or any other non compactly supported statistics) unless the particles number goes to infinity. This is exactly what happens if we use the so-called Nanbu-Babovsky [1] strategy of performing collisions by pairs so that the Monte Carlo methods are exactly conservative and not conservative in the mean. We report in Figure 1 (left) the numerical distribution of the finite sets of particles in the case of the one-dimensional Maxwell model (A.3). The results have been obtained taking initially Maxwellian samples with zero mean and energy 4 and then averaging in time over the Monte Carlo solutions to the equation. For very small numbers of particles it is remarkable that the computed distribution differ considerably from the expected Maxwellian. The different fourth order moments of the corresponding steady solutions are then plotted in Figure 1 (right) against the exact fourth order moment of the Maxwellian. We point out that such small particle numbers can be present is some cells when one consider fully non homogeneous rarefied gas flow simulations and thus, even if the transport part can affect the nature of these correlations, a particular care has to be taken when averaging over such small numbers. Similar conclusion are valid also for different kinetic models where the steady state statistics is not compactly supported like in granular gases, plasma physics, quantum kinetic theory, traffic flows and economic models.

Bibliography

[1] H. Babovsky, *On a simulation scheme for the Boltzmann equation*, Math. Methods Appl. Sci., 8 (1986), pp. 223–233.

[2] G. A. Bird, *Molecular Gas Dynamics*, Oxford University Press, London, 1976.

[3] C. Cercignani, *Rarefied Gas Dynamics: From Basic Concepts to Actual Calculations*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge (2000).

[4] M. Kac, *Probability and Related Topics in Physical Sciences*, Lectures in Appl. Math., Interscience Publishers, London, New York, 1959.

[5] K. Nanbu, *Direct simulation scheme derived from the Boltzmann equation*, J. Phys. Soc. Japan, 49 (1980), pp. 2042–2049.

[6] L. Pareschi, G. Russo, S. Trazzi, A. Shevryn, Ye. Bondar, M. Ivanov, *Comparison betweenTime Relaxed Monte Carlo Method and Majorant Frequency Scheme methods for the space homogeneous Boltzmann equation*, RAREFIED GAS DYNAMICS: 24th International Symposium on Rarefied Gas Dynamics, AIP Conference Proceedings, 762, (2005), pp. 571–576.

[7] L. Pareschi, G. Toscani, *Self-similarity and power-like tails in nonconservative kinetic models*, J.Stat. Phys., 124, (2006), 747–779.

[8] D. I. Pullin, *Generation of normal variates with given sample*, J. Statist. Comput. Simulation, 9 (1979), pp. 303–309.

[9] V. Ya. Rudyak, *Correlations in a finite number of particles system simulating a rarefied gas*, Fluid Dynamics, 26, (1991), pp. 909–914.

Appendix B

Simulation graphs

Performed simulations cover altogether 35 different scenarios (while separately handling heavy and light traffic). In the previous relation only those showing results of some interest were shown, and they're analysed and discussed in more detail in the following.
Given the peculiar geometry of the examined network, it was possible to split the overall traffic according to four different movement modes, corresponding to different destination directions (thereby excluding improbable and in any case neglectable U-shaped routes): in order they are from north to south, from south to north, from west to east and finally from east to west. In the cases of north-south axis the Cispadana has been modeled as bidirectional since it is in practive orthogonal and lacks a preferential directionality. On the other hand the directional selection was used in the cases of traffic on the west-east axis.
These four cases have themselves been further split in more scenarios each, accordint to the realization of the Ferrara-A22 connection in the two variants (near Mantova or near Reggiolo) and further according to the inclusion of Ferrara Mare and E55.

Figure B.1: Scenario 1

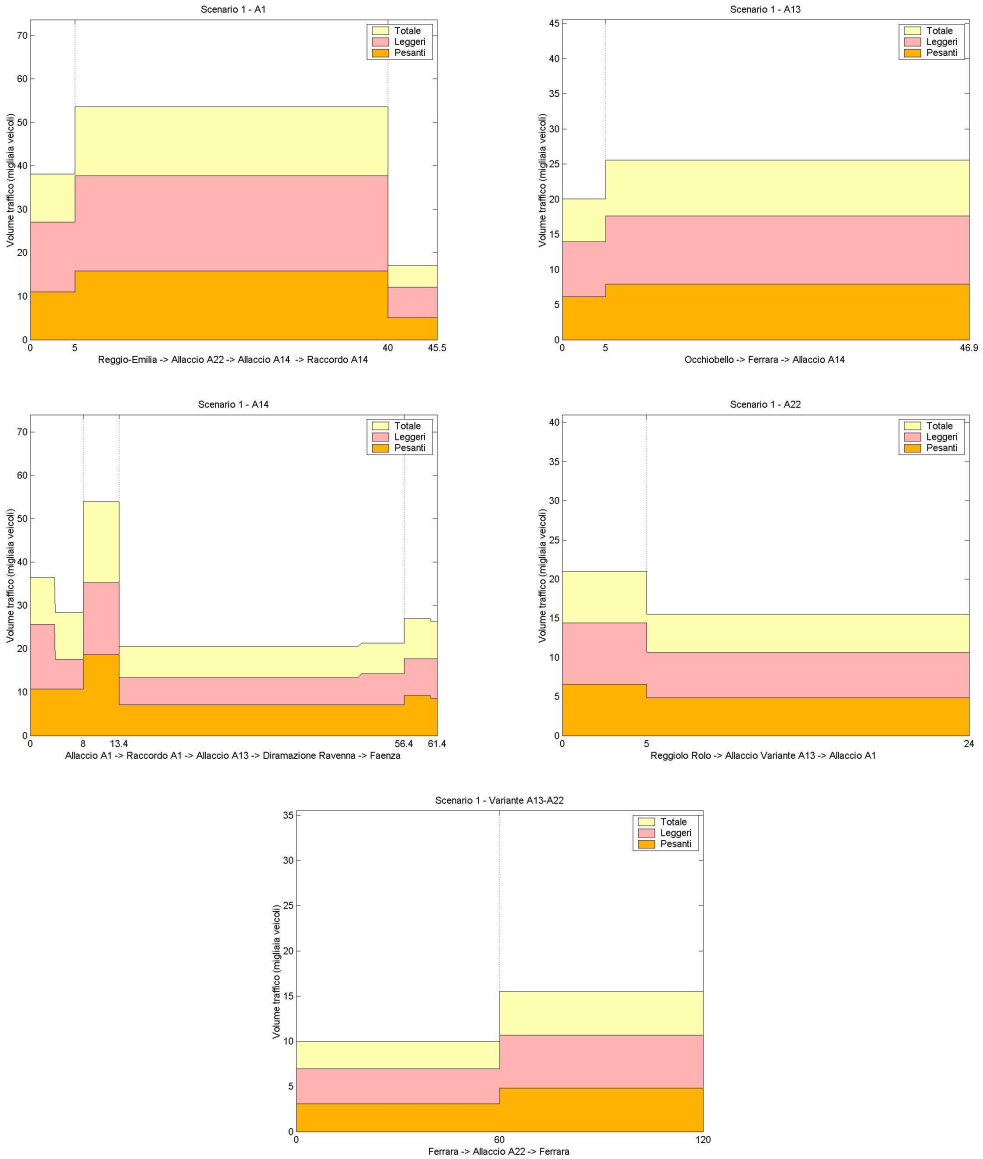


Figure B.2: Scenario 2

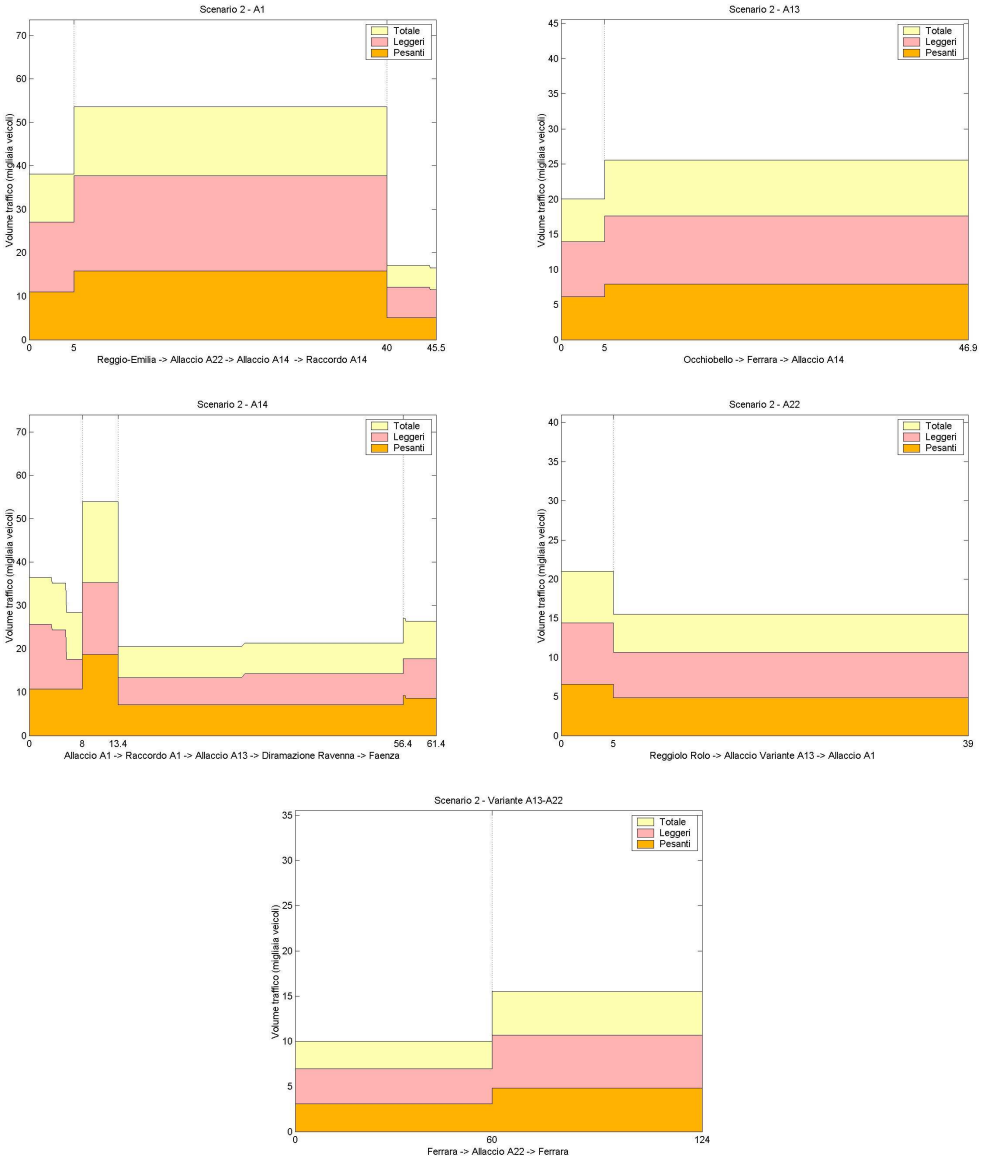


Figure B.3: Scenario 3

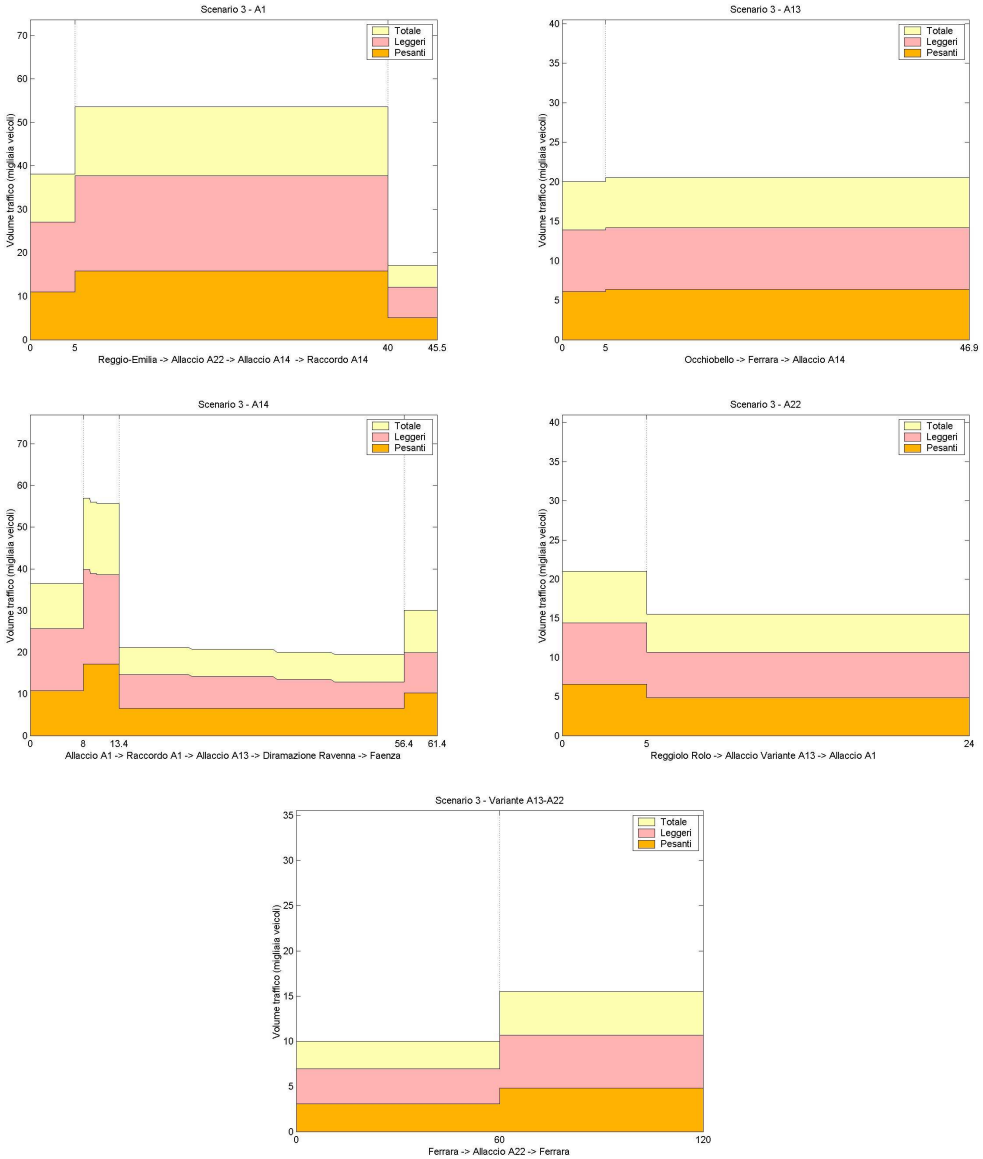


Figure B.4: Scenario 4

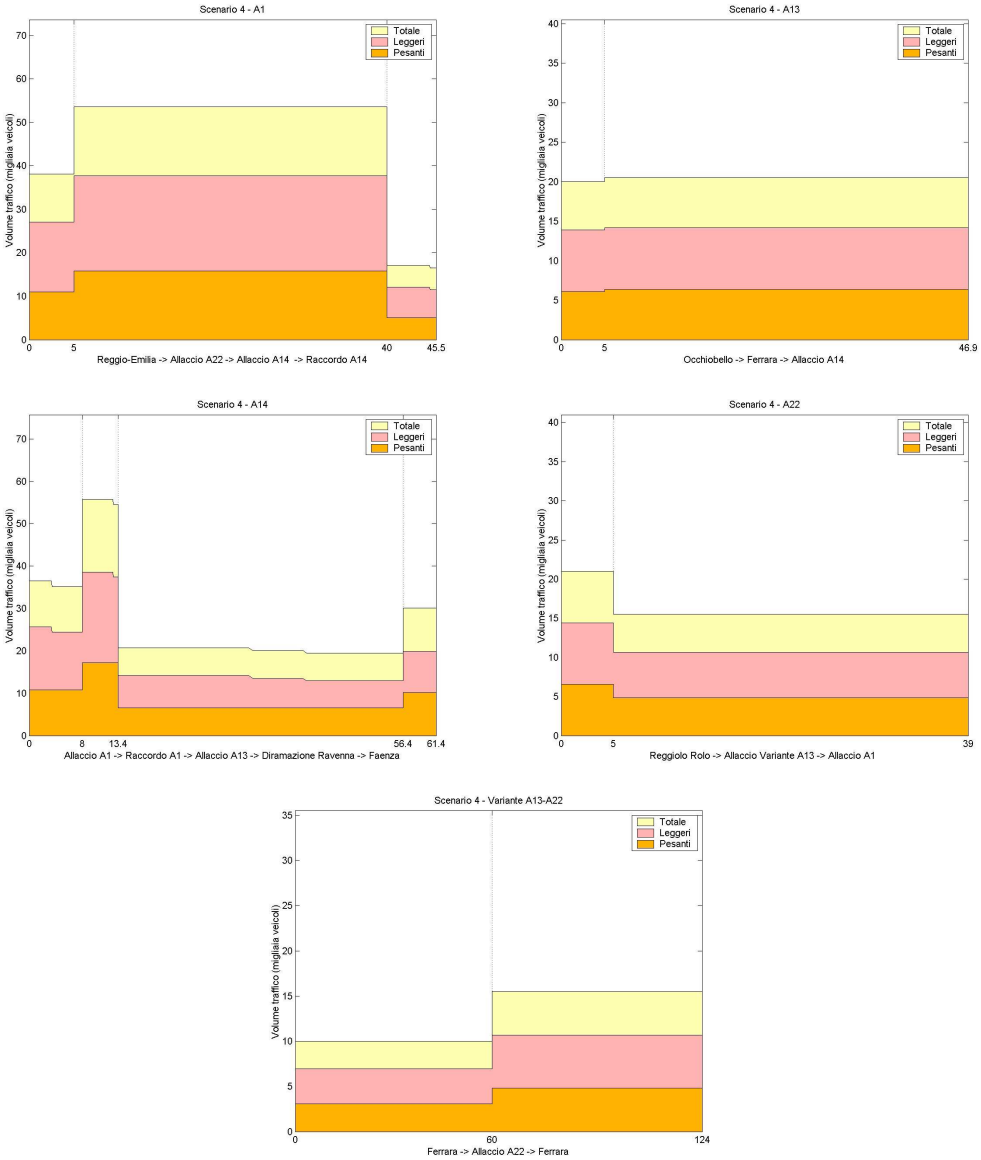


Figure B.5: Scenario 5

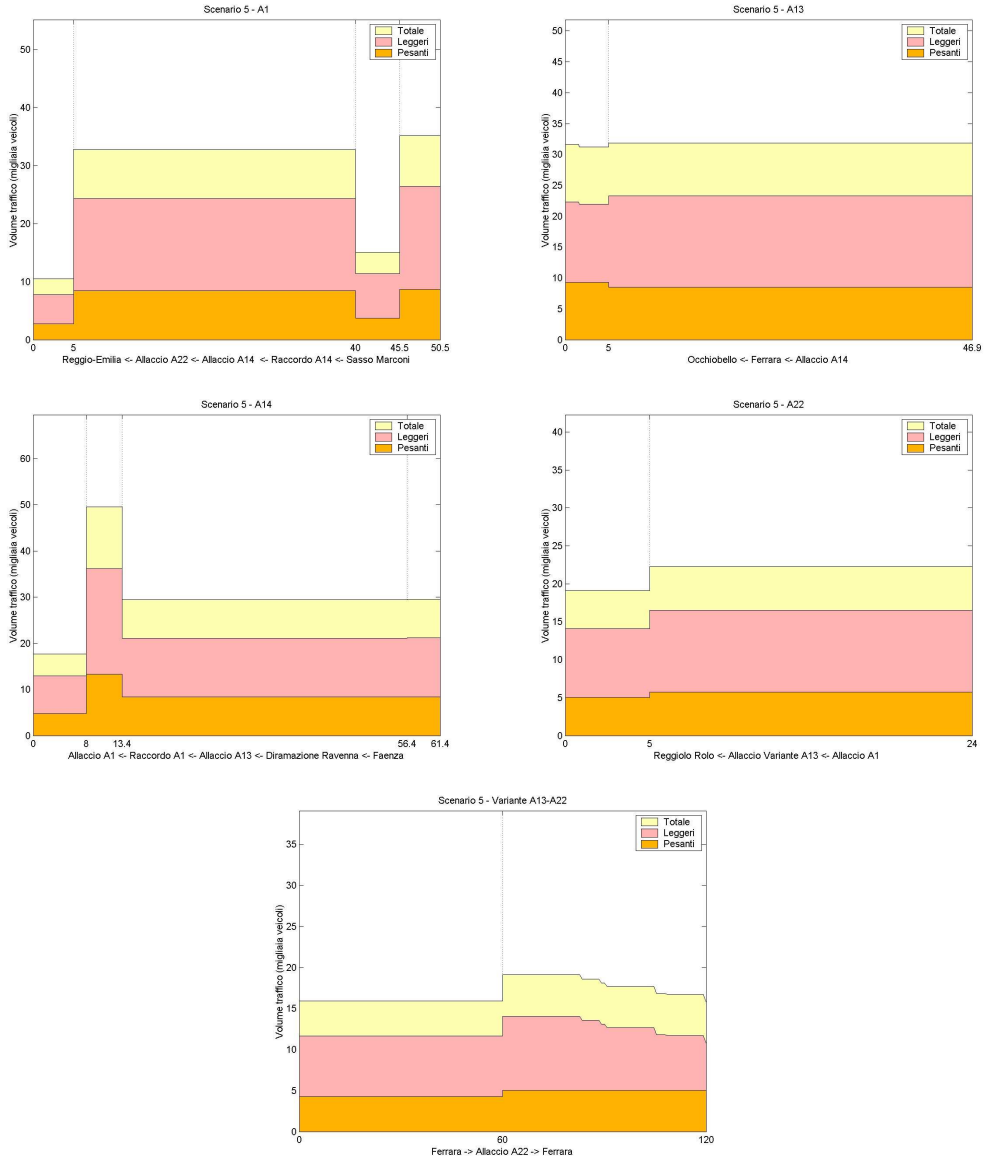


Figure B.6: Scenario 6

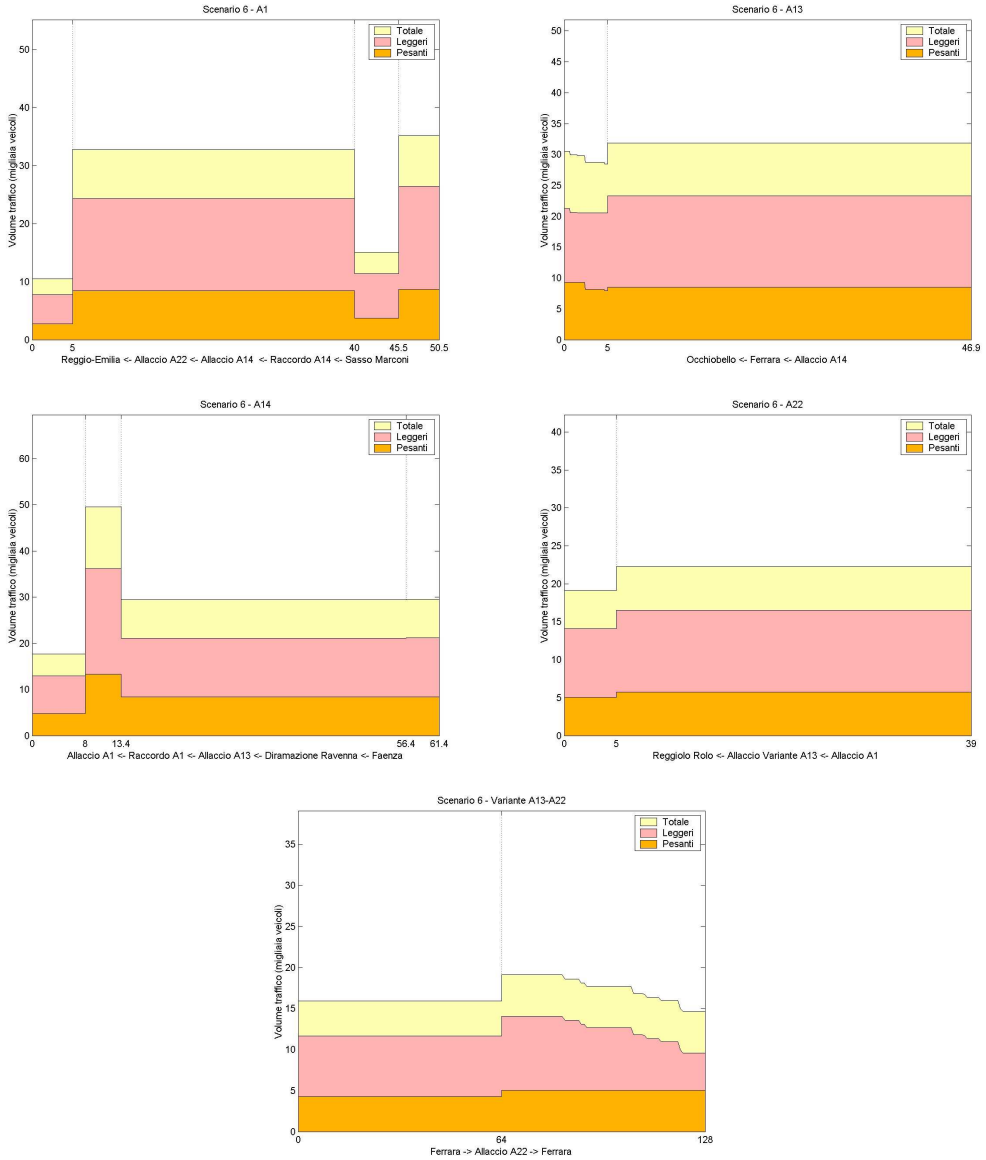


Figure B.7: Scenario 7

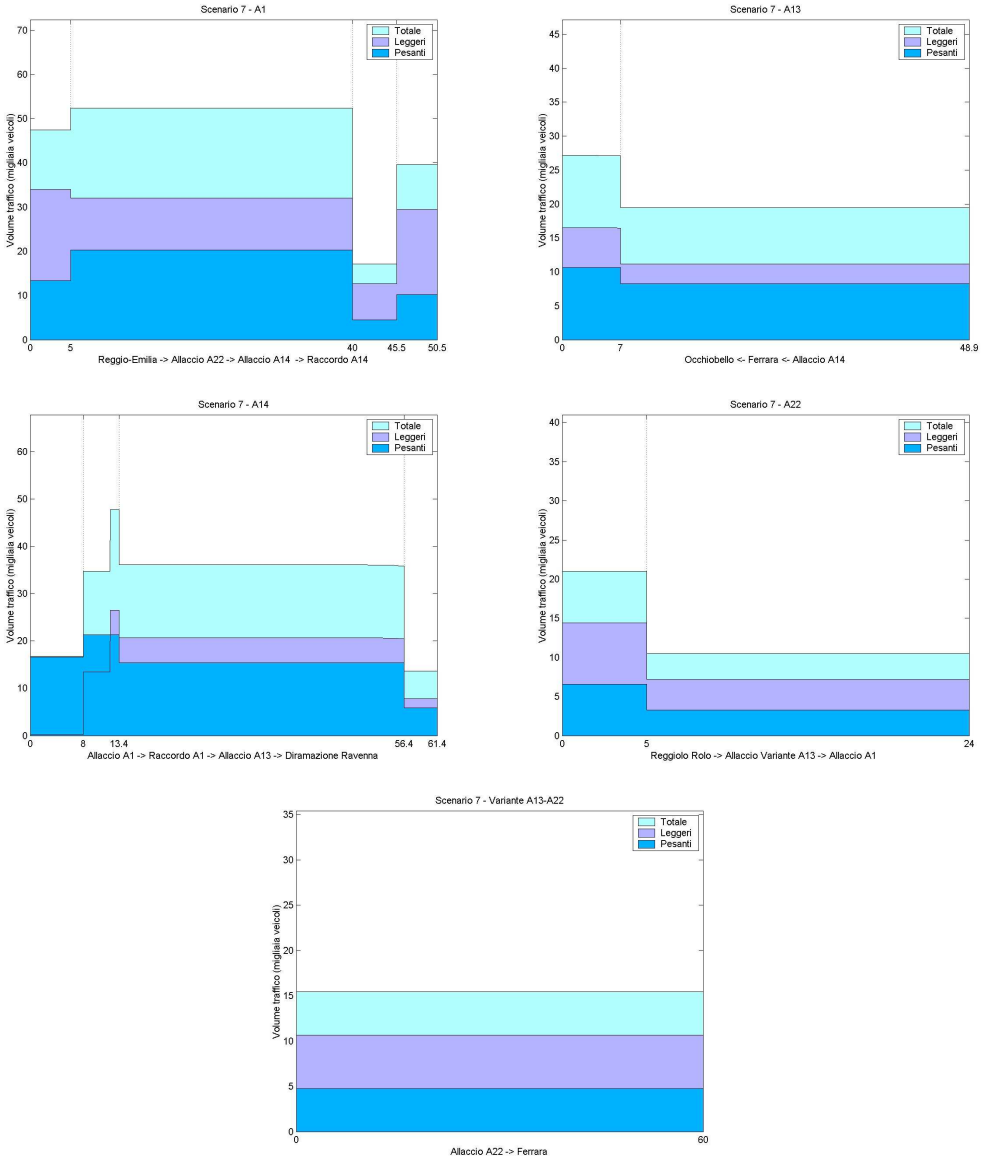


Figure B.8: Scenario 8

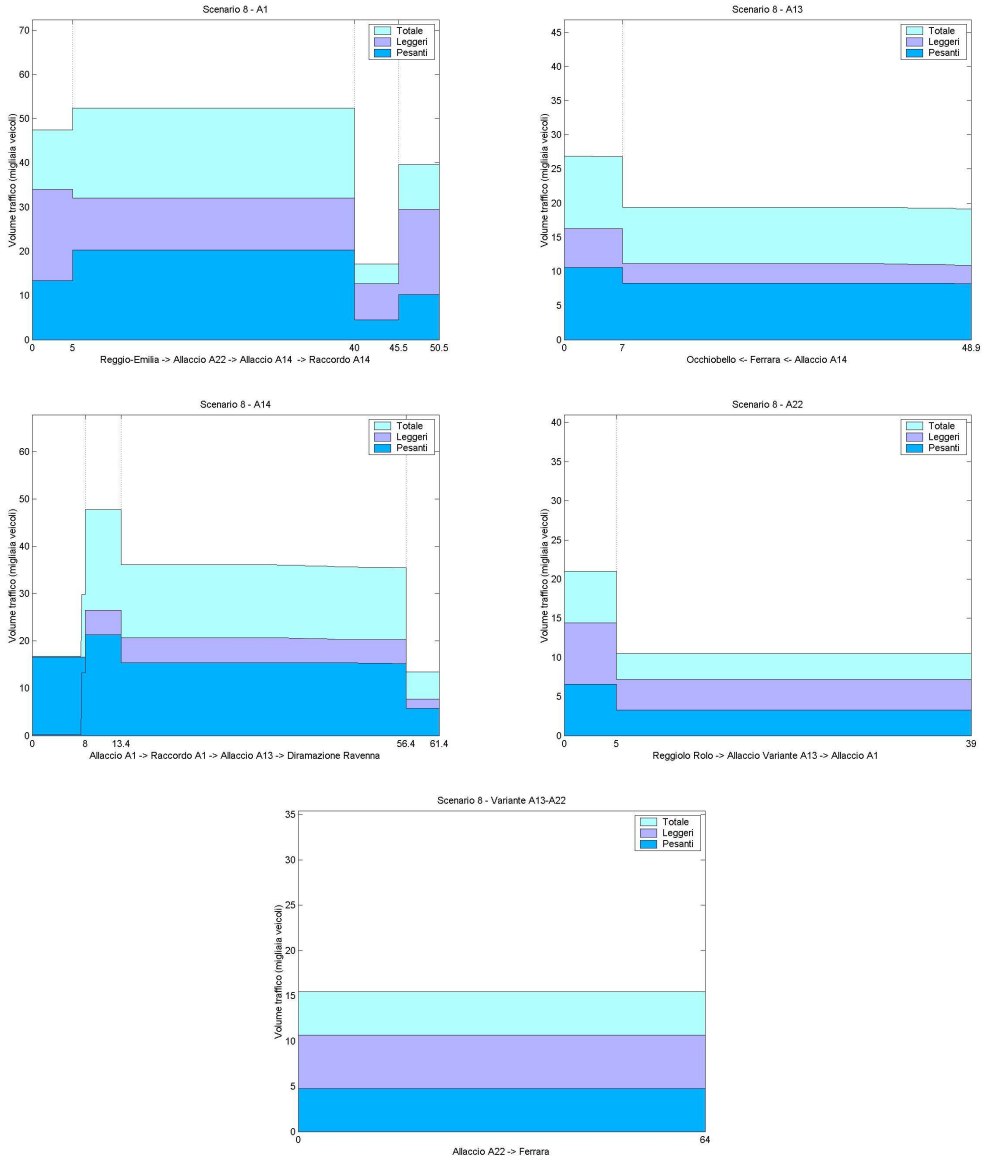


Figure B.9: Scenario 9

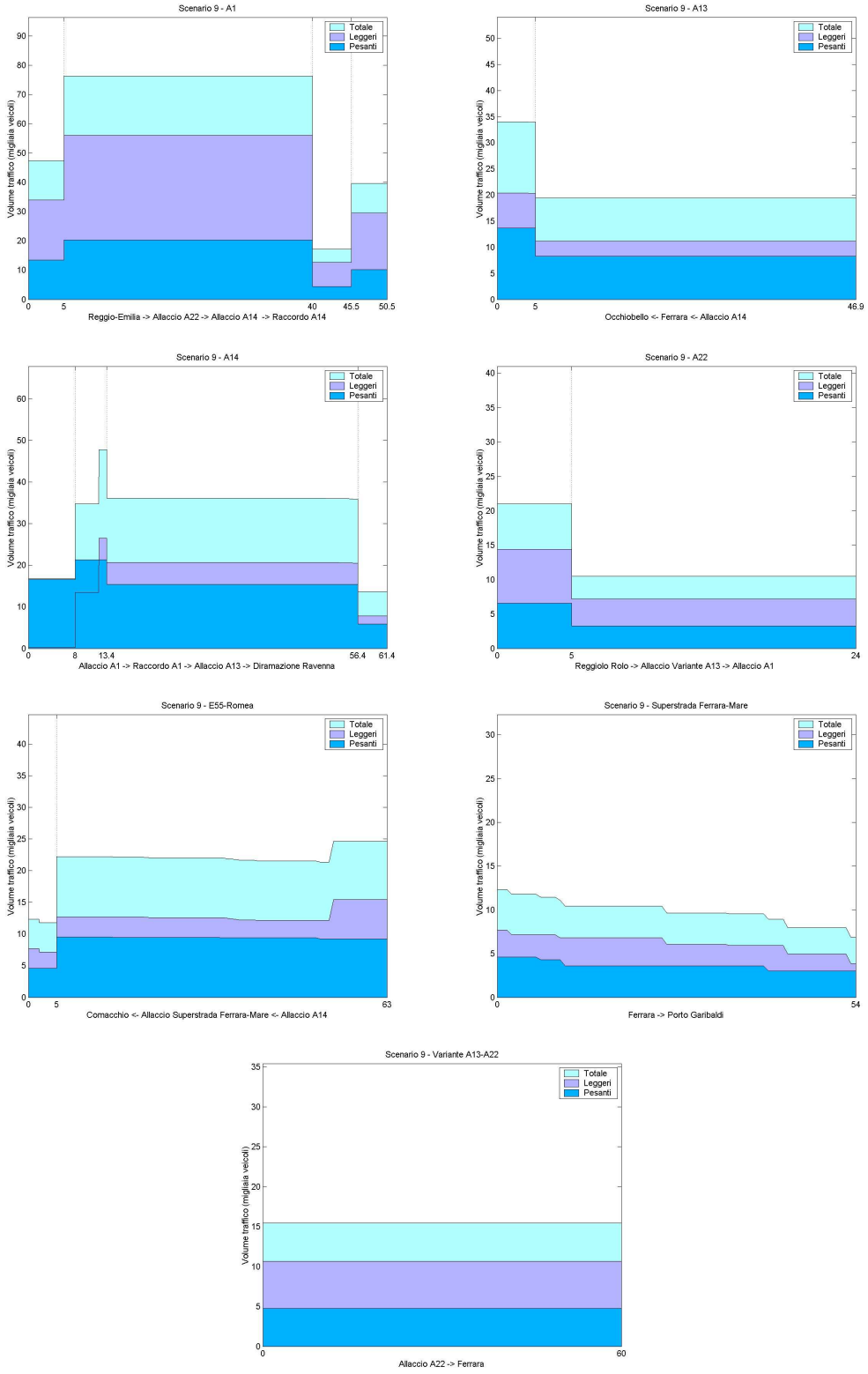


Figure B.10: Scenario 10

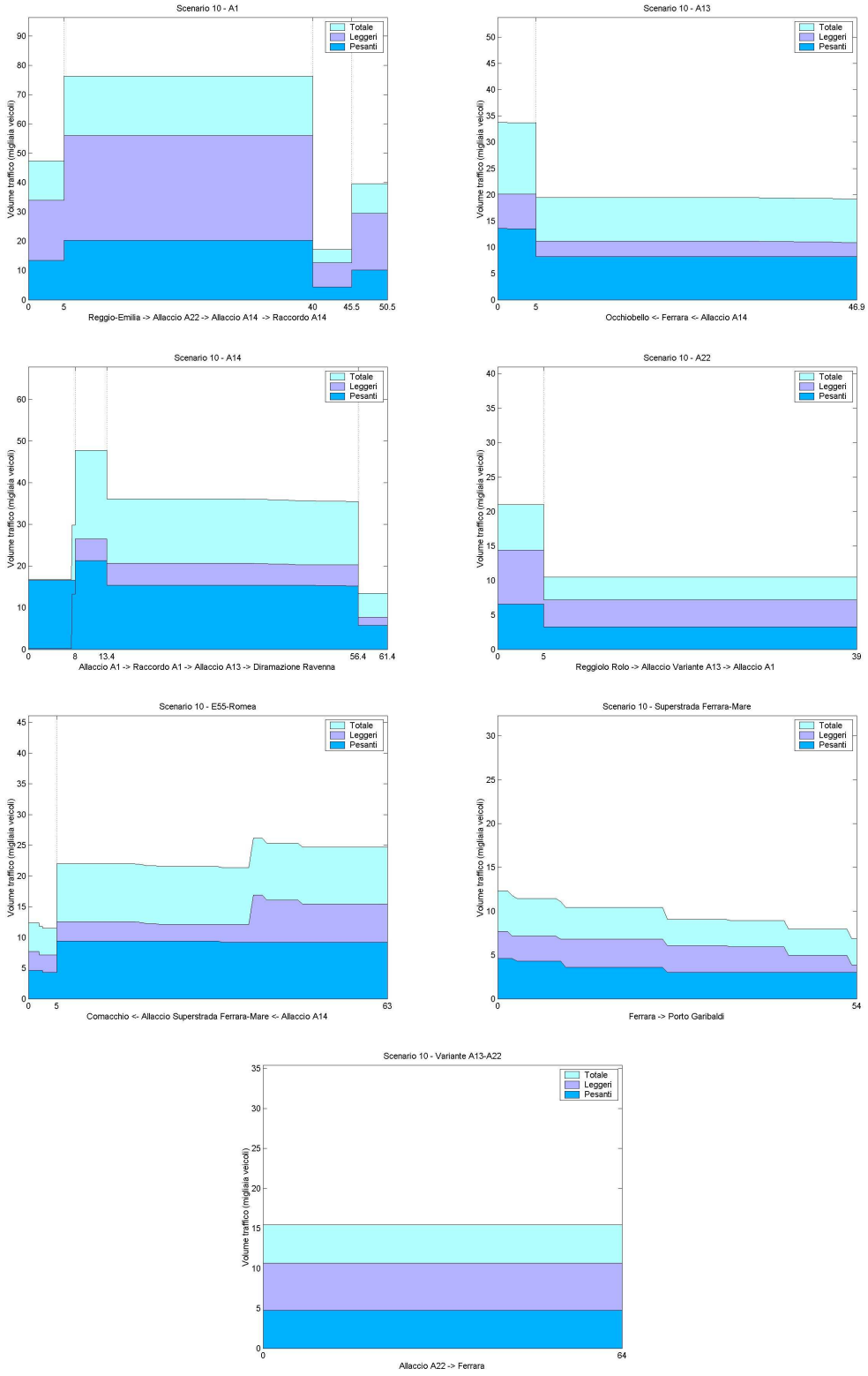


Figure B.11: Scenario 11

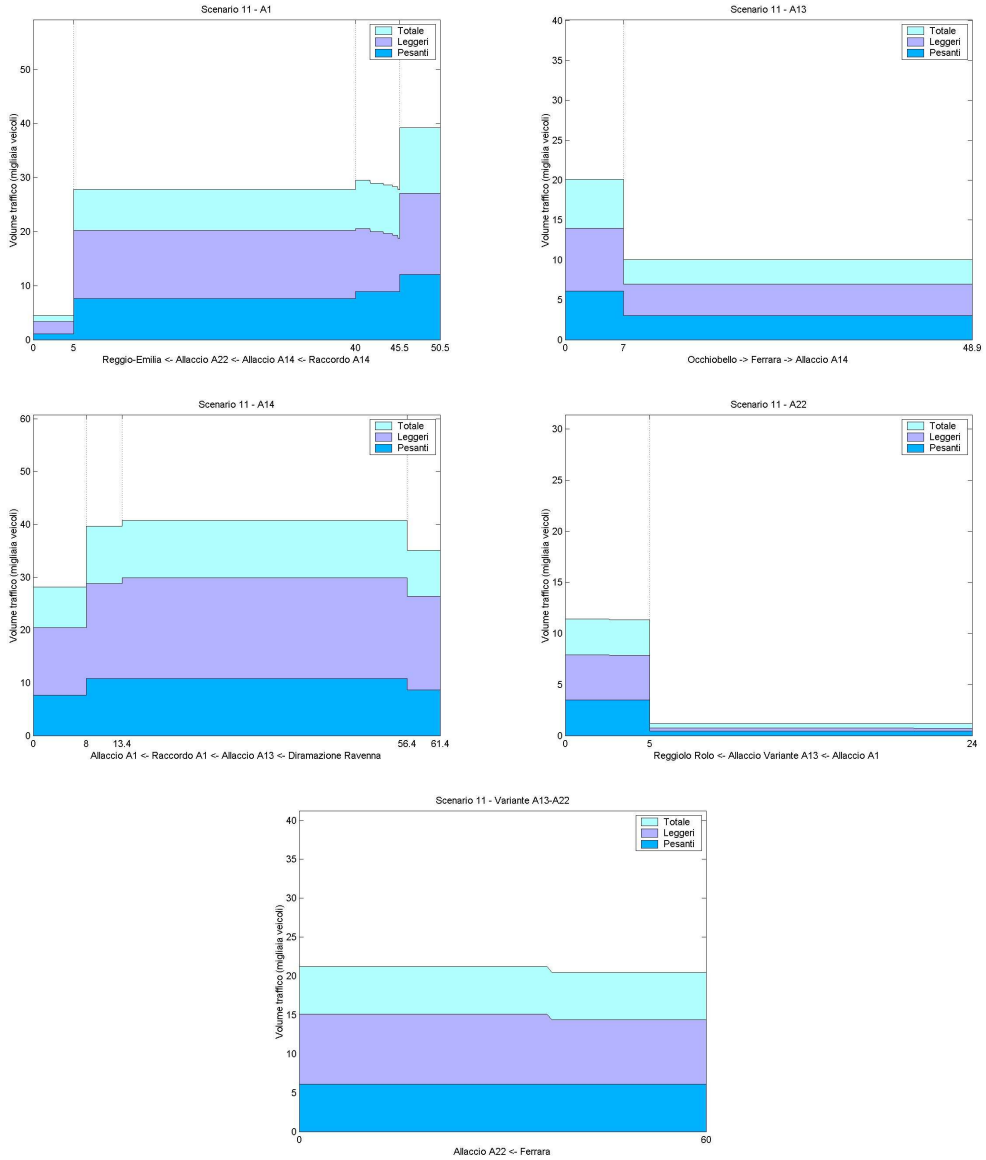


Figure B.12: Scenario 12

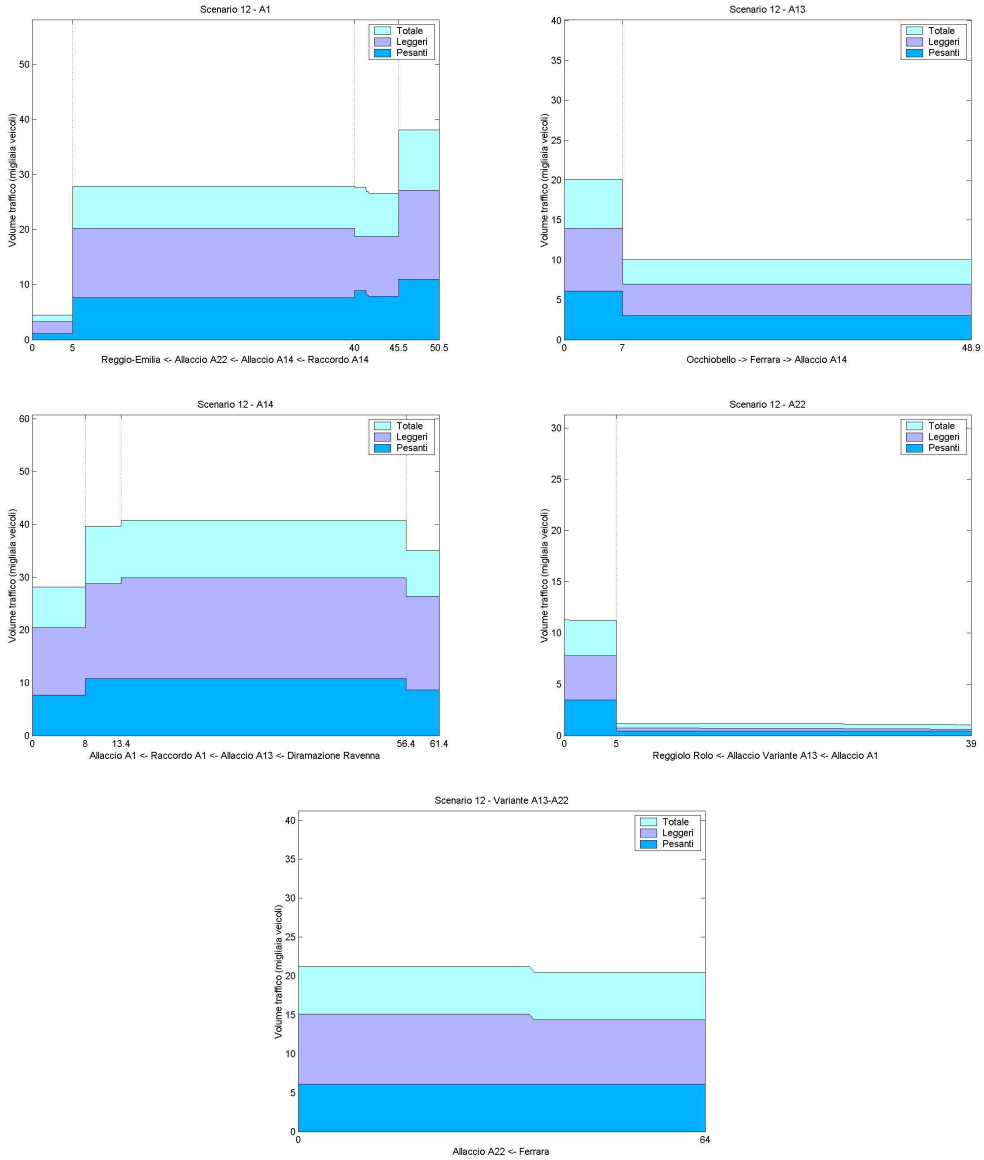


Figure B.13: Scenario 13

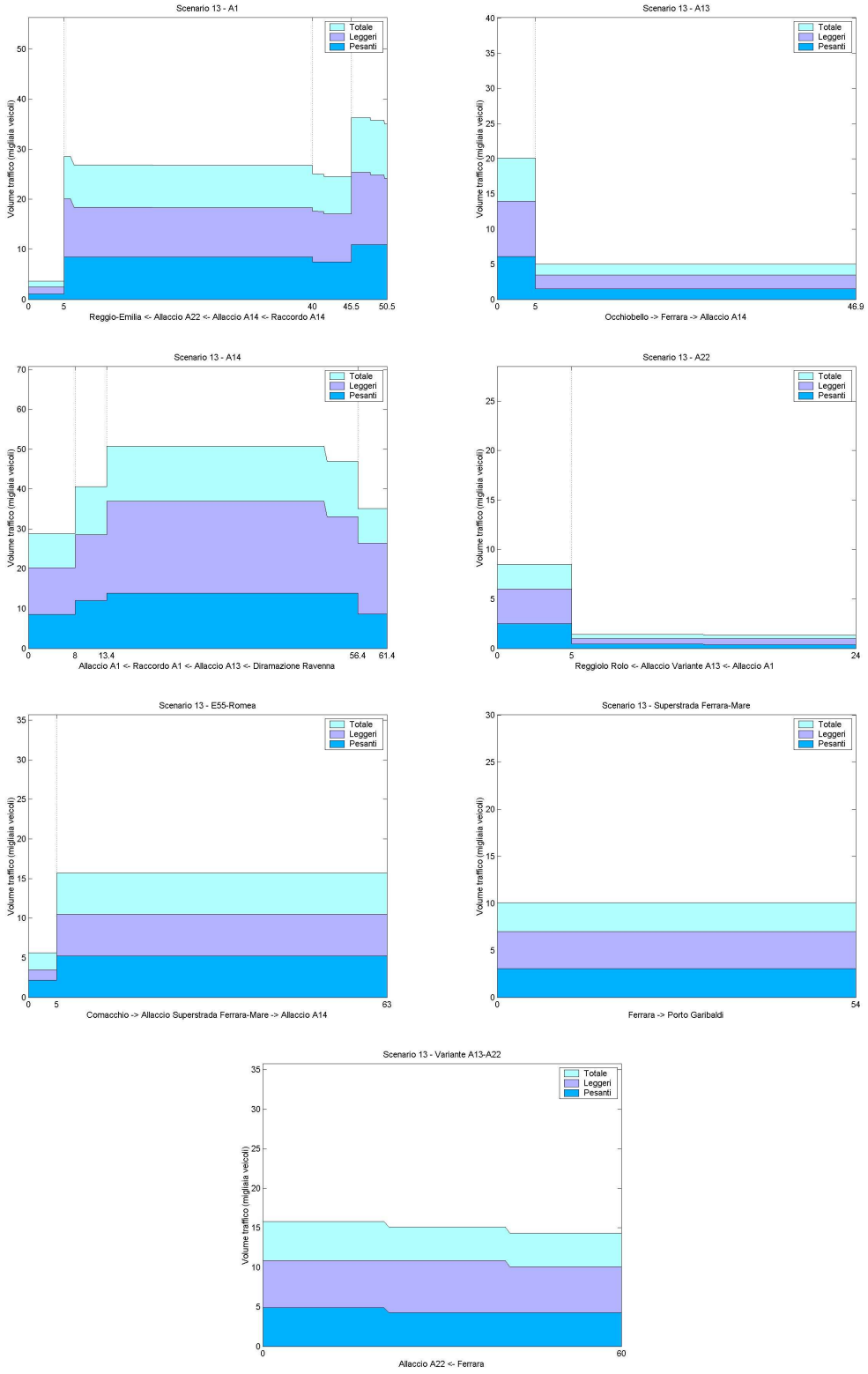
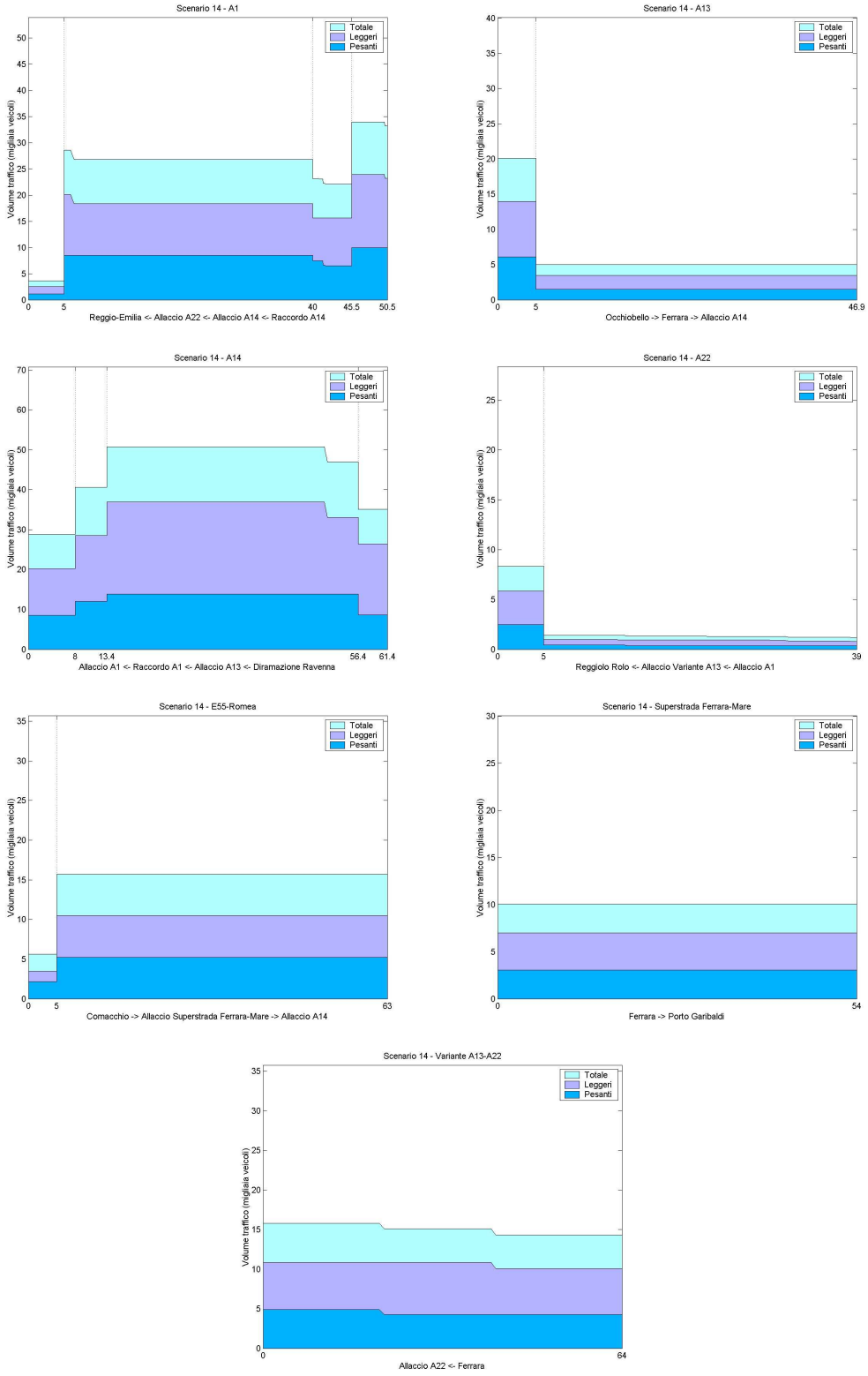


Figure B.14: Scenario 14



Appendix C

General traffic dataset used for simulations

C.1 Province data

The data for provinces (which we obtained in a more detailed form than that on regional scale) have been obtained through a series of measurements on particularly important road points of the Ferrara area, which have been carried through the period from 1992 and 2003.

From such data we are able to estimate various further parameters, such as days with the most pronounced traffic, the distribution of rush hours during the week and the corresponding ratios relative to mean traffic flows, The most critical days turn out being friday and thursday, while the worst hour seem to be that from 18 to 19, followed by the preceding hour.

The heavy transport percentage stays most of the time under 16(in particular at friday) decreasing to a range from 1.5high on a few roads).

C.2 Regions data

The datas for regions shown in the following have been selected according to relevance in the studies and simulations of an alternative highway route for Cispadana and E55, which means data for Ferrara, Ravenna, Modena and Reggio Emilia are included. We also grouped the Modena and Reggio Emilia provinces into a single area, as the Cispadana starts inbetween the two.

Maps are depicting data in a normalized way, so that dark green colour will always

be referring to the maximum vehicle quantity for each data set considered; therefore the different distribution of relationships between fluxes are shown, independently of global variations in the traffic volume which would make inspection less comfortable. Therefore any comparison of absolute volumes in different time slots must be done through the tables exclusively.

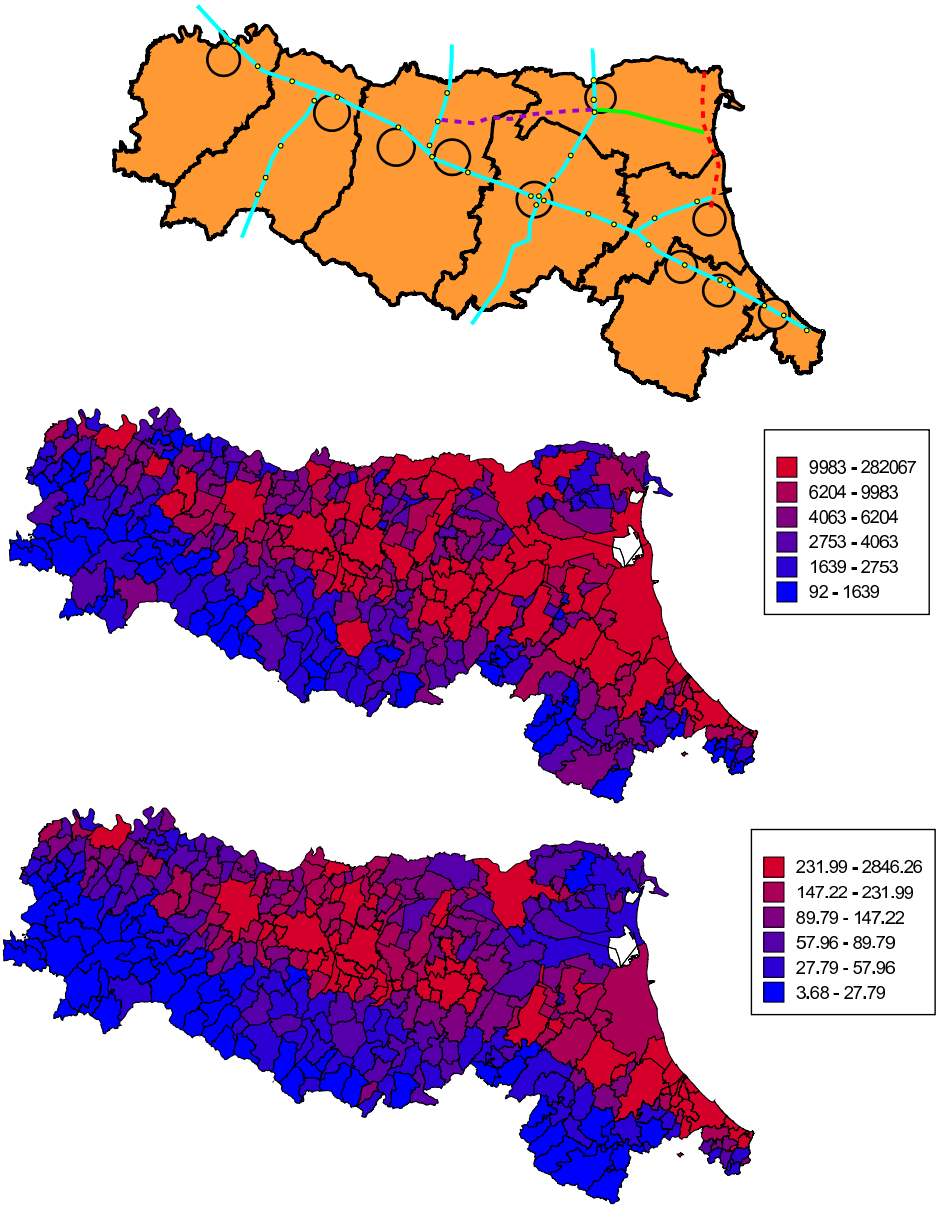


Table C.1: Movements in time zone 7-9

	to FE city	to FE prov.	from FE c.	from FE p.	to RA city	to RA prov.	from RA c.	from RA p.
Piacenza province	1	3	0	2	0	4	1	3
Piacenza city	0	0	0	1	0	1	0	2
Parma province	1	2	0	1	0	2	0	1
Parma city	0	0	1	2	0	1	1	2
Reggio province	2	11	1	1	0	3	1	1
Reggio city	0	0	2	0	0	1	2	0
Modena province	97	348	180	456	1	11	4	9
Modena city	5	10	7	8	0	2	0	0
Bologna province	236	415	700	1011	52	713	92	1001
Bologna city	587	237	161	435	21	188	22	153
Ferrara province	2508	7894	1970	7894	220	219	182	163
Ferrara city	5169	1970	5169	2508	0	40	9	79
Ravenna province	79	163	40	219	1198	7812	665	7812
Ravenna city	9	182	0	220	6812	665	6812	1198
Forlí province	0	1	0	0	62	227	75	287
Cesena	0	1	0	5	35	181	67	283
Forlí city	0	3	0	15	326	555	282	496
Rimini province	0	2	0	1	25	54	15	81
Rimini city	0	1	0	1	6	34	14	76
RE+MO prov.	99	359	181	457	1	14	5	10
	to MO city	to RE city	to prov.	from MO c.	from MO p.	from prov.		
Piacenza province	13	15	6	4	3	40		
Piacenza city	16	12	9	1	1	25		
Parma province	47	176	510	27	154	495		
Parma city	77	223	738	40	272	1011		
Reggio province	486	3903	13646	412	2309	13125		
Reggio city	330	5971	2720	205	5971	4188		
Modena province	4642	285	24536	3800	411	25057		
Modena city	8518	205	4212	8518	330	5128		
Bologna province	252	9	1083	229	13	1293		
Bologna city	234	42	437	160	34	421		
Ferrara province	8	0	457	10	0	359		
Ferrara city	7	2	181	5	0	99		
Ravenna province	0	0	10	2	1	14		
Ravenna city	0	2	5	0	0	1		
Forlí province	0	0	4	0	0	4		
Cesena	0	0	3	0	0	3		
Forlí city	1	0	12	1	0	12		
Rimini province	0	0	8	1	0	7		
Rimini city	0	2	5	1	0	5		
RE+MO prov.	5128	4188	38182	4212	2720	38182		

Table C.2: Movements in time zone 9-13

	to FE city	to FE prov.	from FE c.	from FE p.	to RA city	to RA prov.	from RA c.	from RA p.
Piacenza province	2	6	1	4	0	8	1	6
Piacenza city	0	0	0	3	0	2	0	3
Parma province	2	4	1	2	0	5	1	2
Parma city	0	0	1	3	0	1	1	3
Reggio province	3	20	2	1	0	6	2	1
Reggio city	0	0	4	1	0	1	4	1
Modena province	64	111	69	107	1	22	8	18
Modena city	7	3	7	0	0	5	1	0
Bologna province	174	179	184	179	16	277	23	276
Bologna city	164	151	155	146	2	69	2	57
Ferrara province	889	3170	905	3170	76	59	88	63
Ferrara city	2142	905	2142	889	4	21	5	26
Ravenna province	26	63	21	59	375	3368	417	3368
Ravenna city	5	88	4	76	2387	417	2387	375
Forlí province	0	1	0	0	27	79	30	78
Cesena	0	1	0	0	8	71	15	66
Forlí city	1	5	1	6	112	219	116	201
Rimini province	0	3	1	2	6	19	6	22
Rimini city	0	2	1	2	6	16	9	17
RE+MO prov.	67	131	71	108	1	28	10	19
	to MO city	to RE city	to provinces	from MO c.	from MO P.	from provinces		
Piacenza province	24	29	11	7	6	76		
Piacenza city	30	23	17	2	2	47		
Parma province	3	63	58	5	66	63		
Parma city	6	117	333	7	116	312		
Reggio province	119	1175	4629	131	1206	4646		
Reggio city	105	2481	1336	105	2481	1302		
Modena province	1466	127	8829	1513	130	8812		
Modena city	2461	105	1644	2461	105	1585		
Bologna province	68	17	308	72	25	367		
Bologna city	82	80	49	78	65	173		
Ferrara province	0	1	108	3	0	131		
Ferrara city	7	4	71	7	0	67		
Ravenna province	0	1	19	5	1	28		
Ravenna city	1	4	10	0	0	1		
Forlí province	0	0	7	0	0	7		
Cesena	1	0	5	1	0	4		
Forlí city	2	1	24	2	1	24		
Rimini province	0	1	15	2	1	13		
Rimini city	0	4	9	2	1	11		
RE+MO prov.	1585	1302	13458	1644	1336	13458		

Table C.3: Movements in time zone 13-16

	to FE city	to FE prov.	from FE c.	from FE p.	to RA city	to RA prov.	from RA c.	from RA p.
Piacenza province	2	4	0	3	0	5	1	4
Piacenza city	0	0	0	2	0	1	0	2
Parma province	1	3	0	1	0	3	0	1
Parma city	0	0	1	2	0	1	1	2
Reggio province	2	13	1	1	0	4	1	1
Reggio city	0	0	3	1	0	1	3	1
Modena province	55	139	70	140	1	14	5	12
Modena city	4	3	5	0	0	3	0	0
Bologna province	183	212	211	236	15	332	32	336
Bologna city	196	171	161	140	2	73	3	49
Ferrara province	1026	3632	1060	3632	75	64	115	63
Ferrara city	2506	1060	2506	1026	1	23	5	30
Ravenna province	30	63	23	64	390	3874	523	3874
Ravenna city	5	115	1	75	3019	523	3019	390
Forlí province	0	1	0	0	17	85	37	93
Cesena	0	1	0	0	9	89	19	77
Forlí city	0	3	0	3	127	261	149	215
Rimini province	0	2	0	2	3	23	6	23
Rimini city	0	1	0	1	1	19	8	20
RE+MO prov.	57	152	71	141	1	18	6	13
	to MO city	to RE city	to provinces	from MO c.	from MO P.	from provinces		
Piacenza province	16	19	7	5	4	50		
Piacenza city	20	15	11	1	1	30		
Parma province	0	47	81	6	67	88		
Parma city	10	137	400	14	124	333		
Reggio province	135	1386	5609	176	1475	5628		
Reggio city	121	2909	1619	124	2909	1517		
Modena province	1798	131	10713	1915	144	10694		
Modena city	4113	124	2091	4113	121	1933		
Bologna province	71	11	338	92	16	421		
Bologna city	93	52	204	85	43	164		
Ferrara province	0	1	141	3	0	152		
Ferrara city	5	3	71	4	0	57		
Ravenna province	0	1	13	3	1	18		
Ravenna city	0	3	6	0	0	1		
Forlí province	0	0	4	0	0	4		
Cesena	0	0	3	1	0	3		
Forlí city	1	0	15	1	0	16		
Rimini province	0	1	10	2	0	9		
Rimini city	0	3	5	2	0	7		
RE+MO prov.	1933	1517	16322	2091	1619	16322		

Table C.4: Movements in time zone 16-18

	to FE city	to FE prov.	from FE c.	from FE p.	to RA city	to RA prov.	from RA c.	from RA p.
Piacenza province	2	4	1	3	0	6	1	4
Piacenza city	0	0	0	2	0	1	0	2
Parma province	1	3	0	1	0	3	0	1
Parma city	0	0	1	2	0	1	1	2
Reggio province	2	15	1	1	0	4	1	1
Reggio city	0	0	3	1	0	1	3	1
Modena province	106	209	79	175	1	16	6	13
Modena city	7	0	7	2	0	3	0	0
Bologna province	361	407	193	227	39	527	24	396
Bologna city	160	262	324	176	4	91	5	96
Ferrara province	1284	4910	1481	4910	108	80	131	104
Ferrara city	3289	1481	3289	1284	6	44	5	33
Ravenna province	33	104	44	80	503	5189	694	5189
Ravenna city	5	131	6	108	3908	694	3908	503
Forlí province	0	1	0	0	41	131	38	111
Cesena	0	1	0	0	29	130	17	95
Forlí city	0	4	1	11	172	314	187	333
Rimini province	0	2	0	2	7	39	9	28
Rimini city	0	1	0	1	9	34	5	16
RE+MO prov.	108	224	80	176	1	20	7	14
	to MO city	to RE city	to provinces	from MO c.	from MO P.	from provinces		
Piacenza province	17	20	8	5	5	55		
Piacenza city	22	17	13	1	1	34		
Parma province	5	89	130	6	89	127		
Parma city	10	183	553	29	165	452		
Reggio province	199	1606	7539	224	2167	7733		
Reggio city	139	3821	2335	189	3821	1824		
Modena province	2278	218	14508	2574	168	14314		
Modena city	5419	189	2798	5419	139	2477		
Bologna province	111	12	552	124	188	531		
Bologna city	115	57	271	136	47	259		
Ferrara province	2	1	176	0	0	224		
Ferrara city	7	3	80	7	0	108		
Ravenna province	0	1	14	3	1	20		
Ravenna city	0	3	7	0	0	1		
Forlí province	0	0	4	0	0	5		
Cesena	1	0	4	1	0	3		
Forlí city	1	0	17	1	0	17		
Rimini province	0	1	11	2	1	9		
Rimini city	0	3	6	2	1	8		
RE+MO prov.	2477	1824	22047	2798	2335	22047		

Table C.5: Movements in time zone 18-22

	to FE city	to FE prov.	from FE c.	from FE p.	to RA city	to RA prov.	from RA c.	from RA p.
Piacenza province	2	5	1	3	0	7	1	5
Piacenza city	0	0	0	2	0	1	0	3
Parma province	1	3	0	1	0	4	0	1
Parma city	0	0	1	3	0	1	1	3
Reggio province	2	17	1	1	0	5	1	1
Reggio city	0	0	3	1	0	1	3	1
Modena province	130	255	111	217	1	18	7	15
Modena city	9	7	10	5	0	4	0	0
Bologna province	426	501	265	300	44	648	48	494
Bologna city	216	318	383	240	5	120	8	114
Ferrara province	1561	5950	1825	5950	130	100	171	127
Ferrara city	4055	1825	4055	1561	6	55	6	41
Ravenna province	41	127	55	100	615	6291	863	6291
Ravenna city	6	171	6	130	4819	863	4819	615
Forlí province	0	1	0	0	49	168	57	147
Cesena	0	1	0	0	31	162	26	118
Forlí city	0	3	1	13	205	413	235	403
Rimini province	0	3	0	2	9	43	13	34
Rimini city	0	1	0	2	9	42	10	25
RE+MO prov.	132	272	112	218	1	23	8	16
	to MO city	to RE city	to prov.	from MO c.	from MO p.	from prov.		
Piacenza province	20	24	9	6	5	64		
Piacenza city	25	19	14	2	1	39		
Parma province	7	113	200	14	130	205		
Parma city	19	227	714	35	202	563		
Reggio province	256	1983	9336	298	2686	9581		
Reggio city	178	4744	2916	228	4744	2252		
Modena province	2786	269	17888	3206	230	17643		
Modena city	6748	228	3504	6748	178	3042		
Bologna province	138	14	700	174	21	704		
Bologna city	145	67	352	165	54	320		
Ferrara province	5	1	218	7	0	272		
Ferrara city	10	3	112	9	0	132		
Ravenna province	0	1	16	4	1	23		
Ravenna city	0	3	8	0	0	1		
Forlí province	0	0	5	0	0	5		
Cesena	1	0	4	1	0	4		
Forlí city	2	0	20	2	1	20		
Rimini province	0	1	13	2	1	12		
Rimini city	0	3	7	2	1	9		
RE+MO prov.	3042	2252	27224	3504	2916	27224		

Table C.6: Movements in time zone 22-7

	to FE city	to FE prov.	from FE c.	from FE p.	to RA city	to RA prov.	from RA c.	from RA p.
Piacenza province	1	3	0	2	0	5	1	3
Piacenza city	0	0	0	1	0	1	0	2
Parma province	1	2	0	1	0	3	0	1
Parma city	0	0	1	2	0	1	1	2
Reggio province	2	12	1	1	0	4	1	1
Reggio city	0	0	2	0	0	1	2	1
Modena province	30	75	34	83	1	13	5	10
Modena city	0	0	2	0	0	3	0	0
Bologna province	106	114	109	134	15	190	6	201
Bologna city	101	84	116	98	1	25	2	38
Ferrara province	672	2181	606	2181	73	36	31	25
Ferrara city	1613	606	1613	672	2	16	1	10
Ravenna province	10	25	16	36	359	2357	201	2357
Ravenna city	1	31	2	73	1927	201	1927	359
Forlí province	0	1	0	0	23	47	3	45
Cesena	0	1	0	0	8	41	7	55
Forlí city	0	3	1	4	99	108	70	154
Rimini province	0	2	0	1	4	11	0	14
Rimini city	0	1	0	1	2	11	0	10
RE+MO prov.	32	87	35	84	1	17	6	11
	to MO city	to RE city	to prov.	from MO c.	from MO p.	from prov.		
Piacenza province	14	16	6	4	4	44		
Piacenza city	18	14	10	1	1	27		
Parma province	1	32	47	0	20	37		
Parma city	10	77	152	3	92	235		
Reggio province	93	982	3408	65	824	3364		
Reggio city	79	1916	906	70	1916	1067		
Modena province	1267	85	6811	1099	82	6855		
Modena city	2718	70	1164	2718	79	1360		
Bologna province	44	10	221	33	14	190		
Bologna city	47	46	92	48	38	95		
Ferrara province	0	0	84	0	0	87		
Ferrara city	2	2	35	0	0	32		
Ravenna province	0	1	11	3	1	17		
Ravenna city	0	2	6	0	0	1		
Forlí province	0	0	4	0	0	4		
Cesena	0	0	3	1	0	3		
Forlí city	1	0	13	1	0	13		
Rimini province	0	0	9	1	0	8		
Rimini city	0	2	5	1	0	6		
RE+MO prov.	1360	1067	10219	1164	906	10219		

Appendix D

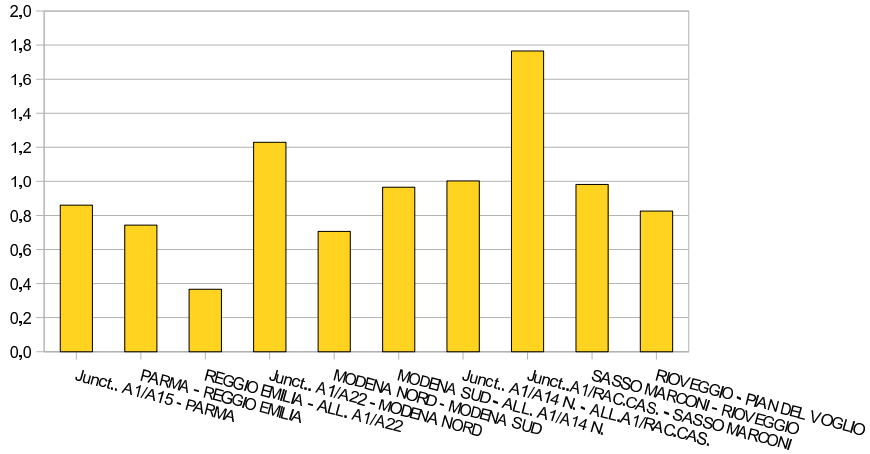
Highway dataset used for simulation

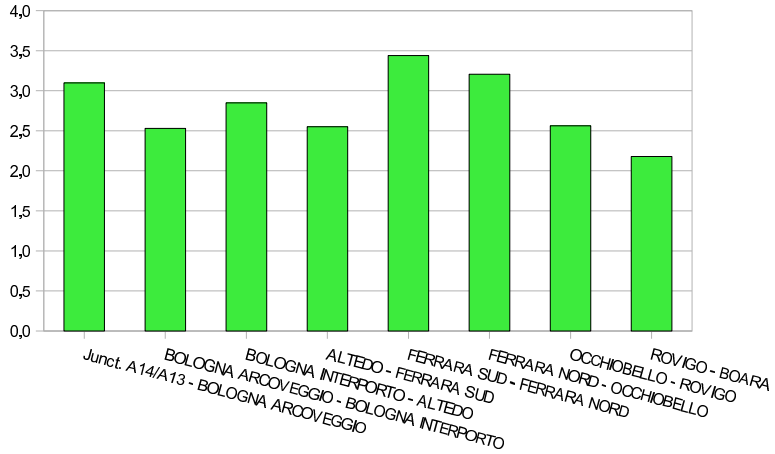
Societa' Autostrade per l'Italia provided the daily mean fluxes for the areas touched by the simulation, with the exception of A22 Modena Brennero. The whole data relative to A22 has been provided by Aiscat. In the graph depicted below the base scenario covering the interested region is shown. All values refer to each stretch in-between two nearby entrances or crossings, and have been calculated over the yearly data for year 2003.
These informations potentially allow for extrapolations over the following years by exploiting both the earlier historical trend as well as criterias for logistic change, encompassing growth and saturation.
Moreover the usage of further aggregated data regarding the hourly fluxes - shown later on - can help obtaining an approximation of mean hourly values on each particular stretch.

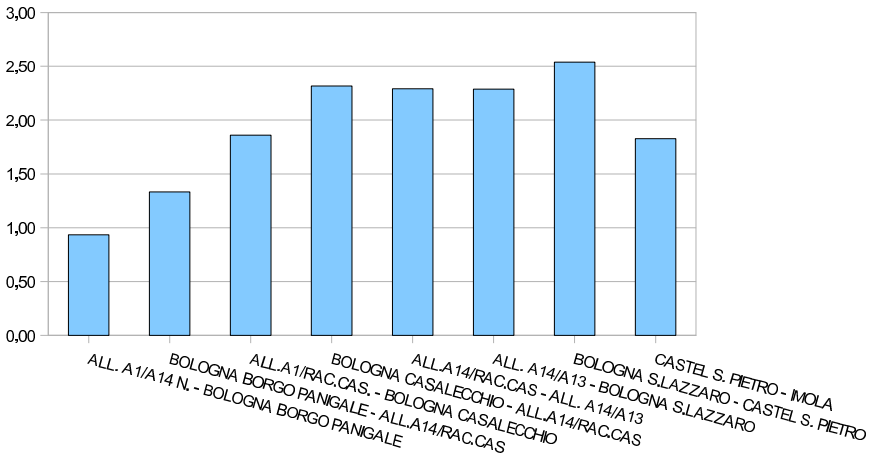
	Light	Heavy	Total
A1 - Milano Napoli			
All. A1/A15 Parma	27.160	11.268	38.427
Parma - Reggio Emilia	25.397	10.670	36.067
Reggio Emilia All. A1/A22	26.866	11.053	37.919
All. A1/A22 - Modena Nord	36.688	15.205	51.893
Modena Nord - Modena Sud	33.037	13.659	46.697
Modena Sud - All. A1/A14 N.	36.017	14.133	50.150
All. A1/A14 N. - All. A1/Rac.Cas.	11.285	4.876	16.161
All. A1/Rac.Cas. - Sasso Marconi	21.043	8.723	29.766
Sasso Marconi Rioveggio	17.472	7.972	25.444
A13 - Bologna Padova			
All. A14/A13 - Bologna Arcoveggio	14.569	5.697	20.266
Bologna Arcoveggio Bologna Interporto	18.259	6.457	24.716
Bologna Interporto Altedo	17.407	5.913	23.320
Altedo - Ferrara Sud	16.664	5.816	22.480
Ferrara Sud - Ferrara Nord	13.726	6.004	19.730
Ferrara Nord Occhiobello	14.118	6.029	20.147
Occhiobello Rovigo	12.123	4.929	17.052
Rovigo Boara	12.131	5.027	17.158
A 14 - Bologna Taranto			
All. A1/A14 N. - Bologna Borgo Panigale	24.777	9.273	34.050
Bologna Borgo Panigale All. A14/Rac.Cas.	18.642	7.874	26.516
All. A1/Rac.Cas. Bologna Casalecchio	14.128	4.453	18.581
Bologna Casalecchio All. A14/Rac.Cas.	8.657	3.233	11.890
All. A14/Rac.Cas. - All. A14/A13	27.300	11.107	38.406
All. A14/A13 - Bologna S.Lazzaro	22.907	9.192	32.100
Bologna S.Lazzaro - Castel S.Pietro	33.927	10.786	44.712
Castel S.Pietro - Imola	32.826	10.647	43.473
Imola - All. A14/Diram. Ravenna	32.055	10.166	42.222
All. A14/Diram. Ravenna - Lugo Cotignola	6.458	1.731	8.189
Lugo Cotignola - Ravenna	6.207	1.501	7.707
All. A14/Diram. Ravenna - Faenza	26.537	8.751	35.287

	Light	Heavy	Total
A1 - Milano Napoli			
All. A1/A15 Parma	26.936	10.932	37.869
Parma - Reggio Emilia	25.171	10.419	35.590
Reggio Emilia All. A1/A22	26.849	10.841	37.690
All. A1/A22 - Modena Nord	36.909	14.881	51.790
Modena Nord - Modena Sud	33.574	13.711	47.285
Modena Sud - All. A1/A14 N.	36.605	14.250	50.855
All. A1/A14 N. - All. A1/Rac.Cas.	11.459	4.641	16.101
All. A1/Rac.Cas. - Sasso Marconi	21.172	8.388	29.561
Sasso Marconi Rioveggio	17.442	7.697	25.139
A13 - Bologna Padova			
All. A14/A13 - Bologna Arcoveggio	13.251	5.708	18.959
Bologna Arcoveggio Bologna Interporto	18.405	6.574	24.979
Bologna Interporto Altedo	17.406	6.010	23.415
Altedo - Ferrara Sud	16.765	5.957	22.722
Ferrara Sud - Ferrara Nord	13.641	6.150	19.791
Ferrara Nord Occhiobello	13.987	6.131	20.118
Occhiobello Rovigo	11.968	4.839	16.807
Rovigo Boara	12.232	5.015	17.247
A 14 - Bologna Taranto			
All. A1/A14 N. - Bologna Borgo Panigale	25.191	9.624	34.815
Bologna Borgo Panigale All. A14/Rac.Cas.	19.045	8.194	27.239
All. A1/Rac.Cas. Bologna Casalecchio	14.173	4.553	18.726
Bologna Casalecchio All. A14/Rac.Cas.	7.729	3.101	10.830
All. A14/Rac.Cas - All. A14/A13	26.774	11.295	38.069
All. A14/A13 - Bologna S.Lazzaro	23.700	9.369	33.070
Bologna S.Lazzaro - Castel S.Pietro	33.295	10.810	44.105
Castel S.Pietro - Imola	32.678	10.762	43.441
Imola - All. A14/Diram. Ravenna	31.900	10.336	42.236
All. A14/Diram. Ravenna - Lugo Cotignola	6.387	1.985	8.372
Lugo Cotignola - Ravenna	6.370	1.806	8.176
All. A14/Diram. Ravenna - Faenza	26.452	8.667	35.119

D.1 Increments 2002-2003







D.2 Modena-Brennero

For the A22 Modena-Brennero highway (in particular the initial Verona-Modena stretch) we have used the aggregate values provided by Aiscat, which are available every three months and here shown merged over the year 2003.

	Daily vehicles mean		Theorical daily vehicles mean		Vehicles-Km in millions		
	2003	2002	2003	2002	2003	2002	Percent variation
Light	49789	47320	28038	26426	921.1	868.1	6,1
Heavy	21595	20694	12369	11694	406.4	384.2	5,78
Total	71384	68014	40407	38120	1327.4	1252.3	6

Entrance and exit fluxes at the A22 toll gates of Pegognaga, Mantova Sud and Mantova Nord have been extrapolated through environmental statistics from the Mantova province over years 1991-1999 by a linear growth model.
Values at toll gates over the lower stretch (Modena Campogalliano, Carpi and Reggiolo-Rolo, from south to north) can be estimated through the statistical informations for the surrounding areas, which has similar characteristics to that of Pegognaga.

D.3 Traffic fluxes at toll gates

Gate	Label	Light in	Heavy in	Light out	Heavy out
Modena North	AB1	10011	4000	10326	4375
Modena South	AB2	6790	2167	6842	2231
Sasso Marconi	2C1	4933	963	4773	1023
Bologna B. Panigale	BD1	8207	2281	8195	2250
Bologna Casalecchio	CD1	12152	2581	11178	2348
Bologna S. Lazzaro	EF1	15883	2640	14458	2488
Bologna Arcoveggio	EH1	7753	1425	9217	1530
Bologna Interporto	EH2	4194	1969	4046	1949
Altedo	EH3	2465	632	2568	677
Ferrara South	H	7391	1877	7205	1882
Ferrara North	5H1	3243	1134	3197	1090
Castel S. Pietro	EF2	3949	1114	4433	1204
Imola	EF3	7057	2141	7049	2196
Lugo Cotignola	FG1	2497	602	2730	653
Ravenna	FG2	6376	1808	6212	1503
Pegognaga	6A4	1828	729	1840	734
Mantova South	6A5	2011	749	1922	715
Mantova North	6A6	4935	2059	4893	2042

D.4 Hourly graphs

Here we show some data pertaining the evolution of highway traffic intensity at different times of the day; from the graph the different behaviours of heavy and light traffic are clearly visible. Values refer to the stretches in the Bologna Area as identified in the feasibility study ”Riorganizzazione del sistema autostradale - tangenziale del nodo di Bologna”.

Time	Light	Heavy	Total	Equivalent vehicles
0	1,73%	2,29%	1,86%	1,97%
1	1,19%	2,06%	1,39%	1,57%
2	0,80%	2,08%	1,10%	1,35%
3	0,61%	2,12%	0,96%	1,26%
4	0,61%	3,03%	1,18%	1,66%
5	1,12%	4,13%	1,82%	2,42%
6	2,29%	5,71%	3,09%	3,76%
7	5,64%	5,35%	5,57%	5,52%
8	7,28%	4,62%	6,66%	6,14%
9	6,28%	6%	6,22%	6,16%
10	5,29%	7,10%	5,72%	6,07%
11	5,08%	6,73%	5,46%	5,79%
12	4,99%	5,46%	5,10%	5,19%
13	4,93%	5,31%	5,02%	5,09%
14	5,33%	5,86%	5,45%	5,56%
15	5,80%	5,54%	5,74%	5,69%
16	6,07%	5,59%	5,96%	5,86%
17	7%	4,81%	6,49%	6,06%
18	7,57%	3,75%	6,68%	5,92%
19	6,96%	2,85%	6,01%	5,19%
20	5,12%	2,46%	4,50%	3,97%
21	3,45%	2,30%	3,18%	2,95%
22	2,60%	2,57%	2,59%	2,59%
23	2,26%	2,29%	2,27%	2,28%

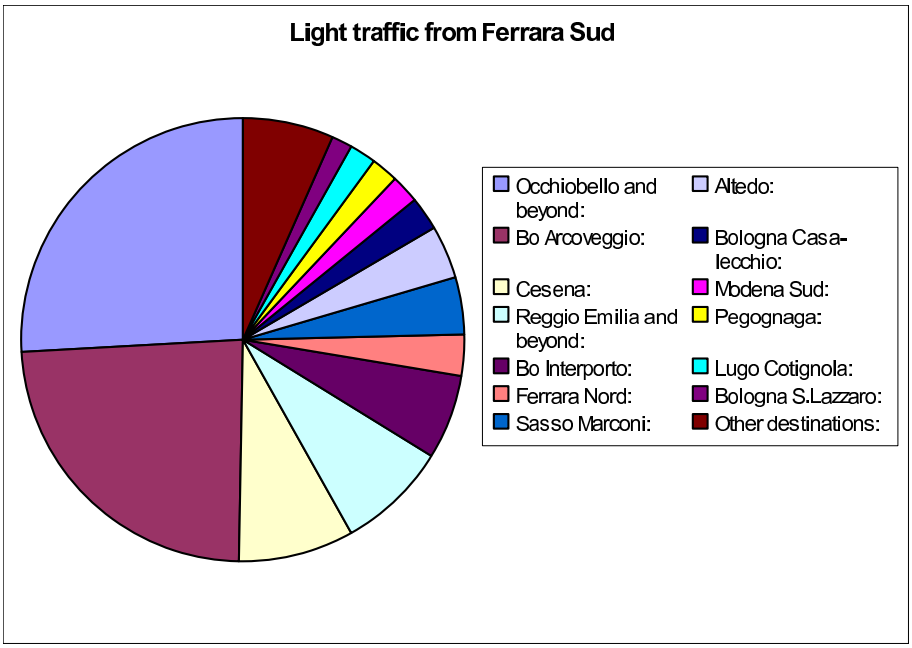
D.5 Junction points analysis

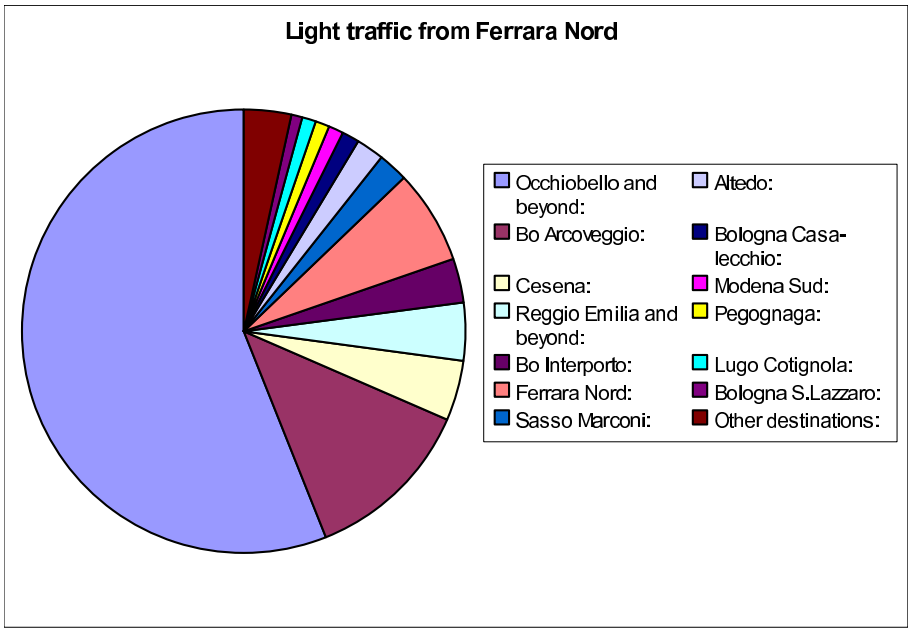
The data for highway fluxes that were given with respect to each stretch can be reorganized according to junctions (in the table ”I” refers to entrance fluxes, ”O” to exiting fluxes). Then with these one can obtain the turning coefficients for each direction pair, that is the percentages in which traffic from each source direction splits.

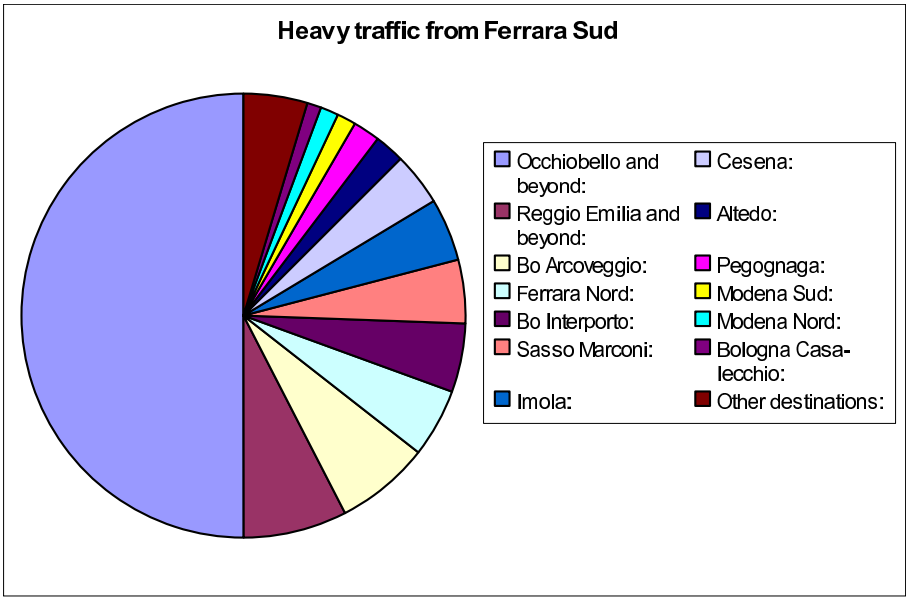
	I - Light	I - Heavy	I - Total	O - Light	O - Heavy	O - Total
A - A1/A22						
Reggio Emilia	26.866	11.053	37.919	26.849	10.841	37.690
Mantova	13.900	6.240	20.140	14.138	6.128	20.266
Modena Nord	36.909	14.881	51.790	36.688	15.205	51.893
	77.675	32.174	109.849	77.675	32.174	109.849
B - A1-A14						
Modena Sud	36.017	14.133	50.150	36.605	14.250	50.855
A14 - Borgo Panigale	25.191	9.624	34.815	24.777	9.273	34.050
Raccordo Casalecchio	11.459	4.641	16.100	11.285	4.876	16.161
	72.667	28.398	101.065	72.667	28.399	101.066
C - A1-Raccordo Casalecchio						
Modena Sud	11.285	4.876	16.161	11.459	4.641	16.100
Bologna Casalecchio	14.173	4.553	18.726	14.128	4.453	18.581
Sasso Marconi	21.172	8.388	29.560	21.043	8.723	29.766
	46.630	17.817	64.447	46.630	17.817	64.447
D - A14-Raccordo Casalecchio						
Bologna Borgo Panigale	18.642	7.874	26.516	19.045	8.194	27.239
Bologna Casalecchio	8.657	3.233	11.890	7.729	3.101	10.830
Raccordo A13-A14	26.774	11.295	38.069	27.300	11.107	38.407
	54.073	22.402	76.475	54.074	22.402	76.476
E - A13 - A14						
Bologna Arcoveggio	13.251	5.708	18.959	14.569	5.697	20.266
BO Raccordo Casal.	27.300	11.107	38.406	26.774	11.295	38.069
Bologna S.Lazzaro	23.700	9.369	33.069	22.907	9.192	32.100
	64.251	26.184	90.434	64.250	26.184	90.435
F - A14-Ravenna branching						
Imola	32.055	10.166	42.221	31.900	10.336	42.236
Ravenna	6.387	1.985	8.372	6.458	1.731	8.189
Faenza	26.452	8.667	35.119	26.537	8.751	35.288
	64.894	20.818	85.712	64.895	20.818	85.713

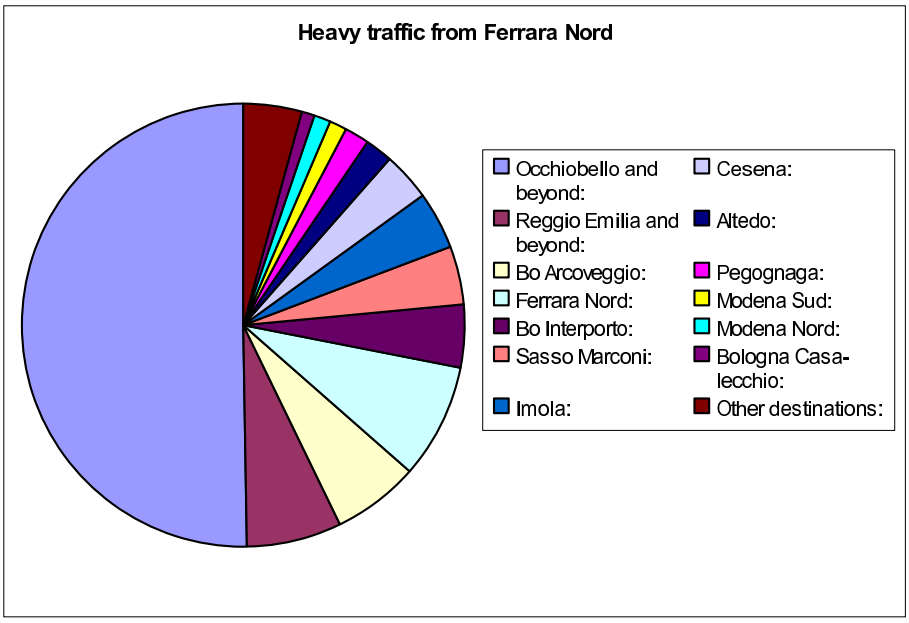
D.6 Distribution of traffic fluxes around Ferrara

Through the coefficients describind turning shares at each intersection (given in Appendix A) it is possible to obtain statistical data for actual composite paths, thus combining on the way effects of the different junctions passed by. This technique allows to obtain clues on the absolute and relative importances of each destination, and here the analysis is performed considering paths starting at Ferrara and ending in the highway network considered.
Since the available data amounts to daily averages the number of vehicle entrances and exits balance each other, so with the same coefficients one can obtain percentages showing not only importance of path starting at Ferrara, but also of the corresponding opposite paths with Ferrara as final destination.
In all calculations the mean values for coefficient intervals were used as representative.


Light traffic from Ferrara Sud
Occhiobello and beyond:
Altedo:
Bo Arcoveggio:
Bologna Casa-lecchio:
Cesena:
Modena Sud:
Reggio Emilia and beyond:
Pegognaga:
Bo Interporto:
Lugo Cotignola:
Ferrara Nord:
Bologna S.Lazzaro:
Sasso Marconi:
Other destinations:


Light traffic from Ferrara Nord
Occhiobello and beyond:
Altedo:
Bo Arcoveggio:
Bologna Casa-lecchio:
Cesena:
Modena Sud:
Reggio Emilia and beyond:
Pegognaga:
Bo Interporto:
Lugo Cotignola:
Ferrara Nord:
Bologna S.Lazzaro:
Sasso Marconi:
Other destinations:


Heavy traffic from Ferrara Sud
Occhiobello and beyond:
Cesena:
Reggio Emilia and beyond:
Altedo:
Bo Arcoveggio:
Pegognaga:
Ferrara Nord:
Modena Sud:
Bo Interporto:
Modena Nord:
Sasso Marconi:
Bologna Casa-lecchio:
Imola:
Other destinations:


Heavy traffic from Ferrara Nord
Occhiobello and beyond:
Cesena:
Reggio Emilia and beyond:
Altedo:
Bo Arcoveggio:
Pegognaga:
Ferrara Nord:
Modena Sud:
Bo Interporto:
Modena Nord:
Sasso Marconi:
Bologna Casa-lecchio:
Imola:
Other destinations:

D.7 Traffic over the Romea/E55 attraction area

Within a study of Romea and E55 carried out by Regione Emilia Romagna - in 2001 traffic measurements were performed over the affected areas. In the following we list all studied sites and the corresponding data, here again these are daily volumes.
The ”right” direction refers to that given in the first table, and clearly the ”left” one refers to the opposite. The numbered section points can be found on the map. Such data, together with those previously shown, allow for an estimation of possible traffic shifts towards the new highway stretches referred by simulation scenarios.

Measurement point	Road	Position
1	S.S.495 di Codigoro between Portomaggiore and Ostellato	Km 10
2	S.S.16 between Argenta and Alfonsine	Km 129.5
3	S.S.16 between Alfonsine and Ravenna	Km 145
4	S.S.309 Romea between Ravenna and Comacchio	Km 12.5
5	S.S.16 adriatica between Polesella and Rovigo	Km 50.5
6	S.S. 495 between Ariano and Adria	Km 64.6
7	S.S.309 Romea between Mesola and Taglio di Po	Km 55.2
8	S.S.16 adriatica between Rovigo and Stanghella	Km 35.8
9	S.S.516 between Adria and Cavarzere	Km38.8
10	S.S.309 Romea between Rosolina and S.Anna	Km 74.3
11	S.S.309 between Chioggia and Venezia	Km 94.2
12	S.P. between Comacchio and Ostellato	
13	Autostrada dei lidi Ferraresi between Comacchio and Ostellato	
14	S.P. between Tresigallo and Massa Fiscaglia	
15	S.P. between Iolanda di Savoia and S.S.309 Romea	
16	S.P. between Crespino and Villanova	
17	S.S.443 between Rovigo and Adria	
18	S.P. between Agna and Cona	
19	S.P. between Arre and Candiana	
20	S.S.516 between Piove di Sacco and Padova	
21	S.S.11 between Dolo and Mira	

Measurement point	A - Light	A - Heavy	B - Light	B - Heavy
1	2020	191	1814	150
2	5443	844	5594	819
3	7074	968	7374	820
4	3611	2327	4140	2237
5	8385	1051	8023	918
6	2265	347	2262	305
7	3263	2443	3251	1907
8	6516	1012	6639	835
9	4110	529	3990	471
10	6029	2631	6148	2132
11	7049	3007	6642	2912
12	2042	101	2072	72
13	2675	474	2813	457
14	2177	95	2278	92
15	1733	272	1581	256
16	1950	314	1870	331
17	3564	366	3553	397
18	1178	75	1178	73
19	2093	165	1940	154
20	8602	934	9229	1128
21	9962	666	9665	601

Appendix E

Turning coefficient

From inflow and outflow fluxes at a crossing it's possible to determine how traffic coming from each direction partitions. Turning fractions, even in the case of simple ternary junction, are not uniquely determined but there is an interval of possibilities; for these we'll give in the following both the extremes and the pair mean/width.
Each coefficient value - named α_{ij} - corresponds to the share of vehicles coming from direction "i" and turning towards direction "j", so that all combinations with "i" different from "j" are shown.

Calculations were performed both at actual junctions and at toll gates, which were renamed like this:

- AB1: Modena Nord
- AB2: Modena Sud
- BD1: Bologna Borgo Panigale
- 2C1: Sasso Marconi
- CD1: Bologna Casalecchio
- EF1: Bologna S. Lazzaro
- EH1: Bologna Arcoveggio
- EH2: Bologna Interporto
- EH3: Altedo

- H: Ferrara Sud
- 5H1: Ferrara Nord
- EF2: Castel S. Pietro
- EF3: Imola
- FG1: Lugo Cotignola
- 6A4: Pegognaga
- 6A5: Mantova Sud
- 6A6: Mantova Nord

The three directions corresponding to the α indices refer to those two of the stretch determined by the first two letters plus the third at the entrance of the toll gate itself. Letters point to the nodes in the first table, and furthermore:

- 1: A1 direction Milan
- 2: A1 direction Florence
- 3: A13 direction Taranto
- 4: SS. 309 direction Venice
- 5: A13 direction Padova
- 6: A44 direction Brennero

Table E.1: Turn coefficients for light traffic

A	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.1518	1	0.72744	5.9684E-06	0.27256	0.84821
2. Extr	3E-07	0.70662	0.61695	0.29338	0.38305	1
mean	0.0759	0.85331	0.672195	0.14669298	0.327805	0.924105
width/2	0.0759	0.14669	0.055245	0.14668702	0.055245	0.075895
B	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.68793	0.00178	1	0.99821	0	0.31207
2. Extr	0.68668	0	0.99607	1	0.003927	0.31332
mean	0.687305	0.00089	0.998035	0.999105	0.0019635	0.312695
width/2	0.000625	0.00089	0.001965	0.000895	0.0019635	0.000625
C	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.3912	1	0.54123	0	0.45877	0.60877
2. Extr	0	0.68849	0.3327	0.31151	0.6673	1
mean	0.1956	0.84425	0.436965	0.155755	0.563035	0.804385
width/2	0.1956	0.15576	0.104265	0.155755	0.104265	0.195615
D	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0	1	0.71132	0	0.28868	1
2. Extr	0	1	0.71132	0	0.28868	1
mean	0	1	0.71132	0	0.28868	1
width/2	0	0	0	0	0	0
E	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.2321	0.46634	0	0.53366	1	0.76794
2. Extr	1	0.83908	0.42939	0.16092	0.57061	0
mean	0.616	0.65271	0.214695	0.34729	0.785305	0.38397
width/2	0.384	0.18637	0.214695	0.18637	0.214695	0.38397
F	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.2015	0.14717	1	0.85283	0	0.79853
2. Extr	0.1721	0.13982	0.99822	0.86018	0.0017767	0.82786
mean	0.1868	0.1435	0.99911	0.856505	0.00088835	0.813195
width/2	0.0147	0.00367	0.00089	0.003675	0.00088835	0.014665

Table E.2: This table refers to heavy traffic

A	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.1889	0.99999	0.72851	6.8414E-06	0.27149	0.8111
2. Extr	3.799E-07	0.66539	0.5882	0.33461	0.4118	1
mean	0.09445019	0.83269	0.658355	0.16730842	0.341645	0.90555
width/2	0.09444981	0.1673	0.070155	0.16730158	0.070155	0.09445
B	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.65502	5.4736E-06	0.99678	1	0.0032207	0.34498
2. Extr	0.65608	0.0015622	1	0.99844	0	0.34392
mean	0.65555	0.00078384	0.99839	0.99922	0.00161035	0.34445
width/2	0.00053	0.00077836	0.00161	0.00078	0.00161035	0.00053
C	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0	0.84492	0.46912	0.15507	0.53088	1
2. Extr	0.14479	1	0.55329	0	0.44671	0.85521
mean	0.072395	0.92246	0.511205	0.077535	0.488795	0.927605
width/2	0.072395	0.07754	0.042085	0.077535	0.042085	0.072395
D	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0	1	0.72545	0	0.27455	1
2. Extr	0	1	0.72545	0	0.27455	1
mean	0	1	0.72545	0	0.27455	1
width/2	0	0	0	0	0	0
E	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.33182	0.4856	0	0.5144	1	0.66818
2. Extr	1	0.82998	0.40709	0.17002	0.59291	0
mean	0.66591	0.65779	0.203545	0.34221	0.796455	0.33409
width/2	0.33409	0.17219	0.203545	0.17219	0.203545	0.33409
F	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.17027	0.15918	1	0.84082	0	0.82973
2. Extr	0.13919	0	0.96354	1	0.036459	0.86081
mean	0.15473	0.07959	0.98177	0.92041	0.0182295	0.84527
width/2	0.01554	0.07959	0.01823	0.07959	0.0182295	0.01554

Table E.3: Turning coefficients α_{ij} at toll gates for light traffic

AB1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.7186	0	0.33325	1	0.66675	0.28145
2. Extr	0.9	0.19882	1	0.80118	0	0.09951
mean	0.8093	0.09941	0.666625	0.90059	0.333375	0.19048
width/2	0.0907	0.09941	0.333375	0.09941	0.333375	0.09097
AB2	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	1	0.18691	0.56112	0.81309	0.43888	0
2. Extr	0.8847	0.08283	0	0.91717	1	0.11532
mean	0.9423	0.13487	0.28056	0.86513	0.71944	0.05766
width/2	0.0577	0.05204	0.28056	0.05204	0.28056	0.05766
2C1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.7732	0	0.75634	1	0.24366	0.22682
2. Extr	0.8303	0.06891	1	0.93109	0	0.1697
mean	0.8017	0.03446	0.87817	0.965545	0.12183	0.19826
width/2	0.0286	0.03446	0.12183	0.034455	0.12183	0.02856
BD1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.6693	0	0.749	1	0.251	0.33075
2. Extr	0.7524	0.10816	1	0.89184	0	0.24761
mean	0.7108	0.05408	0.8745	0.94592	0.1255	0.28918
width/2	0.0416	0.05408	0.1255	0.05408	0.1255	0.04157
CD1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.2088	0	0.53037	1	0.46963	0.38725
2. Extr	0.6128	0.73838	1	0.26162	0	0.79119
mean	0.4108	0.36919	0.765185	0.63081	0.234815	0.58922
width/2	0.202	0.36919	0.234815	0.36919	0.234815	0.20197
EF1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	1	0.43424	0.30618	0.56576	0.69382	0
2. Extr	0.7877	0.28818	0	0.71182	1	0.2123
mean	0.8939	0.36121	0.15309	0.63879	0.84691	0.10615
width/2	0.1062	0.07303	0.15309	0.07303	0.15309	0.10615

EH1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	1	0.50079	0.52406	0.49921	0.47594	0
2. Extr	0.7211	0.28003	0	0.71997	1	0.27888
mean	0.8606	0.39041	0.26203	0.60959	0.73797	0.13944
width/2	0.1394	0.11038	0.26203	0.11038	0.26203	0.13944
EH2	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.7784	0	0.23843	1	0.76157	0.22159
2. Extr	0.9533	0.1835	1	0.8165	0	0.04666
mean	0.8659	0.09175	0.619215	0.90825	0.380785	0.13413
width/2	0.0875	0.09175	0.380785	0.09175	0.380785	0.08747
EH3	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.8525	0	0.26003	1	0.73997	0.14752
2. Extr	0.9573	0.1088	1	0.8912	0	0.04274
mean	0.9049	0.0544	0.630015	0.9456	0.369985	0.09513
width/2	0.0524	0.0544	0.369985	0.0544	0.369985	0.05239
H	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.5676	0	0.42267	1	0.57733	0.43237
2. Extr	0.8237	0.3128	1	0.6872	0	0.17631
mean	0.6957	0.1564	0.711335	0.8436	0.288665	0.30434
width/2	0.128	0.1564	0.288665	0.1564	0.288665	0.12803
5H1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.7934	0.02474	0	0.97526	1	0.20771
2. Extr	1	0.22857	0.87913	0.77143	0.12087	0
mean	0.8967	0.12665	0.439565	0.873345	0.560435	0.10386
width/2	0.1033	0.10192	0.439565	0.101915	0.439565	0.10386
EF1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.7877	0.28818	0	0.71182	1	0.2123
2. Extr	1	0.43424	0.30618	0.56576	0.69382	0
mean	0.8939	0.36121	0.15309	0.63879	0.84691	0.10615
width/2	0.1062	0.07303	0.15309	0.07303	0.15309	0.10615

EF2	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.8693	0	0.15627	1	0.84373	0.13066
2. Extr	0.96755	0.10197	1	0.89803	0	0.03245
mean	0.918445	0.05099	0.578135	0.949015	0.421865	0.08156
width/2	0.049105	0.05099	0.421865	0.050985	0.421865	0.04911
EF3	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.9765	0.1968	1	0.8032	0	0.02349
2. Extr	0.7853	0	0.11038	1	0.88962	0.21474
mean	0.8809	0.0984	0.55519	0.9016	0.44481	0.11911
width/2	0.0956	0.0984	0.44481	0.0984	0.44481	0.09563
FG1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.5773	0	0.0072125	1	0.99279	0.42273
2. Extr	0.9611	0.38916	1	0.61084	0	0.03887
mean	0.7692	0.19458	0.50360625	0.80542	0.496395	0.2308
width/2	0.1919	0.19458	0.49639375	0.19458	0.496395	0.19193

Table E.4: Turning coefficients α_{ij} at toll gates for heavy traffic

AB1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.71227	0	0.29277	1	0.70723	0.28773
2. Extr	0.89832	0.20633	1	0.793067	0	0.10168
mean	0.805295	0.103165	0.646385	0.8965335	0.353615	0.194705
width/2	0.093025	0.103165	0.353615	0.1034665	0.353615	0.093025
AB2	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.87613	0.037825	0	0.96218	1	0.12387
2. Extr	1	0.15656	0.78079	0.84344	0.21921	0
mean	0.938065	0.097193	0.390395	0.90281	0.609605	0.061935
width/2	0.061935	0.059368	0.390395	0.05937	0.390395	0.061935
2C1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.98711	0.098959	1	0.90104	0	0.012887
2. Extr	0.88837	0	0.043949	1	0.95605	0.11163
mean	0.93774	0.04948	0.5219745	0.95052	0.478025	0.0622585
width/2	0.04937	0.04948	0.4780255	0.04948	0.478025	0.0493715

BD1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.75736	0	0.62692	1	0.37308	0.24264
2. Extr	0.84913	0.10385	1	0.89615	0	0.15087
mean	0.803245	0.051925	0.81346	0.948075	0.18654	0.196755
width/2	0.045885	0.051925	0.18654	0.051925	0.18654	0.045885
CD1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.47272	0	0.56297	1	0.43703	0.52728
2. Extr	0.72603	0.36376	1	0.63624	0	0.27397
mean	0.599375	0.18188	0.781485	0.81812	0.218515	0.400625
width/2	0.126655	0.18188	0.218515	0.18188	0.218515	0.126655
EF1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	1	0.13338	0.39623	0.76986	0.60377	0
2. Extr	0.8862	0.23014	0	0.86662	1	0.1138
mean	0.9431	0.18176	0.198115	0.81824	0.801885	0.0569
width/2	0.0569	0.04838	0.198115	0.04838	0.198115	0.0569
EH1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	1	0.23274	0.46599	0.76726	0.53401	0
2. Extr	0.88345	0.13173	0	0.86827	1	0.11655
mean	0.941725	0.182235	0.232995	0.817765	0.767005	0.058275
width/2	0.058275	0.050505	0.232995	0.050505	0.232995	0.058275
EH2	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.69816	0	0.28645	1	0.71355	0.30184
2. Extr	0.91575	0.23378	1	0.76622	0	0.08425
mean	0.806955	0.11689	0.643225	0.88311	0.356775	0.193045
width/2	0.108795	0.11689	0.356775	0.11689	0.356775	0.108795
EH3	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.88551	0	0.082278	1	0.91772	0.11449
2. Extr	0.9836	0.097348	1	0.90265	0	0.016405
mean	0.934555	0.048674	0.541139	0.951325	0.45886	0.0654475
width/2	0.049045	0.048674	0.458861	0.048675	0.45886	0.0490425

H	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.70959	0.031382	0	0.96862	1	0.29041
2. Extr	1	0.30602	0.89985	0.69398	0.10015	0
mean	0.854795	0.168701	0.449925	0.8313	0.550075	0.145205
width/2	0.145205	0.137319	0.449925	0.13732	0.449925	0.145205
5H1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.81845	0	0.016755	1	0.98325	0.18155
2. Extr	1	0.17779	0.97798	0.82221	0.022019	0
mean	0.909225	0.088895	0.4973675	0.911105	0.5026345	0.090775
width/2	0.090775	0.088895	0.4806125	0.088895	0.4806155	0.090775
EF1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.8862	0.13338	0	0.86662	1	0.1138
2. Extr	1	0.23014	0.39623	0.76986	0.60377	0
mean	0.9431	0.18176	0.198115	0.81824	0.801885	0.0569
width/2	0.0569	0.04838	0.198115	0.04838	0.198115	0.0569
EF2	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.88837	0	0.043949	1	0.95605	0.11163
2. Extr	0.98711	0.098956	1	0.90104	0	0.01289
mean	0.93774	0.049478	0.5219745	0.95052	0.478025	0.06226
width/2	0.04937	0.049478	0.4780255	0.04948	0.478025	0.04937
EF3	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.79374	0	0.19895	1	0.80105	0.20626
2. Extr	0.95482	0.16592	1	0.83408	0	0.04518
mean	0.87428	0.08296	0.599475	0.91704	0.400525	0.12572
width/2	0.08054	0.08296	0.400525	0.08296	0.400525	0.08054
FG1	alpha 12	alpha 23	alpha 31	alpha 21	alpha 32	alpha 13
1. Extr.	0.62276	0	0.29734	1	0.70266	0.37724
2. Extr	0.86713	0.23422	1	0.76578	0	0.13287
mean	0.744945	0.11711	0.64867	0.88289	0.35133	0.255055
width/2	0.122185	0.11711	0.35133	0.11711	0.35133	0.122185