



# Decoupling Passenger Flows for Improved Load Prediction

Stefan Haar, Simon Theissing

► **To cite this version:**

Stefan Haar, Simon Theissing. Decoupling Passenger Flows for Improved Load Prediction. 2016. <hal-01294498>

**HAL Id: hal-01294498**

**<https://hal.inria.fr/hal-01294498>**

Submitted on 29 Mar 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Decoupling Passenger Flows for Improved Load Prediction

Stefan Haar and Simon Theissing

MEXiCo team, INRIA and LSV, CNRS & ENS de Cachan,  
Cachan, France

We elaborate the approximate computation of a stochastic hybrid automaton (SHA) model, which we have developed for the analysis of perturbations in modern multi-modal transportation networks (TNs); where passengers spread the perturbations between the different modes and lines through transfers. In particular, we focus on one major bottleneck, which may arise in the approximate computation of the SHA model: the high-dimensionality of all stochastic differential equations (SDEs). They define how all considered fluid passenger loads evolve in time in a particular mode of the SHA model, which latter might exhibit jumps between its different modes only at equidistantly-spaced discrete points in time. In this context, we replace all high-dimensional SDEs set up for a particular mode of the SHA model by a set of lower-dimensional SDEs; in that we decouple all passenger flows in a mode. We proof that the resulting approximating dynamics converges to the original model dynamics if the fixed time interval between two jump layers of the SHA model approaches zero.

**Keywords:** Stochastic Hybrid Automata, Transportation Networks, Fluid Petri Nets, Stochastic Differential Equations Modelling

## 1 Introduction

We elaborate here the approximate computation of the SHA model from [5], which we have developed [4, 6, 5] for the analysis of perturbations in modern multi-modal TNs; perturbations injected into the TNs by uncertain passenger arrival flows, and transmitted between their different modes and lines by passenger transfers.

The SHA model from [5] defines the time evolution of the fluid passenger loads and the vehicle operation considered in an extract of a TN at hand in form of a continuous-time hybrid-state stochastic process; which may exhibit jumps between its different modes only at equidistantly-spaced discrete points in time: Every mode refers to a particular discrete state in the operation of all considered vehicles, including - but not limited to - their positions and driving conditions (parked, stopped, or driving). In doing so, every mode defines which passenger flow between the stations and the vehicles docked to them are possible for the purpose of boarding & alighting. This latter piece of information is crucial in that it uniquely defines the continuous dynamics of all passenger flows in form of a set of systems of SDEs (one system of coupled SDEs for every station). Mode

transitions are either triggered by deterministic-timed events (DEs), such as the scheduled dispatch of a vehicle, or passenger load-driven events (PEs), such as the departure of a vehicle from a station because no more passenger want to alight from or board it.

The system of coupled SDEs set up for a particular station in a particular mode of the above SHA model has in general as many dimensions as the number of different places accommodating the passengers that we do consider in it; multiplied by the number of different passenger groups, which define distinct routes and preferences for the transportation services in the model. This high-dimensionality of a coupled system of SDEs renders the quick but still fairly accurate computation or simulation of the passenger flows - with uncertain initial states - in the different modes of the SHA model intractable; and this independently of the computation or simulation scheme chosen.

Based on simulation or not, the algorithms we have found so far are prone to what is known as the curse of dimensionality in the literature: Monte Carlo simulations [8] require to sample realizations of the considered random variable's (RV's) uncertain initial state. For one-dimensional RVs (subjected to one-dimensional SDEs) this sampling might be trivial (e.g. by employing the inverse transform sampling). However, it seems that sampling the uncertain initial state of multidimensional RVs is a non trivial and still open problem; although a lot of research producing different algorithms in this area has been undertaken in the past. Among these algorithms, there are notably the Metropolis-Hastings and the Gibbs sampler, which can be integrated into what is called a Markov Chain Monte Carlo simulation<sup>1</sup> [1]. Other more exotic sampling techniques might involve e.g. neural networks [7]. Instead of sampling the uncertain initial state of a multidimensional RV, and being exposed to all constraints and weaknesses - including the usage of a computer-implemented random number generator [9, pp. 275 - 279] - inherent to the subsequent simulation of several state trajectories, one could also try to numerically integrate a multivariate Fokker-Planck equation; which latter system of partial ordinary differential equations is derived from the original multidimensional SDE, and describes the time evolution of the RV's initial probability density function (PDF). However, many computational drawbacks also come along with this method, or more specifically with the numerical integrations required. First, not all numerical integration schemes can ensure the conservation of the probability flux in their basic set up; with the Finite Volume method [2] being one exception. Second those schemes which can ensure the conservation of the probability flux are not easily extendible from common two or three dimensional applications to higher-dimensional problems, due the use of special operators such as the cross product or the representation of the RV's state space.

Alternatives to the computation or simulation of high-dimensional SDEs might involve their discrete approximation (not pursued here), or techniques

---

<sup>1</sup> Stationary distribution of a Markov chain converges to PDF of RV's initial state; simulation of the latter Markov chain produces sequence of realizations for initial state

which try to decouple the dynamics in the SDEs so as to produce an alternative set of lower-dimensional SDEs (pursued here); reproducing or approximating the original model dynamics. For instance, the authors of [3] mention the *local* specification of flows in a fluid stochastic Petri net model as a means for the decoupling. However, in contrast to our approach, they do not look at vectorial (passenger) flows, but at scalar flows.

In the rest of this paper, we shortly review our SHA model from [5] in Sec. 2 including the set up of all original high-dimensional SDEs for the passenger flow dynamics in its different modes. We then explain in Sec. 3 how the passenger flows can be systematically decoupled so as to replace them by lower-dimensional SDEs that approximate the original passenger flow dynamics. Last but not least, we summarize the contribution of our approach, and give a brief outlook on future work in Sec. 4.

## 2 Our SHA Model

### 2.1 Model Structure

**Infrastructure.** Basic modelling blocks of the SHA model are place/transition nets (= Petri nets with the token flow left out), which capture the structure of a finite set of stations  $\mathcal{S}$  and a finite set of transportation grids  $\mathcal{G}$  (TGs).

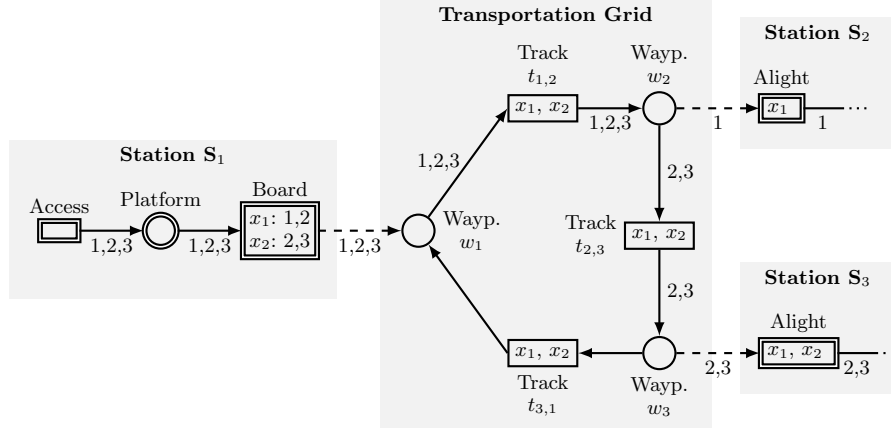
Every station  $s \in \mathcal{S}$  is made up of a finite set  $P_s$  of gathering points  $p \in P_s$  (= places; represented by double circles) that can accommodate a limited number of passengers, and a finite set  $T_s$  of corridors  $t \in T_s$  (= transitions; represented by double boxes) connecting (i) GPs to other GPs, or (ii) GPs to the station's exterior (cf. Fig. 1 below). Here, connected means “possibility of a passenger flow” in the direction of the edges that connect the corridors with the GPs.

Every TG  $g \in \mathcal{G}$  captures the structure of a particular mode or line; and in doing so, all possible vehicle movements between its finite set  $W_g$  of discrete waypoints  $w \in W_g$  (= places; represented by simple circles) which accommodate the vehicle tokens (at maximum one vehicle per waypoint) via tracks (= transitions; represented by simple boxes).

A finite set of tuples  $(a, b) \in \mathcal{I}$ , with  $\mathcal{I} \subseteq (T \times W) \cup (W \times T)$ ,

$$T := \bigcup_{s \in \mathcal{S}} T_s \text{ and } W := \bigcup_{g \in \mathcal{G}} W_g,$$

composed of a transition in a station and a waypoint in a TG, defines the interface between the stations and the TGs (represented by dashed arcs in Fig. 1 above): Every tuple  $(a, b) \in \mathcal{I}$  either connects some GP in a station  $s \in \mathcal{S}$  to a waypoint in a TG  $g \in \mathcal{G}$ , in which case  $a \in P_s$  and  $b \in W_g$ ; or vice versa. In this way, every tuple defines which passenger flow between a vehicle stopped at a waypoint in a TG and a GP (= platform) in a station is possible for the purpose of boarding & alighting; see below.



**Fig. 1.** Representation of the infrastructure of a sample TN (adopted from [5]) in our SHA model, together with (i) the paths of two different vehicle missions  $x_1$  and  $x_2$ , and (ii) an indication of the stops along these paths for the specification of three different trip profiles (TPs) 1, 2, and 3: the labels inscribed to the transitions of the stations indicate that a vehicle which executes  $x_2$  is supposed to stop at S<sub>1</sub> and S<sub>3</sub>, but not at S<sub>2</sub>. Moreover, they indicate that all passengers of TP 3 prefer to board a vehicle that executes  $x_2$  in order to travel from S<sub>1</sub> to S<sub>3</sub> over a vehicle that executes  $x_1$ .

**Vehicle Operation.** At the heart of the operation of a finite set  $\mathcal{V}$  of all vehicle tokens  $v \in \mathcal{V}$  considered in the SHA model are missions: Every mission defines a path in a particular transportation grid, together with (i) a sequence of stops at the waypoints along that path; (ii) deterministic-timed (minimum & maximum dwell times) and passenger load-dependent departure conditions from the stops which might state that a vehicle cannot depart from a stop as long as some passengers still want to alight from or board it; and (iii) driving times between all waypoints which might be functions of the positions of all vehicle tokens.

**Passenger Routing.** We group all passengers into a finite set  $\mathcal{Y} := \{1, 2, \dots, n\}$  of  $n \in \mathbb{N}$  different trip profiles (TPs): Every  $y \in \mathcal{Y}$  defines a particular path in TN's infrastructure, together with the passengers' preferences for the different vehicle missions (cf. Fig. 1 above). However, this does not mean that the passengers cannot change their TPs as we will highlight in a short (see Sec. 2.3).

## 2.2 Hybrid State

As common in the literature of hybrid automata, we refer to the discrete state of our SHA model at any time  $\tau \geq 0$  as its mode: A particular mode  $q \in Q$  from the finite set of all different modes  $Q$  defines for every  $v \in \mathcal{V}$  (i) the position of  $v$  in form of a waypoint in a TG; (ii) the driving condition of  $v$  which is either parked, stopped or driving; (iii) the operational state of  $v$  in form of a

mission to be executed, a discrete state therein, and a sequence of missions to be accomplished. Thus, every  $q \in Q$  tells us which vehicle is docked to which station; and in doing so, defines the (continuous) passenger flow dynamics in TN.

*Remark 1.* We say that a vehicle  $v \in \mathcal{V}$  is docked to a station  $s \in \mathcal{S}$  iff (i)  $v$  is stopped at a waypoint  $w \in W_g$  in some TG  $g \in \mathcal{G}$ ; (ii) acc. to  $\mathcal{I}$ , either passengers can board  $v$  stopped at  $w$  from some GP in  $s$ , or alight from it to some GP in  $s$ ; and (iii) acc. to the specification of all TPs, passengers of at least one TP might want to board  $v$  or alight from it. Moreover, we denote by  $\mathcal{V}(s, q) \subseteq \mathcal{V}$  the subset of all vehicles that are docked to  $s$  in  $q$ .

*Remark 2.* If  $k$  is a row vector, then we denote by  $k[i]$  the element in its  $i$ -th column. Accordingly, if  $k$  is a column vector, then we denote by  $k[i]$  the element in its  $i$ -th row.

The continuous state of the SHA model at any  $\tau \geq 0$ , defines (i) the elapsed dwell times of all stopped vehicles, (ii) the elapsed driving times of all moving vehicles, and (iii) the passenger load  $M(b, \tau)$ , with

$$M : (b, \tau) \in (P \cup \mathcal{V}) \times \mathbb{R}_{\geq 0} \rightarrow \mathcal{M}(b)$$

and

$$\mathcal{M}(b) := \left\{ k \in (\mathbb{R}_{\geq 0})^{|\mathcal{Y}|} : \sum_{i=1}^{|\mathcal{Y}|} k[i] \leq c(b) \right\}, \quad (1)$$

for every vehicle  $b \in \mathcal{V}$  and every GP in a station  $b \in P$ . Therein,

$$P := \bigcup_{s \in \mathcal{S}} P_s,$$

$M(b, \tau)[i]$  gives the number of passenger at/on-board  $b$ , who travel acc. to the TP  $i \in \mathcal{Y}$ , and  $c(b)$ , with  $c : P \cup \mathcal{V} \rightarrow \mathbb{R}_{> 0}$ , gives the maximum number of passengers  $b$  can accommodate at the same time.

### 2.3 Balance Equations

For any  $q \in Q$ , we adapt the notation  $\bullet b(q)$  for the preset and  $b^\bullet(q)$  for the postset of any  $b \in P \cup \mathcal{V}'(q)$ , with

$$\mathcal{V}'(q) := \bigcup_{s \in \mathcal{S}} \mathcal{V}(s, q),$$

from the Petri nets literature for our purposes:  $\bullet b(q)$  denotes the set of all corridors in the stations that are connected by an arc pointing towards  $b$ . Accordingly,  $b^\bullet(q)$  denotes the set of all corridors in the stations that are connected by an arc pointing away from  $b$ . For  $b \in \mathcal{V}'(q)$ , those arcs (dashed arcs in Fig. 1 above) point towards/away from the waypoint which accommodates  $b$ .

Note that all corridors in the stations of our SHA model are connected in a special way to the rest of the modelled infrastructure (GPs in the stations and waypoints in the TGs).

*Remark 3.* For any  $t \in T$ , we denote by  ${}^*t(q) := b$  the single GP in a station or vehicle docked to a station  $b \in P \cup \mathcal{V}'(q)$  which is connected to  $t$  in  $q$  by an arc pointing towards  $t$  iff  $t \in b^\bullet(q)$ . Accordingly, we denote by  $t^*(q) := a$ , for any  $t \in T$ , the single GP or vehicle docked to a station  $a \in P \cup \mathcal{V}'(q)$  which is connected to  $t$  in  $q$  by an arc pointing away from  $t$  iff  $t \in {}^\bullet a(q)$ .

This special structure allows us to decompose all corridors in  $q \in Q$  into three disjoint sets; implementing inflows, transfer flows, and outflows: Inflows model the arrival processes of the passengers who join the SHA model from TN's exterior.

**Definition 1 (Inflow).** *An inflow is a passenger flow assigned to any  $t \in T_1$ , with*

$$T_1 := \left\{ t \in T : \exists p \in P \text{ s.t. } t \in {}^\bullet p \wedge \nexists p' \in P \text{ s.t. } t \in p^\bullet \wedge \nexists w \in W \text{ s.t. } (w, t) \in \mathcal{I} \right\}. \quad (2)$$

Transfer flows model passenger flows within the SHA model; including passenger transfers between the GPs in the stations, as well as passenger transfers between GPs in the stations and vehicles docked to the stations.

**Definition 2 (Transfer Flow).** *A transfer flow in  $q \in Q$  is a passenger flow assigned to any  $t \in T_2(q)$ , with*

$$T_2(q) := \left\{ t \in T : \exists b \in P \cup \mathcal{V}(q) \text{ s.t. } t \in {}^\bullet b \wedge \exists b' \in P \cup \mathcal{V}'(q) \text{ s.t. } t \in (b')^\bullet \right\}. \quad (3)$$

Finally, outflows model the SHA model's drain of passengers to TN's exterior.

**Definition 3 (Outflow).** *An outflow is a passenger flow assigned to any  $t \in T_3$ , with*

$$T_3 := \left\{ t \in T : \exists p \in P \text{ s.t. } t \in p^\bullet \wedge \nexists p' \in P \text{ s.t. } t \in p^\bullet \wedge \nexists w \in W \text{ s.t. } (t, w) \in \mathcal{I} \right\}. \quad (4)$$

With that said, we denote by  $T'(q)$ , with  $T'(q) := T_1 \cup T_2(q) \cup T_3$ , the set of all corridors active in  $q \in Q$ ; and by  $\gamma(\tau)$ , with  $\gamma : \mathbb{R}_{\geq 0} \rightarrow Q$ , the mode of our SHA model at time  $\tau \geq 0$ .

$$\begin{aligned} dM(b, \tau) := & \sum_{t \in {}^\bullet b \cap T'(\gamma(\tau))} R(t) \overbrace{[\phi(t, \tau) d\tau + \delta(t) dW(\tau)]}^{\text{Passenger flow into } b} - \\ & \sum_{t \in {}^\bullet b \cap T'(\gamma(\tau))} \underbrace{[\phi(t, \tau) d\tau + \delta(t) dW(\tau)]}_{\text{Passenger flow leaving } b} \end{aligned} \quad (5)$$

then defines the time evolution of the passenger load of every GP in a station and of every vehicle docked to a station  $b \in P \cup \mathcal{V}'(q)$  at any time  $\tau \geq 0$  when the SHA model is in  $q \in Q$ . This balance equation relates  $M(b, \tau)$  to all passenger flows into  $b$  and leaving it: We capture the routing of all passengers along the different TPs as well as their local re-routing among these TPs in so-called routing matrices.

*Remark 4.* We denote by  $\Psi^{d_1 \times d_2}$ , for some  $d_1, d_2 \in \mathbb{N}_{>0}$  and any set  $\Psi$ , the set of all matrices with  $d_1$  rows and  $d_2$  columns, whose elements are from  $\Psi$ . In the case that  $d_2 = 1$ , we drop  $d_2$  in  $\Psi^{d_1 \times d_2}$  and write  $\Psi^{d_1}$  instead.

The  $i$ -th row and the  $j$ -th column of a particular routing matrix  $R(t)$  assigned to  $t \in T$ , with

$$R : T \rightarrow \left\{ K \in (\mathbb{R}_{\geq 0})^{|\mathcal{Y}| \times |\mathcal{Y}|} : \sum_{i=1}^{|\mathcal{Y}|} K[i, j] = 1, \forall j = \mathcal{Y} \right\},$$

defines the relative amount of the flow of passengers who join  $t$  acc. to the TP  $j \in \mathcal{Y}$ , and who leave  $t$  acc. to the TP  $i \in \mathcal{Y}$ ; and the fact that every column of  $R(t)$  must either sum up to one or to zero, implies that all passenger flows are conserved.

*Remark 5.* Time could be included in the domain of the routing matrices above so that they might change values during mode transitions of the SHA model depending on the hybrid state; so as to account e.g. for loudspeaker announcements.

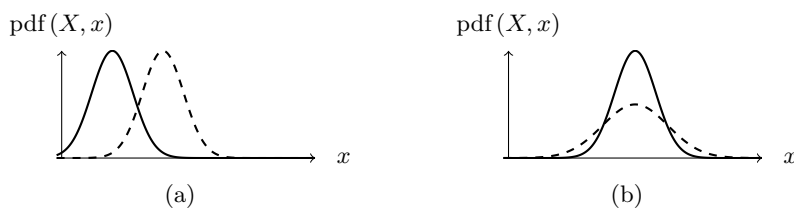
We next write down the passenger flow assigned to every corridor  $t \in T(q)$  in  $q$  acc. to its impact on  $M(p, \tau)$  as the sum of a drift term  $\phi(t, \tau)$ , with

$$\phi : (t, \tau) \in \bigcup_{q \in Q} T'(q) \times \mathbb{R}_{\geq 0} \rightarrow \left\{ v \in (\mathbb{R}_{\geq 0})^{|\mathcal{Y}|} : \sum_{i=1}^{|\mathcal{Y}|} v[i] \leq \phi_{\max}(q, t) \right\},$$

and a constant diagonal diffusion term

$$\delta : \bigcup_{q \in Q} T'(q) \rightarrow \left\{ K \in \mathbb{R}^{|\mathcal{Y}| \times |\mathcal{Y}|} : K[i, j] = 0, \forall i \neq j \right\}.$$

Therein,  $\phi_{\max}(q, t)$ , with  $\phi_{\max} : q \in Q \times T'(q) \rightarrow \mathbb{R}_{\geq 0}$ , is the maximum passenger throughput of the corridor  $t \in T'(q)$ , when the SHA model is in  $q \in Q$ .



**Fig. 2.** Schematic representation of the impact of the drift (a) and the diffusion (b) term of a one-dimensional passenger flow into or leaving some GP in a station or vehicle docked to a station, on its passenger load, say  $X$ ; reflecting boundaries are ignored



*Remark 6.* Let  $X$  be a continuous RV. Then,  $\text{pdf}(X)$  denotes its PDF;  $\sigma(X)$  denotes its state space; and  $\text{pdf}(X, x)$  denotes the evaluation of  $\text{pdf}(X)$  at  $x$  for some  $x \in \sigma(X)$ .

We discuss the specification of  $\phi(\cdot)$  and  $\delta(\cdot)$  in more detail in the rest of this paper. Here, only note that the drift term of a flow into some  $b \in P \cup \mathcal{V}'(q)$  shifts the density of  $M(b, \tau)$  in its domain (cf. Fig. 2a above). The flow's diffusion term narrows or broadens the density of  $M(b, \tau)$  (cf. Fig. 2b above).

## 2.4 Grouping of Balance Equations

In principle, the passenger flows in (5) can be defined as any functions of the SHA model's complete hybrid state as long as they are capacity- and demand-sensitive; crucial properties that we assume for all passenger flows in our SHA model: We say that some passenger flow is capacity-sensitive iff its drift does not cause the passenger load of some GP or vehicle to exceed the capacity limit of that GP or vehicle.

**Definition 4 (Capacity-Sensitive Flow).** *A passenger flow assigned to some  $t \in T'(q)$  in  $q \in Q$  is capacity-sensitive iff  $t \in T_3$  or*

$$\sum_{i=1}^{|\mathcal{Y}|} M(t^*, \tau)[i] \rightarrow c(t^*)$$

*implies that  $\phi(t, \tau) \rightarrow 0$  for any  $\tau \geq 0$ .*

Additionally, we say that a passenger flow is demand-sensitive iff its drift does not cause any passenger load to become negative.

**Definition 5 (Demand-Sensitive Flow).** *A passenger flow assigned to some  $t \in T'(q)$  in  $q \in Q$  is demand-sensitive iff  $t \in T_1$  or*

$$M(*t, \tau)[j] \sum_{i=1}^{|\mathcal{Y}|} R(t)[i, j] \rightarrow 0$$

*implies that  $\phi(t, \tau)[j] \rightarrow 0$  for all  $j \in \mathcal{Y}$  and for any  $\tau \geq 0$ .*

*Remark 7.* Def. 4 and 5 taken alone cannot ensure the non-negativity and capacity limits of the passenger loads assuming non-zero diffusion terms in (5). Instead both properties must be explicitly ensured during the computation or simulation of (5) in form of reflecting boundary conditions. See e.g. [6], where we derive reflecting boundary conditions for the numerical integration of a multivariate Fokker-Planck equation obtained from (5).

For our purposes however, we do not need this kind of global inclusion of the SHA model's complete hybrid state into the specification of the passenger flows: We restrict the domains of their drift terms to the passenger loads in their presets and postsets.

**Definition 6 (Local Flow).** A passenger flow assigned to some  $t \in T'(q)$  in  $q \in Q$  is local iff for any  $\tau \geq 0$ ,

- $t \in T_1$ , and the flow's drift term only depends on  $M(t^*, \tau)$ , or
- $t \in T_2(q)$ , and the flow's drift term only depends on  $M(*t, \tau)$  and  $M(t^*, \tau)$ ,  
or
- $t \in T_3$ , and the flow's drift term only depends on  $M(*t, \tau)$ .

This *local specification* of all passenger flows produces a natural decomposition of all SDEs set up for any  $q \in Q$ : The balance equations in form of (5) set up for the passenger loads of all GPs  $p \in P_s$  and vehicles  $v \in \mathcal{V}(s, q)$ , for some station  $s \in \mathcal{S}$ , are independent from the passenger loads of all GPs outside  $s$  and vehicles not docked to  $s$ . We can thus group them into one common system of coupled SDEs of dimension  $k := (|P_s| + |\mathcal{V}(s, q)|) |\mathcal{Y}|$ , which latter system is decoupled from those systems set up for all other stations.

*Remark 8.* In practice, we do only have to consider all those TPs in the domain specification for the passenger load of a particular GP or vehicle, whose paths cover this GP or vehicle. Thus,  $k$  as defined above only defines an upper bound for the dimension of the system of SDEs set up for  $s$  in  $q$ .

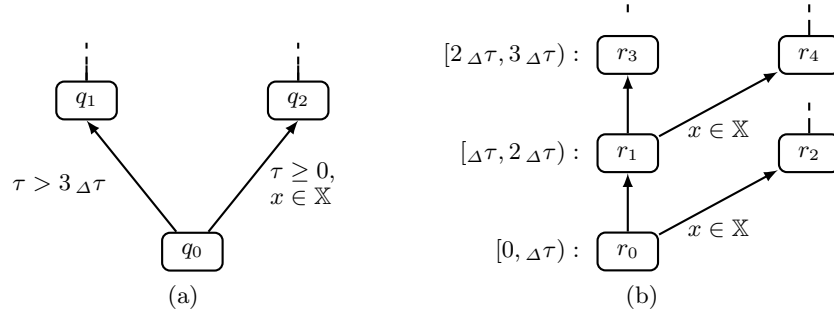
## 2.5 Mode Transitions

We assume that at simulation time  $\tau = 0$ , with  $\tau \geq 0$ , our SHA model is in one particular mode with marginal probability one, and we know the elapsed driving & dwell times of all vehicles. We then let our SHA model transition between its discrete modes only at discrete time steps  $\tau = i \Delta\tau$ , with  $i \in \mathbb{N}_{>0}$ , of fixed length  $\Delta\tau > 0$ . In this context, we also let the elapsed driving & dwell times of all vehicles only evolve at  $\tau = i \Delta\tau$  by  $\Delta\tau$ . We then capture the time evolution of our SHA model's vehicle load (= particular mode and particular realization of all elapsed discrete driving & dwell times) in what we refer to as its vehicle load tree (VLT); cf. Fig. 3b below. We do not go into details of its computation here, but only stress some important points. Refer to [5] for more information: Every node, say  $m$ , in this VLT, say  $G$ , represents a particular vehicle load for our SHA model in the half-closed time interval  $[h_m \Delta\tau, (h_m + 1) \Delta\tau)$  iff  $h_m \in \mathbb{N}_{\geq 0}$  is the height of  $m$  in  $G$ . Thus, two nodes with the same height  $h' \in \mathbb{N}_{>0}$  in  $G$  represent two alternatives for our SHA model's vehicle load in  $[h' \Delta\tau, (h' + 1) \Delta\tau)$ . Two or more branches away from  $m$  indicate the possibility of mode transitions; with one branch for every alternative mode transition, and one additional branch for the continuation of  $m$ -th mode. Several nodes with the same height in  $G$  can have the same mode and thus the same passenger flow dynamics in common.

## 2.6 Propagation of Passenger Loads

At any simulation time  $\tau = i \Delta\tau$ , with  $i \in \mathbb{N}_{\geq 0}$  and  $\Delta\tau > 0$ , one single marginal joint PDF, say  $\text{pdf}(i)$ , defines the passenger loads of all GPs in the stations and of all vehicles. For  $i = 0$ , we assume that  $\text{pdf}(i)$  is known with marginal probability

one. Then, starting from  $i = 0$ , all passenger loads have to be propagated forward in time from one time layer in the VLT to the next: For the computation of  $\text{pdf}(i + 1)$ , for some  $i \in \mathbb{N}_{\geq 0}$ , all high-dimensional systems of SDEs defined by our SHA model's different modes in the time layer  $[i \Delta\tau, (i + 1) \Delta\tau)$  of the VLT, must be computed from  $\tau = i \Delta\tau$  to  $\tau = (i + 1) \Delta\tau$  with  $\text{pdf}(i)$  as common initial PDF. Depending on the particular use case at hand so as to e.g. forecast the risk of overcrowded platforms, this forward propagation is normally terminated once the simulation time exceeds some constant threshold. Refer to [5] for more details.



**Fig. 3.** Schematic comparison of a (classical) mode graph (a) and a VLT (b) for our SHA model:  $\mathbb{X}$  denotes a compact region in the SHA model's complete passenger load space as entrance condition for a not further specified passenger load-driven mode transition, and  $\Delta\tau > 0$  is the fixed time step that separates every pair of two consecutive time layers in the VLT when the SHA model updates its vehicle load

### 3 The Decoupling of All Passenger Flows

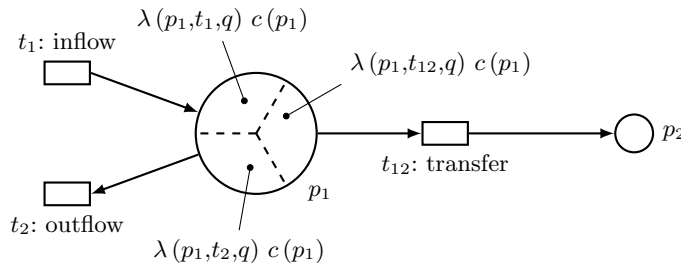
#### 3.1 Overview

Our decoupling approach is perhaps best described by the following sequence of images: We assume that every GP in a station and every vehicle  $b \in P \cup \mathcal{V}$  has the shape of a circular area, say  $A_b$ . We next assume that the passenger load of  $b$  is equally distributed on  $A_b$  at any simulation time step  $\tau = i \Delta\tau$ , with  $\tau \geq 0$ ,  $i \in \mathbb{N}_{\geq 0}$ , and  $\Delta\tau > 0$ ; in which  $\Delta\tau$  is the fixed time step that separates every pair of two consecutive time layers in the VLT of our SHA model.

*Remark 9.* We denote by  $\Gamma(\tau)$ , with  $\Gamma : \mathbb{R}_{\geq 0} \rightarrow 2^Q \setminus \emptyset$ , the subset of all modes our SHA model can be in at time  $\tau \in \mathbb{R}_{\geq 0}$ .

For any time  $\tau \in \mathcal{H}_i$ , from the time interval  $\mathcal{H}_i := [i \Delta\tau, (i + 1) \Delta\tau)$ , any mode  $q \in \Gamma(\tau)$ , and any  $b \in P \cup \mathcal{V}'(q)$ , we divide  $A_b$  into  $|(\bullet b \cup b \bullet) \cap T'(q)|$  non-overlapping slices (cf. Fig 4 below); in which one slice is attributed to every

passenger flow into or leaving  $b$ , i.e., the passenger flow assigned to every corridor  $t \in (\bullet b \cup b \bullet) \cap T'(q)$ . Our assumptions above then imply that at  $\tau = i \Delta\tau$  (i) the surface area of a particular slice defines how many passengers it accommodates at  $\tau$ , and (ii) the distribution of this latter number of passengers w.r.t. the passengers' different TPs is identical to the distribution of the total number of passengers at  $b$  and  $\tau$  w.r.t. the different TPs. We moreover assume that a retractable wall is installed along every frontier separating two neighbouring slices (dashed lines in Fig. 4 below). These walls prevent the equidistant re-distribution of the slices' passenger loads at any  $\tau \in \mathcal{H}_i$ , which *diffusion* is restricted to the discrete time step  $\tau = (i + 1) \Delta\tau$  ( $\rightarrow$  walls are not present any more at  $\tau = (i + 1) \Delta\tau$ ).



**Fig. 4.** Schematic representation of our decoupling approach: all GPs and vehicles docked to the stations in a particular mode of the SHA model are divided into slices, with impenetrable walls separating neighbouring slices during two consecutive jump layers; passenger flows join and leave the slices instead of joining and leaving the original GPs and vehicles; re-distribution of slices' passenger loads only at the SHA model's discrete jump times

So in our physically-touched model above, the slices' passenger loads are decoupled at any  $\tau \in \mathcal{H}_i$ , which implies that they might be filled and emptied at different rates if we assume that the passengers flow into and leave the slices of  $b$ ; instead of flowing into and leaving  $b$  itself. For the specification of the slices' surface areas, we use the maximum passenger throughputs assigned to the corridors for the different modes; see below.

### 3.2 Decoupled Balance Equations

**General Structure.** The system of SDEs that we will set up for the decoupled passenger flow assigned to every  $t \in T'(q)$  for any  $q \in Q$  next, defines how this flow manipulates the passenger load  $M_{q,t}(*t, \tau)$  of the isolated slice from  $*t$  attributed to  $t$  in  $q$  and/or the passenger load  $M_{q,t}(t^*, \tau)$  of the isolated slice from  $t^*$  attributed to  $t$  in  $q$ ; when our SHA model is in  $q$ . We write it down in the very general form of

$$dX_{q,t}(\tau) := A_{q,t}(X_{q,t}(\tau)) d\tau + B_{q,t}(X_{q,t}(\tau)) dW(\tau), \quad (6)$$

with the state vector  $X_{q,t}$ , the drift vector  $A_{q,t}$ , the diffusion matrix  $B_{q,t}$ , and the vector of  $|\mathcal{Y}|$  uncorrelated Wiener processes  $\mathcal{W}$ .

*Remark 10.* We write the tuple of a mode  $q \in Q$  and a transition  $t \in T'(q)$  in form of subscript separating both in the given order by a comma next to a variable or constant iff we refer to the projection of that variable or constant in (6) set up for the decoupled passenger flow assigned to  $t$  in  $q$ .

**Projection of Passenger Loads & Flows.** As outlined in the figurative overview of our decoupling approach above, we project  $M(b, \tau)$ , for any  $b \in P \cup \mathcal{V}'(q)$  and  $q \in Q$ , to  $M_{q,t}(t, \tau)$ , with  $M_{q,t} : T'(q) \times \mathbb{R}_{\geq 0} \rightarrow \mathcal{M}_{q,t}(b)$  and

$$\mathcal{M}_{q,t}(b) := \left\{ k \in (\mathbb{R}_{\geq 0})^{|\mathcal{Y}|} : \sum_{i=1}^{|\mathcal{Y}|} k[i] \leq \lambda(b, t, q) c(b) \right\},$$

at  $\tau = i_{\Delta}\tau$ , with  $i \in \mathbb{N}_{\geq 0}$ , acc. to

$$M_{q,t}(b, i_{\Delta}\tau) := \lambda(b, t, q) M(b, i_{\Delta}\tau) \quad (7)$$

iff our SHA model is in mode  $q$  at  $\tau = i_{\Delta}\tau$ . Therein,  $\lambda(b, t, q)$ , with

$$\lambda(b, t, q) := \frac{\phi_{\max}(q, t)}{\sum_{t' \in (\bullet b \cup b \bullet) \cap T'(q)} \phi_{\max}(q, t')}, \quad (8)$$

defines the maximum number of passengers  $\lambda(b, t, q) c(b)$  the isolated slice from  $b \in P \cup \mathcal{V}'(q)$  attributed to  $t \in (\bullet b \cup b \bullet) \cap T'(q)$  in  $q$  can accommodate (cf. Fig. 4 above). This simple projection implies

$$\text{pdf}(M_{q,t}(b, i_{\Delta}\tau) = \lambda(b, t, q) k) = \text{pdf}(M(b, i_{\Delta}\tau) = k), \forall k \in \mathcal{M}(b), \quad (9)$$

with  $\mathcal{M}(b)$  from (1). We also use (8) to project  $\phi(t, \tau)$  - which we assume to be local, demand- & capacity sensitive - to  $\phi_{q,t}(t, \tau)$  acc. to Tab. 1 below, which implies that all qualitative properties of  $\phi(t, \tau)$  such as demand-sensitiveness are adopted by  $\phi_{q,t}(t, \tau)$ .

**Table 1.** Specification of  $\phi_{q,t}(t, \tau)$  assigned to  $t \in T'(q)$  in  $q \in Q$

Inflow:	$\phi(\lambda^{-1}(t^*, t, q) M_{q,t}(t^*, \tau))$
Transfer Flow:	$\phi(\lambda^{-1}(*t, t, q) M_{q,t}(*t, \tau), \lambda^{-1}(t^*, t, q) M_{q,t}(t^*, \tau))$
Outflow:	$\phi(\lambda^{-1}(*t, t, q) M_{q,t}(*t, \tau))$

**Inflows.** In general, we neither know the passengers' exact arrival times, nor the TPs of the new arriving passengers. However, in most situations we know some reference values, and we can estimate quite reasonably fluctuations around them (e.g. from statistical considerations); which latter knowledge we can then map to the systems of SDEs set up for all decoupled inflows. More specifically, we set up for every  $t \in T_1$  a balance equation in form of (5), which defines the impact of the inflow assigned to  $t$ , to the passenger load of  $t^*$ ; and integrate this balance equation into (6). Tab. 2 lists the corresponding ingredients.

**Transfer Flows.** Once having joined the SHA model, we assume that the passenger transfer dynamics regarded in isolation within the SHA model in a particular mode is deterministic; which implies zero diffusion terms for the specification of all decoupled passenger transfer flows: For every  $t \in T_2(q)$  in  $q \in Q$ , we set up two balance equations in form of (5). The first balance equation defines the impact of the transfer flow assigned to  $t$ , to the passenger load of  $^*t$ . Accordingly, the second balance equation relates the passenger load of  $^*t$  to the same decoupled transfer flow. We then integrate both balance equations into (6) acc. to Tab. 2.

**Outflows.** Similar to the specification of all transfer flows above, we demand zero diffusion terms for all passenger outflows: For every  $t \in T_3$ , we set up a balance equation in form of (5) and integrate it into (6). This balance equation relates the passenger load of  $^*t$ , to the outflow assigned to  $t$  (cf. Tab. 2).

**Table 2.** Specification of the system of SDEs set up for the decoupled inflow, transfer flow, or outflow assigned to  $t \in T'(q)$  in mode  $q \in Q$  of our SHA model

	Inflow	Transfer Flow	Outflow
Schematic structure	$t \square \rightarrow \bigcirc t^*$	$^*t \bigcirc \xrightarrow{t} \square \rightarrow \bigcirc t^*$	$^*t \bigcirc \rightarrow \square t$
$X_{q,t}(\tau)$	$M_{q,t}(t^*, \tau)$	$\begin{bmatrix} M_{q,t}(^*t, \tau) \\ M_{q,t}(t^*, \tau) \end{bmatrix}$	$M_{q,t}(^*t, \tau)$
$A_{q,t}(\tau)$	$R(t) \phi_{q,t}(t, \tau)$	$\begin{bmatrix} -\phi_{q,t}(t, \tau) \\ R(t) \phi_{q,t}(t, \tau) \end{bmatrix}$	$-\phi_{q,t}(t, \tau)$
$B_{q,t}$	$\delta(t)$	0	0

### 3.3 Correctness of Our Decoupling Approach

Assume that our SHA model is in mode  $q \in Q$  at time  $\tau = i_{\Delta}\tau$ , for some  $i \in \mathbb{N}_{\geq 0}$ ; in which  $\Delta\tau > 0$  is the fixed time step that separates every pair

of two consecutive time layers in its VLT. Moreover, assume that we like to compute the probability of a particular mode transition of the SHA model at time  $\tau = (i + 1) \Delta\tau$ ; which is triggered by the passenger load trajectory of some GP in a station or vehicle docked to a station  $b \in P \cup \mathcal{V}'(q)$  taking a value from  $k \in K$ , with  $K \subseteq \mathcal{M}(b)$  and  $\mathcal{M}(b)$  from (1). More formally speaking, we thus like to compute the probability

$$\mathbb{P}(M(b, (i + 1) \Delta\tau) \in K) := \int_K \text{pdf}(M(b, (i + 1) \Delta\tau) = k) dk \quad (10)$$

with  $M(b, \tau)$  specified at  $\tau = i \Delta\tau$  by  $\text{pdf}(M(b, i \Delta\tau))$  acc. to (9).

*Remark 11.* Let  $X_1, X_2, \dots, X_n$  be a vector of  $n \in \mathbb{N}_{>0}$  continuous RVs. Then,  $\text{pdf}(X_i; i \in \{1, 2, \dots, n\})$  denotes the joint PDF of  $X_1, X_2, \dots, X_n$ ;  $\text{pdf}(X_i = x_i; i \in \{1, 2, \dots, n\})$  denotes the evaluation of  $\text{pdf}(X_i; i \in \{1, 2, \dots, n\})$  at  $(x_1, x_2, \dots, x_n)$ , with  $x_i \in \sigma(X_i), \forall i \in \{1, 2, \dots, n\}$ .

Look at

$$\mathbb{P} \left( \sum_{t \in (\bullet b \cup b \bullet) \cap T'(q)} M_{q,t}(b, (i + 1) \Delta\tau) \in K \right) = \int_K \text{pdf}(M_{q,t}(b, (i + 1) \Delta\tau) = k) dk \quad (11)$$

instead, which is the probability that the sum of the decoupled passenger loads of the different isolated slices from  $b$  (isolated in  $q$ ) takes a value from  $K$  at  $\tau = (i + 1) \Delta\tau$ . Let

$$l := |(\bullet b \cup b \bullet) \cap T'(q)|, \quad (12)$$

and introduce the set  $\overline{\mathcal{M}}(b, k)$ , with

$$\overline{\mathcal{M}}(b, k) := \left\{ (k_1, k_2, \dots, k_l) \in (\mathcal{M}(b))^l : \sum_{i=1}^l k_i = k \right\} \quad (13)$$

Moreover, let  $\{t_1, t_2, \dots, t_l\} := (\bullet b \cup b \bullet) \cap T'(q)$ . Then, write down (11) in form of

$$\mathbb{P} \left( \sum_{t \in (\bullet b \cup b \bullet) \cap T'(q)} M_{q,t}(b, (i + 1) \Delta\tau) \in K \right) = \int_K \int_{\overline{\mathcal{M}}(b, k)} \text{pdf}(M_{q,t_i}(b, (i + 1) \Delta\tau) = k_i; i \in \{1, 2, \dots, l\}) d(k_1, k_2, \dots, k_l) dk \quad (14)$$

Therein, note that  $M_{q,t_1}(b, (i+1) \Delta\tau), \dots, M_{q,t_l}(b, (i+1) \Delta\tau)$  are independent RVs. Thus, (14) simplifies to

$$\begin{aligned} \mathbb{P} \left( \sum_{t \in (\bullet \cup b \bullet) \cap T'(q)} M_t(b, (i+1) \Delta\tau) \in K \right) = \\ \int_K \int_{\overline{\mathcal{M}(b,k)}} \prod_{t \in (\bullet \cup b \bullet) \cap T'(q)} \text{pdf}(M_{t_i}(b, (i+1) \tau) = k_i) \, d(k_1, k_2, \dots, k_l) \, dk \end{aligned} \quad (15)$$

**Theorem 1.** *For any  $q \in Q$ ,  $b \in P \cup \mathcal{V}'(q)$ , and  $k \in \mathcal{M}(b)$ , the integral*

$$\int_{\overline{\mathcal{M}(b,k)}} \prod_{t \in (\bullet \cup b \bullet) \cap T'(q)} \text{pdf}(M_{t_i}(b, (i+1) \tau) = k_i) \, d(k_1, k_2, \dots, k_l)$$

from (15) converges to  $\text{pdf}(M(b, (i+1) \Delta\tau) = k)$  from (10) for  $\Delta\tau \xrightarrow{\Delta\tau > 0} 0$ .

The simple proof of Thm. 1 can be found in the appendix. Here, only note that Thm. 1 implies that our decoupling approach produces a set of SDEs (one for every decoupled flow) for the different modes of our SHA model, which approximates the original passenger flow dynamics very well if we let the fixed time step  $\Delta\tau$  separating the different time layers in its VLT approach zero; since  $K$  above can be chosen arbitrarily. Compared to the original high-dimensional systems of SDEs, the new replacing systems of SDEs have at maximum  $2 |\mathcal{Y}|$  dimensions (= systems of SDEs set up for transfer flows).

## 4 Summary & Outlook

In this paper, we have considered one major bottleneck that may arise in the approximate computation of our SHA model from [5]: the numerical computation of the many high-dimensional SDEs, which define the passenger flow dynamics in its different modes. More specifically, we have shown how all passenger flows can be systematically decoupled in the different modes of our SHA model, which produces a set of lower-dimensional SDEs replacing the original ones. We have also proofed correctness of this decoupling approach.

In future work, we intend to further improve the approximate computation of our SHA model, and to apply it to some prediction use cases.

## A Proof

**Common Initial State** From (7), note that

$$\begin{aligned} \sum_{t \in (\bullet \cup b \bullet) \cap T'(q)} M_{q,t}(b, i \Delta\tau) &= \sum_{t \in (\bullet \cup b \bullet) \cap T'(q)} \lambda(b, t, q) M(b, i \Delta\tau) \\ &= M(b, i \Delta\tau) \sum_{t \in (\bullet \cup b \bullet) \cap T'(q)} \lambda(b, t, q). \end{aligned} \quad (16)$$



From (9) follows

$$\sum_{t \in (\bullet b \cup b \bullet) \cap T'(q)} \lambda(b, t, q) = 1, \quad (17)$$

which in turn implies

$$\sum_{t \in (\bullet b \cup b \bullet) \cap T'(q)} M_{q,t}(b, i_{\Delta\tau}) = M(b, i_{\Delta\tau}). \quad (18)$$

**Common Differential Dynamics.** The time evolution of

$$\sum_{t \in (\bullet b \cup b \bullet) \cap T'(q)} M_{q,t}(b, \tau)$$

for  $\tau \geq i_{\Delta\tau}$ , with initial state

$$M_{q,t}(b, i_{\Delta\tau}),$$

in the time interval  $[i_{\Delta\tau}, (i+1)_{\Delta\tau})$  is defined by

$$d \left( \sum_{t \in (\bullet b \cup b \bullet) \cap T'(q)} M_{q,t}(b, \tau) \right) = \sum_{t \in (\bullet b \cup b \bullet) \cap T'(q)} dM_{q,t}(b, \tau), \quad (19)$$

which is identical to (5) for  $\Delta\tau \rightarrow 0$ .

## References

1. Brooks, S.e.a.: Handbook of Markov Chain Monte Carlo (Chapman & Hall/CRC Handbooks of Modern Statistical Methods). Chapman and Hall/CRC (2011)
2. Causon, D.M., Mingham, C.G.: Introductory Finite Volume Methods for PDEs. Ventus Publishing ApS (2011)
3. Ciardo, G., Nicol, D., Trivedi, K.S.: Discrete-event simulation of fluid stochastic Petri nets. IEEE Transactions on Software Engineering 25, 207 – 217 (1997)
4. Haar, S., Theissing, S.: A hybrid-dynamical model for passenger-flow in transportation systems. In: 5th IFAC Conference on Analysis and Design of Hybrid Systems (2015)
5. Haar, S., Theissing, S.: Forecasting Passenger Loads in Transportation Networks (2016), <https://hal.inria.fr/hal-01259585>, working paper
6. Haar, S., Theissing, S.: Predicting traffic load in public transportation networks (2016), <https://hal.archives-ouvertes.fr/hal-01286476>, working paper
7. Hoogerheide, L., Kaashoek, J., van Dijk, H.: Functional approximations to posterior densities: a neural network approach to efficient sampling (Dec 2002), <http://hdl.handle.net/1765/1727>
8. MacKay, D.J.C.: Introduction to monte carlo methods. In: Proceedings of the NATO Advanced Study Institute on Learning in Graphical Models (1998)
9. Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T.: Numerical Recipes in C: The Art of Scientific Computing, Second Edition. Cambridge University Press (1992)