

## ERROR BOUNDS FOR KALMAN FILTERS ON TRAFFIC NETWORKS

YE SUN

University of Illinois at Urbana Champaign  
205 N. Mathews Ave.  
Urbana, IL 61801, USA

DANIEL B. WORK

Vanderbilt University  
1025 16th Ave. S., Suite 102  
Nashville, TN 37212, USA

**ABSTRACT.** This work analyzes the estimation performance of the *Kalman filter* (KF) on transportation networks with junctions. To facilitate the analysis, a hybrid linear model describing traffic dynamics on a network is derived. The model, referred to as the *switching mode model for junctions*, combines the discretized *Lighthill-Whitham-Richards partial differential equation* with a junction model. The system is shown to be unobservable under nearly all of the regimes of the model, motivating attention to the estimation error bounds in these modes. The evolution of the estimation error is investigated via exploring the interactions between the update scheme of the KF and the intrinsic physical properties embedded in the traffic model (e.g., conservation of vehicles and the flow–density relationship). It is shown that the state estimates of all the cells in the traffic network are ultimately bounded inside a physically meaningful interval, which cannot be achieved by an open-loop observer.

**1. Introduction.** The unprecedented growth of sensing and computational capabilities have advanced the development of real-time traffic estimation techniques. However, many estimation algorithms proposed to monitor traffic conditions are only verified through experiments (e.g., [5, 13, 26, 27, 42, 43]), while a theoretical analysis on the estimator performance is lacking. This is mainly due to the complexity of the models describing traffic (e.g., non-linearity and non-differentiability) [2], and, more importantly, the non-observability [30, 37] of the systems subject to conservation laws. When a system is not observable, the available sensor measurements (in conjunction with the model describing system dynamics) are insufficient to correctly reconstruct the full state to be estimated. In traffic estimation problems, the non-observability issue is inevitable when a shock exists and the sensors cannot measure every state variable in the road network of interest.

This work aims at addressing the issue of the lack of a theoretical performance analysis for traffic estimation problems under unobservable scenarios occurring at

---

2010 *Mathematics Subject Classification.* Primary: 93B07; Secondary: 35L65, 93C20.

*Key words and phrases.* network traffic flow, junction models, Kalman filtering.

This material is based upon work supported by the National Science Foundation under Grant No. CMMI-1351717.

junctions. Note that in classical estimation / filtering theory, an unobservable system is very likely to result in an estimation error that diverges [1, 17]. Nevertheless, in this work we derive theoretical error bounds for traffic estimation on unobservable transportation networks (with junctions), where the state estimate is provided by the *Kalman Filter* (KF). Specifically, when the system is not observable, the error bounds are derived through leveraging the intrinsic properties of traffic models (e.g., mass conservation and the flow-density relationship), and exploring the interactions between these intrinsic properties and the update scheme in the Kalman filtering algorithm.

**1.1. Related work.** The classical conservation law describing the evolution of traffic density on a road link (i.e., a road section without junctions) is the *Lighthill-Whitham-Richards partial differential equation* (LWR PDE) [25, 32]. The *cell transmission model* (CTM) [7, 8, 23] is a discretization of the LWR PDE using the *Godunov scheme* [11]. In [30, 35], the CTM is transformed to a hybrid linear system known as the *switching mode model* (SMM) that switches among different linear modes, and the observability of each mode is analyzed.

To extend the traffic model on road links to networks, a model is needed to describe how the traffic exiting the road links on the upstream side of a junction is received by the road links on the downstream side of the junction. A well known issue is that the conservation of vehicles across the junction is insufficient to uniquely define the flows at the junction. To address this issue, a number of junction models [8–10, 12, 14, 15, 18–20, 22] have been proposed via additional rules governing the distribution or priority of the flows on different road links.

In parallel to the ongoing development of traffic models, a number of sequential traffic estimation algorithms have been proposed to integrate model predictions with real-time sensor measurements. For example, the *mixture Kalman Filter* is applied to the SMM in [35] to estimate traffic densities for ramp metering. The *parallelized particle filters* and the *parallelized Gaussian sum particle filter* are designed in [27] for computational scalability. In [40], an *efficient multiple model particle filter* is proposed for joint traffic estimation and incident detection on freeways. Other treatments of traffic estimation include [5, 26, 29, 38, 41–43]. A comprehensive survey of sequential estimation techniques for traffic models can be found in [34].

Although many traffic estimation algorithms proposed in the existing literature are verified experimentally, very few theoretical results exist that analyze the performance of traffic estimators (e.g., bounds on the estimation error), especially under unobservable scenarios. The main results to date are as follows. In [16], the KF is applied on a Gaussian approximation of a stochastic traffic model, and the stochastic observability of the system is proved. To ensure observability of the system, a warm-up period is required where the initial conditions are restricted to be free-flow traffic conditions. In [6], the local observability around an equilibrium traffic state is studied using a Lagrangian formulation of the traffic model. The authors of [28] prove the performance of a noise-free Luenberger observer for traffic estimation based on the SMM, which is the first work to provide theoretical performance analysis for any traffic estimator under both observable and unobservable scenarios. The Luenberger observer in [28] discards measurement feedback in unobservable modes, ensuring the spatial integral of the estimation error is conserved. Similarly in [39], the estimator runs an open-loop predictor under unobservable cases to ensure that the estimation error does not diverge. Although conserving the spatial integral of the estimation error, it is illustrated in [36] that dropping measurement feedback

can lead to physically unreasonable estimates. In our earlier work [37], we showed that incorporating measurement feedback under unobservable scenarios is essential in maintaining physically reasonable estimates, i.e., due to the interactions of the model prediction and the measurement feedback in the filtering algorithm, the mean estimates of all state variables are guaranteed to be ultimately bounded inside the physically meaningful interval under unobservable modes. Note that the results in [6, 16, 28, 37, 39] are restricted to road stretches without considering junction dynamics, which is the focus of the work in this article.

**1.2. Contributions and outline of the article.** The main contribution of this article is the theoretical analysis of the estimation error bounds when using the KF to estimate traffic densities on transportation networks with junctions. To support a performance study of the KF, we first propose a switched linear system to model the evolution of traffic densities on road networks with junctions, namely the *switching mode model for junctions* (SMM-J). The SMM-J is an extension of the SMM, in the sense that the SMM-J describes the traffic density evolution on a road section with a junction, while the SMM only considers one-dimensional (i.e., without junctions) road sections. The SMM-J combines a switched linear system representation of the CTM with the junction model developed in [24], however, other junction models can also be incorporated in a similar fashion to derive linear traffic models on junctions. We also provide the observability result on each mode of the SMM-J, and show that nearly all modes are unobservable due to the irreversibility of the conservation laws in the presence of shocks and junctions. Compared to the one-dimensional road sections, the issue of non-observability is more frequently encountered when junctions exist, motivating attentions to the error bounds under unobservable scenarios. Next, we prove the estimation error properties of the KF that uses the SMM-J to estimate an unobservable freeway section with a junction inside. Although an unobservable system typically results in diverging estimation errors [17], we show that by combining the update scheme of the KF with the physical properties embedded in the traffic model, the following properties can be derived for traffic estimation under unobservable modes:

1. For an unobservable road section, the Kalman gain has bounded infinity norm. This is a necessary condition to ensure bounded mean estimation error for the KF.
2. When a road section stays in an unobservable mode, the mean estimate of each state variable is ultimately bounded inside a physically meaningful interval, thus the mean estimation error of the entire state is also ultimately bounded.

This work complements our previous work studying the performance of traffic estimators that use the SMM to monitor unobservable one-dimensional freeway traffic conditions [37]. Thus, when estimating the traffic conditions in a large-scale road network, we can partition the traffic network into local sections which are either one-dimensional, or having a junction inside. The traffic condition in each local section is estimated by a local estimator based on the KF and the SMM-J (or the SMM). Under this distributed computing manner, the estimation errors for the sections without junctions reside below the bounds derived in [37], and the error bounds studied in this work are used to justify the estimation accuracy in the sections with junctions.

This work is organized as follows. Section 2 reviews the KF and its error properties under observable and unobservable systems, and briefly summarizes the CTM.

Section 3 introduces the SMM-J, its observability under different modes, and the properties of the state transition matrices that reflect the intrinsic physical properties of the traffic model. The properties of the state transition matrices of the SMM-J are applied in Section 4, where we prove the boundedness of the Kalman gain and the ultimate bound of the mean estimation error under unobservable scenarios. Finally, some concluding remarks are provided in Section 5.

*Notations.* Let  $I_n$  and  $\mathbf{0}_{n,m}$  be the  $n \times n$  identity matrix and the  $n \times m$  zero matrix, respectively. The subscripts of  $I_n$  and  $\mathbf{0}_{n,m}$  are sometimes omitted, when dimensionality is clear from the context. Denote as  $E[\cdot]$  the expectation operator, and the “bar” symbol is used to denote the expected value of random vector  $\mathcal{X}$ , i.e.,  $\bar{\mathcal{X}} = E[\mathcal{X}]$ . The symbol  $^\top$  is used to denote the transpose operator.

**2. Preliminaries.** In this section, we first review the Kalman filter and the properties of its error covariance as well as mean estimation error under observable and unobservable scenarios. We also provide a brief overview of the cell transmission model used to construct the SMM-J.

**2.1. Kalman filter.** In this subsection, we briefly review the KF and introduce necessary notations in this article. Consider a general linear time-varying system

$$x_{k+1} = A_k x_k + u_k + w_k; \quad x_k \in \mathbb{R}^n; \quad (1)$$

$$z_k = H_k x_k + v_k; \quad z_k \in \mathbb{R}^m; \quad (2)$$

where  $x_k$  and  $z_k$  are the state vector and sensor measurement vector at time  $k \in \mathbb{N}$ , respectively. The matrices  $A_k$  and  $H_k$  are the state transition matrix and the observation matrix at time  $k$ . The term  $u_k$  in (1) is a deterministic system input. The noise terms  $w_k \sim N(\mathbf{0}; Q_k)$  and  $v_k \sim N(\mathbf{0}; R_k)$  are the white Gaussian model and measurement noise, where  $Q_k$  and  $R_k$  denote the model and the measurement error covariance matrices at time  $k$ . Given the sensor data up to time  $k$  denoted by  $Z_k = \{z_0, \dots, z_k\}$ , the *prior estimate* and *posterior estimate* of the state can be expressed as  $\hat{x}_{k|k-1} = E[x_k | Z_{k-1}]$  and  $\hat{x}_{k|k} = E[x_k | Z_k]$ , respectively. Let  $\hat{x}_{k|k-1} = \hat{x}_{k|k-1}$  and  $\hat{x}_{k|k} = \hat{x}_{k|k}$  denote the prior and posterior estimation errors. The estimation error covariance matrices associated with  $\hat{x}_{k|k-1}$  and  $\hat{x}_{k|k}$  are given by  $\Gamma_{k|k-1} = E[\hat{x}_{k|k-1} \hat{x}_{k|k-1}^\top | Z_{k-1}]$  and  $\Gamma_{k|k} = E[\hat{x}_{k|k} \hat{x}_{k|k}^\top | Z_k]$ . The KF sequentially computes  $\hat{x}_{k|k}$  from  $\hat{x}_{k-1|k-1}$  as follows:

$$\text{Prediction:} \quad \begin{aligned} \hat{x}_{k|k-1} &= A_{k-1} \hat{x}_{k-1|k-1} + u_k \\ \Gamma_{k|k-1} &= A_{k-1} \Gamma_{k-1|k-1} A_{k-1}^\top + Q_{k-1}; \end{aligned} \quad (3)$$

$$\text{Correction:} \quad \begin{aligned} \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k (z_k - H_k \hat{x}_{k|k-1}) \\ \Gamma_{k|k} &= \Gamma_{k|k-1} - K_k H_k \Gamma_{k|k-1} \\ K_k &= \Gamma_{k|k-1} H_k^\top (R_k + H_k \Gamma_{k|k-1} H_k^\top)^{-1}; \end{aligned} \quad (4)$$

In (4), the matrix  $K_k$  is denoted as the *Kalman gain* at time  $k$ . Note that for all  $k$ , the state estimates  $\hat{x}_{k|k-1}$  and  $\hat{x}_{k|k}$  are random vectors. The mean posterior estimate and the mean posterior estimation error<sup>1</sup> are denoted as  $\bar{x}_{k|k}$  and  $\bar{e}_{k|k}$ , respectively.

<sup>1</sup>For the remainder of this article, the term state estimates/estimation errors/error covariance refers to the posterior estimates, unless specified otherwise.

*Observability.* The discrete system (1)-(2) is *uniformly completely observable* if there exists a positive integer  $N$  and positive constants  $\alpha, \beta$  such that

$$I_n < I_{k;k-N} < I_n; \quad \text{for all } k \geq N, \quad (5)$$

where  $I_{k_1;k_0}$  is defined as the information matrix for time interval  $k \geq [k_0; k_1]$ :

$$I_{k_1;k_0} = \sum_{k=k_0}^{k_1} \Xi_{k;k_1}^\top H_k^\top R_k^{-1} H_k \Xi_{k;k_1};$$

where

$$\Xi_{k;k_1} = \prod_{i=k}^{k_1-1} A^{-1}; \quad \text{and} \quad \Xi_{k_1;k} = \Xi_{k;k_1}^{-1} = \prod_{i=k_1-1}^k A; \quad (6)$$

The observability of a system characterizes whether the sensor measurements of the system is sufficient for the KF to correctly estimate the state vector. Given the positive definiteness of the information matrix, the observability grammian matrix linking the the initial state  $x_0$  to the sensor measurements up to time  $N$  (i.e.,  $Z_N$ ) is invertible. The next lemma states the boundedness (both from below and above) of the estimation error covariance and the convergence (to zero) of the estimation error, when using the KF (3)-(4) to estimate a uniformly completely observable system.

**Lemma 1** (Chapter 7.6 in [17]). *If the dynamical system (1)-(2) is uniformly completely observable and the following conditions hold:*

- (C.1): *there exist positive constants  $q_1, q_2, r_1$  and  $r_2$  such that  $q_1 I_n < Q_k < q_2 I_n$  and  $r_1 I_m < R_k < r_2 I_m$  for all  $k$ ;*
- (C.2): *the initial error covariance is positive definite, i.e.,  $\Gamma_{0|0} > \mathbf{0}$ ;*
- (C.3): *the state transition matrix  $A_k$  is nonsingular for all  $k$ ,*

*then there exist positive constants  $c_1 > 0$  and  $c_2 > 0$  such that the error covariance of the KF (3)-(4) satisfies*

$$c_1 I_n < \Gamma_{k|k} < c_2 I_n; \quad \text{for all } k \geq 0;$$

*Moreover, there exists positive constants  $a > 0$  and  $0 < q < 1$  such that the 2-norm<sup>2</sup> of the mean estimation error satisfies*

$$\|\bar{x}_{k|k}\|_2 < a q^k \|\bar{x}_{0|0}\|_2; \quad \text{for all } k \geq 0;$$

When system (1)-(2) is not observable, the mean estimation error  $\bar{x}_{k|k}$  will diverge, unless the unobservable part of the state is bounded or converges to zero automatically. In [Appendix.1](#), we present an example illustrating the evolution of the mean estimation error given by the KF when tracking an unobservable system.

**2.2. Cell transmission model.** The classical conservation law describing the evolution of traffic density  $\rho(t; x)$  on a road at location  $x$  and time  $t$  is the *Lighthill-Whitham-Richards partial differential equation* (LWR PDE) [25, 32]:

$$\partial_t \rho + \partial_x F(\rho) = 0; \quad (7)$$

The function  $F(\rho) = v(\rho) \rho$  is called the flux function, where  $v(\rho)$  is an empirical velocity function used to close the model.

The *cell transmission model* (CTM) [7, 8, 23] is a discretization of (7) using a Godunov scheme [11]. Consider a uniformly sized discretization grid defined by

<sup>2</sup>For the remainder of this article, we denote as  $\|\cdot\|_2$  the 2-norm of a matrix or a vector.

a space step  $\Delta x > 0$  and a time step  $\Delta t > 0$ . Let  $l$  index the cell defined by  $x \in [l\Delta x; (l+1)\Delta x)$ , and denote as  $\rho_k^l$  the density at time  $k\Delta t$  in cell  $l$ , where  $k \geq \mathbb{N}$  and  $l \geq \mathbb{N}^+$ . Moreover, denote as  $f(\rho_k^{l-1}; \rho_k^l)$  the flux between cell  $l-1$  and  $l$ . In the CTM, the discretized model (7) becomes

$$\rho_{k+1}^l = \rho_k^l + \frac{\Delta t}{\Delta x} (f(\rho_k^{l-1}; \rho_k^l) - f(\rho_k^l; \rho_k^{l+1})); \quad (8)$$

where  $f(\rho_k^{l-1}; \rho_k^l)$  is computed by

$$f(\rho_k^{l-1}; \rho_k^l) = \min \{s(\rho_k^{l-1}); r(\rho_k^l)\}; \quad (9)$$

In (9),  $s(\rho_k^{l-1})$  is the *sending capacity* (i.e., maximum sending flow) of cell  $l-1$  at time  $k$ , which is a function of  $\rho_k^{l-1}$ , and  $r(\rho_k^l)$  is the *receiving capacity* (i.e., maximum receiving flow) of cell  $l$  at time  $k$ , which is a function of  $\rho_k^l$ . Note that if the *Courant-Friedrichs-Lewy* (CFL) condition is satisfied, the solution of the CTM converges in  $L^1$  to the weak solution of the LWR PDE as  $\Delta x \rightarrow 0$  [33].

**Remark 1.** Note that the terminologies sending capacity and receiving capacity are equivalent to the notions of demand and supply. Both sending/receiving and demand/supply terminologies are widely used in the traffic community, with the former introduced in [7, 8], and the latter introduced in [23]. In this work we use the sending/receiving terminology, which is consistent throughout the article, and is also consistent with our previous work [37].

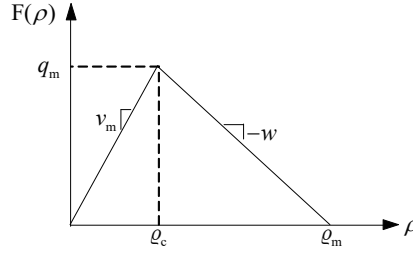


FIGURE 1. The triangular fundamental diagram in (10).

The flux function [8] used in this work is the triangular flux function (shown in Figure 1) given by

$$F(\rho) = \begin{cases} v_m \rho & \text{if } \rho \in [0; \rho_c] \\ w(\rho_m - \rho) & \text{if } \rho \in [\rho_c; \rho_m], \end{cases} \quad (10)$$

where  $w = \frac{\rho_c v_m}{\rho_m - \rho_c}$ ,  $v_m$  denotes the *free flow speed*, and  $\rho_m$  denotes the *maximum density*. The parameter  $\rho_c$  is the *critical density* at which the maximum flux is realized. For the triangular fundamental diagram, the flux function has different slopes in *free flow* ( $0 < \rho < \rho_c$ ) and *congestion* ( $\rho_c < \rho < \rho_m$ ). In freeflow, the slope is  $v_m$ , and in congestion, it is  $w$ . Under the triangular flux function, the sending and receiving capacities are determined by:

$$s(\rho) = \begin{cases} v_m & \text{if } \rho \in [0; \rho_c] \\ q_m & \text{if } \rho \in [\rho_c; \rho_m] \end{cases} \quad r(\rho) = \begin{cases} q_m & \text{if } \rho \in [0; \rho_c] \\ w(\rho_m - \rho) & \text{if } \rho \in [\rho_c; \rho_m], \end{cases} \quad (11)$$

where  $q_m$  is the maximum flow given by  $q_m = v_m \rho_c$ .

**3. The switching mode model for junctions.** In this section, we derive the hybrid linear model describing traffic density evolution on a road section with a junction, namely the *switching mode model for junctions* (SMM-J). The SMM-J combines a switched linear system representation of the CTM (i.e., the CTM where the flux function is the triangular fundamental diagram) with a junction solver. This section starts with a review of the applied junction solver. Next, we introduce the SMM-J and provide examples regarding its explicit formulas. Finally, the observability of the SMM-J is discussed.

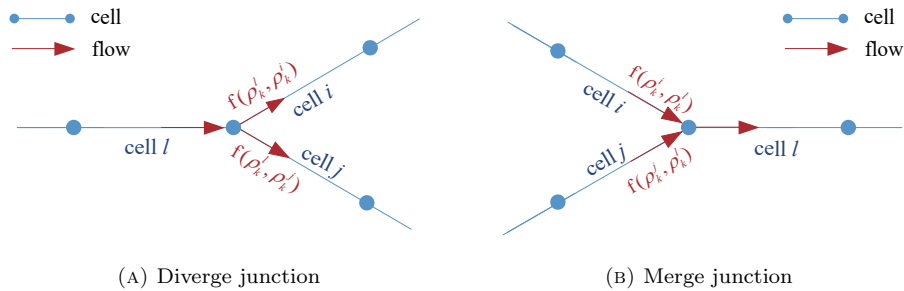


FIGURE 2. A diverge and a merge junction connected by three cells indexed by  $i, j$ , and  $l$ .

**3.1. Junction solver on traffic networks.** This subsection introduces the junction solver proposed in [24]. As shown in Figure 2a, the junction solver computes flux  $f(\begin{smallmatrix} l \\ k \\ i \end{smallmatrix})$  and  $f(\begin{smallmatrix} l \\ k \\ j \end{smallmatrix})$  between the connecting cells  $l, i$  and  $j$  forming a diverge junction. In the merge junction shown in Figure 2b, the junction solver computes  $f(\begin{smallmatrix} i \\ k \\ l \end{smallmatrix})$  and  $f(\begin{smallmatrix} j \\ k \\ l \end{smallmatrix})$  between the connecting cells. This solver is later applied in Section 3.2 to develop the SMM-J.

**3.1.1. Diverge junction.** At a diverge junction in Figure 2a, the junction solver is governed by the following rules:

- (R.1): The mass across the junction is conserved.
- (R.2): The throughput flow  $f(\begin{smallmatrix} l \\ k \\ i \end{smallmatrix}) + f(\begin{smallmatrix} l \\ k \\ j \end{smallmatrix})$ , i.e., the outgoing flow of cell  $l$ , is maximized subject to the maximum flow that can be sent or received on each connecting cell.
- (R.3): The distribution of flux  $f(\begin{smallmatrix} l \\ k \\ i \end{smallmatrix})$  and  $f(\begin{smallmatrix} l \\ k \\ j \end{smallmatrix})$  satisfies  $f(\begin{smallmatrix} l \\ k \\ j \end{smallmatrix}) = \alpha f(\begin{smallmatrix} l \\ k \\ i \end{smallmatrix})$ , where  $\alpha$  is the prescribed distribution parameter that models the routing preference to the downstream cells. When (R.3) conflicts (R.2), i.e., the flow solution that satisfies the distribution equation does not maximize the throughput, then (R.3) is relaxed, such that the solution satisfies (R.2) and minimizes the deviation from the prescribed distribution parameter, e.g.,  $f(\begin{smallmatrix} l \\ k \\ j \end{smallmatrix}) = f(\begin{smallmatrix} l \\ k \\ i \end{smallmatrix}) \alpha$ .

The diverge junction solver is posed as a convex program with a carefully constructed objective function to accommodate the throughput maximization (R.2) and the flow distribution (R.3). The mathematical formula of the diverge junction solver is given below (see [24, Section 3.2] for more details).

**Definition 1** (Convex program for the diverge junction solver). Define the objective function  $J(f_1; f_2)$  as:

$$J(f_1; f_2) = (1 - \alpha)(f_1 + f_2) - \alpha(f_2 - \alpha f_1)^2;$$

where  $0 < \alpha < 1$  and  $\alpha$  is chosen such that  $\frac{\partial J}{\partial f_1} > 0$  and  $\frac{\partial J}{\partial f_2} > 0$ <sup>3</sup>. The conditions on  $\alpha$  is used to prioritize maximizing the throughput  $f(\frac{l}{k}; \frac{i}{k}) + f(\frac{l}{k}; \frac{j}{k})$  (as stated in (R.2)), then minimizing the deviation from the prescribed distribution parameter  $\alpha$  (as stated in (R.3)). The fluxes  $f(\frac{l}{k}; \frac{i}{k})$  and  $f(\frac{l}{k}; \frac{j}{k})$  are obtained by solving the following convex program:

$$f(\frac{l}{k}; \frac{i}{k}); f(\frac{l}{k}; \frac{j}{k}) = \arg \max_{f_1; f_2} J(f_1; f_2)$$

$$\text{s.t. } f_1 \leq r(\frac{i}{k}) \quad (12)$$

$$f_2 \leq r(\frac{j}{k}) \quad (13)$$

$$f_1 + f_2 \leq s(\frac{l}{k}) \quad (14)$$

Figure 3 provides a graphical illustration for the solutions of the convex program defined in Definition 1. The blue vertical solid line denotes the receiving capacity of cell  $i$ , i.e.,  $r(\frac{i}{k})$ , and the blue horizontal solid line denotes the receiving capacity of cell  $j$ , i.e.,  $r(\frac{j}{k})$ . The intercepts of the blue dashed line (with slope -1) denote the sending capacity of cell  $l$ , i.e.,  $s(\frac{l}{k})$ . The shaded area denotes the feasible values of the flux from cell  $l$  to  $i$  and the flux from cell  $l$  to  $j$ , the feasible area is obtained based on (12)-(14). The slope of the black dotted line is the prescribed distribution ratio  $\alpha$ . The fluxes computed by the junction solver in Definition 1 is marked by the red dot, whose horizontal axis and vertical axis values are the obtained  $f(\frac{l}{k}; \frac{i}{k})$  and  $f(\frac{l}{k}; \frac{j}{k})$ , respectively.

According to the convex program in Definition 1, to obtain  $f(\frac{l}{k}; \frac{i}{k})$  and  $f(\frac{l}{k}; \frac{j}{k})$  we need to find the solution point (i.e., the red dot) in Figure 3 that satisfies the following requirements:

The point lies in the shaded feasible area, so that conditions (12)-(14) are satisfied;

The point is as close as possible to the blue dashed line (with slope -1), so that the throughput  $f(\frac{l}{k}; \frac{i}{k}) + f(\frac{l}{k}; \frac{j}{k})$  is maximized; this is due to the fact that the distance  $d_0$  between the point and the blue dashed line is proportional to the disparity between the sending capacity of cell  $l$  and the throughput, i.e.,  $j s(\frac{l}{k}) - (f(\frac{l}{k}; \frac{i}{k}) + f(\frac{l}{k}; \frac{j}{k}))j = \sqrt{2} d_0$  (as illustrated in diverge case I of Figure 3);

Conditioned on prioritizing maximizing the throughput, the point is as close as possible to the black dotted line (with slope  $\alpha$ ), so that the distribution ratio is respected; this means that when maximizing the throughput conflicts with the distribution ratio, the requirement  $f(\frac{l}{k}; \frac{j}{k}) = \alpha f(\frac{l}{k}; \frac{i}{k})$  can be relaxed.

As shown in Figure 3, there are in total three scenarios depending on the values of  $s(\frac{l}{k})$ ,  $r(\frac{i}{k})$  and  $r(\frac{j}{k})$ . The three scenarios are: (i) diverge case I, when  $s(\frac{l}{k}) \leq r(\frac{i}{k}) + r(\frac{j}{k})$ ; (ii) diverge case II, when  $s(\frac{l}{k}) < r(\frac{i}{k}) + r(\frac{j}{k})$ , and the prescribed distribution ratio  $\alpha$  can be followed exactly; and (iii) diverge case III, when  $s(\frac{l}{k}) <$

<sup>3</sup>One possible choice of  $\alpha$  is  $\alpha = \min(1 + 2\sqrt{2}q_m + \epsilon)^{-1}; (1 + 2q_m + \epsilon)^{-1}$ , where  $\epsilon > 0$  can be any positive value.



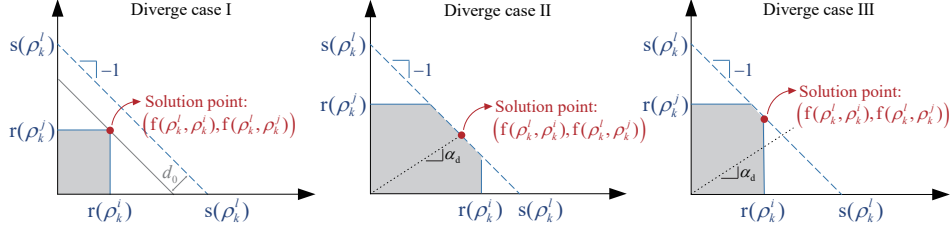


FIGURE 3. Three scenarios in the junction solver [24] for the diverge junction shown in Figure 2a, where cell  $l$  diverges to cell  $i$  and cell  $j$ . The blue vertical (resp. horizontal) solid line denotes the receiving capacity of cell  $i$  (resp. cell  $j$ ). The intercepts of the blue dashed line denote the sending capacity of cell  $l$ . The shaded area denotes the feasible values of the flux from cell  $l$  to  $i$  and the flux from cell  $l$  to  $j$ . The slope of the black dotted line is the prescribed distribution ratio  $d$ . The fluxes computed by the junction solver is marked by the red dot, whose horizontal axis and vertical axis values are the obtained  $f(\frac{l}{k}; \frac{i}{k})$  and  $f(\frac{l}{k}; \frac{j}{k})$ , respectively. Note that in diverge case II and diverge case III, the receiving capacities of cell  $i$  and cell  $j$  are not necessarily smaller than the sending capacity of cell  $l$ , and the graphical illustration of the flux solutions is also applicable for  $r(\frac{i}{k}) = s(\frac{l}{k})$  and/or  $r(\frac{j}{k}) = s(\frac{l}{k})$ .

$r(\frac{i}{k}) + r(\frac{j}{k})$ , but (due to throughput maximization) the prescribed distribution ratio  $d$  cannot be followed exactly.

Under diverge case I, the fluxes  $f(\frac{l}{k}; \frac{i}{k})$  and  $f(\frac{l}{k}; \frac{j}{k})$  are computed by:

$$f(\frac{l}{k}; \frac{i}{k}) = r(\frac{i}{k}); \quad f(\frac{l}{k}; \frac{j}{k}) = r(\frac{j}{k}); \quad (15)$$

Under diverge case II, the fluxes  $f(\frac{l}{k}; \frac{i}{k})$  and  $f(\frac{l}{k}; \frac{j}{k})$  are given as follows:

$$f(\frac{l}{k}; \frac{i}{k}) = \frac{1}{d+1}s(\frac{l}{k}); \quad f(\frac{l}{k}; \frac{j}{k}) = \frac{d}{d+1}s(\frac{l}{k}); \quad (16)$$

Under diverge case III, the junction solver first finds the two vertices of the shaded area that lie on the dashed line, and next define the point closer to the dotted line as the flux solutions  $f(\frac{l}{k}; \frac{i}{k})$  and  $f(\frac{l}{k}; \frac{j}{k})$ . Hence, depending on the magnitude of  $s(\frac{l}{k})$ ,  $r(\frac{i}{k})$  and  $r(\frac{j}{k})$ , the solutions could be either

$$f(\frac{l}{k}; \frac{i}{k}) = r(\frac{i}{k}); \quad f(\frac{l}{k}; \frac{j}{k}) = s(\frac{l}{k}) - r(\frac{i}{k}); \quad (17)$$

or

$$f(\frac{l}{k}; \frac{i}{k}) = s(\frac{l}{k}) - r(\frac{j}{k}); \quad f(\frac{l}{k}; \frac{j}{k}) = r(\frac{j}{k}); \quad (18)$$

Note that the diverge case III shown in Figure 3 illustrates the flux solutions presented in (17).

**Remark 2.** The diverge junction solver described above is a non-*First-In-First-Out* (FIFO) model [8, 10]. The FIFO diverge model maximizes the outgoing flow of cell  $l$  subject to the distribution ratio  $f(\frac{l}{k}; \frac{j}{k}) = d f(\frac{l}{k}; \frac{i}{k})$ . Although the FIFO model circumvents the conflicts between throughput maximization and flow distribution, it produces unrealistic solutions in some circumstances. Several diverge junction

models have been proposed to resolve this issue [14, 20, 22]. In the same spirit of these models, the diverge junction solver applied in this article is developed to produce similar traffic condition dependent solutions without introducing additional complexity on the traffic dynamics [24]. As a related note, the results proved in this article can be extended to FIFO models with minor changes to the proof.

3.1.2. *Merge junction.* At a merge junction in Figure 2b, the junction solver conserves mass, and maximizes the throughput while minimizing the deviation from a prescribed priority parameter  $\rho_p$  denoting the flow assignment ratio  $f(\frac{j}{k}; \frac{l}{k}) = \rho_p f(\frac{l}{k}; \frac{l}{k})$ . This priority equation is relaxed if it conflicts with flow maximization. The reader is referred to [10, 12] for a detailed description of the merge model.

The structure of the diverge and merge models are similar, in the sense that both maximize the throughput while minimizing the deviation from the prescribed distribution or priority parameters. Therefore, the remainder of this article focuses on deriving the linear traffic model and analyzing the performance of the KF on the road network with a diverge junction. The analysis can be transferred to the merge case via combining the merge junction solver with the switched linearized representation of the CTM, exploring the properties of the resulting state transition matrices as in the diverge case, and analyzing the effect of these properties on the boundedness of the Kalman gain and the mean estimation error.

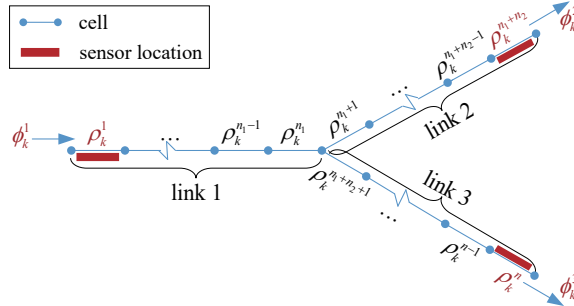


FIGURE 4. A local section with  $n$  cells, three links and a junction.

3.2. **The switching mode model for junctions.** As stated in Section 1.2, when monitoring traffic on large-scale networks, the road network is partitioned into local sections which are either one-dimensional, or having a junction inside. The traffic states on the one-dimensional local sections evolve according to the SMM, and the SMM-J is used to describe the evolution of traffic states on local sections with junctions.

As shown in Figure 4, consider a local section with  $n$  cells, three links and a junction. The number of cells on each link is  $n_1$ ,  $n_2$ , and  $n_3$ , respectively, with  $n_1 + n_2 + n_3 = n$ . The state variable at time  $k \geq N$  is constructed as  $x_k = [\rho_k^1; \dots; \rho_k^{n_1}; \rho_k^{n_1+1}; \dots; \rho_k^{n_1+n_2}; \rho_k^{n_1+n_2+1}; \dots; \rho_k^n]^T$ . As a common treatment [28, 30, 35, 38, 39], the boundary flows, denoted by  $\phi_k^1$ ,  $\phi_k^2$ , and  $\phi_k^3$ , are considered to be deterministic system inputs (please refer to [3] for the concept of using ghost cells to compute boundary flows using boundary state measurements). The SMM-J describes the evolution of  $x_k$  using a switched linear dynamics, and is derived under the following assumptions:

- (A.1): For each local section, there is at most one transition between freeflow and congestion in each of the three links.
- (A.2): The three connecting cells forming the junction (i.e., cell  $n_1$ ,  $n_1 + 1$  and  $n_1 + n_2 + 1$  in Figure 4) are either all in freeflow or all in congestion.

Assumption (A.1) is practically meaningful since the local sections are usually partitioned sufficiently short such that at most one queue is growing or dissipating within each link, which is also a commonly used assumption in the SMM [28, 30, 35, 37, 39]. Assumption (A.2) is imposed to simplify the model by reducing the number of modes considered. Note that when (A.2) is relaxed, the number of modes is greatly increased, but without yielding new insights into the estimation performance at junctions. The analysis in this work still holds, the only difference is to consider an enormously increased number of modes.

Given the assumptions stated above, on each local section with a junction, the SMM-J may switch among 32 modes (defined in Table 1) depending on (i) the freeflow/congestion status of the boundary cells and the connecting cells near the junction, and (ii) the flux solution of the junction solver characterized by the three scenarios shown in Figure 3.

In the SMM-J, the density update scheme for the interior cells in each link (i.e., cells indexed by  $l \in \{1, 2, \dots, n_1-1, n_1, n_1+1, \dots, n_1+n_2-1, n_1+n_2, n_1+n_2+1, \dots, n-1, n\}$ ) is given by (8), where the flow between two adjacent cells is computed according to (9). For the three boundary cells, their density updates are given as follows:

$$\begin{aligned} n_{k+1}^1 &= n_k^1 + \frac{\Delta t}{\Delta X} f(n_k^1; n_k^2); \\ n_{k+1}^{n_1+n_2} &= n_k^{n_1+n_2} + \frac{\Delta t}{\Delta X} f(n_k^{n_1+n_2-1}; n_k^{n_1+n_2}); \\ n_{k+1}^n &= n_k^n + \frac{\Delta t}{\Delta X} f(n_k^{n-1}; n_k^n); \end{aligned}$$

where  $f(n_k^1; n_k^2)$ ,  $f(n_k^{n_1+n_2-1}; n_k^{n_1+n_2})$  and  $f(n_k^{n-1}; n_k^n)$  are also obtained from (9). The density update scheme for the three cells near the junction reads:

$$\begin{aligned} n_{k+1}^{n_1} &= n_k^{n_1} + \frac{\Delta t}{\Delta X} f(n_k^{n_1-1}; n_k^{n_1}) - f(n_k^{n_1}; n_k^{n_1+1}) + f(n_k^{n_1}; n_k^{n_1+n_2+1}); \\ n_{k+1}^{n_1+1} &= n_k^{n_1+1} + \frac{\Delta t}{\Delta X} f(n_k^{n_1}; n_k^{n_1+1}) - f(n_k^{n_1+1}; n_k^{n_1+2}); \\ n_{k+1}^{n_1+n_2+1} &= n_k^{n_1+n_2+1} + \frac{\Delta t}{\Delta X} f(n_k^{n_1}; n_k^{n_1+n_2+1}) - f(n_k^{n_1+n_2+1}; n_k^{n_1+n_2+2}); \end{aligned}$$

where  $f(n_k^{n_1-1}; n_k^{n_1})$ ,  $f(n_k^{n_1+1}; n_k^{n_1+2})$ , and  $f(n_k^{n_1+n_2+1}; n_k^{n_1+n_2+2})$  are computed by (9), and the flows between neighboring cells (i.e.,  $f(n_k^{n_1}; n_k^{n_1+1})$  and  $f(n_k^{n_1}; n_k^{n_1+n_2+1})$ ) are computed based on the junction solver discussed in Section 3.1.

When applying the triangular fundamental diagram (10) to compute the flow across the cells, the traffic state  $n_k$  in a local section evolves with linear dynamics in each mode stated in Table 1, forming a hybrid system:

$$n_{k+1} = A_k n_k + B_k \mathbf{1} q_m + B_k^q \mathbf{1} q_m + B_k n_k; \quad (19)$$

where  $\mathbf{1}$  is the vector of all ones, the vector  $n_k = (n_k^1; n_k^2; n_k^3)^\top$ , and  $A_k \in \mathbb{R}^{n \times n}$ ,  $B_k \in \mathbb{R}^{n \times n}$ ,  $B_k^q \in \mathbb{R}^{n \times n}$ ,  $B_k \in \mathbb{R}^{n \times 3}$  are constructed based on the current mode of the local section, the locations of the shocks and expansion fans (if they exist), and the moving direction of the shocks.

TABLE 1. Mode definition and observability of the SMM-J

Mode	F/C <sup>1</sup> status of cell(s)				Transition <sup>3</sup> on link			Diverge case	Observability <sup>4</sup>
	1	$n_1 + n_2$	$n$	near junction <sup>2</sup>	1	2	3		
1	F	F	F	F	none	none	none	II	O
2	F	F	F	C	Sh.	Ep.	Ep.	I	U
3	F	F	F	C	Sh.	Ep.	Ep.	II	U
4	F	F	F	C	Sh.	Ep.	Ep.	III	U
5	C	C	C	C	none	none	none	I	U
6	C	C	C	C	none	none	none	II	U
7	C	C	C	C	none	none	none	III	U
8	C	C	C	F	Ep.	Sh.	Sh.	II	U
9	F	C	C	C	Sh.	none	none	I	U
10	F	C	C	C	Sh.	none	none	II	U
11	F	C	C	C	Sh.	none	none	III	U
12	F	C	C	F	none	Sh.	Sh.	II	U
13	C	C	F	C	none	none	Ep.	I	U
14	C	C	F	C	none	none	Ep.	II	U
15	C	C	F	C	none	none	Ep.	III	U
16	C	C	F	F	Ep.	Sh.	none	II	U
17	C	F	C	C	none	Ep.	none	I	U
18	C	F	C	C	none	Ep.	none	II	U
19	C	F	C	C	none	Ep.	none	III	U
20	C	F	C	F	Ep.	none	Sh.	II	U
21	C	F	F	F	Ep.	none	none	II	O
22	C	F	F	C	none	Ep.	Ep.	I	U
23	C	F	F	C	none	Ep.	Ep.	II	U
24	C	F	F	C	none	Ep.	Ep.	III	U
25	F	C	F	F	none	Sh.	none	II	U
26	F	C	F	C	Sh.	none	Ep.	I	U
27	F	C	F	C	Sh.	none	Ep.	II	U
28	F	C	F	C	Sh.	none	Ep.	III	U
29	F	F	C	F	none	none	Sh.	II	U
30	F	F	C	C	Sh.	Ep.	none	I	U
31	F	F	C	C	Sh.	Ep.	none	II	U
32	F	F	C	C	Sh.	Ep.	none	III	U

<sup>1</sup> “F” and “C” stand for freeflow and congestion, respectively.

<sup>2</sup> Cells indexed by  $n_1$ ,  $n_1 + 1$  and  $n_1 + n_2 + 1$ .

<sup>3</sup> “Sh.” and “Ep.” stand for shock (i.e., transition from freeflow to congestion) and expansion fan (i.e., transition from congestion to freeflow), respectively.

<sup>4</sup> “O” stands for uniformly completely observable and “U” stands for unobservable. Note that the observability results are derived under sensor locations shown in Figure 4.

The next two examples demonstrate the construction of matrices  $A_k$ ,  $B_k$ ,  $B_k^q$  and  $B_k$ . Specifically, the construction of matrices  $A_k$  provides insights on the properties of the state transition matrices of the SMM-J that reflect the intrinsic physical properties of the traffic model. These properties are critical in proving

the estimation error bounds of the KF on traffic networks. Before showing the examples, we first introduce some notation which will be used as elements of the matrices to be constructed. For  $p \in \mathbb{N}^+$ , define  $\Theta_p \in \mathbb{R}^{p \times p}$  and  $\Delta_p \in \mathbb{R}^{p \times p}$  by their  $(i;j)^{\text{th}}$  entries as

$$\Theta_p(i;j) = \begin{cases} < 1 & \frac{v_m \Delta t}{\Delta x} & \text{if } i = j \\ \frac{v_m \Delta t}{\Delta x} & & \text{if } i = j + 1 \\ 0 & & \text{otherwise,} \end{cases} \quad (20)$$

$$\Delta_p(i;j) = \begin{cases} < 1 & \frac{w \Delta t}{\Delta x} & \text{if } i = j \\ \frac{w \Delta t}{\Delta x} & & \text{if } i = j - 1 \\ 0 & & \text{otherwise.} \end{cases} \quad (21)$$

For  $p_1 \in \mathbb{N}^+$ ,  $p_2 \in \mathbb{N}^+$ ,  $p_3 \in \mathbb{N}^+$  and  $p_4 \in \mathbb{N}^+$ , define  $E_{p_1; p_2}^{p_3; p_4} \in \mathbb{R}^{p_1 \times p_2}$  as the  $p_1 \times p_2$  matrix with all entries zero except its  $(p_3; p_4)^{\text{th}}$  entry, which is one. Explicitly,

$$E_{p_1; p_2}^{p_3; p_4}(i;j) = \begin{cases} 1 & \text{if } i = p_3 \text{ and } j = p_4 \\ 0 & \text{otherwise.} \end{cases} \quad (22)$$

Moreover, define  $\tilde{\Theta}_p \in \mathbb{R}^{p \times p}$  and  $\tilde{\Delta}_p \in \mathbb{R}^{p \times p}$  as:

$$\tilde{\Theta}_p = \begin{cases} < & \Theta_{p-1} & \mathbf{0}_{p-1;1} & & \\ & \frac{v_m \Delta t}{\Delta x} E_{1;p-1}^{1;p-1} & 1 & & \\ & 1 & & & \end{cases} \quad \begin{cases} \text{if } p = 2, \\ \text{if } p = 1, \end{cases}$$

and

$$\tilde{\Delta}_p = \begin{cases} < & 1 & \frac{w \Delta t}{\Delta x} E_{1;p-1}^{1;1} & & \\ & \mathbf{0}_{p-1;1} & \Delta_{p-1} & & \\ & 1 & & & \end{cases} \quad \begin{cases} \text{if } p = 2, \\ \text{if } p = 1. \end{cases}$$

**Example 1** (System dynamics of the SMM-J under Mode 1). Consider the local section in Figure 4 where all the cells are in freeflow. Within a single link, the flux between two neighboring cells indexed by  $i$  and  $i+1$  is given by  $f(\frac{i}{k}; \frac{i+1}{k}) = v_m \frac{i}{k}$ . At the junction, it holds that  $s(\frac{n_1}{k}) = v_m \frac{n_1}{k} < r(\frac{n_1+1}{k}) + r(\frac{n_1+n_2+1}{k}) = 2q_m$ . Also note that  $s(\frac{n_1}{k}) - r(\frac{n_1+1}{k}) = q_m$  and  $s(\frac{n_1}{k}) - r(\frac{n_1+n_2+1}{k}) = q_m$ , the distribution ratio  $d$  can be followed exactly. Hence, the state is under Mode 1 at time  $k$ , and the junction solver computes fluxes  $f(\frac{n_1}{k}; \frac{n_1+1}{k})$  and  $f(\frac{n_1}{k}; \frac{n_1+n_2+1}{k})$  according to diverge case II (16) as follows:

$$f(\frac{n_1}{k}; \frac{n_1+1}{k}) = \frac{\frac{n_1}{k} v_m}{1 + d}; \quad f(\frac{n_1}{k}; \frac{n_1+n_2+1}{k}) = \frac{d \frac{n_1}{k} v_m}{1 + d}.$$

Substituting the flows computed above into the update scheme of the traffic density on each cell, it follows that the explicit forms of  $A_k$ ,  $B_k$ ,  $B_k^q$ , and  $B_k$  in (19) are

$$\begin{aligned} A_k &= \begin{matrix} \text{B} \\ \text{O} \end{matrix} \begin{matrix} \Theta_{n_1} \\ \frac{v_m \Delta t}{(1+d)\Delta x} E_{n_2;n_1}^{1;n_1} \\ \frac{d v_m \Delta t}{(1+d)\Delta x} E_{n_3;n_1}^{1;n_1} \end{matrix} \tilde{\Theta}_{n_2} \begin{matrix} \text{C} \\ \text{A} \end{matrix}; \quad B_k = B_k^q = \mathbf{0}_{n;n}; \quad (23) \\ B_k &= \frac{\Delta t}{\Delta x} \begin{matrix} E_{n;3}^{1;1} & E_{n;3}^{n_1+n_2;2} & E_{n;3}^{n;3} \end{matrix} \end{aligned}$$

in Mode 1. Note that in the above definitions and for the remainder of this subsection, blocks in the matrices which are left blank are zeros everywhere. 2

**Example 2** (System dynamics of the SMM-J under Modes 2-4). Consider the local section in Figure 4 where the three boundary cells indexed by 1,  $n_1 + n_2$  and  $n$  are all in freeflow, and the three cells near the junction, i.e., the cells indexed by  $n_1$ ,  $n_1 + 1$  and  $n_1 + n_2 + 1$  are in congestion. Given the assumption that there is at most one transition between freeflow and congestion in each of the three links connecting the junction, it can be concluded that there is a shock (i.e., transition from freeflow to congestion) on link 1, while link 2 and link 3 each has an expansion fan. Let  $l_1$  be the location of the shock on link 1, i.e., the transition from freeflow to congestion on link 1 occurs between cell  $l_1$  and  $l_1 + 1$ . Moreover, we define

$$\tilde{l}_1 = \begin{cases} l_1 & \text{if the shock has positive velocity or is stationary} \\ l_1 - 1 & \text{if the shock has negative velocity,} \end{cases}$$

and

$$\hat{l}_1 = n_1 - 1 - \tilde{l}_1;$$

which are later used to simplify the notations to define matrices  $A_k$ ,  $B_k$ ,  $B_k^q$ , and  $B_k$ . Similarly, denote as  $l_2$  (resp.  $l_3$ ) the location of the expansion fan on link 2 (resp. link 3), i.e., the transition from congestion to freeflow on link 2 (resp. link 3) occurs between cells  $l_2$  and  $l_2 + 1$  (resp. cells  $l_3$  and  $l_3 + 1$ ). To simplify the notation, we also define

$$\tilde{l}_2 = l_2 - n_1; \quad \hat{l}_2 = n_2 - \tilde{l}_2; \quad \tilde{l}_3 = l_3 - n_1 - n_2; \quad \hat{l}_3 = n_3 - \tilde{l}_3;$$

At the junction, the sending capacity of cell  $n_1$  is  $s(\frac{n_1}{k}) = q_m$ , and the receiving capacities of cell  $n_1 + 1$  and cell  $n_1 + n_2 + 1$  are  $r(\frac{n_1+1}{k}) = w(\frac{\%_m}{k}, \frac{n_1+1}{k})$  and  $r(\frac{n_1+n_2+1}{k}) = w(\frac{\%_m}{k}, \frac{n_1+n_2+1}{k})$ , respectively. Depending on the magnitudes of  $s(\frac{n_1}{k})$ ,  $r(\frac{n_1+1}{k})$ , and  $r(\frac{n_1+n_2+1}{k})$ , the junction solver follows one of the three possible scenarios shown in Figure 3.

1. Diverge case I: when  $s(\frac{n_1}{k}) < r(\frac{n_1+1}{k}) + r(\frac{n_1+n_2+1}{k})$ . In this case, the state is under Mode 2 at time  $k$ , and the junction solver computes fluxes  $f(\frac{n_1}{k}; \frac{n_1+1}{k})$  and  $f(\frac{n_1}{k}; \frac{n_1+n_2+1}{k})$  according to diverge case I (15) as follows:

$$\begin{aligned} f(\frac{n_1}{k}; \frac{n_1+1}{k}) &= r(\frac{n_1+1}{k}) = w(\frac{\%_m}{k}, \frac{n_1+1}{k}); \\ f(\frac{n_1}{k}; \frac{n_1+n_2+1}{k}) &= r(\frac{n_1+n_2+1}{k}) = w(\frac{\%_m}{k}, \frac{n_1+n_2+1}{k}). \end{aligned}$$

Hence in Mode 2, the explicit forms of  $A_k$ ,  $B_k$ ,  $B_k^q$ , and  $B_k$  in (19) are

$$A_k = \begin{array}{c} \textcircled{\circ} \\ \text{⋮} \\ \text{⊗} \end{array} \begin{array}{c} \Theta_{\tilde{l}_1} \\ \frac{v_m \Delta t}{\Delta x} E_{1;\tilde{l}_1}^{1;\tilde{l}_1} \\ 1 \\ \Delta_{\hat{l}_1} \\ \frac{w \Delta t}{\Delta x} E_{\hat{l}_1;\tilde{l}_2}^{\hat{l}_1;1} \\ \Delta_{\tilde{l}_2} \\ \tilde{\Theta}_{\tilde{l}_2} \\ \Delta_{\tilde{l}_3} \\ \tilde{\Theta}_{\tilde{l}_3} \end{array} \begin{array}{c} 1 \\ \text{⋮} \\ \text{⊗} \end{array};$$



construction of the system dynamics when the fluxes are given by (25) can be done in a similar fashion. The matrices  $A_k$ ,  $B_k$ ,  $B_k^q$ , and  $B_k$  read

$$A_k = \begin{array}{c} \textcircled{0} \\ \vdots \\ \textcircled{1} \end{array} \begin{array}{ccccccc} & \Theta_{\tilde{l}_1} & & & & & \\ & \frac{v_m \Delta t}{\Delta x} E_{1;\tilde{l}_1}^{1;\tilde{l}_1} & 1 & \frac{w \Delta t}{\Delta x} E_{1;\tilde{l}_1}^{1;1} & & & \\ & & & \Delta_{\tilde{l}_1} & & & \\ & & & & \frac{w \Delta t}{\Delta x} E_{\tilde{l}_1;\tilde{l}_2}^{\tilde{l}_1;1} & & \\ & & & & \Delta_{\tilde{l}_2} & & \\ & & & & & \tilde{\Theta}_{\tilde{l}_2} & \\ & & & & & & \tilde{\Delta}_{\tilde{l}_3} \\ & & & & & & \tilde{\Theta}_{\tilde{l}_3} \end{array} \begin{array}{c} 1 \\ \vdots \\ \textcircled{A} \end{array};$$

$$B_k = \frac{w \Delta t}{\Delta x} \begin{array}{cc} E_{n;n}^{\tilde{l}_1+1;\tilde{l}_1+2} + E_{n;n}^{l_2;l_2} & \\ E_{n;n}^{n_1+n_2+1;n_1+n_2+2} & E_{n;n}^{n_1+n_2+1;n_1+1} + E_{n;n}^{l_3;l_3} \end{array};$$

$$B_k^q = \frac{\Delta t}{\Delta x} \begin{array}{cc} E_{n;n}^{l_2;l_2} + E_{n;n}^{l_2+1;l_2} & \\ + E_{n;n}^{n_1+n_2+1;n_1+1} & E_{n;n}^{l_3;l_3} + E_{n;n}^{l_3+1;l_3} \end{array};$$

$$B_k = \frac{\Delta t}{\Delta x} \begin{array}{ccc} E_{n;3}^{1;1} & E_{n;3}^{n_1+n_2;2} & E_{n;3}^{n;3} \end{array};$$

2

**3.3. Properties of the state transition matrices of the SMM-J.** Some properties of the state transition matrix  $A_k$  are summarized below, which will be demonstrated in Section 4.1 to assume important roles in proving the boundedness of the Kalman gain (a necessary condition to ensure the ultimate boundedness of the mean estimation error) when using the KF to estimate traffic conditions based on the SMM-J.

(P.1): For  $A_k$  in all modes, each element satisfies

$$0 \leq A_k(r; c) \leq 1; \quad \text{for all } k \geq N \text{ and } r; c \geq \tilde{l}_1; \quad ; ng.$$

This property is due to the CFL condition [33] in the discretization scheme (8).

(P.2): When  $A_k$  is derived under diverge case I and diverge case II, the sum of the elements in  $A_k$  at the same column  $c$  satisfies

$$\sum_{r=1}^{\mathcal{X}} A_k(r; c) \leq 1; \quad \text{for all } k \geq N \text{ and } c \geq \tilde{l}_1; \quad ; ng.$$

This property is due to the CFL condition as in (P.1) and the conservation law embedded in the traffic model.

(P.3): When  $A_k$  is derived under diverge case III, the sum of the elements in  $A_k$  at the same column  $c$  satisfies

$$\sum_{r=1}^{\mathcal{X}} A_k(r; c) \leq 1; \quad \text{for all } k \geq N \text{ and } c \geq \tilde{l}_1; \quad ; ng; c \notin \tilde{g},$$

where  $\tilde{g} = n_1 + 1$  if  $f(\binom{n_1}{k}; \binom{n_1+1}{k}) = r(\binom{n_1+1}{k})$  and  $\tilde{g} = n_1 + n_2 + 1$  if  $f(\binom{n_1}{k}; \binom{n_1+1}{k}) = s(\binom{n_1}{k}) - r(\binom{n_1+n_2+1}{k})$ . Moreover, it also holds that for  $A_k$



under diverge case III,

$$A_k(r; c) = 0; \quad \text{for all } k, r \geq n_1 + 1; \quad ; ng \text{ and } c \geq f_1; \quad n_1 g.$$

This is due to the facts that (i) the flows from cell  $n_1$  at the junction to the two downstream cells (i.e., cell  $n_1 + 1$  and cell  $n_1 + n_2 + 1$ ) do not depend on the densities of the cells on link 1, as shown in (24)-(25), and (ii) the internal flows between adjacent cells on link 2 and link 3 are also independent of the densities of the cells on link 1.

Based on the above properties, the following lemma derives the bounds on each entry of the product of the state transition matrices, which will later be applied to prove the boundedness of the Kalman gain (see Lemma 3).

**Lemma 2.** *Consider the local section in Figure 4 that stays inside a SMM-J mode while  $k \geq (k_0; k_1]$ <sup>4</sup>, where  $0 \leq k_0 < k_1 + 1$ . Recall from (6) that the product of the state transition matrices is denoted as*

$$\Xi_{k+1; k_0+1} = \prod_{i=k}^{k_0+1} A_i; \quad \text{for } k \geq (k_0; k_1]; \quad (26)$$

The  $(i; j)$ <sup>th</sup> entry of  $\Xi_{k+1; k_0+1}$  satisfies

$$0 \leq \Xi_{k+1; k_0+1}(i; j) \leq 1 \quad \text{for all } k \geq (k_0; k_1] \text{ and } i; j \geq f_1; \quad ; ng. \quad (27)$$

*Proof.* The proof applies properties (P.1)-(P.3), and is reported in Appendix.2.  $\square$

**3.4. Observability of the SMM-J.** From an estimation point of view, it is assumed that the sensors are located on the far ends of the three links connecting the junction (as illustrated in Figure 4), measuring the densities of the boundary cells of the local section (i.e.,  $\frac{1}{k}$ ,  $\frac{n_1+n_2}{k}$  and  $\frac{n}{k}$ ).

Incorporating model noise in the SMM-J (19) yields:

$$x_{k+1} = A_k x_k + u_k + !_k; \quad x_k \geq \mathbb{R}^n; \quad (28)$$

where  $!_k \sim N(0; Q_k)$  is the white Gaussian model noise, and define the deterministic system input as:

$$u_k = B_k \mathbf{1}_{\%_m} + B_k^q \mathbf{1}_{q_m} + B_k x_{k-1};$$

The sensor measurements are modeled as follows:

$$z_k = H_k x_k + v_k; \quad z_k \geq \mathbb{R}^3; \quad (29)$$

where the observation matrix  $H_k = E_{3;n}^{1;1} + E_{3;n}^{2;n_1+n_2} + E_{3;n}^{3;n}$ , and  $v_k \sim N(0; R_k)$  is the white Gaussian measurement noise. Hence, as shown in (28)-(29), the system dynamics of the SMM-J is rewritten in the form of (1)-(2).

The observability of system (28)-(29) under different modes are listed in Table 1, which can be derived directly from the definition of observability stated in Section 2.1, i.e., checking the boundedness of the information matrix. According to Table 1, most of the modes are not observable except (i) when all cells in the local section are in freeflow, and (ii) when an expansion fan sits on link 1 and no other transitions between freeflow and congestion exist in the local section. From a physical viewpoint, the non-observability of the SMM-J is due to the irreversibility of the vehicle conservation law given the available sensor measurements in the presence of shocks, and the presence of the junction. It is indicated that compared to the

<sup>4</sup>Recall that the time instant  $k \geq N$ , hence  $k \geq (k_0; k_1]$  denotes  $k \geq f_{k_0} + 1; \quad ; k_1 g.$

observability of the SMM [30], the issue of non-observability is more critical when junctions exist. For example, a one-dimensional road section where the traffic is in congestion everywhere is observable given measurements of the upstream boundary cell [30], while a congested local section with a junction is not observable even with the measurements of all the three boundary cells (as shown by the cases for Modes 5-7). The performance of the KF under uniformly completely observable systems is widely studied (as summarized in Lemma 1). In this article, we focus on analysing the theoretical performance of the KF under the unobservable modes of the SMM-J.

**4. Performance analysis of the KF for unobservable traffic network estimation.** Challenges for estimating an unobservable system stem from the fact that the estimation error covariance could grow unbounded, thus the mean estimation error also potentially diverges (as shown in Example 3 in the appendix). In this subsection, we show that when combining the physical properties of the traffic model (i.e., vehicle conservation and the flow–density relationship) with the update scheme of the KF, the mean estimation errors of all the cells in an unobservable local section are ultimately bounded inside  $(\epsilon; \epsilon_m + \epsilon)$  for all  $\epsilon > 0$ , provided that the density measurements of the three boundary cells are available. This ensures that the mean estimates of the KF for unobservable modes are always physically meaningful to within  $\epsilon$ . Comparatively, it is shown in [36] that an open-loop observer may result in non-physical state estimates in unobservable modes.

Note that the three conditions (C.1)-(C.3) in Lemma 1 are necessary for proving the properties of the KF for traffic estimation under unobservable systems. For system (28)-(29), we assume that conditions (C.1) and (C.2) can be ensured when setting up the parameters in the KF. It can also be directly verified that condition (C.3) always holds for all the modes of the SMM-J.

**4.1. Boundedness of the Kalman gain.** Let  $(k_0^U; k_1^U]$  be the time interval inside which the local section shown in Figure 4 stays in an unobservable mode of the SMM-J. In this subsection, we present a lemma stating the boundedness of the Kalman gain for  $k \geq (k_0^U; k_1^U]$ , which is obtained based on the boundedness of the cross-covariance of the observable and unobservable subsystems in the *Kalman observability canonical form* (see Appendix.3). According to the KF update scheme in (4), the boundedness of the Kalman gain is a necessary condition for the boundedness of the state estimate.

**Lemma 3.** *Consider an unobservable local section shown in Figure 4. The local section stays inside an unobservable mode of the SMM-J while  $k \geq (k_0^U; k_1^U]$ , where  $0 < k_0^U < k_1^U < +1$ . Given density measurements of the three boundary cells, the infinity norm<sup>5</sup> of the Kalman gain computed by the KF (3)-(4) satisfies*

$$\|K_k\|_{\infty} \leq k \Gamma_{k_0^U | k_0^U}; \quad \text{for all } k \geq (k_0^U; k_1^U], \quad (30)$$

where  $k(\cdot)$  is a function of the error covariance at time  $k_0^U$ .

*Proof.* The explicit formula of  $k \Gamma_{k_0^U | k_0^U}$  is presented in Appendix.4.1, and is derived in Appendix.4.2.  $\square$

<sup>5</sup>Recall that for matrix  $M \in \mathbb{R}^{p \times q}$ , its infinity norm is defined as  $\|M\|_{\infty} = \max_{r \in \{1, \dots, p\}} \sum_{c=1}^q |M(r, c)|$ .

In the proof of Lemma 3, the Kalman gain is partitioned into blocks corresponding to the observable and unobservable subsystems (as shown in (49)). The part corresponding to the observable subsystem is a function of the estimation error covariance of the observable subsystem (see (50)), thus its boundedness is relatively straightforward to justify. On the other hand, the block of the Kalman gain that corresponds to the unobservable subsystem is a function of the cross-covariance of the observable and the unobservable subsystems (see (51)). By exploring the interaction between the evolution equation of the cross-covariance (shown in (48)) and the physical properties of the traffic model (reflected in (P.1)-(P.3)), we also derive the boundedness of the unobservable block of the Kalman gain. In summary, the combination of the update scheme of the KF and the intrinsic properties of the traffic model is critical in showing the boundedness of the Kalman gain under unobservable modes.

**4.2. Ultimate boundedness of the mean error.** In this subsection, we show that for an unobservable local section, the mean estimate of each cell will ultimately achieve the physically meaningful interval, thus the mean estimation error is also ultimately bounded. Unlike typical unobservable scenarios where the mean estimation error diverges (as shown in Example 3), the boundedness of the mean error here is ensured by the intrinsic physical properties of the traffic model, i.e., the relationship between the density and the sending/receiving capacity for each cell, as illustrated in the proof of the next proposition.

**Proposition 1.** *Consider an unobservable local section as shown in Figure 4 that stays inside an unobservable mode while  $k \geq (k_0^U; 1)$ . For all  $\epsilon > 0$ , a finite time  $t(\epsilon)$  exists such that  $\bar{k}_{k|k} \geq (\bar{k}_m + \epsilon)$  for all  $k > k_0^U + t(\epsilon)$  and for all  $l \geq \bar{n}_1; \bar{n}_2; ng$ , independent of the initial estimate. Moreover, the mean estimation error satisfies  $k_{k|k}^- < \bar{n}(\bar{k}_m + \epsilon)$  for all  $k > k_0^U + t(\epsilon)$ , independent of the initial estimate.*

*Proof.* Denote as  $\bar{k}_{k|k}^b = (\bar{k}_{k|k}^1, \bar{k}_{k|k}^{n_1+n_2}, \bar{k}_{k|k}^n)^\top$  the mean error of the three boundary cells. Since the three boundary cells are all inside the observable subsystem, it follows that  $\bar{k}_{k|k}^b \rightarrow 0$  as  $k \rightarrow \infty$ .

The proof is by induction. In Step 1, we use an induction from cell 1 to the downstream cells to show that if the estimate of cell 1 converges to the true state, then the estimate of all cells will ultimately be greater than  $\bar{k}_m + \epsilon$  for all  $\epsilon > 0$ . In Step 2, we use an induction from the two downstream boundary cells (i.e., cell  $n_1 + n_2$  and cell  $n$ ) to their upstream cells to show that if the estimates of the two downstream boundary cells converge to the true state, then the estimate of all cells will ultimately be smaller than  $\bar{k}_m + \epsilon$  for all  $\epsilon > 0$ . In Step 3, we combine Step 1 and Step 2 to derive an ultimate bound for the mean estimation error.

**Step 1:** We use induction to show that for all  $\epsilon > 0$  and  $l \geq \bar{n}_1; \bar{n}_2; ng$ , there exists a finite time  $t_1^l(\epsilon)$  such that  $\bar{k}_{k|k}^1 > \bar{k}_m + \epsilon$  for all  $k > k_0^U + t_1^l(\epsilon)$ .

Since the upstream boundary cell (i.e., cell 1) is in the observable subsystem, we have  $\bar{k}_{k|k}^1 \rightarrow 0$  and  $\bar{k}_{k|k}^1 \rightarrow \bar{k}_m$ , where  $\bar{k}_m > 0$ . Hence a finite time  $t_1^1(\epsilon)$  exists such that  $\bar{k}_{k|k}^1 > \bar{k}_m + \epsilon$  for all  $k > k_0^U + t_1^1(\epsilon)$ .

For all interior cells on link 1, i.e., cells indexed by  $l \geq 2; 3; \dots; n_1 g$ , suppose  $\frac{-l-1}{k|k} > \frac{(l-1)}{n}$ , if  $\frac{-l}{k|k} < \frac{(l-1)}{n}$ , we obtain from (9) that

$$f_{k|k}^{-l-1, -l} = v_m \frac{-l-1}{k|k} > v_m \frac{(l-1)}{n}; \quad (31)$$

$$f_{k|k}^{-l, -l+1} = v_m \frac{-l}{k|k}; \quad (32)$$

It follows that the estimate of cell  $l$  satisfies

$$\begin{aligned} \frac{-l}{k+1|k+1} &= \frac{-l}{k|k} + \frac{\Delta t}{\Delta x} f_{k|k}^{-l-1, -l} - f_{k|k}^{-l, -l+1} \\ &= K_{k+1}(l; 1) \frac{-l}{k+1|k} - K_{k+1}(l; 2) \frac{-n_1+n_2}{k+1|k} - K_{k+1}(l; 3) \frac{-n}{k+1|k} \\ &> \frac{-l}{k|k} + \frac{v_m \Delta t}{\Delta x} \frac{-l}{k|k} + \frac{(l-1)}{n} - k \Gamma_{k_0^U | k_0^U} - \frac{-n}{k|k} \infty; \end{aligned} \quad (33)$$

where the inequality is due to  $k \Gamma_{k_0^U | k_0^U}$  given in Lemma 3. Thus there

exists a scalar  $v_1 > \frac{\Delta x k \Gamma_{k_0^U | k_0^U}}{v_m \Delta t}$  such that

$$\begin{aligned} \frac{-l}{k+1|k+1} - \frac{-l}{k|k} &> v_0 \frac{-l}{k|k} + \frac{(l-1)}{n}; \\ \text{for all } \frac{-l}{k|k} + \frac{(l-1)}{n} &> v_1 \frac{-n}{k|k} \infty, \end{aligned} \quad (34)$$

where  $v_0 = \frac{v_m \Delta t}{\Delta x} - \frac{k \Gamma_{k_0^U | k_0^U}}{v_1} > 0$ . Also note that  $\frac{-n}{k|k} \infty \neq 0$  as  $k \neq 1$ , which indicates that the one-step change of the estimates is ultimately positive, and large enough so that a finite time  $t_1^l(\cdot)$  exists such that  $\frac{-l}{k|k} > \frac{l}{n} > \frac{-n}{k|k}$  for all  $k > k_0^U + t_1^l(\cdot)$  [21].

We now show that for the two cells on the downstream side of the junction, i.e., cell  $n_1 + 1$  and cell  $n_1 + n_2 + 1$ , there exist finite times  $t_1^{n_1+1}(\cdot)$  and  $t_1^{n_1+n_2+1}(\cdot)$  such that  $\frac{-n_1+1}{k|k} > \frac{(n_1+1)}{n} > \frac{-n_1+n_2+1}{k|k}$  for all  $k > k_0^U + t_1^{n_1+1}(\cdot)$  and  $\frac{-n_1+n_2+1}{k|k} > \frac{(n_1+1)}{n} > \frac{-n_1+1}{k|k}$  for all  $k > k_0^U + t_1^{n_1+n_2+1}(\cdot)$ . Suppose  $\frac{-n_1}{k|k} > \frac{n_1}{n}$ , if  $\frac{-n_1+1}{k|k} < \frac{n_1}{n}$ , the junction

solver follows diverge case II or diverge case III. Hence, the flow from cell  $n_1$  to cell  $n_1 + 1$  satisfies:

$$f_{k|k}^{-n_1, -n_1+1} = \begin{cases} < \frac{1}{d+1} s \frac{-n_1}{k|k} > v_m \frac{n_1}{n} & \text{diverge case II} \\ : s \frac{-n_1}{k|k} - r \frac{-n_1+n_2+1}{k|k} > v_m \frac{n_1}{n} & \text{diverge case III,} \end{cases}$$

and the outgoing flow for cell  $n_1 + 1$  satisfies:

$$f_{k|k}^{-n_1+1, -n_1+2} = v_m \frac{-n_1+1}{k|k};$$

Following the similar arguments as in (33)-(34), it can be concluded that there exist a finite time  $t_1^{n_1+1}(\cdot)$  such that  $\frac{-n_1+1}{k|k} > \frac{(n_1+1)}{n} > \frac{-n_1+n_2+1}{k|k}$  for all  $k > k_0^U + t_1^{n_1+1}(\cdot)$ . Applying the same analysis for cell  $n_1 + n_2 + 1$ , it can be concluded that there exist a finite time  $t_1^{n_1+n_2+1}(\cdot)$  such that  $\frac{-n_1+n_2+1}{k|k} > \frac{(n_1+1)}{n} > \frac{-n_1+1}{k|k}$  for all  $k > k_0^U + t_1^{n_1+n_2+1}(\cdot)$ . Continuing the induction on link 2 from cell  $n_1 + 1$  to cell  $n_1 + n_2$ , we obtain that for all  $l \geq n_1 + 1; n_1 + 2; \dots; n_1 + n_2 g$ , there exists a finite time  $t_1^l(\cdot)$  such that  $\frac{-l}{k|k} > \frac{l}{n} > \frac{-n}{k|k}$  for all  $k > k_0^U + t_1^l(\cdot)$ . As for the cells on link 3, we process the same induction from cell  $n_1 + n_2 + 1$  to cell  $n$ ,

which yields that for all  $\delta > 0$  and  $l \geq \bar{n}n_1 + n_2 + 1; n_1 + n_2 + 2; \dots; ng$ , there exists a finite time  $t_1^l(\cdot)$  such that  $\frac{-l}{k|k} > \frac{(l-n_2)}{n} > \delta$  for all  $k > k_0^U + t_1^l(\cdot)$ .

Let  $t_1(\cdot) = \max_{l \in \{1, \dots, n\}} \bar{t}_1^l(\cdot)g$ , it is concluded that  $\frac{-l}{k|k} > \delta$  for all  $k > k_0^U + t_1(\cdot)$  and  $l \geq \bar{n}n_1; \dots; ng$ . This proves the ultimate lower bound of the estimates.

**Step 2:** We use induction to show that for all  $\delta > 0$  and  $l \geq \bar{n}n_1; \dots; ng$ , there exists a finite time  $t_2^l(\cdot)$  such that  $\frac{-l}{k|k} < \mathcal{W}_m + \delta$  for all  $k > k_0^U + t_2^l(\cdot)$ .

Since the two downstream boundary cells (indexed by  $n_1 + n_2$  and  $n$ ) are in the observable subsystem, we have  $\frac{-n_1+n_2}{k|k} \neq 0$  and  $\frac{-n_1+n_2}{k|k} \neq \frac{n_1+n_2}{k}$ , as well as  $\frac{-n}{k|k} \neq 0$  and  $\frac{-n}{k|k} \neq \frac{n}{k}$ . Given the facts that  $\frac{n_1+n_2}{k} \neq \mathcal{W}_m$  and  $\frac{n}{k} \neq \mathcal{W}_m$ , there exist finite times  $t_2^{n_1+n_2}(\cdot)$  and  $t_2^n(\cdot)$  such that  $\frac{-n_1+n_2}{k|k} < \mathcal{W}_m + \frac{\delta}{n}$  for all  $k > k_0^U + t_2^{n_1+n_2}(\cdot)$ , and  $\frac{-n}{k|k} < \mathcal{W}_m + \frac{\delta}{n}$  for all  $k > k_0^U + t_2^n(\cdot)$ .

For all interior cells on link 3, i.e., cells indexed by  $l \geq \bar{n}n_1 + n_2 + 1; n_1 + n_2 + 2; \dots; n-1g$ , suppose  $\frac{-l+1}{k|k} < \mathcal{W}_m + \frac{(n-l)}{n}$ , if  $\frac{-l}{k|k} > \mathcal{W}_m + \frac{(n-l)}{n}$ , we obtain from (9) that

$$f_{\frac{-l-1}{k|k}; \frac{-l}{k|k}} = W \mathcal{W}_m \frac{-l}{k|k}; \quad (35)$$

$$f_{\frac{-l}{k|k}; \frac{-l+1}{k|k}} = W \mathcal{W}_m \frac{-l+1}{k|k} > W \frac{(n-l)}{n}; \quad (36)$$

It follows that the estimate of cell  $l$  satisfies

$$\begin{aligned} \frac{-l}{k+1|k+1} &= \frac{-l}{k|k} + \frac{\Delta t}{\Delta x} f_{\frac{-l-1}{k|k}; \frac{-l}{k|k}} - f_{\frac{-l}{k|k}; \frac{-l+1}{k|k}} \\ &= K_{k+1}(l;1) \frac{-l}{k+1|k} - K_{k+1}(l;2) \frac{-n_1+n_2}{k+1|k} - K_{k+1}(l;3) \frac{-n}{k+1|k} \\ &< \frac{-l}{k|k} - \frac{w\Delta t}{\Delta x} \frac{-l}{k|k} - \mathcal{W}_m \frac{(n-l)}{n} + k \Gamma_{k_0^U|k_0^U} \frac{-b}{k|k} \infty. \end{aligned} \quad (37)$$

Thus there exists scalar  $w_1$  such that  $\frac{\Delta x k \Gamma_{k_0^U|k_0^U}}{w\Delta t} < w_1$ , and

$$\begin{aligned} \frac{-l}{k+1|k+1} - \frac{-l}{k|k} &< w_0 \frac{-l}{k|k} - \mathcal{W}_m \frac{(n-l)}{n}; \\ \text{for all } \frac{-l}{k|k} &> \mathcal{W}_m \frac{(n-l)}{n} \quad w_1 \frac{-b}{k|k} \infty. \end{aligned} \quad (38)$$

Also note that  $\frac{-b}{k|k} \infty \neq 0$  as  $k \neq 1$ , which indicates that the one-step change of the estimates is ultimately negative, and large enough so that a finite time  $t_2^l(\cdot)$  exists such that  $\frac{-l}{k|k} < \mathcal{W}_m + \frac{(n-l+1)}{n} < \mathcal{W}_m + \delta$  for all  $k > k_0^U + t_2^l(\cdot)$ .

The above arguments can be generalized for all cells on link 2. Hence for all  $\delta > 0$  and  $l \geq \bar{n}n_1 + 1; n_1 + 2; \dots; n_1 + n_2 - 1g$ , there exists a finite time  $t_2^l(\cdot)$  such that  $\frac{-l}{k|k} < \mathcal{W}_m + \frac{(n_1+n_2-l+1)}{n} < \mathcal{W}_m + \delta$  for all  $k > k_0^U + t_2^l(\cdot)$ .

We now show that for the cell on the upstream side of the junction, i.e., cell  $n_1$ , there exists finite time  $t_2^{n_1}(\cdot)$  such that  $\frac{-n_1}{k|k} < \mathcal{W}_m + \frac{(n-n_1+1)}{n} < \mathcal{W}_m + \delta$  for all  $k > k_0^U + t_2^{n_1}(\cdot)$ . Suppose  $\frac{-n_1+1}{k|k} < \mathcal{W}_m + \frac{n_2}{n}$  and  $\frac{-n_1+n_2+1}{k|k} < \mathcal{W}_m + \frac{n_3}{n}$ , if  $\frac{-n_1}{k|k} > \mathcal{W}_m + \frac{(n-n_1)}{n} = \mathcal{W}_m + \frac{(n_2+n_3)}{n}$ , the incoming and outgoing flows of cell  $n_1$  satisfy

$$f_{\frac{-n_1-1}{k|k}; \frac{-n_1}{k|k}} = W \mathcal{W}_m \frac{-n_1}{k|k};$$

and

$$\begin{aligned}
& \begin{matrix} f \\ \infty \\ < \\ = \\ : \end{matrix} \begin{matrix} -n_1 & -n_1+1 \\ k|k' & k|k \end{matrix} + f \begin{matrix} -n_1 & -n_1+n_2+1 \\ k|k' & k|k \end{matrix} \\
& \begin{matrix} s \\ r \end{matrix} \begin{matrix} -n_1 \\ k|k \end{matrix} = q_m > W \frac{n_2+n_3}{n} & \text{diverge case II or III} \\
& \begin{matrix} -n_1+1 \\ k|k \end{matrix} + r \begin{matrix} -n_1+n_2+1 \\ k|k \end{matrix} > W \frac{n_2+n_3}{n} & \text{diverge case I.}
\end{aligned}$$

Following the similar arguments as in (37)-(38) it can be concluded that there exist finite time  $t_2^{n_1}(\cdot)$  such that  $\frac{-n_1}{k|k} < \%_m + \frac{(n-n_1+1)}{n} < \%_m +$  for all  $k > k_0^U + t_2^{n_1}(\cdot)$ . Continuing the induction from cell  $n_1$  to cell 1, it follows that for all  $l \geq n_1; 2; \dots; n_1 g$  there exists a finite time  $t_2^l(\cdot)$  such that  $\frac{-l}{k|k} < \%_m + \frac{(n-l+1)}{n} < \%_m +$  for all  $k > k_0^U + t_2^l(\cdot)$ .

Let  $t_2(\cdot) = \max_{l \in \{1, \dots, n\}} t_2^l(\cdot) g$ , we obtain  $\frac{-l}{k|k} < \%_m +$  for all  $k > k_0^U + t_2(\cdot)$  and  $l \geq n_1; 2; \dots; n g$ . This proves the ultimate upper bound of the estimates.

**Step 3:** Combining Steps 1 and 2, and define  $t(\cdot) = \max_l t_2^l(\cdot); t_2(\cdot) g$ , we obtain  $\frac{-l}{k|k} < \%_m +$  for all  $l \geq n_1; 2; \dots; n g$  and  $k > k_0^U + t(\cdot)$ . Consequently,  $k^-_{k|k} k < \bar{n}(\%_m +)$  for all  $k > k_0^U + t(\cdot)$ .  $\square$

Proposition 1 indicates that when the mean estimation error of the three boundary cells converges to zero, it will drive the state estimate of all the interior cells inside  $[0; \%_m]$  due to the conservation law and the flow-density relationship embedded in the traffic model. For example, when the state estimate  $\frac{-l}{k|k}$  is smaller than zero, the sending capacity of cell  $l$  is much smaller than the receiving capacity of cell  $l$  (as shown in equations (31)-(32)). Consequently, the update equation of the estimate (33) ensures that the one-step change of the state estimate of cell  $l$  is always positive, and the magnitude of the one-step change is proportional to the distance between zero and the current state estimate  $\frac{-l}{k|k}$ . This ensures that the estimate of cell  $l$  is constantly pushed towards zero unless it is sufficiently close to zero. The ultimate upper bound can also be derived under the same fashion.

**5. Conclusions.** In this article, we establish the theoretical performance of the KF applied to estimate the traffic density on transportation networks under unobservable scenarios. To facilitate the performance analysis of the KF, a linear SMM-J model is introduced which combines a junction solver with the switched linear system representation of the CTM. It is shown that in addition to the existence of shocks, the presence of junctions contributes significantly to the non-observability of the system.

To derive the error bounds for the KF under unobservable traffic estimation problems, we analyze several properties of the state transition matrices of the SMM-J, which reflect the intrinsic physical properties (e.g., vehicle conservation and the CFL condition in the discretization scheme) of the traffic model. Based on the above properties of the SMM-J, we show that the infinity norm of the Kalman gain is uniformly bounded under unobservable modes. Finally, we show that the mean estimate of each cell is ultimately bounded inside the physically meaningful interval, provided that the density measurements of the boundary cells are available. The ultimate lower and upper bounds are derived based on the convergence (to zero) of the mean estimation error of the boundary cells, the boundedness of the Kalman gain, and the flow-density relationship embedded in the model prediction step of the KF. As indicated in the proof, feeding sensor data back to the estimator is

critical to ensure physically meaningful estimates under unobservable system, which cannot be naturally achieved by an open-loop observer. These results provide some theoretical insights into the performance of sequential estimation algorithms widely used in traffic monitoring applications.

## REFERENCES

- [1] B. Anderson and J. Moore, “Optimal Filtering,” Prentice-Hall, inc, Englewood Cliffs, N.J., 1979.
- [2] S. Blandin, A. Couque, A. Bayen, and D. B. Work, *On sequential data assimilation for scalar macroscopic traffic flow models*, *Physica D: Nonlinear Phenomena*, **241** (2012), 1421–1440.
- [3] G. Bretti, R. Natalini and B. Piccoli, *Numerical approximations of a traffic flow model on networks*, *Networks and Heterogeneous Media*, **1** (2006), 57–84.
- [4] C. Chen, “Linear System Theory and Design,” 3rd edition, Oxford University Press, 1999.
- [5] H. Chen and H. A. Rakha, *Prediction of dynamic freeway travel times based on vehicle trajectory construction*, in “Proceedings of the 15th IEEE Conference on Intelligent Transportation Systems”, (2012), 576–581.
- [6] S. Contreras, S. Agarwal, and P. Kachroo, *Quality of traffic observability on highways with Lagrangian sensors*, *IEEE Transactions on Automation Science and Engineering*, (2017).
- [7] C. F. Daganzo, *The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory*, *Transportation Research Part B: Methodological*, **28** (1994), 269–287.
- [8] C. F. Daganzo, *The cell transmission model, part II: network traffic*, *Transportation Research Part B: Methodological*, **29** (1995), 79–93.
- [9] M. Garavello, K. Han, and B. Piccoli, “Models for Vehicular Traffic on Networks,” American Institute of Mathematical Sciences, Springfield, MO, 2016.
- [10] M. Garavello and B. Piccoli, “Traffic Flow on Networks,” American Institute of Mathematical Sciences, Springfield, MO, 2006.
- [11] S. Godunov, *A difference method for the numerical calculation of discontinuous solutions of hydrodynamic equations*, *Mathematics Sbornik*, **47** (1959), 271–306.
- [12] K. Han, B. Piccoli, and W. Szeto, *Continuous-time link-based kinematic wave model: formulation, solution existence, and well-posedness*, *Transportmetrica B: Transport Dynamics*, **3** (2016), 187–222.
- [13] J. C. Herrera, D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, and A. M. Bayen, *Evaluation of traffic data obtained via GPS-enabled mobile phones: the mobile century field experiment*, *Transportation Research Part C: Emerging Technologies*, **18** (2010), 568–583.
- [14] M. Herty and A. Klar, *Modeling, simulation, and optimization of traffic flow networks*, *SIAM Journal on Scientific Computing*, **25** (2003), 1066–1087.
- [15] H. Holden and N. H. Risebro, *A mathematical model of traffic flow on a network of unidirectional roads*, *SIAM Journal on Mathematical Analysis*, **26** (1995), 999–1017.
- [16] S. Jabari and H. Liu, *A stochastic model of traffic flow: Gaussian approximation and estimation*, *Transportation Research Part B: Methodological*, **47** (2013), 15–41.
- [17] A. H. Jazwinski, “Stochastic Process and Filtering Theory,” Academic Press, Cambridge, MA, 1970.
- [18] W. Jin, *Continuous kinematic wave models of merging traffic flow*, *Transportation Research Part B: Methodological*, **44** (2010), 1084–1103.
- [19] W. Jin, *A Riemann solver for a system of hyperbolic conservation laws at a general road junction*, *Transportation Research Part B: Methodological*, **98** (2017), 21–41.
- [20] W. Jin and H. M. Zhang, *On the distribution schemes for determining flows through a merge*, *Transportation Research Part B: Methodological*, **37** (2003), 521–540.
- [21] H. Khalil, “Nonlinear Systems,” 3rd edition, Prentice Hall, 2002.
- [22] J. Lebacque, *Intersection modeling, application to macroscopic network traffic flow models and traffic management*, in “Traffic and Granular Flow’03”, Springer, (2005), 261–278.
- [23] J.-P. Lebacque, *The Godunov scheme and what it means for first order traffic flow models*, in “International Symposium on Transportation and Traffic Theory”, (1996), 647–677.
- [24] Y. Li, C. G. Claudel, B. Piccoli, and D. B. Work, *A convex formulation of traffic dynamics on transportation networks*, *arXiv: 1702. 03908*.

- [25] M. Lighthill and G. Whitham, *On kinematic waves. II) a theory of traffic flow on long crowded roads*, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, **229** (1955), 317–345.
- [26] L. Mihaylova, R. Boel, and A. Hegyi, *Freeway traffic estimation within particle filtering framework*, Automatica, **43** (2007), 290–300.
- [27] L. Mihaylova, A. Hegyi, A. Gning, and R. Boel, *Parallelized particle and Gaussian sum particle filters for large-scale freeway traffic systems*, IEEE Transactions on Intelligent Transportation Systems, **13** (2012), 36 – 48.
- [28] I. Morescu and C. Canudas de Wit, *Highway traffic model-based density estimation*, in “Proceedings of the American Control Conference”, **3** (2011), 2012–2017.
- [29] F. Morbidi, L. L. Ojeda, C. Canudas de Wit, and I. Bellicot, *A new robust approach for highway traffic density estimation*, in “Proceedings of the 13th European Control Conference”, (2014), 1–6.
- [30] L. Munoz, X. Sun, R. Horowitz, and L. Alvarez, *Piecewise-linearized cell transmission model and parameter calibration methodology*, Transportation Research Record, **1965** (2006), 183–191.
- [31] R. Olfati-Saber, *Kalman-consensus filter: optimality, stability, and performance*, in “Proceedings of the 48th IEEE Conference on Decision and Control”, (2009), 7036–7042.
- [32] (MR0075522) P. I. Richards, *Shock waves on the highway*, Operations Research, **4** (1959), 42–51.
- [33] A. Y. L. Roux, *On the convergence of the Godunov’s scheme for first order quasi linear equations*, Proceedings of the Japan Academy, **52** (1976), 488–491.
- [34] T. Seo, A. M. Bayen, T. Kusakabe, and Y. Asakura, *Traffic state estimation on highway: a comprehensive survey*, Annual Reviews in Control, **43** (2017), 128–151.
- [35] X. Sun, L. Munoz, and R. Horowitz, *Highway traffic state estimation using improved mixture Kalman filters for effective ramp metering control*, in “Proceedings of the 42nd IEEE Conference on Decision and Control”, (2003), 6333–6338.
- [36] Y. Sun, “A Distributed Local Kalman Consensus Filter for Traffic Estimation: Design, Analysis and Validation,” Master’s thesis, University of Illinois at Urbana-Champaign, 2015.
- [37] Y. Sun and D. B. Work, *Scaling the Kalman filter for large-scale traffic estimation*, IEEE Transactions on Control of Network Systems, (2017).
- [38] J. Thai and A. M. Bayen, *State estimation for polyhedral hybrid systems and applications to the Godunov scheme for highway traffic estimation*, IEEE Transactions on Automatic Control, **60** (2015), 311–326.
- [39] C. Vivas, S. Siri, A. Ferrara, S. Sacone, G. Cavanna, and F. R. Rubio, *Distributed consensus-based switched observers for freeway traffic density estimation*, in “Proceedings of the 54th IEEE Conference on Decision and Control”, (2015), 3445–3450.
- [40] R. Wang, S. Fan, and D. B. Work, *Efficient multiple model particle filtering for joint traffic state estimation and incident detection*, Transportation Research Part C: Emerging Technologies, **71** (2016), 521–537.
- [41] Y. Wang and M. Papageorgiou, *Real-time freeway traffic state estimation based on extended Kalman filter: a general approach*, Transportation Research Part B: Methodological, **39** (2005), 141–167.
- [42] D. B. Work, S. Blandin, O.-P. Tossavainen, B. Piccoli, and A. Bayen, *A traffic model for velocity data assimilation*, Applied Mathematics Research eXpress, **2010** (2010), 1–35.
- [43] Y. Yuan, J. van Lint, R. Wilson, F. van Wageningen-Kessels, and S. Hoogendoorn, *Real-time Lagrangian traffic state estimator for freeways*, IEEE Transactions on Intelligent Transportation Systems, **13** (2012), 59–70.

## Appendix.

### Appendix.1. Example of the divergence of the mean estimation error under an unobservable system.

**Example 3.** Consider a linear discrete system describing the evolution of a moving object. The state vector is constructed as  $x_k = \begin{bmatrix} x_{k|k} \\ v_{k|k} \\ a_{k|k} \end{bmatrix}$ , where  $x_{k|k}$ ,  $v_{k|k}$ , and  $a_{k|k}$  are the location, speed, and acceleration of the moving object at time  $k \geq N$ ,



respectively. The moving object travels with a constant acceleration. The system dynamics is given by

$$x_{k+1} = A_k x_k + w_k; \quad x_k \in \mathbb{R}^3; \quad (39)$$

where  $w_k \sim N(0; Q_k)$ , the state transition matrix and the model error covariance matrices are given as follows:

$$A_k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad Q_k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; \quad \text{for all } k.$$

The initial state is  $x_0 = (2; 1; 0.05)^\top$ . The sensor measures the acceleration of the

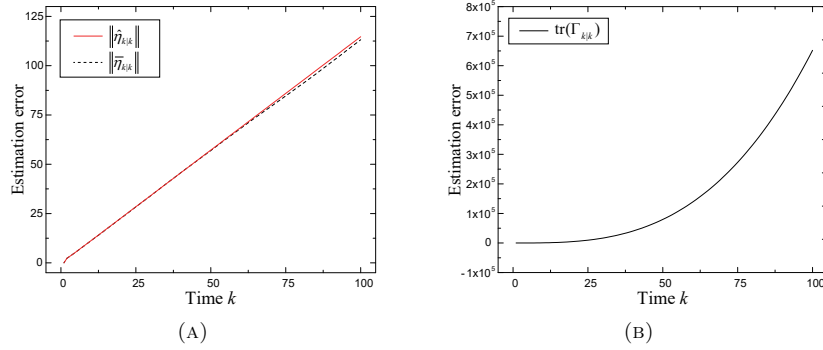


FIGURE 5. The evolutions of the estimation errors (A) and the trace of the error covariance (B) when using the KF to track the unobservable system (39)-(40).

moving object, i.e., the measurement is modeled by

$$z_k = H_k x_k + v_k; \quad z_k \in \mathbb{R}; \quad (40)$$

where

$$H_k = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}; \quad v_k \sim N(0; R_k) \text{ with } R_k = 1; \quad \text{for all } k.$$

We use the KF (3)-(4) to estimate the state, where the initial condition is set to be  $x_{0|0} = (3; 2; 0.2)^\top$  and  $\Gamma_{0|0} = I_3$ .

The system (39)-(40) is not observable, which can be concluded by computing its observability matrix [4, Theorem 6.01], and showing that the observability matrix is not full rank. The mean estimation error evolves as the following equation:

$$\bar{x}_{k|k} = (I - K_k H_k) A_{k-1} \bar{x}_{k-1|k-1}. \quad (41)$$

In Figure 5a, the solid curve shows the analytical evolution of  $\bar{x}_{k|k}$  which follows (41). A Monte Carlo test of  $N_r = 10,000$  realizations of the KF is also conducted, and the dashed curve in Figure 5a shows the empirical evolution of the estimation error  $\hat{x}_{k|k}$ , where  $\hat{x}_{k|k} = \frac{1}{N_r} \sum_{r=1}^{N_r} x_{r,k|k}$ , and  $x_{r,k|k}$  is the posterior estimation error at time  $k$  on the  $r^{\text{th}}$  realization. We also plot in Figure 5b the trace of the estimation error covariance  $\text{tr}(\Gamma_{k|k})$ . It is shown that unlike the observable scenarios described in Lemma 1, the error covariance and the estimation error diverge as  $k$  increases in this example, which is typical for unobservable systems.  $\square$

Appendix.2. **Proof of Lemma 2.** The proof is divided into two steps. In Step 1, we apply properties (P.1) and (P.2) to show that (27) holds for the modes where the junction solver follows diverge case I or diverge case II. In Step 2, properties (P.1) and (P.3) are applied to show that (27) holds for the modes where the junction solver follows diverge case III.

For all  $k_0 + 1 \leq k \leq k_1$  and  $i; j \geq 1; \dots; ng$ , the  $(i; j)^{\text{th}}$  entry of  $\bigcirc_{=k} A$  is given by

$$\bigvee_{=k} A(i; j) = \bigvee_{r=1}^{\infty} A(i; r) \bigwedge_{=k} A(r; j): \quad (42)$$

**Step 1:** Suppose  $A_k$  is under a mode where the junction solver follows diverge case I or diverge case II. Recall from (P.2) that

$$\bigvee_{r=1}^{\infty} A_k(r; j) = 1; \quad \text{for all } k \geq (k_0; k_1] \text{ and } j \geq 1; \dots; ng.$$

Hence, the  $(i; j)^{\text{th}}$  entry of  $\bigcirc_{=k} A$  is no greater than the convex combination of all the entries on the  $i^{\text{th}}$  row of  $\bigcirc_{=k}^{+1} A$ . Moreover, recall from (P.1) that

$$0 \leq A_k(r; c) = 1; \quad \text{for all } k \geq (k_0; k_1] \text{ and } r; c \geq 1; \dots; ng,$$

it follows that

$$0 \leq \bigvee_{=k} A(i; j) = 1; \quad \text{for all } k \geq (k_0; k_1], \leq \geq [k_0 + 1; k] \text{ and } i; j \geq 1; \dots; ng,$$

thus (27) follows directly by setting  $\leq \geq = k_0 + 1$  in the above equation.

**Step 2:** Suppose  $A_k$  is under a mode where the junction solver follows diverge case III. We prove for the case where  $f\left(\begin{smallmatrix} n_1 \\ k \end{smallmatrix}; \begin{smallmatrix} n_1+1 \\ k \end{smallmatrix}\right) = r\left(\begin{smallmatrix} n_1+1 \\ k \end{smallmatrix}\right)$ , and the proof for the case where  $f\left(\begin{smallmatrix} n_1 \\ k \end{smallmatrix}; \begin{smallmatrix} n_1+1 \\ k \end{smallmatrix}\right) = s\left(\begin{smallmatrix} n_1 \\ k \end{smallmatrix}\right) - r\left(\begin{smallmatrix} n_1+n_2+1 \\ k \end{smallmatrix}\right)$  follows by symmetry.

Recall from (P.3) that

$$\bigvee_{r=1}^{\infty} A_k(r; j) = 1; \quad \text{for all } k \geq (k_0; k_1] \text{ and } j \geq 1; \dots; ng; j \notin n_1 + 1g. \quad (43)$$

For  $j = n_1 + 1$ , the sum of all entries of  $A_k$  on column  $j$  is given by

$$\begin{aligned} \bigvee_{r=1}^{\infty} A_k(r; n_1 + 1) &= A_k(n_1; n_1 + 1) + A_k(n_1 + 1; n_1 + 1) + A_k(n_1 + n_2 + 1; n_1 + 1) \\ &= \frac{w\Delta t}{\Delta X} + 1 - \frac{w\Delta t}{\Delta X} + \frac{w\Delta t}{\Delta X} = 1 + \frac{w\Delta t}{\Delta X}; \quad \text{for all } k \geq (k_0; k_1]: \end{aligned}$$

Additionally, one may note that for all  $k \geq (k_0; k_1]$ ,

$$A_k(r; n_1 + n_2 + 1) = \begin{cases} 1 & \text{if } r = n_1 + n_2 + 1 \\ 0 & \text{otherwise.} \end{cases}$$

It follows that for all  $k_0 + 1 \leq k \leq k_1$ ,

$$\bigvee_{=k} A(r; n_1 + n_2 + 1) = \begin{cases} 1 & \text{if } r = n_1 + n_2 + 1 \\ 0 & \text{otherwise.} \end{cases} \quad (44)$$

Combining (43) and (44) with (42), we obtain that for all  $(i; j) \notin (n_1 + n_2 + 1; n_1 + 1)$ , the  $(i; j)^{\text{th}}$  entry of  $\bigcirc_{=k} A$  is no greater than the convex combination of all the

(non-zero) entries on the  $i^{\text{th}}$  row of  $\overset{\circ}{=}_{=k}^{+1} A$ . Also recall from (P.3) that for all  $k \geq (k_0; k_1]$ ,

$$A_k(r; c) = 0; \quad \text{for all } r \geq \bar{r}n_1 + 1; \quad ; ng \text{ and } c \geq \bar{r}1; \quad n_1g,$$

which yields that for all  $k_0 + 1 \leq k \leq k_1$ ,

$$\forall_{=k}^{+1} A_k(r; c) = 0; \quad \text{for all } r \geq \bar{r}n_1 + 1; \quad ; ng \text{ and } c \geq \bar{r}1; \quad n_1g,$$

thus

$$\forall_{=k}^{+1} A_k(n_1 + n_2 + 1; n_1) = 0; \quad \text{for all } k_0 + 1 \leq k \leq k_1.$$

Hence for  $(i; j) = (n_1 + n_2 + 1; n_1 + 1)$ , the  $(i; j)^{\text{th}}$  entry of  $\overset{\circ}{=}_{=k} A$  is also no greater than the convex combination of all the (non-zero) entries on the  $i^{\text{th}}$  row of  $\overset{\circ}{=}_{=k}^{+1} A$ . Moreover, according to (P.1) it holds that

$$0 \leq A_k(r; c) \leq 1; \quad \text{for all } k \geq (k_0; k_1] \text{ and } r; c \geq \bar{r}1; \quad ; ng.$$

It can be concluded that

$$0 \leq \forall_{=k} A_k(i; j) \leq 1; \quad \text{for all } k \geq (k_0; k_1], \quad \bar{r} \geq [k_0 + 1; k] \text{ and } i; j \geq \bar{r}1; \quad ; ng,$$

thus (27) follows directly by setting  $\bar{r} = k_0 + 1$  in the above equation.

**Appendix.3. Observable and unobservable subsystems in the unobservable modes.** In an unobservable mode, the SMM-J can be transformed to the *Kalman observability canonical form*. The transformed state consists of the observable and the unobservable parts of the system, i.e.,

$$\overset{(t)}{x}_k = U \overset{(1)}{x}_k = \begin{pmatrix} \overset{(1)}{x}_k \\ \overset{(2)}{x}_k \end{pmatrix};$$

where  $U$  is an orthogonal matrix,  $\overset{(1)}{x}_k \in \mathbb{R}^{d_1}$  and  $\overset{(2)}{x}_k \in \mathbb{R}^{d_2}$  are the state in the observable and unobservable subsystems, respectively, with  $d_1 + d_2 = n$ . Moreover, since the densities of the three boundary cells are directly measured, it holds that  $d_1 \geq 3$ . As a consequence, system (28)-(29) is transformed to the following formula:

$$\begin{aligned} \overset{(t)}{x}_{k+1} &= A_k^{(t)} \overset{(t)}{x}_k + U_k^{(t)} + \overset{(t)}{w}_k; \quad \overset{(t)}{x}_k \in \mathbb{R}^n; \\ z_k &= H_k^{(t)} \overset{(t)}{x}_k + v_k; \quad z_k \in \mathbb{R}^3; \end{aligned}$$

where the transformed state transition matrix  $A_k^{(t)}$  and the transformed observation matrix  $H_k^{(t)}$  can also be partitioned according to the observable and unobservable subsystems, i.e.,

$$A_k^{(t)} = U A_k U^T = \begin{pmatrix} A_k^{(1)} & \mathbf{0}_{d_1; d_2} \\ A_k^{(21)} & A_k^{(2)} \end{pmatrix}; \quad H_k^{(t)} = H_k U^T = \begin{pmatrix} H_k^{(1)} & \mathbf{0} \end{pmatrix}; \quad (45)$$

with  $H_k^{(1)} \in \mathbb{R}^{3 \times d_1}$  defined as follows:

$$H_k^{(1)} = \begin{cases} I_3 & \text{if } d_1 = 3 \\ I_3 \quad \mathbf{0}_{3; d_1-3} & \text{if } d_1 > 3, \end{cases} \quad \text{for all } k.$$

Moreover, the transformed system input is given by  $u_k^{(t)} = Uu_k$ , and the transformed model noise is given by  $w_k^{(t)} = Uw_k \sim N(\mathbf{0}; Q_k^{(t)})$ , where the transformed model error covariance  $Q_k^{(t)}$  can be partitioned to blocks corresponding to the observable and unobservable subsystems, i.e.,

$$Q_k^{(t)} = UQ_kU^\top = \begin{array}{cc} Q_k^{(1)} & Q_k^{(12)} \\ Q_k^{(21)} & Q_k^{(2)} \end{array} ;$$

The prior estimation error covariance matrix partitioned into the observable and unobservable subsystems is constructed as follows:

$$\Gamma_{k|k-1}^{(t)} = \begin{array}{cc} \Gamma_{k|k-1}^{(1)} & \Gamma_{k|k-1}^{(12)} \\ \Gamma_{k|k-1}^{(21)} & \Gamma_{k|k-1}^{(2)} \end{array} ;$$

In the KF, the prior error covariance matrix is computed recursively by the Riccati equation

$$\begin{aligned} \Gamma_{k+1|k}^{(t)} = & A_k^{(t)} \Gamma_{k|k-1}^{(t)} A_k^{(t)\top} + H_k^{(t)\top} \Gamma_{k|k-1}^{(t)} H_k^{(t)} \Gamma_{k|k-1}^{(t)} H_k^{(t)\top} + R_k^{-1} \\ & H_k^{(t)\top} \Gamma_{k|k-1}^{(t)} A_k^{(t)\top} + Q_k^{(t)}; \end{aligned} \quad (46)$$

Define

$$\begin{aligned} \Upsilon_k^{(1)} = & A_k^{(1)} A_k^{(1)} \Gamma_{k|k-1}^{(1)} H_k^{(1)\top} + H_k^{(1)\top} \Gamma_{k|k-1}^{(1)} H_k^{(1)} \Gamma_{k|k-1}^{(1)} H_k^{(1)\top} + R_k^{-1} H_k^{(1)} \\ = & A_k^{(1)} A_k^{(1)} K_k^{(1)} H_k^{(1)}; \end{aligned}$$

and apply partition into observable and unobservable subsystems to both sides of (46), we obtain the following two blocks of equations describing the evolutions of  $\Gamma_{k+1|k}^{(1)}$  and  $\Gamma_{k+1|k}^{(12)}$ :

$$\Gamma_{k+1|k}^{(1)} = \Upsilon_k^{(1)} \Gamma_{k|k-1}^{(1)} A_k^{(1)\top} + Q_k^{(1)}; \quad (47)$$

$$\Gamma_{k+1|k}^{(12)} = \Upsilon_k^{(1)} \Gamma_{k|k-1}^{(12)} A_k^{(2)\top} + \Upsilon_k^{(1)} \Gamma_{k|k-1}^{(1)} A_k^{(21)\top} + Q_k^{(12)}. \quad (48)$$

**Appendix.4. Explicit formula of the Kalman gain bound and proof of Lemma 3.** In this section, we present the explicit formula of  $k \Gamma_{k_0^U|k_0^U}$  and prove Lemma 3. As detailed in [Appendix.3](#), we transform the state vector according to observable and unobservable subsystems, i.e.,

$$\begin{pmatrix} x_k^{(1)} \\ x_k^{(2)} \end{pmatrix} = U \begin{pmatrix} x_k \\ x_k \end{pmatrix};$$

where  $\begin{pmatrix} x_k^{(1)} \end{pmatrix} \in \mathbb{R}^{d_1}$  and  $\begin{pmatrix} x_k^{(2)} \end{pmatrix} \in \mathbb{R}^{d_2}$  are the state in the observable and unobservable subsystems, respectively. The transformed Kalman gain is given by

$$K_k^{(t)} = UK_k = \begin{array}{c} K_k^{(1)} \\ K_k^{(21)} \end{array}; \quad (49)$$

where  $K_k^{(1)}$  and  $K_k^{(21)}$  correspond to the observable and unobservable subsystems, respectively, with

$$K_k^{(1)} = \Gamma_{k|k-1}^{(1)} \begin{bmatrix} H_k^{(1)\top} & \\ & R_k + H_k^{(1)}\Gamma_{k|k-1}^{(1)} \end{bmatrix}^{-1} H_k^{(1)\top}; \quad (50)$$

$$K_k^{(21)} = \Gamma_{k|k-1}^{(21)} \begin{bmatrix} H_k^{(1)\top} & \\ & R_k + H_k^{(1)}\Gamma_{k|k-1}^{(1)} \end{bmatrix}^{-1}; \quad (51)$$

Appendix.4.1. *Explicit formula of the Kalman gain bound.* Define

$$a_1 = \max_{k \in (k_0^U; k_1^U]} f k A_k k_\infty g; \quad a_2 = \max_{k \in (k_0^U; k_1^U]} A_k^\top; \quad \infty;$$

and

$$\begin{aligned} \tilde{a}_1 &= \max_{k \in (k_0^U; k_1^U]} A_k^{(1)\top}; & \tilde{a}_2 &= \max_{k \in (k_0^U; k_1^U]} A_k^{(1)\top}; \\ \tilde{a}_3 &= \max_{k \in (k_0^U; k_1^U]} \max A_k^{(1)}; & \tilde{a}_4 &= \max_{k \in (k_0^U; k_1^U]} A_k^{(21)\top}; \end{aligned}$$

Moreover, define as  $\tilde{c}_1$  and  $\tilde{c}_2$  the lower and upper bounds of the error covariance of the observable subsystem, i.e.,

$$\tilde{c}_1 I < \Gamma_{k|k}^{(1)} < \tilde{c}_2 I; \quad \text{for all } k \geq (k_0^U; k_1^U], \quad (52)$$

and let

$$\begin{aligned} \tilde{c}_3 &= \tilde{c}_2 + q_1^{-1} \tilde{c}_2^2 \tilde{a}_3^{-1}; & \tilde{t} &= n^2 \overline{d_1} \tilde{c}_2 \tilde{c}_1^{-1 \frac{1}{2}}; & \tilde{q} &= (1 - \tilde{c}_3 \tilde{c}_1)^{\frac{1}{2}} \\ \tilde{\rho} &= d_1 \tilde{c}_2 \tilde{a}_4 (\tilde{a}_1 \tilde{a}_2 \tilde{c}_2 + q_2) q_1^{-1} + q_2; & \tilde{\tau} &= n^{\rho_{\tilde{n}}} \Gamma_{k_0^U | k_0^U} (a_1 a_2)^2 + n^{\rho_{\tilde{n}}} \tilde{n} a_1 a_2 q_2 + \rho_{\tilde{n}} \tilde{n} q_2. \end{aligned}$$

The upper bound of  $k K_k k_\infty$  for  $k \geq (k_0^U; k_1^U]$  in (30) is defined as

$$\begin{aligned} & k \Gamma_{k_0^U | k_0^U} \\ &= \frac{\rho_{\tilde{3}d_1}}{r_1} \max \frac{\rho_{\tilde{n}}}{d_1} \Gamma_{k_0^U | k_0^U} (a_1 a_2 + q_2); \frac{1}{\tilde{\rho}} (\tilde{a}_1 \tilde{a}_2 \tilde{c}_2 + q_2); \tilde{\tau} \tilde{q} + \tilde{\rho}; \frac{\tilde{t} \tilde{\rho} \tilde{q}}{1 - \tilde{q}} + \tilde{\rho} \end{aligned}$$

Appendix.4.2. *Proof of Lemma 3.* The proof consists of the following five steps.

Step 1 derives an upper bound for  $K_{k_0^U+1}$ . Step 2 derives an upper bound of

$K_k^{(1)}$  for  $k \geq (k_0^U + 1; k_1^U]$ . In Step 3, we study the convergence rate of the error dynamics of the observable subsystem, which is also related to the boundedness of  $K_k^{(21)}$ . Based on the convergence rate obtained in Step 3, Step 4 derives an upper bound of  $K_k^{(21)}$  for  $k \geq (k_0^U + 1; k_1^U]$ . Step 5 combines the above steps together and concludes the proof.

**Step 1:** At time step  $k_0^U + 1$ , the Kalman gain is computed as follows:

$$K_{k_0^U+1} = \Gamma_{k_0^U+1 | k_0^U} H_{k_0^U+1}^\top \begin{bmatrix} R_{k_0^U+1} + H_{k_0^U+1} \Gamma_{k_0^U+1 | k_0^U} H_{k_0^U+1}^\top \\ \end{bmatrix}^{-1};$$

where  $\Gamma_{k_0^U+1 | k_0^U} = A_{k_0^U} \Gamma_{k_0^U | k_0^U} A_{k_0^U}^\top + Q_{k_0^U}$ . Given that  $\Gamma_{k_0^U | k_0^U} \leq \rho_{\tilde{n}} \Gamma_{k_0^U | k_0^U}$ ,  $k Q_k k_\infty < \rho_{\tilde{n}} \tilde{n} q_2$ , and define

$$a_1 = \max_{k \in (k_0^U; k_1^U]} f k A_k k_\infty g; \quad a_2 = \max_{k \in (k_0^U; k_1^U]} A_k^\top; \quad \infty;$$

the prior error covariance at time  $k_0^U + 1$  satisfies

$$\begin{aligned} \Gamma_{k_0^U+1|k_0^U} & \leq A_{k_0^U} \Gamma_{k_0^U|k_0^U} A_{k_0^U}^\top + Q_{k_0^U} \\ & < \frac{\rho_-}{n} \Gamma_{k_0^U|k_0^U} a_1 a_2 + \frac{\rho_-}{n} q_2. \end{aligned}$$

Moreover, since

$$\begin{aligned} & R_{k_0^U+1} + H_{k_0^U+1} \Gamma_{k_0^U+1|k_0^U} H_{k_0^U+1}^\top \\ & \leq \frac{\rho_-}{3} R_{k_0^U+1} + H_{k_0^U+1} \Gamma_{k_0^U+1|k_0^U} H_{k_0^U+1}^\top \\ & = \frac{\rho_-}{3} \min R_{k_0^U+1} + H_{k_0^U+1} \Gamma_{k_0^U+1|k_0^U} H_{k_0^U+1}^\top \\ & \leq \frac{\rho_-}{3} \min R_{k_0^U+1} \\ & < \frac{\rho_-}{r_1}; \end{aligned}$$

it follows that

$$\begin{aligned} K_{k_0^U+1} & \leq \Gamma_{k_0^U+1|k_0^U} H_{k_0^U+1}^\top R_{k_0^U+1} + H_{k_0^U+1} \Gamma_{k_0^U+1|k_0^U} H_{k_0^U+1}^\top \\ & < \frac{\rho_-}{r_1} \Gamma_{k_0^U|k_0^U} a_1 a_2 + q_2. \end{aligned}$$

**Step 2:** As stated in (50), Kalman gain associated with the observable subsystem is given by

$$K_k^{(1)} = \Gamma_{k|k-1}^{(1)} H_k^{(1)\top} R_k + H_k^{(1)} \Gamma_{k|k-1}^{(1)} H_k^{(1)\top}{}^{-1};$$

According to Lemma 1, there exist constants  $\tilde{c}_1$  and  $\tilde{c}_2$  such that the error covariance of the observable subsystem satisfies

$$\tilde{c}_1 I < \Gamma_{k|k}^{(1)} < \tilde{c}_2 I; \quad \text{for all } k \geq (k_0^U; k_1^U]. \quad (53)$$

Given that

$$\Gamma_{k|k-1}^{(1)} = A_k^{(1)} \Gamma_{k-1|k-1}^{(1)} A_k^{(1)\top} + Q_{k-1}^{(1)};$$

we have

$$\Gamma_{k|k-1}^{(1)} \leq \tilde{a}_1 \tilde{a}_2 \Gamma_{k-1|k-1}^{(1)} + Q_{k-1}^{(1)} < \frac{\rho_-}{d_1} (\tilde{a}_1 \tilde{a}_2 \tilde{c}_2 + q_2);$$

with  $\tilde{a}_1$  and  $\tilde{a}_2$  defined as

$$\tilde{a}_1 = \max_{k \in (k_0^U; k_1^U]} A_k^{(1)}; \quad \tilde{a}_2 = \max_{k \in (k_0^U; k_1^U]} A_k^{(1)\top};$$

Following the similar argument as in Step 1, we obtain

$$K_k^{(1)} < \frac{\rho_-}{r_1} (\tilde{a}_1 \tilde{a}_2 \tilde{c}_2 + q_2); \quad \text{for all } k \geq (k_0^U + 1; k_1^U].$$

**Step 3:** Define the Lyapunov function of the observable subsystem as

$$V_k^{(1)} = \Gamma_{k|k}^{(1)}{}^{-1}.$$

According to Lemma 3 in [31], the one-step change of  $V_k^{(1)}$  is given by:

$$\begin{aligned}
& V_{k+1}^{(1)} - V_k^{(1)} \\
= & \begin{pmatrix} - \\ k|k \end{pmatrix}^\top \left( \Gamma_{k|k}^{(1)} + \Gamma_{k|k}^{(1)} A_k^{(1)\top} \right. \\
& \left. Q_k^{(1)} + \Gamma_{k+1|k}^{(1)} H_k^{(1)\top} R_{k+1}^{-1} H_k^{(1)} \Gamma_{k+1|k}^{(1)} - A_k^{(1)} \Gamma_{k|k}^{(1)} \right) \\
& \begin{pmatrix} - \\ k|k \end{pmatrix}^2 \left( \Gamma_{k|k}^{(1)} + \Gamma_{k|k}^{(1)} A_k^{(1)\top} \right. \\
& \left. Q_k^{(1)} + \Gamma_{k+1|k}^{(1)} H_k^{(1)\top} R_{k+1}^{-1} H_k^{(1)} \Gamma_{k+1|k}^{(1)} - A_k^{(1)} \Gamma_{k|k}^{(1)} \right)^{-1};
\end{aligned}$$

where

$$\begin{aligned}
& \Gamma_{k|k}^{(1)} + \Gamma_{k|k}^{(1)} A_k^{(1)\top} \left( Q_k^{(1)} + \Gamma_{k+1|k}^{(1)} H_k^{(1)\top} R_{k+1}^{-1} H_k^{(1)} \Gamma_{k+1|k}^{(1)} - A_k^{(1)} \Gamma_{k|k}^{(1)} \right)^{-1} \\
& < \tilde{c}_2 + q_1^{-1} \Gamma_{k|k}^{(1)} A_k^{(1)\top} A_k^{(1)} \Gamma_{k|k}^{(1)} \\
& \tilde{c}_2 + q_1^{-1} \Gamma_{k|k}^{(1)} A_k^{(1)\top} A_k^{(1)} \Gamma_{k|k}^{(1)} \\
& < \tilde{c}_2 + q_1^{-1} \tilde{c}_2^2 \tilde{\alpha}_3;
\end{aligned}$$

with  $\tilde{\alpha}_3$  defined as

$$\tilde{\alpha}_3 = \max_{k \in (k_0^U, k_1^U]} \max A_k^{(1)};$$

and  $\max(M)$  is the maximum singular value of matrix  $M$ . It follows that for all  $k \geq (k_0^U, k_1^U]$ , the Lyapunov function  $V_k^{(1)}$  satisfies

$$\tilde{c}_2^{-1} \begin{pmatrix} - \\ k|k \end{pmatrix}^2 < V_k^{(1)} < \tilde{c}_1^{-1} \begin{pmatrix} - \\ k|k \end{pmatrix}^2; \quad \text{and} \quad V_{k+1}^{(1)} - V_k^{(1)} < \tilde{c}_3 \begin{pmatrix} - \\ k|k \end{pmatrix}^2;$$

where

$$\tilde{c}_3 = \tilde{c}_2 + q_1^{-1} \tilde{c}_2^2 \tilde{\alpha}_3;$$

Hence, for all  $k \geq (k_0^U + 1, k_1^U]$ , the 2-norm of the mean estimation error of the observable subsystem satisfies

$$\begin{aligned}
\begin{pmatrix} - \\ k|k \end{pmatrix} & < \tilde{c}_2 V_k^{(1)}^{\frac{1}{2}} \\
& < \tilde{c}_2 V_{k_0^U+1}^{(1)} (1 - c_3 \tilde{c}_1)^{k-k_0^U-1} \frac{1}{2} \\
& < \frac{\tilde{c}_2}{\tilde{c}_1} \begin{pmatrix} - \\ k_0^U+1|k_0^U+1 \end{pmatrix}^2 (1 - \tilde{c}_3 \tilde{c}_1)^{k-k_0^U-1} \frac{1}{2} \\
& = \frac{\tilde{c}_2}{\tilde{c}_1} \frac{1}{2} \begin{pmatrix} - \\ k_0^U+1|k_0^U+1 \end{pmatrix} (1 - \tilde{c}_3 \tilde{c}_1)^{\frac{1}{2} (k-k_0^U-1)};
\end{aligned} \tag{54}$$

Moreover, for all  $k \geq (k_0^U + 1; k_1^U]$ , the mean estimation error of the observable subsystem is given as follows:

$$\underset{=k}{\overset{k_Y^U+2}{\Gamma}} \overset{-(1)}{k|k} = \underset{=k}{\overset{k_Y^U+2}{\Gamma}} \overset{-(1)}{k_0^U+1|k_0^U+1}; \quad (55)$$

where  $\Upsilon^{(1)} = \Gamma_{|}^{(1)} \Gamma_{|-1}^{(1)} \overset{-1}{A}_{|-1}^{(1)}$ . Combining (54) and (55), it is concluded based on the definition of matrix induced norm that

$$\underset{=k}{\overset{k_Y^U+2}{\Gamma}} \overset{(1)}{\Gamma} = \frac{\tilde{c}_2}{\tilde{c}_1} \overset{\frac{1}{2}}{1} \left( \tilde{c}_3 \tilde{c}_1 \right)^{\frac{1}{2}} \overset{k-k_0^U-1}{1}; \quad \text{for } k \geq (k_0^U + 1; k_1^U]. \quad (56)$$

**Step 4:** Vectorizing both sides of (48) yields that for  $k \geq (k_0^U; k_1^U]$ ,

$$\begin{aligned} \text{vec } \overset{\circ}{\Gamma}_{k+1|k}^{(12)} &= \overset{\circ}{A}_k^{(2)} \overset{\circ}{\Upsilon}_k^{(1)} \text{vec } \overset{\circ}{\Gamma}_{k|k-1}^{(12)} \\ &+ \text{vec } \overset{\circ}{\Upsilon}_k^{(1)} \overset{\circ}{\Gamma}_{k|k-1}^{(1)} \overset{\circ}{A}_k^{(21)\top} + \text{vec } \overset{\circ}{Q}_k^{(12)\circ}; \end{aligned}$$

which implies that for all  $k \geq (k_0^U + 1; k_1^U]$ ,

$$\text{vec } \overset{\circ}{\Gamma}_{k+1|k}^{(12)\circ} = \overset{\circ}{A}_k^{(2)} \overset{\circ}{\Upsilon}_k^{(1)} \overset{\circ}{A}_k \text{vec } \overset{\circ}{\Gamma}_{k_0^U+2|k_0^U+1}^{(12)\circ} + \Phi_k; \quad (57)$$

where

$$\begin{aligned} \Phi_k &= \text{vec } \overset{\circ}{\Upsilon}_k^{(1)} \overset{\circ}{\Gamma}_{k|k-1}^{(1)} \overset{\circ}{A}_k^{(21)\top} + \overset{\circ}{Q}_k^{(12)} \\ &+ \overset{\circ}{A}_k^{(2)} \overset{\circ}{\Upsilon}_k^{(1)} \text{vec } \overset{\circ}{\Upsilon}_{k-1}^{(1)} \overset{\circ}{\Gamma}_{k-1|k-2}^{(1)} \overset{\circ}{A}_{k-1}^{(21)\top} + \overset{\circ}{Q}_{k-1}^{(12)} \\ &+ \overset{\circ}{A}_k^{(2)} \overset{\circ}{\Upsilon}_k^{(1)} \overset{\circ}{A}_{k-1}^{(2)} \overset{\circ}{\Upsilon}_{k-1}^{(1)} \text{vec } \overset{\circ}{\Upsilon}_{k-2}^{(1)} \overset{\circ}{\Gamma}_{k-2|k-3}^{(1)} \overset{\circ}{A}_{k-2}^{(21)\top} + \overset{\circ}{Q}_{k-2}^{(12)} \\ &+ \overset{\circ}{A}_k^{(2)} \overset{\circ}{\Upsilon}_k^{(1)} \overset{\circ}{A}_{k-1}^{(2)} \overset{\circ}{\Upsilon}_{k-1}^{(1)} \text{vec } \overset{\circ}{\Upsilon}_{k_0^U+2}^{(1)} \overset{\circ}{\Gamma}_{k_0^U+2|k_0^U+1}^{(1)} \overset{\circ}{A}_{k_0^U+2}^{(21)\top} + \overset{\circ}{Q}_{k_0^U+2}^{(12)}; \end{aligned}$$

The explicit form of  $\overset{\circ}{A}_k^{(2)} \overset{\circ}{\Upsilon}_k^{(1)}$  reads

$$\overset{\circ}{A}_k^{(2)} \overset{\circ}{\Upsilon}_k^{(1)} = \overset{\circ}{\text{vec}} \begin{pmatrix} \overset{\circ}{A}_k^{(2)}(1;1)\overset{\circ}{\Upsilon}_k^{(1)} & \dots & \overset{\circ}{A}_k^{(2)}(1;d_2)\overset{\circ}{\Upsilon}_k^{(1)} \\ \vdots & \ddots & \vdots \\ \overset{\circ}{A}_k^{(2)}(d_2;1)\overset{\circ}{\Upsilon}_k^{(1)} & \dots & \overset{\circ}{A}_k^{(2)}(d_2;d_2)\overset{\circ}{\Upsilon}_k^{(1)} \end{pmatrix};$$

hence

$$\underset{=k}{\overset{k_Y^U+2}{\overset{\circ}{A}} \overset{\circ}{\Upsilon}}^{(1)} = \overset{\circ}{\text{vec}} \begin{pmatrix} \#_k(1;1) \overset{\circ}{Q}_{=k}^{k_0^U+2} \overset{\circ}{\Upsilon}^{(1)} & \dots & \#_k(1;d_2) \overset{\circ}{Q}_{=k}^{k_0^U+2} \overset{\circ}{\Upsilon}^{(1)} \\ \vdots & \ddots & \vdots \\ \#_k(d_2;1) \overset{\circ}{Q}_{=k}^{k_0^U+2} \overset{\circ}{\Upsilon}^{(1)} & \dots & \#_k(d_2;d_2) \overset{\circ}{Q}_{=k}^{k_0^U+2} \overset{\circ}{\Upsilon}^{(1)} \end{pmatrix};$$

where  $\#_k(i;j)$  is the  $(i;j)$ <sup>th</sup> element of  $\overset{\circ}{Q}_{=k}^{k_0^U+2} \overset{\circ}{A}^{(2)}$ . Define

$$P = \mathbf{0}_{d_2;n-d_2} \quad I_{d_2};$$



and given that the top right block of  $A_k^{(t)}$  is a zero matrix  $\mathbf{0}_{d_1; d_2}$  (as shown in (45)), it can be concluded that

$$\begin{aligned} \overset{k_0^U+2}{=} A^{(2)} = P @ \overset{0}{=} \overset{k_0^U+2}{=} A^{(t)} A P^\top = P @ \overset{0}{=} \overset{k_0^U+2}{=} U A U^\top A P^\top = P U @ \overset{0}{=} \overset{k_0^U+2}{=} A A U^\top P^\top; \end{aligned} \quad (58)$$

Based on Lemma 2, the  $(i; j)^{th}$  element of  $\overset{0}{=} \overset{k_0^U+2}{=} A$  satisfies

$$0 @ \overset{0}{=} \overset{k_0^U+2}{=} A A (i; j) = 1; \quad \text{for all } i; j \geq 1; \quad ; ng.$$

Hence, it can be derived from (58) that

$$\overset{k_0^U+2}{=} \check{A}^{(2)} \overset{k_0^U+2}{=} \check{A} \overset{0}{=} \overset{k_0^U+2}{=} U @ \overset{0}{=} \overset{k_0^U+2}{=} A A U^\top \overset{P}{=} \overset{k_0^U+2}{=} A \quad n^2;$$

Consequently,

$$\begin{aligned} \overset{k_0^U+2}{=} A^{(2)} \quad \Upsilon^{(1)} \quad n^2 \quad \overset{k_0^U+2}{=} \Upsilon^{(1)} \\ = k \quad \infty \quad = k \quad \infty \quad = k \quad \infty \quad = k \quad \infty \\ n^2 \overset{P}{=} \frac{1}{d_1} \frac{\tilde{c}_2}{\tilde{c}_1} \frac{1}{2} (1 - \tilde{c}_3 \tilde{c}_1)^{\frac{1}{2}} \quad k - k_0^U - 1 \quad (59) \\ = \tilde{q}^{k - k_0^U - 1}; \end{aligned}$$

where the last inequality is due to (56). Recall from (53) that  $\tilde{c}_1 l < \Gamma_{k|k}^{(1)} < \tilde{c}_2 l$  for  $k \geq (k_0^U; k_1^U]$ . Since

$$\Gamma_{k|k-1}^{(1)} = A_k^{(1)} \Gamma_{k-1|k-1}^{(1)} A_k^{(1)\top} + Q_{k-1}^{(1)};$$

it follows that  $\Gamma_{k|k-1}^{(1)} \overset{P}{=} \frac{1}{d_1} (\tilde{a}_1 \tilde{a}_2 \tilde{c}_2 + q_2)$  for all  $k \geq (k_0^U + 1; k_1^U]$ , and the prior error covariance of the observable subsystem satisfies  $q_1 l < \Gamma_{k|k-1}^{(1)}$  for  $k \geq (k_0^U; k_1^U]$ . As a consequence,

$$\Upsilon_k^{(1)} \overset{P}{=} \frac{1}{d_1} \Gamma_{k|k}^{(1)} \Gamma_{k|k-1}^{(1)-1} < \overset{P}{=} \frac{1}{d_1} \tilde{c}_2 q_1^{-1}; \quad \text{for } k \geq (k_0^U; k_1^U]. \quad (60)$$

Define  $\tilde{a}_4$  as

$$\tilde{a}_4 = \max_{k \in (k_0^U; k_1^U]} A_k^{(21)} \quad \infty;$$

it follows that<sup>6</sup> for  $k \geq (k_0^U + 1; k_1^U]$ ,

$$\begin{aligned}
& \text{vec } \Upsilon_k^{(1)} \Gamma_{k|k-1}^{(1)} A_k^{(21) \top} + \text{vec } \Gamma_{k+1|k}^{(12) \circ} \\
& \quad \Upsilon_k^{(1)} \Gamma_{k|k-1}^{(1)} A_k^{(21) \top} + \Gamma_{k+1|k}^{(12) \circ} \\
& < \frac{\rho^-}{d_1 \tilde{c}_2 q_1^{-1} \tilde{a}_4} \Gamma_{k|k-1}^{(1)} + q_2 \\
& < d_1 \tilde{c}_2 \tilde{a}_4 (\tilde{a}_1 \tilde{a}_2 \tilde{c}_2 + q_2) q_1^{-1} + q_2 \\
& = \tilde{\rho}:
\end{aligned} \tag{61}$$

Substituting (59) and (61) into (57), we obtain that for  $k \geq (k_0^U + 1; k_1^U]$ ,

$$\text{vec } \Gamma_{k+1|k}^{(12) \circ} b(k), \tilde{t}\tilde{q}^{k-k_0^U-1} \text{vec } \Gamma_{k_0^U+2|k_0^U+1}^{(12) \circ} + \tilde{\rho} + \tilde{t}\tilde{\rho} \tilde{q}^{k-k_0^U-2};$$

where  $b(k)$  is either a non-increasing or a non-decreasing function of  $k$ . Hence, we obtain that for  $k \geq (k_0^U; k_1^U]$ ,

$$\begin{aligned}
& \text{vec } \Gamma_{k+1|k}^{(12) \circ} \\
\max & \text{vec } \Gamma_{k_0^U+2|k_0^U+1}^{(12) \circ}; b(k_0^U + 2); \lim_{k \rightarrow \infty} b(k) \\
\max & \text{vec } \Gamma_{k_0^U+2|k_0^U+1}^{(12) \circ}; \tilde{t}\tilde{q} \text{vec } \Gamma_{k_0^U+2|k_0^U+1}^{(12) \circ} + \tilde{\rho}; \frac{\tilde{t}\tilde{\rho}\tilde{q}}{1-\tilde{q}} + \tilde{\rho};
\end{aligned}$$

where

$$\begin{aligned}
& \text{vec } \Gamma_{k_0^U+2|k_0^U+1}^{(12) \circ} \\
& < \frac{\rho^-}{n} a_1 a_2 \Gamma_{k_0^U+1|k_0^U+1}^{(12) \circ} + \frac{\rho^-}{n} n q_2 \\
& \quad \frac{\rho^-}{n} a_1 a_2 \Gamma_{k_0^U+1|k_0^U}^{(12) \circ} + \frac{\rho^-}{n} n q_2 \\
& < n \frac{\rho^-}{n} \Gamma_{k_0^U|k_0^U}^{(12) \circ} (a_1 a_2)^2 + n \frac{\rho^-}{n} a_1 a_2 q_2 + \frac{\rho^-}{n} n q_2 \\
& = \tilde{\rho}:
\end{aligned}$$

Also since

$$K_k^{(21)} = \Gamma_{k|k-1}^{(12) \circ} H_k^{(1) \top} R_k + H_k^{(1)} \Gamma_{k|k-1}^{(1)} H_k^{(1) \top}^{-1};$$

it follows that for  $k \geq (k_0^U + 1; k_1^U]$ ,

$$K_k^{(21)} \frac{\rho^-}{r_1} \Gamma_{k|k-1}^{(12) \circ} \frac{d_1 \rho^-}{r_1} \max \tilde{\rho}; \tilde{t}\tilde{q} + \tilde{\rho}; \frac{\tilde{t}\tilde{\rho}\tilde{q}}{1-\tilde{q}} + \tilde{\rho};$$

<sup>6</sup>Recall that for matrix  $M \geq R^p \times q$ ,  $kMk_{\max} = kMk_2 = \max_{1 \leq r \leq p, 1 \leq c \leq q} |M(r; c)|$ .

**Step 5:** Combining Steps 1, 2 and 4, it can be concluded that for  $k \in (k_0^U; k_1^U]$

$$\|K_k\|_\infty = U^\top K_k^{(t)} \leq n \|K_k^{(t)}\|_\infty$$

$$\frac{\rho_{\bar{3}d_1}}{r_1} \max \left\{ \frac{\rho_{\bar{n}}}{d_1} \Gamma_{k_0^U|k_0^U} a_1 a_2 + q_2; \frac{1}{d_1} (\tilde{a}_1 \tilde{a}_2 \tilde{c}_2 + q_2); \tilde{t} \tilde{q} + \tilde{\rho}; \frac{\tilde{t} \tilde{\rho} \tilde{q}}{1 - \tilde{q}} + \tilde{\rho} \right\}$$

$$\|k\|_\infty \Gamma_{k_0^U|k_0^U} ;$$

which completes the proof.

*E-mail address:* [yesun08@gmail.com](mailto:yesun08@gmail.com)

*E-mail address:* [dan.work@vanderbilt.edu](mailto:dan.work@vanderbilt.edu)