STABILITY OF NONLINEAR FILTERS

P. Chigansky

The Weizmann Institute of Science, Israel

January 2, 2005
Nonlinear filtering

Given the statistical description of the Markov process \((X_t, Y_t)_{t\geq 0}\), find the optimal in the mean square sense recursive estimate of the signal component \(f(X_t)\) for a fixed \(f\), given a trajectory of the observation process \(\{Y_s, s \leq t\}\), for any \(t \geq 0\).

In other words, find a recursive realization for

\[
\pi_t(f) := \mathbf{E}(f(X_t) | \mathcal{F}^Y_t), \quad t \geq 0
\]

where \(\mathcal{F}^Y_t = \sigma\{Y_s, s \leq t\}\).
General discrete time setting

* $X = (X_n)_{n \in \mathbb{Z}_+}$ is a Markov process with values in $S_x \subseteq \mathbb{R}$, the transition kernel $\Lambda(x, du)$ and the initial distribution $\nu$

$$P(X_n \in A | \mathcal{F}_n^{X}) = \int_A \Lambda(X_{n-1}, du), \quad P(X_0 \in A) = \nu(A).$$

* $Y = (Y_n)_{n \in \mathbb{Z}_+}$ is an i.i.d. sequence with values in $S_y \subseteq \mathbb{R}$, conditioned on $X$

$$P(Y_n \in A | \mathcal{F}_n^{X} \vee \mathcal{F}_n^{Y}) = \int_A g(X_n, u) \lambda(du)$$

The solution: Solve the filtering equation

$$\pi_n(dx) = \frac{g(u, Y_n) \int_{S_x} \Lambda(x, du) \pi_{n-1}(dx)}{\int_{S_x} g(u, Y_n) \int_{S_x} \Lambda(x, du) \pi_{n-1}(dx)}, \quad \pi_0(du) = \nu(du),$$

for the conditional distribution $\pi_n(du)$ and calculate

$$\pi_n(f) = E(f(X_n) | \mathcal{F}_n^{Y}) = \int_{S_x} f(u) \pi_n(du).$$
The stability problem

* take a probability distribution $\bar{\nu}$ on $\mathbb{S}_x$, different from $\nu$ and so that the solution of the filtering equation is well defined, if started from $\bar{\nu}$ (such pair $(\bar{\nu}, \nu)$ is admissible).

* generate the solution $\bar{\pi}_n(dx)$ of the filtering equation, started from $\bar{\nu}$ (the observation process, driving the equation, corresponds to $\nu$!)

The following notions of stability with respect to initial conditions are usually considered

1. $\lim_{n \to \infty} E((\pi_n(f) - \bar{\pi}_n(f))^2) = 0$ for any continuous and bounded $f$

2. $\liminf_{n \to \infty} n^{-1} \log |\pi_n - \bar{\pi}_n| < 0$, $P$-a.s., where $| \cdot |$ denotes the total variation distance for measures (densities)

[Q]: What are the conditions on the signal/observation model parameters (transition density, etc.) for the filter to be stable?
Rough chronology

1971 H.Kunita

\[ X \text{ is Markov-Feller} \implies \lim_{t \to \infty} E(f(X_t) - \pi_t(f))^2 \text{ is independent of } \nu \forall f \in C_b \]

1996 D.Ocone & E.Pardoux

\[ \lim_{t \to \infty} E(\pi_t(f) - \bar{\pi}_t(f))^2 = 0, \forall f \in C_b \text{ and } \nu \ll \bar{\nu} \]

1997 R.Atar & O.Zeitouni considered the stability index

\[ \gamma = \lim_{t \to \infty} \frac{1}{t} \log |\pi_t - \bar{\pi}_t| \leq 0 \]

and derived strictly negative upper bounds via:

- Oseledec’s multiplicative ergodic theorem
- Birkhoff contraction inequality for positive operators
Rough chronology (cntd.)

1999 P. Del Moral & A. Guionnet
- bounds on $\gamma$ via Dobrushin ergodic coefficient

2004 P. Baxendale, P. Ch. & R. Liptser
- spotted a serious gap in H. Kunita’s proof
- bounds on $\gamma$ by ”native” filtering arguments

Remark: The chronology is rough, omitting many interesting results, which essentially use the aforementioned methods or are applicable to very specific settings (as Kalman-Bucy, Benes filters, etc.) Some to be mentioned as the story unfolds.
Generality drop diagram

-General filtering problem

- Continuous time
- Discrete time

- Ergodic $X$

- Compact signal state space

- Finite signal state space

The least general case turns to be rich enough to
- exhibit the main difficulties
- demonstrate the known techniques
The study case: Hidden Markov Models

The signal: $X_n$ takes values in $S = \{a_1, ..., a_d\}$, has the transition matrix $\Lambda$ with the entries $\lambda_{ij} = P(X_n = a_j | X_{n-1} = a_i)$ and initial distribution $\nu$.

The observation: $Y_n = \sum_{i=1}^d 1_{\{X_n = a_i\}} \xi_n(i)$, where $\xi_n$ are i.i.d. vectors with independent entries and

$$P(\xi_1(i) \in B) = \int_B g_i(u) \lambda(du), \quad i = 1, ..., d$$

The filter: The vector $\pi_n$ with the entries $\pi_n(i) = P(X_n = a_i | \mathcal{F}_n^Y)$ satisfies

$$\pi_n = \frac{g_j(Y_n) \sum_{i=1}^d \lambda_{ij} \pi_{n-1}(i)}{\sum_{j=1}^d g_j(Y_n) \sum_{i=1}^d \lambda_{ij} \pi_{n-1}(i)} = \frac{G(Y_n) \Lambda^* \pi_{n-1}}{|G(Y_n) \Lambda^* \pi_{n-1}|}, \quad \pi_0 = \nu.$$ 

where $G(y)$, $y \in \mathbb{R}$ is a diagonal matrix with entries $g_i(y)$.

Stability: What are the conditions on $\Lambda$, $g_i(u)$’s and $(\bar{\nu}, \nu)$ so that $\lim_{n \to \infty} |\pi_n - \bar{\pi}_n| = 0$ is some sense?
**Which \((\nu, \bar{\nu})\) are admissible?**

Let

* \((\bar{X}, \bar{Y})\) denote a copy of \((X, Y)\), when \(X_0\) is sampled from \(\bar{\nu}\)
* \(Q\) and \(\bar{Q}\) are distributions induced by \((X, Y)\) and \((\bar{X}, \bar{Y})\) respectively
* \(Q^Y\) and \(\bar{Q}^Y\) are \(Y\)-marginals of \(Q\) and \(\bar{Q}\)

\[
\mathbb{P}(|G(Y_n)\Lambda^*\pi_{n-1}| = 0) = 0 \implies \text{the "correct" filtering sequence is always well defined}
\]

The "wrong" filtering sequence may not be well defined

\[
\mathbb{P}(|G(Y_n)\Lambda^*\bar{\pi}_{n-1}| = 0) \neq 0
\]

**Solution:** if \(Q^Y \ll \bar{Q}^Y\) (at least when restricted to any \([0, n]\)), then

\[
\mathbb{P}(|G(\bar{Y}_n)\Lambda^*\bar{\pi}_{n-1}| = 0) = 0 \implies \mathbb{P}(|G(Y_n)\Lambda^*\pi_{n-1}| = 0) = 0
\]

This will be the case if either of the conditions holds

1. \(\nu \ll \bar{\nu}\)
2. all entries of \(\Lambda\) are positive
3. the distribution of the noises are equivalent (mutually absolutely continuous)
I. Cul-de-sac ...?

Assume for simplicity $\nu \sim \bar{\nu}$: by standard change of measure argument

$$\bar{\pi}_n(f) = \frac{\mathbb{E}(f(X_n) \frac{d\bar{\nu}}{d\nu}(X_0)|F_n^Y)}{\mathbb{E}(\frac{d\bar{\nu}}{d\nu}(X_0)|F_n^Y)}$$

for any bounded $f$ and so

$$\mathbb{E}|\pi_n(f) - \bar{\pi}_n(f)| = \mathbb{E} \left| \mathbb{E}(f(X_n)|F_n^Y) - \frac{\mathbb{E}(f(X_n) \frac{d\bar{\nu}}{d\nu}(X_0)|F_n^Y)}{\mathbb{E}(\frac{d\bar{\nu}}{d\nu}(X_0)|F_n^Y)} \right| \leq 
\text{const.} \mathbb{E} \left| \mathbb{E} \left( \frac{d\bar{\nu}}{d\nu}(X_0)|F_n^Y \right) \mathbb{E}(f(X_n)|F_n^Y) - \mathbb{E}\left( f(X_n) \frac{d\bar{\nu}}{d\nu}(X_0)|F_n^Y \right) \right| \leq 
\text{const.} \|f\|_\infty \mathbb{E} \left| \mathbb{E} \left( \frac{d\bar{\nu}}{d\nu}(X_0)|F_n^Y \right) |F_n^Y \right) - \mathbb{E}\left( \frac{d\bar{\nu}}{d\nu}(X_0)|F_n^Y \lor \sigma\{X_n\} \right) \right| = 
\text{const.} \mathbb{E} \left| \mathbb{E} \left( \frac{d\bar{\nu}}{d\nu}(X_0)|F_n^Y \right) - \mathbb{E}\left( \frac{d\bar{\nu}}{d\nu}(X_0)|F_{[0,\infty)}^Y \lor F_{[n,\infty)}^X \right) \right| = 
$$

where $\mathcal{F}_{[n,\infty)}^X = \bigvee_{m \geq n} \sigma\{X_n, \ldots, X_m\}$, etc.
By martingale convergence

\[ \lim_{n \to \infty} E\left( \frac{d\bar{\nu}}{d\nu}(X_0) \bigg| \mathcal{F}_n^Y \right) = E\left( \frac{d\bar{\nu}}{d\nu}(X_0) \bigg| \mathcal{F}_{0,\infty}^Y \right) \]

\[ \lim_{n \to \infty} E\left( \frac{d\bar{\nu}}{d\nu}(X_0) \bigg| \mathcal{F}_{0,\infty}^Y \vee \mathcal{F}_{n,\infty}^X \right) = E\left( \frac{d\bar{\nu}}{d\nu}(X_0) \bigg| \bigcap_{n \geq 1} \mathcal{F}_{0,\infty}^Y \vee \mathcal{F}_{n,\infty}^X \right) \]

so the filter would be stable if

\[ \mathcal{F}_{0,\infty}^Y \vee \bigcap_{n \geq 1} \mathcal{F}_{n,\infty}^X \quad ? \quad \bigcap_{n \geq 1} \mathcal{F}_{0,\infty}^Y \vee \mathcal{F}_{n,\infty}^X \]

For ergodic signals (originally considered by H.Kunita, 1971) with trivial tail \( \sigma \)-algebra \( \bigcap_{n \geq 1} \mathcal{F}_{n,\infty}^X \), this reduces to the question

\[ \mathcal{F}_{0,\infty}^Y \quad ? \quad \bigcap_{n \geq 1} \mathcal{F}_{0,\infty}^Y \vee \mathcal{F}_{n,\infty}^X \]
Can \( \lor \) and \( \cap \) be interchanged?

No! Kaijser’s counterexample: consider a chain \( X_n \) with \( S = \{1, 2, 3, 4\} \) and transition matrix (this is an ergodic chain!)

\[
\Lambda = \frac{1}{2} \begin{pmatrix}
1 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 \\
0 & 0 & 1 & 1 \\
1 & 0 & 0 & 1 \\
\end{pmatrix}
\]

and let \( Y_n = 1_{\{X_n=1\}} + 1_{\{X_n=3\}} \).

Given \( \mathcal{F}^Y_{[0,\infty)} \) all the transitions \( \{1, 3\} \leftrightarrow \{2, 4\} \) can be recovered. If \( X_n \) is added at some \( n \), then the whole trajectory of \( X \) is fixed. In particular \( X_0 \) is \( \mathcal{F}^Y_{[0,\infty)} \lor \mathcal{F}^X_{[n,\infty)} \) - measurable \( \forall n \geq 1 \) and so \( \bigcap_{n \geq 0} \mathcal{F}^Y_{[0,\infty)} \lor \mathcal{F}^X_{[n,\infty)} \) - measurable.

However \( X_0 \) is not measurable w.r.t. \( \mathcal{F}^Y_{[0,\infty)} \) alone, since \( X_n \) can not be resolved within the pairs \( \{1, 3\} \) and \( \{2, 4\} \).

Remark: in this example the filter is unstable \( |\pi_n - \bar{\pi}_n| \geq C(\nu, \bar{\nu}) > 0 \) for all \( n \geq 0 \), while the signal is ergodic!
II. Stability as contraction of positive operators

For a pair of measures \( p \) and \( q \) on \( \mathbb{S} \) (i.e. vectors in the simplex \( \mathcal{S}^{d-1} \)), the Hilbert projective metric \( h(p, q) \) is defined as

\[
h(p, q) = \log \frac{\max_{q_j > 0}(p_j / q_j)}{\min_{q_i > 0}(p_i / q_i)}, \quad \text{when } p \sim q
\]

and \( h(p, q) := \infty \) for \( p \not\sim q \).

This (pseudo) metric has the following properties:

1. \( h(c_1p, c_2q) = h(p, q) \) for any positive constants \( c_1 \) and \( c_2 \).

2. For a matrix \( A \) with nonnegative entries \((A_{ij})\)

\[
h(Ap, Aq) \leq \tau(A)h(p, q)
\]

where \( \tau(A) = \frac{1 - \sqrt{\psi(A)}}{1 + \sqrt{\psi(A)}} \) is the Birkhoff contraction coefficient with

\[
\psi(A) = \min_{i,j,k,\ell} \frac{A_{ik}A_{j\ell}}{A_{i\ell}A_{j_k}}.
\]

3. \( |p - q| \leq \frac{2}{\log 3} h(p, q) \)
Some facts to recall

Ergodic chains

**Definition:** $X$ is ergodic if the limits $\lim_{n \to \infty} P(X_n = a_i) > 0$ exist and are independent of $\nu$.

**Fact:** $X$ is ergodic if and only if $\Lambda$ is $r$-primitive: there is an integer such that all the entries of $\Lambda^r$ are positive.

The Zakai equation

Both $\pi_n$ and $\bar{\pi}_n$ can be obtained by solving linear Zakai type equation,

$$\rho_n = G(Y_n)\Lambda^* \rho_{n-1}, \quad \rho_0 = \pi_0$$

and normalizing $\pi_n = \rho_n/|\rho_n|$ (and $\bar{\pi}_n = \bar{\rho}_n/|\bar{\rho}_n|$).
Back to stability

\[ h(\pi_n, \bar{\pi}_n) = h(\rho_n, \bar{\rho}_n) \leq \tau \left( \prod_{m=n-r+1}^{n} G(Y_m)\Lambda^* \right) h(\rho_{n-r}, \bar{\rho}_{n-r}) \]

and thus (note that \(|\pi_n - \bar{\pi}_n|\) is a nonincreasing sequence)

\[
\lim_{n \to \infty} \frac{1}{n} \log |\pi_n - \bar{\pi}_n| = \lim_{k \to \infty} \frac{1}{k} \log |\pi_{kr} - \bar{\pi}_{kr}| \leq \lim_{k \to \infty} \frac{1}{k} \log h(\pi_{kr}, \bar{\pi}_{kr}) \leq \\
\lim_{k \to \infty} \frac{1}{k} \sum_{\ell=1}^{k} \frac{1}{r} \log \tau \left( \prod_{m=\ell-r+1}^{\ell} G(Y_m)\Lambda^* \right) = \frac{1}{r} E_s \log \tau \left( \prod_{m=1}^{r} G(Y_m)\Lambda^* \right)
\]

where \(E_s\) is the expectation with respect to the stationary measure of \((X, Y)\).

If \(\Lambda^r\) has positive entries and the densities \(g_i(u)\) vanish only simultaneously, then the entries of \(\prod_{m=1}^{r} G(Y_m)\Lambda^*\) are positive \(P\)-a.s. and

\[
\lim_{n \to \infty} \frac{1}{n} \log |\pi_n - \bar{\pi}_n| < 0. \quad \square
\]
**Limitations**

1. If $r = 1$, i.e. if $\Lambda$ has all positive entries, then

\[
\lim_{n \to \infty} \frac{1}{n} \log |\pi_n - \bar{\pi}_n| \leq \tau(\Lambda) \leq -\lambda^*/\lambda_*, \quad \text{where } \lambda_* \leq \lambda_{ij} \leq \lambda^*
\]

independently of the noise densities. This *mixing condition* is stronger than just ergodicity of $X$.

2. The Hilbert metric approach typically fails for signals with noncompact state space: the metric can be infinite.

3. Usually requires ergodicity of the signal

**Extensions**


**Remark** The method due to Del Moral & Guionnet uses Dobrushin’s ergodic coefficient (instead of Birkhoff’s) and leads to essentially the same mixing condition
III. Oseledec’s Multiplicative ergodic theorem

Incomplete formulation: Let $A_n(\omega)$ be a stationary sequence of random $d \times d$ matrices, such that $\mathbb{E} \log^+ \|A_1\| < \infty$ and let $x_n$ be the solution of

$$x_n = A_n x_{n-1}, \quad x_0 = x \in \mathbb{R}^d.$$ 

Then there are $d$ constants $-\infty \leq \lambda_d \leq \ldots \leq \lambda_1 < \infty$ (the Lyapunov exponents) such that

$$\lim_{n \to \infty} n^{-1} \log |x_n| = \lambda_i, \quad \mathbb{P} - a.s.$$ 

for some $i$, depending on the initial vector $x$.

Moreover the norm of exterior product $x_n \wedge \bar{x}_n$ of two solutions $x_n$ and $\bar{x}_n$ (i.e. area between the vectors) corresponding to the initial conditions $x \neq \bar{x}$, grows exponentially, so that

$$\lim_{n \to \infty} n^{-1} \log |x_n \wedge \bar{x}_n| \leq \lambda_1 + \lambda_2. \quad \square$$
Application to filtering

The key is the inequality (used already by Delyon & Zeitouni, 1992)

\[ |\pi_n - \bar{\pi}_n| = \left| \frac{\rho_n}{|\rho_n|} - \frac{\bar{\rho}_n}{|\bar{\rho}_n|} \right| = \frac{\sum_{i=1}^{d} \sum_{j=1}^{d} (\rho_n(i)\bar{\rho}_n(j) - \bar{\rho}_n(i)\rho_n(j))}{|\rho_n||\bar{\rho}_n|} \leq \frac{\sum_{i=1}^{d} \sum_{j=1}^{d} |\rho_n(i)\bar{\rho}_n(j) - \bar{\rho}_n(i)\rho_n(j)|}{|\rho_n||\bar{\rho}_n|} := \frac{|\rho_n \wedge \bar{\rho}_n|}{|\rho_n||\bar{\rho}_n|}, \]

where \( a \wedge b \) is the exterior product of vectors \( a, b \) in \( \mathbb{R}^d \), i.e. the matrix with the entries \((a_i b_j - a_j b_i)\)

By the Oseledec’s MET

\[ \lambda_1 = \lim_{n \to \infty} \frac{1}{n} \log |\rho_n| = \lim_{n \to \infty} \frac{1}{n} \log |\bar{\rho}_n|, \quad \mathbb{P} - a.s. \]

where \( \lambda_1 \) is non-random top (largest) Lyapunov exponent (of the Zakai equation), independent of \( \nu \) and \( \bar{\nu} \).
By the second part of the Oseledec theorem

$$\lim_{n \to \infty} \frac{1}{n} \log |\rho_n \land \bar{\rho}_n| \leq \lambda_1 + \lambda_2$$

where $\lambda_2$ is the second Lyapunov exponent (of the Zakai equation).

**Conclusion:** the stability is controlled by the **Lyapunov spectral gap**:

$$\lim_{n \to \infty} \frac{1}{n} \log |\pi_n - \bar{\pi}_n| \leq \lim_{n \to \infty} |\rho_n \land \bar{\rho}_n| - \lim_{n \to \infty} \frac{1}{n} \log |\rho_n| -$$

$$\lim_{n \to \infty} \frac{1}{n} \log |\bar{\rho}_n| \leq (\lambda_1 + \lambda_2) - \lambda_1 - \lambda_1 = \lambda_2 - \lambda_1 \leq 0.$$  

The Lyapunov exponents are hard to calculate in general (vector) case!
High signal-to-noise asymptotic (Atar & Zeitouni, 1997)

Assume $X$ is ergodic and

$$Y_n = h(X_n) + \sigma \xi_n$$

where $\xi$ is a standard i.i.d. Gaussian sequence.

Then $\gamma(\sigma) = \lim_{n \to \infty} n^{-1} \log |\pi_n - \bar{\pi}_n|$ has the following asymptotic

$$\lim_{\sigma \to 0} \sigma^2 \gamma(\sigma) \leq -\frac{1}{2} \sum_{i=1}^{d} \mu_i \min_{j \neq i} (h(a_i) - h(a_j))^2.$$

The main tool: Kallianpur-Striebel representation for $\rho_n$ and $\rho_n \wedge \bar{\rho}_n$ (a Feynman-Kac type formula for conditional expectations)
Slow signal asymptotic (Ch., to appear)

For a small parameter $\varepsilon > 0$, let $X_n^\varepsilon$ be the Markov chain with transition probabilities

$$P(X_n = a_j | X_{n-1} = a_i) = \begin{cases} 
1 - \varepsilon \lambda_{ii}, & i = j \\
\varepsilon \lambda_{ij}, & i \neq j
\end{cases}$$

When $\varepsilon$ is small, the transition frequency drops down.

Let $Y_n^\varepsilon = \sum_{i=1}^d 1\{X_n^\varepsilon = a_i\} \xi_n(i)$ where $\xi_n$ are i.i.d. vectors with independent entries distributed with densities $g_i(u)$. Then

$$\lim_{\varepsilon \to 0} \gamma(\varepsilon) \leq - \sum_{i=1}^d \mu_i \min_{j \neq i} \mathcal{D}(g_i \| g_j)$$

where $\mathcal{D}(g_i \| g_j) = \int_{\mathbb{R}} g_i(u) \log \frac{g_i(u)}{g_j(u)} \lambda(du)$ (Kullback-Leibler divergences).

The main tool: Furstenberg-Khasminskii formulae
Limitations

1. The estimates on the Lyapunov exponents are usually impossible to calculate exactly and the obtained results are typically asymptotic.

Extensions

IV. A "native" filtering approach

The bound
\begin{align*}
E|\pi_n(f) - \bar{\pi}_n(f)| & \leq \\
& \text{const.}E\left|E\left(\frac{d\bar{\nu}}{d\nu}(X_0) \mid \mathcal{F}_n^Y\right) - E\left(\frac{d\bar{\nu}}{d\nu}(X_0) \mid \mathcal{F}_n^Y \vee \sigma\{X_n\}\right)\right|
\end{align*}

hints to consider the "reversed" filtering probabilities (Ch. & Liptser, 2004)

\[\rho_{ij}(n) = P(X_0 = a_i \mid \mathcal{F}_n^Y, X_n = a_j).\]

These satisfy linear equations, driven by \(\pi_n\), \(n \geq 0\)

\[\rho_{ij}(n) = \frac{\sum_{i=1}^{d} \lambda_{ij} \rho_{i\ell}(n - 1) \pi_{n-1}(\ell)}{\sum_{i=1}^{d} \lambda_{ij} \pi_{n-1}(\ell)}, \quad \rho_{ij}(0) = \delta_{ij}\]

The filter is stable if \(\rho_{ij}(n)\) becomes independent of \(j\) as \(n \to \infty\), i.e. if

\[\delta_i(n) := \max_j \rho_{ij}(n) - \min_m \rho_{im}(n) \xrightarrow{n \to \infty} 0, \quad \forall i = 1, ..., d\]
The sequence \( \delta_i(n) \) satisfies the inequality (recall \( \lambda^* = \max_{i,j} \lambda_{ij} \))

\[
\delta_i(n) \leq \delta_i(n - 1) \left( 1 - \frac{1}{\lambda^*} \sum_{j=1}^{d} \pi_{n-1}(j) \min_{r} \lambda_{jr} \right)
\]

and by the law of large numbers

\[
\lim_{n \to \infty} n^{-1} \log |\pi_n - \bar{\pi}_n| \leq \lim_{n \to \infty} n^{-1} \log \max_{i} \delta_i(n) \leq -\frac{1}{\lambda^*} \sum_{j=1}^{d} \mu_j \min_{r} \lambda_{jr}.
\]

Note that \( \mu_j > 0 \) and so the filter is stable if \( \Lambda \) has at least one row with all positive entries (independently of the noise densities). This is a relaxed "mixing condition" (still mixing!).

**Limitation**

Does not seem to be easily extendible to nonergodic or noncompact cases in a direct way (still can be used as a building block as in the approach due to LeGland & Oudjane).
Some open problems

1. T.Kaijser (1974) addressed the question of ergodicity (=stability) of $\pi_n$ for the following setting

   - $X$ is an ergodic Markov chain
   - $Y_n = h(X_n)$ (noiseless partial observations)

   giving only sufficient conditions on $\Lambda$ and $h$. Remarkably the sufficient and necessary conditions for this simple setting elude all the aforementioned methods (in a certain sense the presence of noise makes the problem easier!)

2. What is the weakest condition on $\Lambda$, so that the filter is stable, regardless of the noise densities? ($\{\text{ergodic } X\} + \{\min_r \lambda_{jr} > 0 \text{ for some } j\}$, is the weakest known, but not necessary ...?)

3. The noncompact state space: $\{\text{ergodic } X\} + \{\text{nowhere vanishing noise densities (or equivalent densities)}\}$ imply (non-exponential?) stability

4. Non-ergodic signals are still mysterious ...