On hypoellipticity of degenerate operators in testing and detection problems

Yuqiong Wang

yuqw@umich.edu

Joint work with Erhan Bayraktar

Department of Mathematics
University of Michigan

9th Eastern Conference on Mathematical Finance

October 25, 2025

Outline

1. Introduction and motivation

2. Problem formulation

3. Main Results

4. Examples

Outline

The classic 1D sequential testing

- A stopper observes a 1D Brownian motion with drift

$$dX_t = \theta t + dW_t, \quad X_0 = 0.$$

with $P(\theta = 1) = 1 - \theta = 0 = \pi \in [0, 1]$.

- Want to test hypotheses for its drift, e.g., $H_0 : \theta = 0$ vs $H_1 : \theta = 1$.
- We are penalized for making a mistake, and have a constant observation cost c per unit time.

- **"Sequential testing problem"**, can be formulated as an optimal stopping problem in the posterior probability process, as

$$V(\pi) = \inf_{\tau} \mathbb{E}[\Pi_\tau \wedge (1 - \Pi_\tau) + c\tau]$$

where $\Pi_t := \mathbb{P}(\theta = 1 | \mathcal{F}_t^X)$.

- Can be formulated as a free-boundary problem and solved explicitly [cf. Shiryaev (1978)].

The classic 1D sequential testing

- A stopper observes a 1D Brownian motion with drift

$$dX_t = \theta t + dW_t, \quad X_0 = 0.$$

with $P(\theta = 1) = 1 - \theta = 0 = \pi \in [0, 1]$.

- Want to test hypotheses for its drift, e.g., $H_0 : \theta = 0$ vs $H_1 : \theta = 1$.
- We are penalized for making a mistake, and have a constant observation cost c per unit time.
- **"Sequential testing problem"**, can be formulated as an optimal stopping problem in the posterior probability process, as

$$V(\pi) = \inf_\tau \mathbb{E}[\Pi_\tau \wedge (1 - \Pi_\tau) + c\tau]$$

where $\Pi_t := \mathbb{P}(\theta = 1 | \mathcal{F}_t^X)$.

- Can be formulated as a free-boundary problem and solved explicitly [cf. Shiryaev (1978)].

A challenge when in higher dimensions: degeneracy

Multi-dimension setting? many possibilities. Let's for example, look at the following problem:

- Consider now the Brownian motion has 3 possible drifts instead of 2:
  $P(\theta = i) = \pi_i, i \in \{0, 1, 2\}$.
- The sufficient statistics in this case is (Π^1, Π^2).
- There are two underlying coordinates but only one underlying Brownian source. The operator is degenerate elliptic.
- Π^i_t are functions of (t, X_t). Can formulate it in the (t, x)−coordinate: uniformly parabolic [cf. Zhitlukhin and Shiryaev(2011)].

A challenge when in higher dimensions: degeneracy

What about the following cases when they are degenerate in π?

- When θ can change its value at exponential times. e.g. classic quickest detection:

$$dX_t = 1_{\theta \geq t}dt + dW_t.$$

The posterior process $\Pi_t := \mathbb{P}(\theta > t | \theta > 0)$ depends on the whole path. No longer possible to formulate in (t, x).

- When the problem is $X-$dependent. e.g., a "hiring problem" application:

$$X_t = \theta t + W_t,$$

and the payoff upon stopping at τ being

$$e^{-r\tau}X_\tau.$$

After filtering, the problem has (X, Π) as its state.

These motivate us to study properties of these degenerate cases.

A challenge when in higher dimensions: degeneracy

What about the following cases when they are degenerate in π?

- When θ can change its value at exponential times. e.g. classic quickest detection:

$$dX_t = 1_{\theta \geq t} dt + dW_t.$$

The posterior process $\Pi_t := \mathbb{P}(\theta > t | \theta > 0)$ depends on the whole path. No longer possible to formulate in (t, x).

- When the problem is $X-$dependent. e.g., a "hiring problem" application:

$$X_t = \theta t + W_t,$$

and the payoff upon stopping at τ being

$$e^{-r\tau} X_\tau.$$

After filtering, the problem has (X, Π) as its state.

These motivate us to study properties of these degenerate cases.

A challenge when in higher dimensions: degeneracy

What about the following cases when they are degenerate in π?

- When θ can change its value at exponential times. e.g. classic quickest detection:

$$dX_t = 1_{\theta \geq t}dt + dW_t.$$

  The posterior process $\Pi_t := \mathbb{P}(\theta > t | \theta > 0)$ depends on the whole path. No longer possible to formulate in (t, x).

- When the problem is $X-$dependent. e.g., a "hiring problem" application:

$$X_t = \theta t + W_t,$$

  and the payoff upon stopping at τ being

$$e^{-r\tau}X_\tau.$$

  After filtering, the problem has (X, Π) as its state.

These motivate us to study properties of these degenerate cases.

Outline

Problem formulation

The ingredients of our problem:

- A continuous time Markov chain $(\theta_t)_{t \geq 0}$ taking values in $\{0, 1, \ldots, n\}$ with generator $Q = (q_{i,j})_{i,j \in \bar{n}}$,
- $k-$dimensional Brownian motion $W = (W^1, \ldots, W^k)$ independent of θ.
- $k < n$.

We consider

$$dX_t = \sum_{j=0}^{n} 1_{\theta_t = j} \lambda_j dt + dW_t, \quad X_0 = 0. \tag{1}$$

and problems of the form

$$V = \sup_{\tau \in \mathcal{T}} \mathbb{E}_{\pi, x} \left[e^{-\int_0^\tau r(\Pi_s, X_s) ds} g(\Pi_\tau, X_\tau) + \int_0^\tau e^{-\int_0^t r(\Pi_s, X_s) ds} h(\Pi_t, X_t) dt \right]. \tag{2}$$

- Here g, h, r are continuous, $r \geq 0$, $\lambda_i \in \mathbb{R}^k$, $i \in \{0, 1, \ldots, n\}$.
- The posterior Pi lives on the n-dimensional simplex P_{n+1}

$$\Pi_t^i = \mathbb{P}_\pi(\theta_t = i \mid \mathcal{F}_t^X) \quad \text{for } i \in \{0, \ldots n\}$$

Problem formulation

The ingredients of our problem:

- A continuous time Markov chain $(\theta_t)_{t \geq 0}$ taking values in $\{0, 1, \ldots, n\}$ with generator $Q = (q_{i,j})_{i,j \in \bar{n}}$,
- k−dimensional Brownian motion $W = (W^1, \ldots, W^k)$ independent of θ.
- $k < n$.

We consider

$$dX_t = \sum_{j=0}^{n} 1_{\theta_t = j} \lambda_j dt + dW_t, \quad X_0 = 0. \tag{1}$$

and problems of the form

$$V = \sup_{\tau \in \mathcal{T}} \mathbb{E}_{\pi, x} \left[e^{- \int_0^\tau r(\Pi_s, X_s) ds} g(\Pi_\tau, X_\tau) + \int_0^\tau e^{- \int_0^t r(\Pi_s, X_s) ds} h(\Pi_t, X_t) dt \right]. \tag{2}$$

- Here g, h, r are continuous, $r \geq 0$, $\lambda_i \in \mathbb{R}^k$, $i \in \{0, 1, \ldots, n\}$.
- The posterior Pi lives on the n-dimensional simplex P_{n+1}

$$\Pi_t^i = \mathbb{P}_\pi(\theta_t = i \mid \mathcal{F}_t^X) \quad \text{for } i \in \{0, \ldots n\}$$

Problem formulation

- By standard filtering theory, Π_t has dynamics:

$$d\Pi_t^j = \underbrace{\sum_{i=0}^n q_{ij}\Pi_t^i dt}_{\text{Drift from } Q} + \underbrace{\Pi_t^j(\lambda_j - \overline{\lambda}_t) \cdot d\tilde{W}_t}_{\text{Diffusion from } W_t}$$

- where $\overline{\lambda}_t = \sum_{i=0}^n \lambda_i \Pi_t^i$, and $\tilde{W}_t$ is the innovation process.
- The stopping problem is governed by the infinitesimal generator $\mathcal{L}_\pi$ for this Π_t process

$$\mathcal{L}_\pi = \underbrace{\frac{1}{2}\sum_{i,j=0}^n \pi_i\pi_j(\lambda_i - \overline{\lambda}) \cdot (\lambda_j - \overline{\lambda})\frac{\partial^2}{\partial\pi_i\partial\pi_j}}_{\text{Diffusion (degenerate)}}$$

$$+ \underbrace{\sum_{i,j=0}^n q_{ij}\pi_i\frac{\partial}{\partial\pi_j}}_{\text{Drift (from } Q\text{)}}$$

- **It degenerates everywhere.**

Hypoellipticity and the Hörmander's condition

- **Question:** Even if the operator is not elliptic, can we still recover regularity (e.g., smoothness) of the value function?
- **Our hope:** Hypoellipticity (the property that u smooth if $\mathcal{L}u$ smooth).
- **Intuition:** The operator may be degenerate, but the randomness "spreads" through the system.
- The "missing directions" from the k-dimensional diffusion might be restored via iterated **Lie brackets**.

Hörmander (1967)

Write $\mathcal{L}_\pi = \sum_{r=1}^k D_r^2 + D_0$, where D_i's are C^∞ vector fields. If

$$Lie(D_0, D_1, \ldots, D_K)$$

spans the tangent space of the simplex at every point in $int(P_{n+1})$, the Hörmander's condition is satisfied, and the operator $\mathcal{L}_\pi$ is hypoelliptic.

- [cf. Caffarelli and Friedman (1981)] studied the problem where $n = k$, with

$$g(\pi) = a_0(1 - \pi_0) \wedge \cdots \wedge a_n(1 - \pi_n) \quad h(\pi) = \sum_{i=0}^{n} c_i \pi_i.$$

- They commented on the case where $k < n$ and gave the 1D, 3 drift example.
- Few literature in the filtering field: [cf. Peskir (2022), Ernst et al (2022)]

Our goal: characterize when the Hörmander's condition holds for $\mathcal{L}_\pi$.

A better coordinate: Φ_t

- We do a change of coordinate to the **posterior likelihood process** Φ_t:

$$\Phi_t^i = \frac{\Pi_t^i}{\Pi_t^0} \quad \text{for } i = 1, ..., n \quad (\text{Note: } \Phi_t^0 \equiv 1)$$

- This map is a C^∞-diffeomorphism from $int(P_{n+1}) \to (0, \infty)^n$.
  Hypoellipticity is preserved.

- Define $a_i \in \mathbb{R}^k$ for $i = 1, \ldots, n$: $a_i := \lambda_i - \lambda_0$, and $\Sigma_{ij} = a_i \cdot a_j$.

Denote by $y(\phi) = \sum_{i=0}^n \phi_i$, the generator $\mathcal{L}$ for the Φ process is:

$$\mathcal{L} = \frac{1}{2} \sum_{i,j=1}^n \Sigma_{ij} \phi_i \phi_j \frac{\partial^2}{\partial \phi_i \partial \phi_j} + \frac{1}{y(\phi)} \sum_{i,j=1}^n \Sigma_{ij} \phi_i \phi_j \frac{\partial}{\partial \phi_j}$$

$$+ \sum_{j=1}^n \sum_{i=0}^n (q_{ij} - q_{i0} \phi_j) \phi_i \frac{\partial}{\partial \phi_j}$$

Outline

We first write $\mathcal{L}$ in the "sum of squares" form $\mathcal{L} = D_0 + \frac{1}{2}\sum_{r=1}^{k} D_r^2$.

- **Key observation:**
  1. The diffusion fields commute:
  $$[D_r, D_u] = 0 \quad \text{for all } r, u \in \{1, ..., k\}$$

  2. The bracket of the drift and a diffusion field **stays in the diffusion span**:
  $$[D_0, D_u] = \sum_{s=1}^{k} c_s(\phi) D_s \in \text{span}\{D_1, ..., D_k\}$$

Theorem 0

Let $Q = 0$. Let $A = (a_1, ..., a_n) \in \mathbb{R}^{k \times n}$.
If $\mathbf{n > k + 1}$, the Hömander's condition FAILS.
If $\mathbf{n = k + 1}$, the Hömander's condition HOLDS if and only if $rank(A) = k$ and the vector $(\|a_1\|^2, ..., \|a_n\|^2)$ is not in the rowspace of A.

Sufficient conditions for the detection case ($Q \neq 0$)

- Define "switch" field $J := \sum_{j=1}^{n} \sum_{i=0}^{n} (q_{i,j} - q_{i,0}\phi_j)\phi_i \partial_{\phi_j}$
- **The Key:** The Lie bracket of J with the diffusion fields D_r creates *new* vector fields.

$$[J, D_r], \quad [D_s, [J, D_r]], \quad \text{etc.}$$

Theorem 1 (Sufficient Condition 1)

The Hömander's condition holds if: (1) The drift-difference vectors $a_1, ..., a_n$ are pairwise distinct, and (2) **For each** coordinate $i \in \{1, ..., n\}$, there exists some state $m \neq i$ such that $\mathbf{q_{mi} > 0}$.

Remark: This is much weaker than Q being irreducible.

Sufficient conditions for the detection case ($Q \neq 0$)

Theorem 2 (Sufficient Condition 2)

Assume $q_{mj} = 0$ for all $m \neq j$ and that there is at least one j such that $q_{j0} > 0$ (at least one state can jump back to 0).
Define the $n \times (k + 2)$ augmented matrix $\tilde{A}$:

$$\tilde{A} = \begin{bmatrix} a_{11} & \dots & a_{k1} & \|a_1\|^2 & 1 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{1n} & \dots & a_{kn} & \|a_n\|^2 & 1 \end{bmatrix}$$

Then, dim $Lie(D_0^J, ..., D_k) = \min(rank(\tilde{A}), n)$. Hypoellipticity holds if $rank(\tilde{A}) = n$, which can **never** hold if $\mathbf{n > k + 2}$.

Parabolic Hörmander's condition is immediate:

- Define $\bar{D}_0 := -\partial_t + D_0$, dim $\text{Lie}\{D_0^J, D_1, D_2, \ldots, D_k\} = n$, then dim $\text{Lie}\{\bar{D}_0^J, D_1, D_2, \ldots, D_k\} = n + 1$.
- It follows that the process $(\Phi_t)_{t \geq 0}$ is strong Feller.
- This allows us to deal with $t-$dependent problems.

Hörmander holds in the (ϕ, x) coordinate:

- The operator $\mathcal{L}_{\phi,x}$ for $(\phi, x)-$dependent problems is hypoelliptic on $(0, \infty)^n \times \mathbb{R}^k$ if and only if the operator $\mathcal{L}$ is hypoelliptic on $(0, \infty)^n$.
- This allows us to deal with $x-$dependent problems.

Parabolic Hörmander's condition is immediate:

- Define $\bar{D}_0 := -\partial_t + D_0$, dim Lie$\{D_0^J, D_1, D_2, \ldots, D_k\} = n$, then dim Lie$\{\bar{D}_0^J, D_1, D_2, \ldots, D_k\} = n + 1$.
- It follows that the process $(\Phi_t)_{t \geq 0}$ is strong Feller.
- This allows us to deal with $t-$dependent problems.

Hörmander holds in the (ϕ, x) coordinate:

- The operator $\mathcal{L}_{\phi,x}$ for $(\phi, x)-$dependent problems is hypoelliptic on $(0, \infty)^n \times \mathbb{R}^k$ if and only if the operator $\mathcal{L}$ is hypoelliptic on $(0, \infty)^n$.
- This allows us to deal with $x-$dependent problems.

Let C be the **continuation region** (where $V > g$).

- If the Hömander condition holds, and the running payoff $h \in C^\infty(C)$, then the value function V is also $C^\infty(C)$.
- If $r, h \in C^{0,\alpha}(C)$ for $\alpha \in (0,1)$, then the value function $V \in \mathbf{C^{2,\alpha}(C)}$.

But no "smooth fit" implied:

- All these regularity results (C^∞ or $C^{2,\alpha}$) hold **only** in the $\mathcal{C}$.
- In particular, hypoellipticity alone **does not imply** the "smooth fit" condition.

Let C be the **continuation region** (where $V > g$).

- If the Hömander condition holds, and the running payoff $h \in C^\infty(C)$, then the value function V is also $C^\infty(C)$.
- If $r, h \in C^{0,\alpha}(C)$ for $\alpha \in (0, 1)$, then the value function $V \in \mathbf{C^{2,\alpha}(C)}$.

But no "smooth fit" implied:

- All these regularity results (C^∞ or $C^{2,\alpha}$) hold **only** in the $\mathcal{C}$.
- In particular, hypoellipticity alone **does not imply** the "smooth fit" condition.

Outline

Example 1: Detecting the change time in multiple coordinate

- We observe an N-dim BM. At time η, K out of N coordinates gain a drift μ.
- There are N Brownian coordinates and $\binom{N}{K}$ possible drifts.
- The generator Q has a specific structure: only the first row is non-zero.

$$Q = \begin{bmatrix} -\binom{N}{K}\lambda & \lambda & \dots & \lambda \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

  1. The drift-difference vectors a_j are all distinct.   ✓
  2. All $q_{0i} > 0$ ✓ .

- The operator **is hypoelliptic** by Theorem 1.

[cf. Ernst et al (2022)]

Example 2: Signal tracking with regime switching

- We observe $X_t = \sum_{j=0}^{n} 1_{\theta_t = j} \lambda_j t + W_t$. with

$$
Q = \begin{bmatrix}
-\sum_{i=1}^{n} q_i & q_1 & q_2 & \ldots & q_n \\
p_1 & -p_1 & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
p_k & 0 & 0 & \ldots & -p_k
\end{bmatrix}
\tag{3}
$$

with $p_i, q_j > 0$ for all $i \in \{1, k\}, j \in \{1, n\}$.

- Application-wise: monitoring signals from a radar with different levels of disorder. We can consider, e.g.,

$$
\inf_d \mathbb{E}\left[\int_0^T e^{-rt} c(d_t, \theta_t) dt\right]
$$

with

$$
c(d, \theta) = 1_{d=\theta} \sum_{i=1}^{n} c^1(\theta) + 1_{d \neq \theta} \sum_{i=1}^{n} c^2(\theta).
$$

- The operator **is hypoelliptic** by Theorem 1 (same as Example 1).

Example 3: Byzantine detection

A detection problem with possible corrupted sensor:

- We observe two processes X^1, X^2:

$$X_t^1 = \mu_0 1_{\eta < t} + \mu_1 1_{\eta \leq t} + W_t^1,$$
$$X_t^2 = m_0 1_{\eta < t} + m_1 1_{\eta \leq t} + W_t^2,$$

with $\mathbb{P}_0(\eta = 0) = p, \quad \mathbb{P}_0(\eta > t) = (1-p)e^{-rt}, \quad t \geq 0.$

- The unknown state θ has 4 possiblilities:

$$\theta_0 = \begin{cases} (\mu_1, m_1), & \text{both channels are affected at } t = 0 \\ (\mu_1, m_0), & X^1 \text{ is affected at time } t = 0, \\ (\mu_0, m_1), & X^2 \text{ is affected at time } t = 0, \\ (\mu_0, m_0), & \text{no channels are affected at time } t = 0, \end{cases} \quad (4)$$

We call it Byzantine because of the "Byzantine general problem".

Example 3: Byzantine detection

- The drift matrix

$$A = - \begin{bmatrix} \mu_1 - \mu_0 & 0 & \mu_1 - \mu_0 \\ 0 & m_1 - m_0 & m_1 - m_0 \end{bmatrix}$$

  has full rank (2), but the vector $(\|a_1\|^2, \|a_2\|^2, \|a_3\|^2)$ is always in the row-space of A.

- The generator matrix Q as

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \lambda & -\lambda & 0 & 0 \\ \lambda & 0 & -\lambda & 0 \\ \lambda & 0 & 0 & -\lambda \end{bmatrix}.$$

Theorem 3 applies, the argumented matrix

$$\tilde{A} = - \begin{bmatrix} \mu_1 - \mu_0 & 0 & (\mu_1 - \mu_0)^2 & 1 \\ 0 & m_1 - m_0 & (m_1 - m_0)^2 & 1 \\ \mu_1 - \mu_0 & m_1 - m_0 & (\mu_1 - \mu_0)^2 + (m_1 - m_0)^2 & 1 \end{bmatrix}$$

has rank 3, Hörmander condition holds. In working progress.

Thank you!