

Supplementary Materials for ”Optimal control for parameter estimation in partially observed hypoelliptic stochastic differential equations”

Quentin Clairon

University of Bordeaux, Inria Bordeaux Sud-Ouest,
Inserm, Bordeaux Population Health Research Center, SISTM Team, UMR1219,
F-33000 Bordeaux, France

Vaccine Research Institute, F-94000 Créteil, France
and

Adeline Samson *

Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP*, LJK,
38000 Grenoble, France

* Institute of Engineering Univ. Grenoble Alpes

August 12, 2021

*The author gratefully acknowledges the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01) and MIAI@Grenoble Alpes, (ANR-19-P3IA-0003).

Appendix 1: proof of propositions 1 and 2

In this section we give the proof of propositions 1 and 2 on which we build our lagged constrast function H^{m_B} . For the sake of clarity we restrict ourselves in a region $\mathcal{Z} \subseteq \mathbb{R}^d$

for each variable $j \in \llbracket 1, d \rrbracket$ we got either:

1. $\forall Z \in \mathcal{Z}, \frac{\partial g_\theta(Z)}{\partial Z_j} = 0$

or

2. $\forall Z \in \mathcal{Z}, \frac{\partial g_\theta(Z)}{\partial Z_j} \neq 0.$

Proposition 1. *If the SDE (1) is hypoelliptic then for each $l \in \llbracket 1, d_V \rrbracket$ it exists $j \in \llbracket 1, d_U \rrbracket$ and a sequence $(q_1, \dots, q_{m_l}) \in \llbracket 1, d_V \rrbracket^{m_l}$ with $m_l \leq d_V$ such that*

$$\frac{\partial g_{\theta,l}(Z_t, t)}{\partial V_{q_1}} \frac{\partial g_{\theta,q_1}(Z_t, t)}{\partial V_{q_2}} \dots \frac{\partial g_{\theta,q_{m_l-1}}(Z_t, t)}{\partial V_{q_{m_l}}} \frac{\partial g_{\theta,q_{m_l}}(Z_t, t)}{\partial U_j} \neq 0.$$

Proof. In this proof, we drop the dependence in t for the sake of clarity. Since $B_\sigma(Z_t)$ is non-singular for every possible values $Z_t \in \mathcal{Z}$, then the vectors Γ_j already span \mathbb{R}^{d_U} . Therefore hypoellipticity is ensured as soon as there are vectors belonging to \mathcal{L} which have their first d_V entries spanning \mathbb{R}^{d_V} .

Let us construct recursively the vectors defined by the generation rule of the Hörmander condition, we denote $P^{(k)}$ a vector constructed at the k -th iteration. They are defined by either $P^{(k)} = [f(Z_t), P^{(k-1)}]$ or $P^{(k)} = [\Gamma_j(Z_t), P^{(k-1)}]$ for $1 \leq j \leq d_U$ with $[\cdot, \cdot]$ the Lie bracket. Each $P^{(k)}$ is a d vector composed of two subvectors $P^{(k)} = \begin{pmatrix} P_V^{(k)} \\ P_U^{(k)} \end{pmatrix}$ and we denote $\{P_V^{(k)}\}$ the set of all generated subvectors at iteration k . Hypoellipticity implies that there is an index k such that the recursively defined matrix $M^{(k)} = \left[\left\{ P_V^{(k)} \right\}, M^{(k-1)} \right]$ is of rank d_V and especially the matrix $M^{(k)}$ has only non null rows.

We want to prove the proposition for a fixed index l . Let $P_{l,V}^{(k)}$ denote the l -th entry of $P_V^{(k)}$ which is located in the l -th row of $M^{(k)}$. We now prove recursively the proposition $\mathcal{P}(k)$: if the sequence $\left\{P_{l,V}^{(k)}\right\}_{k \in \mathbb{N}}$ becomes non null at the $k+1$ -th iteration i.e if there is $(k+1)$ such that $P_{l,V}^{(k+1)} \neq 0$ and $P_{l,V}^{(k)} = 0$ for all k , then it exists $j \in \llbracket 1, d_U \rrbracket$ and a sequence (q_1, \dots, q_{m_l}) with $m_l \leq k$ such that $\frac{\partial g_l(Z)}{\partial V_{q_1}} \frac{\partial g_{q_1}(Z)}{\partial V_{q_2}} \dots \frac{\partial g_{q_{m_l}}(Z)}{\partial U_j} \neq 0$.

Initialization: for $k=0$, $m_l=0$ and we only have to prove the existence of j such that $\frac{\partial g_l(Z)}{\partial U_j} \neq 0$. The only vectors in \mathcal{L} before any application of the generation step are the columns Γ_h with $h=1, \dots, d_U$. From this, there are 1) d_V^2 vectors $P^{(1)}$ given by $P^{(1)} = [\Gamma_h, \Gamma_j]$ and 2) d_U vectors given by $P^{(1)} = \begin{pmatrix} P_V^{(1)} \\ P_U^{(1)} \end{pmatrix} = [f(Z), \Gamma_h]$. In case 1), the d_V first entries are given by $[\Gamma_h, \Gamma_j]_r = \left(\frac{\partial \Gamma_{r,j}(Z)}{\partial z}\right) \Gamma_h - \left(\frac{\partial \Gamma_{r,h}(Z)}{\partial z}\right) \Gamma_j = 0$ for $r \in \llbracket 1, d_V \rrbracket$ and so $P_V^{(1)} = 0$. So let us focus on case 2) to find non-zero $P_V^{(1)}$. In case 2), we got $P_{l,V}^{(1)} = [f(Z), \Gamma_h]_l = -\frac{\partial g_l(Z)}{\partial z} \begin{pmatrix} 0_{d_V,1} \\ B_h(Z) \end{pmatrix} = -\frac{\partial g_l(Z)}{\partial U} B_h(Z) = -\sum_j \frac{\partial g_l(Z)}{\partial U_j} B_{j,h}(Z)$. So if $P_{l,V}^{(1)} \neq 0$, it imposes the existence of at least one j and h such that $\frac{\partial g_l(Z)}{\partial U_j} B_{j,h}(Z) \neq 0$ and so $\frac{\partial g_l(Z)}{\partial U_j} \neq 0$ which concludes the initialization.

Recursion: let us assume that \mathcal{P} is true up to iteration $k-1$. Here, we consider that $\frac{\partial g_l(Z)}{\partial U_j} = 0$ for all $j \in \llbracket 1, d_U \rrbracket$ otherwise the proof is trivial. Again, we consider two cases either 1) $P^{(k)} = [\Gamma_j(Z_t), P^{(k-1)}]$ or 2) $P^{(k)} = [f(Z_t), P^{(k-1)}]$. We start with the case 1) with vectors $P^{(k)} = [\Gamma_j(Z_t), P^{(k-1)}]$, now let us denote $r \in \llbracket 1, d_V \rrbracket$ such that $P_r^{(k-1)} = 0$. Then $P_r^{(k)}$, the r -th entry of $P_r^{(k)}$ is given by:

$$P_r^{(k)} = [\Gamma_j(Z), P^{(k-1)}]_r = \left(\frac{\partial P_r^{(k-1)}}{\partial z}\right)^T \begin{pmatrix} 0_{d_V,1} \\ B_j(Z) \end{pmatrix} - \left(\frac{\partial \Gamma_{r,j}(Z)}{\partial z}\right)^T P^{(k-1)} = 0.$$

Thus $P^{(k)}$ has the same null entries as $P^{(k-1)}$. Now let us focus on case 2), let us consider

$P_{l,V}^{(k+1)}$ such that $P_{l,V}^{(k)} = 0$. The recursive definition of $P_{l,V}^{(k+1)}$ gives:

$$\begin{aligned} P_{l,V}^{(k+1)} &= \left(\frac{\partial P_{l,V}^{(k)}}{\partial z} \right)^T f(Z) - \frac{\partial g_l(Z)^T}{\partial V} P_V^{(k)} - \frac{\partial g_l(Z)^T}{\partial U} P_U^{(k)} \\ &= - \sum_p \frac{\partial g_l(Z)}{\partial V_p} P_{p,V}^{(k)}. \end{aligned}$$

Thus $P_{l,V}^{(k+1)} \neq 0$ imposes the existence of p such that $\frac{\partial g_l(Z)}{\partial V_p} \neq 0$ and $P_{p,V}^{(k)} \neq 0$. Let us denote $k_1 < k$ the first index such that $P_{p,V}^{(k_1)} = 0$ as well as $P_{p,V}^{(k_2+1)} \neq 0$ for $k_1 \leq k_2 < k$. By using the recursion hypothesis, we derive the existence of a sequence (q_1, \dots, q_{m_p}) with $m_p \leq k_1$ such that $\frac{\partial g_p(Z)}{\partial V_{q_1}} \frac{\partial g_{q_1}(Z)}{\partial V_{q_2}} \dots \frac{\partial g_{q_{m_p-1}}(Z)}{\partial V_{q_{m_p}}} \frac{\partial g_{q_{m_p}}(Z)}{\partial U_j} \neq 0$. By multiplying the last quantity with $\frac{\partial g_l(Z)}{\partial V_p}$, the sequence (p, q_1, \dots, q_{m_p}) of length less than k is such that $\frac{\partial g_l(Z)}{\partial V_p} \frac{\partial g_p(Z)}{\partial V_{q_1}} \frac{\partial g_{q_1}(Z)}{\partial V_{q_2}} \dots \frac{\partial g_{q_{m_p-1}}(Z)}{\partial V_{q_{m_p}}} \frac{\partial g_{q_{m_p}}(Z)}{\partial U_j} \neq 0$ which completes the proof for the recursion. \square

We demonstrated the existence of paths between each smooth variable to the rough ones in the continuous setting. Now, we derive the consequence of this existence on the discretized model used for parameter estimation.

Proposition 2. *Let us assume that for each $l \in \llbracket 1, d_V \rrbracket$ it exists a unique minimal length sequence (q_1, \dots, q_{m_l}) such that the connexity property holds. Then the Euler-Maruyama approximation $(\tilde{Z}_i, \text{ for } i = 1, \dots, n)$ is such that for each $l \in \llbracket 1, d_V \rrbracket$ it exists $j \in \llbracket 1, d_U \rrbracket$ such that:*

$$\frac{\partial \tilde{V}_{l,i+k}}{\partial \tilde{U}_{j,i}} = 0, \quad \text{for } k = 1, \dots, m_l \quad \text{and} \quad \frac{\partial \tilde{V}_{l,i+m_l+1}}{\partial \tilde{U}_{j,i}} \neq 0$$

Proof. Again for the sake of notational simplicity, we drop the dependence in θ and t in drift functions g_θ and h_θ . The proof is based on the general relationship:

$$\begin{aligned} \frac{\partial \tilde{V}_{l,i+r+1}}{\partial \tilde{U}_{j,i}} &= \frac{\partial \tilde{V}_{l,i+r}}{\partial \tilde{U}_{j,i}} + \Delta \left(\sum_p \frac{\partial g_l}{\partial V_p}(\tilde{Z}_{i+r}) \frac{\partial \tilde{V}_{p,i+r}}{\partial \tilde{U}_{j,i}} + \sum_{pU} \frac{\partial g_l}{\partial U_{pU}}(\tilde{Z}_{i+r}) \frac{\partial \tilde{U}_{pU,i+r}}{\partial \tilde{U}_{j,i}} \right) \\ \frac{\partial \tilde{U}_{lU,i+r+1}}{\partial \tilde{U}_{j,i}} &= \frac{\partial \tilde{U}_{lU,i+r}}{\partial \tilde{U}_{j,i}} + \Delta \left(\sum_p \frac{\partial h_{lU}}{\partial V_p}(\tilde{Z}_{i+r}) \frac{\partial \tilde{V}_{p,i+r}}{\partial \tilde{U}_{j,i}} + \sum_{pU} \frac{\partial h_{lU}}{\partial U_{pU}}(\tilde{Z}_{i+r}) \frac{\partial \tilde{U}_{pU,i+r}}{\partial \tilde{U}_{j,i}} \right) \\ &+ \sqrt{\Delta} \frac{\partial B_{lU}}{\partial U_{j,i}}(\tilde{Z}_{i+r}) u_{i+r} \end{aligned}$$

obtained by differentiation of the discretized SDE for $l \in \llbracket 1, d_V \rrbracket$ and $l_U \in \llbracket 1, d_U \rrbracket$. Since the discretized variables can not be instantaneously affected by other ones, we get the initial values: $\frac{\partial \tilde{V}_{p,i}}{\partial \tilde{U}_{j,i}} = 0$, $\frac{\partial \tilde{U}_{p_U,i}}{\partial \tilde{U}_{j,i}} = 0$ and $\frac{\partial \tilde{V}_{j,i}}{\partial \tilde{V}_{j,i}} = 1$ for $p \in \llbracket 1, d_V \rrbracket$ and $p_U \in \llbracket 1, d_U \rrbracket \setminus \{j\}$. Then we derive:

$$\begin{aligned} \frac{\partial \tilde{V}_{l,i+1}}{\partial \tilde{U}_{j,i}} &= \Delta \left(\sum_p \frac{\partial g_l}{\partial V_p}(\tilde{Z}_i) \frac{\partial \tilde{V}_{p,i}}{\partial \tilde{U}_{j,i}} + \sum_{p_U} \frac{\partial g_l}{\partial U_{p_U}}(\tilde{Z}_i) \frac{\partial \tilde{U}_{p_U,i}}{\partial \tilde{U}_{j,i}} \right) = \Delta \frac{\partial g_l}{\partial U_j}(\tilde{Z}_i) \\ \frac{\partial \tilde{U}_{l_U,i+1}}{\partial \tilde{U}_{j,i}} &= \mathbf{1}_{\{l_U=j\}} + \Delta \left(\sum_p \frac{\partial h_{l_U}}{\partial V_p}(\tilde{Z}_i) \frac{\partial \tilde{V}_{p,i}}{\partial \tilde{U}_{j,i}} + \sum_{p_U} \frac{\partial h_{l_U}}{\partial U_{p_U}}(\tilde{Z}_i) \frac{\partial \tilde{U}_{p_U,i}}{\partial \tilde{U}_{j,i}} \right) + \sqrt{\Delta} \frac{\partial B_{l_U}}{\partial U_j}(\tilde{Z}_i) u_i \\ &= \mathbf{1}_{\{l_U=j\}} + \Delta \frac{\partial h_{l_U}}{\partial U_j}(\tilde{Z}_i) + \sqrt{\Delta} \frac{\partial B_{l_U}}{\partial U_j}(\tilde{Z}_i) u_i. \end{aligned}$$

Based on the previous equalities, we can prove our claim for the special case $m_l = 0$, we have $\frac{\partial \tilde{V}_{l,i+m_l+1}}{\partial \tilde{U}_{j,i}} = \frac{\partial \tilde{V}_{l,i+1}}{\partial \tilde{U}_{j,i}} \neq 0$ if and only if $\frac{\partial g_l}{\partial U_j} \neq 0$. For the general case $m_l > 0$, we prove by recursion the following proposition $\mathcal{P}(m)$: “for each $l \in \llbracket 1, d_V \rrbracket$ if it exists a unique minimal length sequence $(q_1^l, \dots, q_{m_l}^l)$ of length $m_l \leq m$ such that $\frac{\partial g_l(Z)}{\partial V_{q_1^l}} \dots \frac{\partial g_{q_{m_l}^l}(Z)}{\partial U_j} \neq 0$, then $\frac{\partial \tilde{V}_{l,i+k}}{\partial \tilde{U}_{j,i}} = 0$ for $k \leq m_l$ and $\frac{\partial \tilde{V}_{l,i+m_l+1}}{\partial \tilde{U}_{j,i}} \neq 0$ ”

Let us prove $\mathcal{P}(m)$ for $m = 1$ for the initialization. Based on the general relationship for $\frac{\partial \tilde{V}_{l,i+r+1}}{\partial \tilde{U}_{j,i}}$ and the results we got for $\frac{\partial \tilde{V}_{l,i+1}}{\partial \tilde{U}_{j,i}}$ and $\frac{\partial \tilde{U}_{l_U,i+1}}{\partial \tilde{U}_{j,i}}$, we derive:

$$\begin{aligned} \frac{\partial \tilde{V}_{l,i+2}}{\partial \tilde{U}_{j,i}} &= \frac{\partial \tilde{V}_{l,i+1}}{\partial \tilde{U}_{j,i}} + \Delta \left(\sum_p \frac{\partial g_l}{\partial V_p}(\tilde{Z}_{i+1}) \frac{\partial \tilde{V}_{p,i+1}}{\partial \tilde{U}_{j,i}} + \sum_{p_U} \frac{\partial g_l}{\partial U_{p_U}}(\tilde{Z}_{i+1}) \frac{\partial \tilde{U}_{p_U,i+1}}{\partial \tilde{U}_{j,i}} \right) \\ &= \Delta \frac{\partial g_l}{\partial U_j}(\tilde{Z}_i) + \Delta^2 \sum_p \frac{\partial g_l}{\partial V_p}(\tilde{Z}_{i+1}) \frac{\partial g_p}{\partial U_j}(\tilde{Z}_i) \\ &+ \Delta \sum_{p_U} \frac{\partial g_l}{\partial U_{p_U}}(\tilde{Z}_{i+1}) \frac{\partial \tilde{U}_{p_U,i+1}}{\partial \tilde{U}_{j,i}} \end{aligned}$$

Since $m_l > 0$ and $m = 1$, the length of the minimal sequence m_l is necessary 1. So there is no direct link between \tilde{V}_l and the rough part. That means $\frac{\partial g_l}{\partial U_j}(\tilde{Z}_i) = 0$ and $\frac{\partial g_l}{\partial U_{p_U}}(\tilde{Z}_{i+1}) = 0$ and we obtain $\frac{\partial \tilde{V}_{l,i+1}}{\partial \tilde{U}_{j,i}} = \Delta \frac{\partial g_l}{\partial V_j}(\tilde{Z}_i) = 0$ as well as the simplified expression $\frac{\partial \tilde{V}_{l,i+2}}{\partial \tilde{U}_{j,i}} = \Delta^2 \sum_p \frac{\partial g_l}{\partial V_p}(\tilde{Z}_{i+1}) \frac{\partial g_p}{\partial U_j}(\tilde{Z}_i)$. By assumption, the minimal sequence is unique, this imposes $\frac{\partial \tilde{U}_{l_U,i+2}}{\partial \tilde{V}_{j,i}} = \Delta^2 \sum_p \frac{\partial g_l}{\partial V_p}(\tilde{Z}_{i+1}) \frac{\partial g_p}{\partial U_j}(\tilde{Z}_i) = \Delta^2 \frac{\partial g_l}{\partial V_{q_1^l}}(\tilde{Z}_{i+1}) \frac{\partial g_{q_1^l}}{\partial U_j}(\tilde{Z}_i) \neq 0$ which ends the initialization. Now let us assume $\mathcal{P}(m)$ is true. Let us choose an arbitrary $l \in \llbracket 1, d_U \rrbracket$ and

let us assume there is a unique sequence $(q_1^l, \dots, q_{m_l+1}^l)$ of minimal length $m_l + 1$ such that $\frac{\partial g_l(Z)}{\partial V_{q_1^l}} \dots \frac{\partial g_{q_{m_l+1}^l}(Z)}{\partial U_j} \neq 0$. The general relationship gives us:

$$\begin{aligned} \frac{\partial \tilde{V}_{l,i+m_l+2}}{\partial \tilde{U}_{j,i}} &= \frac{\partial \tilde{V}_{l,i+m_l+1}}{\partial \tilde{U}_{j,i}} + \Delta \left(\sum_p \frac{\partial g_l}{\partial V_p}(Z_{i+m_l+1}) \frac{\partial \tilde{V}_{p,i+m_l+1}}{\partial \tilde{U}_{j,i}} + \sum_{pU} \frac{\partial g_l}{\partial U_{pU}}(\tilde{Z}_{i+m_l+1}) \frac{\partial \tilde{U}_{pU,i+m_l+1}}{\partial \tilde{U}_{j,i}} \right) \\ \frac{\partial \tilde{V}_{l,i+m_l+1}}{\partial \tilde{U}_{j,i}} &= \frac{\partial \tilde{V}_{l,i+m_l}}{\partial \tilde{U}_{j,i}} + \Delta \left(\sum_p \frac{\partial g_l}{\partial V_p}(\tilde{Z}_{i+m_l}) \frac{\partial \tilde{V}_{p,i+m_l}}{\partial \tilde{U}_{j,i}} + \sum_{pU} \frac{\partial g_l}{\partial U_{pU}}(\tilde{Z}_{i+m_l}) \frac{\partial \tilde{U}_{pU,i+m_l}}{\partial \tilde{U}_{j,i}} \right) \end{aligned}$$

Necessarily $\frac{\partial g_l}{\partial U_{pU}} = 0$ and by recursion hypothesis, we got $\frac{\partial \tilde{V}_{l,i+k}}{\partial \tilde{U}_{j,i}} = 0$ for every $k \leq m_l$ since there is no minimal sequence of length m_l or lower, so:

$$\begin{aligned} \frac{\partial \tilde{V}_{l,i+m_l+2}}{\partial \tilde{U}_{j,i}} &= \frac{\partial \tilde{V}_{l,i+m_l+1}}{\partial \tilde{U}_{j,i}} + \Delta \sum_p \frac{\partial g_l}{\partial V_p}(\tilde{Z}_{i+m_l+1}) \frac{\partial \tilde{V}_{p,i+m_l+1}}{\partial \tilde{U}_{j,i}} \\ \frac{\partial \tilde{V}_{l,i+m_l+1}}{\partial \tilde{U}_{j,i}} &= \Delta \sum_p \frac{\partial g_l}{\partial V_p}(\tilde{Z}_{i+m_l}) \frac{\partial \tilde{V}_{p,i+m_l}}{\partial \tilde{U}_{j,i}}. \end{aligned}$$

Now let us assume $\frac{\partial \tilde{V}_{l,i+m_l+1}}{\partial \tilde{U}_{j,i}} \neq 0$, this implies there is at least one p such that $\frac{\partial g_l}{\partial V_p}(Z) \neq 0$ and $\frac{\partial \tilde{V}_{p,i+m_l}}{\partial \tilde{U}_{j,i}} \neq 0$. By contraposition of the recursion hypothesis if there is no minimal sequence $(q_1^p, \dots, q_{m_l-1}^p)$ such that $\frac{\partial g_p(Z)}{\partial V_{q_1^p}} \dots \frac{\partial g_{q_{m_l-1}^p}(Z)}{\partial U_j} \neq 0$, then $\frac{\partial \tilde{V}_{p,i+m_l}}{\partial \tilde{U}_{j,i}} = 0$. So necessarily there is a sequence such that $\frac{\partial g_p(Z)}{\partial V_{q_1^p}} \dots \frac{\partial g_{q_{m_l-1}^p}(Z)}{\partial U_j} \neq 0$ to ensure $\frac{\partial \tilde{V}_{p,i+m_l}}{\partial \tilde{U}_{j,i}} \neq 0$. By multiplying this last sequence by $\frac{\partial g_l}{\partial V_p}(Z)$, we conclude the existence of the sequence $(p, q_1^p, \dots, q_{m_l-1}^p)$ of length m_l such that $\frac{\partial g_l}{\partial V_p}(Z) \frac{\partial g_p(Z)}{\partial V_{q_1^p}} \dots \frac{\partial g_{q_{m_l-1}^p}(Z)}{\partial U_j} \neq 0$ in contradiction with the minimality of $(q_1^l, \dots, q_{m_l+1}^l)$. So by contradiction $\frac{\partial \tilde{V}_{l,i+m_l+1}}{\partial \tilde{U}_{j,i}} = 0$. From this, we get: $\frac{\partial \tilde{V}_{l,i+m_l+2}}{\partial \tilde{U}_{j,i}} = \Delta \sum_p \frac{\partial g_l}{\partial V_p}(\tilde{Z}_{i+m_l}) \frac{\partial \tilde{V}_{p,i+m_l}}{\partial \tilde{U}_{j,i}}$, assuming $\frac{\partial \tilde{V}_{l,i+m_l+2}}{\partial \tilde{U}_{j,i}} = 0$ leads to have for all $p \in \llbracket 1, d_V \rrbracket$ and sequences $(q_1^p, \dots, q_{m_l}^p)$ the equality $\frac{\partial g_p(Z)}{\partial V_{q_1^p}} \dots \frac{\partial g_{q_{m_l}^p}(Z)}{\partial U_j} = 0$ and so $\frac{\partial g_l}{\partial V_p}(Z) \frac{\partial g_p(Z)}{\partial V_{q_1^p}} \dots \frac{\partial g_{q_{m_l}^p}(Z)}{\partial U_j} = 0$ in contradiction with the existence of $(q_1^l, \dots, q_{m_l+1}^l)$, thus $\frac{\partial \tilde{V}_{l,i+m_l+2}}{\partial \tilde{U}_{j,i}} \neq 0$. We proved that if $\mathcal{P}(m)$ is true for sequence $m_l \leq m$, the property holds as well for sequences of length $m_l + 1$ and so $\mathcal{P}(m + 1)$ is true as well. This concludes the recursion and the proof. \square

Appendix 2: \bar{u}_{Z_0} and \widehat{Z}_0 formal derivation

Let us introduce the following matrices for $i = 0, \dots, n$:

$$\bar{\mathbf{A}}_{\theta,i} = \begin{pmatrix} \mathbf{A}_{\theta,i} & \Delta_i r_{\theta}(t_i) \\ 0_{1,d} & 1 \end{pmatrix}, Q_i = \begin{pmatrix} C^T C & -C^T Y_i \\ -Y_i^T C & Y_i^T Y_i \end{pmatrix}, \mathbf{\Gamma}_{\sigma}(t_i) = \begin{pmatrix} \sqrt{\Delta} \mathbf{\Gamma}_{\sigma}(t_i) \\ 0_{1,d_u} \end{pmatrix}, P_i = \frac{1}{w} I_m.$$

Let us extend the process Z by introducing $\mathbf{Z}_0^d = (Z_0^T, 1)^T$ and by iteration, $\mathbf{Z}_i^d = (\tilde{Z}_i, 1)$ the extended state-space variable:

$$\left\{ \begin{array}{l} \mathbf{Z}_{i+1}^d = \bar{\mathbf{A}}_{\theta,i} \mathbf{Z}_{i+1}^d + \mathbf{\Gamma}_{\sigma}(t_i) u_i \end{array} \right. \quad (1)$$

as well as the reformulation of the control problem for the extended process \mathbf{Z}^d :

$$\begin{array}{l} \text{Minimize in } u: C_w(u|Y; Z_0) = (\mathbf{Z}_n^d)^T Q_n \mathbf{Z}_n^d + \sum_{i=1}^{n-1} \left((\mathbf{Z}_i^d)^T Q_i \mathbf{Z}_i^d + u_i^T P_i u_i \right) \\ \text{Subject to: } \left\{ \begin{array}{l} \mathbf{Z}_{i+1}^d = \bar{\mathbf{A}}_{\theta,i} \mathbf{Z}_{i+1}^d + \mathbf{\Gamma}_{\sigma}(t_i) u_i \\ \mathbf{Z}_0^d = (Z_0, 1) \end{array} \right. \end{array} \quad (2)$$

By introducing $E_{\theta,k}^d$ the solution of the discrete Riccati equation:

$$\left\{ \begin{array}{l} R_i = Q_i + \bar{\mathbf{A}}_{\theta,i}^T R_{i+1} \bar{\mathbf{A}}_{\theta,i} \\ -\bar{\mathbf{A}}_{\theta,i}^T R_{i+1} \mathbf{\Gamma}_{\sigma}(t_i) (P_i + \mathbf{\Gamma}_{\sigma}(t_i)^T R_{i+1} \mathbf{\Gamma}_{\sigma}(t_i))^{-1} \mathbf{\Gamma}_{\sigma}(t_i)^T R_{j+1} \bar{\mathbf{A}}_{\theta,i} \\ R_n = Q_n \end{array} \right. \quad (3)$$

we know that the solution of (2) is equal to

$$\bar{u}_{Z_0,i} = -(P_i + \mathbf{\Gamma}_{\sigma}(t_i)^T R_{i+1} \mathbf{\Gamma}_{\sigma}(t_i))^{-1} \mathbf{\Gamma}_{\sigma}(t_i)^T R_{i+1} \bar{\mathbf{A}}_{\theta,i} \bar{\mathbf{Z}}_i^d.$$

The corresponding cost value is given by $C_w(\bar{u}_{Z_0}|Y; Z_0) = (\mathbf{Z}_0^d)^T R_0 \bar{\mathbf{Z}}_0^d$ (see Bertsekas (2005)). We now derive by reversed time induction that the solution of (3) is a sequence of

symmetric matrices i.e. R_i has the block structure $R_i^d = \begin{pmatrix} E_i & h_i \\ (h_i)^T & \alpha_i^d \end{pmatrix}$ where each block respects a specific finite difference equation. Let us assume the decomposition holds for $i + 1$ i.e. $R_{i+1} = \begin{pmatrix} E_{i+1} & h_{i+1} \\ (h_{i+1})^T & \alpha_{i+1}^d \end{pmatrix}$ Let us compute $\overline{\mathbf{A}}_{\theta,i}^T R_{i+1} \overline{\mathbf{A}}_{\theta,i}$. We have:

$$\begin{aligned}
& \overline{\mathbf{A}}_{\theta,i}^T R_{i+1} \overline{\mathbf{A}}_{\theta,i} \\
&= \begin{pmatrix} \mathbf{A}_{\theta,i} & \Delta_i r_\theta(t_i) \\ 0_{1,d} & 1 \end{pmatrix}^T \begin{pmatrix} E_{i+1} & h_{i+1} \\ (h_{i+1})^T & \alpha_{i+1}^d \end{pmatrix} \begin{pmatrix} \mathbf{A}_{\theta,i} & \Delta_i r_\theta(t_i) \\ 0_{1,d} & 1 \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{A}_{\theta,i} & \Delta_i r_\theta(t_i) \\ 0_{1,d} & 1 \end{pmatrix}^T \begin{pmatrix} E_{i+1} \mathbf{A}_{\theta,i} & \Delta_i E_{i+1} r_\theta(t_i) + h_{i+1} \\ (h_{i+1})^T \mathbf{A}_{\theta,i} & \Delta_i (h_{i+1})^T r_\theta(t_i) + \alpha_{i+1}^d \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{A}_{\theta,i}^T & 0_{d,1} \\ \Delta_i r_\theta(t_i)^T & 1 \end{pmatrix} \begin{pmatrix} E_{i+1} \mathbf{A}_{\theta,i} & \Delta_i E_{i+1} r_\theta(t_i) + h_{i+1} \\ (h_{i+1})^T \mathbf{A}_{\theta,i} & \Delta_i (h_{i+1})^T r_\theta(t_i) + \alpha_{i+1}^d \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{A}_{\theta,i}^T E_{i+1} \mathbf{A}_{\theta,i} & \Delta_i \mathbf{A}_{\theta,i}^T E_{i+1} r_\theta(t_i) + \mathbf{A}_{\theta,i}^T h_{i+1} \\ \Delta_i r_\theta(t_i)^T E_{i+1} \mathbf{A}_{\theta,i} + (h_{i+1})^T \mathbf{A}_{\theta,i} & \Delta_i r_\theta(t_i)^T (\Delta_i E_{i+1} r_\theta(t_i) + 2h_{i+1}) + \alpha_{i+1}^d \end{pmatrix} \\
&= F(E_{i+1}) := \begin{pmatrix} F_1(E_{i+1}) & F_2(E_{i+1}, h_{i+1}) \\ F_2(E_{i+1}, h_{i+1})^T & F_3(E_{i+1}, h_{i+1}, \alpha_{\theta,k+1}^d) \end{pmatrix}
\end{aligned}$$

where the symmetric matrix F is easily derivable by identification and the d dimensional square matrix F_1 only depends on E_{i+1} . Now let us compute the expression in $[P_i + \mathbf{\Gamma}_\sigma(t_i)^T R_{i+1} \mathbf{\Gamma}_\sigma(t_i)]^{-1}$. We obtain:

$$\begin{aligned}
[P_i + \mathbf{\Gamma}_\sigma(t_i)^T R_{i+1} \mathbf{\Gamma}_\sigma(t_i)]^{-1} &= \left[\frac{1}{w} I_m + \begin{pmatrix} \sqrt{\Delta} \mathbf{\Gamma}_\sigma(t_i) \\ 0_{1,d_u} \end{pmatrix}^T \begin{pmatrix} E_{i+1} & h_{i+1} \\ (h_{i+1})^T & \alpha_{i+1}^d \end{pmatrix} \begin{pmatrix} \sqrt{\Delta} \mathbf{\Gamma}_\sigma(t_i) \\ 0_{1,d_u} \end{pmatrix} \right]^{-1} \\
&= \left[\frac{1}{w} I_m + \Delta \begin{pmatrix} E_{i+1} \mathbf{\Gamma}_\sigma(t_i) \\ (h_{i+1})^T \mathbf{\Gamma}_\sigma(t_i) \end{pmatrix} \right]^{-1} \\
&= \left[\frac{1}{w} I_m + \Delta \begin{pmatrix} \mathbf{\Gamma}_\sigma(t_i)^T & 0_{d_u,1} \end{pmatrix} \begin{pmatrix} E_{i+1} \mathbf{\Gamma}_\sigma(t_i) \\ (h_{i+1})^T \mathbf{\Gamma}_\sigma(t_i) \end{pmatrix} \right]^{-1} \\
&= \left[\frac{1}{w} I_m + \Delta \mathbf{\Gamma}_\sigma(t_i)^T E_{i+1} \mathbf{\Gamma}_\sigma(t_i) \right]^{-1} := G(E_{i+1})
\end{aligned}$$

with G a symmetric matrix valued function (as the inverse of a symmetric square matrix).

Moreover, we have:

$$\begin{aligned}
&\overline{\mathbf{A}}_{\theta,i}^T R_{i+1} \mathbf{\Gamma}_\sigma(t_i) \\
&= \begin{pmatrix} \mathbf{A}_{\theta,i} & \Delta_i r_\theta(t_i) \\ 0_{1,d} & 1 \end{pmatrix}^T \begin{pmatrix} E_{i+1} & h_{i+1} \\ (h_{i+1})^T & \alpha_{i+1}^d \end{pmatrix} \begin{pmatrix} \sqrt{\Delta} \mathbf{\Gamma}_\sigma(t_i) \\ 0_{1,d_u} \end{pmatrix} \\
&= \sqrt{\Delta} \begin{pmatrix} \mathbf{A}_{\theta,i}^T & 0_{d,1} \\ \Delta_i r_\theta(t_i)^T & 1 \end{pmatrix} \begin{pmatrix} E_{i+1} \mathbf{\Gamma}_\sigma(t_i) \\ (h_{i+1})^T \mathbf{\Gamma}_\sigma(t_i) \end{pmatrix} \\
&= \sqrt{\Delta} \begin{pmatrix} \mathbf{A}_{\theta,i}^T E_{i+1} \mathbf{\Gamma}_\sigma(t_i) \\ \Delta_i r_\theta(t_i)^T E_{i+1} \mathbf{\Gamma}_\sigma(t_i) + (h_{i+1})^T \mathbf{\Gamma}_\sigma(t_i) \end{pmatrix}
\end{aligned}$$

so we can compute

$$\begin{aligned}
& \bar{\mathbf{A}}_{\theta,i}^T R_{i+1} \mathbf{\Gamma}_\sigma(t_i) G(E_{i+1}) \mathbf{\Gamma}_\sigma(t_i)^T R_{j+1} \bar{\mathbf{A}}_{\theta,i} \\
&= \sqrt{\Delta} \begin{pmatrix} \mathbf{A}_{\theta,i}^T E_{i+1} \mathbf{\Gamma}_\sigma(t_i) \\ \Delta_i r_\theta(t_i)^T E_{i+1} \mathbf{\Gamma}_\sigma(t_i) + (h_{i+1})^T \mathbf{\Gamma}_\sigma(t_i) \end{pmatrix} G(E_{i+1}) \sqrt{\Delta} \begin{pmatrix} \mathbf{A}_{\theta,i}^T E_{i+1} \mathbf{\Gamma}_\sigma(t_i) \\ \Delta_i r_\theta(t_i)^T E_{i+1} \mathbf{\Gamma}_\sigma(t_i) + (h_{i+1})^T \mathbf{\Gamma}_\sigma(t_i) \end{pmatrix}^T \\
&= \Delta \begin{pmatrix} \mathbf{A}_{\theta,i}^T E_{i+1} \mathbf{\Gamma}_\sigma(t_i) \\ (\Delta_i E_{i+1} r_\theta(t_i) + h_{i+1})^T \mathbf{\Gamma}_\sigma(t_i) \end{pmatrix} G(E_{i+1}) \begin{pmatrix} \mathbf{\Gamma}_\sigma(t_i)^T E_{i+1} \mathbf{A}_{\theta,i} & \mathbf{\Gamma}_\sigma(t_i)^T (h_{i+1} + \Delta_i E_{i+1} r_\theta(t_i)) \end{pmatrix} \\
&:= \Delta \begin{pmatrix} H_1(E_{i+1}) & H_2(E_{i+1}, h_{i+1}) \\ H_2(E_{i+1}, h_{i+1})^T & H_3(E_{i+1}, h_{i+1}) \end{pmatrix}.
\end{aligned}$$

By re-injecting all the derived expression in (3), we obtain:

$$\begin{aligned}
R_i &= \begin{pmatrix} F_1(E_{i+1}) & F_2(E_{i+1}, h_{i+1}) \\ F_2(E_{i+1}, h_{i+1})^T & F_3(E_{i+1}, h_{i+1}, \alpha_{\theta,k+1}^d) \end{pmatrix} + \begin{pmatrix} C^T C & -C^T Y_i \\ -Y_i^T C & Y_i^T Y_i \end{pmatrix} \\
&\quad - \Delta \begin{pmatrix} H_1(E_{i+1}) & H_2(E_{i+1}, h_{i+1}) \\ H_2(E_{i+1}, h_{i+1})^T & H_3(E_{i+1}, h_{i+1}) \end{pmatrix}
\end{aligned}$$

Thus R_i is indeed symmetric and has the required form. Hence the recursion. We also obtain the following finite difference equation

$$\begin{aligned}
E_i^d &= \mathbf{A}_{\theta,i}^T E_{i+1} \mathbf{A}_{\theta,i} + C^T C - \Delta \mathbf{A}_{\theta,i}^T E_{i+1} \mathbf{\Gamma}_\sigma(t_i) G(E_{i+1}) \mathbf{\Gamma}_\sigma(t_i)^T E_{i+1} \mathbf{A}_{\theta,i} \\
h_i &= \Delta_i \mathbf{A}_{\theta,i}^T E_{i+1} r_\theta(t_i) + \mathbf{A}_{\theta,i}^T h_{i+1} - C^T Y_i \\
&\quad - \Delta \mathbf{A}_{\theta,i}^T E_{i+1} \mathbf{\Gamma}_\sigma(t_i) G(E_{i+1}) \mathbf{\Gamma}_\sigma(t_i)^T (h_{i+1} + \Delta_i E_{i+1} r_\theta(t_i)).
\end{aligned}$$

From the previous expression for R_i , we derive for the related optimal control \bar{u}_i^d :

$$\begin{aligned}
\bar{u}_i &= -(P_i + \mathbf{\Gamma}_\sigma(t_i)^T R_{i+1} \mathbf{\Gamma}_\sigma(t_i))^{-1} \mathbf{\Gamma}_\sigma(t_i)^T R_{i+1} \bar{\mathbf{A}}_{\theta,i} \begin{pmatrix} \bar{Z}_i \\ 1 \end{pmatrix} \\
&= -G(E_{i+1}) \sqrt{\Delta} \begin{pmatrix} \mathbf{\Gamma}_\sigma(t_i)^T & 0_{d_u,1} \end{pmatrix} \begin{pmatrix} E_{i+1} & h_{i+1} \\ (h_{i+1})^T & \alpha_{i+1}^d \end{pmatrix} \begin{pmatrix} \mathbf{A}_{\theta,i} & \Delta_i r_\theta(t_i) \\ 0_{1,d} & 1 \end{pmatrix} \begin{pmatrix} \bar{Z}_i \\ 1 \end{pmatrix} \\
&= -\sqrt{\Delta} G(R_{\theta,k+1}^d) \begin{pmatrix} \mathbf{\Gamma}_\sigma(t_i)^T & 0_{d_u,1} \end{pmatrix} \begin{pmatrix} E_{i+1} & h_{i+1} \\ (h_{i+1})^T & \alpha_{i+1}^d \end{pmatrix} \begin{pmatrix} \mathbf{A}_{\theta,i} \bar{Z}_i + \Delta_i r_\theta(t_i) \\ 1 \end{pmatrix} \\
&= -\sqrt{\Delta} G(R_{\theta,k+1}^d) \begin{pmatrix} \mathbf{\Gamma}_\sigma(t_i)^T & 0_{d_u,1} \end{pmatrix} \begin{pmatrix} E_{i+1} (\mathbf{A}_{\theta,i} \bar{Z}_i + \Delta_i r_\theta(t_i)) + h_{i+1} \\ (h_{i+1})^T (\mathbf{A}_{\theta,i} \bar{Z}_i + \Delta_i r_\theta(t_i)) + \alpha_{i+1}^d \end{pmatrix} \\
&= -\sqrt{\Delta} G(R_{\theta,k+1}^d) \mathbf{\Gamma}_\sigma(t_i)^T (E_{i+1} (\mathbf{A}_{\theta,i} \bar{Z}_i + \Delta_i r_\theta(t_i)) + h_{i+1}).
\end{aligned}$$

From these equations, we easily derive that

$$\begin{aligned}
\widehat{Z}_0 &= \arg \min_{Z_0} C_w(\bar{u}_{Z_0} | Y; Z_0) \\
&= \arg \min_{Z_0} \left\{ (\mathbf{Z}_0^d)^T \begin{pmatrix} E_0^d & h_0^d \\ (h_0^d)^T & \alpha_0^d \end{pmatrix} \bar{\mathbf{Z}}_0^d \right\} \\
&= \arg \min_{Z_0} \left\{ Z_0^T R_0^d Z_0 + 2 (h_0^d)^T Z_0 + \alpha_0^d \right\} \\
&= - (E_0^d)^{-1} h_0^d.
\end{aligned}$$

References

Bertsekas, D. (2005). *Dynamic Programming and Optimal Control*. Athena Scientific.