Delft University of Technology

Delft Center for Systems and Control

Technical report 15-025

Properties of applying a resource allocation coordination algorithm to optimization problems with discrete decision variables*

R. Luo, R. Bourdais, T.J.J. van den Boom, and B. De Schutter

September 2015

Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft
The Netherlands
phone: +31-15-278.24.73 (secretary)
URL: https://www.dcsc.tudelft.nl

Properties of applying a resource allocation coordination algorithm to optimization problems with discrete decision variables

Renshi Luo, Romain Bourdais, Ton J.J. van den Boom, Bart De Schutter

Abstract

This addendum contains the proof of the optimality conditions and the proof of the oscillation detecting conditions for discrete decision variables when the resource allocation coordination algorithm is directly applied to optimization problems with discrete decision variables. Those conditions are used in the paper "Multi-agent model predictive control based on resource allocation coordination for a class of hybrid systems with limited information sharing" by R. Luo, R. Bourdais, T.J.J. van den Boom and B. De Schutter, *Engineering Applications of Artificial Intelligence*, vol. 58, pp. 123–133, 2017.

1 Questions needed to be answered

In this document, we will explain:

- Whether the global optimum is found if none of the decision variables oscillates
- How to detect oscillations of discrete decision variables

2 Answer to question 1

We consider

$$\min_{u} \sum_{i=1}^{N} f_i(u_i) \tag{1}$$

subject to

$$u_i \in D_i$$

$$\sum_{i=1}^{N} g_i(u_i) \leq r$$

We assume:

- u_i is a scalar
- $f_i(\cdot)$ is a convex function
- $g_i(\cdot)$ is a monotonically strictly increasing function.
- D_i is a finite discrete set

By defining

- if $f_i(\cdot)$ is strictly convex, then

$$x_i^{\text{best}} = \arg\min_{u_i \in D_i} f_i(u_i)$$

- if $f_i(\cdot)$ is not strictly convex, then

$$f_i^{\text{best}} = \min_{u_i \in D_i} f_i(u_i)$$
$$u_i^{\text{best}} = \min_{u_i \in D_i,\ f_i(u_i) = f_i^{\text{best}}} u_i$$

three cases can occur:

case 1: $\sum_{i=1}^{N} g_i(u_i^{\text{best}}) < r$

case 2: $\sum_{i=1}^{N} g_i(u_i^{\text{best}}) = r$

case 3: $\sum_{i=1}^{N} g_i(u_i^{\text{best}}) > r$

Decompose the overall problem into N subproblems:

- if $f_i(\cdot)$ is strictly convex, then subproblem i is defined by

$$\begin{aligned} &\min_{u_i} f_i(u_i) \\ &\text{subject to} \\ &\qquad u_i \in D_i \\ &\qquad g_i(u_i) \leq \theta_i \end{aligned} \tag{2}$$

- if $f_i(\cdot)$ is not strictly convex, then subproblem i is defined by

$$\begin{aligned} &\min_{x_i} f_i(u_i) \\ &\text{subject to} \\ &\qquad u_i \in D_i \\ &\qquad g_i(u_i) \leq \theta_i \\ &\qquad u_i \leq u_i^{\text{best}} \end{aligned}$$

with $\sum_{i=1}^{N} \theta_i = r$.

Before applying the resource allocation coordination method to the three cases, we want to stress for any subproblem that given θ_i

- if $\theta_i \geq g_i(u_i^{\text{best}})$, constraint $g_i(u_i) \leq \theta_i$ does not pose any restriction. Therefore, $u_i^* = u_i^{\text{best}}$ and[1] $\lambda_i = 0$.

[1]When $\theta_i = g_i(u_i^{\text{best}})$, the corresponding λ_i is free. However, in that case we set it equal to 0 by definition.

- $\theta_i < g_i(u_i^{\text{best}}) \to u_i^* < u_i^{\text{best}} \to \lambda_i = -\frac{f_i^{'}(u_i^*)}{g_i^{'}(u_i^*)} > 0$ with $f_i^{'}(u_i^*) < 0$ since $f_i(u_i^*) > f_i(u_i^{\text{best}})$ and $f_i(\cdot)$ is convex and with $g_i^{'}(u_i^*) > 0$ since $g_i(u_i^*)$ is a monotonically strictly increasing function.

where u_i^* is the solution of subproblem i with θ_i given and λ_i is the Lagrange multiplier corresponding to the constraint $g_i(u_i) \leq \theta_i$ in the subproblem.

Note that if the resource allocation coordination method is used, the resource allocation at each iteration is updated by

$$\theta_i^{(z+1)} = \theta_i^{(z)} + \xi^{(z)} \left(\lambda_i^{(z)} - \frac{1}{N} \sum_{j=1}^{N} \lambda_j^{(z)} \right) \tag{3}$$

with

$$\xi^{(z)} > 0, \lim_{z \to +\infty} \xi^{(z)} = 0, \sum_{z=1}^{+\infty} \xi^{(z)} = +\infty, \quad \sum_{z=1}^{+\infty} (\xi^{(z)})^2 < +\infty$$

In the following, I will start with $N = 2$ to prove some properties of the evolution of $\theta_i^{(z)}$ and $u_i^{*,(z)}$ when the resource allocation coordination is used.

2.1 For case 1

In this case, $\sum_{i=1}^{2} g_i(u_i^{\text{best}}) < r$. Since $\sum_{i=1}^{2} \theta_i = r$, we consider the following three modes:

- mode 1.1: $\theta_1^{(z)} > g_1(u_1^{\text{best}})$ and $\theta_2^{(z)} < g_2(u_2^{\text{best}})$
- mode 1.2: $\theta_1^{(z)} < g_1(u_1^{\text{best}})$ and $\theta_2^{(z)} > g_2(u_2^{\text{best}})$
- mode 1.3: $\theta_1^{(z)} \geq g_1(u_1^{\text{best}})$ and $\theta_2^{(z)} \geq g_2(u_2^{\text{best}})$

The mode transition diagram of case 1 is shown in Figure 1.

Definition: A persistent mode is such a mode that once it has been reached, the system stays in that mode.

Proposition 2.1.1: Let $\delta_2 = \min_{u_2 \in D_2,\, u_2 < u_2^{\text{best}}} -\frac{f_2^{'}(u_2)}{g_2^{'}(u_2)}$, then $\lambda_2^{(z)} \geq \delta_2 > 0$ holds for all z in mode 1.1.
Proof:
Since $-\frac{f_2^{'}(u_2)}{g_2^{'}(u_2)} > 0$ holds for all $u_2 < u_2^{\text{best}}$ and $u_2 \in D_2$, it is directly proved that $\delta_2 > 0$.

If mode 1.1 is active at any step z, then $\theta_2^{(z)} < g_2(u_2^{\text{best}})$. Therefore, $u_2^{*,(z)} < u_2^{\text{best}}$. Also note that $u_2^{*,(z)} \in D_2$. Hence, $\lambda_2^{(z)} = -\frac{f_2^{'}(u_2^{*,(z)})}{g_2^{'}(u_2^{*,(z)})} \geq \delta_2$ always holds. $\square$

Proposition 2.1.2: Mode 1.1 is not persistent.
Proof:
Let us now assume that the system stays in mode 1.1 from step z_0 on and show that this leads to a contradiction.

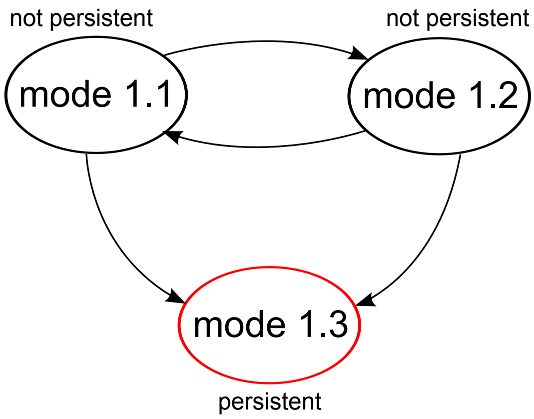

Figure 1: Mode transition diagram of case 1 with the final mode marked in red

If the system in mode 1.1 at step z_0, then $\theta_1^{(z_0)} > g_1(u_1^{\text{best}})$ and $\theta_2^{(z_0)} < g_2(u_2^{\text{best}})$. In this mode, $u_1^{*,(z_0)} = u_1^{\text{best}}$, $\lambda_1^{(z_0)} = 0$ and $u_2^{*,(z_0)} < u_2^{\text{best}}$, $\lambda_2^{(z_0)} > 0$. According to (3), in this mode, at step $z_0 + 1$

$$\theta_1^{(z_0+1)} = \theta_1^{(z_0)} - \frac{1}{2} \cdot \xi^{(z_0)} \cdot \lambda_2^{(z_0)}$$
$$\theta_2^{(z_0+1)} = \theta_2^{(z_0)} + \frac{1}{2} \cdot \xi^{(z_0)} \cdot \lambda_2^{(z_0)}$$

Since $\sum_{z=z_0}^{+\infty} \xi^{(z)} = +\infty$ and $\lambda_2^{(z)} \geq \delta_2 > 0$ holds for all z in mode 1.1, it is straightforwardly derived that at a certain step $z_0 + K$ with $K \geq 1$, either $\theta_1^{(z_0+K)} > g_1(x_1^{\text{best}})$ or $\theta_2^{(z_0+K)} < g_2(x_2^{\text{best}})$ does not hold. This contradicts the condition of mode 1.1. Therefore, mode 1.1 is not persistent and the system will definitely switch to other modes. □

Proposition 2.1.3: Let $\delta_1 = \min\limits_{u_1 \in D_1,\ u_1 < u_1^{\text{best}}} -\frac{f_1'(u_1)}{g_1'(u_1)}$, then $\lambda_1^{(z)} \geq \delta_1 > 0$ holds for all z in mode 1.2.
Proof:
The proof of this proposition is similar to the one of Proposition 2.1.1.

Proposition 2.1.4: Mode 1.2 is not persistent.
Proof:
The proof of this proposition is similar to the one of Proposition 2.1.2.

Proposition 2.1.5: Mode 1.3 is persistent.
Proof:
Suppose the system is in mode 1.3 at step z_0, then we show that the system stay in mode 1.3 for all $z > z_0$.

If the system is in mode 1.3 at step z_0, then $\theta_1^{(z_0)} \geq g_1(u_1^{\text{best}})$ and $\theta_2^{(z_0)} \geq g_2(u_2^{\text{best}})$ at step z_0. In this mode,

$u_1^{*,(z_0)} = u_1^{\text{best}}$, $\lambda_1^{(z_0)} = 0$ and $u_2^{*,(z_0)} = u_2^{\text{best}}$, $\lambda_2^{(z_0)} = 0$. According to (3), at step z_0+1

$$\theta_1^{(z_0+1)} = \theta_1^{(z_0)}$$
$$\theta_2^{(z_0+1)} = \theta_2^{(z_0)}$$

It is clear that $\theta_1^{(z)} = \theta_1^{(z_0)}$ and $\theta_2^{(z)} = \theta_2^{(z_0)}$ holds for all $z > z_0$. Therefore, mode 1.3 is persistent. Besides, the overall optimal solution $[u_1^{\text{best}} \;\; u_2^{\text{best}}]^{\mathrm{T}}$ is directly attained in this mode.

Proposition 2.1.6: There exists an $M > 0$ such that at $z \geq M$, mode 1.3 is active i.e., $\theta_1^{(z)} \geq g_1(u_1^{\text{best}})$ and $\theta_2^{(z)} \geq g_2(u_2^{\text{best}})$.
Proof:
Let us assume mode 1.3 will never be reached and show that this leads to a contradiction.

Define $\lambda_2^{\max} = \max\limits_{x_2 \in D_2,\ u_2 < u_2^{\text{best}}} -\frac{f_2'(u_2)}{g_2'(u_2)}$. Since $\lim_{z\to+\infty} \xi^{(z)} = 0$, given $\varepsilon = \frac{2\left(r - \Sigma_{i=1}^{2} g_i(u_i^{\text{best}})\right)}{\lambda_2^{\max}}$, there exists an $M > 0$ such that $\xi^{(z)} < \varepsilon$ hold for all $z \geq M$.

If mode 1.3 will never be reached, since mode 1.1 and mode 1.2 have been proved to be not persistent, there are always mode switches either from mode 1.1 to mode 1.2 or from mode 1.2 to mode 1.1. Assume there is a switch from mode 1.1 to mode 1.2 at some step z_1 with $z_1 > M$ (i.e. at step z_1 mode 1.1 is active and at step z_1+1 mode 1.2 is active). Then according to the conditions of mode 1.1 and mode 1.2, we have

$$\theta_1^{(z_1)} > g_1(u_1^{\text{best}}),\;\; \theta_2^{(z_1)} < g_2(u_2^{\text{best}})$$
$$\theta_1^{(z_1+1)} < g_1(u_1^{\text{best}}),\;\; \theta_2^{(z_1+1)} > g_2(u_2^{\text{best}})$$

Hence we have

$$\theta_1^{(z_1+1)} - \theta_1^{(z_1)} < \theta_1^{(z_1+1)} - g_1(u_1^{\text{best}}) < 0$$
$$0 < \theta_2^{(z_1+1)} - g_2(u_2^{\text{best}}) < \theta_2^{(z_1+1)} - \theta_2^{(z)}$$

Since mode 1 is active at step z_1, we have $\theta_1^{(z_1+1)} - \theta_1^{(z_1)} = -\frac{1}{2}\cdot\xi^{(z_1)}\cdot\lambda_2^{(z_1)}$ and $\theta_2^{(z_1+1)} - \theta_2^{(z_1)} = \frac{1}{2}\cdot\xi^{(z_1)}\cdot\lambda_2^{(z_1)}$ (see proof of Proposition 2.1.1). As a consequence we have

$$-\frac{1}{2}\cdot\xi^{(z_1)}\cdot\lambda_2^{(z_1)} < \theta_1^{(z_1+1)} - g_1(u_1^{\text{best}}) < 0$$
$$0 < \theta_2^{(z_1+1)} - g_2(u_2^{\text{best}}) < \frac{1}{2}\cdot\xi^{(z_1)}\cdot\lambda_2^{(z_1)}$$

So

$$-\frac{1}{2}\cdot\xi^{(z_1)}\cdot\lambda_2^{(z_1)} < \theta_1^{(z_1+1)} + \theta_2^{(z_1+1)} - g_1(u_1^{\text{best}}) - g_2(u_2^{\text{best}}) < \frac{1}{2}\cdot\xi^{(z_1)}\cdot\lambda_2^{(z_1)}$$

Since $\theta_2^{(z_1)} < g_2(u_2^{\text{best}})$, $u_2^{*,(z_1)} < u_2^{\text{best}}$. Then $\lambda_2^{(z_1)} = -\frac{f_2'(u_2^{*,(z_1)})}{g_2'(u_2^{*,(z_1)})} < \lambda^{\max}$. Therefore, we have

$$\theta_1^{(z_1+1)} + \theta_2^{(z_1+1)} - g_1(u_1^{\text{best}}) - g_2(u_2^{\text{best}}) < \frac{1}{2}\cdot\xi^{(z_1)}\cdot\lambda_2^{\max}$$

If the switch happens at $z_1 > M$, we have

$$\theta_1^{(z_1+1)} + \theta_2^{(z_1+1)} - g_1(u_1^{\text{best}}) - g_2(u_2^{\text{best}}) < \frac{1}{2} \cdot \varepsilon \cdot \lambda_2^{\max}$$

so

$$\theta_1^{(z_1+1)} + \theta_2^{(z_1+1)} - g_1(u_1^{\text{best}}) - g_2(u_2^{\text{best}}) < r - \sum_{i=1}^{2} g_i(u_i^{\text{best}})$$

Since $\theta_1^{(z_1+1)} + \theta_2^{(z_1+1)} = r$, we have

$$g_1(u_1^{\text{best}}) + g_2(u_2^{\text{best}}) > \sum_{i=1}^{2} g_i(u_i^{\text{best}})$$

Clearly, it is a contradiction. Therefore, the assumption that mode 1.3 is never reached does not hold. □

Proposition 2.1.7: There exists an $M > 0$ such that the global optimum $[u_1^{\text{best}}\ u_2^{\text{best}}]^{\mathrm{T}}$ is attained at $z = M$.
Proof:
According to Proposition 2.1.6, there exists an $M > 0$ such that at any $z \geq M$, $\theta_1^{(z)} \geq g_1(u_1^{\text{best}})$ and $\theta_2^{(z)} \geq g_2(u_2^{\text{best}})$. Therefore, we have $\theta_1^{(M)} \geq g_1^{(M)}(u_1^{\text{best}})$, $\theta_2^{(M)} \geq g_2(u_2^{\text{best}})$ and $u_1^{*,(M)} = u_1^{\text{best}}$, $u_2^{*,(M)} = u_2^{\text{best}}$. Since $f_1(u_1^{\text{best}}) \leq f_1(u_1)$ holds for all $u_1 \in D_1$ and $f_2(u_2^{\text{best}}) \leq f_2(u_2)$ holds for all $u_2 \in D_2$, it is directly derived that $[u_1^{\text{best}}\ u_2^{\text{best}}]^{\mathrm{T}}$ is the global optimum. Finally, since $u_1^{*,(z)} = u_1^{\text{best}}$, $u_2^{*,(z)} = u_2^{\text{best}}$ holds for all $z \geq M$, the global optimum is attained at $z = M$. □

Graph-aided explanation
Given finite discrete set D_i, for any $u_i \in D_i$, if $u_i \geq u_i^{\text{best}}$, we have the corresponding $\lambda_i = 0$. Actually, by definition, we have

$$\lambda_i = \begin{cases} -\frac{f_i'(u_i)}{g_i'(u_i)}, & \text{if } u_i \in D_i, u_i < u_i^{\text{best}} \\ 0, & \text{if } u_i \in D_i, u_i \geq u_i^{\text{best}} \end{cases}$$

Without loss of generality, let

$$D_1 = \{u_{1,n_1},\ u_{1,n_1-1},\ \ldots,\ u_{1,1},\ u_1^{\text{best}},\ u_{1,n_1+2},\ \ldots,\ u_{1,n_1+m_1},\ u_{1,n_1+m_1+1}\}, \qquad n_1 \geq 0,\ m_1 \geq 0$$
$$D_2 = \{u_{2,n_2},\ u_{2,n_2-1},\ \ldots,\ u_{2,1},\ u_2^{\text{best}},\ u_{2,n_2+2},\ \ldots,\ u_{2,n_2+m_2},\ u_{2,n_2+m_2+1}\}, \qquad n_2 \geq 0,\ m_2 \geq 0$$

with

$$u_{1,n_1} < u_{1,n_1-1} < u_{1,1} < x_1^{\text{best}} < u_{1,n_1+2} < u_{1,n_1+m_1} < u_{1,n_1+m_1+1}$$
$$u_{2,n_2} < u_{2,n_2-1} < u_{2,1} < u_2^{\text{best}} < u_{2,n_2+2} < u_{2,n_2+m_2} < u_{2,n_2+m_2+1}$$

Depending on different $f_i(\cdot)$, $g_i(\cdot)$, D_i, values of λ_i can be different. Without loss of generality, the values of λ_1 and λ_2 along with θ_1 and θ_2 are shown in Figure 2.

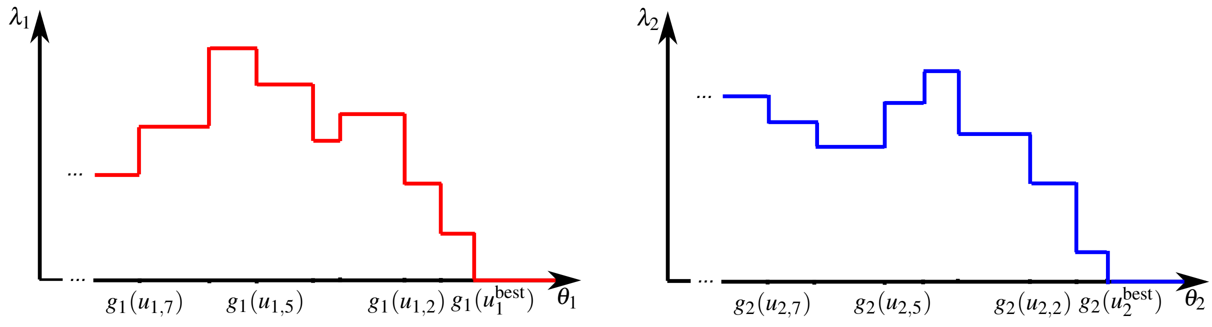

Figure 2: Values of λ_1 and λ_2 along the axis of θ_1 and θ_2

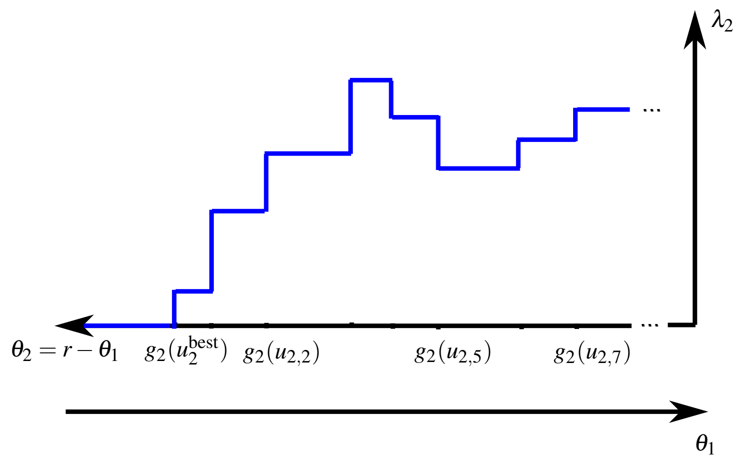

Figure 3: Values of λ_2 along the axis of θ_1

In the resource allocation coordination method, at any step z, no matter what is the value of $\theta_1^{(z)}$, we have

$$\theta_2^{(z)} = r - \theta_1^{(z)}$$

Therefore, given r, the values of λ_2 along with θ_2 can be expressed as that shown in Figure 3.

Further, in this case $g_1(u_1^{\text{best}}) + g_2(u_2^{\text{best}}) < r$, the values of λ_1 and λ_2 along with θ_1 is shown in Figure 4.

Since in the resource allocation coordination method, the update of $\theta_1^{(z+1)}$ and $\theta_2^{(z+1)}$ is done by

$$\theta_1^{(z+1)} = \theta_1^{(z)} + \varepsilon^{(z)} \frac{\lambda_1^{(z)} - \lambda_2^{(z)}}{2}$$
$$\theta_2^{(z+1)} = \theta_2^{(z)} + \varepsilon^{(z)} \frac{\lambda_2^{(z)} - \lambda_1^{(z)}}{2}$$

with diminishing step size $\varepsilon^{(z)}$, it can be easily derived from Figure 4 that no matter what are the values of $\theta_1^{(1)}$ and $\theta_2^{(1)}$, as the iteration step z increases, $\theta_1^{(z)}$ and $\theta_2^{(z)}$ will reach a point within the deepened segment with a finite z and stay at that point afterwards. In the deepened segment, we have $\theta_1^{(z)} \geq g_1^{(z)}(u_1^{\text{best}})$, $\theta_2^{(z)} \geq g_2(u_2^{\text{best}})$ and $u_1^{*,(z)} = u_1^{\text{best}}$, $u_2^{*,(z)} = u_2^{\text{best}}$ with $[u_1^{\text{best}}\ u_2^{\text{best}}]^{\text{T}}$ being the global optimum.

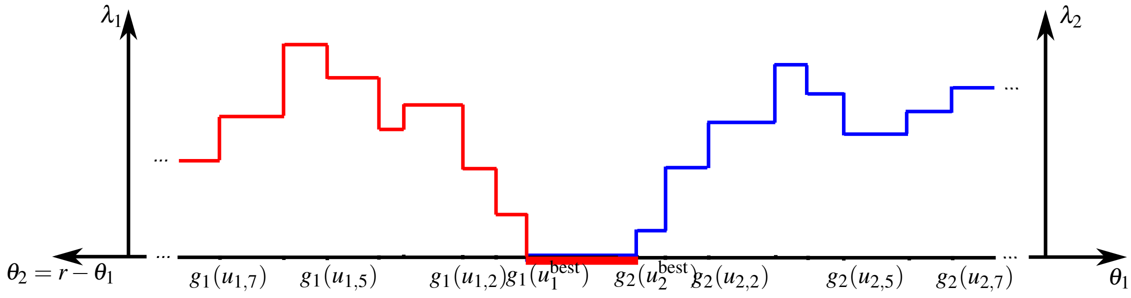

Figure 4: Values of λ_1 and λ_2 along the axis of θ_1

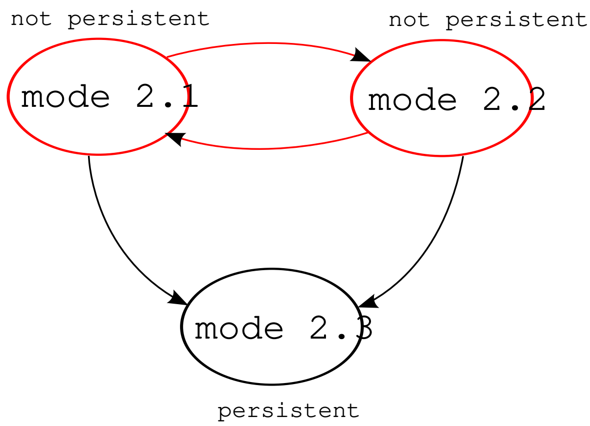

Figure 5: Mode transition diagram of case 3 with the final state marked in red

2.2 For case 2

In this case, $\sum_{i=1}^{2} g_i(u_i^{\text{best}}) = r$. Since $\sum_{i=1}^{2} \theta_i = r$, we consider the following three modes:

- mode 2.1: $\theta_1^{(z)} > g_1(u_1^{\text{best}})$ and $\theta_2^{(z)} < g_2(u_2^{\text{best}})$
- mode 2.2: $\theta_1^{(z)} < g_1(u_1^{\text{best}})$ and $\theta_2^{(z)} > g_2(u_2^{\text{best}})$
- mode 2.3: $\theta_1^{(z)} = g_1(u_1^{\text{best}})$ and $\theta_2^{(z)} = g_2(u_2^{\text{best}})$

The mode transition diagram of case 2 is shown in Figure 5.

Proposition 2.2.1: Mode 2.1 is not persistent.
Proof:
The proof of this proposition is similar to the one of Proposition 2.1.2.

Proposition 2.2.2: Mode 2.2 is not persistent.
Proof:
The proof of this proposition is similar to the one of Proposition 2.1.4.

Proposition 2.2.3: Mode 2.3 is persistent.
Proof:
The proof of this proposition is similar to the one of Proposition 2.1.5. The overall optimal solution $[u_1^{\text{best}} \;\; u_2^{\text{best}}]^{\mathrm{T}}$ is attained in this mode.

Lemma 2.2.4: In mode 2.1, $|\theta_1^{(z)} - g_1(u_1^{\text{best}})|$ and $|\theta_2^{(z)} - g_2(u_2^{\text{best}})|$ are strictly decreasing as functions of z.
Proof:
If mode 2.1 is active at step z_0, we have $\theta_1^{(z_0)} > g_1(u_1^{\text{best}})$ and $\theta_2^{(z_0)} < g_2(u_2^{\text{best}})$ and

$$\begin{aligned}\theta_1^{(z_0+1)} &= \theta_1^{(z_0)} - \frac{1}{2} \cdot \xi^{(z_0)} \cdot \lambda_2^{(z_0)} \\ \theta_2^{(z_0+1)} &= \theta_2^{(z_0)} + \frac{1}{2} \cdot \xi^{(z_0)} \cdot \lambda_2^{(z_0)}\end{aligned}$$

If mode 2.1 is still active at step z_0+1, we have $\theta_1^{(z_0+1)} > g_1(u_1^{\text{best}})$ and $\theta_2^{(z_0+1)} < g_2(u_2^{\text{best}})$. Then we have

$$\begin{aligned}&|\theta_1^{(z_0+1)} - g_1(u_1^{\text{best}})| - |\theta_1^{(z_0)} - g_1(u_1^{\text{best}})| \\ =&\theta_1^{(z_0+1)} - g_1(u_1^{\text{best}}) - \left(\theta_1^{(z_0)} - g_1(u_1^{\text{best}})\right) \\ =&\theta_1^{(z_0+1)} - \theta_1^{(z_0)} \\ =&-\frac{1}{2} \cdot \xi^{(z_0)} \cdot \lambda_2^{(z_0)} \\ <&-\frac{1}{2} \cdot \xi^{(z_0)} \cdot \delta_2 < 0\end{aligned}$$

and

$$\begin{aligned}&|\theta_2^{(z_0+1)} - g_2(u_2^{\text{best}})| - |\theta_2^{(z_0)} - g_2(u_2^{\text{best}})| \\ =&g_2(u_2^{\text{best}}) - \theta_2^{(z_0+1)} - \left(g_2(u_2^{\text{best}}) - \theta_2^{(z_0)}\right) \\ =&\theta_2^{(z_0)} - \theta_2^{(z_0+1)} \\ =&-\frac{1}{2} \cdot \xi^{(z_0)} \cdot \lambda_2^{(z_0)} \\ <&-\frac{1}{2} \cdot \xi^{(z_0)} \cdot \delta_2 < 0\end{aligned}$$

where δ_2 is the same as before. $\square$

Lemma 2.2.5: In mode 2.2, $|\theta_1^{(z)} - g_1(u_1^{\text{best}})|$ and $|\theta_2^{(z)} - g_2(u_2^{\text{best}})|$ are strictly decreasing as functions of z.
Proof: the proof of this lemma is similar to the one of Lemma 2.3.4.

Proposition 2.2.6: Given $\sigma^{\max} = \max\{\sigma_1, \sigma_2\}$, $\theta_1^{(1)}$ and $\theta_2^{(1)}$, a large integer K and a small real number $\varepsilon = \frac{\sigma^{\max}}{2}\xi^{(K)}$, there exists an $M > K$ such that $|\theta_1^{(z)} - g_1(u_1^{\text{best}})| < \varepsilon$ and $|\theta_2^{(z)} - g_2(u_2^{\text{best}})| < \varepsilon$ hold for all $z > M$.
Proof:

If there is a switch from mode 2.1 to mode 2.2 at step z_0, we have $\theta_1^{(z_0)} > g_1(u_1^{\text{best}})$ and $\theta_2^{(z_0)} < g_2(u_2^{\text{best}})$ and $\theta_1^{(z_0+1)} < g_1(u_1^{\text{best}})$ and $\theta_2^{(z_0+1)} > g_2(u_2^{\text{best}})$, and also

$$\begin{aligned}
&|\theta_1^{(z_0)} - g_1(u_1^{\text{best}})| \\
=&\theta_1^{(z_0)} - g_1(u_1^{\text{best}}) \\
<&\theta_1^{(z_0)} - \theta_1^{(z_0+1)} \\
<&\frac{1}{2} \cdot \xi^{(z_0)} \lambda_2^{(z_0)} \\
<&\frac{1}{2} \cdot \xi^{(z_0)} \sigma^{\max}
\end{aligned}$$

and

$$\begin{aligned}
&|\theta_1^{(z_0+1)} - g_1(u_1^{\text{best}})| \\
=&g_1(u_1^{\text{best}}) - \theta_1^{(z_0+1)} \\
<&\theta_1^{(z_0)} - \theta_1^{(z_0+1)} \\
<&\frac{1}{2} \cdot \xi^{(z_0)} \lambda_2^{(z_0)} \\
<&\frac{1}{2} \cdot \xi^{(z_0)} \sigma^{\max}
\end{aligned}$$

and

$$\begin{aligned}
&|\theta_2^{(z_0)} - g_2(u_2^{\text{best}})| \\
=&g_2(u_2^{\text{best}}) - \theta_2^{(z_0)} \\
<&\theta_2^{(z_0+1)} - \theta_2^{(z_0)} \\
<&\frac{1}{2} \cdot \xi^{(z_0)} \lambda_2^{(z_0)} \\
<&\frac{1}{2} \cdot \xi^{(z_0)} \sigma^{\max}
\end{aligned}$$

and

$$\begin{aligned}
&|\theta_2^{(z_0+1)} - g_2(u_2^{\text{best}})| \\
=&\theta_2^{(z_0+1)} - g_2(u_2^{\text{best}}) \\
<&\theta_2^{(z_0+1)} - \theta_2^{(z_0)} \\
<&\frac{1}{2} \cdot \xi^{(z_0)} \lambda_2^{(z_0)} \\
<&\frac{1}{2} \cdot \xi^{(z_0)} \sigma^{\max}
\end{aligned}$$

Likewise, if there is a switch from mode 2.2 to mode 2.1 at step z_0, we also have

$$
\begin{aligned}
|\theta_1^{(z_0)} - g_1(u_1^{\text{best}})| &< \frac{1}{2} \cdot \xi^{(z_0)} \sigma^{\max} \\
|\theta_1^{(z_0+1)} - g_1(u_1^{\text{best}})| &< \frac{1}{2} \cdot \xi^{(z_0)} \sigma^{\max} \\
|\theta_2^{(z_0)} - g_2(u_2^{\text{best}})| &< \frac{1}{2} \cdot \xi^{(z_0)} \sigma^{\max} \\
|\theta_2^{(z_0+1)} - g_2(u_2^{\text{best}})| &< \frac{1}{2} \cdot \xi^{(z_0)} \sigma^{\max}
\end{aligned}
$$

Next, if mode 2.3 is not reached for any $z < \infty$, then there are repeated mode transitions between mode 2.1 and 2.2 since none of mode 2.1 and 2.2 is persistent. Therefore, no matter what is the value of K, there exists an $M > K$ such that a mode (no matter it is from mode 2.1 to mode 2.2 or from mode 2.2 to mode 2.1) switch occurs at step M. Hence, we have

$$
\begin{aligned}
|\theta_1^{(M)} - g_1(u_1^{\text{best}})| &< \frac{1}{2} \cdot \xi^{(M)} \sigma^{\max} < \frac{1}{2} \cdot \xi^{(K)} \sigma^{\max} = \varepsilon \\
|\theta_1^{(M+1)} - g_1(u_1^{\text{best}})| &< \frac{1}{2} \cdot \xi^{(M)} \sigma^{\max} < \frac{1}{2} \cdot \xi^{(K)} \sigma^{\max} = \varepsilon \\
|\theta_2^{(M)} - g_2(u_2^{\text{best}})| &< \frac{1}{2} \cdot \xi^{(M)} \sigma^{\max} < \frac{1}{2} \cdot \xi^{(K)} \sigma^{\max} = \varepsilon \\
|\theta_2^{(M+1)} - g_2(u_2^{\text{best}})| &< \frac{1}{2} \cdot \xi^{(M)} \sigma^{\max} < \frac{1}{2} \cdot \xi^{(K)} \sigma^{\max} = \varepsilon
\end{aligned}
$$

Since we have also proved in Proposition 2.2.4 and 2.2.5 that $|\theta_1^{(z)} - g_1(u_1^{\text{best}})|$ and $|\theta_2^{(z)} - g_2(u_2^{\text{best}})|$ are strictly decreasing in mode 2.1 and 2.2, we can conclude that at any step $z > M$ no matter whether the system in a mode or switch from a mode to another mode, $|\theta_1^{(z)} - g_1(u_1^{\text{best}})| < \varepsilon$ and $|\theta_2^{(z)} - g_2(u_2^{\text{best}})| < \varepsilon$ holds.

Finally, if mode 2.3 is reached at $z_1 < \infty$, no matter what is the value of K, there exists an $M > K$ such that $|\theta_1^{(z)} - g_1(u_1^{\text{best}})| < \varepsilon$ and $|\theta_2^{(z)} - g_2(u_2^{\text{best}})| < \varepsilon$ hold for all $z > M$. More specifically, if $z_1 < K$, then for any $z > z_1$, we have $|\theta_1^{(z)} - g_1(u_1^{\text{best}})| = 0 < \varepsilon$ and $|\theta_2^{(z)} - g_2(u_2^{\text{best}})| = 0 < \varepsilon$. Then, by letting $M = K + 1 > z_1$, for any $z > M$, we have $|\theta_1^{(z)} - g_1(u_1^{\text{best}})| = 0 < \varepsilon$ and $|\theta_2^{(z)} - g_2(u_2^{\text{best}})| = 0 < \varepsilon$. If $z_1 > K$, by letting $M = z_1$, for all $z > M$, we have $|\theta_1^{(z)} - g_1(u_1^{\text{best}})| = 0 < \varepsilon$ and $|\theta_2^{(z)} - g_2(u_2^{\text{best}})| = 0 < \varepsilon$.
□

Proposition 2.2.7: $\lim_{z\to\infty} \theta_1^{(z)} = g_1(u_1^{\text{best}})$ and $\lim_{z\to\infty} \theta_2^{(z)} = g_1(u_2^{\text{best}})$.
Proof:
The proof of this proposition can be directly derived from Proposition 2.2.6 with $K = \infty$.

Graph-aided explanation
In this case $g_1(u_1^{\text{best}}) + g_2(u_2^{\text{best}}) = r$, the values of λ_1 and λ_2 along with θ_1 is shown in Figure 6. It can be easily derived from Figure 6 that no matter what are the values of $\theta_1^{(1)}$ and $\theta_2^{(1)}$, as the iteration step z goes to infinity, $\theta_1^{(z)}$ and $\theta_2^{(z)}$ reach the point where $\theta_1^{(z)} = g_1^{(z)}(u_1^{\text{best}})$, $\theta_2^{(z)} = g_2(u_2^{\text{best}})$. The the global optimum $[u_1^{\text{best}}\ u_2^{\text{best}}]^{\mathrm{T}}$ is attained when z goes to infinity.

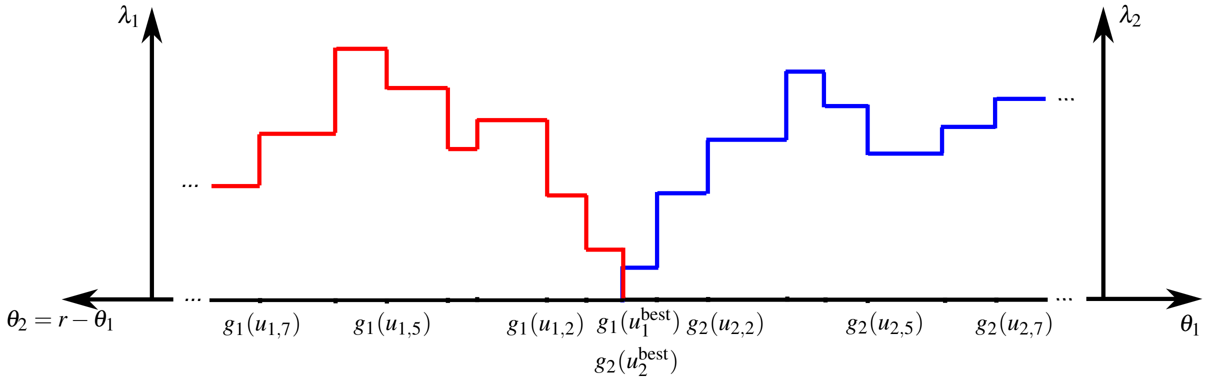

Figure 6: Values of λ_1 and λ_2 along the axis of θ_1

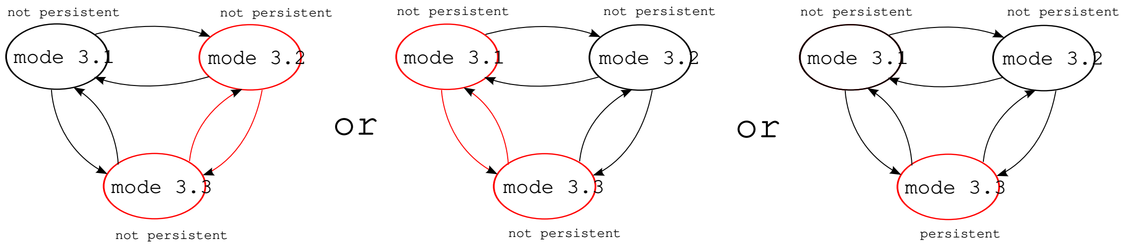

Figure 7: Mode transition diagram of case 2 with the final state marked in red

2.3 For case 3

In this case, $\sum_{i=1}^{2} g_i(u_i^{best}) > r$. Since $\sum_{i=1}^{2} \theta_i = r$, we consider the following three modes:

- mode 3.1: $\theta_1^{(z)} \geq g_1(u_1^{best})$ and $\theta_2^{(z)} < g_2(u_2^{best})$
- mode 3.2: $\theta_1^{(z)} < g_1(u_1^{best})$ and $\theta_2^{(z)} \geq g_2(u_2^{best})$
- mode 3.3: $\theta_1^{(z)} < g_1(u_1^{best})$ and $\theta_2^{(z)} < g_2(u_2^{best})$

The mode transition diagram of case 3 is shown in Figure 7.

Proposition 2.3.1: Mode 3.1 is not persistent.
Proof: the proof of this proposition is similar to the one of Proposition 2.1.2.

Proposition 2.3.2: Mode 3.2 is not persistent.
Proof: the proof of this proposition is similar to the one of Proposition 2.1.4.

Proposition 2.3.3: Given $\frac{f_1'(u_1)}{g_1'(u_1)} \neq \frac{f_2'(u_2)}{g_2'(u_2)}$ holds for all $u_1 \in D_1$ and $u_2 \in D_2$ with $u_1 < u_1^{best}$, $u_2 < u_2^{best}$ and $g_1(u_1) + g_2(u_2) \leq r$, then either local solution u_1^* or u_2^* will not stay constant.

Proof:
Since mode 3.1 is not persistent, given mode 3.1 is active at some step $z_1 > 0$ with $u_1^{*,(z_1)} = u_1^{\text{best}}$, there will be a switch from 3.1 to either mode 3.2 or mode 3.3. Assume the switch happens at step z_2 with $z_2 > z_1$, we have $u_1^{*,(z_2)} < u_1^{\text{best}}$. Therefore, $u_1^{*,(z_2)} \neq u_1^{*,(z_1)}$.

Since mode 3.2 is not persistent, given mode 3.2 is active at some step $z_3 > 0$ with $u_2^{*,(z_3)} = u_2^{\text{best}}$, there will be a switch from 3.2 to either mode 3.1 or mode 3.3. Assume the switch happens at step z_4 with $z_4 > z_3$, we have $u_2^{*,(z_4)} < u_2^{\text{best}}$. Therefore, $u_2^{*,(z_4)} \neq u_2^{*,(z_3)}$.

If mode 3.3 is not persistent, given mode 3.3 is active at some step z_5 with $u_1^{*,(z_5)} < u_1^{\text{best}}$ and $u_2^{*,(z_5)} < u_2^{\text{best}}$, there will be a switch from 3.3 to either mode 3.1 or mode 3.2. If the switch is from mode 2.3 to mode 2.1 and happens at step z_6 with $z_6 > z_5$, we have $u_1^{*,(z_6)} = u_1^{\text{best}}$ and then $u_1^{*,(z_6)} \neq u_1^{*,(z_5)}$. If the switch is from mode 3.3 to mode 3.2 and happens at z_7 with $z_7 > z_5$, we have $u_2^{*,(z_7)} = u_2^{\text{best}}$ and then $u_2^{*,(z_7)} \neq u_2^{*,(z_5)}$.
If mode 3.3 is persistent, given mode 3.3 is active at step some step z_8, we have $u_1^{*,(z_8)} < u_1^{\text{best}}$, $\lambda_1^{(z_8)} > 0$ and $u_2^{*,(z_8)} < u_2^{\text{best}}$, $\lambda_2^{(z_8)} > 0$. According to (3), in this mode, at step $z_8 + 1$

$$\theta_1^{(z_8+1)} = \theta_1^{(z_8)} - \frac{\lambda_2^{(z_8)} - \lambda_1^{(z_8)}}{2} \cdot \xi^{(z_8)}$$
$$\theta_2^{(z_8+1)} = \theta_2^{(z_8)} + \frac{\lambda_2^{(z_8)} - \lambda_1^{(z_8)}}{2} \cdot \xi^{(z_8)}$$

Since $\lambda_1^{(z_8)} = -\frac{f_1'(u_1^{*,(z_8)})}{g_1'(u_1^{*,(z_8)})}$, $\lambda_2^{(z_8)} = -\frac{f_2'(u_2^{*,(z_8)})}{g_2'(u_2^{*,(z_8)})}$ and $\frac{f_1'(u_1^{*,(z_8)})}{g_1'(u_1^{*,(z_8)})} \neq \frac{f_2'(u_2^{*,(z_8)})}{g_2'(u_2^{*,(z_8)})}$, we have $\lambda_1^{(z_8)} \neq \lambda_2^{(z_8)}$. Therefore, also because $\sum_{z=z_8}^{+\infty} \xi^{(z)} = +\infty$ and $\lambda_1^{(z_8+j)} \neq \lambda_2^{(z_8+j)}$ with $j \geq 0$, $\theta_1^{(z_8+j)}$ keeps increasing (or decreasing) and $\theta_2^{(z_8+j)}$ keeps decreasing (or increasing) until at step z_9 with $z_9 > z_8$ either $u_1^{*,(z_9)} \neq u_1^{*,(z_8)}$ or $u_2^{*,(z_9)} \neq u_2^{*,(z_8)}$. $\square$

Proposition 2.3.4: Depending on different $f_i(\cdot)$, $g_i(\cdot)$, D_i and $\theta_i^{(1)}$ with $i = 1,2$, the mode transition diagram of case 2 can be any of the three kinds shown in Figure 7.
Proof:
The proof will be given in the graph-aided explanation.

Graph-aided explanation
In the case $g_1(u_1^{\text{best}}) + g_2(u_2^{\text{best}}) > r$, depending on different $f_i(\cdot)$, $g_i(\cdot)$, D_i, the graphs of showing the values of λ_1 and λ_2 along the axis of θ_1 can be different. Without loss of generality, we present three subcases 3.1, 3.2, and 3.3 in Figure 8, 9 and 10 respectively.

In subcase 3.1 as shown in Figure 8, no matter what are the values of $\theta_1^{(1)}$ and $\theta_2^{(1)}$, as the iteration step z increases, $\theta_1^{(z)}$ gets closer to $g_1(u_1^{\text{best}})$ and $\theta_2^{(z)}$ gets closer to $r - g_1(u_1^{\text{best}})$. However, if $\theta_1^{(z)} < g_1(u_1^{\text{best}})$, then $u_1^{*,(z)} < u_1^{\text{best}}$ and $\lambda_1^{(z)} > \lambda_2^{(z)}$, and hence according to the update equation of θ_1, we have $\theta_1^{(z+1)} > \theta_1^{(z)}$; if $\theta_1^{(z)} \geq g_1(u_1^{\text{best}})$, then $u_1^{*,(z)} = u_1^{\text{best}}$ and $\lambda_1^{(z)} < \lambda_2^{(z)}$, and hence $\theta_1^{(z+1)} < \theta_1^{(z)}$. In this subcase, with z goes to infinity, we have u_1^* oscillating between $u_{1,1}$ and u_1^{best} as indicated a small circle in Figure 8. Besides,

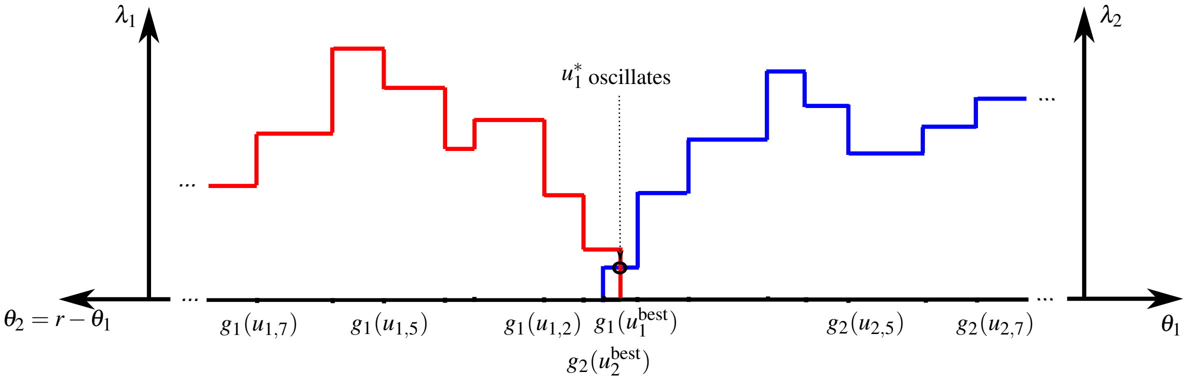

Figure 8: Values of λ_1 and λ_2 along the axis of θ_1 in subcase 3.1

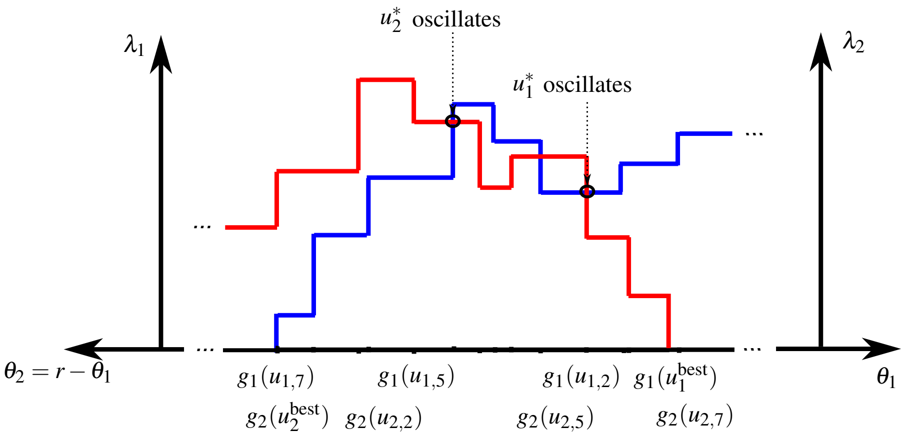

Figure 9: Values of λ_1 and λ_2 along the axis of θ_1 in subcase 3.2

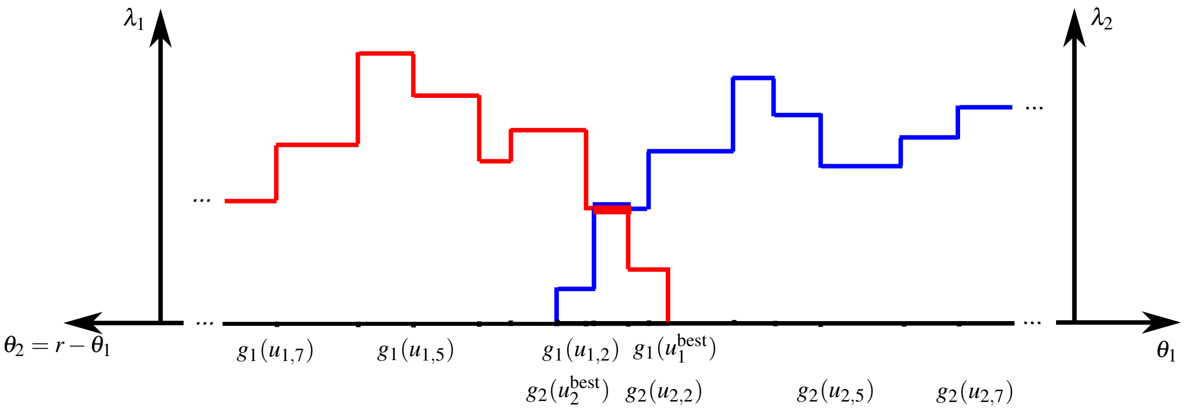

Figure 10: Values of λ_1 and λ_2 along the axis of θ_1 in subcase 3.3

it is directly derived from Figure 8 that the oscillation of u_1^* is characterized by

$$u_1^{*,(z+1)} \neq u_1^{*,(z)}$$
$$\operatorname{sgn}\left(\Delta\theta_1^{(z+1)}\right) \neq \operatorname{sgn}\left(\Delta\theta_1^{(z)}\right)$$

with $\Delta\theta_1^{(z+1)} = \theta_1^{(z+1)} - \theta_1^{(z)}$ and $\Delta\theta_1^{(z)} = \theta_1^{(z)} - \theta_1^{(z-1)}$.

Note that Figure 8 directly corresponds to the second kind of mode transition diagram in Figure 7. Besides, in Figure 8, if we switch the labels for u_1 and u_2 by letting u_1 be u_2 and letting u_2 be u_1, we have Figure 8 corresponds to the first kind of mode transition diagram in Figure 7.

In subcase 3.2, as shown in Figure 9, depending on the initial values $\theta_1^{(1)}$ and $\theta_2^{(1)}$ and the step size, as the iteration step z increases, either u_1^* or u_2^* oscillates. Both the oscillations of u_1^* and u_2^* are characterized by

$$u_i^{*,(z+1)} \neq u_i^{*,(z)}$$
$$\operatorname{sgn}\left(\Delta\theta_i^{(z+1)}\right) \neq \operatorname{sgn}\left(\Delta\theta_i^{(z)}\right)$$

Note that Figure 9 corresponds to the third kind of mode transition diagram in Figure 7.

In subcase 3.3, as shown in Figure 10, no matter what are the values of $\theta_1^{(1)}$ and $\theta_2^{(1)}$, as the iteration step z increases, $\theta_1^{(z)}$ and $\theta_2^{(z)}$ will reach a point within the deepened segment with a finite z and stay at that point afterwards. In the deepened segment, $\lambda_1^{(z)} = \lambda_2^{(z)} \neq 0$ and hence neither of u_1^* and u_2^* oscillate. However, in this subcase, even though neither u_1^* or u_2^* oscillate, there is no guarantee that the global optimum is attained. That is to say, when the resource allocation coordination method is applied to problem (1), even though no oscillation of discrete variables is detected, it is possible the global optimum is not yet attained.

3 Answer to question 2

Proposition 3.1.1: The oscillation of discrete decision variable u_i is characterized by $u_i^{*,(z+1)} \neq u_i^{*,(z)}$ and $\operatorname{sgn}(\Delta\theta_i^{(z+1)}) \neq \operatorname{sgn}(\Delta\theta_i^{(z)})$.

Proof:
The proof of this proposition has been given in the graph-aid explanation in Section 2.3, especially in the discussion of subcase 3.1 and subcase 3.2.

4 General problem

In the previous sections, we have proved some properties of applying the resource allocation coordination to a simple example with $N = 2$. To make it general, in this case, we prove the same properties of the resource allocation coordination method to examples with any $N > 2$ still hold.

Let us define $I_1^{(z)} = \{i|\theta_i \geq g_i(u_i^{\text{best}})\}$ and $I_2^{(z)} = \{j|\theta_j < g_j(u_j^{\text{best}})\}$. If we define $I = \{1, 2, ..., N\}$, it is obvious that $I = I_1^{(z)} \cup I_2^{(z)}$ holds for all z.

4.1 For case 1

Proposition 4.1.1: In the case $\sum_{i=1}^N g_i(u_i^{\text{best}}) < r$, there exists an $M \geq 0$ such that at all step $z \geq M$, $I_2^{(z)} = \emptyset$.

Proof:
In this case, with $\sum_{i=1}^N g_i(u_i^{\text{best}}) < r$, we want to prove that no matter what the values of $\theta_i^{(1)}$ for $i \in I$ are, $\theta_i^{(z)}$ will eventually reach a steady state with $\theta_i^{(z)} \geq g_i(u_i^{\text{best}})$ for all i.

In order to prove I_2 will eventually be empty, I first assume that I_2 will never be empty and then find a contradiction.

At any step $z > 0$, we have $\lambda_1^{(z)} = 0$ for all $i \in I_1^{(z)}$ and $\lambda_j^{(z)} > 0$ for all $j \in I_2^{(z)}$. Now let us define

$$\bar{\lambda}^{(z)} = \frac{1}{N}\Big(\sum_{j\in I_2^{(z)}} \lambda_j^{(z)} + \sum_{i\in I_1^{(z)}} \lambda_i^{(z)}\Big) = \frac{1}{N}\sum_{j\in I_2^{(z)}} \lambda_j^{(z)}$$

If $I_2^{(z)} \neq \emptyset$ at step $z > 0$, we have $\bar{\lambda}^{(z)} > 0$. As defined in Proposition 2.1.1, we define for all $i \in I$

$$\delta_i = \min_{u_i \in D_i,\ u_i < u_i^{\text{best}}} -\frac{f_i'(u_i)}{g_i'(u_i)}$$

Further, define

$$\delta^{\min} = \min_{i\in I} \delta_i$$

Since $f_i(\cdot)$ is convex and $f_i(u_i^{\text{best}}) < f_i(u_i)$ holds for $u_i \in D_i$ with $u_i < u_i^{\text{best}}$, we have $f_i'(u_i) < 0$ for $u_i \in D_i$ with $u_i < u_i^{\text{best}}$. In addition, given $g_i(\cdot)$ is monotonically strictly increasing, we have $g_i'(\cdot) > 0$. Therefore, we have $-\frac{f_i'(u_i)}{g_i'(u_i)} > 0$ holds for $u_i \in D_i$ with $x_i < u_i^{\text{best}}$ and it is directly derived that $\delta_i > 0$ holds for all $i \in I$ and $\delta^{\min} > 0$.

Therefore, if $I_2^{(z)} \neq \emptyset$ at step z, for every $j \in I_2^{(z)}$, we have $u_j^{*,(z)} < u_j^{\text{best}}$ and $\lambda_j^{(z)} = -\frac{f_j'(u_j^{*,(z)})}{g_j'(u_j^{*,(z)})} \geq \delta^{\min}$. Then we have

$$\bar{\lambda}^{(z)} \geq \frac{\delta^{\min}}{N} > 0$$

Now let us define

$$\sigma_i = \max_{x_i \in D_i,\ u_i < u_i^{\text{best}}} -\frac{f_i'(u_i)}{g_i'(u_i)},\ \ \forall\, i \in I$$

$$\sigma^{\max} = \max_{i \in I} \sigma_i$$

Like δ_i and $\delta_i^{\min}$, $\sigma_i > 0$ holds for all $i \in I$ and $\sigma^{\max} > 0$. Besides, $\sigma^{\max}$ is finite since I has finite elements.

Therefore, if $I_2^{(z)} \neq \emptyset$ at step z, we have for all $j \in I_2^{(z)}$

$$\lambda_j^{(z)} - \bar{\lambda}^{(z)} < \lambda_j^{(z)} \leq \sigma^{\max}$$

Now let us define a nonnegative function

$$J(z) = \sum_{i=1}^{N} \left(\theta_i^{(z)} - g_i(u_i^{\text{best}})\right)^2$$

then

$$\begin{aligned} & J(z+1) - J(z) \\ =& \sum_{i=1}^{N} \left(\theta_i^{(z+1)} - \theta_i^{(z)}\right)\left(\theta_i^{(z+1)} + \theta_i^{(z)} - 2g_i(u_i^{\text{best}})\right) \end{aligned}$$

At any step z, if $I_2^{(z)} \neq \emptyset$, we have

$$\begin{aligned} & J(z+1) - J(z) \\ =& \sum_{i \in I_1^{(z)}} -\xi^{(z)}\bar{\lambda}^{(z)}\left(\theta_i^{(z+1)} + \theta_i^{(z)} - 2g_i(u_i^{\text{best}})\right) + \sum_{j \in I_2^{(z)}} \xi^{(z)}\left(\lambda_j^{(z)} - \bar{\lambda}^{(z)}\right)\left(\theta_j^{(z+1)} + \theta_j^{(z)} - 2g_j(u_j^{\text{best}})\right) \\ =& \sum_{i=1}^{N} -\xi^{(z)}\bar{\lambda}^{(z)}\left(\theta_i^{(z+1)} + \theta_i^{(z)} - 2g_i(u_i^{\text{best}})\right) + \sum_{j \in I_2^{(z)}} \xi^{(z)}\lambda_j^{(z)}\left(\theta_j^{(z+1)} + \theta_j^{(z)} - 2g_j(u_j^{\text{best}})\right) \end{aligned}$$

Since $I_1^{(z)} \cup I_2^{(z)} = I = \{1, 2, ..., N\}$ holds for all z.

Further, since

$$\begin{aligned} & \sum_{i=1}^{N} -\xi^{(z)}\bar{\lambda}^{(z)}\left(\theta_i^{(z+1)} + \theta_i^{(z)} - 2g_i(u_i^{\text{best}})\right) \\ =& -\xi^{(z)}\bar{\lambda}^{(z)}\left(\sum_{i=1}^{N}\theta_i^{(z+1)} + \sum_{i=1}^{N}\theta_i^{(z)} - 2\sum_{i=1}^{N} g_i(u_i^{\text{best}})\right) \\ =& -2\xi^{(z)}\bar{\lambda}^{(z)}\left(r - \sum_{i=1}^{N} g_i(u_i^{\text{best}})\right) \\ \leq& -2\xi^{(z)}\frac{\delta^{\min}}{N}\left(r - \sum_{i=1}^{N} g_i(u_i^{\text{best}})\right) \end{aligned}$$

and

$$\begin{aligned}
&\sum_{j\in I_2^{(z)}} \xi^{(z)}\lambda_j^{(z)}\Big(\theta_j^{(z+1)}+\theta_j^{(z)}-2g_j(u_j^{\text{best}})\Big)\\
&=\sum_{j\in I_2^{(z)}} \xi^{(z)}\lambda_j^{(z)}\Big(\theta_j^{(z)}+\xi^{(z)}\Big(\lambda_j^{(z)}-\bar{\lambda}^{(z)}\Big)+\theta_j^{(z)}-2g_j(u_j^{\text{best}})\Big)\\
&=\sum_{j\in I_2^{(z)}} \big(\xi^{(z)}\big)^2\lambda_j^{(z)}\Big(\lambda_j^{(z)}-\bar{\lambda}^{(z)}\Big)+\sum_{j\in I_2^{(z)}} 2\xi^{(z)}\lambda_j^{(z)}\Big(\theta_j^{(z)}-g_j(u_j^{\text{best}})\Big)\\
&<\sum_{j\in I_2^{(z)}} \big(\xi^{(z)}\big)^2\lambda_j^{(z)}\Big(\lambda_j^{(z)}-\bar{\lambda}^{(z)}\Big)\\
&<\sum_{j\in I_2^{(z)}} \big(\xi^{(z)}\big)^2(\sigma^{\max})^2 < N\cdot\big(\xi^{(z)}\big)^2(\sigma^{\max})^2
\end{aligned}$$

we have

$$J(z+1)-J(z) < -2\xi^{(z)}\frac{\delta^{\min}}{N}\Big(r-\sum_{i=1}^{N} g_i(u_i^{\text{best}})\Big)+N\cdot\big(\xi^{(z)}\big)^2(\sigma^{\max})^2$$

First, let K be an arbitrary integer. Then, if $I_2^{(z)}\neq\emptyset$ at any of step $z\in\{1,2,...,K\}$, we have

$$J(K+1) < J(1)-\frac{2\delta^{\min}}{N}\Big(r-\sum_{i=1}^{N} g_i(u_i^{\text{best}})\Big)\sum_{z=1}^{K}\xi^{(z)}+N\cdot(\sigma^{\max})^2\sum_{z=1}^{K}\big(\xi^{(z)}\big)^2$$

Since $\sum_{z=1}^{+\infty}\xi^{(z)}=+\infty$ and $\sum_{z=1}^{+\infty}(\xi^{(z)})^2<+\infty$, we can always select K such that

$$J(1)-\frac{2\delta^{\min}}{N}\Big(r-\sum_{i=1}^{N} g_i(u_i^{\text{best}})\Big)\sum_{z=1}^{K}\xi^{(z)}+N\cdot(\sigma^{\max})^2\sum_{z=1}^{K}\big(\xi^{(z)}\big)^2<0$$

then we reach

$$J(K+1)<0$$

However, this contradicts the fact that $J(\cdot)$ is a nonnegative function. Therefore, the assumption that $I_2^{(z)}\neq\emptyset$ at any of step $z\in\{1,2,...,K\}$ does not hold. That is to say, $I_2^{(M)}=\emptyset$ at some step $M\leq K$.

Since we have proved that $I_2^{(M)}=\emptyset$ at some step $M\leq K$, then we have $\theta_i^{(M)}\geq g_i(u_i^{\text{best}})$, $u_i^{*,(M)}=u_i^{\text{best}}$ and $\lambda_i^{*,(M)}=0$ for all $i\in I$. Therefore, we have for all $i\in I$

$$\theta_i^{(z)}=\theta_i^{(M)},\ \ \forall z\geq M$$

and $\theta_i^{(z)}\geq g_i(u_i^{\text{best}})$ holds for all $z\geq M$. It is proved that $I_2^{(z)}$ is empty for all $z\geq M$. □

Proposition 4.1.2: In the case $\sum_{i=1}^{N} g_i(u_i^{\text{best}})<r$, there exists an $M\geq 0$ such that the overall optimal solution is attained at step $z=M$.

Proof:
According to Proposition 4.1.1, there exists an $M \geq 0$ such that $I_2^{(M)} = \emptyset$. Then, $\theta_i^{(M)} \geq g_i(u_i^{\text{best}})$ holds for all $i \in I$ and we have $u_i^{*,(M)} = u_i^{\text{best}}$ for all $i \in I$. Since $f_i(u_i) \geq f_i(u_i^{\text{best}})$ holds for all $i \in I$, it is directly prove that $u_i^{*,(M)} = u_i^{\text{best}}$ with $i \in I$ is the overall optimal solution.

Note that since $\theta_i^{(z)} = \theta_i^{(M)}$ for all $z > M$, we have $u_i^{*,(z)} = u_i^{*,(M)} = u_i^{\text{best}}$. Therefore, the overall optimal solution is also attained at step $z > M$. $\square$

4.2 For case 2

Proposition 4.2.1: $\forall \varepsilon > 0$, given $\sum_{i=1}^{N} g_i(x_i^{\text{best}}) = r - \varepsilon$, there exists an $M \geq 0$ such that at all steps $z \geq M$, $I_2^{(z)} = \emptyset$.

Proof:
We first assume that I_2 will never be empty and then find a contradiction. Next, we show that once I_2 is empty, it stays empty afterwards.

In the proof of Proposition 4.1.1, we have derived that at any step z, if $I_2^{(z)} \neq \emptyset$, we have

$$J(z+1) - J(z) < -2\xi^{(z)} \frac{\delta^{\min}}{N} \Big(r - \sum_{i=1}^{N} g_i(u_i^{\text{best}}) \Big) + N \cdot \left(\xi^{(z)}\right)^2 (\sigma^{\max})^2$$

Then since $\sum_{i=1}^{N} g_i(u_i^{\text{best}}) = r - \varepsilon$, we have

$$J(z+1) - J(z) < -2\xi^{(z)} \frac{\delta^{\min}}{N} \cdot \varepsilon + N \cdot \left(\xi^{(z)}\right)^2 (\sigma^{\max})^2$$

First, let K be an arbitrary integer. Then, if $I_2^{(z)} \neq \emptyset$ at each of step $z \in \{1, 2, ..., K\}$, we have

$$J(K+1) < J(1) - \frac{2\delta^{\min}}{N} \cdot \varepsilon \cdot \sum_{z=1}^{K} \xi^{(z)} + N \cdot (\sigma^{\max})^2 \sum_{z=1}^{K} \left(\xi^{(z)}\right)^2$$

Since $\sum_{z=1}^{+\infty} \xi^{(z)} = +\infty$ and $\sum_{z=1}^{+\infty} (\xi^{(z)})^2 < +\infty$, for any $\varepsilon > 0$, we can always select K such that

$$J(1) - \frac{2\delta^{\min}}{N} \cdot \varepsilon \cdot \sum_{z=1}^{K} \xi^{(z)} + N \cdot (\sigma^{\max})^2 \sum_{z=1}^{K} \left(\xi^{(z)}\right)^2 < 0$$

then we reach

$$J(K+1) < 0$$

However, this contradicts the fact that $J(\cdot)$ is a nonnegative function. Therefore, the assumption that $I_2^{(z)} \neq \emptyset$ at each step $z \in \{1, 2, ..., K\}$ does not hold. That is to say, $I_2^{(M)} = \emptyset$ at some step $M \leq K$.

Since we have proved that $I_2^{(M)} = \emptyset$ at some step $M \leq K$, then we have $\theta_i^{(M)} \geq g_i(u_i^{\text{best}})$, $u_i^{*,(M)} = u_i^{\text{best}}$ and $\lambda_i^{*,(M)} = 0$ for all $i \in I$. Therefore, we have for all $i \in I$

$$\theta_i^{(z)} = \theta_i^{(M)}, \ \ \forall z \geq M$$

and $\theta_i^{(z)} \geq g_i(u_i^{\text{best}})$ holds for all $z \geq M$. Hence, it has been proved that $I_2^{(z)}$ is empty for all $z \geq M$. $\square$

Proposition 4.2.2: $\forall \varepsilon > 0$, given $\sum_{i=1}^{N} g_i(u_i^{\text{best}}) = r - \varepsilon$, there exists an $M \geq 0$ such that at all steps $z \geq M$, $0 \leq \theta_i^{(z)} - g_i(u_i^{\text{best}})$ for all $i \in I$ and $\sum_{i=1}^{N} \left(\theta_i^{(z)} - g_i(u_i^{\text{best}})\right) = \varepsilon$.

Proof:
According to proposition 4.2.1, there exists an $M \geq 0$ such that at all step $z \geq M$, $I_2^{(z)} = \emptyset$. Therefore, at all step $z \geq M$, we have $\theta_i^{(z)} \geq g_i(u_i^{\text{best}})$ for all $i \in I$.

Since $\theta_i^{(z)} - g_i(u_i^{\text{best}}) \geq 0$ at all step $z \geq M$, we have for all $i \in I$, $\theta_i^{(z)} - g_i(u_i^{\text{best}}) \leq \sum_{i=1}^{N} \left(\theta_i^{(z)} - g_i(u_i^{\text{best}})\right)$. Further, since $\sum_{i=1}^{N} g_i(u_i^{\text{best}}) = r - \varepsilon$ and $\sum_{i=1}^{N} \theta_i^{(z)} = r$, we have $\sum_{i=1}^{N} \left(\theta_i^{(z)} - g_i(u_i^{\text{best}})\right) = \varepsilon$ and $\theta_i^{(z)} - g_i(u_i^{\text{best}}) \leq \varepsilon$. $\square$