The Carbon Removal Obligation
Updated analytical model and scenario analysis

Johannes Bednar, Artem Baklanov, Justin Macinante

Supplementary information

The partial derivative of a function $f(x_1, ..., x_n)$ with respect to the variable x_i is denoted by $\mathrm{D}_i f$.

Theorem 1 Consider an *original problem*:

$$\min_{x(\cdot)\in\mathcal{X}} \int_{\tau}^{\theta} f(t, x(t))dt \tag{1}$$

s.t. some additional constraints on $x(\cdot)$. Here $\mathcal{X}$ denotes the set of all continuous functions $x(\cdot) : [\tau, \theta] \longrightarrow X$, the set of admissible solutions. We assume that

1. X is a non-empty convex compact subset of $\mathbb{R}$ with non-empty interior $\mathrm{Int}\ X$;

2. f is continuous in $[\tau, \theta] \times X$;

3. function $f(t, \cdot)$ is strictly convex on X and differentiable on $\mathrm{Int}\ X$ for all $t \in [\tau, \theta]$;

4. $\mathrm{D}_2 f$, the derivative of f w.r.t to the second variable, is continuous in $[\tau, \theta] \times \mathrm{Int}\ X$ and has continuous extension to $[\tau, \theta] \times X$;

5. the optimal solution $x^*(\cdot)$ of (1) exists in the class of continuous functions such that $x^*(t) \in \mathrm{Int}\ X\ \forall t \in [\tau, \theta]$.

We claim that the same optimal solution holds in the following *modified (constraint-free)* problem:

$$\min_{x(\cdot)\in\mathcal{X}^{pw}} \int_{\tau}^{\theta} [f(t, x(t)) - \mathrm{D}_2 f(t, x^*(t))x(t)]dt \tag{2}$$

where x^* is the optimal solution in the original problem and $\mathcal{X}^{pw}$ denotes the set of all piece-wise continuous functions $x(\cdot) : [\tau, \theta] \longrightarrow X$. Note that now additional constraints are omitted and admissible solutions can be piece-wise continuous.

Proof.

- From a necessary and sufficient condition for a minimum in an elementary optimal control problem[1], it follows that any piece-wise continuous function $\hat{x} : [\tau, \theta] \to X$ that satisfies

$$\min_{x \in X} f(t,x) - \mathrm{D}_2 f(t, x^*(t))x = f(t, \hat{x}(t)) - \mathrm{D}_2 f(t, x^*(t))\hat{x}(t) \tag{3}$$

  for every point of continuity of $\hat{x}$ yields the global minimum in the modified problem.

- Note that for any $t \in [\tau, \theta]$

$$x \longmapsto f(t,x) - \mathrm{D}_2 f(t, x^*(t))x : X \longrightarrow \mathbb{R} \tag{4}$$

  is a strictly convex function defined on convex set X. Hence, we have two properties: (i) any local minimum is also a global one in X and (ii) the point of the global minimum is unique.

- By properties (i) and (ii), any stationary point of function (4) that is admissible (belongs to X) is the unique solution of the problem in the left-hand side of (3).

- By combining strict convexity and the first-order condition for local extrema, we obtain that for any t, $x^*(t)$ is a unique stationary point of function (4), which is also admissible by the definition since $x^*(t) \in \mathrm{Int}\ X$. Hence, x^* satisfies property (3): $\forall t \in [\tau, \theta]$

$$\min_{x \in X} f(t,x) - \mathrm{D}_2 f(t, x^*(t))x = f(t, x^*(t)) - \mathrm{D}_2 f(t, x^*(t))x^*(t).$$

- Finally, x^* is the unique solution of (2) in the following sense: any solution of (2) will be equal almost everywhere to x^* (in the sense of Lebesgue measure). For example, the solutions may differ in a finite number of points but this will not influence the optimal result in (2). □

Note that Theorem 1 allows a straightforward extension to the case of two variables (x_1 and x_2):

Theorem 2 Consider an *original problem*:

$$\min_{x_1(t), x_2(t) \in X} \int_\tau^\theta f(t, x_1(t), x_2(t))dt \tag{5}$$

s.t. some additional constraints on x_1, x_2. Here $\mathcal{X}$ denotes the set of all continuous functions $x(\cdot) : [\tau, \theta] \longrightarrow X$ such that $x(t) \in X$. We assume that

1. X is non-empty convex compact subset of $\mathbb{R}$ with non-empty interior;
2. f is continuous in $[\tau, \theta] \times X \times X$;

[1] see p.160 in Alekseev V.M., Tikhomirov V.M., Fomin S.V. *Optimal Control*. Springer Science + Business Media, New York (1987)

3. $f(t,\cdot,\cdot)$ is strictly convex on $X \times X$ and differentiable on Int $X \times$ Int X for all $t \in [\tau, \theta]$;

4. $\mathrm{D}_2 f$, $\mathrm{D}_3 f$ are continuous in $[\tau,\theta] \times$ Int $X \times$ Int X and have continuous extensions to $[\tau,\theta] \times X \times X$;

5. the optimal solution $x_1^*(\cdot), x_2^*(\cdot)$ of (5) exists in the class of continuous functions such that $x_1^*(\cdot), x_2^*(\cdot) \in$ Int X $\forall t \in [\tau,\theta]$.

We claim that the same optimal solution holds in the following *modified (constraint-free) problem*:

$$\min_{x_1,x_2 \in \mathcal{X}^{pw}} \int_\tau^\theta [f(t, x_1(t), x_2(t)) \\ - x_1(t)\, \mathrm{D}_2 f(t, x_1^*(t), x_2^*(t)) - x_2(t)\, \mathrm{D}_3 f(t, x_1^*(t), x_2^*(t))]dt \quad (6)$$

where x^* is the optimal solution in the original problem and $\mathcal{X}^{pw}$ denotes the set of all piece-wise continuous functions $x(\cdot) : [\tau,\theta] \longrightarrow X$ such that $x(t) \in X$. Note that now additional constraints are omitted and admissible solutions can be piece-wise continuous.

How we use Theorem 2. In the paper, we dial with the problem

$$\min_{\alpha,\mu} \int_{T_0}^{T} f_c(t, \alpha(t), \mu(t)) e^{-R(t-T_0)} dt, \quad (7)$$

s.t. the additional scenario-specific constraints. For applying Theorem 2, we need to transform variables $\alpha(t)$ and $\mu(t)$ such that their ranges of possible values do not depend on time, i.e., we will introduce normalised versions of this variables with the constant range.

First, let us recall the definitions of cost functions:

$$f_c(t, \alpha(t), \mu(t)) = E(t)\Big(\int_0^{\alpha(t)} f_\alpha(t,a)\, da + \int_0^{\mu(t)} f_\mu(t,a)\, da\Big)$$

$$f_i(t,a) = Q_i(t)\Big(\frac{a}{U_i(t) - a}\Big)^{C_i}, C_i > 0, i = \alpha, \mu,$$

where $Q_i, U_i : [T_0, T] \to (0, \infty)$ are given continuous functions.

Second, by using the following transformations for variables

$$\tilde{\alpha}(t) = \frac{\alpha(t)}{U_\alpha(t)}, \tilde{\mu}(t) = \frac{\mu(t)}{U_\mu(t)}, \quad (8)$$

we modify the cost functions:

$$\tilde{f}_i(t,a) = Q_i(t)\Big(\frac{a}{1-a}\Big)^{C_i}, i = \alpha, \mu;$$

$$\tilde{f}_c(t, \tilde{\alpha}(t), \tilde{\mu}(t)) = E(t)\Big(U_\alpha(t) \int_0^{\tilde{\alpha}(t)} \tilde{f}_\alpha(t,a)\, da + U_\mu(t) \int_0^{\tilde{\mu}(t)} \tilde{f}_\mu(t,a)\, da\Big).$$

The relationship between functions is quite straightforward:

$$\tilde{f}_c(t, \tilde{\alpha}(t), \tilde{\mu}(t)) = f_c(t, \tilde{\alpha}(t)U_\alpha(t), \tilde{\mu}(t)U_\mu(t));$$

$$\tilde{f}_i(t, a) = f_i(t, aU_i(t)), i = \alpha, \mu.$$

Note that tilde sign ($\tilde{}$) indicates that the object operates with transformed variables.

Third, we specify Theorem 2 for our case:

$$f(\hat{t}, \hat{x}_1, \hat{x}_2) = \tilde{f}_c(\hat{t}, \hat{x}_1, \hat{x}_2)e^{-R(\hat{t}-T_0)}; x_1 = \tilde{\alpha}, x_2 = \tilde{\mu}, \tau = T_0, \theta = T.$$

We assume that $X = [\epsilon, 1-\epsilon]$, where $\epsilon > 0$ is small enough to not influence the optimal solution. Given the form of the function of total costs, the existence of optimal solutions is sufficient for such ϵ to be well defined. Hence the normalised abatement rate at instant t can be in the range $[\epsilon, 1-\epsilon]$. Introduction of such ϵ is also a standard practice in numerical optimisation.

Then

$$\mathrm{D}_2 f(t, \tilde{\alpha}, \tilde{\mu}) = e^{-R(t-T_0)}\,\mathrm{D}_2\tilde{f}_c(t, \tilde{\alpha}, \tilde{\mu}) = e^{-R(t-T_0)}E(t)U_\alpha(t)\tilde{f}_\alpha(t, \tilde{\alpha}),$$

$$\mathrm{D}_3 f(t, \tilde{\alpha}, \tilde{\mu}) = e^{-R(t-T_0)}\,\mathrm{D}_3\tilde{f}_c(t, \tilde{\alpha}, \tilde{\mu}) = e^{-R(t-T_0)}E(t)U_\mu(t)\tilde{f}_\mu(t, \tilde{\mu}).$$

Thus, the original problem is

$$\min_{\tilde{\alpha}(t), \tilde{\mu}(t) \in [\epsilon, 1-\epsilon]} \int_{T_0}^{T} \tilde{f}_c(t, \tilde{\alpha}(t), \tilde{\mu}(t))e^{-R(t-T_0)}dt, \tag{9}$$

s.t. the same constraints as in (7) that are accordingly adjusted. Note that solutions of (7) after normalisation (8) are exactly the solutions of (9). This follows from our definitions of all functions with tilde sign[2]. The modified problem for (9), in the context of Theorem 2, is as follows

$$\begin{aligned}\min_{\tilde{\alpha}(t), \tilde{\mu}(t) \in [\epsilon, 1-\epsilon]} \int_{T_0}^{T} [\tilde{f}_c(t, \tilde{\alpha}(t), \tilde{\mu}(t)) \\ - E(t)\tilde{f}_\alpha(t, \tilde{\alpha}^*(t))U_\alpha(t)\tilde{\alpha}(t) \\ - E(t)\tilde{f}_\mu(t, \tilde{\mu}^*(t))U_\mu(t)\tilde{\mu}(t)]e^{-R(t-T_0)}dt \end{aligned} \tag{10}$$

By using (8), we restore our original variables obtaining the equivalent of (10):

$$\begin{aligned}\min_{\substack{\alpha(t) \in [\epsilon U_\alpha(t), (1-\epsilon)U_\alpha(t)], \\ \mu(t) \in [\epsilon U_\mu(t), (1-\epsilon)U_\mu(t)]}} \int_{T_0}^{T} [f_c(t, \alpha(t), \mu(t)) \\ - E(t)f_\alpha(t, \alpha^*(t))\alpha(t) \\ - E(t)f_\mu(t, \mu^*(t))\mu(t)]e^{-R(t-T_0)}dt \end{aligned} \tag{11}$$

[2]We thank Mikhail Gomoyunov for hinting us to use the elegant transformation instead of proving Theorem 2 for a more general case.

Recall that $e(t) = E(t)(1-\alpha(t))$ and $r(t) = E(t)\mu(t)$. Since $P_\alpha(t) = f_\alpha(t, \alpha^*(t))$, $P_\mu(t) = f_\mu(t, \mu^*(t))$, we have that (11) takes the form

$$\begin{aligned}\min_{\substack{\alpha(t)\in[\epsilon U_\alpha(t),(1-\epsilon)U_\alpha(t)],\\ \mu(t)\in[\epsilon U_\mu(t),(1-\epsilon)U_\mu(t)]}} \int_{T_0}^{T} [f_c(t, \alpha(t), \mu(t)) \\ + P_\alpha(t) \underbrace{E(t)(1-\alpha(t))}_{e(t)} \\ - P_\mu(t) \underbrace{E(t)\mu(t)}_{r(t)} - \Delta(t)] e^{-R(t-T_0)} dt \qquad (12)\end{aligned}$$

where $\Delta(t) = P_\alpha(t)E(t)$ does not depend on α, μ hence can be omitted since we aim only for optimal α, μ. In this sense, problem (12) is equivalent to the following problem

$$\min_{\alpha,\mu} \int_{T_0}^{T} [f_c(t, \alpha(t), \mu(t)) + P_\alpha(t)e(t) - P_\mu(t)r(t)] e^{-R(t-T_0)} dt. \qquad (13)$$

Finally, we conclude that (13) is indeed solved by optimal α^*, μ^* of (7); this result was used in the main text.

Remark: Theorem 2 requires for $\tilde{f}_c(t, \cdot, \cdot)e^{-R(t-T_0)}$, in particular, to have continuous partial derivatives and to be strictly convex in $[\epsilon, 1-\epsilon] \times [\epsilon, 1-\epsilon]$ for any $t \in [T_0, T]$. Note that the multiplier $e^{-R(t-T_0)}$ does not play an important role since it is a positive constant for any instant t, so we focus only on $\tilde{f}_c(t, \cdot, \cdot)$. The costs are defined such that for a fixed $t \in [T_0, T]$, instantaneous cost function $\tilde{f}_c(t, \cdot, \cdot)$ is twice continuously differentiable in $[\epsilon, 1-\epsilon] \times [\epsilon, 1-\epsilon]$, which implies continuous differentiability. This allows us to consider the Hessian matrix for $\tilde{f}_c(t, \cdot, \cdot)$ being a function of two variables:

$$\mathbf{H}(\tilde{\alpha}, \tilde{\mu}) = \begin{pmatrix} E(t)U_\alpha(t)\,\mathrm{D}_2\tilde{f}_\alpha(t, \tilde{\alpha}) & 0 \\ 0 & E(t)U_\mu(t)\,\mathrm{D}_2\tilde{f}_\mu(t, \tilde{\mu}) \end{pmatrix}.$$

For $\tilde{\alpha}, \tilde{\mu} \in [\epsilon, 1-\epsilon]$, this matrix is positive definite since $E(t)U_i(t) > 0, i = \alpha, \mu$, and both derivatives are positive. Hence, this implies strict convexity of $\tilde{f}_c(t, \cdot, \cdot)$.