Optimal Boundary Control of Hyperbolic Balance Laws with State Constraints

JOHANN M. SCHMITT* AND STEFAN ULBRICH †

Abstract. In this paper we analyze the optimal control of initial-boundary value problems for entropy solutions of scalar hyperbolic balance laws with pointwise state constraints. Hereby, we suppose that the initial and the boundary data switch between different $C^1$-functions at certain switching points, where the $C^1$-functions and the switching points are considered as the control. For a class of cost functionals, we prove first order necessary optimality conditions for the corresponding optimal control problem with state constraints. Furthermore, we use a Moreau-Yosida type regularization to approximate the optimal control problem with state constraints. We derive optimality conditions for the regularized problems and finally prove convergence to the solution of the optimal control problem with state constraints.

Key words. optimal control, scalar conservation law, state constraints, Moreau-Yosida

AMS subject classifications. 49K20, 35L65, 35R05, 49J50

1. Introduction. In this paper we derive necessary optimality conditions for state constrained optimal control problems of the form

$$\begin{equation}
\min J(y(w)) := \int_a^b \psi(y(\tilde{t}, x; w), y_d(x)) \, dx + R(w),
\end{equation}$$

where $\psi \in C^{1,1}_{loc}(\mathbb{R}^2)$, $y_d \in BV_{loc}(\mathbb{R}) \cap L^\infty(\mathbb{R})$ is the desired state, $R : W \to [0, \infty)$ is a Fréchet differentiable regularization term and $y$ is given by the solution of the following initial boundary value problem (IBVP)

$$\begin{align}
(1a) \quad y_t + f(y)_x &= g(\cdot, y, u_1), & \text{on } \Omega_1 := (0, \bar{t}) \times (a, b), \\
(1b) \quad y(0, \cdot) &= u_0(\cdot; w), & \text{on } \Omega := (a, b), \\
(1c) \quad y(\cdot, a+) &= u_{B,a}(\cdot; w), & \text{in the sense of } (6), \\
(1d) \quad y(\cdot, b-) &= u_{B,b}(\cdot; w), & \text{in the sense of } (6),
\end{align}$$

where $-\infty < a < b < \infty$, and the following state constraints are satisfied

$$\begin{equation}
y(\tilde{t}, x) \leq \bar{y}(x) \quad \text{for all } x \in [a, b].
\end{equation}$$

Here, we assume $\bar{y} \in C^1([a, b])$. One could restrict the cost functional in (P) to an interval $[a_1, b_1] \subset [a, b]$ by multiplying $\psi$ with a weight function $\gamma$. We associate with $w = (u^0, u^{B,a}, u^{B,b}, x^0, u_{a}, u_{b}) \in W_{ad}$, where $W_{ad} \subset C^1(\Omega_{n+1} \times \Omega_{m+1} \times \Omega_{m+1} \times \Omega_{m+1} \times \Omega_{m+1} \times \Omega_{m+1} \times \Omega_{m+1} \times \Omega_{m+1} \times \Omega_{m+1} \times \Omega_{m+1} \times \Omega_{m+1} \times \Omega_{m+1} \times \Omega_{m+1})$, the initial and boundary data

$$\begin{equation}
u_0(x; w) = \begin{cases}
u_0^0(x) & \text{if } x \in (a, x_0], \\nu_j^0(x) & \text{if } x \in (x_{j-1}^0, x_j^0], 2 \leq j \leq n_x, \\nu_{n_x+1}^0(x) & \text{if } x \in (x_{n_x}^0, b),
\end{cases}
\end{equation}$$

*Submitted to the editors DATE.

Funding: Supported by the German Research Foundation (DFG) within the Collaborative Research Center TRR 154 "Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks", project A02, and by the DFG within the Collaborative Research Center SFB 1194 "Interaction between Transport and Wetting Processes", project B04.

†Department of Mathematics, TU Darmstadt, Germany (jschmitt@mathematik.tu-darmstadt.de).

‡Department of Mathematics, TU Darmstadt, Germany (stefan.ulbrich@tu-darmstadt.de).

This manuscript is for review purposes only.
\[ u_B,l(t;w) = \begin{cases} u_{B,l}^1(t) & \text{if } t \in [0,t_1^l], \\ u_{B,l}^j(t) & \text{if } t \in (t_{j-1}^l,t_j^l], \\ u_{n_{l,t}+1}^l(t) & \text{if } t \in (t_{n_{l,t}},\bar{t}). \end{cases} \] (4)

**Notation 1.** In order to treat the boundaries at \( x = a \) and \( x = b \) with a uniform notation, we define for a function \( v : (a,b) \to \mathbb{R} \) of bounded variation

\[
\sigma(l) = \begin{cases} 1 & \text{if } l = a, \\ -1 & \text{if } l = b, \end{cases} \quad \text{and} \quad v(l_0(t)) = \begin{cases} v(a^+) & \text{if } l = a, \\ v(b^-) & \text{if } l = b. \end{cases}
\]

**Related literature.** Since hyperbolic conservation laws do not admit unique weak solutions (see [1]), one has to consider entropy solutions of (1) in the sense of [24] in order to guarantee uniqueness, see (5) below. Optimal control problems with state constraints have been studied in several papers, e.g., [7,19,21,23]. But to the best of the authors knowledge the optimal control of hyperbolic balance laws with state constraints has not been discussed so far.

The derivation of optimality conditions for the optimal control of hyperbolic balance laws with pointwise state constraints is involved, since in order to guarantee a constraint qualification one has usually to assume that the control-to-state mapping is continuously differentiable to \( L^\infty \). Since it is well known that entropy solutions develop shocks after finite time (see, e.g., [1]), the control-to-state mapping is in this case not even continuous to \( L^\infty \).

Optimal control of hyperbolic balance laws has been considered in several papers, e.g., [2,5,6,8,25,26,34,35,38,39]. The authors in [2,5,6] suppose that the state is piecewise smooth, respectively piecewise Lipschitz continuous, and only consider structure-preserving variations of the initial data. In contrast to these papers, the authors in [34,35,38], who restrict themselves to the scalar case, do not assume a priori knowledge of the shock structure. The developments in this paper are mainly based on [34], where the concept of shift-differentiability [38,39] has been extended to initial-boundary value problems of the form (1). The proofs of the results in [34,35,39,40] are based on the special structure of solutions of hyperbolic balance laws, that was derived by using the concept of generalized characteristic in [10]. In this paper we want to exploit this structural properties and introduce new state variables with \( C^1 \) regularity as a technical tool. This will allow us to derive first order necessary optimality conditions for (P) in terms of the new state variables. These optimality conditions can be transformed back to the original formulation of problem (P). Furthermore, using a Moreau-Yosida regularization approach to handle state constraints algorithmically, which was first introduced in [21], we can show that the optimal control and state of the regularized problems converge strongly to the optimal control and state of problem (P) (see, e.g., [11,30]), while the corresponding Lagrange multiplier approximations obtained form the regularized problem converge weakly to the Lagrange multipliers of (P) for the state constraints.

**1.1. Contributions.** To the best of the authors’ knowledge, this is the first attempt to derive necessary optimality conditions for optimal control problems governed by hyperbolic balance laws and state constraints. In contrast to [2,5,6], we do not assume a priori knowledge of the shock structure. In addition, we prove that the Robinson Constraint Qualification is satisfied in an optimal solution under
mild assumptions. Moreover, we carry out a convergence analysis of Moreau-Yosida regularizations, which provides a theoretical basis for conveniently computing solutions of the optimality system without any prior knowledge of the shock structure. In contrast to [5,6], the introduction of shock curves as additional state variables is not required. Moreover, these results form the basis of numerical approaches for solving optimal control problems with hyperbolic balance laws and state constraints. We further present a practical example dealing with the optimal control of traffic flow. In this example, we demonstrate that the results of this paper are applicable to real world problems. Finally, the present paper extends the sensitivity and adjoint calculus of [34,38,39] to the case where the shifting of rarefaction centers is allowed.

1.2. Organization. The paper is organized as follows. In section 2 we introduce basic assumptions and collect results concerning the well-posedness of the IBVP and structural properties of its solution. In section 3 we discuss the optimal control problem (P), prove existence of a global optimum and derive first order necessary optimality conditions. In section 4 we apply the Moreau-Yosida regularization approach and prove convergence to the solution of the optimal control problem with state constraints.

2. The initial-boundary value problem. In this section we collect some results for the IBVP (1). The norm of \( L^r(D) \) with a measurable domain \( D \) will be denoted by \( \| \cdot \|_{r,D} \), \( 1 \leq r \leq \infty \), and the scalar product of \( L^2(D) \) by \( \langle \cdot, \cdot \rangle_{2,D} \).

2.1. Notion of a solution for the IBVP and basic assumptions. We consider entropy solutions of (1) in the sense of [24], i.e., for every (Kružkov-) entropy \( \eta_c(\lambda) := |\lambda - c|, c \in \mathbb{R} \), and associated entropy flux \( q_c(\lambda) := \text{sgn}(\lambda - c) (f(\lambda) - f(c)) \) \( y \) has to satisfy

\[
(5a) \quad \lim_{t \to 0^+} \eta_c(y(t)) + (q_c(y(t)))_x - \eta'_c(y(t))g(\cdot, y, u_1) \leq 0 \quad \text{in } \mathcal{D}'(\Omega_t),
\]

\[
(5b) \quad \lim_{t \to 0^+} \|y(t, \cdot) - u_0\|_{L^1(\Omega_t)} = 0 \quad \text{for all } R > 0.
\]

In order to guarantee that the problem is well-posed, the boundary conditions in (1) have to be understood in the sense of [1], i.e. for \( t \in (a, b) \) it holds

\[
(6) \quad \min_{k \in I} \sigma(l)[u_{B,l} - y(\cdot, l^1)](f(y(\cdot, l^1)) - f(k)) = 0 \quad \text{a.e. on } [0, t],
\]

with \( I(\alpha, \beta) := \lfloor \min(\alpha, \beta), \max(\alpha, \beta) \rfloor \).

Remark 2. We observe that condition (6) involves boundary traces \( y(\cdot, a^+) \) and \( y(\cdot, b^-) \). These limits exist if, e.g., \( y(t, \cdot) \in BV(a, b) \), which holds under suitable assumptions, see Proposition 4. In [28,29,32,41], characterizations of entropy solutions are presented, which are even valid in the \( L^\infty \)-setting. More precisely, the authors in [32] and [28] introduce so-called boundary entropy-entropy-flux pairs to replace (6) for the homogeneous case \( g = 0 \). Vovelle uses in [41] semi-Kružkov-entropies to characterize entropy solutions by entropy inequalities, which is extended to the inhomogeneous case by Martin in [29]. We emphasize that in the case of \( BV \)-regularity, all these characterizations of entropy solutions are equivalent. Since we will restrict ourselves to the \( BV \)-setting in the present paper, we use (5) and (6) to characterize entropy solutions.

We will work under the following assumptions. 

(A1) The flux function satisfies \( f \in C^{3}_{\text{loc}}(\mathbb{R}) \), \( f^{\beta-1} \in C^{2,\beta}_{\text{loc}}(\mathbb{R}) \) for some \( \beta \in (0, 1] \) and is strongly convex, i.e., there exists a positive constant \( m_{f''} > 0 \) such that
Finally, let $g$ be Lipschitz w.r.t. $x$ and affine linear w.r.t. $y$.

(A2) The set of admissible controls $W_{ad}$ is nonempty, convex and compact in $W := C^1([\bar{t}, \bar{t}]^{n_x+1} \times U_{B,a}^0 \times U_{B,b}^0 \times \mathcal{X} \times \mathcal{T} \times \mathcal{C}([0, \bar{t}]; C^1(\mathbb{R})^m)$, where $\mathcal{X} := \{ \bar{x} \in \Omega_{nx} : a < x_1 < \ldots < x_{nx} < b \}$, $\mathcal{T}^t := \{ \bar{t} \in [0, \bar{t}]^{n_t} : 0 < t_1 < \ldots < t_{n_t,1} < \bar{t} \}$, $U_{B,t} := \{ u^{B,t} \in C^1([0, \bar{t}]^{n_x+1} : \sigma(l)f'(u^{B,t}_j) \geq \alpha, j = 1, \ldots, n_{t,l}+1 \}$, for $l \in \{ a, b \}$ and some $\alpha > 0$ and $W$ is equipped with the norm

$\|w\| := \|u_0\|_{C^1([\bar{t}, \bar{t}]^{n_x+1})} + \|u^{B,a}\|_{C^1([0, \bar{t}]^{n_x+1})} + \|u^{B,b}\|_{C^1([0, \bar{t}]^{n_x+1})}$

For all $w \in W_{ad}$, it holds that $\|u_1\|_{L^\infty([0, \bar{t}], C^1([\bar{t}, \bar{t}]^{n_x+1}))} \leq M_u$ for some $M_u > 0$.

**Notation 3.** Given some $w \in W$, we will use the abbreviations $[\psi(x)] := \psi(x-) - \psi(x+)$, $I_{j,0}(w) := (a_{k-1}^j, x_0^j)$ for $j = 1, \ldots, n_x + 1$ and $I_{t,l,j}(w) := (t_{j-1,l}^t, t_j^t)$ for $j = 1, \ldots, n_{t,l} + 1$ and set $x_0^j := a, x_0^{n_x+1} := b, t_0^t := 0$ and $t_{n_{t,l}+1}^t := \bar{t}$ for $l \in \{ a, b \}$. Furthermore, we define $B^W_p(w) := \{ \tilde{w} \in W : \|w - \tilde{w}\|_W \leq \rho \}$ and the sets of indices

$I_{x,0}(w) := \{ j \in \{ 1, \ldots, n_x \} : |u_0(x_0^j)| \geq 0 \}$,

$I_{x,l}(w) := \{ j \in \{ 1, \ldots, n_{x,l} \} : \sigma(l) |u_{B,l}(t_j^l)| \leq 0 \}$, $l \in \{ a, b \}$,

$I_{t,0}(w) := \{ 1, \ldots, n_x \} \setminus I_{x,0}(w)$, $I_{t,l}(w) := \{ 1, \ldots, n_t \} \setminus I_{x,l}(w)$, $l \in \{ a, b \}$.

**2.2. Existence and uniqueness of a solution for the IBVP.** We obtain the following existence and uniqueness result for entropy solution of (1) [1,9,32]. Oleinik’s entropy condition (7) can be found in [33, Lemma 3.1.13].

**Proposition 4.** Let (A1) and (A2) hold. Then for every $w \in W$ (1) has a unique entropy solution $y = y(w) \in L^\infty(\Omega_t)$. Moreover, after a modification on a set of measure zero $y$ satisfies $y \in C([0, \bar{t}]; L^1_{loc}(\Omega))$ and there exist constants $M_y, L_y > 0$ such that for all $w, \tilde{w} \in W_{ad}$ and all $t \in (0, \bar{t})$ the following estimates hold.

$\|y(t, \cdot; w)\|_{L^\infty([0,a], b)} \leq M_y$,

$\|y(t, \cdot; w) - y(t, \cdot; \tilde{w})\|_{L^1([0,a], b)} \leq L_y \left( \|u_0(w) - u_0(\tilde{w})\|_{L^1([0,a], b)} + \|u_{B,a}(w) - u_{B,a}(\tilde{w})\|_{L^1([0,a], b)} + \|\tilde{u}_1 - u_{B,b}(\tilde{w})\|_{L^1([0,a], b)} \right)$,

where $a < b$, $I_1 := [a - tM_{f'}, b + tM_{f'}] \cap \Omega$ and $M_{f'} := \max_{|y| \leq M_y} |f'(y)|$.

Finally, $y$ satisfies $y \in BV(\Omega_t)$ and there exists a constant $C > 0$ such that for all $t \in (0, \bar{t})$ and all $\varepsilon_1, \varepsilon_2 > 0$ the solution $y$ satisfies

$\|y(t, \cdot; \tilde{w}) - C\|_{L^\infty([0,a], b)} \leq \frac{C}{1 - e^{-m_{f'} C \min(t, \varepsilon_1, \varepsilon_2)}} |a + \varepsilon_1, b - \varepsilon_2|$

in the sense of distributions yielding $y(t, x-) \geq y(t, x+)$ for all $(t, x) \in \Omega_t$.
CONVENTION 5. We consider the representative of \( y \) satisfying \( y \in C([0, t]; L^1(\Omega)) \) and \( y(t, x) = y(t, x-) \) for all \( (t, x) \in [0, t] \times (a, b) \), \( y(t, a) = y(t, a+) \) for all \( t \in [0, t] \).

LEMMA 6. Under Assumptions (A1) and (A2) for all \( \hat{w}, \hat{w} \in W_{ad} \) it holds for \( u_0, u_{B,a}, u_{B,b} \) defined in (3) and (4) that
\[
\| u_0(w) - u_0(\hat{w}) \|_{1,[a,b]} \leq (|b - a| + \| w \|_W + \| \hat{w} \|_W) \| w - \hat{w} \|_W \\
\| u_{B,1}(w) - u_{B,1}(\hat{w}) \|_{1,[0,\hat{t}]} \leq (\hat{t} + \| w \|_W + \| \hat{w} \|_W) \| w - \hat{w} \|_W, \quad l \in \{a, b\}.
\]

Proof. Since by (3) the locations of the \( i \)th discontinuity differ by \( |x^0_i - \hat{x}_i^0| \) and \( |u_0(x; w) - u_0(x; \hat{w})| \) can be bounded by \( \| u_0^0 \|_{C(\Omega)}^{\#_{x+1}} + \| \hat{u}_0^0 \|_{C(\Omega)}^{\#_{x+1}} \), we obtain
\[
\| u_0(w) - u_0(\hat{w}) \|_{1,[a,b]} \leq |b - a| \| u_0^0 - \hat{u}_0^0 \|_{C(\Omega)}^{\#_{x+1}} \\
+ \| u_0^0 \|_{C(\Omega)}^{\#_{x+1}} + \| \hat{u}_0^0 \|_{C(\Omega)}^{\#_{x+1}} \| x^0_i - \hat{x}_i^0 \|_1 \leq (|b - a| + \| w \|_W + \| \hat{w} \|_W) \| w - \hat{w} \|_W.
\]
Analogously, one can show the remaining two inequalities. \( \square \)

Using Proposition 4, Lemma 6 and the boundedness of \( W_{ad} \) we obtain

**Corollary 7.** Let (A1) and (A2) hold. Then there exists \( L_y > 0 \) such that
\[
\| y(t, \cdot; w) - y(t, \cdot; \hat{w}) \|_{1,[a,b]} \leq L_y \| w - \hat{w} \|_W \quad \text{for all} \ w, \hat{w} \in W_{ad}.
\]

**Corollary 7.** \( \psi \in C_{loc}^{1,1}(\mathbb{R}^2) \) and \( y_{ad} \in BV([a, b]) \) yield (cf. [40]):

**Corollary 8.** Let (A1) and (A2) hold. Then there exists \( L > 0 \) such that
\[
| J(y(w)) - J(y(\hat{w})) | \leq L \| w - \hat{w} \|_W \quad \text{for all} \ w, \hat{w} \in W_{ad}.
\]

2.3. Structure of the solution of the IBVP. In this section we will summarize some results of [34] concerning the structure of entropy solutions of (1) which can be obtained by Dafermos’ theory of generalized characteristics [10], see [33, 34, 39, 40].

**Definition 9.** A Lipschitz curve \( [\alpha, \beta] \subset [0, \hat{t}] \rightarrow \Omega, t \rightarrow (t, \xi(t)) \) satisfying
\[
(8) \quad \dot{\xi}(t) \in \{ f'(y(t, \xi(t)+)), f'(y(t, \xi(t)\)) \} \quad \text{a.e. on} \ [\alpha, \beta]
\]
is called a generalized characteristic on \([\alpha, \beta]\). If the lower and upper bounds in (8) coincide for almost all \( t \in [\alpha, \beta] \), then \( \xi \) is called genuine. \( \xi^{\pm} \) is called a maximal/minimal characteristic, if it satisfies \( \xi^{\pm}(t) = f'(y(t, \xi(t)\pm)) \).

According to [10, Theorem 3.3], if \( \xi \) is a genuine characteristic, \( y(\cdot, \xi(\cdot)) = v(\cdot) \) holds almost everywhere on some interval \([\alpha, \beta]\), where \( v \) is given by the solution of the characteristic equation
\[
(9) \quad \dot{\xi}(t) = f'(v(t)), \quad \dot{\upsilon}(t) = g(t, \xi(t), v(t), u_1(t, \xi(t))),
\]
and \( \xi(\cdot) \) coincides with \( \xi(\cdot) \) on \([\alpha, \beta]\) almost everywhere. For given initial conditions \( (\zeta, v)(\tau) = (z, w) \) with \( z \in [a, b] \) and \( w \in \mathbb{R} \), we denote the solution of (9) by \( (\zeta, v)(\cdot, \tau, z, w, u_1) \). The solutions of the characteristic equation (9) and its linearized version form the basis for the sensitivity and adjoint calculus which is developed in [34] for problems of the form (P), see also [38–40].

**Notation 10.** Let
\[
T_t := \{ \theta \in [0, t] : \sigma(l)f'(y(\theta+), l_{\sigma(l)}) > 0 \land \sigma(l)f'(y(\theta-), l_{\sigma(l)}) < 0 \}, \quad l \in \{a, b\}
\]
and denote for all \( \theta \in T^a \cup T^b \) the corresponding maximal/minimal backward characteristic through \((\theta, a/b)\) by \( \xi^0_{a/b} \) and the corresponding timepoint where the characteristic leaves the domain \( \Omega \) by \( \vartheta_{a/b}^\theta \). Analogously to [34], we define

\[
D_- := \bigcup_{\theta \in T^a} \{(t, x) : t \in]0, \xi^0_{a}(t)[\} \cup \bigcup_{\theta \in T^b} \{(t, x) : t \in]0, \xi^0_{b}(t), b[\}.
\]

If \( \xi^0_{a/b} \) ends in a point \((\vartheta^0_{a/b}, b/a)\), we extend \( D_- \) by adding \([0, \vartheta^0_{a/b}] \times \Omega \) to the right hand side of (10).

In [33, Lemma 3.1.17] the following result is shown.

**Lemma 11.** Let (A1) and (A2) hold and consider some \( w \in W \) satisfying

\[
\inf_{t : u_{B,l}(t,w) \neq y(l,\sigma(t);w)} |f(u_{B,l}(t;w)) - f(y(t,\sigma(t);w))| > 0, \quad l \in \{a,b\}.
\]

Then the sets \( T^a/b \) are finite and it holds that \( u_{B,a}(\theta+) > u_{B,a}(\theta-) \) for all \( \theta \in T^a \) and \( u_{B,b}(\theta+) < u_{B,b}(\theta-) \) for all \( \theta \in T^b \).

We define nondegeneracy of shocks according to [34]:

**Definition 12.** A discontinuity \( \bar{x} \) of \( y(\bar{t},\cdot;w) \) is called nondegenerated, if it is not the center of a centered compression wave (see [10, Definition 4.3]) and the corresponding minimal/maximal backward characteristic through \( (\bar{x}, \bar{t}) \) end in some point \((\bar{x}, 0), (a, \bar{t})\) or \((b, \bar{t})\) where \( u_0, u_{B,a} \) or \( u_{B,a} \) is continuously differentiable, respectively, or lie in the interior of a rarefaction wave which is created either by a discontinuity of \( u_0 \) or \( u_{B,a} \).

We will work under the following nondegeneracy condition.

**(ND)** A control \( w \in W \) is called nondegenerated if the following holds: There is no point \( x \in \Omega \) or \( t \in [0, \bar{t}] \) where \( u_0(\cdot;w), u_{B,a}(\cdot;w) \) or \( u_{B,b}(\cdot;w) \) is continuous, but not differentiable. The corresponding solution \( y(\bar{t}, \cdot;w) \) has no shock generation points on \([a,b]\) and a finite number of discontinuities \( a < x_1 < \cdots < x_N < b \) that are no shock interaction points and nondegenerated according to Definition 12. Moreover, (11) is satisfied and for all \( t^j_l \in T^l \) with \( l \in \{a,b\} \) the following is satisfied: The extreme backward characteristic \( \xi^{l\prime}_l \) through \((l, t^j_l)\) ends in the interior of a rarefaction wave or in a point \((0, \bar{z})\) or \((\bar{\theta}, l)\) where \( u_0 \) respectively \( u_{B,l} \) is smooth. In the latter case we require that

\[
\frac{d}{dz}\zeta(t;0,z,u_0(z),u_1)|_{z=\bar{z}} \geq \beta \quad \forall t \in [0, t^j_l]
\]

respectively

\[
\sigma(l) \frac{d}{d\theta}\zeta(t;\theta,l,u_{B,l}(\theta),u_1)|_{\theta=\bar{\theta}} \leq -\beta \quad \forall t \in [\bar{\theta}, t^j_l]
\]

holds for some constant \( \beta > 0 \). Conversely, if \( \xi \) ends in the interior of a rarefaction wave, it has to hold

\[
\frac{d}{dw}\zeta(t;0,\bar{z},w,u_1)|_{w=\bar{w}} \geq \beta \quad \forall t \in [0, t^j_l] \quad \text{with } \bar{w} \in ]u_0(z-),u_0(z+)]
\]

respectively

\[
\frac{d}{dw}\zeta(t;\theta,l,w,u_1)|_{w=\bar{w}} \geq \beta(t-\bar{\theta}) \quad \forall t \in [\bar{\theta}, t^j_l]
\]

with \( \bar{w} \in ]u_{B,a}(\bar{\theta}+), u_{B,a}(\bar{\theta}-)[ \) respectively \( \bar{w} \in ]u_{B,b}(\bar{\theta}-), u_{B,b}(\bar{\theta}+)[ \).
Remark 13. The conditions in (ND) imply that the points $t_j^{a/b} \in T^{a/b}$ are nondegenerated in the sense of [34, Definition 4.12].

We extend now the results [33, Lemma 6.3.1, Theorem 6.3.8] such that shifts of rarefaction centers are also allowed by following the ideas of [33, Chapter 8.2].

Lemma 14 (Stability of the shock position). Let (A1) and (A2) hold and consider some $\bar{w} \in W$ satisfying (ND) and a nondegnerated shock point $\bar{x}$ of $y(\bar{t}, \cdot; \bar{w})$. Then there exist functions $Y_{f/v}$ constructed around the minimal and maximal characteristics through $(\bar{t}, \bar{x})$, an interval $(x, x_r) \ni \bar{x}$ and a Lipschitz continuous function

\[(12) \quad x_\ast : w \in B^W_\rho(\bar{w}) \mapsto x_\ast(w) \in (x, x_r)\]

such that for all $w \in B^W_\rho(\bar{w})$ with $\rho > 0$ small enough, $y(\bar{t}, \cdot; w)$ is given by

\[y(\bar{t}, x; w)|_{(x, x_r)} = Y_f(\bar{t}, x; w), \quad y(\bar{t}, x; w)|_{(x_r, x)} = Y_v(\bar{t}, x; w).\]

Proof. The first case where the extreme characteristics through $(\bar{t}, \bar{x})$ end in points where the initial or boundary data are smooth has been shown in [33, Lemma 6.3.1].

We now consider the second case where $\xi_-$ ends in the interior of a rarefaction wave produced in $(\bar{t}_j^{a/a}, a)$ and note that all remaining cases can be treated analogously.

We choose some $\bar{t} > \bar{t}_j^{a/a}$ satisfying $M_f(\bar{t} - \bar{t}_j^{a/a}) < \frac{\varepsilon}{2}$. Since $g$ is equal to zero for all $t \in [0, \varepsilon]$ and $x \in [a, a + \varepsilon] \cup [b - \varepsilon, b]$, defining $I = (a + f'(u_{j+1}^{B,a}))(\bar{t} - \bar{t}_j^{a/a} + \varepsilon, a + f'(u_{j+1}^{B,a}))(\bar{t} - \bar{t}_j^{a/a} - \varepsilon)$ yields

\[\forall \varepsilon > 0 \exists \rho > 0 : \quad y(\bar{t}, x; w)|_I = f^{\rho-1}\left(\frac{x - a}{\bar{t} - \bar{t}_j^{a/a}}\right) \quad \forall w \in B^W_\rho(\bar{w}).\]

Since the mapping $B^W_\rho(\bar{w}) \ni w \mapsto y(\bar{t}, \cdot; w) \in C^1(I)$ is continuously differentiable and $\xi_-(\bar{t}) \in I$ holds for $\varepsilon$ small enough, we can consider (1) on the truncated space-time cylinder $[\bar{t}, \bar{t}] \times \Omega$ with initial data $y(\bar{t}, \cdot, \bar{w})$ and proceed as in the first case. ⬜

In the next we will prove Fréchet-differentiability of the shock position (12) and derive an adjoint-based representation of the gradient. Formally, for a given solution $y$ of (1) and given end data $p^f$, the corresponding adjoint equation reads

\[(13) \quad p_t + f'(y)p_x = -g_p(\cdot, y, u_1)p \quad \text{on} \quad \Omega, \quad p(\bar{t}, \cdot) = p^f(\cdot) \quad \text{on} \quad \Omega, \quad p(t, l_{\sigma(t)}) = 0 \quad \text{on} \quad \{t \in (0, \bar{t}) : \sigma(t)f'(y(t, l_{\sigma(t)}; w)) \leq 0\}.\]

The adjoint state associated with (13) is defined as in [33, Definition 3.3]:

Definition 15. Consider a bounded function $p^f$ that is pointwise everywhere the limit of a sequence $(p^f_n)_{n \in \mathbb{N}} \subset C_0^{0,1}(\Omega)$ which is bounded in $C(\Omega) \cap W^{1,1}_{loc}(\Omega)$. Then the adjoint state associated with (13) is defined by the following requirements:

i) For all $x \in \Omega$ and all generalized backward characteristics $\xi$ of $y$ through $(\bar{t}, \bar{x})$ the function $p^\xi(t) := p(t, \xi(t))$ is given by the characteristic equation

\[p^\xi(t) = -g_p(t, \xi(t), y(t, \xi(t)), u_1(t, \xi(t)))p^\xi(t), \quad t \in (0, \bar{t}) : \xi(t) \in \Omega, \quad p^\xi(\bar{t}) = p^f(\bar{x}).\]

ii) $p(t, x) = 0$ for all $(t, x) \in D_\ast$. 

This manuscript is for review purposes only.
THEOREM 16 (Differentiability of shock points). Let (A1) and (A2) hold. Consider any \( \bar{\omega} \in W \) satisfying (ND) and a nondegenerated shock point \( \bar{x} \) of \( y(\bar{t}, \cdot; \bar{\omega}) \).

Then the mapping (12) is continuously differentiable with derivative (cf. Notations 1, 3)
\[
\frac{\partial}{\partial \omega} x_\alpha(w) \delta w = \langle p, g_u(\cdot, y, u_1) \delta u_1 \rangle_{2, \Omega} - \sum_{i \in I_{r,0}(w)} p_i^{r,0} \delta x_i^0 + \sum_{i=1}^{n_r+1} (p(0, \cdot), \delta u_i^0)_{2, I_i(w)}
\]
\[
+ \sum_{i \in I_{r,0}(w)} p(0, x_i^0)[u_0(x_i^0)] \delta x_i^0 + \sum_{l \in \{a,b\}} \sigma(l) \left( \sum_{i \in I_{r,1}(w), t_i^l \leq t} p_i^{r,l} \delta t_i^l + \sum_{i=1}^{n_{r,l}+1} (p(\cdot, t_{\sigma(l)}), f'(u_i^{B,l}) \delta u_i^{B,l})_{2, I_{b,l}(w) \cap [a,b]} + \sum_{i \in I_{r,1}(w), t_i^l \leq t} p(t_i^l, l_{\sigma(l)})[f(y(t_i^l, l_{\sigma(l)}; w))] \delta t_i^l \right)
\]
for \( \delta w \in W \), where \( p \) denotes the adjoint state with \( p(\bar{t}, \cdot) = \mathbb{1}_\delta(\cdot) \frac{1}{\|y(t, \bar{x}; \bar{\omega})\|} \) and
\[
(14a) \quad p_j^{r,0} := \int_{f'(u_j^0(x_j^0))}^{f'(u_j^0(x_j^0 + x_j^0))} \lim_{\varepsilon \to 0} p(t, z + x_j^0) \frac{1}{f''(f^{-1}(z))} \, dz, \quad j \in I_{r,0}(w),
\]
\[
(14b) \quad p_j^{r,l} := \int_{f'(u_j^{0,l}(t_j^l))}^{f'(u_j^{0,l}(t_j^l + t_j^l))} \lim_{\varepsilon \to 0} p(t, z(t + t_j^l)) \frac{z}{f''(f^{-1}(z))} \, dz, \quad j \in I_{r,l}(w), \quad l = a, b.
\]

Proof. This result is proved in Appendix A.

From Lemma 14 and Theorem 16, we can deduce the following result.

THEOREM 17. Let (A1) and (A2) hold and consider some \( \bar{\omega} \in W \) satisfying (ND).

Then there exist constants \( \rho, \varepsilon > 0 \) and continuously differentiable mappings
\[
Y_k : (x, w) \in I_k^{\varepsilon} \times B_p^W(\bar{\omega}) \mapsto Y_k(\bar{t}, x; w) \in \mathbb{R}
\]
\[
x_k : w \in B_p^W(\bar{\omega}) \mapsto x_k(w) \in (x_k(\bar{\omega}) - \varepsilon, x_k(\bar{\omega}) + \varepsilon),
\]
where \( I_k^{\varepsilon} := (x_k-1(\bar{\omega}) - \varepsilon, x_k(\bar{\omega}) + \varepsilon) \) for \( k = 1, \ldots, K + 1 \), and \( x_0 = a, x_{K+1} = b \).

Finally, for all \( w \in B_p^W(\bar{\omega}) \) it holds
\[
y(\bar{t}, x; w) |_{[a,b]} = Y_1(\bar{t}, x; w) \cdot \mathbb{1}_{[a,x_1(w)]}(x) + \sum_{k=2}^{K+1} Y_k(\bar{t}, x; w) \cdot \mathbb{1}_{[x_{k-1}(w), x_k(w)]}(x)
\]
and the mappings \( w \in B_p^W(\bar{\omega}) \mapsto Y_k(\bar{t}, \cdot; w) \in C(I_k^{\varepsilon}) \) are continuously differentiable.

Remark 18. The derivatives of the mappings \( x_k \) in (16) can be computed as follows: If \( x_k(w) \) is a discontinuity of \( y(\bar{t}, \cdot; \bar{\omega}) \), its derivative can be computed according to Theorem 16. If \( x_k(w) \) lies on the boundary of a rarefaction wave, the corresponding derivative can be computed by solving the linearized characteristic equation (3.36a)-(3.36c) in [33]. The mappings (15) are those described in Lemma 14. The corresponding derivatives can be computed according to [33, Lemma 6.2.1 and 6.2.7], where in the case that the minimal/maximal characteristics end in the inner of rarefaction centers, we consider the truncated IBVPs as described in the proof of Lemma 14.
Analogously to the proof of Theorem 16, one can extend [33, Theorem 5.2.6] to the case that shifts of rarefaction centers are allowed:

**Theorem 19** (Differentiability of the tracking-type functional). Let (A1) and (A2) hold and consider some \( \bar{w} \in W \) satisfying (ND). Further assume that \( y_d \) is approximately continuous in a neighborhood of the discontinuities of \( y(t, \cdot, \bar{w}) \) on \( [a, b] \).

Then the mapping \( w \in B^p_{\rho} (\bar{w}) \mapsto J(w) := J(y(w)) \in \mathbb{R} \) is for sufficiently small \( p \) continuously differentiable. Let \( p \) denote the adjoint state according to Definition 15 with end data

\[
\tilde{\psi}_p(x) := \int_0^1 \psi_p(y(\bar{t}, x; + \bar{w})), + \tau[y(\bar{t}, x; \bar{w})], y_d(x) + \tau[y_d(x)] \, \, \, d\tau.
\]

Using (14), the derivative of \( \bar{J}(w) \) in a direction \( \delta w \in W \) is given by

\[
\bar{J}'(w) \cdot \delta w = R'(w) \delta w + (p, g_{u_1}(\cdot, y, u_1) \delta u_1)_{2, \Omega} + \sum_{i=1}^{n+1} \left( 2 \sigma(\tau(i)) \left( \delta u_i \right) \right),
\]

where the last equality holds due to the continuity of \( \delta u_i \).

Analogously to the proof of Theorem 16, one can extend [33, Theorem 5.2.6] to the case that shifts of rarefaction centers are allowed:

**Theorem 20.** Let (A1) and (A2) hold and assume that there exists \( \bar{w} \in W_{ad} \) such that \( y(\bar{t}, x, \bar{w}) \leq \bar{y}(x) \) is satisfied for all \( x \in [a, b] \). Then (P) admits a globally optimal solution.

**3.1. Existence of globally optimal solutions.** In this section we will prove existence of an optimal solution and necessary optimality conditions for (P).

Proof. We show compactness of the set \( W_{ad} := \{ w \in W_{ad} : y(\bar{t}, x; w) \leq \bar{y}(x) \ \forall x \in [a, b] \} \). Since \( W_{ad} \) is by assumption non-empty, we can consider a sequence \( (w_n)_{n \in \mathbb{N}} \subseteq W_{ad} \subseteq W \). Due to the compactness of \( W_{ad} \) in \( W \), there exists a subsequence, again denoted by \( (w_n)_{n \in \mathbb{N}} \), converging to some \( \bar{w} \in W_{ad} \) w.r.t. \( || \cdot ||_W \). Corollary 7 implies that the sequence \( (y(\bar{t}, \cdot; w_n))_{n \in \mathbb{N}} \) converges in \( L^1 (a, b) \) y(\bar{t}, \cdot, \bar{w}) \) and hence there exists a subsequence, again denoted by \( (y(\bar{t}, \cdot; w_n))_{n \in \mathbb{N}} \), converging pointwise almost everywhere to \( y(\bar{t}, \cdot, \bar{w}) \) on \( [a, b] \). Therefore and since \( y(\bar{t}, x, w_n) \leq \bar{y}(x) \) holds for all \( x \in [a, b] \) and all \( n \in \mathbb{N} \), we obtain that

\[
(18) \quad y(\bar{t}, x, \bar{w}) \leq \bar{y}(x) \quad \text{for a.a. } x \in [a, b].
\]

In order to show that (18) holds for all \( x \in [a, b] \), let \( \hat{x} \in (a, b) \) be arbitrary. Due to (18), we can choose a sequence \( (x_n)_{n \in \mathbb{N}} \) with \( x_n \nrightarrow \hat{x} \) for \( n \to \infty \) such (18) holds for all \( x = x_n \). By Convention 5, we obtain

\[
y(\bar{t}, \hat{x}, \bar{w}) = \lim_{n \to \infty} y(\bar{t}, x_n; \bar{w}) \leq \lim_{n \to \infty} \bar{y}(x_n) = \bar{y}(\hat{x}),
\]

where the last equality holds due to the continuity of \( \bar{y}(\cdot) \). Since \( y(\bar{t}, a, \bar{w}) = y(\bar{t}, a+, \bar{w}) \)

by Convention 5, we can argue analogously with a sequence \( x_n \nrightarrow a \) satisfying (18)
and deduce \( y(\overline{t}, a, \overline{w}) \leq \overline{y}(a) \). Thus, (18) holds for all \( x \in [a, b] \) yielding that \( \overline{w} \in \hat{W}_{ad} \).

Therefore, \( \hat{W}_{ad} \) is compact. We now consider a sequence \( (w_n)_{n \in \mathbb{N}} \subseteq \hat{W}_{ad} \) satisfying

\[
\hat{J}(w_n) \to \inf_{w \in \hat{W}_{ad}} \hat{J}(w) \quad \text{for } k \to \infty.
\]

Since \( \hat{W}_{ad} \) is compact, there exists a convergent subsequence, again denoted by \( (w_n)_{n \in \mathbb{N}} \), with \( w_n \to \overline{w} \in \hat{W}_{ad} \). Since \( \hat{J} \) is Lipschitz-continuous w.r.t. \( w \) by Corollary 8, we obtain \( \hat{J}(\overline{w}) = \inf_{w \in \hat{W}_{ad}} \hat{J}(w) \) and hence, \( \overline{w} \) is a global minimum for (P).

**3.2. Reformulation of the state variable.** In order to derive necessary optimality conditions we need a constraint qualification which requires due to (2) that \( y(\overline{t}, \cdot; w) \) is an element of \( L^\infty([a, b]) \). Difficulties arise from the fact that the control-to-state-mapping is not continuous to \( L^\infty \). Theorem 17 yields that for \( w \in W \) satisfying (ND), \( y(\overline{t}, \cdot; w)|_{[a, b]} \) can be rewritten according to (17). Hence, we introduce \((y_1(w), \ldots, y_{K+1}(w), x_1(w), \ldots, x_K(w))\) as new state variables, where

\[
y_k(\lambda; w) := Y_k(\overline{t}, x_{k-1}(w) + \lambda(x_k(w) - x_{k-1}(w)); w), \quad \lambda \in [0, 1].
\]

\( Y_1, \ldots, Y_{K+1} \) and \( x_1, \ldots, x_K \) are given according to Theorem 17 and Remark 18. In terms of the new state variables the state constraints (2) read

\[
y_k(\lambda; w) \leq \overline{y}(\overline{t}, x_{k-1}(w) + \lambda(x_k(w) - x_{k-1}(w))) =: \overline{y}_k(\lambda; w), \lambda \in [0, 1].
\]

**Lemma 21.** Let (A1) and (A2) hold and consider some \( \overline{w} \in W \) satisfying (ND).

Then we have

\[
y(\overline{t}, \cdot; \overline{w}) \leq \overline{y}(\cdot) \text{ on } [a, b] \iff y_k(\cdot; \overline{w}) \leq \overline{y}_k(\cdot; \overline{w}) \text{ on } [0, 1] \quad \forall k = 1, \ldots, K + 1
\]

**Proof.** This is obvious by using (19), (20), the representation of \( y(\overline{t}, \cdot; \overline{w}) \) in (17) and Convention 5, since \( \overline{y} \) is continuous, \( y(\overline{t}, \cdot; \overline{w}) \) is continuous on \((x_k(w), x_{k+1}(w))\), \( k = 0, \ldots, K \), and admits one-sided traces. \( \square \)

**Theorem 22 (Continuous differentiability of the state).** Let (A1) and (A2) hold and consider some \( \overline{w} \in W \) satisfying (ND). Then the mapping

\[
w \in B^W_{\rho}(\overline{w}) \mapsto (y_1(w), \ldots, y_{K+1}(w), x_1(w), \ldots, x_K(w)) \in C([0, 1])^{K+1} \times \mathbb{R}^K
\]

is for \( \rho > 0 \) small enough well-defined and continuously differentiable, where

\[
\frac{\partial}{\partial w} y_k(\lambda; w) \delta w = \frac{\partial}{\partial w} Y_k(\overline{t}, \lambda x_k(w) + (1 - \lambda)x_{k-1}(w); w) \delta w
\]

\[
+ \frac{\partial}{\partial x} Y_k(\overline{t}, \lambda x_k(w) + (1 - \lambda)x_{k-1}(w); w) \cdot [\lambda \frac{\partial}{\partial w} x_k(w) + (1 - \lambda) \frac{\partial}{\partial w} x_{k-1}(w)] \delta w
\]

and the derivatives of \( x_k(\cdot) \) can be computed according to Remark 18. **Proof.** This theorem is a consequence of Theorem 17 and (19). \( \square \)

**3.3. First order necessary optimality conditions.** Our aim is to derive necessary optimality conditions for (P). To this end, we consider a general problem of the form

\[
imin_{z \in \mathcal{Z}} f(z) \text{ subject to } G(z) \in \mathcal{K}, \quad z \in \mathcal{C}
\]

and recall the following result.
THEOREM 23 (Karush-Kuhn-Tucker conditions). Consider a local solution $\bar{z}$ of (23) and assume that the mappings $f : Z \to \mathbb{R}$ and $G : Z \to V$ are continuously differentiable in $\bar{z} \in Z$ with Banach spaces $Z$ and $V$. We further assume that $C \subset Z$ is closed, convex and non-empty and that Robinson’s CQ is satisfied at $\bar{z}$. Hence, Theorem 23 and Lemma 24 yield the statement of the above theorem.

Proof. This result can be found e.g. in [20,42].

Using the new state variables (21), we can write (P) in the form (23) by setting

$\bar{z} = w$, $V = C ([0,1])^{K+1} \times \mathbb{R}^K$.

$G_k (w) = y_k (\lambda; w) - \bar{y}_k (\lambda; w)$, $k = 1, \ldots, K + 1$,

$G_k (w) = x_k (w)$, $k = 1, \ldots, K$,

$K = C ([0,1], (-\infty, 0])^{K+1} \times \mathbb{R}^K$, and $C = W_{ad}$.

Hereby, $V^*$ is due to Riesz-Radon theorem given by $V^* = M ([0,1])^{K+1} \times \mathbb{R}^K$, where $M ([0,1])$ denotes the space of bounded Radon measures on $[0,1]$ (see e.g. [4]).

LEMMA 24. The polar cone of $K$ can be characterized as follows:

$q \in K^\circ \iff q = (\mu_1, \ldots, \mu_{K+1}, 0, \ldots, 0)$,

where $\mu_1, \ldots, \mu_{K+1} \in M ([0,1])$ are nonnegative.

We are now able to derive first order necessary optimality conditions for (P):

THEOREM 25. Let (A1) and (A2) hold and consider a local solution $\bar{w} \in W_{ad}$ for (P) satisfying (ND) and (24). Then there exist nonnegative $\mu_1, \ldots, \mu_{K+1} \in M ([0,1])$ such that

$\bar{y}_k (\cdot, \bar{w}) \leq \bar{y}_k (\cdot, \bar{w})$ on $[0,1]$ $\forall k = 1, \ldots, K + 1$

$\sum_{k=1}^{K+1} \int_{[0,1]} (y_k (\lambda, \bar{w}) - \bar{y}_k (\lambda, \bar{w})) d\mu_k (\lambda) = 0$

$J^*(\bar{w}) (w - \bar{w}) + \int_{[0,1]} \frac{\partial}{\partial w} (y_k (\lambda, \bar{w}) - \bar{y}_k (\lambda, \bar{w})) (w - \bar{w}) d\mu_k (\lambda) \geq 0 \forall w \in W_{ad}$.

Proof. Setting (25)-(28), we observe that $\bar{w}$ is due to Lemma 21 a locally optimal solution for (23). Due to Theorem 22 and (A2) the assumptions of Theorem 23 are satisfied. Hence, Theorem 23 and Lemma 24 yield the statement of the above theorem.

LEMMA 26. Let (A1) and (A2) hold and consider some $\bar{w} \in W_{ad}$ satisfying (ND).

Then (24) is satisfied if and only if there exists $\bar{w} \in W_{ad}$ such that for a constant $\varepsilon > 0$ the following assertion is valid for all $k = 1, \ldots, K + 1$:

$y_k (\lambda, \bar{w}) - \bar{y}_k (\lambda, \bar{w}) + \frac{\partial}{\partial w} (y_k (\lambda, \bar{w}) - \bar{y}_k (\lambda, \bar{w})) (\bar{w} - w) \leq -\varepsilon \forall \lambda \in [0,1]$.

This manuscript is for review purposes only.
Proof. From (24) we can directly deduce (33) and vice versa.

Now we will reformulate the optimality conditions (30), (31) and (32) in terms of the original state \( y(\bar{t}, \cdots; \bar{w}) \). As a first step, we rewrite (30), (31) and (32) in terms of the mappings \( Y_1, \ldots, Y_{K+1} \) which were introduced in Theorem 17. In (19) we have introduced the new state variables by using the variable transformations

\[
\varphi_{k;\bar{w}} : [0, 1] \rightarrow [x_{k-1}(\bar{w}), x_k(\bar{w})], \quad \lambda \mapsto x_{k-1}(\bar{w}) + \lambda (x_k(\bar{w}) - x_{k-1}(\bar{w})),
\]

\[
\varphi_{k;\bar{w}}^{-1} : [x_{k-1}(\bar{w}), x_k(\bar{w})] \rightarrow [0, 1], \quad x \mapsto \frac{x - x_{k-1}(\bar{w})}{x_k(\bar{w}) - x_{k-1}(\bar{w})}.
\]

Using (34) and (22), the optimality conditions in (30)-(32) can be written as follows

\[
Y_k(\bar{t}, \varphi_{k;\bar{w}}(\cdot), \bar{w}) \leq \bar{y}(\varphi_{k;\bar{w}}(\cdot)) \quad \text{on} \ [0, 1] \ \forall k = 1, \ldots, K + 1
\]

\[
\sum_{k=1}^{K+1} \int_{[0,1]} (Y_k(\bar{t}, \varphi_{k;\bar{w}}(\lambda), \bar{w}) - \bar{y}(\varphi_{k;\bar{w}}(\lambda))) d\mu_\lambda(\lambda) = 0
\]

\[
j'(\bar{w})(w - \bar{w}) + \sum_{k=1}^{K+1} \int_{[0,1]} \left[ \frac{\partial}{\partial x} (Y_k(\bar{t}, \varphi_{k;\bar{w}}(\lambda); \bar{w}) - \bar{y}(\varphi_{k;\bar{w}}(\lambda)))
\right.
\]

\[
\cdot \left( \varphi_{k;\bar{w}}^{-1}(\varphi_{k;\bar{w}}(\lambda)) \frac{\partial}{\partial \bar{w}} x_k(\bar{w})(w - \bar{w}) + (1 - \varphi_{k;\bar{w}}^{-1}(\varphi_{k;\bar{w}}(\lambda))) \frac{\partial}{\partial \bar{w}} x_{k-1}(\bar{w})(w - \bar{w}) \right)
\]

\[
+ \frac{\partial}{\partial \bar{w}} (Y_k(\bar{t}, \varphi_{k;\bar{w}}(\lambda); \bar{w}) - \bar{y}(\varphi_{k;\bar{w}}(\lambda))) \cdot (w - \bar{w}) \right] d\mu_\lambda(\lambda) \geq 0 \ \forall \ w \in W_{ad}.
\]

Using [12, V, §3, (3.1)], one can show that there exist nonnegative measures \( \mu_k \in \mathcal{M}(I_k) \), where \( I_k := [x_{k-1}(\bar{w}), x_k(\bar{w})] \), which are given by \( \mu_k(A) := \mu_k(\varphi_{k;\bar{w}}^{-1}(A)) \) for all measurable \( A \subset I_k \) and all \( k = 1, \ldots, K + 1 \) such that the following holds.

\[
Y_k(\bar{t}, \cdot, \bar{w}) \leq \bar{y}(\cdot) \quad \text{on} \ I_k \ \forall k = 1, \ldots, K + 1
\]

\[
\sum_{k=1}^{K+1} \int_{I_k} (Y_k(\bar{t}, x, \bar{w})) - \bar{y}(x) d\bar{\mu}_k(x) = 0
\]

\[
j'(\bar{w})(w - \bar{w}) + \sum_{k=1}^{K+1} \int_{I_k} \left[ \frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \frac{x - x_{k-1}(\bar{w})}{x_k(\bar{w}) - x_{k-1}(\bar{w})} d\bar{\mu}_k(x)
\right.
\]

\[
\cdot \frac{\partial}{\partial \bar{w}} x_k(\bar{w})(w - \bar{w}) + \int_{I_k} \frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \frac{x_{k}(\bar{w}) - x}{x_k(\bar{w}) - x_{k-1}(\bar{w})} d\bar{\mu}_k(x)
\]

\[
\cdot \int_{I_k} \frac{\partial}{\partial \bar{w}} Y_k(\bar{t}, x, \bar{w})(w - \bar{w}) d\bar{\mu}_k(x) \geq 0 \ \forall w \in W_{ad}.
\]

In connection with Theorem 25, we then obtain the following result.

**Theorem 27.** Let (A1) and (A2) hold and consider a locally optimal solution \( \bar{w} \in W_{ad} \) of (P) satisfying (ND) and (24). Then there exist \( K + 1 \) nonnegative measures \( \bar{\mu}_k \in \mathcal{M}(I_k) \) such that (36)-(38) hold.

To simplify (38), we use the following observation.

**Lemma 28.** Let (A1) and (A2) hold and consider some \( \bar{w} \in W_{ad} \) and nonnegative measures \( \mu_k \in \mathcal{M}(I_k), \ k = 1, \ldots, K + 1 \) such that (36) and (37) are satisfied. Then
for all measurable $A \subseteq (x_{k-1}(\bar{w}), x_k(\bar{w}))$ the following holds

\begin{equation}
\int_A \frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \left( \frac{x - x_{k-1}(\bar{w})}{x_k(\bar{w}) - x_{k-1}(\bar{w})} \right) d\bar{\mu}_k(x) = 0,
\end{equation}

\begin{equation}
\int_A \frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \left( \frac{x_k(\bar{w}) - x}{x_k(\bar{w}) - x_{k-1}(\bar{w})} \right) d\bar{\mu}_k(x) = 0.
\end{equation}

\begin{proof}
For arbitrary $k \in \{1, \ldots, K+1\}$ and measurable $A \subseteq (x_{k-1}(\bar{w}), x_k(\bar{w}))$ we define $A_1 := \{ x \in A : Y_k(\bar{t}, x, \bar{w}) < \bar{y}(x) \}$, $A_2 := \{ x \in A : Y_k(\bar{t}, x, \bar{w}) = \bar{y}(x) \}$ and observe that $A_1$ and $A_2$ are both measurable due to the regularity of $Y_k(\bar{t}, \cdot; \bar{w})$ and $\bar{y}(\cdot)$. From (36) we can deduce that $A = A_1 \cup A_2$. We firstly prove that

\begin{equation}
\bar{\mu}_k(A_1) = 0,
\end{equation}

\begin{equation}
\frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \bigg|_{A_2} = 0.
\end{equation}

To prove the first assertion, we suppose that $\bar{\mu}_k(A_1) \neq 0$. We observe that due to the nonnegativity of $\bar{\mu}_k$, it holds that $\bar{\mu}_k(A_1) > 0$. Then we obtain by (36)

\begin{equation}
\sum_{k=1}^{K+1} \int_{I_k} (Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)) d\bar{\mu}_k(x) \leq \int_{I_k \cap A_1} (Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)) d\bar{\mu}_k(x) < 0.
\end{equation}

This is a contradiction to (37) and hence the first assertion is proved.

To prove the second assertion, we note that by (36) the set $A_2$ consists of global maxima of the differentiable function $Y_k(\bar{t}, \cdot; \bar{w}) - \bar{y}$ on the open set $(x_{k-1}(\bar{w}), x_k(\bar{w}))$. Hence, $\frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \bigg|_{A_2} = 0$. This shows the second assertion.

Recalling $A = A_1 \cup A_2$, and using (41), we obtain

\begin{equation}
\int_A \frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \left( \frac{x - x_{k-1}(\bar{w})}{x_k(\bar{w}) - x_{k-1}(\bar{w})} \right) d\bar{\mu}_k(x) = 0.
\end{equation}

Analogously, we can prove (40).

\begin{flushright}
\small\textbf{\textbullet \ }
\end{flushright}

Using this result, we can further simplify the optimality conditions in Theorem 27:

\begin{corollary}
Let (A1) and (A2) hold and consider some $\bar{w} \in W_{ad}$ and nonnegative measures $\bar{\mu}_k \in \mathcal{M}(I_k)$, $k = 1, \ldots, K+1$, satisfying (36), (37) and (38). Then (38) can also be written as:

\begin{equation}
\begin{aligned}
J'(\bar{w})(w - \bar{w}) + \sum_{k=1}^{K+1} & \left( \frac{\partial}{\partial x} [Y_k(\bar{t}, x_k(w), \bar{w}) - \bar{y}(x_k(\bar{w}))] \cdot \bar{\mu}_k(\{x_k(\bar{w})\}) \\
\cdot \frac{\partial}{\partial w} x_k(\bar{w})(w - \bar{w}) + & \frac{\partial}{\partial x} [Y_k(\bar{t}, x_{k-1}(\bar{w}), \bar{w}) - \bar{y}(x_{k-1}(\bar{w}))] \cdot \bar{\mu}_k(\{x_{k-1}(\bar{w})\}) \\
\cdot & \frac{\partial}{\partial w} x_{k-1}(\bar{w})(w - \bar{w}) + \int_{I_k} \frac{\partial}{\partial w} Y_k(\bar{t}, x, \bar{w})(w - \bar{w}) d\bar{\mu}_k(x) \gtrless 0, \forall w \in W_{ad}
\end{aligned}
\end{equation}

Using the representation of $y$ in (17) and Convention 5, we can formulate the optimality conditions from Theorem 27 also in terms of the original state $y(\bar{t}, \cdot; \bar{w})$:

\begin{theorem}
Let (A1) and (A2) hold and let $\bar{w} \in W_{ad}$ be a local solution of (P) satisfying (ND) and (24). Then (36)-(38) and (42) are still valid, if we replace $Y_k(\bar{t}, x; \bar{w})$, $\frac{\partial}{\partial w} Y_k(\bar{t}, x; \bar{w})$ and $\frac{\partial}{\partial x} y(\bar{t}, x; \bar{w})$ by $y(\bar{t}, x; \bar{w})$, $\frac{\partial}{\partial w} y(\bar{t}, x; \bar{w})$ and $\frac{\partial}{\partial x} y(\bar{t}, x; \bar{w})$, respectively.
\end{theorem}

This manuscript is for review purposes only.
Lemma 31. Let (A1) and (A2) hold and consider some \( \tilde{w} \in W_{ad} \) satisfying (ND). Then (24) is satisfied if and only if there exists a constant \( \varepsilon > 0 \) and \( \delta w \in W_{ad} - \tilde{w} \) such that for all \( x \in I_k \) and \( k = 1, \ldots, K + 1 \) it holds that

\[
Y_k(\tilde{t}, x, \tilde{w}) - \tilde{y}(x) + \frac{\partial}{\partial w} Y_k(\tilde{t}, x, \tilde{w}) \delta w + \frac{\partial}{\partial x} Y_k(\tilde{t}, x, \tilde{w}) \delta w - \tilde{y}(x) \\
\left( x - x_{k-1}(\tilde{w}) \right) \frac{\partial}{\partial w} x_k(\tilde{w}) \delta w + \frac{x_k(\tilde{w}) - x}{x_k(\tilde{w}) - x_{k-1}(\tilde{w})} \frac{\partial}{\partial w} x_{k-1}(\tilde{w}) \delta w \leq -\varepsilon
\]

Proof. This result follows directly from Lemma 26 and (19).

Theorem 32. Let (A1) and (A2) hold and consider some \( \tilde{w} \in W_{ad} \) satisfying (ND), (2) and assume that \( \tilde{w} + \delta w \in W_{ad} \) holds for all \( \delta w \in B_{w}^{\tilde{w}}(\tilde{w}) \cap \{ w \in W : w \leq 0 \} \) \((w \leq 0 \) has to be understood component-wise), if \( \rho > 0 \) is small enough. We further assume that \( g|_{\Omega_1} = 0 \), the upper bound \( \tilde{y}(\cdot) \) in (2) is a constant and \( a \) and \( b \) are points of continuity of \( y(\cdot, \cdot; \tilde{w}) \). Then (24) is satisfied in \( \tilde{w} \).

Proof. We will prove Theorem 32 by applying Lemma 31. More precisely, we show that there exist \( \varepsilon > 0 \) and \( \delta w \in W_{ad} - \tilde{w} \) such that for all \( x \in [x_{k-1}(\tilde{w}), x_k(\tilde{w})] \) and \( k = 1, \ldots, K + 1 \) (43) holds. By assumption, \( x_k(\tilde{w}) \in [a, b] \) with \( k = \{ 1, \ldots, K + 1 \} \) is either a discontinuity of \( y(\cdot, \cdot; \tilde{w}) \) or lies on the boundary of a rarefaction wave. Moreover, all discontinuities of \( y(\cdot, \cdot; \tilde{w}) \) on \( [a, b] \) are nondegenerated according to Definition 12. Therefore, we can w.l.o.g. restrict ourselves to the case \( K = 2 \), in which we can discuss all relevant cases that can occur. For the case \( K > 2 \), the subsequent procedure can just be continued.

We assume that \((\tilde{t}, x_1(\tilde{w}))\) lies on the right boundary of a rarefaction wave being created in a discontinuity \( t^0 \) of the left boundary data \( u_{B,a}(:; \tilde{w}) \). In addition, let the genuine backward characteristic through \((\tilde{t}, a)\) also end in \((t^0, a)\). We further assume that \((\tilde{t}, x_2(\tilde{w}))\) is a nondegenerated shock, where the minimal backward characteristic \( \xi_- \) through \((\tilde{t}, x_2(\tilde{w}))\) ends in a continuity point \( t^a \in (t^0_{-1}, t^0_1) \) of \( u_{B,a}(:; \tilde{w}) \) and the maximal backward characteristic \( \xi_+ \) ends in a continuity point \( \tilde{x} \in (\tilde{x}^{0}_{m-1}, \tilde{x}^{0}_m) \) of the initial data \( u_0(:; \tilde{w}) \). Finally, let the genuine backward characteristic through \((\tilde{t}, b)\) also end within the interval \((\tilde{x}^{0}_{m-1}, \tilde{x}^{0}_m)\).

The proof consists of two steps: In Step 1, we will derive representations for the terms \( \frac{\partial}{\partial x} Y_1(\tilde{t}, x, \tilde{w}) \delta x \) and \( \frac{\partial}{\partial w} Y_1(\tilde{t}, x, \tilde{w}) \delta w \) in (43) for all \( k = 1,2,3 \).

**Step 1:** Due to \( g|_{\Omega_1} = 0 \) it holds that \( Y_1(\tilde{t}, x, \tilde{w}) = f^{-1}(\tilde{x}/t^0_{-1}) \) yielding

\[
\frac{\partial}{\partial x} Y_1(\tilde{t}, x, \tilde{w}) \delta x = \frac{\delta x}{(t - t^0_{-1}) \cdot f''(f^{-1}(\tilde{x}/t^0_{-1}))},
\]

\[
\frac{\partial}{\partial w} Y_1(\tilde{t}, x, \tilde{w}) \delta w = \frac{(x - a) \cdot \delta t^0_{-1}}{(t - t^0_{-1})^2 f''(f^{-1}(\tilde{x}/t^0_{-1}))}.
\]

Since the minimal backward characteristic through \((\tilde{t}, x_2(\tilde{w}))\) ends in a continuity point \( t^a \in (t^0_{-1}, t^0_1) \) of \( u_{B,a}(:; \tilde{w}) \) and \( \tilde{w} \in W_{ad} \) satisfies (ND), [33, Lemma 6.2.7] yields

\[
\frac{\partial}{\partial x} Y_2(\tilde{t}, x, \tilde{w}) \delta x = \frac{(\tilde{u}^{B,a}_{j}(\Phi(\cdot))) \cdot \delta x}{f''((\tilde{u}^{B,a}_{j}(\Phi(\cdot))) (\Phi(\cdot)) (\tilde{t} - \Phi(\cdot)) - f'((\tilde{u}^{B,a}_{j}(\Phi(\cdot))))},
\]

\[
\frac{\partial}{\partial w} Y_2(\tilde{t}, x, \tilde{w}) \delta w = \frac{(\tilde{u}^{B,a}_{j}(\Phi(\cdot))) \cdot \delta x}{f''((\tilde{u}^{B,a}_{j}(\Phi(\cdot))) (\Phi(\cdot)) (\tilde{t} - \Phi(\cdot)) - f'((\tilde{u}^{B,a}_{j}(\Phi(\cdot))))}.
\]
where \( \cdot = (\bar{t}, x, \bar{w}) \), \((\bar{u}_j^{B,a})'()\) denotes the derivative of \( \bar{u}_j^{B,a}(\cdot) \) and \( \Phi(\bar{t}, x, \bar{w}) \) is the unique solution of the equation \( x = f'((\bar{u}_j^{B,a}(\phi))(\bar{t} - \phi) + a \) w.r.t. \( \phi \). We observe that
\[
\Phi(\bar{t}, x, \bar{w}) \in [\bar{\rho}, \bar{\rho}^2] \quad \text{for all } x \in [x_1(\bar{w}), x_2(\bar{w})].
\]

From \([33, \text{Lemma } 6.2.7 \ (i)]\), we can further deduce that there exists constants \( \delta_0, \beta > 0 \) such that for all \( x \in (x_1(\bar{w}) - \delta_0, x_2(\bar{w}) + \delta_0) \) the following is satisfied:
\[
q_1(\cdot) := f''((\bar{u}_j^{B,a}(\Phi(\cdot)))((\bar{u}_j^{B,a}(\Phi(\cdot))(\bar{t} - \Phi(\cdot)) - f'((\bar{u}_j^{B,a}(\Phi(\cdot))) < -\beta
\]

Now, we will have a closer look at the term \( Y_3 \). According to \([33, \text{Lemma } 6.2.1]\), since the maximal backward characteristic through \((\bar{t}, x_2(\bar{w})) \) ends in a point \( \bar{x} \in (\bar{x}_m^{01}, \bar{x}_m^{02}) \) where the initial data \( u_0(\cdot, \bar{w}) \) are smooth, it holds that
\[
\frac{\partial}{\partial x} Y_3(\bar{t}, x, \bar{w}) \delta x = \frac{(\bar{u}_m^{0})'(Z(\cdot)) \cdot \delta x}{f''(\bar{u}_m^{0}(Z(\cdot)))((\bar{u}_m^{0})'(Z(\cdot)))t + 1},
\]
\[
\frac{\partial}{\partial w} Y_3(\bar{t}, x, \bar{w}) \delta w = \frac{\delta u_m^{0}(Z(\cdot))}{f''(\bar{u}_m^{0}(Z(\cdot)))((\bar{u}_m^{0})'(Z(\cdot)))t + 1},
\]
where \( Z(\cdot) = Z(\bar{t}, x, \bar{w}) \) denotes the unique solution of the equation \( x = f'(\bar{u}_m^{0}(z))\bar{t} + z \) w.r.t. \( z \). Then for (a possibly smaller) \( \delta_0 > 0 \), we can deduce from \([33, \text{Lemma } 6.2.1 \ (i)]\) that for all \( x \in (x_2(\bar{w}) - \delta_0, b) \) it holds that
\[
q_2(\cdot) := f''((\bar{u}_m^{0}(Z(\cdot)))((\bar{u}_m^{0})'(Z(\cdot)))\bar{t} + 1 > \beta
\]
for a (possibly smaller) positive constant \( \beta > 0 \).

Next, we examine the term \( x_1(\bar{w}) \). Since \((\bar{t}, x_1(\bar{w})) \) lies on the right boundary of a rarefaction wave and the source term \( g \) is by assumption equal to zero, we obtain that
\[
x_1(\bar{w}) = f'((\bar{u}_j^{B,a}(\bar{t}_j^0))) \cdot (\bar{t} - \bar{t}_j^0) + a \quad \text{and hence}
\]
\[
\frac{\partial}{\partial w} x_1(\bar{w}) \delta w = f''((\bar{u}_j^{B,a}(\bar{t}_j^0))(\bar{t} - \bar{t}_j^0)[\delta u_j^{B,a}((\bar{t}_j^0)) + (\bar{u}_j^{B,a}(\bar{t}_j^0))\delta t_j^0] - f'((\bar{u}_j^{B,a}(\bar{t}_j^0))\delta t_j^0.
\]

The derivative of the shock position \( x_2(\bar{w}) \) w.r.t. \( w \) is due to \textbf{Theorem 16} given by
\[
\frac{\partial}{\partial w} x_2(\bar{w}) \delta w = \sum_{k=1}^{m} p(0, \cdot) \delta u_k^{0}/2, t_k^0 \cap [0, \bar{x}] + \sum_{j=1}^{j} (p(\cdot, a), f'((\bar{u}_k^{B,a})\delta u_k^{B,a}(2, t_k^a \cap [0, \bar{x}] [\delta x_k + \sum_{k \in I_n, \omega}(\bar{u}_0^{0}(\bar{t}_k^0) = \bar{u}_k^{0}(\bar{t}_k^0)) \delta t_k^a]
\]
\[
- \sum_{k \in I_n, \omega}(p(0, \bar{t}_k^0) = \bar{u}_k^{0}(\bar{t}_k^0) = \bar{u}_k^{0}(\bar{t}_k^0)) \delta x_k + \sum_{k \in I_n, \omega}(p(\bar{t}_k^0, a) = f(y(\bar{t}_k^0, a; \bar{w})) \delta t_k^a,
\]
where \( p \) is the adjoint state with end data \( p(\bar{t}, \cdot) = 1_{x_2(\bar{w})}(\cdot) \) and is equal to zero on \( \Omega_t \setminus D \). \( \Omega_t \setminus D \) denotes the domain \( \{t-x \in \Omega_t \setminus D\} \), and equal to zero on \( \Omega_t \setminus D \).

\textbf{Step 2:} Our goal is to choose \( \delta w \) such that (43) is satisfied for all \( k \in \{1, 2, 3\} \).

To this end, we choose all components of \( \delta w \), except \( \delta u_j^{B,a} \) and \( \delta u_m^{0} \), equal to zero.
Let \( \delta u^B_{j,a} \) and \( \delta u^0_m \) be given by

\[
\begin{align*}
\phi_1(t) & \quad \text{if } \bar{t}_j - 1 \leq t < \bar{t}_a - \rho_1, \\
\phi_2(t) & \quad \text{if } \bar{t}_a - \rho_1 \leq t < \bar{t}_a, \\
\phi_3(t) & \quad \text{if } \bar{t}_a - \rho_1 \leq t < \bar{t}_a.
\end{align*}
\]

(55) \( \delta u^B_{j,a}(t) = \begin{cases} 0 & \text{if } x < \bar{x} - \rho_3, \\ -\varepsilon_0 & \text{if } \bar{x} - \rho_3 \leq x < \bar{x}, \\ \phi_3(x) & \text{if } \bar{x} - \rho_3 \leq x < \bar{x}, \\ \phi_3(x) & \text{if } x \leq \bar{x}, \\ \frac{\partial \bar{x} u^B_{j,a}(t)}{\partial t^N} & \text{if } t = \bar{t}_j,
\end{cases} \]

with positive constants \( \varepsilon_0, \rho_1, \rho_2 \) and \( \rho_3 \). Moreover, let \( N \in \mathbb{N} \) be chosen such that \( f''(u^B,\alpha((\bar{t}_j))) (\bar{t} - \bar{t}_j) \cdot N > 1 \) holds. The constants \( \rho_1, \rho_2 \) and \( \rho_3 \) will be identified later.

Finally, let \( \phi_1, \phi_2 \) and \( \phi_3 \) be chosen such that \( -\varepsilon_0 \leq \phi \leq 0 \).

We consider (43) and start with \( k = 1 \). We note that since \( f''(\bar{x}) \geq M_f' \), the function \( Y_1(\bar{t}, \cdot; \bar{w}) \) is strictly increasing and hence, the only point where it may touch the upper bound on the interval \([a, x_1(\bar{w})]\) is at \( x_1(\bar{w}) \). From (53) and (55) we obtain that \( \frac{\partial \bar{x} u^B_{j,a}(\bar{w})}{\partial x} \delta w = \frac{-\varepsilon_0}{N} (2), (44), (45) \), the fact that \( \frac{\partial \bar{x} u^B_{j,a}(\bar{w})}{\partial x} \delta w = \frac{-\varepsilon_0}{N} \) and the choice of \( \delta w \) yield that the left term of (43) is at \( x = x_1(\bar{w}) \) bounded from above by

\[
\begin{align*}
\frac{\partial \bar{x} Y_1(\bar{t}, x_1(\bar{w}), \bar{w})}{\partial x} \frac{\partial \bar{x} u^B_{j,a}(\bar{w})}{\partial x} \delta w &= -\varepsilon_0 \\
&= -\varepsilon_0 \cdot \frac{-\varepsilon_0}{N M_f'} = -\varepsilon_{11},
\end{align*}
\]

where the inequality holds due to \( 0 < m_{f''} \leq f''(\nu') \leq M_{f''} \). (56) and the continuity of the left term of (43) w.r.t. \( x \) yield that there exists a constant \( \delta > 0 \) such that the left term of (43) is smaller than \( -\varepsilon_{12} \) for all \( x \in (x_1 - \delta_1, x_1] \). Since \( Y_1(\bar{t}, \cdot; \bar{w}) \) is strictly monotone increasing, for some \( \varepsilon_{12} > 0 \) it holds that

\[
Y_1(\bar{t}, x; \bar{w}) - \bar{y} \leq -\varepsilon_{12} \quad \text{for all } x \in [a, x_1(\bar{w}) - \delta_1].
\]

Hence, due to (44), (55) and the fact that \( \frac{\partial \bar{x} u^B_{j,a}(\bar{w})}{\partial x} \delta w = \frac{-\varepsilon_0}{N} \), the left term of (43) is on \([a, x_1(\bar{w}) - \delta_1] \) bounded from above by

\[
\frac{-\varepsilon_{12}}{N} \cdot \frac{\partial \bar{x} Y_1(\bar{t}, x_1(\bar{w}), \bar{w})}{\partial x} \frac{\partial \bar{x} u^B_{j,a}(\bar{w})}{\partial x} \delta w = -\varepsilon_{12}.
\]

Choosing \( \varepsilon = \varepsilon_1 := \min\{\frac{\varepsilon_1}{2}, \varepsilon_{12}\} \) yields that (43) is satisfied for \( k = 1 \).

Considering \( k = 2 \), (2) yields that the left term of (43) is bounded from above by

\[
\frac{\partial \bar{x} Y_2(\bar{t}, x, \bar{w})}{\partial x} \frac{\partial \bar{x} u^B_{j,a}(\bar{w})}{\partial x} \delta w + \frac{\partial \bar{x} Y_2(\bar{t}, x, \bar{w})}{\partial x} = \frac{x - a}{x_1(\bar{w}) - x_1(\bar{w})} \delta w \leq -\varepsilon_{12}.
\]

Inserting (46), (47) and (53) in (59), we obtain that (59) is in \( x = x_1(\bar{w}) \) equal to

\[
\text{Hence, the continuity of (59) w.r.t. } x \text{ yields that (59) is on } [x_1(\bar{w}), x_1(\bar{w}) + \delta] \text{ bounded from above by }
\]

\[
\frac{\partial \bar{x} u^B_{j,a}(\bar{w})}{\partial x} = \frac{\partial \bar{x} Y_1(\bar{t}, x_1(\bar{w}), \bar{w})}{\partial x} \frac{\partial \bar{x} u^B_{j,a}(\bar{w})}{\partial x} \delta w \leq -\varepsilon_{21} < 0, \text{ where } \delta > 0 \text{ is sufficiently small. We choose } \rho_2 \text{ such that } \Phi(x_1(\bar{w}) + \delta) = \bar{t}_j - \rho_2. \text{ We further choose}
\]

\[
\rho_1 = \frac{[y(\bar{t}, x_2(\bar{w}) ; \bar{w})]}{2N f'([\bar{\bar{u}}]^{B,a}(\cdot)) ||_{\infty,[0,\bar{\bar{w}}]}} > 0 \quad \text{and} \quad \rho_3 = \frac{[y(\bar{t}, x_2(\bar{w}) ; \bar{w})]}{2N} > 0,
\]

This manuscript is for review purposes only.
where $N$ has to be chosen sufficiently large, such that $\rho_1$ and $\rho_2$ are sufficiently small so that (55) makes sense. Then (54), (55) and $\|\phi_i(\cdot)\|_\infty \leq \varepsilon_0$ yield that $|\frac{\partial}{\partial w} x_2(\bar{w})\delta w| \leq \frac{\varepsilon_0}{N}$.

Using (47), (49) and (55), we obtain

$$\frac{\partial}{\partial w} Y_2(t, \bar{w})\delta w \leq \frac{-\alpha \varepsilon_0}{\|q_1(\cdot)\|_\infty, [x_1(\bar{w}) + \delta, x_2(\bar{w})]} \quad \text{on} \quad [x_1(\bar{w}) + \delta, x_2(\bar{w})],$$

where $\alpha$ is the constant in (A2). Due to the boundedness of $\frac{\partial}{\partial w} Y_2(t, x, \bar{w})$ and the fact that $|\frac{\partial}{\partial w} x_1(\bar{w})\delta w|, |\frac{\partial}{\partial w} x_2(\bar{w})\delta w| \leq \frac{\varepsilon_0}{N}$, one can choose $N$ large enough such that (59) is on $[x_1(\bar{w}) + \delta, x_2(\bar{w})]$ bounded from above by $-\varepsilon_{2,2} := 2\|q_1(t, \bar{w})\|_\infty, [x_1(\bar{w}) + \delta, x_2(\bar{w})]$. Choosing $\varepsilon = \varepsilon_2 := \min\{\varepsilon_{21}, \varepsilon_{22}\}$, (43) is satisfied for $k = 2$.

Finally, we consider the case $k = 3$. Using (51), (52) and (55), we obtain

$$\frac{\partial}{\partial w} Y_3(t, \cdot, \bar{w})\delta w \leq \frac{-\varepsilon_0}{\|q_2(t, \cdot, \bar{w})\|_\infty, [x_2(\bar{w}), b]} \quad \text{on} \quad [x_2(\bar{w}), b].$$

Due to the boundedness of $\frac{\partial}{\partial x} Y_3(t, x, \bar{w})$ and the fact that $|\frac{\partial}{\partial w} x_2(\bar{w})\delta w| \leq \frac{\varepsilon_0}{2}$, one can choose $N$ large enough such that (2) and (60) yield that the left term of (43) is on $[x_2(\bar{w}), b]$ smaller than $-\varepsilon_{2,2} := -\frac{\varepsilon_0}{2}\|q_2(t, \cdot, \bar{w})\|_\infty, [x_2(\bar{w}), b]$. Hereby, $\varepsilon_3$ is finite since $q_2((t, \cdot, \bar{w})$ is continuous on the compact interval $[x_2(\bar{w}), b]$. Hence, (43) is satisfied for $k = 3$ if we choose $\varepsilon = \varepsilon_3$. Choosing $\varepsilon := \varepsilon_{21}$, (43) is satisfied for $k = 1, 2, 3$ for the choice of $\delta w$ in (55). This also remains valid (with $\varepsilon > 0$), if we replace $\delta w$ by $\tau \delta w$ with a sufficiently small constant $\tau > 0$, such that $\bar{w} + \tau \delta w \in W$. More precisely: If we replace $\delta w$ by $\tau \delta w$, (56) holds with $-\tau \varepsilon_{11}$ and (58) remains valid for the same $\varepsilon_{12}$ such that choosing $\varepsilon = \min\{\varepsilon_{11}, \varepsilon_{12}\}$ yields that (43) is satisfied for $k = 1$. Since the term (59) depends linearly on $\tau$, for $k = 2$ (43) is satisfied for the choice $\varepsilon = \varepsilon_{2} := \min\{\varepsilon_{21}, \varepsilon_{22}\}$. For the case $k = 3$, choosing $\varepsilon = \varepsilon_{3}$, one can see that (43) is valid. Hence, if we choose $\varepsilon := \min\{\varepsilon_{21}, \varepsilon_{12}, \varepsilon_{2}, \varepsilon_{3}\}$, (43) remains satisfied for $k = 1, 2, 3$ if we replace $\delta w$ in (55) by $\tau \delta w$.

**Remark 33.** A careful study of [33, Lemma 6.2.1, Lemma 6.2.2] shows that Theorem 32 also holds for certain source terms which are not equal to zero on $\Omega$, e.g., for source terms which only depend on the state $y$ and satisfy (A1). Furthermore, using the same arguments as in the proof above, one can show that Theorem 32 is still valid if the upper bound $\bar{y}$ is not constant.

**Example 34.** We consider as example a unidirectional road, where the number of incoming cars at $x = 0$ can be controlled. Let the traffic density $\rho$ on the road $[0, \infty)$ be modeled by the so-called LWR-model, see [27] and [36],

$$\begin{align*}
\rho_t + \hat{f}(\rho)_x &= 0, & \text{on} & & \Omega_t := (0, \bar{t}) \times (0, \infty), \\
\rho(0, \cdot) &= \rho_0(\cdot), & \text{on} & & \Omega := (0, \infty), \\
\rho(\cdot, 0^+) &= \rho_{B,0}(\cdot), & \text{in the sense of} & & (6)
\end{align*}$$

with $\hat{f}(\rho) := \rho v(\rho)$ and $v(\rho) = v_{max} \left(1 - \left(\frac{\rho}{\rho_{max}}\right)^n\right)$ describing the speed. Here, we assume for convenience that $n = 1$, $v_{max} = 1$ and $\rho_{max} = 1$. Since our results require convex flux functions, but $\hat{f}$ is concave, we set

$$\begin{align*}
y := -\rho & \quad \text{and} \quad f(y) := -\hat{f}(-y) = yv(-y) = y^2 + y.
\end{align*}$$

This manuscript is for review purposes only.
We note that $f$ is uniformly convex. Applying (62) to (61) gives
\begin{align}
    y_n + f'(y)x &= 0, \quad \text{on } \Omega, \\
    y(0, \cdot) &= u_0(\cdot) := -\rho_0(\cdot), \quad \text{on } \Omega, \\
    y(\cdot, 0+) &= u_{B,0}(\cdot) := -\rho_{B,0}(\cdot), \quad \text{in the sense of (6)}.
\end{align}

Our goal is to maximize the flow on the interval $[0, L]$ at time $t = \bar{t}$, i.e. \( \int_0^L \hat{f}(\rho(\bar{t}, x)) \, dx \)
under the condition that the state constraints $\rho(\bar{t}, \cdot) \leq \bar{\rho} \in \mathbb{R}$ hold on $[0, L]$. By (62),
the corresponding optimal control problem in terms of $y$ reads
\begin{align}
    \min_{w \in W_{ad}} J(y(w)) := \int_0^L f(y(\bar{t}, x; w)) \, dx, \\
    \text{s.t. } w \in W_{ad}, y \text{ satisfies } (63) \text{ and } y(\bar{t}, \cdot; w) \geq -\bar{\rho} \text{ on } [0, L],
\end{align}
where the constant lower bound $-\bar{\rho} < 0$, the control $w$ and the set of admissible
controls $W_{ad}$ will be further specified below. We consider fixed piecewise continuously
differentiable initial data $u_0$ and boundary data of the form
\begin{align}
    u_{B,0}(t; w) = \begin{cases}
    0 & \text{if } t \in [0, t^{on}], \\
    u(t) & \text{if } t \in (t^{on}, t^{off}], \\
    0 & \text{if } t \in (t^{off}, \bar{t}]
    \end{cases}
\end{align}
with control $w = (u, t^{on})$ and fixed $0 < t^{off} < \bar{t}$. We note that this example can be
extended to the case of more than two switching points, where in addition also the
off-switching points can be shifted. Introducing suitable node conditions, it would also
be possible to extend this example to traffic networks.

Let
\begin{align}
    W_{ad} = W = \{ (u, t^{on}) \in C^1([t^{on}, t^{off}]) \times (0, t^{off}) : -\frac{1}{2} + \alpha \leq u \leq 0 \}
\end{align}
for some arbitrary small $\alpha > 0$ and
\begin{align}
    \frac{1}{2} + \alpha < -\bar{\rho} < 0.
\end{align}

To ensure the existence of a control $w$ for which the state constraints in (64) are
satisfied, we suppose that the initial data are chosen such that for any $x \in [0, L]$ and
$w \in W_{ad}$, $y(\bar{t}, x; w)$ does not depend on $u_0$.

We first observe that by our assumptions above the conditions in (A1) and (A2),
except the compactness of $W_{ad}$ in $W$, are satisfied. We will see that there exists
nevertheless an optimal solution in $W_{ad}$.

Next, we note that there is a shock curve $\eta(t; w)$ emanating from $(t^{off}, 0)$, where
we set $x_s(w) := \eta(t; w)$. Our goal is to construct a control $\bar{w}$ for which the necessary
optimality conditions (36)-(38) as well as Robinson’s CQ (24) are satisfied. To this
end, it is intuitive to choose the maximal possible inflow of cars, i.e. $\bar{u} \equiv -\frac{1}{2} + \alpha$.
To satisfy (64), let $\bar{t}^{on} \in (0, t^{off})$ be such that the maximal backward characteristic
through $(\bar{t}, x_s(\bar{w}))$ ends in the inner of the rarefaction wave produced in $(\bar{t}^{on}, 0)$ and
$y(\bar{t}, x_s(\bar{w})+) = -\bar{\rho}$. Hence, $(\bar{t}, x_s(\bar{w}))$ is a non-degenerated shock point. Note that if
$y(\bar{t}, x_s(\bar{w})+) > -\bar{\rho}$, the control $\bar{w}$ will not satisfy (36)-(38), since in this case $(\bar{t}, x_s(\bar{w}))$
can be shifted closer to $x = 0$ by a small change of $\bar{t}^{on}$, which would result in a smaller
cost function value (this can, e.g., be deduced from (69)).
In this setting, \( y(\bar{t}, x; \bar{w}) \) has one discontinuity and no shock generation points on \([0, L]\). Then, the solution to (63) is in a small neighborhood of \( \bar{w} \) of the form

\[
\begin{align*}
Y_1(\bar{t}, x, w) &= 0, & & \text{if } t \in [0, x_s(w)], \\
Y_2(\bar{t}, x, w) &= f^{s-1}\left( \frac{x}{t - t_{on}} \right), & & \text{if } t \in (x_s(w), x_r(w)], \\
Y_3(\bar{t}, x, w) &= 0, & & \text{if } t \in (x_r(w), L],
\end{align*}
\]

where \( x_r(w) \) denotes the right boundary of the rarefaction wave produced in \((t_{on}, 0)\).

In order to make the analysis of this example easier to understand, we provide an illustration of the structure of the solution in Figure 1. In (68), we have used the assumption that \( y(\bar{t}, x; w) \) does not depend on \( u_0 \) for all \( x \in [0, L] \) and \( w \in W_{ad} \).

By Theorem 19, we obtain

\[
\begin{align*}
\hat{r}'(\bar{w})(w - \bar{w}) &= \int_{t_{on}}^{\bar{t}} p(t, 0)f'(\bar{u}(t))(u(t) - \bar{u}(t)) \, dt + p^r \cdot (t_{on} - \bar{t}_{on}),
\end{align*}
\]

where \( p \) denotes the adjoint state with end data \( \bar{\psi}_y(x) \), see Definition 15, and \( p^r \) is given according to (14). Hence, we get \( p(\cdot, 0)|_{(t_{on}, t < \bar{t})} \equiv 1 - \bar{p} \) and

\[
\begin{align*}
p^r &= (1 - \bar{p}) \int_{f'(\bar{t}_{on})}^{f'(\bar{t})} \frac{z^2}{2} \, dz + \int_{f'(\bar{t}_{on})}^{f'(\bar{t})} \frac{z^2}{2} \, dz = \frac{1}{3} \bar{p}^3 - \frac{1}{4} + O(\alpha).
\end{align*}
\]

Inserting this in (69) yields

\[
\begin{align*}
\hat{r}'(\bar{w})(w - \bar{w}) &= (1 - \bar{p}) \cdot \int_{t_{on}}^{\bar{t}} f'(\bar{u}(t))(u(t) - \bar{u}(t)) \, dt \\
&\quad + \left[ \frac{1}{3} \bar{p}^3 - \frac{1}{4} + O(\alpha) \right] \cdot (t_{on} - \bar{t}_{on}).
\end{align*}
\]

By Theorem 16, the derivative \( \frac{\partial}{\partial w} x_s(\bar{w}) \) is given by the right-hand side of (69), but with a different adjoint state, i.e. the adjoint state \( p \) with end data \( p(\bar{t}, \cdot) = 1 - \frac{1}{\bar{p}} \).

For this end data, we get \( p(\cdot, 0)|_{(t_{on}, t < \bar{t})} \equiv 1 - \frac{1}{\bar{p}} \) and

\[
\begin{align*}
p^r &= \frac{1}{\bar{p}} \int_{f'(\bar{t}_{on})}^{f'(\bar{t})} \frac{z}{2} \, dz = \frac{(1 - 2\bar{p})^2}{(4\bar{p})} + O(\alpha)
\end{align*}
\]

such that

\[
\frac{\partial}{\partial w} x_s(\bar{w})(w - \bar{w}) = \frac{1}{\bar{p}} \int_{t_{on}}^{\bar{t}} f'(\bar{u}(t))(u(t) - \bar{u}(t)) \, dt + \left[ \frac{(1 - 2\bar{p})^2}{(4\bar{p})} + O(\alpha) \right] \cdot (t_{on} - \bar{t}_{on}).
\]

We further obtain

\[
\begin{align*}
\frac{\partial}{\partial x} Y_2(\bar{t}, x, \bar{w}) &= \frac{1}{2(t - t_{on})} \quad \text{and} \quad \frac{\partial}{\partial w} Y_2(\bar{t}, x, \bar{w})(w - \bar{w}) = \frac{x(t_{on} - \bar{t}_{on})}{2(t - t_{on})^2}.
\end{align*}
\]
We will prove that there exist nonnegative measures \( \bar{\mu}_1, \bar{\mu}_2 \) and \( \bar{\mu}_3 \) such that the optimality system in (36)-(38) is satisfied for the choice \( w = \bar{w} \), provided that \( \alpha \) is sufficiently small and \( \bar{t} - \bar{t}^{\text{opt}} \geq 2 \). Since we deal with lower bounds in the state constraints, the optimality system has to be modified to

\[
Y_k(\bar{t}, \cdot, \bar{w}) \geq -\bar{\rho} (\cdot) \quad \text{on } I_k \quad \text{for } k = 1, 2, 3
\]

\[
\sum_{k=1}^{3} \int_{I_k} \left( -\bar{\rho} - Y_k(\bar{t}, x, \bar{w}) \right) d\bar{\mu}_k(x) = 0
\]

\[
\frac{\partial}{\partial w} x_k(\bar{w})(w - \bar{w}) + \sum_{k=1}^{3} \frac{\partial}{\partial x} Y_k(\bar{t}, x, \bar{w}) \frac{x - x_{k-1}(\bar{w})}{x_k(\bar{w}) - x_{k-1}(\bar{w})} d\bar{\mu}_k(x)
\]

\[
\cdot \frac{\partial}{\partial w} \bar{w}(w - \bar{w}) + \int_{I_k} \frac{\partial}{\partial x} Y_k(\bar{t}, x, \bar{w}) \frac{x_k(\bar{w}) - x}{x_k(\bar{w}) - x_{k-1}(\bar{w})} d\bar{\mu}_k(x)
\]

\[
\cdot \frac{\partial}{\partial w} x_{k-1}(\bar{w})(w - \bar{w}) - \int_{I_k} \frac{\partial}{\partial w} Y_k(\bar{t}, x, \bar{w})(w - \bar{w}) d\bar{\mu}_k(x) \geq 0 \quad \forall w \in W_{ad}.
\]

Since \( Y_1 = Y_3 \equiv 0 \) and \( Y_2 \) is strictly monotonically increasing with \( Y_2(\bar{t}, x_s(\bar{w}), \bar{w}) = -\bar{\rho} \), (73) is satisfied for the choice of \( \bar{w} \). Next, we observe that (74) yields

\[
\bar{\mu}_1 = \bar{\mu}_3 \equiv 0 \quad \text{and} \quad \bar{\mu}_2(A) = 0 \quad \forall A \subseteq (x_s(\bar{w}), x_r].
\]

Hence, by Corollary 29, (75) reads

\[
\frac{\partial}{\partial w} Y_2(\bar{t}, x_s(\bar{w}), \bar{w})(w - \bar{w}) \sum_{k=1}^{3} \frac{\partial}{\partial x} x_s(\bar{w})(w - \bar{w}) \bar{\mu}_2 \{ x_s(\bar{w}) \} \geq 0 \quad \forall w \in W_{ad}.
\]

Inserting (70)-(72) in (76) gives

\[
\left[ (1 - \bar{\rho} - \frac{\bar{\mu}_2 \{ x_s(\bar{w}) \}}{2\bar{\rho}(t - \bar{t}^{\text{on}})}) \int_{\bar{t}^{\text{on}}}^{\bar{t}^{\text{eff}}} f'(\bar{u}(t))(u(t) - \bar{u}(t)) dt
\]

\[
+ \left[ \frac{\bar{\rho}^3}{3} + \frac{3\bar{\rho}}{4} + 1 + O(\alpha) - \frac{1}{2(t - \bar{t}^{\text{on}})} \left( \frac{1 - 2\bar{\rho}}{4\bar{\rho}} \right)^2 \right. \left( (1 - 2\bar{\rho})(t - \bar{t}^{\text{on}}) \right)
\]

\[
\cdot \bar{\mu}_2(\{x_s(\bar{w})\}) \left( t^{\text{on}} - \bar{t}^{\text{on}} \right) \geq 0 \quad \forall w \in W_{ad}.
\]

Hence, we find a bounded Radon measure of the form

\[
\bar{\mu}_2(\cdot) = \frac{1}{(1 - 2\bar{\rho})^2 + 4\bar{\rho}(1 - 2\bar{\rho})(t - \bar{t}^{\text{on}})} \delta_{x_s(\bar{w})}(\cdot),
\]

where \( \delta_{x_s(\bar{w})}(\cdot) \) denotes the Dirac measure, such that the second term in (77) vanishes.

Moreover, for \( \alpha > 0 \) sufficiently small \( \bar{\mu}_2 \) is nonnegative and

\[
\left[ (1 - \bar{\rho}) - \frac{\bar{\mu}_2 \{ x_s(\bar{w}) \}}{2\bar{\rho}(t - \bar{t}^{\text{on}})} \right] \geq 0,
\]
if \( t - t_{\text{off}} \geq 2 \). By the choice of \( \bar{u} \), the integral in (77) is nonnegative for all \( w \in W_{\text{ad}} \).

Combining this with (79) and the fact that the second term in (77) is equal to zero for the choice in (78), the optimality system (73)–(75) is satisfied. We will now use Theorem 32 to show that Robinson’s CQ (24) is satisfied. Since we deal with lower bounds in this example, the conditions in Theorem 32 have to be modified to

\[
\bar{w} + \delta w \in W_{\text{ad}} \quad \forall \delta w \in B_{\varepsilon}^W(\bar{w}) \cap \{ w \in W_{\text{ad}} : w \geq 0 \}
\]

for a sufficiently small constant \( \varepsilon > 0 \). Since \( \bar{u} \equiv -\frac{1}{2} + \alpha < 0 \), (80) is satisfied.

\[\begin{array}{c}
\text{Fig. 1. Structure of the solution}
\end{array}\]

4. Moreau-Yosida Regularization. Since it is quite involved to compute a solution of the optimality system in Theorem 27, we will now study the convergence of a penalty method. To this end we approximate the state constraints by adding a penalty function \( \gamma P(y(w)) \) with a penalty parameter \( \gamma > 0 \) to the cost functional and obtain

\[
\min_{w \in W} J_{\gamma}(y(w)) := J(y(w)) + \frac{\gamma}{2} \int_a^b (y(t, x; w) - \bar{y}(x))^2_+ \, dx
\]

s.t. \( w \in W_{\text{ad}} \) and \( y(w) \) solves the IBVP (1),

where \( (\cdot)_+ := \max\{\cdot, 0\} \). This approach is often called Moreau-Yosida regularization, see for example [21], [30]. Other approaches can be found, e.g., in [23], [22].

The goal of this section is to conduct a convergence analysis of \((P_{\gamma})\) yielding a convenient analytical basis for the numerical approximation of solutions to the optimality conditions in Theorem 27. The first theorem in this section yields an adjoint based formulation for the derivative of \( J_{\gamma}(y(w)) \), which gives the theoretical foundation for applying numerical optimization methods. The analysis of penalty approaches of the form \((P_{\gamma})\) has been subject to several publications, e.g., for the case of elliptic problems in [30] and for parabolic problems in [31]. However, the case of
conservation laws requires very different techniques and is more involved than the
aforementioned elliptic and parabolic cases, since one has to deal with discontinuous
solutions.

Concerning the derivative of \( J_\gamma(y(w)) \), since \( J_\gamma(\cdot) \) can be written in the form of
\( J(\cdot) \) in (P), we obtain the following result.

**Theorem 35.** Let (A1) and (A2) hold and consider some \( \bar{w} \in W \) satisfying (ND).
Then there exists a neighborhood \( B^W_{\rho}(\bar{w}) \) of \( \bar{w} \) such that the mapping \( B^W_{\rho}(\bar{w}) \ni w \mapsto J_\gamma(y(w)) \in \mathbb{R} \) is continuously differentiable. The derivative in a direction \( \delta w \in W \) can be computed according to Theorem 19.

**Theorem 36.** Let (A1) and (A2) hold. Then for each penalty parameter \( \gamma \) there exists a globally optimal solution \( w_\gamma \in W_{ad} \) for \( (P_\gamma) \).

**Proof.** The proof is similar to the proof of Theorem 20.

**Theorem 37.** Let (A1) and (A2) hold and consider a sequence \( (w_{\gamma_j})_{j \in \mathbb{N}} \subset W_{ad} \)
of global solutions for \( (P_{\gamma_j}) \) with \( \lim_{j \to \infty} \gamma_j = \infty \). Then \( (w_{\gamma_j})_{j \in \mathbb{N}} \) has at least one accumulation point and any accumulation point is a global solution for (P).

**Proof.** Let \( (w_{\gamma_j})_{j \in \mathbb{N}} \) be a sequence of global optima for \( (P_{\gamma_j}) \). Since the set \( W_{ad} \subset W \) is compact, there exists at least one accumulation point \( w^* \). Let \( w^* \) be any accumulation point and consider a convergent subsequence (again denoted by \( (w_{\gamma_j})_{j \in \mathbb{N}} \)) such that \( w_{\gamma_j} \to w^* \in W_{ad} \) w.r.t. \( \|\cdot\|_W \). In the next step, we prove that \( y(w^*) \) fulfills the state constraints. To this end, we firstly note that
\[
J_{\gamma_j}(y(w_{\gamma_j})) \leq J(y(w)) \quad \text{for all } j \in \mathbb{N},
\]
where \( \bar{w} \) denotes a globally optimal solution for (P), which exists due to Theorem 20.

By (81), the continuity of \( J(y(\cdot)) \) w.r.t. \( w \) and the fact that \( w_{\gamma_j} \to w^* \in W_{ad} \), there exists a constant \( C > 0 \) such that for all \( j \in \mathbb{N} \) it holds
\[
0 \leq \frac{\gamma_j}{2} \int_a^b (y(\bar{t}, x; w_{\gamma_j}) - \bar{y}(x))^2_+ \, dx \leq J(y(w)) - J(y(w_{\gamma_j})) \leq C
\]
Using (82), we prove that \( y(\bar{t}, x; w^*) \leq \bar{y}(x) \) holds for all \( x \in [a, b] \). From (82) and \( \gamma_j \to \infty \) we deduce \( y(\bar{t}, \cdot; w_{\gamma_j}) - \bar{y}(\cdot)_+ \to 0 \) in \( L^2([a, b]) \) and hence pointwise almost everywhere on \( [a, b] \) for a subsequence (again denoted by \( (w_{\gamma_j})_{j \in \mathbb{N}} \)). Considering this subsequence, we know by Corollary 7 that \( y(\bar{t}, \cdot; w_{\gamma_j}) \to y(\bar{t}, \cdot; w^*) \) in \( L^1([a, b]) \) and hence pointwise almost everywhere on \( [a, b] \) for another subsequence, which is again denoted by \( (w_{\gamma_j})_{j \in \mathbb{N}} \). Since \( y(\bar{t}, \cdot; w_{\gamma_j}) - \bar{y}(\cdot)_+ \to 0 \) and \( y(\bar{t}, \cdot; w_{\gamma_j}) \to y(\bar{t}, \cdot; w^*) \) for almost all \( x \in [a, b] \), we obtain \( y(\bar{t}, \cdot; w^*) - \bar{y}(\cdot)_+ = 0 \) almost everywhere on \( [a, b] \) which is equivalent to
\[
y(\bar{t}, x; w^*) \leq \bar{y}(x) \quad \text{for a.a. } x \in [a, b].
\]
Analogously to the proof of Theorem 20, one can show that (83) holds for all \( x \in [a, b] \) and hence \( y(\bar{t}, \cdot; w^*) \) fulfills the state constraints. Furthermore, the inequality \( J_{\gamma_j}(y(w_{\gamma_j})) \leq J(y(\bar{w})) \) for all \( j \in \mathbb{N} \) implies \( J(y(w^*)) \leq J(y(\bar{w})) \). Since \( y(\bar{t}, \cdot; w^*) \) satisfies the state constraints and \( w^* \in W_{ad} \), \( w^* \) is thus a globally optimal solution for (P).

We will now examine the convergence of local solutions of \( (P_\gamma) \). To this end, we introduce for some constant \( r > 0 \) the auxiliary problems
\[
(P^r) \quad \min_{w \in W^r} J(y(w)) \quad \text{s.t.} \quad y(w) \text{ solves the IBVP (1), } w \in W^r := W_{ad} \cap B^W_{\rho}(\bar{w})
\]
\[
y(\bar{t}, x; w) \leq \bar{y}(x) \quad \forall x \in [a, b],
\]
\[(P^r_γ) \quad \min_{w \in W} J_γ(y(w)) \text{ s.t. } y(w) \text{ solves the IBVP (1), } w \in W^r\]

(cf. [7]). Similar to the proof of Theorem 20, one can show that these problems admit global solutions denoted by \(\bar{w}^r\) and \(w^{\gamma}_r\), respectively.

**Theorem 38.** Suppose that (A1) and (A2) hold and let \(\bar{w} \in W_{ad}\) be a local optimum for \((P^r)\) such that for constants \(\varepsilon, \delta > 0\) the quadratic growth condition

\[J(\bar{y}(\bar{w})) + \frac{\delta}{2}\|w - \bar{w}\|_H \leq J(y(w)) \quad \forall w \in \bar{W}_{ad} \text{ with } \|w - \bar{w}\|_W < \varepsilon\]

is satisfied, where \(\bar{W}_{ad} := \{ w \in W_{ad} : y(\bar{t}, x; w) \leq \bar{y}(x) \, \forall x \in [a, b]\} \) and \(H\) is a Hilbert space such that \(W\) is continuously embedded in \(H\). Then there exists a sequence \((w, γ)_j \in \mathbb{N}\) of local solutions for \((P_γ)\) converging to \(\bar{w}\) w.r.t. \(\|\cdot\|_W\).

**Proof.** We consider a sequence of globally optimal solutions for \((P^r_γ)\) denoted by \((w^{\gamma}_j)_j \in \mathbb{N} \subset W^r\) with \(\gamma_j \to \infty\) and \(r = \frac{\varepsilon}{2}\). Since \(W^r\) is compact in \(W\) (see proof of Theorem 20), there is a convergent subsequence, again denoted by \((w^{\gamma}_j)_j \in \mathbb{N}\), converging w.r.t. \(\|\cdot\|_W\) to some \(w^* \in W^r\). As in the proof of Theorem 37, one can show that \(w^*\) is a globally optimal solution for \((P^r)\) and due to (84) we obtain that \(w^* = \bar{w}\). Hence, for \(j\) large enough it holds that \(w^{\gamma}_j \in \text{int} \, B^H_\varepsilon(\bar{w})\). This yields that \(w^{\gamma}_j\) is a locally optimal solution for \((P_γ)\) if \(j\) is chosen large enough (cf. [30, Proof of Theorem 5.2]).

**Theorem 39.** Suppose that (A1) and (A2) hold and consider a sequence \((w^{\gamma}_j)_j \in \mathbb{N}\) of local solutions for \((P_γ)\) such that for constants \(\varepsilon, \delta > 0\) the condition

\[J_γ(y(w^{\gamma}_j)) + \frac{\delta}{2}\|w - w^{\gamma}_j\|_H \leq J_γ(y(w)) \quad \forall w \in W_{ad} \text{ with } \|w - w^{\gamma}_j\|_W < \varepsilon\]

is satisfied for all \(j \in \mathbb{N}\) large enough, where \(\|\cdot\|_H\) is defined as in Theorem 38. Then there exists a subsequence converging to some \(\bar{w} \in W_{ad}\) w.r.t. \(\|\cdot\|_W\), which is a local solution for \((P)\).

**Proof.** Since \(W_{ad}\) is by assumption compact, there exists a convergent subsequence again denoted by \((w^{\gamma}_j)_j \in \mathbb{N}\) with limit \(\bar{w} \in W_{ad}\). We consider the corresponding problems \((P^r)\) and \((P^r_γ)\) with \(r = \frac{\varepsilon}{2}\). Then (85) yields that \(w^{\gamma}_j\) is the unique globally optimal solution for \((P^r_γ)\), if \(j\) is large enough. Using this, one can analogously to the proof of Theorem 37 show that \(\bar{w} \in W_{ad}\) is a globally optimal solution for \((P^2)\) and hence a local solution for \((P)\).

**Lemma 40.** Let (A1) and (A2) hold and consider a sequence \((w^{\gamma}_j)_j \in \mathbb{N} \subset W\) with \(\gamma_j \to \infty\) for \(j \to \infty\), that converges to some \(w^* \in W\) satisfying (ND). Then for \(j\) large enough \(w^{\gamma}_j\) satisfies (ND) and there exists \(\varepsilon > 0\) such that for all \(k = 1, \ldots, K + 1\) it holds that

\[\lim_{j \to \infty} Y_k(\bar{t}, \cdot; w^{\gamma}_j) = Y_k(\bar{t}, \cdot; w^*) \quad \text{in } C^1((x_{k-1}(w^*) - \varepsilon, x_K(w^*) + \varepsilon)),\]

\[\lim_{j \to \infty} \frac{\partial}{\partial w} Y_k(\bar{t}, \cdot; w^{\gamma}_j) = \frac{\partial}{\partial w} Y_k(\bar{t}, \cdot; w^*) \quad \text{in } C((x_{k-1}(w^*) - \varepsilon, x_K(w^*) + \varepsilon)).\]

**Proof.** The first assertion is a consequence of [33, Lemma 3.1.10]. The second one and the fact that for \(j\) large enough \(w^{\gamma}_j\) satisfies (ND) follows from Theorem 17.

**Theorem 41.** Let (A1) and (A2) hold and \(w, γ \in W_{ad}\) be a local solution for \((P_γ)\) satisfying (ND). Then it holds that

\[\frac{d}{dw} J_γ(y(w, γ)) \cdot (w - w, γ) \geq 0 \quad \forall w \in W_{ad}.\]
Proof. This result can be found, e.g., in [20].

For \( w \in W \) satisfying (ND), using Theorem 17 we can rewrite \( J_\gamma(y(w)) \) as

\[
J_\gamma(y(w)) = J(y(w)) + \sum_{k=1}^{K+1} z_k(w) \cdot \frac{\gamma}{2z_k(w)} \int_{x_{k-1}(w)}^{x_k(w)} (Y_k(\bar{t}, x; w) - \bar{y}(x))^2 \, dx,
\]

where \( z_k(w) := (x_k(w) - x_{k-1}(w)) \). We will use the following abbreviations:

\[
\lambda_k(x; w) := \begin{cases} 
\gamma (Y_k(\bar{t}, x; w) - \bar{y}(x))^2, & \text{for } x_{k-1}(w) \leq x \leq x_k(w) \\
0, & \text{for } x \in [a, b] \setminus [x_{k-1}(w), x_k(w)],
\end{cases}
\]

\[
r_k(w) := \frac{\gamma}{2z_k(w)} \int_{x_{k-1}(w)}^{x_k(w)} (Y_k(\bar{t}, x; w) - \bar{y}(x))^2 \, dx
\]

Lemma 42. Let (A1) and (A2) hold. Then the derivative of \( J_\gamma(y(w)) \) in some \( w \in W \) satisfying (ND) in a direction \( \delta w \in W \) is given by

\[
\frac{d}{dw} J_\gamma(y(w)) \delta w = \frac{d}{dw} J(y(w)) \delta w + \sum_{k=1}^{K+1} \left[ \int_{x_{k-1}(w)}^{x_k(w)} \frac{\partial}{\partial w} Y_k(\bar{t}, x; w) \delta w \lambda_k(x; w) \, dx \right.
\]

\[
\left. + \int_{x_{k-1}(w)}^{x_k(w)} \frac{\partial}{\partial x} \left( Y_k(\bar{t}, x; w) - \bar{y}(x) \right) \cdot \frac{x-x_{k-1}(w)}{x_k(w)-x_{k-1}(w)} \lambda_k(x; w) \, dx \cdot \frac{\partial}{\partial w} x_k(w) \delta w \right.
\]

\[
+ \int_{x_{k-1}(w)}^{x_k(w)} \frac{\partial}{\partial x} \left( Y_k(\bar{t}, x; w) - \bar{y}(x) \right) \cdot \frac{x_k(w)-x}{x_k(w)-x_{k-1}(w)} \lambda_k(x; w) \, dx \cdot \frac{\partial}{\partial w} x_{k-1}(w) \delta w
\]

\[
\left. + r_k(w) \frac{\partial}{\partial w} (x_k(w) - x_{k-1}(w)) \delta w \right].
\]

Proof. We obtain (90) by differentiating (87) w.r.t. \( w \). The first term of the right hand side of (90) is obvious. The first and the last term in the sum in (90) result from differentiating the integrand and the first factor under the sum in (87) w.r.t. \( w \).

Finally, we observe that the derivative

\[
\frac{d}{dw} \left( z_k(w) \cdot \frac{\gamma}{2z_k(w)} \int_{x_{k-1}(w)}^{x_k(w)} (Y_k(\bar{t}, x; w) - \bar{y}(x))^2 \, dx \right) \bigg|_{w=w} \cdot \delta w = \frac{\gamma}{2} \left[ -\frac{1}{z_k(w)} \right.
\]

\[
\cdot \int_{x_{k-1}(w)}^{x_k(w)} (Y_k(\bar{t}, x; w) - \bar{y}(x))^2 \, dx \cdot \frac{\partial}{\partial w} z_k(w) \delta w + \left( Y_k(\bar{t}, x_k(w); w) - \bar{y}(x_k(w))) \right)^2
\]

\[
\cdot \frac{\partial}{\partial w} x_k(w) \delta w - \left( Y_k(\bar{t}, x_{k-1}(w); w) - \bar{y}(x_{k-1}(w))) \right)^2 \frac{\partial}{\partial w} x_{k-1}(w) \delta w
\]

coincides with the sum of the second and the third integral in (90), where we use that

\[
\frac{\partial}{\partial x} \left( Y_k(\bar{t}, x; w) - \bar{y}(x) \right) \lambda_k(x; w) = \frac{\gamma}{2} \frac{\partial}{\partial x} \left( Y_k(\bar{t}, x; w) - \bar{y}(x) \right)^2
\]

and apply integration by parts.

Lemma 43. Let (A1) and (A2) hold and consider a sequence \( (w_{\gamma_j})_{j \in \mathbb{N}} \) of local solutions \( w_{\gamma_j} \) for \( (P_{\gamma_j}) \) with \( \gamma_j \to \infty \) for \( j \to \infty \) that converges to a local solution \( \bar{w} \in W_{ad} \) for \( (P) \) and satisfies

\[
J_{\gamma_j}(y(w_{\gamma_j})) \leq J_{\gamma_j}(y(w)) \quad \forall w \in W_{ad} \text{ with } \|w - w_{\gamma_j}\|_W < \varepsilon
\]
for some constant $\varepsilon > 0$ and sufficiently large $j$. If $\bar{w}$ satisfies in addition (ND) then
\begin{equation}
\lim_{j \to \infty} r_k(w_{\gamma_j}) = 0 \quad \text{for all } k = 1, \ldots, K + 1.
\end{equation}

**Proof.** We note that since $\bar{w}$ satisfies (ND), Lemma 40 yields that $w_{\gamma_j}$ satisfies (ND) for $j$ large enough such that (87) is valid for all $w = w_{\gamma_j}$ if $j$ is large enough. Using (87), (91), the same arguments as in the proof of Theorem 37 show that
\begin{equation}
\lim_{j \to \infty} \sum_{k=1}^{K+1} z_k(w_{\gamma_j}) r_k(w_{\gamma_j}) = \lim_{j \to \infty} \frac{\gamma_j}{2} \int_a^b (y(\bar{t}, x; w_{\gamma_j}) - \bar{y}(x))^2 \, dx = 0.
\end{equation}

Since $x_{k-1}(w_{\gamma_j})$ and $x_k(w_{\gamma_j})$ are uniformly bounded away from each other for $j \in \mathbb{N}$ large enough and the integrands in (93) are nonnegative, we obtain (92).

**Remark 44.** The quadratic growth condition (84) implies (91). Considering a sequence of local solutions $(w_{\gamma_j})_{j \in \mathbb{N}}$ of $(P_\gamma)$ satisfying (ND), we want to analyze in which sense the terms in (88) converge. In the next lemma we prove that the sequences $(\lambda_k(x;x_{\gamma_j}))_{j \in \mathbb{N}}$ are uniformly bounded in $L^1$.

**Lemma 45.** Let (A1) and (A2) hold and consider a sequence $(w_{\gamma_j})_{j \in \mathbb{N}}$ of local solutions for $(P_\gamma)$ with $\gamma_j \to \infty$ for $j \to \infty$ that satisfies (91) for sufficiently large $j$ and converges to a local solution $\bar{w} \in W_{ad}$ for (P). If $\bar{w}$ satisfies (ND) and (24), then the sequences $(\lambda_k(x;x_{\gamma_j}))_{j \in \mathbb{N}}$ are uniformly bounded in $L^1([a,b])$.

**Proof.** Lemma 40 ensures that $w_{\gamma_j}$ satisfies (ND) for sufficiently large $j$. Therefore (86) is satisfied in $w_{\gamma_j}$, which together with the representation of the gradient in (90), Theorem 19, Lemma 43 and the compactness of $W_{ad}$ yields the existence of a positive constant $C > 0$ such that
\begin{equation}
\int_{x_{k-1}(w_{\gamma_j})}^{x_k(w_{\gamma_j})} \left[ \frac{\partial}{\partial w} Y_k(\bar{t}, x; w_{\gamma_j})(w - w_{\gamma_j}) + \frac{\partial}{\partial x} (Y_k(\bar{t}, x; w_{\gamma_j}) - \bar{y}(x)) \right] \lambda_k(x;x_{\gamma_j}) \, dx \leq C.
\end{equation}

We show that there is $\bar{j} > 0$ such that $\lambda_k(\cdot;x_{\gamma_j})$, $k = 1, \ldots, K + 1$, are uniformly bounded in $L^1([a,b])$ for all $j \geq \bar{j}$. Since (24) is satisfied in $\bar{w}$ by assumption, Lemma 31 yields the existence of $\bar{w} \in W_{ad}$ and $\varepsilon_1 > 0$ such that
\begin{equation}
- \varepsilon_1 \geq Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x) + \frac{\partial}{\partial w} Y_k(\bar{t}, x, \bar{w})(\bar{w} - \bar{w}) + \frac{\partial}{\partial x} (Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)) \cdot (\bar{w} - \bar{w})
\end{equation}
holds for all $x \in [x_{k-1}(\bar{w}), x_k(\bar{w})]$ and $k = 1, \ldots, K + 1$. By Lemma 40 we find $\varepsilon > 0$ and $\bar{j} > 0$ such that for all $j \geq \bar{j}$
\begin{equation}
- \frac{\varepsilon}{2} \geq Y_k(\bar{t}, x, w_{\gamma_j}) - \bar{y}(x) + \frac{\partial}{\partial w} Y_k(\bar{t}, x, w_{\gamma_j})(\bar{w} - w_{\gamma_j}) + \frac{\partial}{\partial x} (Y_k(\bar{t}, x, w_{\gamma_j}) - \bar{y}(x)) \cdot (\bar{w} - w_{\gamma_j}).
\end{equation}
for all \( x \in (x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon) \), since the right hand side of (95) is continuous on \((x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon)\) and the right hand side of (96) converges on this set by Lemma 40 uniformly to the right hand side of (95).

Possibly after increasing \( j \) we have \( \text{supp}(\lambda_k(x; w_{\gamma_j})) = [x_{k-1}(w_{\gamma_j}), x_k(w_{\gamma_j})] \subset (x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon) \) for all \( j \geq j \). Using (96) in connection with (94) and the nonnegativity of \( \lambda_k(x; w_{\gamma_j}) \) we obtain for all \( j \geq j \)

\[
\sum_{k=1}^{K+1} \int_{\text{supp}(\lambda_k(x; w_{\gamma_j}))} \lambda_k(x; w_{\gamma_j}) \, dx \leq \frac{2C}{\varepsilon_1}.
\]

Thus, the sequences \( (\lambda_k(x; w_{\gamma_j}))_{j \in \mathbb{N}} \) are uniformly bounded in \( L^1([a, b]) \).

\[\square\]

**Theorem 46.** Let (A1) and (A2) hold and consider a sequence \( (w_{\gamma_j})_{j \in \mathbb{N}} \) of local solutions for \((P_{\gamma_j})\) with \( \gamma_j \to \infty \) for \( j \to \infty \), that satisfies (91) for sufficiently large \( j \) and converges to a local solution \( \bar{w} \in W_{ad} \) for \((P)\). If \( \bar{w} \) satisfies (ND) and (24), then there exists a subsequence (again denoted by \( (w_{\gamma_j})_{j \in \mathbb{N}} \)) such that \( \lambda_k(\cdot; w_{\gamma_j}) \rightharpoonup \mu_k \) in \( \mathcal{M}([a, b]) \)-weak* for \( k = 1, \ldots, K+1 \). Moreover, \( \text{supp}(\mu_k) \subset I_k = [x_{k-1}(\bar{w}), x_k(\bar{w})] \), the measures \( \mu_k \in \mathcal{M}([a, b]) \) are nonnegative and the optimality conditions in Theorem 27 are satisfied in \( \bar{w} \) for \( \mu_k = \mu_k|_{I_k} \) for all \( k = 1, \ldots, K+1 \). Finally, for all \( k = 1, \ldots, K+1 \) it holds

\[
(97)
\hat{\lambda}_k(\cdot; w_{\gamma_j}) := \lambda_k(x_{k-1}(w_{\gamma_j}) + (\cdot - x_{k-1}(\bar{w}))/x_k(\bar{w}) - x_{k-1}(\bar{w})/x_k(\bar{w}); w_{\gamma_j}) \rightharpoonup \mu_k(\cdot)
\]

in \( \mathcal{M}(I_k) \)-weak* for \( j \to \infty \).

**Proof.** Since the sequences of Lagrange multiplier estimates \( (\lambda_k(\cdot; w_{\gamma_j}))_{j \in \mathbb{N}} \) are uniformly bounded in \( L^1([a, b]) \) by Lemma 45, the Banach-Alaoglu Theorem yields a subsequence, again denoted by \( (\lambda_k(\cdot; w_{\gamma_j}))_{j \in \mathbb{N}} \), such that \( \lambda_k(\cdot; w_{\gamma_j}) \rightharpoonup \mu_k \) in \( \mathcal{M}([a, b]) \)-weak* for \( k = 1, \ldots, K+1 \). Recalling (88) and using the continuity of \( x_1(w), \ldots, x_K(w) \) w.r.t. \( w \), we obtain that \( \text{supp}(\mu_k) \subset I_k \) for all \( k = 1, \ldots, K+1 \). The measures \( \mu_k \in \mathcal{M}([a, b]) \) are nonnegative, since \( (\lambda_k(\cdot; w_{\gamma_j})) \geq 0 \) for all \( j \in \mathbb{N} \). We will prove that the optimality conditions in Theorem 27, i.e. (36), (37) and (38) are satisfied for \( \mu_k = \mu_k|_{I_k} \), \( k = 1, \ldots, K+1 \).

To this end we will use the following argument repeatedly: Let \( k \in \{1, \ldots, K+1\} \) be arbitrary and let \( \psi : [x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon] \times B^W_{\rho}(\bar{w}) \ni (x, w) \mapsto \psi(x, w) \in \mathbb{R} \) be a function satisfying

\[
(98) \quad \lim_{j \to \infty} \psi(\cdot, w_{\gamma_j}) = \psi(\cdot, \bar{w}) \quad \text{in } C ([x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon]),
\]

where \( \varepsilon, \rho > 0 \) are some constants. Then for all \( j \in \mathbb{N} \) sufficiently large we obtain

\[
(99) \quad \text{supp}(\lambda_k(\cdot; w_{\gamma_j})) = [x_{k-1}(w_{\gamma_j}), x_k(w_{\gamma_j})] \subset [x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon] =: I_k^* \]

and thus also \( \lambda_k(\cdot; w_{\gamma_j}) \rightharpoonup \mu_k \) in \( \mathcal{M}(I_k^*) \)-weak*, \( k = 1, \ldots, K+1 \). By using this together with the strong convergence (98) in \( C(I_k^*) \) we conclude that

\[
\lim_{j \to \infty} \int_{I_k^*} \psi(x; w_{\gamma_j}) \lambda_k(x; w_{\gamma_j}) \, dx = \lim_{j \to \infty} \int_{I_k^*} \psi(x, w_{\gamma_j}) \lambda_k(x; w_{\gamma_j}) \, dx
\]

\[
= \int_{I_k^*} \psi(x, \bar{w}) \, d\mu_k(x) = \int_{I_k} \psi(x, \bar{w}) \, d\mu_k(x).
\]
We note that the optimal solutions \( \psi_k \) of the problem \( (P) \) are given by the \( \bar{J} \) transformation formula. For the sake of simplicity, we consider the case where \( \int_{\gamma} \bar{J}(\bar{y}) \) is finite. The last equality follows from \( \text{supp}(\mu_k) = \bar{I}_k \).

Now \( J_{\gamma}(w_{\gamma}) \leq J_{\gamma}(\bar{w}) = J(y(\bar{w})) \) which holds by (91), \( w_{\gamma} \to \bar{w} \) w.r.t. \( \| \cdot \|_W \) and the definitions of \( \lambda_k(\cdot; w) \) and \( J_{\gamma}(y(w)) \), we obtain

\[
0 \leq \lim_{j \to \infty} \sum_{k=1}^{K+1} \int_{x_k(w_{\gamma})}^{x_{k-1}(w_{\gamma})} (Y_k(\tilde{t}, x; w_{\gamma}) - \bar{y}(x))\lambda_k(x; w_{\gamma}) \, dx 
\leq 2 \cdot \lim_{j \to \infty} (J(\bar{w}) - J(w_{\gamma})) = 0.
\]

In order to prove (37), we use (101) in connection with (100). More precisely, we set

\[
\psi_k(x, w; \bar{w}) := \begin{cases}
\frac{\partial}{\partial x} \bar{y}(x) - Y_k(\tilde{t}, x; w) & \text{for } k = 1, \ldots, K + 1, \\
0 & \text{for other } k.
\end{cases}
\]

We note that the optimal solutions \( w_{\gamma} \) for \( (P_{\gamma}) \) satisfy (ND) for sufficiently large \( j \) due to Lemma 40. Therefore, using Theorem 41, (90) and Lemma 43, we obtain for all \( w \in W_{ad} \)

\[
\lim_{j \to \infty} \left( \frac{d}{dw} J(y(w_{\gamma})) \right)(w - w_{\gamma}) \leq 0.
\]

In order to compute the limit in (102), we first note that \( \frac{d}{dw} J(y(w_{\gamma}))(w - w_{\gamma}) \to \frac{d}{dw} J(y(\bar{w}))(w - \bar{w}) \) for \( j \to \infty \) due to Theorem 19. Moreover, \( \frac{\partial}{\partial x} x_k(w_{\gamma}) \to \frac{\partial}{\partial x} x_k(\bar{w}) \) for \( j \to \infty \) by Theorem 17 and together with Lemma 40 we obtain that (98) holds for all functions in the bracket under the integral. Hence, by (100) the integrals in (102) converge to the corresponding integrals with \( \bar{w} \) instead of \( w_{\gamma} \) and \( d\mu_k(x) \) instead of \( \lambda_k(x; w_{\gamma}) \, dx \), which is (38) for \( \mu_k = \mu_k|_{\bar{I}_k} \).

To prove (97), let \( k \in \{1, \ldots, K + 1\} \) and \( \varphi \in C([x_{k-1}(\bar{w}), x_k(\bar{w})]) \) be arbitrarily chosen. For \( \varepsilon > 0 \) we can extend \( \varphi \) to \( C([x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon]) \). Then the transformation formula yields

\[
\lim_{j \to \infty} \int_{x_{k-1}(\bar{w})}^{x_k(\bar{w})} \hat{\lambda}_k(x; w_{\gamma}) \varphi(x) \, dx = \lim_{j \to \infty} \int_{x_{k-1}(w_{\gamma})}^{x_k(w_{\gamma})} \lambda_k(x; w_{\gamma}) \tilde{\varphi}(x, w_{\gamma}) \, dx
\]

with

\[
\tilde{\varphi}(x, w_{\gamma}) = \varphi \left( x_{k-1}(\bar{w}) + (x_k(\bar{w}) - x_{k-1}(\bar{w})) \frac{x - x_{k-1}(w_{\gamma})}{x_k(w_{\gamma}) - x_{k-1}(w_{\gamma})} \right) - \frac{x_k(w_{\gamma}) - x_{k-1}(w_{\gamma})}{x_k(\bar{w}) - x_{k-1}(\bar{w})}.
\]

Now clearly \( \tilde{\varphi}(\cdot, w_{\gamma}) \to \varphi(\cdot, \bar{w}) = \varphi \) in \( C([x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon]) \). Hence, (98) holds for \( \psi = \tilde{\varphi} \) and (100) yields

\[
\lim_{j \to \infty} \int_{x_{k-1}(\bar{w})}^{x_k(\bar{w})} \hat{\lambda}_k(x; w_{\gamma}) \varphi(x) \, dx = \int_{x_{k-1}(\bar{w})}^{x_k(\bar{w})} \varphi(x) \, d\mu_k(x).
\]
and thus (97), since $\varphi$ was arbitrary.

5. Conclusion and possible extensions. We have considered the optimal control of initial-boundary value problems with pointwise state constraints. Hereby, we have proved the existence of an optimal control and have derived necessary optimality conditions. In addition, we have discussed the Moreau-Yosida regularization approach where we have proved that each sequence of optimal solutions for the regularized problems converges strongly to an optimal solution of the original problem with pointwise state constraints. Furthermore, we have derived optimality conditions for the regularized problems and have shown that if the optimal solutions of the regularized problems converge to a local optimum for (P) in which Robinson’s CQ is satisfied, then there exists a sequence of Lagrange multiplier estimates that converges in $\mathcal{M}([a,b])$-weak* to Lagrange multipliers for the optimality system for (P).

The optimality conditions and Robinson’s regularity condition provide a basis for stopping criteria, the proof of convergence of discretizations and the derivation of error estimators. In particular, on the basis of the results in this paper, we obtain that a convergent discretization of the state equation and the adjoint equation, see for example [13–18, 37], are sufficient to obtain a convergent algorithm using the Moreau-Yosida regularization approach.

The method, that was developed in this paper to derive necessary optimality conditions, could be extended to networks with node conditions, e.g., traffic light problems, where the traffic flow is modeled by the LWR-model (see [27, 36]) by using the results of [35]. Another possible extension is to consider systems of balance laws, e.g., the Euler equations modeling the gas flow in a pipe. Hereby, if the solution has similar structural properties, that we have collected in section 2 (see also condition H2 in [6]), then one can introduce new state variables and derive optimality conditions and prove convergence of Moreau-Yosida type regularizations. Hence, if a variational calculus for systems is available, the results of this paper can be generalized to more complex problems. Some results dealing with a variational calculus for systems can be found, e.g., in [5, 6].

Acknowledgments. We thank the referees for their constructive comments. We would like to acknowledge the support by German Research Foundation (DFG) within the collaborative research center TRR 154 "Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks", project A02. Moreover, the work of Stefan Ulbrich was supported by the DFG within the collaborative research center SFB 1194 "Interaction between Transport and Wetting Processes", project B04.
Appendix A. Proof of Theorem 16.

Proof. In the following we will prove Fréchet-differentiability of the mapping (12) in \( u = \bar{w} \), where the continuous differentiability follows from the stability of genuine characteristics and of the adjoint state (cf. [33, Proof of Lemma 6.2.7]). Denote by \( \xi_{l/r} \) the minimal/maximal backward characteristic through \((\bar{t}, x_s(\bar{w}))\). We will restrict ourselves in the proof to the case that \( \xi \) ends in the interior of a rarefaction wave created by a discontinuity of the boundary data \( u_{B,a} \) in \( t = \bar{t}_m \) and \( \xi \) ends in a point \((0, z)\) where the initial data \( u_0 \) is smooth. We further assume that \((\bar{t}_m, a)\) and \((0, \bar{x}^0)\) with \( \bar{x}^0 < z \) are the only rarefaction wave creating discontinuities of the initial and boundary data. Moreover, we assume that

\[(103)\]

\[T^a = \emptyset,\]

where the treatment of the case that \( T^a \not= \emptyset \) is explained in [34, Lemma 4.13]. We note that (6) and (103) together yield

\[(104)\]

\[y(\cdot, a+; u) = u_{B,a}(\cdot) \quad \text{on } [0, \bar{t}]\]

for all \( w \in B_r^W(\bar{w}) \) with \( r > 0 \) small enough. We use these assumptions for the sake of simplicity and in order to avoid technical effort. Nevertheless, the proof can easily be extended to the general case as we will show later.

Consider some \( \delta w \in W \) and set \( w := \bar{w} + \delta w \). Furthermore, let \( \bar{y} := y(\bar{w}) \) and \( y := y(w) \) denote the entropy solutions of (1) for the controls \( u_1 \) and \( u_1 \) in the source term and initial and boundary values \( \bar{u}_0 := u_0(\bar{w}), \bar{u}_{B,a/b} := u_{B,a/b}(\bar{w}), u_0 := u_0(w) \) and \( u_{B,a/b} := u_{B,a/b}(w) \), respectively. We define \( \delta u_0 := u_0 - \bar{u}_0, \delta u_{B,a/b} := u_{B,a/b} - \bar{u}_{B,a/b}, \delta u_1 = u_1 - \bar{u}_1 \) and \( \Delta y := y - \bar{y} \).

In the following, \( C \) and \( \rho \) denote large/small constants, respectively, which possibly change their values throughout the proof. For the sake of simplicity, we set \( a = 0 \).

As in [33, Proof of Lemma 6.3.7], one can show that for all \( \varepsilon > 0 \) it holds that

\[(105)\]

\[\int_{x_s(\bar{w})+\varepsilon}^{x_s(\bar{w})-\varepsilon} \Delta y(\bar{t}, x) \, dx = (x_s(w) - x_s(\bar{w}))(\bar{y}(\bar{t}, x_s(\bar{w}))) + O((\varepsilon + \|\delta w\|_W)\|\delta w\|_W).\]

In the remaining part of the proof, we will derive an adjoint representation for the term

\[(106)\]

\[\frac{1}{[\bar{y}(\bar{t}, x_s(\bar{w}))]} \int_{x_s(\bar{w})-\varepsilon}^{x_s(\bar{w})+\varepsilon} \Delta y(\bar{t}, x) \, dx.\]

As in [34, Proof of Lemma 4.10], we define for \((t, x) \in \Omega_\bar{t} := (0, \bar{t}) \times (0, b)\)

\[(107)\]

\[c(t, x) := f'(\bar{y}(t, x)), \quad d(t, x) := g_y(t, x, \bar{y}(t, x), \bar{u}_1),\]

and observe that \( \Delta y \) is a weak solution of

\[(108)\]

\[\Delta y_t + (\tilde{c} \Delta y)_x = \tilde{d} \Delta y + g(\cdot, \tilde{y}, u_1) - g(\cdot, \bar{y}, \bar{u}_1)\]
on $\Omega$. As described in [34, Lemma 4.10], the functions $c, \tilde{c}, d, \tilde{d}$ can be extended to $[0, \bar{t}] \times \mathbb{R}$ by setting

$$c(t, x) = \tilde{c}(t, x) = M_f \text{ if } x < 0, \quad (d, \tilde{d})(t, x) = (d, \tilde{d})(t, 0+) \text{ if } x < 0,$$

$$c(t, x) = \tilde{c}(t, x) = -M_f \text{ if } x > b \quad \text{ and } \quad (d, \tilde{d})(t, x) = (d, \tilde{d})(t, b-) \text{ if } x > b.$$  

We observe that since $g_y$ does not depend on $y$ by (A1) and due to the regularity of $g$, it holds that

$$\tilde{d}, d \in L^\infty(0, \bar{t}; C^{0,1}(\mathbb{R})) \quad \text{and} \quad \tilde{d} \to d \text{ in } L^\infty(0, \bar{t}; C(\mathbb{R})) \text{ as } \|\delta w\|_W \to 0.$$  

(cf. [40]). Moreover, from (7) and [33, Lemma 6.3.3] we can deduce that there exists a constant $\rho > 0$ such that for all $w \in B_p^W(\bar{w})$ the coefficients $\tilde{c}(\cdot)$ and $c(\cdot)$ satisfy

$$\|\tilde{c}\|_{\infty, [t_1, t_2] \times \mathbb{R}}, \|c\|_{\infty, [t_1, t_2] \times \mathbb{R}} \leq M_y, \quad \|\tilde{c} - c\|_{L^\infty([t_1, t_2] \times \mathbb{R})} \text{ and in } L^\infty([t_1, t_2] \times \mathbb{R})\text{-weak}^* \text{ as } \|\delta w\|_W \to 0.$$  

(cf. [40]). Moreover, from (7) and [33, Lemma 6.3.3] we can deduce that there exists a constant $\rho > 0$ such that for all $w \in B_p^W(\bar{w})$ the coefficients $\tilde{c}(\cdot)$ and $c(\cdot)$ satisfy

$$\|\tilde{c}\|_{\infty, [t_1, t_2] \times \mathbb{R}}, \|c\|_{\infty, [t_1, t_2] \times \mathbb{R}} \leq M_y, \quad \|\tilde{c} - c\|_{L^\infty([t_1, t_2] \times \mathbb{R})} \text{ and in } L^\infty([t_1, t_2] \times \mathbb{R})\text{-weak}^* \text{ as } \|\delta w\|_W \to 0.$$  

We now consider

$$\begin{align*}
\rho_t + \tilde{c} p_x &= -\tilde{d} p \quad \text{on } (t_1, t_2) \times \mathbb{R}, \\
p(t_2, \cdot) &= p_{t_2}(\cdot) = p_{t_2}^s(\cdot) \quad \text{on } \mathbb{R}.
\end{align*}$$  

Since (109), (110) and (111) hold, we can apply [40, Theorem 14] yielding that (112) and (113) admit reversible solutions

$$\begin{align*}
\hat{p} &\in C^{0,1}([t_1, t_2] \times \mathbb{R}) \\
p &\in C^{0,1}([t_1, t_2] \times \mathbb{R}),
\end{align*}$$

respectively. In what follows, a reversible solution has to be understood according to [40, Definition 12]. Moreover, (109), (110) and (111) yield

$$\begin{align*}
\hat{p} &\to p \text{ in } C([t_1, t_2] \times [-R, R])
\end{align*}$$

for all $R > 0$ (cf. [40, Theorem 16]).

Before we start to derive an adjoint representation for the term in (106), we first give a brief overview of the main steps and basic concepts of the proof.

Since $\tilde{w}$ satisfies (ND), for sufficiently small $\bar{\varepsilon}$ the points $x = x_\alpha(\tilde{w}) - \bar{\varepsilon}$ and $x = x_\alpha(\tilde{w}) + \bar{\varepsilon}$ are points of continuity of $\tilde{y}(\tilde{t}, \cdot)$. Therefore, there exist unique backward characteristics $\zeta_{t, r}$ through the points $x_\alpha(\tilde{w}) \mp \bar{\varepsilon}$ such that we can define the set

$$D_{\bar{\varepsilon}} := \{(t, x) \in [0, \bar{t}], x : 0 \leq x \leq \zeta_r(t)\}$$

$\cup \{(t, x) \in [\bar{t}, \tilde{t}], x : \tilde{\zeta}_l(t) \leq x \leq \zeta_r(t)\}.$
Due to the stability of backward characteristics it holds that

\[ D^0 = \{(t, x) \in [0, \tilde{t}_m^n] \times [0, b] : 0 \leq x \leq \xi_r(t)\} \]

\[ \cup \{(t, x) \in [\tilde{t}_m^n, \tilde{t}] \times [0, b] : \xi(t) \leq x \leq \xi_r(t)\}. \]

(118)

The main idea of this proof is equal to the proof of Lemma 4.10 in [34], but since in the present case the shifting of rarefaction centers in the initial and boundary data is allowed, we have to find a slightly different method to derive an adjoint representation of (106). Hereby, we will follow the ideas that are used in [35] and based on the fact that if the source term \( g \) is equal to zero, then the local solution near a rarefaction center is explicitly known. The proof consists of the following main steps (we highly recommend the reader to have a look at Figure 2):

In step 1 we will choose some sufficiently small \( \tilde{t} > \tilde{t}_m^n \) with \( \tilde{t}_m^n < \tilde{t} < \tilde{t}_m^{n+1} \) such that

\[ \tilde{x} := M_f(\tilde{t} - \tilde{t}_m^n) < \frac{\varepsilon}{2} \quad \text{and} \quad \tilde{y}(\tilde{t}, \cdot) \in C^{0,1}(0, 2\tilde{x}) \]

holds and define the set

\[ D_{1,\varepsilon} := \{(t, x) \in [\tilde{t}, \tilde{t}] \times \Omega : \xi_l(t) \leq x \leq \xi_r(t)\}. \]

(120)

Then we consider (108) on \( [\tilde{t}, \tilde{t}] \times \Omega \), multiply it with the reversible solution \( \tilde{p}_1 \) of (112) for \([t_1, t_2] = [\tilde{t}, \tilde{t}] \) and enddata \( \tilde{p}^\varepsilon = [\tilde{p}(\tilde{t}, x, t_0^{(w)})] \) and apply integration by parts on \( D_{1,\varepsilon} \).

Since \( [\tilde{t}, \tilde{t}] \times \Omega \) contains no rarefaction center of \( \tilde{y} \), (111) holds such that using (116) yields that (106) is up to some \( o(\|\delta w\|_W) \) and \( \varepsilon O(\|\delta w\|_W) \) terms equal to some terms \( I_{1,\varepsilon} + I_{2,\varepsilon} \), which are defined in (130). The remaining steps are concerned with further simplifying \( I_{1,\varepsilon} \) and \( I_{2,\varepsilon} \).

In step 2, using the explicitly known local solution near the rarefaction center \( \tilde{t}_m^n \), we will derive a representation of \( I_{1,\varepsilon} \) in (141), which depends on the local solution near the rarefaction center and the adjoint state in Definition 15.

Step 3 is concerned with the simplification of \( I_{2,\varepsilon} \). To this end, we define the domains

\[ D_{2,\varepsilon} := \{(t, x) \in [\tilde{t}, \tilde{t}] \times [0, b] : \max\{0, f'(\tilde{u}_m^{B,a}(\tilde{t}_m^n - \varepsilon))(t-\tilde{t}_m^n + \varepsilon)\} \leq x \leq \xi_r(t)\}, \]

(121)

\[ D_{2} := \{(t, x) \in [\tilde{t}, \tilde{t}] \times [0, b] : \max\{0, f'(\tilde{u}_m^{B,a}(\tilde{t}_m^n))(t-\tilde{t}_m^n)\} \leq x \leq \xi_r(t)\}, \]

(122)

where \( \tilde{t} > \tilde{t}_m^{n-1} \) and \( \varepsilon \) is chosen according to (137). We note that due to the rarefaction center of \( \tilde{y} \) in \((\tilde{t}_m^n, 0)\) the OSLC (111) is violated. To solve this problem, we replace the coefficients \( c \) and \( \tilde{c} \) on \([\tilde{t}, \tilde{t}] \times \mathbb{R}\) by suitable \( \tilde{c}_{\text{loc}}, c_{\text{loc}} \) such that the OSLC (111) is satisfied and in addition \( \tilde{c}_{\text{loc}}|D_{2,\varepsilon} \equiv \tilde{c} \) and \( c_{\text{loc}}|D_{2,\varepsilon} \equiv c \) hold for all \( w \in B^W_{\rho}(\tilde{w}) \) with a sufficiently small \( \rho > 0 \). Let (108) be considered on \([\tilde{t}, \tilde{t}] \times \Omega\), multiply it with the reversible solution \( \tilde{p}_2 \) of (112) for \([t_1, t_2] = [\tilde{t}, \tilde{t}] \), \( \tilde{c} \) replaced by \( \tilde{c}_{\text{loc}} \) and enddata \( \tilde{p}^\varepsilon = \tilde{p}_1(\tilde{t}, \cdot) \). Then we apply integration by parts on \( D_{2,\varepsilon} \). Observing that (109), (110) and (111) are satisfied, we can use (116) to rewrite \( I_{2,\varepsilon} \) in terms of the boundary data, the reversible solution \( p_2 \) of (113) with \( c \) replaced by \( c_{\text{loc}} \) and the term \( I_{21} \), which is defined in (156).

In step 4, we have a closer look at the term \( I_{21} \). Choosing a time point \( s \) with \( 0 < s < \varepsilon_y \), we define the set

\[ D_3 := \{(t, x) \in [s, \tilde{t}] \times [0, b] : 0 \leq x \leq \xi_r(t)\}. \]

(123)
Since \( D_3 \) contains no rarefaction center, we can proceed as in step 1 and derive a representation for \( I_{21} \) (up to some \( o(\|\delta w\|_W) \) terms) depending on the boundary data, the reversible solution \( p_3 \) of (113) for \([t_1, t_2] = [s, \hat{t}]\) and enddata \( p_1 = \hat{p}(\hat{t}) \) and the terms \( I_{1,\varepsilon}^0, I_{2,\varepsilon}^0 \) and \( I_{3,\varepsilon}^0 \) which are defined in (172).

In step 5, we use the same methods as in step 2 to derive a representation of \( I_{2,\varepsilon}^0 \).

Concerning \( I_{0,\varepsilon}^1 \) and \( I_{0,\varepsilon}^3 \), we can use similar arguments as in step 3.

Finally, setting \( \bar{\varepsilon} := \bar{\varepsilon}(\delta w) \) as a function depending on \( \delta w \) and satisfying \( \bar{\varepsilon}(\delta w) \to 0 \) if \( \|\delta w\|_W \to 0 \), we obtain the desired result.

**Fig. 2. Proof of Theorem 16**

**Step 1:** We recall the set \( D_{1,\varepsilon} \) in (120) and observe that defining

\[
D_1 := \{(t, x) \in [\tilde{t}, \bar{t}] \times \Omega : \xi_l(t) \leq x \leq \xi_r(t)\},
\]

it holds that \( D_{1,0} = D_1 \). Since \( y(\cdot; w) \) has no rarefaction center on \( D_{1,\varepsilon} \) for all \( w \in B_\rho^W(\bar{w}) \) with \( \rho > 0 \) small enough, (109), (109) and (111) hold on \([\tilde{t}, \bar{t}] \times \mathbb{R}\).

Therefore,

\[
\begin{align*}
(124) \quad p_t + \xi \partial_x p_x &= -\ddot{d}p & \text{on } (\tilde{t}, \bar{t}) \times \mathbb{R}, & p(t, \cdot) = \frac{1}{[y(t, \xi(t); \bar{w})]} & \text{on } \mathbb{R} \\
(125) \quad p_t + \xi \partial_x p_x &= -\ddot{d}p & \text{on } (\tilde{t}, \bar{t}) \times \mathbb{R}, & p(t, \cdot) = \frac{1}{[y(t, \xi(t); \bar{w})]} & \text{on } \mathbb{R}
\end{align*}
\]

admit reservible solutions \( \hat{p}_1, p_1 \in C^0(\tilde{t}, \bar{t}) \times \mathbb{R} \) satisfying

\[
(126) \quad \hat{p}_1 \to p_1 \text{ in } C([\tilde{t}, \bar{t}] \times [-R, R]) \quad \text{if } \|\delta w\|_W \to 0
\]

for all \( R > 0 \). We consider (108) on \((\tilde{t}, \bar{t}) \times \mathbb{R},\) multiply it with \( \hat{p}_1 \) and apply integration.
by parts on the domain \( D_{1,\varepsilon} \) yielding

\[
\frac{1}{|\hat{y}(t, x_0(\omega))|} \int_{x_0(\omega) - \varepsilon}^{x_0(\omega) + \varepsilon} \Delta y(t, x) \, dx = \int_{\xi(t)}^{\xi(t)} \hat{p}_1(t, x) \Delta y(t, x) \, dx \\
+ \int_{D_{1,\varepsilon}} \hat{p}_1(t, x)(g(t, x, \bar{y}, u_1) - g(t, x, \bar{y}, \bar{u}_1)) \, dx \, dt \\
- \int_{t}^{\hat{t}} \hat{p}_1(t, \zeta(t)) \Delta y(t, \zeta(t))(\hat{c}(t, \zeta(t)) - c(t, \zeta(t))) \, dt \\
(127) + \int_{t}^{\hat{t}} \hat{p}_1(t, \zeta_r(t)) \Delta y(t, \zeta_r(t))(\hat{c}(t, \zeta_r(t)) - c(t, \zeta_r(t))) \, dt.
\]

Using (126) and the regularity of \( g \) w.r.t. \( u_1 \), we can rewrite the second term on the right hand side of (127) by

\[
\int_{D_{1,\varepsilon}} \hat{p}_1(t, x)(g(t, x, \bar{y}, u_1) - g(t, x, \bar{y}, \bar{u}_1)) \, dx \, dt \\
= \int_{D_{1,\varepsilon}} p_1(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 \, dx \, dt + o(\|\delta w\|_W) \\
= \int_{D_{1,\varepsilon} \setminus D_1} p_1(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 \, dx \, dt \\
+ \int_{D_1} p_1(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 \, dx \, dt + o(\|\delta w\|_W)
\]

Hereby, the term \( o(\|\delta w\|_W) \) is uniform w.r.t. \( \varepsilon > 0 \). From [33, Lemma 3.1.15] we can deduce that

\[
\|\xi_{t/r}(\cdot) - \zeta_{t/r}(\cdot)\|_{C([\hat{t}, \hat{t}])} \leq C \varepsilon.
\]

This and the boundedness of \( g_{u_1}(\cdot, \bar{y}, \bar{u}_1) \) yield

\[
\int_{D_{1,\varepsilon} \setminus D_1} p_1(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 \, dx \, dt = \bar{\varepsilon} O(\|\delta w\|_W)
\]

and hence

\[
\int_{D_{1,\varepsilon}} \hat{p}_1(t, x)(g(t, x, \bar{y}, u_1) - g(t, x, \bar{y}, \bar{u}_1)) \, dx \, dt = \int_{D_1} p_1(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 \, dx \, dt \\
+ \bar{\varepsilon} O(\|\delta w\|_W) + o(\|\delta w\|_W).
\]

Since \( \xi_r \) ends in a point where \( \bar{a}_0 \) is smooth and \( \xi_t \) ends in the inner of a rarefaction center we can use [33, Lemma 6.2.1] yielding smooth local solutions \( w \mapsto Y_{t/r}(\cdot, w) \)

defined on some stripes \( S_{t/r} \). Using those local solutions, the definitions of \( \hat{c} \) and \( c \)

and the uniform boundedness of \( \hat{p}_1 \), we can show that the last two integrals on the right hand side of (127) are equal to \( O(\|\delta w\|_W^2) \):

\[
- \int_{t}^{\hat{t}} \hat{p}_1(t, \zeta(t)) \Delta y(t, \zeta(t))(\hat{c}(t, \zeta(t)) - c(t, \zeta(t))) \, dt \\
+ \int_{t}^{\hat{t}} \hat{p}_1(t, \zeta_r(t)) \Delta y(t, \zeta_r(t))(\hat{c}(t, \zeta_r(t)) - c(t, \zeta_r(t))) \, dt = O(\|\delta w\|_W^2)
\]
In order to derive $Y_t$, we apply [33, Lemma 6.2.1] to IBVP on the truncated space-time cylinder $[\tilde{t}, \bar{t}] \times \Omega$ as in the proof of Lemma 14. We note that the term $O(\|\delta w\|_W^2)$ in (130) is uniform w.r.t. $\tilde{\varepsilon}$.

Considering the first integral on the right hand side of (127), we can use (126) and Corollary 7 to show that

$$
\int_{\xi(t)}^{\xi(i)} \tilde{p}_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx = \int_{\xi(i)}^{\xi(t)} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx + o(\|\delta w\|_W),
$$

where $o(\|\delta w\|_W)$ is uniform w.r.t. $\tilde{\varepsilon}$. We further obtain

$$
\int_{\xi(t)}^{\xi(i)} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx = \int_{\xi(i)}^{\xi(t)} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx
$$

+ \int_{\xi(t)}^{\xi(i)} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx + \int_{\xi(i)}^{\xi(t)} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx.

With the help of the local solutions $Y_{i/r}$ that we used to prove (130) and the boundedness of $p_1$, we can show the estimation

$$
\left| \int_{\xi(i)}^{\xi(t)} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx \right| \leq \left( |\xi(i) - \xi(t)| + |\xi(t) - \xi(i)| \right) \cdot O(\|\delta w\|_W),
$$

where the term $O(\|\delta w\|_W)$ is uniform w.r.t. $\tilde{\varepsilon}$. Using this result and (128), we can deduce that

$$
\int_{\xi(i)}^{\xi(t)} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx \leq \tilde{\varepsilon} O(\|\delta w\|_W).
$$

Inserting (133) in (132), we obtain that (131) can be rewritten as

$$
\int_{\xi(i)}^{\xi(t)} \tilde{p}_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx = \int_{\xi(i)}^{\xi(t)} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx + \tilde{\varepsilon} O(\|\delta w\|_W) + o(\|\delta w\|_W).
$$

Inserting (129), (130) and (134) in (106) yields

$$
\int_{\xi(i)}^{\xi(t)} \tilde{p}_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx = \int_{\xi(i)}^{\xi(t)} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx + o(\|\delta w\|_W) + \tilde{\varepsilon} O(\|\delta w\|_W).
$$

We observe that due to [33, Lemma 6.3.5], $p_1$ coincides on $D_1$ with the adjoint state $p$ defined in Definition 15 with enddata $\bar{p}(\cdot) = 1_{x_s(\bar{\omega})} (\cdot) \sum_{s \in (\bar{\omega})} [0, \bar{x}(s,x_s(\bar{\omega}))]$. We will now have a closer look at the first term of the right hand side of (135). For all $\tilde{\varepsilon} \geq 0$ small enough it holds that

$$
\int_{\xi(i)}^{\xi(t)} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx = \int_{\xi(i)}^{\xi(t)} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx
$$

+ \int_{f_t^{\bar{w}_m}(\bar{t}_{m} - \varepsilon) - \tilde{\varepsilon}}^{f_t^{\bar{w}_m}(\bar{t}_{m} - \varepsilon) - \tilde{\varepsilon}} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx = I_{1,\tilde{\varepsilon}} + I_{2,\tilde{\varepsilon}}.
In what follows, we choose $\varepsilon$ as a function of $\delta w$ satisfying
\begin{equation}
(137) \quad \varepsilon(\delta w) \geq 2\|\delta w\|_W \quad \text{and} \quad \varepsilon(\delta w) \to 0 \quad \text{if} \quad \|\delta w\|_W \to 0.
\end{equation}

**Step 2:** Considering $I_{1,\varepsilon}$, we note that since $\xi_t$ ends in the inner of a rarefaction wave, for all $w \in B^W_\rho(\bar{w})$ with $\rho > 0$ small enough it holds that
\begin{equation}
(138) \quad y(\tilde{t}, x, w) = \begin{cases} 
 f^{-1} \left( \frac{x}{t - \bar{t}_m} \right) & \text{if} \quad x \in [f'(u^B_m(t \bar{a}_m^{-} \cdot \bar{m}))(\tilde{t} - t_{\bar{a}_m}^a), f'(u^B_m(t \bar{a}_m^{-} \cdot \bar{m}))(\tilde{t} - t_{\bar{a}_m}^a)], \\
 Y(\tilde{t}, x, w) & \text{if} \quad x \in [f'(u^B_m(t \bar{a}_m^{-} \cdot \bar{m}))(\tilde{t} - t_{\bar{a}_m}^a), f'(u^B_m(t \bar{a}_m^{-} \cdot \bar{m}))(\tilde{t} - t_{\bar{a}_m}^a) + \delta),
\end{cases}
\end{equation}
where $\delta > 0$ is a small constant and $w \in B^W_\rho(\bar{w}) \ni w \to Y(\tilde{t}, \cdot, w) \in C(I)$ is a smooth function obtained by [33, Lemma 6.2.7] with
\begin{equation}
I = (f'(u^B_m(t \bar{a}_m^{-} \cdot \bar{m}))(\tilde{t} - t_{\bar{a}_m}^a) - \delta, f'(u^B_m(t \bar{a}_m^{-} \cdot \bar{m}))(\tilde{t} - t_{\bar{a}_m}^a) + \delta).
\end{equation}

We note that $y(\tilde{t}, x, w)$ is continuous in $x = f'(u^B_m(t \bar{a}_m^{-} \cdot \bar{m}))(\tilde{t} - t_{\bar{a}_m}^a)$ for all $w \in B^W_\rho(\bar{w})$.

Using a Taylor approximation of $f'(u^B_m(t \bar{a}_m^{-} \cdot \bar{m}))(\tilde{t} - t_{\bar{a}_m}^a)$ in $u^B_m = \bar{u}^B_m$ and $t^a = \bar{t}_m$,
we can deduce that there exists a constant $C > 0$ such that
\begin{equation}
\xi_t(\tilde{t}) < f'(u^B_m(t \bar{a}_m^{-} \cdot \bar{m}))(\tilde{t} - t_{\bar{a}_m}^a) - C\|\delta w\|_W
\end{equation}
and
\begin{equation}
(139) \quad \xi_t(\tilde{t}), f'(u^B_m(t \bar{a}_m^{-} \cdot \bar{m}))(\tilde{t} - t_{\bar{a}_m}^a) - C\|\delta w\|_W \subset [\xi_t(\tilde{t}), f'(u^B_m(t \bar{a}_m^{-} \cdot \bar{m}))(\tilde{t} - t_{\bar{a}_m}^a)]
\end{equation}
hold for all $w \in B^W_\rho(\bar{w})$ with $\rho > 0$ small enough. Then (138) and (139) yield
\begin{equation}
(140) \quad \Delta y(\tilde{t}, x)|_{[\xi_t(\tilde{t}), f'(u^B_m(t \bar{a}_m^{-} \cdot \bar{m}))(\tilde{t} - t_{\bar{a}_m}^a) - C\|\delta w\|_W]} = \frac{x \cdot \delta t_{\bar{a}_m}^a}{f'(f^{-1}(\frac{x}{t - \bar{t}_m})) \cdot (t - \bar{t}_m)^2} + o(\|\delta w\|_W).
\end{equation}

Using (140), we can rewrite the term $I_{1,\varepsilon}$ in (136) as follows:
\begin{equation}
I_{1,\varepsilon} = \int_{\xi_t(\tilde{t})} \int_{\tilde{t}_1(\tilde{t})} \frac{x \cdot \delta t_{\bar{a}_m}^a}{f'(f^{-1}(\frac{x}{t - \bar{t}_m})) \cdot (t - \bar{t}_m)^2} dx
\end{equation}
\begin{equation}
+ \int_{\xi_t(\tilde{t})} \int_{\tilde{t}_1(\tilde{t})} \frac{x \cdot \delta t_{\bar{a}_m}^a}{f'(f^{-1}(\frac{x}{t - \bar{t}_m})) \cdot (t - \bar{t}_m)^2} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) dx + o(\|\delta w\|_W)
\end{equation}
\begin{equation}
+ \int_{\xi_t(\tilde{t})} \int_{\tilde{t}_1(\tilde{t})} \frac{x \cdot \delta t_{\bar{a}_m}^a}{f'(f^{-1}(\frac{x}{t - \bar{t}_m})) \cdot (t - \bar{t}_m)^2} dx
\end{equation}
\begin{equation}
- \int_{\xi_t(\tilde{t})} \int_{\tilde{t}_1(\tilde{t})} \frac{x \cdot \delta t_{\bar{a}_m}^a}{f'(f^{-1}(\frac{x}{t - \bar{t}_m})) \cdot (t - \bar{t}_m)^2} dx
\end{equation}
\begin{equation}
+ \int_{\xi_t(\tilde{t})} \int_{\tilde{t}_1(\tilde{t})} \frac{x \cdot \delta t_{\bar{a}_m}^a}{f'(f^{-1}(\frac{x}{t - \bar{t}_m})) \cdot (t - \bar{t}_m)^2} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) dx + o(\|\delta w\|_W)
\end{equation}
\begin{equation}
\int_{\xi_t(\tilde{t})} \int_{\tilde{t}_1(\tilde{t})} \frac{x \cdot \delta t_{\bar{a}_m}^a}{f'(f^{-1}(\frac{x}{t - \bar{t}_m})) \cdot (t - \bar{t}_m)^2} dx
\end{equation}
\begin{equation}
+ \int_{\xi_t(\tilde{t})} \int_{\tilde{t}_1(\tilde{t})} \frac{x \cdot \delta t_{\bar{a}_m}^a}{f'(f^{-1}(\frac{x}{t - \bar{t}_m})) \cdot (t - \bar{t}_m)^2} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) dx + o(\|\delta w\|_W)
\end{equation}
\begin{equation}
Due to the regularity of $f$ and $p_1$, the second integral on the right hand side is equal to $o(\|\delta w\|_W)$. Using again the regularity of $p_1$, the choice of $\varepsilon$ in (137) and the

\text{This manuscript is for review purposes only.}
representation of $y(\tilde{t}, x, w)$ in (138), which is in particular continuous, we obtain that the third integral is also equal to $o(\|\delta w\|_W)$. Hence, we can conclude

$$I_{1, \varepsilon} = \int_{\Xi(\tilde{t})} f'(\tilde{u}_{m,a}^{B,a}(\tilde{t}, \tilde{x})) \cdot (\tilde{t} - \tilde{t}_m) p_1(\tilde{t}, x) f''(f^{-1}(\tilde{t} - \tilde{t}_m)) \cdot (\tilde{t} - \tilde{t}_m) \frac{x \cdot \delta t_m^a}{(\tilde{t} - \tilde{t}_m)^2} \, dx + o(\|\delta w\|_W).$$

Since $p_1$ coincides with the adjoint $p$ on $D_1$, we further obtain

$$I_{1, \varepsilon} = \int_{\Xi(\tilde{t})} f'(\tilde{u}_{m,a}^{B,a}(\tilde{t}, \tilde{x})) \cdot (\tilde{t} - \tilde{t}_m) p(\tilde{t}, x) f''(f^{-1}(\tilde{t} - \tilde{t}_m)) \cdot (\tilde{t} - \tilde{t}_m) \frac{x \cdot \delta t_m^a}{(\tilde{t} - \tilde{t}_m)^2} \, dx + o(\|\delta w\|_W)$$

$$= \lim_{\tilde{t}_m \to \tilde{t}} \int_{\Xi(\tilde{t})} f'(\tilde{u}_{m,a}^{B,a}(\tilde{t}, \tilde{x})) p(\tilde{t}, z(\tilde{t} - \tilde{t}_m)) f''(f^{-1}(z)) \, dz + o(\|\delta w\|_W)$$

(141)

The last equality holds due to the fact that for all $z \in [f'(\tilde{u}_{m,a}^{B,a}(\tilde{t}, \tilde{x})), f'(\tilde{u}_{m,a}^{B,a}(\tilde{t}, \tilde{x}))]$ the term $p(\cdot, z(\cdot - \tilde{t}_m))$ is constant on $(\tilde{t}_m, \tilde{t})$ and is equal to zero for all $z \in [f'(\tilde{u}_{m,a}^{B,a}(\tilde{t}, \tilde{x})), \Xi(\tilde{t})/(\tilde{t} - \tilde{t}_m)]$.

**Step 3:** In order to compute $I_{2, \varepsilon}$, we define suitable coefficients $\tilde{c}_{\text{loc}}$, $c_{\text{loc}}$ such that $\tilde{c}_{\text{loc}}|_{D_{2,\varepsilon}} \equiv \tilde{c}$ and $c_{\text{loc}}|_{D_{2,\varepsilon}} \equiv c$ hold for all $w \in B^W_{\rho}(\tilde{w})$ with a sufficiently small $\rho > 0$ such that (109), (110) and (111) are satisfied. To this end, we choose some $\tilde{t} \in (\tilde{t}_m+1, \tilde{t}_m)$ and define $\tilde{c}_{\text{loc}}$ and $c_{\text{loc}}$ on $[\tilde{t}, \tilde{t}] \times \mathbb{R}$ by

$$\tilde{c}_{\text{loc}}(t, x) = M_{f'} \text{, if } (t, x) \in [\tilde{t}, \tilde{t}] \times [0, \xi_{\varepsilon}(t)] \setminus D_{2, \varepsilon}, \quad \tilde{c}_{\text{loc}}(t, x) = \tilde{c}(t, x), \text{ else}$$

$$c_{\text{loc}}(t, x) = M_{f'} \text{, if } (t, x) \in [\tilde{t}, \tilde{t}] \times [0, \xi_{\varepsilon}(t)] \setminus D_{2, \varepsilon}, \quad c_{\text{loc}}(t, x) = c(t, x), \text{ else}.$$

Recalling the sets $D_{2, \varepsilon}$ and $D_{2}$ defined in (121) and (122), we choose the function $\varepsilon(\delta w)$ in (137) such that

$$D_{2, \varepsilon} \subset D_{2, \varepsilon} \subset D_{2} \text{ for all } w \in B^W_{\rho}(\tilde{w})$$

for sufficiently small $\rho > 0$.

Due to the construction of $\tilde{c}_{\text{loc}}$ and $c_{\text{loc}}$ and the choice of $\varepsilon$ in (137), one can show that (109), (110) and (111) hold on $[\tilde{t}, \tilde{t}] \times \mathbb{R}$. Therefore,

$$p_t + \tilde{c}_{\text{loc}} p_x = -\dot{\tilde{p}}_2 \text{ on } (\tilde{t}, \tilde{t}) \times \mathbb{R}, \quad p(\tilde{t}, \cdot) = p_1(\tilde{t}, \cdot) \text{ on } \mathbb{R}$$

(143)

$$p_t + c_{\text{loc}} p_x = -d \text{ on } (\tilde{t}, \tilde{t}) \times \mathbb{R}, \quad p(\tilde{t}, \cdot) = p_1(\tilde{t}, \cdot) \text{ on } \mathbb{R}$$

(144)

admit reversible solutions $\tilde{p}_2, p_2 \in C^{0,1}([\tilde{t}, \tilde{t}] \times \mathbb{R})$, respectively, satisfying

$$\tilde{p}_2 \to p_2 \text{ in } C([\tilde{t}, \tilde{t}] \times [-R, R]) \text{ if } \|\delta w\|_W \to 0$$

(145)

for all $R > 0$. We note that (142) yields that $\tilde{c}_{\text{loc}}(\cdot)$ coincides with $\tilde{c}(\cdot)$ on $D_{2, \varepsilon}$. Hence, if we consider (108) on $(\tilde{t}, \tilde{t}) \times \mathbb{R}$, multiply it with $p_2$ and apply integration by parts
on $D_{2, \varepsilon}$, then we obtain

\begin{align*}
I_{2,\varepsilon} &= \int_{D_{2,\varepsilon}} \hat{p}_2(t, x)(g(t, x, \bar{y}, u_1) - g(t, x, \bar{y}, \bar{u}_1)) \, dx \, dt + \int_0^{\xi(t)} \hat{p}_2(\hat{t}, x) \Delta y(\hat{t}, x) \, dx \\
+ \int_{\bar{r}_m - \varepsilon}^{\bar{r}_m} \hat{p}_2(t, \gamma(t, \varepsilon)) \Delta y(t, \gamma(t, \varepsilon))(f'(\bar{u}^{B,a}(\bar{r}_m - \varepsilon) - \hat{c}(t, \gamma(t, \varepsilon))) dt \\
- \int_{\bar{r}_m}^{\bar{r}_m+\varepsilon} \hat{p}_2(t, \xi_r(t)) \Delta y(t, \xi_r(t))(f'(\hat{y}(t, \xi_r(t))) - \hat{c}(t, \xi_r(t))) dt \\
(146) + \int_{\bar{r}_m - \varepsilon}^{\bar{r}_m+\varepsilon} \hat{p}_2(t, 0) \hat{c}(t, 0+) \Delta y(t, 0+) \, dt &=: I_{20} + I_{21} + I_{22} + I_{23} + I_{24}.
\end{align*}

where $\gamma(t, \varepsilon) := f'(\bar{u}^{B,a}(\bar{r}_m - \varepsilon))(t - \bar{r}_m + \varepsilon)$.

Concerning (146), we note that $g(\cdot, y(w), u_1) = 0$ holds on $D_{2,0} \setminus D_{2,\varepsilon}$ for all $w \in B_{\rho}^W(\tilde{w})$. From this result, the regularity of $g$ and (145) we obtain

\begin{equation}
I_{20} = (p_2(\cdot), g_{u_1}(\cdot, \bar{y}, \bar{u}_1)\delta u_1)_{2, D_{2,0}} + o(\|\delta w\|_W).
\end{equation}

Now, we show that

\begin{equation}
\|I_{22}\| \leq C\|\delta w\|_W^2
\end{equation}

and note that

\begin{equation}
\|I_{23}\| \leq C\|\delta w\|_W^2
\end{equation}

can be proved analogously.

To this end, we note that since the point $(\hat{t}, f'(\bar{u}^{B,a}(\bar{r}_m))(\hat{t} - \bar{r}_m))$ lies on the right boundary of the rarefaction wave emanating from $(\bar{r}_m, 0)$, [33, Lemma 6.2.7] implies that there exists a stripe $\hat{S}$ with

\begin{equation}
\{(t, x) \in [\bar{r}_m, \hat{t}] \times \Omega : x = f'(\bar{u}^{B,a}(\bar{r}_m))(t - \bar{r}_m)\} \subset S
\end{equation}

\begin{equation}
\{(t, x) \in [\bar{r}_m - \delta, \hat{t}] \times \Omega : x = M_f(t - \bar{r}_m + \delta)\} \subset S,
\end{equation}

and a continuously Fréchet-differentiable mapping $B_{\rho}^W(\tilde{w}) \ni w \mapsto Y_m(\cdot; w) \in C(S)$ such that

\begin{equation}
y(\cdot; w)|_{\hat{S}} = Y_m(\cdot; w) \text{ for all } w \in B_{\rho}^W(\tilde{w})
\end{equation}

holds, where

\begin{equation}
\hat{S} := \{(t, x) \in S : x > \max\{0, f'(\bar{u}^{B,a}(\bar{r}_m)) \cdot (t - \bar{r}_m + \frac{\varepsilon}{2})\}\}
\end{equation}

provided that $\rho > 0$ and $\delta$ are small enough and $\hat{t} > \bar{r}_m$ is sufficiently close to $\bar{r}_m$.

Hereby, (150) implies that $\hat{S}$ is nonempty.

We observe that we can choose $\varepsilon > 0$ in (137) such that $(t, \gamma(t, \varepsilon)) \in \hat{S}$ holds for all $t \in [\bar{r}_m - \varepsilon, \hat{t}]$ and all $w \in B_{\rho}^W(\tilde{w})$. Therefore, (151) yields for all $w \in B_{\rho}^W(\tilde{w})$

\begin{align*}
\|\Delta y(t, \gamma(t, \varepsilon))\|_{C([\bar{r}_m - \varepsilon, \hat{t}])} &= \|Y_m(t, \gamma(t, \varepsilon); w) - Y_m(t, \gamma(t, \varepsilon); \tilde{w})\|_{C([\bar{r}_m - \varepsilon, \hat{t}])} \\
&\leq \|Y_m(\cdot; w) - Y_m(\cdot; \tilde{w})\|_{C(\hat{S})} = \|\frac{\partial}{\partial w} Y_m(\cdot; \tilde{w})\|_{C(\hat{S})} \|\delta w\|_W + o(\|\delta w\|_W) \\
&\leq C\|\delta w\|_W
\end{align*}

This manuscript is for review purposes only.
Next, we note that $I_{22}$ in (146) can be estimated from above by

$$
|I_{22}| \leq C \int_{t_{m-\varepsilon}}^{\bar{t}} |\Delta y(t, \gamma(t, \varepsilon))|^2 \, dt,
$$

where we have used the uniform boundedness of $\hat{p}_2$ and the regularity of $f$. From (153) and (154) we obtain that (148) holds. Using similar arguments and (104), we obtain

$$
I_{24} = (p_2(\cdot, 0), f(\hat{y}(\cdot, 0+))\delta u_m(B, a(\cdot))_{2,[\bar{t}, t_{m}]} + o(\delta w_{\infty})).
$$

Concerning $I_{21}$, from (145) and Corollary 7, we can deduce that

$$
I_{21} = \int_{0}^{t_{(\hat{t})}} p_2(\hat{t}, x) \Delta y(\hat{t}, x) \, dx + o(\delta w_{\infty}).
$$

Due to [33, Lemma 6.3.5], $p_2$ and the adjoint state $p$ coincide on $D_{2,0}$. Therefore, we can replace $p_2$ by $p$ in the terms $I_{20}, I_{21}$ and $I_{24}$.

**Step 4:** Next, we want to further simplify the term $I_{21}$ in (156). To this end, we choose some $s \in (0, \varepsilon_{g})$ and observe that since $\hat{c}$ has no rarefaction centers on the domain $[s, \hat{t}] \times \mathbb{R}$, (109), (110) and (111) hold. Therefore,

$$
p_t + c p_x = -\hat{d}p \quad \text{on} \quad (s, \hat{t}) \times \mathbb{R}, \quad p(\hat{t}, \cdot) = p_2(\hat{t}, \cdot) \quad \text{on} \quad \mathbb{R}
$$

admit reversible solutions $\hat{p}_3, p_3 \in C^{0,1}(s, \hat{t}) \times \mathbb{R}$, respectively. Moreover,

$$
\hat{p}_3 \to p_3 \quad \text{in} \quad C([s, \hat{t}] \times [-R, R]) \quad \text{if} \quad \delta w_{\infty} \to 0
$$

holds for all $R > 0$. Now, we consider (108) on $(s, \hat{t}) \times \Omega$, multiply it with $\hat{p}_3$ and apply integration by parts on $D_3$ defined in (123). This yields

$$
I_{21} = \int_{D_3} \hat{p}_3(t, x)(g(t, x, \hat{y}, u_1) - g(t, x, \hat{y}, \hat{u}_1)) \, dx \, dt
$$

$$
+ \int_{s}^{\hat{t}} \hat{p}_3(t, 0+)\hat{c}(t, 0+)\Delta y(t, 0+) \, dt
$$

$$
- \int_{s}^{\hat{t}} \hat{p}_3(t, \xi_r(t))\Delta y(t, \xi_r(t))(c(t, \xi_r(t)) - \hat{c}(t, \xi_r(t))) \, dt
$$

$$
+ \int_{0}^{\xi_r(s)} \hat{p}_3(s, x) \Delta y(s, x) \, dx + o(\delta w_{\infty})
$$

$$
= I_{31} + I_{32} + I_{33} + I_{34} + o(\delta w_{\infty}).
$$

Analogously to the simplification of (146) in **Step 2**, one can also simplify the terms on the right hand side of (160). Using the regularity of $g$ and (159), we obtain that

$$
I_{31} = \int_{D_3} p_3(t, x) g_{u_1}(t, x, \hat{y}, \hat{u}_1) \delta u_1 \, dx + o(\delta w_{\infty}).
$$

Using that $\xi_r$ ends in a point where the initial data is smooth and $\hat{w}$ satisfies (ND), we can use the same arguments as in the estimation of (148) and obtain

$$
I_{33} = O(\delta w_{\infty}).
$$
Moreover, using (159) and Corollary 7, we obtain

\[ I_{34} = \int_0^{\xi(s)} p_3(s, x) \Delta y(s, x) \, dx + o(\|\delta w\|_W). \]

Now we want to have a closer look at the term \( I_{32} \). Observing that \( \Delta y(\cdot, 0+) = \delta u_{B,a} \) holds on \([s, t]\) due to (104), Lemma 6 yields that \( \|\Delta y(\cdot, 0+)\|_{1,(s,t)} = O(\|\delta w\|_W) \). Using this result, (159) and the uniform boundedness of \( \hat{c} \), we can rewrite \( I_{32} \) by

\[ I_{32} = \int_s^t p_3(t, 0) \hat{c}(t, 0+) \Delta y(t, 0+) \, dt + o(\|\delta w\|_W). \]

Moreover, it holds that

\[ \left\| \Delta y(\cdot, 0+) - \sum_{i=1}^{n_{t,a}+1} \delta u_{i}^{B,a}(\cdot) \mathbb{I}_{B,a}(\tilde{w}_i)(\cdot) - \sum_{i=1}^{n_{t,a}} \text{sgn}(\delta t_i^a)[\tilde{u}_{B,a}(\tilde{t}_i^a)] \right\|_{1,(s,t)} = o(\|\delta w\|_W). \]

Using this result and the uniform boundedness of \( \hat{c} \) and \( p_3 \), we can rewrite the right hand side of (164) by

\[ \int_s^t p_3(t, 0+) \hat{c}(t, 0+) \Delta y(t, 0+) \, dt + o(\|\delta w\|_W) \]

\[ = \sum_{i=1}^{n_{t,a}+1} (p_3(\cdot, 0), \hat{c}(t, 0+) \delta u_{i}^{B,a}(\cdot))_{2, I^{B,a}(s,t)} + \sum_{i \in I_{s,a}, s < \tilde{t}_i < t} (p_3(\cdot, 0), \hat{c}(t, 0+) \cdot \text{sgn}(\delta t_i^a)[\tilde{u}_{B,a}(\tilde{t}_i^a)])_{2, I^{B,a}(s,t)} + o(\|\delta w\|_W) \]

\[ =: I_{321} + I_{322} + o(\|\delta w\|_W). \]

Since \( \hat{c}(\cdot, 0+ \to c(\cdot, 0+) \in L^1([s, t]) \) due to Corollary 7 and \( p_3 \) is bounded, we obtain that

\[ I_{321} = \sum_{i=1}^{n_{t,a}+1} (p_3(\cdot, 0), c(t, 0+) \delta u_{i}^{B,a}(\cdot))_{2, I^{B,a}(s,t)} + o(\|\delta w\|_W). \]

In order to simplify \( I_{322} \) we assume w.l.o.g. that \( \delta t_i^a > 0 \), where the case \( \delta t_i^a < 0 \) can be treated analogously and the case \( \delta t_i^a = 0 \) is trivial. We observe that for all \( i \in I_{s,a} \) with \( s < \tilde{t}_i < t \) it holds that

\[ \|\tilde{u}_{B,a}(\tilde{t}_i^a) - (u_{i}^{B,a}(\cdot) - \tilde{u}_{i}^{B,a}(\cdot))\|_{L^1(I^{B,a}(s,t))} = o(\|\delta w\|_W). \]

This together with the uniform boundedness of \( \hat{c} \) and \( p_3 \) yields

\[ I_{322} = \sum_{i \in I_{s,a}, s < \tilde{t}_i < t} (p_3(\cdot, 0), \hat{c}(\cdot, 0+) \cdot (u_{i}^{B,a}(\cdot) - \tilde{u}_{i}^{B,a}(\cdot)))_{2, I^{B,a}(s,t)} + o(\|\delta w\|_W) \]

\[ = \sum_{i \in I_{s,a}, s < \tilde{t}_i < t} (p_3(\cdot, 0), f(u_{i}^{B,a}(\cdot) - f(\tilde{u}_{i}^{B,a}(\cdot)))_{2, I^{B,a}(s,t)} + o(\|\delta w\|_W), \]
where the equality follows from the definition of $\breve{c}$. We further observe that due to the regularity of $p_3$, $f$ and $v^B_j$ for $j = 1, \ldots, n_{t,a}$ it holds that

$$I_{322} = \sum_{i \in I_{n,a}, s < t_i^\alpha} p_3(\tilde{r}_i^\alpha, 0)(f(\tilde{u}_{i+1}^B(\tilde{r}_i^\alpha))) - f(\tilde{u}_i^B(\tilde{r}_i^\alpha)) \delta t_i^\alpha + o(\|\delta w\|_W)$$

(168)

Inserting (166) and (168) in (165) and (165) in (164), we obtain

$$I_{32} = \sum_{i \in I_{n,a}, s < t_i^\alpha} p_3(\tilde{r}_i^\alpha, 0)[f(\tilde{y}(\tilde{r}_i^\alpha, 0+)) \delta t_i^\alpha + o(\|\delta w\|_W).$$

(169)

Inserting (161), (162), (163) and (169) in (160) finally yields

$$I_{21} = (p(\cdot), g_{u_1}(\cdot, \bar{y}, \bar{u}_1) \delta u_1)_{2,D_3}$$

$$+ (p(s, \cdot), \Delta y(s, \cdot))_{2,0, \xi,(s)} + \sum_{i=1}^{n_{t,a}+1} (p(\cdot, 0), f'(\tilde{u}_i^B, \cdot) \delta u_i^B)_{2, I_{t,a} \cap [s, \bar{t}]}$$

(170)

$$+ \sum_{i \in I_{n,a}, s \leq t_i^\alpha < \bar{t}} p_3(\tilde{r}_i^\alpha, 0)[f(\tilde{y}(\tilde{r}_i^\alpha, 0+)) \delta t_i^\alpha + o(\|\delta w\|_W).$$

We note that $p_3$ coincides with the adjoint state $p$ on $D_3$ such that we obtain from

$$I_{21} = (p(\cdot), g_{u_1}(\cdot, \bar{y}, \bar{u}_1) \delta u_1)_{2,D_3}$$

$$+ (p(s, \cdot), \Delta y(s, \cdot))_{2,0, \xi,(s)} + \sum_{i=1}^{n_{t,a}+1} (p(\cdot, 0), f'(\tilde{u}_i^B, \cdot) \delta u_i^B)_{2, I_{t,a} \cap [s, \bar{t}]}$$

(171)

$$+ \sum_{i \in I_{n,a}, s \leq t_i^\alpha < \bar{t}} p(\tilde{r}_i^\alpha, 0)[f(\tilde{y}(\tilde{r}_i^\alpha, 0+)) \delta t_i^\alpha + o(\|\delta w\|_W).$$

The second term of the right side of (171) can be rewritten as follows:

$$p(s, \cdot, \Delta y(s, \cdot))_{2,0, \xi,(s)}$$

$$= (p(s, \cdot, \Delta y(s, \cdot))_{2,0, \xi, \cdot + f'(\tilde{u}_0^B(\tilde{x}_0^\cdot - \xi)), s}$$

$$+ (p(s, \cdot, \Delta y(s, \cdot))_{2,0, \xi, \cdot + f'(\tilde{u}_0^B(\tilde{x}_0^\cdot - \xi)), s} \cdot \tilde{r}_1^\alpha + f'(\tilde{u}_1^B(\tilde{x}_1^\cdot + \xi)), s}$$

$$+ (p(s, \cdot, \Delta y(s, \cdot))_{2,0, \xi, \cdot + f'(\tilde{u}_1^B(\tilde{x}_1^\cdot + \xi)), s})$$

(172)

Step 5: Using similar arguments as in the estimation of $I_{2,\varepsilon}$, one can firstly show that

$$I_{1,\varepsilon}^0 = (p(0, \cdot, \delta u_0(\cdot))_{2,0, \xi, \cdot} + o(\|\delta w\|_W).$$
where the term $o(\|\delta w\|_W)$ is uniform w.r.t. $\varepsilon$. Next, we observe that

$$\left\| \delta u_0(\cdot) - \sum_{i=1}^l \delta u_0(\cdot) \mathbf{1}_{I_i^0}(\bar{w})(\cdot) - \sum_{i=1}^{l-1} \text{sgn}(\delta x_i^0) \mathbf{1}_{I_i^0}(\bar{x}_i^0+\delta x_i^0)(\cdot) \| \tilde{u}_0(\bar{x}_i^0) \delta x_i^0 \right\|_{L^1(0,\bar{x}_i^0-\varepsilon)} = o(\|\delta w\|_W),$$

where the term $o(\|\delta w\|_W)$ is uniform w.r.t. $\varepsilon$ again. Hence, by the regularity of $p$ and the choice of $\varepsilon$ in (137), we obtain

$$I_{1,\varepsilon}^0 = \sum_{i=1}^l (p(0,\cdot), \delta u_i^0(\cdot))_{2,T_i^0(\bar{w})} + \sum_{i=1}^{l-1} (0, \bar{x}_i^0)(x_i^0) p(0, x_i^0) \tilde{u}_0(\bar{x}_i^0) \delta x_i^0 + o(\|\delta w\|_W)$$

where the second equality holds due to the fact that the adjoint state $p$ is by definition equal to zero on $\{ (t,x) \in [0,\bar{t}] \times [0,\bar{x}] : x > \xi_r(t) \}$. Analogously, one can show that

$$I_{3,\varepsilon}^0 = \sum_{i=1}^{n+1} (p(0,\cdot), \delta u_i^0(\cdot))_{2,T_i^0(\bar{w})} + \sum_{i=1}^{n+1} p(0, x_i^0) \tilde{u}_0(\bar{x}_i^0) \delta x_i^0 + o(\|\delta w\|_W).$$

Similar to the estimation of $I_{1,\varepsilon}$, using that

$$y(s,x,w) \| f'(u_i^0(x_i^0)) \cdot s + x_i^0, f'(u_{i+1}^0(x_{i+1}^0)) \cdot s + x_{i+1}^0 = f'^{-1} \left( \frac{x - x_i^0}{s} \right) \quad \forall w \in B^W_{\rho}(\bar{w})$$

for sufficiently small $\rho > 0$, we can further show that

$$I_{2,\varepsilon}^0 = - \int f'(u_i^0(x_i^0)) \lim_{t \to 0} p(t, zt + x_i^0) \frac{\delta x_i^0}{f'(f'^{-1}(z))} dz \quad o(\|\delta w\|_W),$$

where the term $o(\|\delta w\|_W)$ is uniform w.r.t. $\varepsilon$. Inserting (137), (174) and (175) in (172), we obtain

$$\left( p(s,\cdot), \Delta y(s,\cdot) \right)_{2,(0,\xi_r(s),\nu)} = \sum_{i=1}^{n+1} (p(0,\cdot), \delta u_i^0(\cdot))_{2,T_i^0(\bar{w})}$$

$$- \int f'(u_i^0(x_i^0)) \lim_{t \to 0} p(t, zt + x_i^0) \frac{\delta x_i^0}{f'(f'^{-1}(z))} dz$$

$$+ \sum_{i \in I_{1,\varepsilon}(\bar{w})} p(0, x_i^0) \tilde{u}_0(\bar{x}_i^0) \delta x_i^0 + o(\|\delta w\|_W).$$

This manuscript is for review purposes only.
Inserting (176) in (171), we can deduce

\[ I_{21} = (p(\cdot), g_{u_1}(\cdot, y, u_1)\delta u_1)_{2, D_2} + \sum_{i=1}^{n_{t, a} + 1} (p(0, \cdot), \delta u_i^0(\cdot))_{2, t_i^0(\bar{w})} \]

\[ \int_{f'((\bar{u}_i^0(\bar{\bar{x}}_i^0)))}^{f'((\bar{u}_i^0(\bar{\bar{x}}_i^0)))} \lim_{t \to 0} p(t, zt + \bar{x}_i^0(t)) \frac{\delta x_i^0}{f'(f^{-1}(z))} \, dz \]

\[ + \sum_{i \in I_{s, o}(\bar{w}), s \leq t_i^0 < t_i^1} p(0, \bar{x}_i^0(\bar{\bar{u}}_0(\bar{\bar{x}}_i^0)))\delta x_i^0 + \sum_{i=1}^{n_{t, a} + 1} (p(0, \cdot), f'(\bar{u}_i^0(\cdot))\delta u_i^0(\cdot))_{2, t_i^0(\bar{w})} \cap [0, \bar{t}_i] \]

\[ (177) + \sum_{i \in I_{s, a}(\bar{w}), s \leq t_i^0 < t_i^1} p(t_i^0, 0)[f(\bar{\bar{x}}_i^0, 0)] \delta t_i^0 + o(\|\delta w\| W). \]

As already mentioned at the end of step 2, \( p_2 \) and the adjoint state \( p \) coincide on \( D_{2, 0} \) such that we can replace \( p_2 \) by \( p \) in (147) and (155) yielding

\[ I_{20} = (p(\cdot), g_{u_1}(\cdot, \bar{y}, \bar{u}_1)\delta u_1)_{2, D_{2, 0}} + o(\|\delta w\| W) \]

\[ I_{24} = (p(\cdot), f'(\bar{y}_i(\cdot, 0+))\delta u_i^0(\cdot))_{2, t_i^0(\bar{w})} + o(\|\delta w\| W). \]

Inserting (148), (149), (177), (178) and (179) in (146), we obtain

\[ I_{2, e} = (p(\cdot), g_{u_1}(\cdot, \bar{y}, \bar{u}_1)\delta u_1)_{2, D_{2, 0}, o, D_3} \]

\[ + \sum_{i=1}^{n_{t, a} + 1} (p(0, \cdot), f'(\bar{u}_i^0(\cdot))\delta u_i^0(\cdot))_{2, t_i^0(\bar{w})} \cap [0, \bar{t}_i] \]

\[ + \sum_{i \in I_{s, a}(\bar{w}), s \leq t_i^0 \leq t_i^1} p(t_i^0, 0)[f(\bar{\bar{y}}_i^0, 0+)] \delta t_i^0 + \sum_{i=0}^{n_{t, a} + 1} (p(0, \cdot), \delta u_i^0(\cdot))_{2, t_i^0(\bar{w})} \]

\[ - \int_{f'((\bar{u}_i^0(\bar{\bar{x}}_i^0)))}^{f'((\bar{u}_i^0(\bar{\bar{x}}_i^0)))} \lim_{t \to 0} p(t, zt + \bar{x}_i^0(t)) \frac{\delta x_i^0}{f'(f^{-1}(z))} \, dz \]

\[ + \sum_{i \in I_{s, o}(\bar{w})} p(0, \bar{x}_i^0(\bar{\bar{u}}_0(\bar{\bar{x}}_i^0)))\delta x_i^0 + o(\|\delta w\| W). \]

Since \( p(\cdot, 0)|_{[\bar{t}_m, \bar{t}]} \equiv 0 \), choosing \( s > 0 \) sufficiently small, we obtain that (180) can be simplified to

\[ I_{2, e} = (p(\cdot), g_{u_1}(\cdot, \bar{y}, \bar{u}_1)\delta u_1)_{2, D_{2, 0}, o, D_3} \]

\[ + \sum_{i=1}^{n_{t, a} + 1} (p(0, \cdot), f'(\bar{u}_i^0(\cdot))\delta u_i^0(\cdot))_{2, t_i^0(\bar{w})} \cap [0, \bar{t}_i] \]

\[ + \sum_{i \in I_{s, a}(\bar{w}), t_i^0 \leq t_i^1} p(t_i^0, 0)[f(\bar{\bar{y}}_i^0, 0+)] \delta t_i^0 + \sum_{i=0}^{n_{t, a} + 1} (p(0, \cdot), \delta u_i^0(\cdot))_{2, t_i^0(\bar{w})} \]

\[ - \int_{f'((\bar{u}_i^0(\bar{\bar{x}}_i^0)))}^{f'((\bar{u}_i^0(\bar{\bar{x}}_i^0)))} \lim_{t \to 0} p(t, zt + \bar{x}_i^0(t)) \frac{\delta x_i^0}{f'(f^{-1}(z))} \, dz \]

\[ + \sum_{i \in I_{s, o}(\bar{w})} p(0, \bar{x}_i^0(\bar{\bar{u}}_0(\bar{\bar{x}}_i^0)))\delta x_i^0 + o(\|\delta w\| W). \]
Inserting (181) and (141) in (136) and further (136) in (135) yields
\[
\frac{1}{[y(t, x, y)]} \int_{x,0}^{x,0} \Delta y(t, x) \, dx = \int_{D_2 \cup D_3 \cup D_1} p(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 \, dx \, dt \\
+ \int_{f'(u_{m+1}^m(\bar{r}_m^n))} \lim_{n_m \to 1} p(t, z(t - \bar{r}_m^n)) \frac{\delta t_m^n \cdot z}{f''(f^{-1}(z))} \, dz \\
+ \sum_{i=1}^{n_m-1+1} (p(\cdot, 0), f'(u_1^B,\cdot(\cdot)) B_{u_1^B,\cdot(\cdot)} 2, I_{\bar{u}_1^B(\bar{y})} \cap [0, \bar{\eta}]) + \sum_{i \in I_{\bar{u}_1^B}, i^* \leq i} p(\bar{r}_i^n, 0) [f(\bar{y}(t_i, 0+))] \delta t_i^n \\
+ \sum_{i=0}^{n_m+1} (p(0, \cdot), \delta u_i^0(\cdot)) 2, I_0 \cap [0, \bar{\eta}] \\
+ \sum_{i \in I_{\bar{u}_1^B}, i^* \leq i} p(\bar{r}_i^n, 0) [f(\bar{y}(t_i, 0+))] \delta t_i^n \\
+ \sum_{i=0}^{n_m+1} (p(0, \cdot), \delta u_i^0(\cdot)) 2, I_0 \cap [0, \bar{\eta}] \\
+ \sum_{i \in I_{\bar{u}_1^B}, i^* \leq i} p(\bar{r}_i^n, 0) [f(\bar{y}(t_i, 0+))] \delta t_i^n \\
+ \sum_{i=0}^{n_m+1} (p(0, \cdot), \delta u_i^0(\cdot)) 2, I_0 \cap [0, \bar{\eta}] \\
+ \sum_{i \in I_{\bar{u}_1^B}, i^* \leq i} p(\bar{r}_i^n, 0) [f(\bar{y}(t_i, 0+))] \delta t_i^n \\
(182) \\
- \int_{f'(\bar{u}_i^0(\bar{x}_i^0))} \lim_{I_{\bar{u}_1^B} \cap [0, \bar{\eta}]} p(t, z + \bar{x}_i^0) \frac{\delta x_i^0}{f''(f^{-1}(z))} \, dz.
\]

Since it holds \(\text{supp}_n(p) \subset D_2 \cup D_3 \cup D_1\), we can finally conclude that the representation in (182) coincides with the representation in Theorem 16. In the case of \(\tilde{\eta}\) rarefaction center \(\{i_1, \ldots, i_\tilde{n}\}\), the term \((p(s, \cdot), \Delta y(s, \cdot)) 2, \Omega\) in (172) is splitted as follows:
\[
(p(s, \cdot), \Delta y(s, \cdot)) 2, \Omega = (p(s, \cdot), \Delta y(s, \cdot))_{2,(-\infty, x_1^0 - \epsilon + f'(\bar{u}_i^0 (x_1^0 - \epsilon)) \cdot s}] \\
+ \sum_{j=2}^{n-1} (p(s, \cdot), \Delta y(s, \cdot))_{2, x_j^0 - \epsilon + f'(\bar{u}_i^0 (x_j^0 - \epsilon)) \cdot s, x_j^0 + \epsilon + f'(\bar{u}_i^0 (x_j^0 + \epsilon)) \cdot s}] \\
+ (p(s, \cdot), \Delta y(s, \cdot))_{2, x_n^0 + \epsilon + f'(\bar{u}_i^0 (x_n^0 + \epsilon)) \cdot s, \infty} = \sum_{k=1}^{n+1} I_k^0 \\
\]

Then we can treat all terms \(I_k^0\) analogously to \(I_1^0, I_2^0\) and \(I_3^0\), respectively and obtain the desired result. In the case that also the right boundary is within shock

This manuscript is for review purposes only.
funnel and there are several rarefaction centers, we introduce time points \( \hat{t}_i > \hat{t}_0 \) for \( i \in I_{\alpha}(\hat{w}) \) and \( \bar{t}_i > \bar{t}_0 \) for \( i \in I_{\beta}(\hat{w}) \) satisfying (119), respectively. The treatment of the case that \( |T^a|, |T^b| > 1 \) is basically the same as it was shown in Step 3. \( \square \)

REFERENCES


OPTIMAL BOUNDARY CONTROL OF BALANCE LAWS


This manuscript is for review purposes only.