

# Optimal Boundary Control of Hyperbolic Balance Laws with State Constraints \*

JOHANN M. SCHMITT<sup>†</sup> AND STEFAN ULBRICH<sup>‡</sup>

**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

$$(P) \quad \min J(y(w)) := \int_a^b \psi(y(\bar{t}, x; w), y_d(x)) \, dx + R(w),$$

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

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

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

$$(2) \quad y(\bar{t}, x) \leq \bar{y}(x) \quad \text{for all } x \in [a, b].$$

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  $w = (u^0, u^{B,a}, u^{B,b}, x^0, t^a, t^b, u_1) \in W_{ad}$ , where  $W_{ad} \subset C^1(\bar{\Omega})^{n_x+1} \times C^1([0, \bar{t}])^{n_t, a+1} \times C^1([0, \bar{t}])^{n_t, b+1} \times \mathbb{R}^{n_x} \times [0, \bar{t}]^{n_t, a} \times [0, \bar{t}]^{n_t, b} \times C([0, \bar{t}]; C^1(\mathbb{R})^m)$ , the initial and boundary data

$$(3) \quad u_0(x; w) = \begin{cases} u_1^0(x) & \text{if } x \in (a, x_1^0], \\ u_j^0(x) & \text{if } x \in (x_{j-1}^0, x_j^0], 2 \leq j \leq n_x, \\ u_{n_x+1}^0(x) & \text{if } x \in (x_{n_x}^0, b), \end{cases}$$

\*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.

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

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

36

$$(4) \quad u_{B,l}(t; w) = \begin{cases} u_1^{B,l}(t) & \text{if } t \in [0, t_1^l], \\ u_j^{B,l}(t) & \text{if } t \in (t_{j-1}^l, t_j^l], 2 \leq j \leq n_{t,l}, l \in \{a, b\}, \\ u_{n_{t,l}+1}^{B,l}(t) & \text{if } t \in (t_{n_{t,l}}, \bar{t}]. \end{cases}$$

38

39

40

41

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

$$(42) \quad \sigma(l) = \begin{cases} 1 & \text{if } l = a, \\ -1 & \text{if } l = b, \end{cases} \quad \text{and} \quad v(l_{\sigma(l)}) = \begin{cases} v(a+) & \text{if } l = a, \\ v(b-) & \text{if } l = b. \end{cases}$$

43

44

45

46

47

48

**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.

49

50

51

52

53

54

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  $L^\infty$ .

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

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 from the regularized problem converge weakly to the Lagrange multipliers of (P) for the state constraints.

75

76

77

78

79

**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

80 mild assumptions. Moreover, we carry out a convergence analysis of Moreau-Yosida  
 81 regularizations, which provides a theoretical basis for conveniently computing solutions  
 82 of the optimality system without any prior knowledge of the shock structure. In  
 83 contrast to [5, 6], the introduction of shock curves as additional state variables is not  
 84 required. Moreover, these results form the basis of numerical approaches for solving  
 85 optimal control problems with hyperbolic balance laws and state constraints. We  
 86 further present a practical example dealing with the optimal control of traffic flow.  
 87 In this example, we demonstrate that the results of this paper are applicable to real  
 88 world problems. Finally, the present paper extends the sensitivity and adjoint calculus  
 89 of [34, 38, 39] to the case where the shifting of rarefaction centers is allowed.

90 **1.2. Organization.** The paper is organized as follows. In section 2 we introduce  
 91 basic assumptions and collect results concerning the well-posedness of the IBVP and  
 92 structural properties of its solution. In section 3 we discuss the optimal control problem  
 93 (P), prove existence of a global optimum and derive first order necessary optimality  
 94 conditions. In section 4 we apply the Moreau-Yosida regularization approach and prove  
 95 convergence to the solution of the optimal control problem with state constraints.

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

99 **2.1. Notion of a solution for the IBVP and basic assumptions.** We con-  
 100 sider entropy solutions of (1) in the sense of [24], i.e., for every (Kruřkov-) entropy  
 101  $\eta_c(\lambda) := |\lambda - c|$ ,  $c \in \mathbb{R}$ , and associated entropy flux  $q_c(\lambda) := \text{sgn}(\lambda - c)(f(\lambda) - f(c))$   $y$   
 102 has to satisfy

$$\begin{aligned} 103 \quad (5a) \quad & (\eta_c(y))_t + (q_c(y))_x - \eta'_c(y)g(\cdot, y, u_1) \leq 0 && \text{in } \mathcal{D}'(\Omega_{\bar{t}}), \\ 104 \quad (5b) \quad & \text{esslim}_{t \rightarrow 0+} \|y(t, \cdot) - u_0\|_{1, \Omega \cap (-R, R)} = 0 && \text{for all } R > 0. \end{aligned}$$

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

$$108 \quad (6) \quad \min_{k \in I(y(\cdot, l_{\sigma(l)}), u_{B,l})} \sigma(l) \text{sgn}(u_{B,l} - y(\cdot, l_{\sigma(l)}))(f(y(\cdot, l_{\sigma(l)})) - f(k)) = 0 \text{ a.e. on } [0, \bar{t}],$$

109 with  $I(\alpha, \beta) := [\min(\alpha, \beta), \max(\alpha, \beta)]$ .

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

122 We will work under the following assumptions.

123 (A1) The flux function satisfies  $f \in C_{\text{loc}}^3(\mathbb{R})$ ,  $f'^{-1} \in C_{\text{loc}}^{2,\beta}(\mathbb{R})$  for some  $\beta \in (0, 1]$   
 124 and is strongly convex, i.e., there exists a positive constant  $m_{f''} > 0$  such that

125  $f'' \geq m_{f''}$ . Moreover, we assume that  $g \in C([0, \bar{t}]; C_{\text{loc}}^1(\mathbb{R} \times \mathbb{R} \times \mathbb{R}^m))$  and  
 126 that there exists  $\varepsilon_g > 0$  such that for all  $(y, u_1) \in \mathbb{R} \times \mathbb{R}^m$

$$127 \quad g(\cdot, y, u_1) = 0 \quad \text{on } [0, \varepsilon_g] \times \mathbb{R} \cup [0, \bar{t}] \times \mathbb{R} \setminus (a + \varepsilon_g, b - \varepsilon_g).$$

129 Finally, let  $g$  be Lipschitz w.r.t.  $x$  and affine linear w.r.t.  $y$ .

130 (A2) The set of admissible controls  $W_{\text{ad}}$  is nonempty, convex and compact in

$$131 \quad W := C^1(\bar{\Omega})^{n_x+1} \times U_{B,a}^\alpha \times U_{B,b}^\alpha \times \mathcal{X} \times \mathcal{T}^a \times \mathcal{T}^b \times C([0, \bar{t}]; C^1(\mathbb{R}^m)),$$

132 where  $\mathcal{X} := \{\bar{x} \in \Omega^{n_x} : a < x_1 < \dots < x_{n_x} < b\}$ ,

$$133 \quad \mathcal{T}^l := \{\bar{t} \in [0, \bar{t}]^{n_{t,l}} : 0 < t_1 < \dots < t_{n_{t,l}} < \bar{t}\},$$

$$134 \quad U_{B,l}^\alpha := \{u^{B,l} \in C^1([0, \bar{t}])^{n_{t,l}+1} : \sigma(l)f'(u_j^{B,l}) \geq \alpha, j = 1, \dots, n_{t,l} + 1\},$$

136 for  $l \in \{a, b\}$  and some  $\alpha > 0$  and  $W$  is equipped with the norm

$$137 \quad \|w\|_W := \|u^0\|_{C^1(\bar{\Omega})^{n_x+1}} + \|u^{B,a}\|_{C^1([0, \bar{t}])^{n_{t,a}+1}} + \|u^{B,b}\|_{C^1([0, \bar{t}])^{n_{t,b}+1}} \\ 138 \quad + \|x^0\|_1 + \|t^a\|_1 + \|t^b\|_1 + \|u_1\|_{C([0, \bar{t}]; C^1(\mathbb{R}^m))}$$

140 For all  $w \in W_{\text{ad}}$ , it holds that  $\|u_1\|_{L^\infty(0, \bar{t}; C^1(\bar{\Omega})^m)} \leq M_u$  for some  $M_u > 0$ .

141 NOTATION 3. Given some  $w \in W$ , we will use the abbreviations  $[\psi(x)] := \psi(x-) -$   
 142  $\psi(x+)$ ,  $I_0^j(w) := (x_{k-1}^0, x_k^0)$  for  $j = 1, \dots, n_x + 1$  and  $I_{B,l}^j(w) := (t_{j-1}^l, t_j^l)$  for  $j =$   
 143  $1, \dots, n_{t,l} + 1$  and set  $x_0^0 := a, x_{n_x+1}^0 := b, t_0^l = 0$  and  $t_{n_{t,l}+1}^l = \bar{t}$  for  $l \in \{a, b\}$ .  
 144 Furthermore, we define  $B_\rho^W(w) := \{\hat{w} \in W : \|w - \hat{w}\|_W \leq \rho\}$  and the sets of indices

$$145 \quad I_{s,0}(w) := \{j \in \{1, \dots, n_x\} : [u_0(x_j^0)] \geq 0\},$$

$$146 \quad I_{s,l}(w) := \{j \in \{1, \dots, n_{t,l}\} : \sigma(l)[u_{B,l}(t_j^l)] \leq 0\}, \quad l \in \{a, b\},$$

$$147 \quad I_{r,0}(w) := \{1, \dots, n_x\} \setminus I_{s,0}(w), \quad I_{r,l}(w) := \{1, \dots, n_l\} \setminus I_{s,l}(w), \quad l \in \{a, b\}.$$

149 **2.2. Existence and uniqueness of a solution for the IBVP.** We obtain the  
 150 following existence and uniqueness result for entropy solution of (1) [1, 9, 32]. Oleinik's  
 151 entropy condition (7) can be found in [33, Lemma 3.1.13].

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

$$156 \quad \|y(t, \cdot; w)\|_{\infty, [a, b]} \leq M_y,$$

$$157 \quad \|y(t, \cdot; w) - y(t, \cdot; \hat{w})\|_{1, [a, b]} \leq L_y (\|u_0(w) - u_0(\hat{w})\|_{1, I_t} + \|u_{B,a}(w) - u_{B,a}(\hat{w})\|_{1, [0, t]} \\ 158 \quad + \|u_{B,b}(w) - u_{B,b}(\hat{w})\|_{1, [0, t]} + \|u_1 - \hat{u}_1\|_{1, [0, t] \times I_t}),$$

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

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

$$163 \quad (7) \quad y_x(t, \cdot) \leq \frac{C}{1 - e^{-m_{f''} C \min(t, C\varepsilon_1, C\varepsilon_2)}} \quad \text{on } [a + \varepsilon_1, b - \varepsilon_2]$$

165 in the sense of distributions yielding  $y(t, x-) \geq y(t, x+)$  for all  $(t, x) \in \Omega_{\bar{t}}$ .

166 CONVENTION 5. We consider the representative of  $y$  satisfying  $y \in C([0, \bar{t}]; L^1(\Omega))$   
 167 and  $y(t, x) = y(t, x-)$  for all  $(t, x) \in [0, \bar{t}] \times (a, b]$ ,  $y(t, a) = y(t, a+)$  for all  $t \in [0, \bar{t}]$ .

168 LEMMA 6. Under Assumptions (A1) and (A2) for all  $w, \hat{w} \in W_{\text{ad}}$  it holds for  
 169  $u_0, u_{B,a}, u_{B,b}$  defined in (3) and (4) that

$$\begin{aligned} 170 \quad & \|u_0(w) - u_0(\hat{w})\|_{1,[a,b]} \leq (|b-a| + \|w\|_W + \|\hat{w}\|_W) \|w - \hat{w}\|_W \\ 171 \quad & \|u_{B,l}(w) - u_{B,l}(\hat{w})\|_{1,[0,\bar{t}]} \leq (\bar{t} + \|w\|_W + \|\hat{w}\|_W) \|w - \hat{w}\|_W, \quad l \in \{a, b\}. \end{aligned}$$

173 *Proof.* Since by (3) the locations of the  $i$ th discontinuity differ by  $|x_i^0 - \hat{x}_i^0|$  and  
 174  $|u_0(x; w) - u_0(x; \hat{w})|$  can be bounded by  $\|u^0\|_{C(\bar{\Omega})^{n_x+1}} + \|\hat{u}^0\|_{C(\bar{\Omega})^{n_x+1}}$ , we obtain

$$\begin{aligned} 175 \quad & \|u_0(w) - u_0(\hat{w})\|_{1,[a,b]} \leq |b-a| \|u^0 - \hat{u}^0\|_{C(\bar{\Omega})^{n_x+1}} \\ 176 \quad & + (\|u^0\|_{C(\bar{\Omega})^{n_x+1}} + \|\hat{u}^0\|_{C(\bar{\Omega})^{n_x+1}}) \|x^0 - \hat{x}^0\|_1 \leq (|b-a| + \|w\|_W + \|\hat{w}\|_W) \|w - \hat{w}\|_W. \end{aligned}$$

178 Analogously, one can show the remaining two inequalities.  $\square$

179 Using Proposition 4, Lemma 6 and the boundedness of  $W_{\text{ad}}$  we obtain

180 COROLLARY 7. Let (A1) and (A2) hold. Then there exists  $L_y > 0$  such that

$$181 \quad \|y(t, \cdot; w) - y(t, \cdot; \hat{w})\|_{1,[a,b]} \leq L_y \|w - \hat{w}\|_W \quad \text{for all } w, \hat{w} \in W_{\text{ad}}.$$

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

184 COROLLARY 8. Let (A1) and (A2) hold. Then there exists  $L > 0$  such that

$$185 \quad |J(y(w)) - J(y(\hat{w}))| \leq L \|w - \hat{w}\|_W \quad \text{for all } w, \hat{w} \in W_{\text{ad}}.$$

187 **2.3. Structure of the solution of the IBVP.** In this section we will summarize  
 188 some results of [34] concerning the structure of entropy solutions of (1) which can be  
 189 obtained by Dafermos' theory of generalized characteristics [10], see [33, 34, 39, 40].

190 DEFINITION 9. A Lipschitz curve  $[\alpha, \beta] \subset [0, \bar{t}] \rightarrow \Omega_{\bar{t}}, t \mapsto (t, \xi(t))$  satisfying

$$191 \quad (8) \quad \dot{\xi}(t) \in [f'(y(t, \xi(t)+)), f'(y(t, \xi(t)-))] \quad \text{a.e. on } [\alpha, \beta]$$

193 is called a generalized characteristic on  $[\alpha, \beta]$ . If the lower and upper bounds in (8) coin-  
 194 cide for almost all  $t \in [\alpha, \beta]$ , then  $\xi$  is called genuine.  $\xi_{\pm}$  is called a maximal/minimal  
 195 characteristic, if it satisfies  $\dot{\xi}_{\pm}(t) = f'(y(t, \xi(t) \pm))$ .

196 According to [10, Theorem 3.3], if  $\xi$  is a genuine characteristic,  $y(\cdot, \xi(\cdot)) = v(\cdot)$  holds  
 197 almost everywhere on some interval  $[\alpha, \beta]$ , where  $v$  is given by the solution of the  
 198 characteristic equation

$$199 \quad (9) \quad \dot{\zeta}(t) = f'(v(t)), \quad \dot{v}(t) = g(t, \zeta(t), v(t), u_1(t, \zeta(t))),$$

201 and  $\xi(\cdot)$  coincides with  $\zeta(\cdot)$  on  $[\alpha, \beta]$  almost everywhere. For given initial conditions  
 202  $(\zeta, v)(\tau) = (z, w)$  with  $z \in [a, b]$  and  $w \in \mathbb{R}$ , we denote the solution of (9) by  
 203  $(\zeta, v)(\cdot, \tau, z, w, u_1)$ . The solutions of the characteristic equation (9) and its linearized  
 204 version form the basis for the sensitivity and adjoint calculus which is developed in [34]  
 205 for problems of the form (P), see also [38–40].

206 NOTATION 10. Let

$$207 \quad \mathbb{T}^l := \{\theta \in [0, \bar{t}] : \sigma(l)f'(y(\theta+, l_{\sigma(l)})) > 0 \wedge \sigma(l)f'(y(\theta-, l_{\sigma(l)})) < 0\}, \quad l \in \{a, b\}$$

208 and denote for all  $\theta \in \mathbb{T}^a \cup \mathbb{T}^b$  the corresponding maximal/minimal backward character-  
 209 istic through  $(\theta, a/b)$  by  $\xi_{a/b}^\theta$  and the corresponding timepoint where the characteristic  
 210 leaves the domain  $\Omega_{\bar{t}}$  by  $\vartheta_{a/b}^\theta$ . Analogously to [34], we define

$$211 \quad (10) \quad D_- := \bigcup_{\theta \in \mathbb{T}^a} \{(t, x) : t \in ]\vartheta_a^\theta, \theta[, x \in ]a, \xi_a^\theta(t)[\} \cup \bigcup_{\theta \in \mathbb{T}^b} \{(t, x) : t \in ]\vartheta_b^\theta, \theta[, x \in ]\xi_b^\theta(t), b[\}. \\ 212$$

213 If  $\xi_{a/b}^\theta$  ends in a point  $(\vartheta_{a/b}^\theta, b/a)$ , we extend  $D_-$  by adding  $[0, \vartheta_{a/b}^\theta] \times \Omega$  to the right  
 214 hand side of (10).

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

216 LEMMA 11. Let (A1) and (A2) hold and consider some  $w \in W$  satisfying

$$217 \quad (11) \quad \operatorname{ess\,inf}_{t : u_{B,l}(t;w) \neq y(t, l_{\sigma(l)}; w)} |f(u_{B,l}(t;w)) - f(y(t, l_{\sigma(l)}; w))| > 0, \quad l \in \{a, b\}. \\ 218$$

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

221 We define nondegeneracy of shocks according to [34]:

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

228 We will work under the following nondegeneracy condition.

229 (ND) A control  $w \in W$  is called nondegenerated if the following holds: There is no  
 230 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,  
 231 but not differentiable. The corresponding solution  $y(\bar{t}, \cdot; w)$  has no shock  
 232 generation points on  $[a, b]$  and a finite number of discontinuities  $a < x_1 <$   
 233  $\dots < x_N < b$  that are no shock interaction points and nondegenerated  
 234 according to Definition 12. Moreover, (11) is satisfied and for all  $t_j^l \in \mathbb{T}^l$  with  
 235  $l \in \{a, b\}$  the following is satisfied: The extreme backward characteristic  $\xi_l^{t_j^l}$   
 236 through  $(l, t_j^l)$  ends in the interior of a rarefaction wave or in a point  $(0, \bar{z})$  or  
 237  $(\bar{\theta}, l)$  where  $u_0$  respectively  $u_{B,l}$  is smooth. In the latter case we require that

$$238 \quad \frac{d}{dz} \zeta(t; 0, z, u_0(z), u_1)|_{z=\bar{z}} \geq \beta \quad \forall t \in [0, t_j^l] \\ 239 \quad \text{respectively } \sigma(l) \frac{d}{d\theta} \zeta(t; \theta, l, u_{B,l}(\theta), u_1)|_{\theta=\bar{\theta}} \leq -\beta \quad \forall t \in [\bar{\theta}, t_j^l] \\ 240$$

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

$$243 \quad \frac{d}{dw} \zeta(t; 0, \bar{z}, w, u_1)|_{w=\bar{w}} \geq \beta \quad \forall t \in [0, t_j^l] \quad \text{with } \bar{w} \in ]u_0(z-), u_0(z+)[ \\ 244 \quad \text{respectively } \frac{d}{dw} \zeta(t; \theta, l, w, u_1)|_{w=\bar{w}} \geq \beta(t - \bar{\theta}) \quad \forall t \in [\bar{\theta}, t_j^l] \\ 245$$

246 with  $\bar{w} \in ]u_{B,a}(\bar{\theta}+), u_{B,a}(\bar{\theta}-)[$  respectively  $\bar{w} \in ]u_{B,b}(\bar{\theta}-), u_{B,b}(\bar{\theta}+)[$ .

247 REMARK 13. The conditions in (ND) imply that the points  $t_j^{a/b} \in \mathbb{T}^{a/b}$  are nonde-  
248 generated in the sense of [34, Definition 4.12].

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

251 LEMMA 14 (Stability of the shock position). *Let (A1) and (A2) hold and consider*  
252 *some  $\bar{w} \in W$  satisfying (ND) and a nondegenerated shock point  $\bar{x}$  of  $y(\bar{t}, \cdot; \bar{w})$ . Then*  
253 *there exist functions  $Y_{l/r}$  constructed around the minimal and maximal characteristics*  
254 *through  $(\bar{t}, \bar{x})$ , an interval  $(x_l, x_r) \ni \bar{x}$  and a Lipschitz continuous function*

$$255 \quad (12) \quad x_s : w \in B_\rho^W(\bar{w}) \mapsto x_s(w) \in (x_l, x_r)$$

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

$$258 \quad y(\bar{t}, x; w)|_{(x_l, x_s(w))} = Y_l(\bar{t}, x; w), \quad y(\bar{t}, x; w)|_{(x_s(w), x_r)} = Y_r(\bar{t}, x; w).$$

260 *Proof.* The first case where the extreme characteristics through  $(\bar{t}, \bar{x})$  end in points  
261 where the initial or boundary data are smooth has been shown in [33, Lemma 6.3.1].  
262 We now consider the second case where  $\xi_-$  ends in the interior of a rarefaction wave  
263 produced in  $(\bar{t}_j^a, a)$  and note that all remaining cases can be treated analogously.  
264 We choose some  $\tilde{t} > \bar{t}_j^a$  satisfying  $M_{f'}(\tilde{t} - \bar{t}_j^a) < \frac{\varepsilon_g}{2}$ . Since  $g$  is equal to zero for all  
265  $t \in [0, \varepsilon_g]$  and  $x \in [a, a + \varepsilon_g] \cup [b - \varepsilon_g, b]$ , defining  $I = (a + f'(\bar{u}_{j+1}^{B,a}(\bar{t}_j^a)) \cdot (\tilde{t} - \bar{t}_j^a) +$   
266  $\varepsilon, a + f'(\bar{u}_j^{B,a}(\bar{t}_j^a)) \cdot (\tilde{t} - \bar{t}_j^a) - \varepsilon)$  yields

$$267 \quad \forall \varepsilon > 0 \exists \rho > 0 : \quad y(\tilde{t}, x; w)|_I = f'^{-1}\left(\frac{x-a}{\tilde{t}-\bar{t}_j^a}\right) \quad \forall w \in B_\rho^W(\bar{w}).$$

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

272 In the next theorem we will prove Fréchet-differentiability of the shock position  
273 (12) and derive an adjoint-based representation of the gradient. Formally, for a given  
274 solution  $y$  of (1) and given end data  $p^{\bar{t}}$ , the corresponding adjoint equation reads

$$275 \quad (13) \quad p_t + f'(y)p_x = -g_y(\cdot, y, u_1)p \quad \text{on } \Omega_{\bar{t}}, \quad p(\bar{t}, \cdot) = p^{\bar{t}}(\cdot) \quad \text{on } \Omega,$$

$$276 \quad p(t, l_{\sigma(t)}) = 0 \quad \text{on } \{t \in (0, \bar{t}) : \sigma(t)f'(y(\tilde{t}, l_{\sigma(t)}; w)) \leq 0\}.$$

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

278 DEFINITION 15. *Consider a bounded function  $p^{\bar{t}}$  that is pointwise everywhere the*  
279 *limit of a sequence  $(p_n^{\bar{t}})_{n \in \mathbb{N}} \subset C^{0,1}(\Omega)$  which is bounded in  $C(\Omega) \cap W_{loc}^{1,1}(\Omega)$ . Then the*  
280 *adjoint state associated with (13) is defined by the following requirements:*

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

$$283 \quad \dot{p}^\xi(t) = -g_y(t, \xi(t), y(t, \xi(t)), u_1(t, \xi(t)))p^\xi(t), \quad t \in (0, \bar{t}) : \xi(t) \in \Omega,$$

$$284 \quad p^\xi(\bar{t}) = p^{\bar{t}}(\bar{x}).$$

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

287 **THEOREM 16** (Differentiability of shock points). *Let (A1) and (A2) hold. Con-*  
 288 *sider any  $\bar{w} \in W$  satisfying (ND) and a nondegenerated shock point  $\bar{x}$  of  $y(\bar{t}, \cdot; \bar{w})$ .*  
 289 *Then the mapping (12) is continuously differentiable with derivative (cf. Notations 1,*  
 290 *3)*

$$\begin{aligned}
 291 \quad & \frac{\partial}{\partial w} x_s(w) \delta w = (p, g_{u_1}(\cdot, y, u_1) \delta u_1)_{2, \Omega_{\bar{t}}} - \sum_{i \in I_{r,0}(w)} p_i^{r,0} \delta x_i^0 + \sum_{i=1}^{n_x+1} (p(0, \cdot), \delta u_i^0)_{2, I_0^i(w)} \\
 292 \quad & + \sum_{i \in I_{s,0}(w)} p(0, x_i^0) [u_0(x_i^0)] \delta x_i^0 + \sum_{l \in \{a,b\}} \sigma(l) \left( \sum_{\substack{i \in I_{r,l}(w): \\ t_i^l \leq \bar{t}}} p_i^{r,l} \delta t_i^l + \right. \\
 293 \quad & \left. \sum_{i=1}^{n_{t,l}+1} (p(\cdot, l_{\sigma(l)}), f'(u_i^{B,l}) \delta u_i^{B,l})_{2, I_{B,l}^i(w) \cap ]0, \bar{t}[} + \sum_{\substack{i \in I_{s,l}(w): \\ t_i^l \leq \bar{t}}} p(t_i^l, l_{\sigma(l)}) [f(y(t_i^l, l_{\sigma(l)}; w))] \delta t_i^l \right) \\
 294
 \end{aligned}$$

295 for  $\delta w \in W$ , where  $p$  denotes the adjoint state with  $p(\bar{t}, \cdot) = \mathbb{1}_{\bar{x}}(\cdot) \frac{1}{[y(\bar{t}, \bar{x}; \bar{w})]}$  and

$$296 \quad (14a) \quad p_j^{r,0} := \int_{f'(u_j^0(x_j^0))}^{f'(u_{j+1}^0(x_j^0))} \lim_{t \searrow 0} p(t, zt + x_j^0) \frac{1}{f''(f'^{-1}(z))} dz, \quad j \in I_{r,0}(w),$$

$$297 \quad (14b) \quad p_j^{r,l} := \int_{f'(u_{j+1}^{B,l}(t_j^l))}^{f'(u_j^{B,l}(t_j^l))} \lim_{t \searrow t_j^l} p(t, z(t - t_j^l)) \frac{z}{f''(f'^{-1}(z))} dz, \quad j \in I_{r,l}(w), \quad l = a, b.$$

299 *Proof.* This result is proved in [Appendix A](#). □

300 From [Lemma 14](#) and [Theorem 16](#), we can deduce the following result.

301 **THEOREM 17.** *Let (A1) and (A2) hold and consider some  $\bar{w} \in W$  satisfying (ND).*  
 302 *Then there exist constants  $\rho, \varepsilon > 0$  and continuously differentiable mappings*

$$303 \quad (15) \quad Y_k : (x, w) \in I_k^\varepsilon \times B_\rho^W(\bar{w}) \mapsto Y_k(\bar{t}, x; w) \in \mathbb{R}$$

$$304 \quad (16) \quad x_k : w \in B_\rho^W(\bar{w}) \mapsto x_k(w) \in (x_k(\bar{w}) - \frac{\varepsilon}{2}, x_k(\bar{w}) + \frac{\varepsilon}{2}),$$

306 where  $I_k^\varepsilon := (x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon)$  for  $k = 1, \dots, K+1$ , and  $x_0 = a, x_{K+1} = b$ .  
 307 Finally, for all  $w \in B_\rho^W(\bar{w})$  it holds

$$308 \quad (17) \quad y(\bar{t}, x; w)|_{[a,b]} = Y_1(\bar{t}, x; w) \cdot \mathbb{1}_{[a, x_1(w)]}(x) + \sum_{k=2}^{K+1} Y_k(\bar{t}, x; w) \cdot \mathbb{1}_{(x_{k-1}(w), x_k(w)]}(x)$$

309 and the mappings  $w \in B_\rho^W(\bar{w}) \mapsto Y_k(\bar{t}, \cdot; w) \in C(I_k^\varepsilon)$  are continuously differentiable.

310 **REMARK 18.** *The derivatives of the mappings  $x_k$  in (16) can be computed as*  
 311 *follows: If  $x_k(w)$  is a discontinuity of  $y(\bar{t}, \cdot; w)$ , its derivative can be computed according*  
 312 *to [Theorem 16](#). If  $x_k(w)$  lies on the boundary of a rarefaction wave, the corresponding*  
 313 *derivative can be computed by solving the linearized characteristic equation (3.36a)-*  
 314 *(3.36c) in [33]. The mappings (15) are those described in [Lemma 14](#). The corresponding*  
 315 *derivatives can be computed according to [33, Lemma 6.2.1 and 6.2.7], where in the*  
 316 *case that the minimal/maximal characteristics end in the inner of rarefaction centers,*  
 317 *we consider the truncated IBVPs as described in the proof of [Lemma 14](#).*

318 Analogously to the proof of [Theorem 16](#), one can extend [\[33, Theorem 5.2.6\]](#) to the  
319 case that shifts of rarefaction centers are allowed:

320 **THEOREM 19** (Differentiability of the tracking-type functional). *Let (A1) and*  
321 *(A2) hold and consider some  $\bar{w} \in W$  satisfying (ND). Further assume that  $y_d$  is*  
322 *approximately continuous in a neighborhood of the discontinuities of  $y(\bar{t}, \cdot, \bar{w})$  on  $[a, b]$ .*  
323 *Then the mapping  $w \in B_\rho^W(\bar{w}) \mapsto \hat{J}(w) := J(y(w)) \in \mathbb{R}$  is for sufficiently small  $\rho$*   
324 *continuously differentiable. Let  $p$  denote the adjoint state according to [Definition 15](#)*  
325 *with end data*

$$326 \quad \bar{\psi}_y(x) := \int_0^1 \psi_y(y(\bar{t}, x+; \bar{w})) + \tau[y(\bar{t}, x; \bar{w}), y_d(x+) + \tau[y_d(x)]] \, d\tau.$$

327 Using [\(14\)](#), the derivative of  $\hat{J}(w)$  in a direction  $\delta w \in W$  is given by

$$328 \quad \hat{J}'(w) \cdot \delta w = R'(w)\delta w + (p, g_{u_1}(\cdot, y, u_1)\delta u_1)_{2, \Omega_{\bar{t}}} + \sum_{i=1}^{n_x+1} (p(0, \cdot), \delta u_i^0)_{2, I_i^0(w)}$$

$$329 \quad + \sum_{i \in I_{s,0}(w)} p(0, x_i^0)[u_0(x_i^0)]\delta x_i^0 - \sum_{i \in I_{r,0}(w)} p_i^{r,0}\delta x_i^0 + \sum_{l \in \{a,b\}} \sigma(l) \left( \sum_{\substack{i \in I_{r,l}(w): \\ t_i^l \leq \bar{t}}} p_i^{r,l}\delta t_i^l + \right.$$

$$330 \quad \left. \sum_{i=1}^{n_{t,l}+1} (p(\cdot, l_{\sigma(l)}), f'(u_i^{B,l})\delta u_i^{B,l})_{2, I_{B,l}^i(w) \cap ]0, \bar{t}[} + \sum_{\substack{i \in I_{s,l}(w): \\ t_i^l \leq \bar{t}}} p(t_i^l, l_{\sigma(l)})[f(y(t_i^l, l_{\sigma(l)}; w))]\delta t_i^l \right)$$

331

332 **3. The optimal control problem.** In this section we will prove existence of an  
333 optimal solution and necessary optimality conditions for [\(P\)](#).

334 **3.1. Existence of globally optimal solutions.**

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

338 *Proof.* We show compactness of the set  $\tilde{W}_{\text{ad}} := \{w \in W_{\text{ad}} : y(\bar{t}, x; w) \leq \bar{y}(x) \, \forall x \in$   
339  $[a, b]\}$ . Since  $\tilde{W}_{\text{ad}}$  is by assumption non-empty, we can consider a sequence  $(w_n)_{n \in \mathbb{N}} \subseteq$   
340  $\tilde{W}_{\text{ad}} \subseteq W$ . Due to the compactness of  $W_{\text{ad}}$  in  $W$ , there exists a subsequence, again  
341 denoted by  $(w_n)_{n \in \mathbb{N}}$ , converging to some  $\bar{w} \in W_{\text{ad}}$  w.r.t.  $\|\cdot\|_W$ . [Corollary 7](#) implies  
342 that the sequence  $(y(\bar{t}, \cdot; w_n))_{n \in \mathbb{N}}$  converges in  $L^1(a, b)$  to  $y(\bar{t}, \cdot, \bar{w})$  and hence there  
343 exists a subsequence, again denoted by  $(y(\bar{t}, \cdot; w_n))_{n \in \mathbb{N}}$ , converging pointwise almost  
344 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  
345  $x \in [a, b]$  and all  $n \in \mathbb{N}$ , we obtain that

$$346 \quad (18) \quad y(\bar{t}, x, \bar{w}) \leq \bar{y}(x) \quad \text{for a.a. } x \in [a, b].$$

348 In order to show that [\(18\)](#) holds for all  $x \in [a, b]$ , let  $\hat{x} \in (a, b)$  be arbitrary. Due to  
349 [\(18\)](#), we can choose a sequence  $(x_n)_{n \in \mathbb{N}}$  with  $x_n \nearrow \hat{x}$  for  $n \rightarrow \infty$  such [\(18\)](#) holds for  
350 all  $x = x_n$ . By [Convention 5](#), we obtain

$$351 \quad y(\bar{t}, \hat{x}; \bar{w}) = \lim_{n \rightarrow \infty} y(\bar{t}, x_n; \bar{w}) \leq \lim_{n \rightarrow \infty} \bar{y}(x_n) = \bar{y}(\hat{x}),$$

352

353 where the last equality holds due to the continuity of  $\bar{y}(\cdot)$ . Since  $y(\bar{t}, a, \bar{w}) = y(\bar{t}, a+, \bar{w})$   
354 by [Convention 5](#), we can argue analogously with a sequence  $x_n \searrow a$  satisfying [\(18\)](#)

355 and deduce  $y(\bar{t}, a, \bar{w}) \leq \bar{y}(a)$ . Thus, (18) holds for all  $x \in [a, b]$  yielding that  $\bar{w} \in \tilde{W}_{\text{ad}}$ .  
 356 Therefore,  $\tilde{W}_{\text{ad}}$  is compact. We now consider a sequence  $(w_n)_{n \in \mathbb{N}} \subseteq \tilde{W}_{\text{ad}}$  satisfying

$$357 \quad \hat{J}(w_n) \rightarrow \inf_{w \in \tilde{W}_{\text{ad}}} \hat{J}(w) \quad \text{for } k \rightarrow \infty.$$

358  
 359 Since  $\tilde{W}_{\text{ad}}$  is compact, there exists a convergent subsequence, again denoted by  
 360  $(w_n)_{n \in \mathbb{N}}$ , with  $w_n \rightarrow \bar{w} \in \tilde{W}_{\text{ad}}$ . Since  $\hat{J}$  is Lipschitz-continuous w.r.t.  $w$  by [Corollary 8](#),  
 361 we obtain  $\hat{J}(\bar{w}) = \inf_{w \in \tilde{W}_{\text{ad}}} \hat{J}(w)$  and hence,  $\bar{w}$  is a global minimum for (P).  $\square$

362 **3.2. Reformulation of the state variable.** In order to derive necessary opti-  
 363 mality conditions we need a constraint qualification which requires due to (2) that  
 364  $y(\bar{t}, \cdot; w)$  is an element of  $L^\infty([a, b])$ . Difficulties arise from the fact that the control-  
 365 to-state-mapping is not continuous to  $L^\infty$ . [Theorem 17](#) yields that for  $w \in W$   
 366 satisfying (ND),  $y(\bar{t}, \cdot; w)|_{[a, b]}$  can be rewritten according to (17). Hence, we introduce  
 367  $(y_1(w), \dots, y_{K+1}(w), x_1(w), \dots, x_K(w))$  as new state variables, where

$$368 \quad (19) \quad y_k(\lambda; w) := Y_k(\bar{t}, x_{k-1}(w) + \lambda(x_k(w) - x_{k-1}(w)); w), \quad \lambda \in [0, 1].$$

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

$$372 \quad (20) \quad y_k(\lambda; w) \leq \bar{y}(\bar{t}, x_{k-1}(w) + \lambda(x_k(w) - x_{k-1}(w))) =: \bar{y}_k(\lambda; w), \lambda \in [0, 1].$$

374

375 **LEMMA 21.** *Let (A1) and (A2) hold and consider some  $\bar{w} \in W$  satisfying (ND).*  
 376 *Then we have*

$$377 \quad y(\bar{t}, \cdot; \bar{w}) \leq \bar{y}(\cdot) \text{ on } [a, b] \iff y_k(\cdot; \bar{w}) \leq \bar{y}_k(\cdot; \bar{w}) \text{ on } [0, 1] \forall k = 1, \dots, K + 1$$

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

382 **THEOREM 22** (Continuous differentiability of the state). *Let (A1) and (A2) hold*  
 383 *and consider some  $\bar{w} \in W$  satisfying (ND). Then the mapping*

$$384 \quad (21) \quad w \in B_\rho^W(\bar{w}) \mapsto (y_1(w), \dots, y_{K+1}(w), x_1(w), \dots, x_K(w)) \in C([0, 1]^{K+1} \times \mathbb{R}^K$$

386 *is for  $\rho > 0$  small enough well-defined and continuously differentiable, where*

$$387 \quad (22) \quad \frac{\partial}{\partial w} y_k(\lambda; w) \delta w = \frac{\partial}{\partial w} Y_k(\bar{t}, \lambda x_k(w) + (1 - \lambda)x_{k-1}(w); w) \delta w$$

$$388 \quad + \frac{\partial}{\partial x} Y_k(\bar{t}, \lambda x_k(w) + (1 - \lambda)x_{k-1}(w); w) \cdot \left[ \lambda \frac{\partial}{\partial w} x_k(w) + (1 - \lambda) \frac{\partial}{\partial w} x_{k-1}(w) \right] \delta w$$

389 *and the derivatives of  $x_k(\cdot)$  can be computed according to [Remark 18](#).*

390 *Proof.* This theorem is a consequence of [Theorem 17](#) and (19).  $\square$

391 **3.3. First order necessary optimality conditions.** Our aim is to derive  
 392 necessary optimality conditions for (P). To this end, we consider a general problem of  
 393 the form

$$394 \quad (23) \quad \min_{z \in Z} f(z) \quad \text{subject to} \quad G(z) \in \mathcal{K}, \quad z \in \mathcal{C}$$

396 and recall the following result.

397 THEOREM 23 (Karush-Kuhn-Tucker conditions). Consider a local solution  $\bar{z}$  of  
 398 (23) and assume that the mappings  $f : Z \rightarrow \mathbb{R}$  and  $G : Z \rightarrow V$  are continuously  
 399 differentiable in  $\bar{z} \in Z$  with Banach spaces  $Z$  and  $V$ . We further assume that  $\mathcal{C} \subset Z$  is  
 400 closed, convex and non-empty and  $\mathcal{K} \subset V$  is a closed convex cone. Finally, we assume  
 401 that Robinson's CQ is satisfied at  $\bar{z} \in Z$ , i.e. it holds that

$$403 \quad (24) \quad 0 \in \text{int} (G(\bar{z}) + G'(\bar{z})(\mathcal{C} - \bar{z}) - \mathcal{K}).$$

404 Then there exists a Lagrange multiplier  $\bar{q} \in V^*$  such that

$$405 \quad G(\bar{z}) \in \mathcal{K}, \quad \bar{q} \in \mathcal{K}^\circ := \{q \in V^* : \langle q, v \rangle_{V^*, V} \leq 0 \quad \forall v \in \mathcal{K}\}$$

$$406 \quad \langle \bar{q}, G(\bar{z}) \rangle_{V^*, V} = 0, \quad \bar{z} \in \mathcal{C}, \quad \langle f'(\bar{z}) + G'(\bar{z})^* \bar{q}, z - \bar{z} \rangle_{Z^*, Z} \geq 0 \quad \forall z \in \mathcal{C},$$

408 where  $V^*$  denotes the dual space of  $V$  and  $\mathcal{K}^\circ$  the polar cone of  $\mathcal{K}$ .

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

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

$$411 \quad (25) \quad Z = W, \quad V = C([0, 1])^{K+1} \times \mathbb{R}^K,$$

$$412 \quad (26) \quad G_k(w) = y_k(\lambda; w) - \bar{y}_k(\lambda; w), \quad k = 1, \dots, K+1,$$

$$413 \quad (27) \quad G_{k+K+1}(w) = x_k(w), \quad k = 1, \dots, K,$$

$$414 \quad (28) \quad \mathcal{K} = C([0, 1], (-\infty, 0])^{K+1} \times \mathbb{R}^K, \quad \text{and } \mathcal{C} = W_{\text{ad}}.$$

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

418 LEMMA 24. The polar cone of  $\mathcal{K}$  can be characterized as follows:

$$419 \quad (29) \quad q \in \mathcal{K}^\circ \iff q = (\mu_1, \dots, \mu_{K+1}, 0, \dots, 0),$$

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

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

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

$$426 \quad (30) \quad y_k(\cdot, \bar{w}) \leq \bar{y}_k(\cdot, \bar{w}) \quad \text{on } [0, 1] \quad \forall k = 1, \dots, K+1$$

$$427 \quad (31) \quad \sum_{k=1}^{K+1} \int_{[0, 1]} (y_k(\lambda, \bar{w}) - \bar{y}_k(\lambda, \bar{w})) d\mu_k(\lambda) = 0$$

$$428 \quad (32) \quad \hat{J}'(\bar{w})(w - \bar{w}) + \sum_{k=1}^{K+1} \int_{[0, 1]} \frac{\partial}{\partial w} (y_k(\lambda, \bar{w}) - \bar{y}_k(\lambda, \bar{w})) (w - \bar{w}) d\mu_k(\lambda) \geq 0 \quad \forall w \in W_{\text{ad}}.$$

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

433 LEMMA 26. Let (A1) and (A2) hold and consider some  $\bar{w} \in W_{\text{ad}}$  satisfying (ND).  
 434 Then (24) is satisfied if and only if there exists  $\tilde{w} \in W_{\text{ad}}$  such that for a constant  $\varepsilon > 0$   
 435 the following assertion is valid for all  $k = 1, \dots, K+1$ :

$$436 \quad (33) \quad y_k(\lambda, \bar{w}) - \bar{y}_k(\lambda, \bar{w}) + \frac{\partial}{\partial w} (y_k(\lambda, \bar{w}) - \bar{y}_k(\lambda, \bar{w})) (\tilde{w} - \bar{w}) \leq -\varepsilon \quad \forall \lambda \in [0, 1]$$

438 *Proof.* From (24) we can directly deduce (33) and vice versa.  $\square$

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

$$443 \quad (34) \quad \varphi_{k;\bar{w}} : [0, 1] \rightarrow [x_{k-1}(\bar{w}), x_k(\bar{w})], \quad \lambda \mapsto x_{k-1}(\bar{w}) + \lambda(x_k(\bar{w}) - x_{k-1}(\bar{w})),$$

$$444 \quad (35) \quad \varphi_{k;\bar{w}}^{-1} : [x_{k-1}(\bar{w}), x_k(\bar{w})] \rightarrow [0, 1], \quad x \mapsto \frac{x - x_{k-1}(\bar{w})}{x_k(\bar{w}) - x_{k-1}(\bar{w})}.$$

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

$$447 \quad Y_k(\bar{t}, \varphi_{k;\bar{w}}(\cdot), \bar{w}) \leq \bar{y}(\varphi_{k;\bar{w}}(\cdot)) \quad \text{on } [0, 1] \quad \forall k = 1, \dots, K+1$$

$$448 \quad \sum_{k=1}^{K+1} \int_{[0,1]} (Y_k(\bar{t}, \varphi_{k;\bar{w}}(\lambda), \bar{w}) - \bar{y}(\varphi_{k;\bar{w}}(\lambda))) d\mu_k(\lambda) = 0$$

$$449 \quad \hat{J}'(\bar{w})(w - \bar{w}) + \sum_{k=1}^{K+1} \int_{[0,1]} \left[ \frac{\partial}{\partial x} (Y_k(\bar{t}, \varphi_{k;\bar{w}}(\lambda); \bar{w}) - \bar{y}(\varphi_{k;\bar{w}}(\lambda))) \right.$$

$$450 \quad \cdot (\varphi_{k;\bar{w}}^{-1}(\varphi_{k;\bar{w}}(\lambda)) \frac{\partial}{\partial w} x_k(\bar{w})(w - \bar{w}) + (1 - \varphi_{k;\bar{w}}^{-1}(\varphi_{k;\bar{w}}(\lambda))) \frac{\partial}{\partial w} x_{k-1}(\bar{w})(w - \bar{w}))$$

$$451 \quad \left. + \frac{\partial}{\partial w} (Y_k(\bar{t}, \varphi_{k;\bar{w}}(\lambda); \bar{w}) - \bar{y}(\varphi_{k;\bar{w}}(\lambda))) \cdot (w - \bar{w}) \right] d\mu_k(\lambda) \geq 0 \quad \forall w \in W_{\text{ad}}.$$

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

$$456 \quad (36) \quad Y_k(\bar{t}, \cdot, \bar{w}) \leq \bar{y}(\cdot) \quad \text{on } I_k \quad \forall k = 1, \dots, K+1$$

$$457 \quad (37) \quad \sum_{k=1}^{K+1} \int_{I_k} (Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)) d\bar{\mu}_k(x) = 0$$

$$458 \quad (38) \quad \hat{J}'(\bar{w})(w - \bar{w}) + \sum_{k=1}^{K+1} \left[ \int_{I_k} \frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \frac{x - x_{k-1}(\bar{w})}{x_k(\bar{w}) - x_{k-1}(\bar{w})} d\bar{\mu}_k(x) \right.$$

$$459 \quad \cdot \frac{\partial}{\partial w} x_k(\bar{w})(w - \bar{w}) + \int_{I_k} \frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \frac{x_k(\bar{w}) - x}{x_k(\bar{w}) - x_{k-1}(\bar{w})} d\bar{\mu}_k(x)$$

$$460 \quad \cdot \left. \frac{\partial}{\partial w} x_{k-1}(\bar{w})(w - \bar{w}) + \int_{I_k} \frac{\partial}{\partial w} Y_k(\bar{t}, x, \bar{w})(w - \bar{w}) d\bar{\mu}_k(x) \right] \geq 0 \quad \forall w \in W_{\text{ad}}.$$

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

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

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

467 **LEMMA 28.** *Let (A1) and (A2) hold and consider some  $\bar{w} \in W_{\text{ad}}$  and nonnegative*  
468 *measures  $\mu_k \in \mathcal{M}(I_k)$ ,  $k = 1, \dots, K+1$  such that (36) and (37) are satisfied. Then*

469 for all measurable  $A \subseteq (x_{k-1}(\bar{w}), x_k(\bar{w}))$  the following holds

$$470 \quad (39) \quad \int_A \frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \left( \frac{x - x_{k-1}(\bar{w})}{x_k(\bar{w}) - x_{k-1}(\bar{w})} \right) d\bar{\mu}_k(x) = 0,$$

$$471 \quad (40) \quad \int_A \frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \left( \frac{x_k(\bar{w}) - x}{x_k(\bar{w}) - x_{k-1}(\bar{w})} \right) d\bar{\mu}_k(x) = 0.$$

473 *Proof.* For arbitrary  $k \in \{1, \dots, K+1\}$  and measurable  $A \subseteq (x_{k-1}(\bar{w}), x_k(\bar{w}))$  we  
 474 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  
 475 observe that  $A_1$  and  $A_2$  are both measurable due to the regularity of  $Y_k(\bar{t}, \cdot; \bar{w})$  and  
 476  $\bar{y}(\cdot)$ . From (36) we can deduce that  $A = A_1 \cup A_2$ . We firstly prove that

$$477 \quad (41) \quad \bar{\mu}_k(A_1) = 0, \quad \frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \Big|_{A_2} \equiv 0.$$

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

$$481 \quad \sum_{k=1}^{K+1} \int_{I_k} (Y_k(\bar{t}, x, \bar{w})) - \bar{y}(x) d\bar{\mu}_k(x) \leq \int_{I_k \cap A_1} (Y_k(\bar{t}, x, \bar{w})) - \bar{y}(x) d\bar{\mu}_k(x) < 0.$$

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

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

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

$$488 \quad \int_A \frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \left( \frac{x - x_{k-1}(\bar{w})}{x_k(\bar{w}) - x_{k-1}(\bar{w})} \right) d\bar{\mu}_k(x) = 0$$

490 Analogously, we can prove (40).  $\square$

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

492 **COROLLARY 29.** *Let (A1) and (A2) hold and consider some  $\bar{w} \in W_{\text{ad}}$  and nonneg-*  
 493 *ative measures  $\mu_k \in \mathcal{M}(I_k)$ ,  $k = 1, \dots, K+1$ , satisfying (36), (37) and (38). Then*  
 494 *(38) can also be written as:*

$$495 \quad (42) \quad \hat{J}'(\bar{w})(w - \bar{w}) + \sum_{k=1}^{K+1} \left[ \frac{\partial}{\partial x} [Y_k(\bar{t}, x_k(\bar{w}), \bar{w}) - \bar{y}(x_k(\bar{w}))] \cdot \bar{\mu}_k(\{x_k(\bar{w})\}) \right. \\
 496 \quad \cdot \frac{\partial}{\partial w} x_k(\bar{w})(w - \bar{w}) + \frac{\partial}{\partial x} [Y_k(\bar{t}, x_{k-1}(\bar{w}), \bar{w}) - \bar{y}(x_{k-1}(\bar{w}))] \cdot \bar{\mu}_k(\{x_{k-1}(\bar{w})\}) \\
 497 \quad \left. \cdot \frac{\partial}{\partial w} x_{k-1}(\bar{w})(w - \bar{w}) + \int_{I_k} \frac{\partial}{\partial w} Y_k(\bar{t}, x, \bar{w})(w - \bar{w}) d\bar{\mu}_k(x) \right] \geq 0, \quad \forall w \in W_{\text{ad}}$$

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

501 **THEOREM 30.** *Let (A1) and (A2) hold and let  $\bar{w} \in W_{\text{ad}}$  be a local solution*  
 502 *of (P) satisfying (ND) and (24). Then (36)-(38) and (42) are still valid, if we*  
 503 *replace  $Y_k(\bar{t}, x; \bar{w})$ ,  $\frac{\partial}{\partial w} Y_k(\bar{t}, x; \bar{w})$  and  $\frac{\partial}{\partial x} Y_k(\bar{t}, x; \bar{w})$  by  $y(\bar{t}, x; \bar{w})$ ,  $\frac{\partial}{\partial w} y(\bar{t}, x-; \bar{w})$  and*  
 504  *$\frac{\partial}{\partial x} y(\bar{t}, x-; \bar{w})$ , respectively.*

505 LEMMA 31. Let (A1) and (A2) hold and consider some  $\bar{w} \in W_{\text{ad}}$  satisfying (ND).  
 506 Then (24) is satisfied if and only if there exists a constant  $\varepsilon > 0$  and  $\delta w \in W_{\text{ad}} - \bar{w}$   
 507 such that for all  $x \in I_k$  and  $k = 1, \dots, K + 1$  it holds that

$$508 \quad (43) \quad Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x) + \frac{\partial}{\partial w} Y_k(\bar{t}, x, \bar{w}) \delta w + \frac{\partial}{\partial x} [Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)] \cdot \\ 509 \quad \left( \frac{x - x_{k-1}(\bar{w})}{x_k(\bar{w}) - x_{k-1}(\bar{w})} \frac{\partial}{\partial w} x_k(\bar{w}) \delta w + \frac{x_k(\bar{w}) - x}{x_k(\bar{w}) - x_{k-1}(\bar{w})} \frac{\partial}{\partial w} x_{k-1}(\bar{w}) \delta w \right) \leq -\varepsilon$$

511 *Proof.* This result follows directly from Lemma 26 and (19).  $\square$

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

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

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

533 The proof consists of two steps: In **Step 1**, we will derive representations for the  
 534 terms  $\frac{\partial}{\partial w} Y_1(\bar{t}, \cdot; \bar{w})$ ,  $\frac{\partial}{\partial w} Y_2(\bar{t}, \cdot; \bar{w})$ ,  $\frac{\partial}{\partial w} Y_3(\bar{t}, \cdot; \bar{w})$ ,  $\frac{\partial}{\partial w} x_1(\bar{w})$  and  $\frac{\partial}{\partial w} x_2(\bar{w})$ . In **Step 2**, we  
 535 will choose  $\delta w \in W_{\text{ad}} - \bar{w}$  such that (43) is satisfied for all  $k = 1, 2, 3$ .

536 **Step 1:** Due to  $g|_{\Omega_{\bar{t}}} \equiv 0$  it holds that  $Y_1(\bar{t}, x, \bar{w}) = f'^{-1}\left(\frac{x-a}{\bar{t}-\bar{t}_j^a}\right)$  yielding

$$537 \quad (44) \quad \frac{\partial}{\partial x} Y_1(\bar{t}, x, \bar{w}) \delta x = \frac{\delta x}{(\bar{t} - \bar{t}_j^a) \cdot f''(f'^{-1}(\frac{x-a}{\bar{t}-\bar{t}_j^a}))},$$

$$538 \quad (45) \quad \frac{\partial}{\partial w} Y_1(\bar{t}, x, \bar{w}) \delta w = \frac{(x-a) \cdot \delta \bar{t}_j^a}{(\bar{t} - \bar{t}_j^a)^2 f''(f'^{-1}(\frac{x-a}{\bar{t}-\bar{t}_j^a}))}.$$

540 Since the minimal backward characteristic through  $(\bar{t}, x_2(\bar{w}))$  ends in a continuity  
 541 point  $\bar{t}^a \in (\bar{t}_{j-1}^a, \bar{t}_j^a)$  of  $u_{B,a}(\cdot; \bar{w})$  and  $\bar{w} \in W_{\text{ad}}$  satisfies (ND), [33, Lemma 6.2.7] yields

$$542 \quad (46) \quad \frac{\partial}{\partial x} Y_2(\bar{t}, x, \bar{w}) \delta x = \frac{(\bar{u}_j^{B,a})'(\Phi(\cdot)) \cdot \delta x}{f''(\bar{u}_j^{B,a}(\Phi(\cdot))) (\bar{u}_j^{B,a})'(\Phi(\cdot)) (\bar{t} - \Phi(\cdot)) - f'(\bar{u}_j^{B,a}(\Phi(\cdot)))}$$

$$543 \quad (47) \quad \frac{\partial}{\partial w} Y_2(\bar{t}, x, \bar{w}) \delta w = - \frac{f'(\bar{u}_j^{B,a}(\Phi(\cdot))) \cdot \delta u_j^{B,a}(\Phi(\cdot))}{f''(\bar{u}_j^{B,a}(\Phi(\cdot))) (\bar{u}_j^{B,a})'(\Phi(\cdot)) (\bar{t} - \Phi(\cdot)) - f'(\bar{u}_j^{B,a}(\Phi(\cdot)))}$$

544

545 where  $(\cdot) = (\bar{t}, x, \bar{w})$ ,  $(\bar{u}_j^{B,a})'(\cdot)$  denotes the derivative of  $\bar{u}_j^{B,a}(\cdot)$  and  $\Phi(\bar{t}, x, \bar{w})$  is the  
546 unique solution of the equation  $x = f'(\bar{u}_j^{B,a}(\phi))(\bar{t} - \phi) + a$  w.r.t.  $\phi$ . We observe that

$$547 \quad (48) \quad \Phi(\bar{t}, x, \bar{w}) \in [\bar{t}^a, \bar{t}_j^a] \quad \text{for all } x \in [x_1(\bar{w}), x_2(\bar{w})].$$

549 From [33, Lemma 6.2.7 (i)], we can further deduce that there exists constants  $\delta_0, \beta > 0$   
550 such that for all  $x \in (x_1(\bar{w}) - \delta_0, x_2(\bar{w}) + \delta_0)$  the following is satisfied:

$$551 \quad (49) \quad q_1(\cdot) := f''(\bar{u}_j^{B,a}(\Phi(\cdot)))(\bar{u}_j^{B,a})'(\Phi(\cdot))(\bar{t} - \Phi(\cdot)) - f'(\bar{u}_j^{B,a}(\Phi(\cdot))) < -\beta$$

553 Now, we will have a closer look at the term  $Y_3$ . According to [33, Lemma 6.2.1], since  
554 the maximal backward characteristic through  $(\bar{t}, x_2(\bar{w}))$  ends in a point  $\tilde{x} \in (\bar{x}_{m-1}^0, \bar{x}_m^0)$   
555 where the initial data  $u_0(\cdot; \bar{w})$  are smooth, it holds that

$$556 \quad (50) \quad \frac{\partial}{\partial x} Y_3(\bar{t}, x, \bar{w}) \delta x = \frac{(\bar{u}_m^0)'(Z(\cdot)) \cdot \delta x}{f''(\bar{u}_m^0(Z(\cdot)))(\bar{u}_m^0)'(Z(\cdot))\bar{t} + 1},$$

$$557 \quad (51) \quad \frac{\partial}{\partial w} Y_3(\bar{t}, x, \bar{w}) \delta w = \frac{\delta u_m^0(Z(\cdot))}{f''(\bar{u}_m^0(Z(\cdot)))(\bar{u}_m^0)'(Z(\cdot))\bar{t} + 1},$$

559 where  $Z(\cdot) = Z(\bar{t}, x, \bar{w})$  denotes the unique solution of the equation  $x = f'(\bar{u}_m^0(z))\bar{t} + z$   
560 w.r.t.  $z$ . Then for (a possibly smaller)  $\delta_0 > 0$ , we can deduce from [33, Lemma 6.2.1  
561 (i)] that for all  $x \in (x_2(\bar{w}) - \delta_0, b)$  it holds that

$$562 \quad (52) \quad q_2(\cdot) = f''(\bar{u}_m^0(Z(\cdot)))(\bar{u}_m^0)'(Z(\cdot))\bar{t} + 1 > \beta$$

564 for a (possibly smaller) positive constant  $\beta > 0$ .

565 Next, we examine the term  $x_1(\bar{w})$ . Since  $(\bar{t}, x_1(\bar{w}))$  lies on the right boundary of a  
566 rarefaction wave and the source term  $g$  is by assumption equal to zero, we obtain that  
567  $x_1(\bar{w}) = f'(\bar{u}_j^{B,a}(\bar{t}_j^a)) \cdot (\bar{t} - \bar{t}_j^a) + a$  and hence

$$568 \quad (53) \quad \frac{\partial}{\partial w} x_1(\bar{w}) \delta w = f''(\bar{u}_j^{B,a}(\bar{t}_j^a))(\bar{t} - \bar{t}_j^a) [\delta u_j^{B,a}(\bar{t}_j^a) + (\bar{u}_j^{B,a})'(\bar{t}_j^a) \delta \bar{t}_j^a] - f'(\bar{u}_j^{B,a}(\bar{t}_j^a)) \delta \bar{t}_j^a.$$

570 The derivative of the shock position  $x_2(\bar{w})$  w.r.t.  $w$  is due to [Theorem 16](#) given by

$$571 \quad (54) \quad \frac{\partial}{\partial w} x_2(\bar{w}) \delta w = \sum_{k=1}^m (p(0, \cdot), \delta u_k^0)_{2, I_k^0 \cap ]0, \tilde{x}[} + \sum_{k=1}^j (p(\cdot, a), f'(\bar{u}_k^{B,a}) \delta u_k^{B,a})_{2, I_{B,a}^k \cap ]0, \tilde{x}[}$$

$$572 \quad + \sum_{k \in I_{s,0}: \bar{x}_k^0 \in ]0, \tilde{x}[} p(0, \bar{x}_k^0) [\bar{u}_0(\bar{x}_k^0)] \delta x_k + \sum_{k \in I_{s,a}(\bar{w}): \bar{t}_k^a \in ]0, \tilde{x}[} p(\bar{t}_k^a, a) [f(y(\bar{t}_k^a, a+; \bar{w}))] \delta \bar{t}_k^a$$

$$573 \quad - \sum_{k \in I_{r,0}: \bar{x}_k^0 \in ]0, \tilde{x}[} p_k^{r,0} \delta x_k^0 + \sum_{k \in I_{r,a}(\bar{w}): \bar{t}_k^a \in ]0, \tilde{x}[} p_k^{r,a} \delta \bar{t}_k^a,$$

575 where  $p$  is the adjoint state with end data  $p(\bar{t}, \cdot) = \mathbf{1}_{x_2(\bar{w})}(\cdot) \frac{1}{[y(\bar{t}, x_2(\bar{w}); \bar{w})]}$ , which is  
576 according to [Definition 15](#) given by  $p(t, x)|_D \equiv \frac{1}{[y(\bar{t}, x_2(\bar{w}); \bar{w})]} > 0$ , with  $D = \{(t, x) \in$   
577  $\Omega_{\bar{t}} \setminus D_- : \xi_-(t) \leq x \leq \xi_+(t)\}$ , and equal to zero on  $\Omega_{\bar{t}} \setminus D$ .

578 **Step 2:** Our goal is to choose  $\delta w$  such that (43) is satisfied for all  $k \in \{1, 2, 3\}$ .  
579 To this end, we choose all components of  $\delta w$ , except  $\delta u_j^{B,a}$  and  $\delta u_m^0$ , equal to zero.

580 Let  $\delta u_j^{B,a}$  and  $\delta u_m^0$  be given by

$$581 \quad (55) \quad \delta u_j^{B,a}(t) = \begin{cases} 0 & \text{if } \bar{t}_{j-1}^a \leq t < \bar{t}^a - \rho_1, \\ \phi_1(t) & \text{if } \bar{t}^a - \rho_1 \leq t < \bar{t}^a, \\ -\varepsilon_0 & \text{if } \bar{t}^a \leq t < \bar{t}_j^a - \rho_2, \\ \phi_2(t) & \text{if } \bar{t}_j^a - \rho_2 \leq t < \bar{t}_j^a, \\ \frac{-\varepsilon_0}{f''(\bar{u}_j^{B,a}(\bar{t}_j^a))(\bar{t} - \bar{t}_j^a)N} & \text{if } t = \bar{t}_j^a, \end{cases} \quad \delta u_m^0(x) = \begin{cases} 0 & \text{if } x < \tilde{x} - \rho_3, \\ \phi_3(x) & \text{if } \tilde{x} - \rho_3 \leq x < \tilde{x}, \\ -\varepsilon_0 & \text{if } \tilde{x} \leq x, \end{cases}$$

582

583 with positive constants  $\varepsilon_0, \rho_1, \rho_2$  and  $\rho_3$ . Moreover, let  $N \in \mathbb{N}$  be chosen such that  
 584  $f''(\bar{u}_j^{B,a}(\bar{t}_j^a))(\bar{t} - \bar{t}_j^a) \cdot N > 1$  holds. The constants  $\rho_1, \rho_2$  and  $\rho_3$  will be identified later.  
 585 Finally, let  $\phi_1, \phi_2$  and  $\phi_3$  be chosen such that  $-\varepsilon_0 \leq \phi_i \leq 0$ .

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

$$591 \quad (56) \quad \frac{\partial}{\partial x} Y_1(\bar{t}, x_1(\bar{w}), \bar{w}) \cdot \frac{\partial}{\partial w} x_1(\bar{w}) \delta w = \frac{-\varepsilon_0}{N(\bar{t} - \bar{t}_j^a) f''(f'^{-1}(\frac{x_1(\bar{w})}{\bar{t} - \bar{t}_j^a}))} \leq \frac{-\varepsilon_0}{N \bar{t} M_{f''}} =: -\varepsilon_{11},$$

592

593 where the inequality holds due to  $0 < m_{f''} \leq f''(\cdot) \leq M_{f''}$ . (56) and the continuity  
 594 of the left term of (43) w.r.t.  $x$  yield that there exists a constant  $\delta_1 > 0$  such that  
 595 the left term of (43) is smaller than  $-\frac{\varepsilon_{11}}{2}$  for all  $x \in (x_1 - \delta_1, x_1]$ . Since  $Y_1(\bar{t}, \cdot; \bar{w})$  is  
 596 strictly monotone increasing, for some  $\varepsilon_{12} > 0$  it holds that

$$597 \quad (57) \quad Y_1(\bar{t}, x; \bar{w}) - \bar{y} \leq -\varepsilon_{12} \quad \text{for all } x \in [a, x_1(\bar{w}) - \delta_1].$$

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

$$601 \quad (58) \quad -\varepsilon_{12} - \frac{\varepsilon_0}{N \cdot (\bar{t} - \bar{t}_j^a) \cdot f''(f'^{-1}(\frac{x}{\bar{t} - \bar{t}_j^a}))} \frac{x - a}{x_1(\bar{w}) - a} \leq -\varepsilon_{12}.$$

602

603 Choosing  $\varepsilon = \varepsilon_1 := \min\{\frac{\varepsilon_{11}}{2}, \varepsilon_{12}\}$  yields that (43) is satisfied for  $k = 1$ .

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

$$605 \quad \frac{\partial}{\partial w} Y_2(\bar{t}, x, \bar{w}) \delta w + \frac{\partial}{\partial x} Y_2(\bar{t}, x, \bar{w}) \left[ \frac{x_2(\bar{w}) - x}{x_2(\bar{w}) - x_1(\bar{w})} \frac{\partial}{\partial w} x_1(\bar{w}) \delta w \right. \\ 606 \quad (59) \quad \left. + \frac{x - x_1(\bar{w})}{x_2(\bar{w}) - x_1(\bar{w})} \frac{\partial}{\partial w} x_2(\bar{w}) \delta w \right].$$

607

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

$$609 \quad \frac{-f'(\bar{u}_j^{B,a}(\bar{t}_j^a)) \cdot \delta u_j^{B,a}(\bar{t}_j^a) + (\bar{u}_j^{B,a})'(\bar{t}_j^a) \cdot f''(\bar{u}_j^{B,a}(\bar{t}_j^a))(\bar{t} - \bar{t}_j^a) \delta u_j^{B,a}(\bar{t}_j^a)}{f''(\bar{u}_j^{B,a}(\bar{t}_j^a))(\bar{u}_j^{B,a})'(\bar{t}_j^a)(\bar{t} - \bar{t}_j^a) - f'(\bar{u}_j^{B,a}(\bar{t}_j^a))} = \delta u_j^{B,a}(\bar{t}_j^a)$$

610

611 Hence, the continuity of (59) w.r.t.  $x$  yields that (59) is on  $[x_1(\bar{w}), x_1(\bar{w}) + \delta]$  bounded  
 612 from above by  $\frac{\delta u_j^{B,a}(\bar{t}_j^a)}{2} = \frac{-\varepsilon_0}{2f''(\bar{u}_j^{B,a}(\bar{t}_j^a))(\bar{t} - \bar{t}_j^a)N} =: -\varepsilon_{21} < 0$ , where  $\delta > 0$  is sufficiently  
 613 small. We choose  $\rho_2$  such that  $\Phi(x_1(\bar{w}) + \delta) = \bar{t}_j^a - \rho_2$ . We further choose

$$614 \quad \rho_1 = \frac{[y(\bar{t}, x_2(\bar{w}); \bar{w})]}{2N \|f'(\bar{u}_j^{B,a}(\cdot))\|_{\infty, [0, \bar{t}^a]}} > 0 \quad \text{and} \quad \rho_3 = \frac{[y(\bar{t}, x_2(\bar{w}); \bar{w})]}{2N} > 0,$$

615

616 where  $N$  has to be chosen sufficiently large, such that  $\rho_1$  and  $\rho_3$  are sufficiently small so  
 617 that (55) makes sense. Then (54), (55) and  $\|\phi_i(\cdot)\|_\infty \leq \varepsilon_0$  yield that  $|\frac{\partial}{\partial w} x_2(\bar{w})\delta w| \leq \frac{\varepsilon_0}{N}$ .  
 618 Using (47), (49) and (55), we obtain

$$619 \quad \frac{\partial}{\partial w} Y_2(\bar{t}, \cdot; \bar{w})\delta w \leq \frac{-\alpha\varepsilon_0}{\|q_1(\cdot)\|_{\infty, [x_1(\bar{w})+\delta, x_2(\bar{w})]}} \quad \text{on } [x_1(\bar{w}) + \delta, x_2(\bar{w})],$$

621 where  $\alpha$  is the constant in (A2). Due to the boundedness of  $\frac{\partial}{\partial x} Y_2(\bar{t}, x, \bar{w})$  and the  
 622 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  
 623 (59) is on  $[x_1(\bar{w}) + \delta, x_2(\bar{w})]$  bounded from above by  $-\varepsilon_{22} := \frac{-\alpha\varepsilon_0}{2\|q_1(\bar{t}, \cdot, \bar{w})\|_{\infty, [x_1(\bar{w})+\delta, x_2(\bar{w})]}}$ .  
 624 Choosing  $\varepsilon = \varepsilon_2 := \min\{\varepsilon_{21}, \varepsilon_{22}\}$ , (43) is satisfied for  $k = 2$ .

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

$$626 \quad (60) \quad \frac{\partial}{\partial w} Y_3(\bar{t}, \cdot; \bar{w})\delta w \leq \frac{-\varepsilon_0}{\|q_2(\bar{t}, \cdot, \bar{w})\|_{\infty, [x_2(\bar{w}), b]}} \quad \text{on } [x_2(\bar{w}), b].$$

628 Due to the boundedness of  $\frac{\partial}{\partial x} Y_3(\bar{t}, x, \bar{w})$  and the fact that  $|\frac{\partial}{\partial w} x_2(\bar{w})\delta w| \leq \frac{\varepsilon_0}{N}$ , one  
 629 can choose  $N$  large enough such that (2) and (60) yield that the left term of (43)  
 630 is on  $[x_2(\bar{w}), b]$  smaller than  $-\frac{\varepsilon_0}{2\|q_2(\bar{t}, \cdot, \bar{w})\|_{\infty, [x_2(\bar{w}), b]}} =: -\varepsilon_3$ . Hereby,  $\varepsilon_3$  is finite since  
 631  $q_2(\bar{t}, \cdot, \bar{w})$  is continuous on the compact interval  $[x_2(\bar{w}), b]$ . Hence, (43) is satisfied  
 632 for  $k = 3$  if we choose  $\varepsilon = \varepsilon_3$ . Choosing  $\varepsilon := \min\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$ , (43) is satisfied for  
 633  $k = 1, 2, 3$  for the choice of  $\delta w$  in (55). This also remains valid (with  $\varepsilon > 0$ ), if we  
 634 replace  $\delta w$  by  $\tau\delta w$  with a sufficiently small constant  $\tau > 0$ , such that  $\bar{w} + \tau\delta w \in W_{\text{ad}}$ .  
 635 More precisely: If we replace  $\delta w$  by  $\tau\delta w$ , (56) holds with  $-\tau\varepsilon_{11}$  and (58) remains valid  
 636 for the same  $\varepsilon_{12}$  such that choosing  $\varepsilon = \min\{\frac{\tau\varepsilon_{11}}{2}, \varepsilon_{12}\}$  yields that (43) is satisfied  
 637 for  $k = 1$ . Since the term (59) depends linearly on  $\tau$ , for  $k = 2$  (43) is satisfied for  
 638 the choice  $\varepsilon = \tau\varepsilon_2 := \min\{\tau\varepsilon_{21}, \tau\varepsilon_{22}\}$ . For the case  $k = 3$ , choosing  $\varepsilon = \tau\varepsilon_3$ , one  
 639 can see that (43) is valid. Hence, if we choose  $\varepsilon := \min\{\min\{\frac{\tau\varepsilon_{11}}{2}, \varepsilon_{12}\}, \tau\varepsilon_2, \tau\varepsilon_3\}$ , (43)  
 640 remains satisfied for  $k = 1, 2, 3$  if we replace  $\delta w$  in (55) by  $\tau\delta w$ .  $\square$

641 **REMARK 33.** *A careful study of [33, Lemma 6.2.1, Lemma 6.2.7] shows that*  
 642 *Theorem 32 also holds for certain source terms which are not equal to zero on  $\Omega_{\bar{t}}$ ,*  
 643 *e.g., for source terms which only depend on the state  $y$  and satisfy (A1). Furthermore,*  
 644 *using the same arguments as in the proof above, one can show that Theorem 32 is still*  
 645 *valid if the upper bound  $\bar{y}$  is not constant.*

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

$$649 \quad (61) \quad \begin{aligned} \rho_t + \hat{f}(\rho)_x &= 0, & \text{on } \Omega_{\bar{t}} &:= (0, \bar{t}) \times (0, \infty), \\ \rho(0, \cdot) &= \rho_0(\cdot), & \text{on } \Omega &:= (0, \infty), \\ \rho(\cdot, 0+) &= \rho_{B,0}(\cdot), & \text{in the sense of (6)} \end{aligned}$$

651 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  
 652 assume for convenience that  $n = 1$ ,  $v_{\max} = 1$  and  $\rho_{\max} = 1$ . Since our results require  
 653 convex flux functions, but  $\hat{f}$  is concave, we set

$$654 \quad (62) \quad y := -\rho \quad \text{and} \quad f(y) := -\hat{f}(-y) = yv(-y) = y^2 + y.$$

656 We note that  $f$  is uniformly convex. Applying (62) to (61) gives

$$\begin{aligned}
 & y_t + f(y)_x = 0, && \text{on } \Omega_{\bar{t}}, \\
 657 \quad (63) \quad & y(0, \cdot) = u_0(\cdot) := -\rho_0(\cdot), && \text{on } \Omega, \\
 658 \quad & y(\cdot, 0+) = u_{B,0}(\cdot) := -\rho_{B,0}(\cdot), && \text{in the sense of (6)}.
 \end{aligned}$$

659 Our goal is to maximize the flow on the interval  $[0, L]$  at time  $t = \bar{t}$ , i.e.  $\int_0^L \hat{f}(\rho(\bar{t}, x)) \, dx$   
 660 under the condition that the state constraints  $\rho(\bar{t}, \cdot) \leq \bar{\rho} \in \mathbb{R}$  hold on  $[0, L]$ . By (62),  
 661 the corresponding optimal control problem in terms of  $y$  reads

$$\begin{aligned}
 662 \quad (64) \quad & \min J(y(w)) := \int_0^L f(y(\bar{t}, x; w)) \, dx, \\
 663 \quad & \text{s.t. } w \in W_{\text{ad}}, y \text{ satisfies (63) and } y(\bar{t}, \cdot; w) \geq -\bar{\rho} \text{ on } [0, L],
 \end{aligned}$$

664 where the constant lower bound  $-\bar{\rho} < 0$ , the control  $w$  and the set of admissible  
 665 controls  $W_{\text{ad}}$  will be further specified below. We consider fixed piecewise continuously  
 666 differentiable initial data  $u_0$  and boundary data of the form

$$667 \quad (65) \quad u_{B,0}(t; w) = \begin{cases} 0 & \text{if } t \in [0, t^{\text{on}}], \\ u(t) & \text{if } t \in (t^{\text{on}}, t^{\text{off}}], \\ 0 & \text{if } t \in (t^{\text{off}}, \bar{t}] \end{cases}$$

669 with control  $w = (u, t^{\text{on}})$  and fixed  $0 < t^{\text{off}} < \bar{t}$ . We note that this example can be  
 670 extended to the case of more than two switching points, where in addition also the  
 671 off-switching points can be shifted. Introducing suitable node conditions, it would also  
 672 be possible to extend this example to traffic networks.

673 Let

$$674 \quad (66) \quad W_{\text{ad}} = W = \{(u, t^{\text{on}}) \in C^1([t^{\text{on}}, t^{\text{off}}]) \times (0, t^{\text{off}}) : -\frac{1}{2} + \alpha \leq u \leq 0\}$$

675 for some arbitrary small  $\alpha > 0$  and

$$677 \quad (67) \quad -\frac{1}{2} + \alpha < -\bar{\rho} < 0.$$

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

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

685 Next, we note that there is a shock curve  $\eta(t; w)$  emanating from  $(t^{\text{off}}, 0)$ , where  
 686 we set  $x_s(w) := \eta(\bar{t}; w)$ . Our goal is to construct a control  $\bar{w}$  for which the necessary  
 687 optimality conditions (36)-(38) as well as Robinson's CQ (24) are satisfied. To this  
 688 end, it is intuitive to choose the maximal possible inflow of cars, i.e.  $\bar{u} \equiv -\frac{1}{2} + \alpha$ .  
 689 To satisfy (64), let  $\bar{t}^{\text{on}} \in (0, \bar{t}^{\text{off}})$  be such that the maximal backward characteristic  
 690 through  $(\bar{t}, x_s(\bar{w}))$  ends in the inner of the rarefaction wave produced in  $(\bar{t}^{\text{on}}, 0)$  and  
 691  $y(\bar{t}, x_s(\bar{w})+) = -\bar{\rho}$ . Hence,  $(\bar{t}, x_s(\bar{w}))$  is a non-degenerated shock point. Note that if  
 692  $y(\bar{t}, x_s(\bar{w})+) > -\bar{\rho}$ , the control  $\bar{w}$  will not satisfy (36)-(38), since in this case  $(\bar{t}, x_s(\bar{w}))$   
 693 can be shifted closer to  $x = 0$  by a small change of  $\bar{t}^{\text{on}}$ , which would result in a smaller  
 694 cost function value (this can, e.g., be deduced from (69)).

695 In this setting,  $y(\bar{t}, x; \bar{w})$  has one discontinuity and no shock generation points on  
 696  $[0, L]$ . Then, the solution to (63) is in a small neighborhood of  $\bar{w}$  of the form

$$697 \quad (68) \quad y(\bar{t}, x; w) = \begin{cases} Y_1(\bar{t}, x, w) = 0 & \text{if } t \in [0, x_s(w)], \\ Y_2(\bar{t}, x, w) = f'^{-1}\left(\frac{x}{t - t^{on}}\right) & \text{if } t \in (x_s(w), x_r(w)], \\ Y_3(\bar{t}, x, w) = 0 & \text{if } t \in (x_r(w), L], \end{cases}$$

698 where  $x_r(w)$  denotes the right boundary of the rarefaction wave produced in  $(t^{on}, 0)$ .  
 699 In order to make the analysis of this example easier to understand, we provide an  
 700 illustration of the structure of the solution in [Figure 1](#). In (68), we have used the  
 701 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  
 702 the non-degeneracy of  $(\bar{t}, x_s(\bar{w}))$  and the structure of  $y(\bar{t}, \cdot)$  in (68),  $\bar{w}$  satisfies (ND).  
 703 Therefore, by [Theorem 19](#)

$$705 \quad (69) \quad \hat{J}'(\bar{w})(w - \bar{w}) = \int_{t^{on}}^{t^{off}} p(t, 0) f'(\bar{u}(t)) (u(t) - \bar{u}(t)) dt + p^r \cdot (t^{on} - \bar{t}^{on}),$$

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

$$709 \quad p^r = (1 - \bar{\rho}) \int_{f'(-\frac{1}{2} + \alpha)}^{f'(-\bar{\rho})} \frac{z}{2} dz + \int_{f'(-\bar{\rho})}^{f'(0)} \frac{z^2}{2} dz = \frac{1}{3} \bar{\rho}^3 - \frac{\bar{\rho}}{4} + \frac{1}{4} + O(\alpha).$$

710 Inserting this in (69) yields

$$712 \quad (70) \quad \hat{J}'(\bar{w})(w - \bar{w}) = (1 - \bar{\rho}) \cdot \int_{t^{on}}^{t^{off}} f'(\bar{u}(t)) (u(t) - \bar{u}(t)) dt$$

$$713 \quad + \left[ \frac{1}{3} \bar{\rho}^3 - \frac{\bar{\rho}}{4} + \frac{1}{4} + O(\alpha) \right] (t^{on} - \bar{t}^{on}).$$

714 By [Theorem 16](#), the derivative  $\frac{\partial}{\partial w} x_s(\bar{w})$  is given by the right-hand side of (69), but  
 715 with a different adjoint state, i.e. the adjoint state  $p$  with end data  $p(\bar{t}, \cdot) = \mathbf{1}_{x_s(\bar{w})}(\cdot) \frac{1}{\bar{\rho}}$ .  
 716 For this end data, we get  $p(\cdot, 0)|_{(t^{on}, t^{off})} \equiv \frac{1}{\bar{\rho}}$  and

$$717 \quad p^r = \frac{1}{\bar{\rho}} \int_{f'(-\frac{1}{2} + \alpha)}^{f'(-\bar{\rho})} \frac{z}{2} dz = \frac{(1 - 2\bar{\rho})^2}{(4\bar{\rho})} + O(\alpha)$$

718 such that

$$719 \quad (71) \quad \frac{\partial}{\partial w} x_s(\bar{w})(w - \bar{w}) = \frac{1}{\bar{\rho}} \int_{t^{on}}^{t^{off}} f'(\bar{u}(t)) (u(t) - \bar{u}(t)) dt + \left[ \frac{(1 - 2\bar{\rho})^2}{(4\bar{\rho})} + O(\alpha) \right] (t^{on} - \bar{t}^{on}).$$

720 We further obtain

$$723 \quad (72) \quad \frac{\partial}{\partial x} Y_2(\bar{t}, x, \bar{w}) = \frac{1}{2(\bar{t} - \bar{t}^{on})} \quad \text{and} \quad \frac{\partial}{\partial w} Y_2(\bar{t}, x, \bar{w})(w - \bar{w}) = \frac{x(t^{on} - \bar{t}^{on})}{2(\bar{t} - \bar{t}^{on})^2}.$$

724

725 We will prove that there exist nonnegative measures  $\bar{\mu}_1$ ,  $\bar{\mu}_2$  and  $\bar{\mu}_3$  such that the  
 726 optimality system in (36)-(38) is satisfied for the choice  $w = \bar{w}$ , provided that  $\alpha$  is  
 727 sufficiently small and  $\bar{t} - \bar{t}^{\text{off}} \geq 2$ . Since we deal with lower bounds in the state  
 728 constraints, the optimality system has to be modified to

$$729 \quad (73) \quad Y_k(\bar{t}, \cdot, \bar{w}) \geq -\bar{\rho}(\cdot) \quad \text{on } I_k \quad \text{for } k = 1, 2, 3$$

$$730 \quad (74) \quad \sum_{k=1}^3 \int_{I_k} (-\bar{\rho} - Y_k(\bar{t}, x, \bar{w})) d\bar{\mu}_k(x) = 0$$

$$731 \quad \hat{J}'(\bar{w})(w - \bar{w}) + \sum_{k=1}^3 \left[ \int_{I_k} \frac{\partial}{\partial x} [-\bar{\rho} - Y_k(\bar{t}, x, \bar{w})] \frac{x - x_{k-1}(\bar{w})}{x_k(\bar{w}) - x_{k-1}(\bar{w})} d\bar{\mu}_k(x) \right.$$

$$732 \quad (75) \quad \cdot \frac{\partial}{\partial w} x_k(\bar{w})(w - \bar{w}) + \int_{I_k} \frac{\partial}{\partial x} [-\bar{\rho} - Y_k(\bar{t}, x, \bar{w})] \frac{x_k(\bar{w}) - x}{x_k(\bar{w}) - x_{k-1}(\bar{w})} d\bar{\mu}_k(x)$$

$$733 \quad \cdot \left. \frac{\partial}{\partial w} x_{k-1}(\bar{w})(w - \bar{w}) - \int_{I_k} \frac{\partial}{\partial w} Y_k(\bar{t}, x, \bar{w})(w - \bar{w}) d\bar{\mu}_k(x) \right] \geq 0 \quad \forall w \in W_{\text{ad}}.$$

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

$$737 \quad \bar{\mu}_1 = \bar{\mu}_3 \equiv 0 \quad \text{and} \quad \bar{\mu}_2(A) = 0 \quad \forall A \subseteq (x_s(\bar{w}), x_r].$$

739 Hence, by Corollary 29, (75) reads

$$740 \quad (76) \quad \hat{J}'(\bar{w})(w - \bar{w}) - \frac{\partial}{\partial x} Y_2(\bar{t}, x_s(\bar{w}), \bar{w}) \bar{\mu}_2(\{x_s(\bar{w})\}) \frac{\partial}{\partial w} x_s(\bar{w})(w - \bar{w})$$

$$741 \quad - \frac{\partial}{\partial w} Y_2(\bar{t}, x_s(\bar{w}), \bar{w})(w - \bar{w}) \bar{\mu}_2(\{x_s(\bar{w})\}) \geq 0 \quad \forall w \in W_{\text{ad}}.$$

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

$$743 \quad (77) \quad \left[ (1 - \bar{\rho}) - \frac{\bar{\mu}_2(\{x_s(\bar{w})\})}{2\bar{\rho}(\bar{t} - \bar{t}^{\text{on}})} \right] \cdot \int_{\bar{t}^{\text{on}}}^{\bar{t}^{\text{off}}} f'(\bar{u}(t))(u(t) - \bar{u}(t)) dt$$

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

$$744 \quad \cdot \bar{\mu}_2(\{x_s(\bar{w})\}) \geq 0 \quad \forall w \in W_{\text{ad}}.$$

745 Hence, we find a bounded Radon measure of the form

$$746 \quad (78) \quad \bar{\mu}_2(\cdot) = \frac{(\frac{1}{4}(1 - \bar{\rho}) + \frac{1}{3}\bar{\rho}^3)8\bar{\rho}(\bar{t} - \bar{t}^{\text{on}}) + O(\alpha)}{(1 - 2\bar{\rho})^2 + 4\bar{\rho}(1 - 2\bar{\rho})(\bar{t} - \bar{t}^{\text{on}})} \delta_{x_s(\bar{w})}(\cdot),$$

748 where  $\delta_{x_s(\bar{w})}(\cdot)$  denotes the Dirac measure, such that the second term in (77) vanishes.  
 749 Moreover, for  $\alpha > 0$  sufficiently small  $\bar{\mu}_2$  is nonnegative and

$$750 \quad (79) \quad \left[ (1 - \bar{\rho}) - \frac{\bar{\mu}_2(\{x_s(\bar{w})\})}{2\bar{\rho}(\bar{t} - \bar{t}^{\text{on}})} \right] \geq 0,$$

751

752 if  $\bar{t} - \bar{t}^{\text{off}} \geq 2$ . By the choice of  $\bar{u}$ , the integral in (77) is nonnegative for all  $w \in W_{\text{ad}}$ .  
 753 Combining this with (79) and the fact that the second term in (77) is equal to zero  
 754 for the choice in (78), the optimality system (73)–(75) is satisfied. We will now use  
 755 *Theorem 32* to show that Robinson’s CQ (24) is satisfied. Since we deal with lower  
 756 bounds in this example, the conditions in *Theorem 32* have to be modified to

$$757 \quad (80) \quad \bar{w} + \delta w \in W_{\text{ad}} \quad \forall \delta w \in B_\varepsilon^W(\bar{w}) \cap \{w \in W_{\text{ad}} : w \geq 0\}$$

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

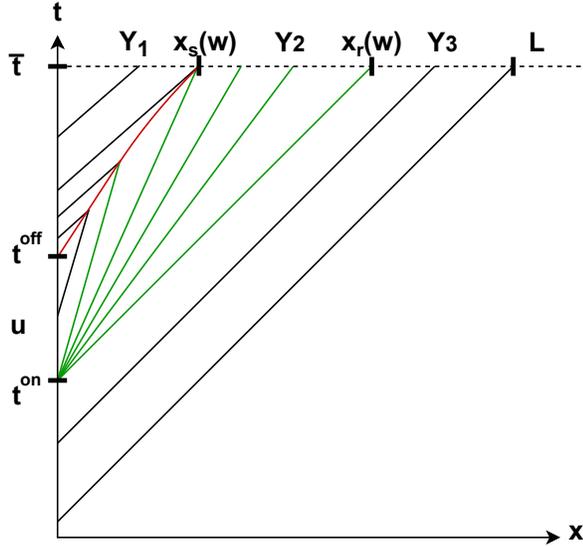


FIG. 1. Structure of the solution

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

$$765 \quad (P_\gamma) \quad \min_{w \in W} J_\gamma(y(w)) := J(y(w)) + \frac{\gamma}{2} \int_a^b (y(\bar{t}, x; w) - \bar{y}(x))_+^2 dx$$

$$766 \quad \text{s.t. } w \in W_{\text{ad}} \text{ and } y(w) \text{ solves the IBVP (1),}$$

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

770 The goal of this section is to conduct a convergence analysis of  $(P_\gamma)$  yielding  
 771 a convenient analytical basis for the numerical approximation of solutions to the  
 772 optimality conditions in *Theorem 27*. The first theorem in this section yields an  
 773 adjoint based formulation for the derivative of  $J_\gamma(y(w))$ , which gives the theoretical  
 774 foundation for applying numerical optimization methods. The analysis of penalty  
 775 approaches of the form  $(P_\gamma)$  has been subject to several publications, e.g., for the case  
 776 of elliptic problems in [30] and for parabolic problems in [31]. However, the case of

777 conservation laws requires very different techniques and is more involved than the  
778 aforementioned elliptic and parabolic cases, since one has to deal with discontinuous  
779 solutions.

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

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

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

788 *Proof.* The proof is similar to the proof of [Theorem 20](#). □

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

792 *Proof.* Let  $(w_{\gamma_j})_{j \in \mathbb{N}}$  be a sequence of global optima for  $(P_{\gamma_j})$ . Since the set  
793  $W_{\text{ad}} \subset W$  is compact, there exists at least one accumulation point  $w^*$ . Let  $w^*$  be  
794 any accumulation point and consider a convergent subsequence (again denoted by  
795  $(w_{\gamma_j})_{j \in \mathbb{N}}$ ) such that  $w_{\gamma_j} \rightarrow w^* \in W_{\text{ad}}$  w.r.t.  $\|\cdot\|_W$ . In the next step, we prove that  
796  $y(w^*)$  fulfills the state constraints. To this end, we firstly note that

$$797 \quad (81) \quad J_{\gamma_j}(y(w_{\gamma_j})) \leq J(y(\bar{w})) \quad \text{for all } j \in \mathbb{N},$$

799 where  $\bar{w}$  denotes a globally optimal solution for (P), which exists due to [Theorem 20](#).  
800 By (81), the continuity of  $J(y(\cdot))$  w.r.t.  $w$  and the fact that  $w_{\gamma_j} \rightarrow w^* \in W_{\text{ad}}$ , there  
801 exists a constant  $C > 0$  such that for all  $j \in \mathbb{N}$  it holds

$$802 \quad (82) \quad 0 \leq \frac{\gamma_j}{2} \int_a^b (y(\bar{t}, x; w_{\gamma_j}) - \bar{y}(x))_+^2 dx \leq J(y(\bar{w})) - J(y(w_{\gamma_j})) \leq C$$

804 Using (82), we prove that  $y(\bar{t}, x; w^*) \leq \bar{y}(x)$  holds for all  $x \in [a, b]$ . From (82) and  
805  $\gamma_j \rightarrow \infty$  we deduce  $(y(\bar{t}, \cdot; w_{\gamma_j}) - \bar{y}(\cdot))_+ \rightarrow 0$  in  $L^2([a, b])$  and hence pointwise almost  
806 everywhere on  $[a, b]$  for a subsequence (again denoted by  $(w_{\gamma_j})_{j \in \mathbb{N}}$ ). Considering this  
807 subsequence, we know by [Corollary 7](#) that  $y(\bar{t}, \cdot; w_{\gamma_j}) \rightarrow y(\bar{t}, \cdot; w^*)$  in  $L^1([a, b])$  and  
808 hence pointwise almost everywhere on  $[a, b]$  for another subsequence, which is again  
809 denoted by  $(w_{\gamma_j})_{j \in \mathbb{N}}$ . Since  $(y(\bar{t}, \cdot; w_{\gamma_j}) - \bar{y}(\cdot))_+ \rightarrow 0$  and  $y(\bar{t}, \cdot; w_{\gamma_j}) \rightarrow y(\bar{t}, \cdot; w^*)$   
810 for almost all  $x \in [a, b]$ , we obtain  $(y(\bar{t}, \cdot; w^*) - \bar{y}(\cdot))_+ = 0$  almost everywhere on  $[a, b]$   
811 which is equivalent to

$$812 \quad (83) \quad y(\bar{t}, x; w^*) \leq \bar{y}(x) \quad \text{for a.a. } x \in [a, b].$$

814 Analogously to the proof of [Theorem 20](#), one can show that (83) holds for all  $x \in$   
815  $[a, b]$  and hence  $y(\bar{t}, \cdot; w^*)$  fulfills the state constraints. Furthermore, the inequality  
816  $J_{\gamma_j}(y(w_{\gamma_j})) \leq J(y(\bar{w}))$  for all  $j \in \mathbb{N}$  implies  $J(y(w^*)) \leq J(y(\bar{w}))$ . Since  $y(\bar{t}, \cdot; w^*)$   
817 satisfies the state constraints and  $w^* \in W_{\text{ad}}$ ,  $w^*$  is thus a globally optimal solution for  
818 (P). □

819 We will now examine the convergence of local solutions of  $(P_\gamma)$ . To this end, we  
820 introduce for some constant  $r > 0$  the auxiliary problems

$$821 \quad (P^r) \quad \min_{w \in W} J(y(w)) \quad \text{s.t.} \quad y(w) \text{ solves the IBVP (1), } w \in W^r := W_{\text{ad}} \cap B_r^W(\bar{w})$$

$$822 \quad y(\bar{t}, x; w) \leq \bar{y}(x) \quad \forall x \in [a, b],$$

824  $(P_\gamma^r)$   $\min_{w \in W} J_\gamma(y(w))$  s.t.  $y(w)$  solves the IBVP (1),  $w \in W^r$   
826

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

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

$$831 \quad (84) \quad J(y(\bar{w})) + \frac{\delta}{2} \|w - \bar{w}\|_H \leq J(y(w)) \quad \forall w \in \tilde{W}_{\text{ad}} \text{ with } \|w - \bar{w}\|_W < \varepsilon$$

833 *is satisfied, where  $\tilde{W}_{\text{ad}} := \{w \in W_{\text{ad}} : y(\bar{t}, x; w) \leq \bar{y}(x) \ \forall x \in [a, b]\}$  and  $H$  is a  
834 Hilbert space such that  $W$  is continuously embedded in  $H$ . Then there exists a sequence  
835  $(w_{\gamma_j})_{j \in \mathbb{N}}$  of local solutions for  $(P_{\gamma_j})$  converging to  $\bar{w}$  w.r.t.  $\|\cdot\|_W$ .*

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

843 **THEOREM 39.** *Suppose that (A1) and (A2) hold and consider a sequence  $(w_{\gamma_j})_{j \in \mathbb{N}}$   
844 of local solutions for  $(P_{\gamma_j})$  such that for constants  $\varepsilon, \delta > 0$  the condition*

$$845 \quad (85) \quad J_{\gamma_j}(y(w_{\gamma_j})) + \frac{\delta}{2} \|w - w_{\gamma_j}\|_H \leq J_{\gamma_j}(y(w)) \quad \forall w \in W_{\text{ad}} \text{ with } \|w - w_{\gamma_j}\|_W < \varepsilon$$

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

850 *Proof.* Since  $W_{\text{ad}}$  is by assumption compact, there exists a convergent subsequence  
851 again denoted by  $(w_{\gamma_j})_{j \in \mathbb{N}}$  with limit  $\bar{w} \in W_{\text{ad}}$ . We consider the corresponding  
852 problems  $(P^r)$  and  $(P_{\gamma_j}^r)$  with  $r = \frac{\varepsilon}{2}$ . Then (85) yields that  $w_{\gamma_j}$  is the unique globally  
853 optimal solution for  $(P_{\gamma_j}^{\frac{\varepsilon}{2}})$ , if  $j$  is large enough. Using this, one can analogously to the  
854 proof of [Theorem 37](#) show that  $\bar{w} \in W_{\text{ad}}$  is a globally optimal solution for  $(P^{\frac{\varepsilon}{2}})$  and  
855 hence a local solution for (P).  $\square$

856 **LEMMA 40.** *Let (A1) and (A2) hold and consider a sequence  $(w_{\gamma_j})_{j \in \mathbb{N}} \subset W$  with  
857  $\gamma_j \rightarrow \infty$  for  $j \rightarrow \infty$ , that converges to some  $w^* \in W$  satisfying (ND). Then for  $j$  large  
858 enough  $w_{\gamma_j}$  satisfies (ND) and there exists  $\varepsilon > 0$  such that for all  $k = 1, \dots, K + 1$  it  
859 holds that*

$$860 \quad \lim_{j \rightarrow \infty} Y_k(\bar{t}, \cdot; w_{\gamma_j}) = Y_k(\bar{t}, \cdot; w^*) \quad \text{in } C^1((x_{k-1}(w^*) - \varepsilon, x_k(w^*) + \varepsilon)),$$

$$861 \quad \lim_{j \rightarrow \infty} \frac{\partial}{\partial w} Y_k(\bar{t}, \cdot; w_{\gamma_j}) = \frac{\partial}{\partial w} Y_k(\bar{t}, \cdot; w^*) \quad \text{in } C((x_{k-1}(w^*) - \varepsilon, x_k(w^*) + \varepsilon)).$$

863 *Proof.* The first assertion is a consequence of [33, Lemma 3.1.10]. The second one  
864 and the fact that for  $j$  large enough  $w_{\gamma_j}$  satisfies (ND) follows from [Theorem 17](#).  $\square$

865 **THEOREM 41.** *Let (A1) and (A2) hold and  $w_\gamma \in W_{\text{ad}}$  be a local solution for  $(P_\gamma)$   
866 satisfying (ND). Then it holds that*

$$867 \quad (86) \quad \frac{d}{dw} J_\gamma(y(w_\gamma)) \cdot (w - w_\gamma) \geq 0 \quad \forall w \in W_{\text{ad}}.$$

868

869 *Proof.* This result can be found, e.g., in [20].  $\square$

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

$$871 \quad (87) \quad J_\gamma(y(w)) = J(y(w)) + \sum_{k=1}^{K+1} z_k(w) \cdot \frac{\gamma}{2z_k(w)} \int_{x_{k-1}(w)}^{x_k(w)} (Y_k(\bar{t}, x; w) - \bar{y}(x))_+^2 dx,$$

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

$$874 \quad (88) \quad \lambda_k(x; w) := \begin{cases} \gamma (Y_k(\bar{t}, x, w) - \bar{y}(x))_+, & \text{for } x_{k-1}(w) \leq x \leq x_k(w) \\ 0, & \text{for } x \in [a, b] \setminus [x_{k-1}(w), x_k(w)], \end{cases}$$

$$875 \quad (89) \quad r_k(w) := \frac{\gamma}{2z_k(w)} \int_{x_{k-1}(w)}^{x_k(w)} (Y_k(\bar{t}, x, w) - \bar{y}(x))_+^2 dx$$

876  
877  
878 **LEMMA 42.** *Let (A1) and (A2) hold. Then the derivative of  $J_\gamma(y(w))$  in some*  
879  *$w \in W$  satisfying (ND) in a direction  $\delta w \in W$  is given by*

$$880 \quad \frac{d}{dw} J_\gamma(y(w)) \delta w = \frac{d}{dw} J(y(w)) \delta w + \sum_{k=1}^{K+1} \left[ \int_{x_{k-1}(w)}^{x_k(w)} \frac{\partial}{\partial w} Y_k(\bar{t}, x; w) \delta w \lambda_k(x; w) dx \right. \\ 881 \quad (90) \quad + \int_{x_{k-1}(w)}^{x_k(w)} \frac{\partial}{\partial x} (Y_k(\bar{t}, x; w) - \bar{y}(x)) \frac{x - x_{k-1}(w)}{x_k(w) - x_{k-1}(w)} \lambda_k(x; w) dx \frac{\partial}{\partial w} x_k(w) \delta w \\ 882 \quad + \int_{x_{k-1}(w)}^{x_k(w)} \frac{\partial}{\partial x} (Y_k(\bar{t}, x; w) - \bar{y}(x)) \frac{x_k(w) - x}{x_k(w) - x_{k-1}(w)} \lambda_k(x; w) dx \frac{\partial}{\partial w} x_{k-1}(w) \delta w \\ 883 \quad \left. + r_k(w) \frac{\partial}{\partial w} (x_k(w) - x_{k-1}(w)) \delta w \right].$$

884  
885 *Proof.* We obtain (90) by differentiating (87) w.r.t.  $w$ . The first term of the right  
886 hand side of (90) is obvious. The first and the last term in the sum in (90) result  
887 from differentiating the integrand and the first factor under the sum in (87) w.r.t.  $w$ .  
888 Finally, we observe that the derivative

$$889 \quad \frac{d}{dv} \left( z_k(w) \cdot \frac{\gamma}{2z_k(v)} \int_{x_{k-1}(v)}^{x_k(v)} (Y_k(\bar{t}, x; w) - \bar{y}(x))_+^2 dx \right) \Big|_{v=w} \cdot \delta w = \frac{\gamma}{2} \left[ -\frac{1}{z_k(w)} \right. \\ 890 \quad \cdot \int_{x_{k-1}(w)}^{x_k(w)} (Y_k(\bar{t}, x; w) - \bar{y}(x))_+^2 dx \frac{\partial}{\partial w} z_k(w) \delta w + (Y_k(\bar{t}, x_k(w); w) - \bar{y}(x_k(w)))_+^2 \\ 891 \quad \cdot \frac{\partial}{\partial w} x_k(w) \delta w - (Y_k(\bar{t}, x_{k-1}(w); w) - \bar{y}(x_{k-1}(w)))_+^2 \frac{\partial}{\partial w} x_{k-1}(w) \delta w \left. \right]$$

892 coincides with the sum of the second and the third integral in (90), where we use that  
893  $\frac{\partial}{\partial x} (Y_k(\bar{t}, x; w) - \bar{y}(x)) \lambda_k(x; w) = \frac{\gamma}{2} \frac{\partial}{\partial x} (Y_k(\bar{t}, x; w) - \bar{y}(x))_+^2$  and apply integration by  
894 parts.  $\square$

895  
896 **LEMMA 43.** *Let (A1) and (A2) hold and consider a sequence  $(w_{\gamma_j})_{j \in \mathbb{N}}$  of local*  
897 *solutions  $w_{\gamma_j}$  for  $(P_{\gamma_j})$  with  $\gamma_j \rightarrow \infty$  for  $j \rightarrow \infty$  that converges to a local solution*  
898  *$\bar{w} \in W_{\text{ad}}$  for (P) and satisfies*

$$899 \quad (91) \quad J_{\gamma_j}(y(w_{\gamma_j})) \leq J_{\gamma_j}(y(w)) \quad \forall w \in W_{\text{ad}} \text{ with } \|w - w_{\gamma_j}\|_W < \varepsilon$$

901 for some constant  $\varepsilon > 0$  and sufficiently large  $j$ . If  $\bar{w}$  satisfies in addition (ND) then

$$902 \quad (92) \quad \lim_{j \rightarrow \infty} r_k(w_{\gamma_j}) = 0 \quad \text{for all } k = 1, \dots, K + 1.$$

903  
904 *Proof.* We note that since  $\bar{w}$  satisfies (ND), [Lemma 40](#) yields that  $w_{\gamma_j}$  satisfies  
905 (ND) for  $j$  large enough such that [\(87\)](#) is valid for all  $w = w_{\gamma_j}$  if  $j$  is large enough.  
906 Using [\(87\)](#), [\(91\)](#), the same arguments as in the proof of [Theorem 37](#) show that

$$907 \quad (93) \quad \lim_{j \rightarrow \infty} \sum_{k=1}^{K+1} z_k(w_{\gamma_j}) r_k(w_{\gamma_j}) = \lim_{j \rightarrow \infty} \frac{\gamma_j}{2} \int_a^b (y(\bar{t}, x; w_{\gamma_j}) - \bar{y}(x))_+^2 dx = 0.$$

908  
909 Since  $x_{k-1}(w_{\gamma_j})$  and  $x_k(w_{\gamma_j})$  are uniformly bounded away from each other for  $j \in \mathbb{N}$   
910 large enough and the integrands in [\(93\)](#) are nonnegative, we obtain [\(92\)](#).  $\square$

911 **REMARK 44.** *The quadratic growth condition [\(84\)](#) implies [\(91\)](#).*

912 Considering a sequence of local solutions  $(w_{\gamma_j})_{j \in \mathbb{N}}$  of  $(P_{\gamma_j})$  satisfying (ND), we want  
913 to analyze in which sense the terms in [\(88\)](#) converge. In the next lemma we prove that  
914 the sequences  $(\lambda_k(x; w_{\gamma_j}))_{j \in \mathbb{N}}$  are uniformly bounded in  $L^1$ .

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

919 *Proof.* [Lemma 40](#) ensures that  $w_{\gamma_j}$  satisfies (ND) for sufficiently large  $j$ . Therefore  
920 [\(86\)](#) is satisfied in  $w_{\gamma_j}$ , which together with the representation of the gradient in [\(90\)](#),  
921 [Theorem 19](#), [Lemma 43](#) and the compactness of  $W_{\text{ad}}$  yields the existence of a positive  
922 constant  $C > 0$  such that

$$923 \quad \sum_{k=1}^{K+1} \int_{x_{k-1}(w_{\gamma_j})}^{x_k(w_{\gamma_j})} - \left[ \frac{\partial}{\partial w} Y_k(\bar{t}, x; w_{\gamma_j})(w - w_{\gamma_j}) + \frac{\partial}{\partial x} (Y_k(\bar{t}, x; w_{\gamma_j}) - \bar{y}(x)) \right.$$

$$924 \quad (94) \quad \cdot \frac{x - x_{k-1}(w_{\gamma_j})}{x_k(w_{\gamma_j}) - x_{k-1}(w_{\gamma_j})} \frac{\partial}{\partial w} x_k(w_{\gamma_j})(w - w_{\gamma_j}) + \frac{\partial}{\partial x} (Y_k(\bar{t}, x; w_{\gamma_j}) - \bar{y}(x))$$

$$925 \quad \cdot \left. \frac{x_k(w_{\gamma_j}) - x}{x_k(w_{\gamma_j}) - x_{k-1}(w_{\gamma_j})} \frac{\partial}{\partial w} x_{k-1}(w_{\gamma_j}) \cdot (w - w_{\gamma_j}) \right] \lambda_k(x; w_{\gamma_j}) dx \leq C.$$

926  
927 We show that there is  $\bar{j} > 0$  such that  $\lambda_k(\cdot; w_{\gamma_j})$ ,  $k = 1, \dots, K + 1$ , are uniformly  
928 bounded in  $L^1([a, b])$  for all  $j \geq \bar{j}$ . Since [\(24\)](#) is satisfied in  $\bar{w}$  by assumption, [Lemma 31](#)  
929 yields the existence of  $\tilde{w} \in W_{\text{ad}}$  and  $\varepsilon_1 > 0$  such that

$$930 \quad (95) \quad -\varepsilon_1 \geq Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x) + \frac{\partial}{\partial w} Y_k(\bar{t}, x, \bar{w})(\tilde{w} - \bar{w}) + \frac{\partial}{\partial x} (Y_k(\bar{t}, x, \bar{w}) - \bar{y}(x)) \cdot$$

$$\left( \frac{x - x_{k-1}(\bar{w})}{x_k(\bar{w}) - x_{k-1}(\bar{w})} \frac{\partial}{\partial w} x_k(\bar{w}) + \frac{x_k(\bar{w}) - x}{x_k(\bar{w}) - x_{k-1}(\bar{w})} \frac{\partial}{\partial w} x_{k-1}(\bar{w}) \right) \cdot (\tilde{w} - \bar{w})$$

931 holds for all  $x \in [x_{k-1}(\bar{w}), x_k(\bar{w})]$  and  $k = 1, \dots, K + 1$ . By [Lemma 40](#) we find  $\varepsilon > 0$   
932 and  $\bar{j} > 0$  such that for all  $j \geq \bar{j}$

$$933 \quad (96) \quad -\frac{\varepsilon_1}{2} \geq Y_k(\bar{t}, x, w_{\gamma_j}) - \bar{y}(x) + \frac{\partial}{\partial w} Y_k(\bar{t}, x, w_{\gamma_j})(\tilde{w} - w_{\gamma_j}) + \frac{\partial}{\partial x} (Y_k(\bar{t}, x, w_{\gamma_j}) - \bar{y}(x)) \cdot$$

$$\left( \frac{x - x_{k-1}(w_{\gamma_j})}{x_k(w_{\gamma_j}) - x_{k-1}(w_{\gamma_j})} \frac{\partial}{\partial w} x_k(w_{\gamma_j}) + \frac{x_k(w_{\gamma_j}) - x}{x_k(w_{\gamma_j}) - x_{k-1}(w_{\gamma_j})} \frac{\partial}{\partial w} x_{k-1}(w_{\gamma_j}) \right) \cdot (\tilde{w} - w_{\gamma_j})$$

934 for all  $x \in (x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon)$ , since the right hand side of (95) is continuous  
 935 on  $(x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon)$  and the right hand side of (96) converges on this set by  
 936 Lemma 40 uniformly to the right hand side of (95).

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

$$940 \quad \sum_{k=1}^{K+1} \int_{\text{supp}(\lambda_k(x; w_{\gamma_j}))} \lambda_k(x; w_{\gamma_j}) dx \leq \frac{2C}{\varepsilon_1}.$$

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

942 **THEOREM 46.** *Let (A1) and (A2) hold and consider a sequence  $(w_{\gamma_j})_{j \in \mathbb{N}}$  of local*  
 943 *solutions for  $(P_{\gamma_j})$  with  $\gamma_j \rightarrow \infty$  for  $j \rightarrow \infty$ , that satisfies (91) for sufficiently large  $j$*   
 944 *and converges to a local solution  $\bar{w} \in W_{\text{ad}}$  for (P). If  $\bar{w}$  satisfies (ND) and (24), then*  
 945 *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*  
 946  *$\mathcal{M}([a, b])$ -weak\* for  $k = 1, \dots, K+1$ . Moreover,  $\text{supp}(\mu_k) \subset I_k = [x_{k-1}(\bar{w}), x_k(\bar{w})]$ , the*  
 947 *measures  $\mu_k \in \mathcal{M}([a, b])$  are nonnegative and the optimality conditions in Theorem 27*  
 948 *are satisfied in  $\bar{w}$  for  $\bar{\mu}_k = \mu_k|_{I_k}$  for all  $k = 1, \dots, K+1$ . Finally, for all  $k = 1, \dots, K+1$*   
 949 *it holds*

(97)

$$950 \quad \hat{\lambda}_k(\cdot; w_{\gamma_j}) := \lambda_k \left( x_{k-1}(w_{\gamma_j}) + (\cdot - x_{k-1}(\bar{w})) \frac{x_k(w_{\gamma_j}) - x_{k-1}(w_{\gamma_j})}{x_k(\bar{w}) - x_{k-1}(\bar{w})}; w_{\gamma_j} \right) \rightharpoonup \mu_k(\cdot)$$

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

953 *Proof.* Since the sequences of Lagrange multiplier estimates  $(\lambda_k(\cdot; w_{\gamma_j}))_{j \in \mathbb{N}}$  are uni-  
 954 formly bounded in  $L^1([a, b])$  by Lemma 45, the Banach-Alaoglu Theorem yields a subse-  
 955 quence, 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\*  
 956 for  $k = 1, \dots, K+1$ . Recalling (88) and using the continuity of  $x_1(w), \dots, x_K(w)$   
 957 w.r.t.  $w$ , we obtain that  $\text{supp}(\mu_k) \subset I_k$  for all  $k = 1, \dots, K+1$ . The measures  
 958  $\mu_k \in \mathcal{M}([a, b])$  are nonnegative, since  $(\lambda_k(\cdot; w_{\gamma_j})) \geq 0$  for all  $j \in \mathbb{N}$ . We will prove  
 959 that the optimality conditions in Theorem 27, i.e. (36), (37) and (38) are satisfied for  
 960  $\bar{\mu}_k = \mu_k|_{I_k}$ ,  $k = 1, \dots, K+1$ .

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

$$964 \quad (98) \quad \lim_{j \rightarrow \infty} \psi(\cdot, w_{\gamma_j}) = \psi(\cdot, \bar{w}) \quad \text{in } C([x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon]),$$

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

$$967 \quad (99) \quad \text{supp}(\lambda_k(\cdot; w_{\gamma_j})) = [x_{k-1}(w_{\gamma_j}), x_k(w_{\gamma_j})] \subset [x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon] =: I_k^\varepsilon$$

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

$$971 \quad (100) \quad \begin{aligned} & \lim_{j \rightarrow \infty} \int_{x_{k-1}(w_{\gamma_j})}^{x_k(w_{\gamma_j})} \psi(x; w_{\gamma_j}) \lambda_k(x; w_{\gamma_j}) dx = \lim_{j \rightarrow \infty} \int_{I_k^\varepsilon} \psi(x, w_{\gamma_j}) \lambda_k(x; w_{\gamma_j}) dx \\ & = \int_{I_k^\varepsilon} \psi(x, \bar{w}) d\mu_k(x) = \int_{I_k} \psi(x, \bar{w}) d\mu_k(x). \end{aligned}$$

972 The last equality follows from  $\text{supp}(\mu_k) = I_k$ .

973 Now (36) holds, since  $\bar{w}$  is a local solution of (P). Concerning (37), using  $0 \leq$   
 974  $J_{\gamma_j}(w_{\gamma_j}) \leq J_{\gamma_j}(\bar{w}) = J(y(\bar{w}))$  which holds by (91),  $w_{\gamma_j} \rightarrow \bar{w}$  w.r.t.  $\|\cdot\|_W$  and the  
 975 definitions of  $\lambda_k(\cdot; w)$  and  $J_{\gamma_j}(y(w))$ , we obtain

$$\begin{aligned} 976 \quad (101) \quad & 0 \leq \lim_{j \rightarrow \infty} \sum_{k=1}^{K+1} \int_{x_{k-1}(w_{\gamma_j})}^{x_k(w_{\gamma_j})} (Y_k(\bar{t}, x; w_{\gamma_j}) - \bar{y}(x)) \lambda_k(x; w_{\gamma_j}) \, dx \\ & \leq 2 \cdot \lim_{j \rightarrow \infty} (J(\bar{w}) - J(w_{\gamma_j})) = 0. \end{aligned}$$

977 In order to prove (37), we use (101) in connection with (100). More precisely, we set  
 978  $\psi_k(x, w) = Y_k(\bar{t}, x; w) - \bar{y}(x)$ ,  $k = 1, \dots, K+1$ , and apply (100) to each integral in  
 979 the sum in (101), which is possible due to the results Lemma 40. Then we obtain (37).  
 980 To show (38), we use for  $i = 0, 1$  the abbreviations

$$981 \quad \psi_{k,i}(x, w; \tilde{w}) := (-1)^i \frac{\partial}{\partial x} (\bar{y}(x) - Y_k(\bar{t}, x; w)) \cdot \frac{x - x_{k-1+i}(w)}{x_k(w) - x_{k-1}(w)} \frac{\partial}{\partial w} x_{k-i}(w) (\tilde{w} - w).$$

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

$$\begin{aligned} (102) \quad & \lim_{j \rightarrow \infty} \left( \frac{d}{dw} J(y(w_{\gamma_j})) \right) (w - w_{\gamma_j}) \\ 986 \quad & + \sum_{k=1}^{K+1} \int_{x_{k-1}(w_{\gamma_j})}^{x_k(w_{\gamma_j})} \left( \frac{\partial}{\partial w} Y_k(\bar{t}, x; w_{\gamma_j}) (w - w_{\gamma_j}) + \sum_{i=0}^1 \psi_{k,i}(x, w_{\gamma_j}; w) \right) \lambda_k(x; w_{\gamma_j}) \, dx \geq 0. \end{aligned}$$

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

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

$$996 \quad \lim_{j \rightarrow \infty} \int_{x_{k-1}(\bar{w})}^{x_k(\bar{w})} \hat{\lambda}_k(x; w_{\gamma_j}) \varphi(x) \, dx = \lim_{j \rightarrow \infty} \int_{x_{k-1}(w_{\gamma_j})}^{x_k(w_{\gamma_j})} \lambda_k(x; w_{\gamma_j}) \tilde{\varphi}(x, w_{\gamma_j}) \, dx$$

998 with

$$\begin{aligned} 999 \quad \tilde{\varphi}(x, w_{\gamma_j}) = & \varphi \left( x_{k-1}(\bar{w}) + (x_k(\bar{w}) - x_{k-1}(\bar{w})) \frac{x - x_{k-1}(w_{\gamma_j})}{x_k(w_{\gamma_j}) - x_{k-1}(w_{\gamma_j})} \right) \\ & \cdot \frac{x_k(w_{\gamma_j}) - x_{k-1}(w_{\gamma_j})}{x_k(\bar{w}) - x_{k-1}(\bar{w})} \end{aligned}$$

1000

1001

1002 Now clearly  $\tilde{\varphi}(\cdot, w_{\gamma_j}) \rightarrow \tilde{\varphi}(\cdot, \bar{w}) = \varphi$  in  $C([x_{k-1}(\bar{w}) - \varepsilon, x_k(\bar{w}) + \varepsilon])$ . Hence, (98) holds  
 1003 for  $\psi = \tilde{\varphi}$  and (100) yields

$$1004 \quad \lim_{j \rightarrow \infty} \int_{x_{k-1}(\bar{w})}^{x_k(\bar{w})} \hat{\lambda}_k(x; w_{\gamma_j}) \varphi(x) \, dx = \int_{x_{k-1}(\bar{w})}^{x_k(\bar{w})} \varphi(x) \, d\mu_k(x)$$

1005 and thus (97), since  $\varphi$  was arbitrary.  $\square$

1006 **5. Conclusion and possible extensions.** We have considered the optimal  
1007 control of initial-boundary value problems with pointwise state constraints. Hereby, we  
1008 have proved the existence of an optimal control and have derived necessary optimality  
1009 conditions. In addition, we have discussed the Moreau-Yosida regularization approach  
1010 where we have proved that each sequence of optimal solutions for the regularized  
1011 problems converges strongly to an optimal solution of the original problem with  
1012 pointwise state constraints. Furthermore, we have derived optimality conditions  
1013 for the regularized problems and have shown that if the optimal solutions of the  
1014 regularized problems converge to a local optimum for (P) in which Robinson's CQ is  
1015 satisfied, then there exists a sequence of Lagrange multiplier estimates that converges  
1016 in  $\mathcal{M}([a, b])$ -weak\* to Lagrange multipliers for the optimality system for (P).

1017 The optimality conditions and Robinson's regularity condition provide a basis  
1018 for stopping criteria, the proof of convergence of discretizations and the derivation of  
1019 error estimators. In particular, on the basis of the results in this paper, we obtain  
1020 that a convergent discretization of the state equation and the adjoint equation, see  
1021 for example [13–18, 37], are sufficient to obtain a convergent algorithm using the  
1022 Moreau-Yosida regularization approach.

1023 The method, that was developed in this paper to derive necessary optimality  
1024 conditions, could be extended to networks with node conditions, e.g., traffic light  
1025 problems, where the traffic flow is modeled by the LWR-model (see [27, 36]) by using  
1026 the results of [35]. Another possible extension is to consider systems of balance laws,  
1027 e.g., the Euler equations modeling the gas flow in a pipe. Hereby, if the solution has  
1028 similar structural properties, that we have collected in section 2 (see also condition H2  
1029 in [6]), then one can introduce new state variables and derive optimality conditions  
1030 and prove convergence of Moreau-Yosida type regularizations. Hence, if a variational  
1031 calculus for systems is available, the results of this paper can be generalized to more  
1032 complex problems. Some results dealing with a variational calculus for systems can be  
1033 found, e.g., in [5, 6].

1034 **Acknowledgments.** We thank the referees for their constructive comments. We  
1035 would like to acknowledge the support by German Research Foundation (DFG) within  
1036 the collaborative research center TRR 154 "Mathematical Modelling, Simulation and  
1037 Optimization using the Example of Gas Networks", project A02. Moreover, the work  
1038 of Stefan Ulbrich was supported by the DFG within the collaborative research center  
1039 SFB 1194 "Interaction between Transport and Wetting Processes", project B04.

1040 **Appendix A. Proof of Theorem 16.**

1041 *Proof.* In the following we will prove Fréchet-differentiability of the mapping (12)  
 1042 in  $w = \bar{w}$ , where the continuous differentiability follows from the stability of genuine  
 1043 characteristics and of the adjoint state (cf. [33, Proof of Lemma 6.2.7]). Denote by  
 1044  $\xi_{l/r}$  the minimal/maximal backward characteristic through  $(\bar{t}, x_s(\bar{w}))$ . We will restrict  
 1045 ourselves in the proof to the case that  $\xi_l$  ends in the interior of a rarefaction wave  
 1046 created by a discontinuity of the boundary data  $u_{B,a}$  in  $t = \bar{t}_m^a$  and  $\xi_r$  ends in a point  
 1047  $(0, z)$  where the initial data  $u_0$  is smooth. We further assume that  $(\bar{t}_m^a, a)$  and  $(0, \bar{x}_l^0)$   
 1048 with  $\bar{x}_l^0 < z$  are the only rarefaction wave creating discontinuities of the initial and  
 1049 boundary data. Moreover, we assume that

$$1050 \quad (103) \quad \mathbb{T}^a = \emptyset,$$

1052 where the treatment of the case that  $\mathbb{T}^a \neq \emptyset$  is explained in [34, Lemma 4.13]. We  
 1053 note that (6) and (103) together yield

$$1054 \quad (104) \quad y(\cdot, a+; w) = u_m^{B,a}(\cdot) \quad \text{on } [0, \bar{t}]$$

1056 for all  $w \in B_\rho^W(\bar{w})$  with  $\rho > 0$  small enough. We use these assumptions for the sake of  
 1057 simplicity and in order to avoid technical effort. Nevertheless, the proof can easily be  
 1058 extended to the general case as we will show later.

1059 Consider some  $\delta w \in W$  and set  $w := \bar{w} + \delta w$ . Furthermore, let  $\bar{y} := y(\bar{w})$   
 1060 and  $y := y(w)$  denote the entropy solutions of (1) for the controls  $\bar{u}_1$  and  $u_1$  in  
 1061 the source term and initial and boundary values  $\bar{u}_0 := u_0(\bar{w})$ ,  $\bar{u}_{B,a/b} := u_{B,a/b}(\bar{w})$ ,  
 1062  $u_0 := u_0(w)$  and  $u_{B,a/b} := u_{B,a/b}(w)$ , respectively. We define  $\delta u_0 := u_0 - \bar{u}_0$ ,  
 1063  $\delta u_{B,a/b} := u_{B,a/b} - \bar{u}_{B,a/b}$ ,  $\delta u_1 = u_1 - \bar{u}_1$  and  $\Delta y := y - \bar{y}$ .

1064 In the following,  $C$  and  $\rho$  denote large/small constants, respectively, which possibly  
 1065 change their values throughout the proof. For the sake of simplicity, we set  $a = 0$ .

1066 As in [33, Proof of Lemma 6.3.7], one can show that for all  $\bar{\varepsilon} > 0$  it holds that

$$1067 \quad (105) \quad \int_{x_s(\bar{w}) - \bar{\varepsilon}}^{x_s(\bar{w}) + \bar{\varepsilon}} \Delta y(\bar{t}, x) \, dx = (x_s(w) - x_s(\bar{w}))[\bar{y}(\bar{t}, x_s(\bar{w}))] + O((\bar{\varepsilon} + \|\delta w\|_W)\|\delta w\|_W).$$

1069 In the remaining part of the proof, we will derive an adjoint representation for the  
 1070 term

$$1071 \quad (106) \quad \frac{1}{[\bar{y}(\bar{t}, x_s(\bar{w}))]} \int_{x_s(\bar{w}) - \bar{\varepsilon}}^{x_s(\bar{w}) + \bar{\varepsilon}} \Delta y(\bar{t}, x) \, dx.$$

1073 As in [34, Proof of Lemma 4.10], we define for  $(t, x) \in \Omega_{\bar{t}} := (0, \bar{t}) \times (0, b)$

$$1074 \quad c(t, x) := f'(\bar{y}(t, x)), \quad d(t, x) := g_y(t, x, \bar{y}(t, x), \bar{u}_1),$$

$$1075 \quad (107) \quad \tilde{c}(t, x) := \int_0^1 f'(\tau y(t, x) + (1 - \tau)\bar{y}(t, x)) \, d\tau, \quad \tilde{d}(t, x) := g_y(t, x, y(t, x), u_1)$$

1077 and observe that  $\Delta y$  is a weak solution of

$$1078 \quad (108) \quad \Delta y_t + (\tilde{c}\Delta y)_x = \tilde{d}\Delta y + g(\cdot, \bar{y}, u_1) - g(\cdot, \bar{y}, \bar{u}_1)$$

1080 on  $\Omega_{\bar{t}}$ . As described in [34, Lemma 4.10], the functions  $c, \tilde{c}, d, \tilde{d}$  can be extended to  
 1081  $[0, \bar{t}] \times \mathbb{R}$  by setting

$$\begin{aligned} 1082 \quad c(t, x) &= \tilde{c}(t, x) = M_{f'} \text{ if } x < 0, & (d, \tilde{d})(t, x) &= (d, \tilde{d})(t, 0+) \text{ if } x < 0, \\ 1083 \quad c(t, x) &= \tilde{c}(t, x) = -M_{f'} \text{ if } x > b & (d, \tilde{d})(t, x) &= (d, \tilde{d})(t, b-) \text{ if } x > b. \end{aligned}$$

1085 We observe that since  $g_y$  does not depend on  $y$  by (A1) and due to the regularity of  $g$ ,  
 1086 it holds that

$$\begin{aligned} 1087 \quad & \tilde{d}, d \in L^\infty(0, \bar{t}; C^{0,1}(\mathbb{R})) \quad \text{and} \\ 1088 \quad (109) \quad & \tilde{d} \rightarrow d \text{ in } L^\infty(0, \bar{t}; C(\mathbb{R})) \text{ as } \|\delta w\|_W \rightarrow 0 \end{aligned}$$

1090 (cf. [40]).

1091 Let the interval  $[t_1, t_2] \subset [0, T]$  be chosen such that  $\bar{y}$  has no rarefaction wave  
 1092 creating discontinuity on  $[t_1, t_2] \times [0, b]$ . Then Proposition 4 and Corollary 7 yield

$$\begin{aligned} 1093 \quad & \|\tilde{c}\|_{\infty, [t_1, t_2] \times \mathbb{R}}, \|c\|_{\infty, [t_1, t_2] \times \mathbb{R}} \leq M_y, \\ 1094 \quad (110) \quad & \tilde{c} \rightarrow c \text{ in } L^1_{\text{loc}}([t_1, t_2] \times \mathbb{R}) \text{ and in } L^\infty([t_1, t_2] \times \mathbb{R})\text{-weak}^* \text{ as } \|\delta w\|_W \rightarrow 0. \end{aligned}$$

1096 (cf. [40]). Moreover, from (7) and [33, Lemma 6.3.3] we can deduce that there exists  
 1097 a constant  $\rho > 0$  such that for all  $w \in B_\rho^W(\bar{w})$  the coefficients  $\tilde{c}(\cdot)$  and  $c(\cdot)$  satisfy  
 1098 on  $[t_1, t_2] \times \mathbb{R}$  the one sided Lipschitz condition (OSLC) (cf. [3]), i.e. there exists  
 1099  $\alpha \in L^1(t_1, t_2)$  such that

$$1100 \quad (111) \quad c_x(t, \cdot) \leq \alpha(t) \text{ for a.a. } t \in [t_1, t_2]$$

1102 holds in the sense of distributions on  $\mathbb{R}$ .

1103 We now consider

$$\begin{aligned} 1104 \quad (112) \quad & p_t + \tilde{c}p_x = -\tilde{d}p \text{ on } (t_1, t_2) \times \mathbb{R}, \quad p(t_2, \cdot) = p^{t_2}(\cdot) \text{ on } \mathbb{R} \\ 1105 \quad (113) \quad & \text{and } p_t + cp_x = -dp \text{ on } (t_1, t_2) \times \mathbb{R}, \quad p(t_2, \cdot) = p^{t_2}(\cdot) \text{ on } \mathbb{R}. \end{aligned}$$

1107 Since (109), (110) and (111) hold, we can apply [40, Theorem 14] yielding that (112)  
 1108 and (113) admit reversible solutions

$$1109 \quad (114) \quad \hat{p} \in C^{0,1}([t_1, t_2] \times \mathbb{R})$$

$$1110 \quad (115) \quad p \in C^{0,1}([t_1, t_2] \times \mathbb{R}),$$

1112 respectively. In what follows, a reversible solution has to be understood according  
 1113 to [40, Definition 12]. Moreover, (109), (110) and (111) yield

$$1114 \quad (116) \quad \hat{p} \rightarrow p \text{ in } C([t_1, t_2] \times [-R, R])$$

1116 for all  $R > 0$  (cf. [40, Theorem 16]).

1117 Before we start to derive an adjoint representation for the term in (106), we first  
 1118 give a brief overview of the main steps and basic concepts of the proof.

1119 Since  $\bar{w}$  satisfies (ND), for sufficiently small  $\bar{\varepsilon}$  the points  $x = x_s(\bar{w}) - \bar{\varepsilon}$  and  
 1120  $x = x_s(\bar{w}) + \bar{\varepsilon}$  are points of continuity of  $\bar{y}(\bar{t}, \cdot)$ . Therefore, there exist unique backward  
 1121 characteristics  $\zeta_{l/r}$  through the points  $x_s(\bar{w}) \mp \bar{\varepsilon}$  such that we can define the set

$$\begin{aligned} 1122 \quad D^{\bar{\varepsilon}} &:= \{(t, x) \in [0, \bar{t}_m^a] \times [0, b] : 0 \leq x \leq \zeta_r(t)\} \\ 1123 \quad (117) \quad & \cup \{(t, x) \in [\bar{t}_m^a, \bar{t}] \times [0, b] : \zeta_l(t) \leq x \leq \zeta_r(t)\}. \end{aligned}$$

1125 Due to the stability of backward characteristics it holds that

$$1126 \quad D^0 = \{(t, x) \in [0, \bar{t}_m^a] \times [0, b] : 0 \leq x \leq \xi_r(t)\} \\ 1127 \quad \cup \{(t, x) \in [\bar{t}_m^a, \bar{t}] \times [0, b] : \xi_l(t) \leq x \leq \xi_r(t)\}.$$

1129 The main idea of this proof is equal to the proof of Lemma 4.10 in [34], but since  
1130 in the present case the shifting of rarefaction centers in the initial and boundary data is  
1131 allowed, we have to find a slightly different method to derive an adjoint representation  
1132 of (106). Hereby, we will follow the ideas that are used in [35] and based on the fact  
1133 that if the source term  $g$  is equal to zero, then the local solution near a rarefaction  
1134 center is explicitly known. The proof consists of the following main steps (we highly  
1135 recommend the reader to have a look at Figure 2):

1136 In **step 1** we will choose some sufficiently small  $\tilde{t} > \bar{t}_m^a$  with  $\bar{t}_m^a < \tilde{t} < \bar{t}_{m+1}^0$  such  
1137 that

$$1138 \quad (119) \quad \tilde{x} := M_{f'}(\tilde{t} - \bar{t}_m^a) < \frac{\varepsilon g}{2} \quad \text{and} \quad \bar{y}(\tilde{t}, \cdot) \in C^{0,1}([0, 2\tilde{x}])$$

1140 holds and define the set

$$1141 \quad (120) \quad D_{1,\varepsilon} := \{(t, x) \in [\tilde{t}, \bar{t}] \times \Omega : \zeta_l(t) \leq x \leq \zeta_r(t)\}.$$

1143 Then we consider (108) on  $[\tilde{t}, \bar{t}] \times \Omega$ , multiply it with the reversible solution  $\hat{p}_1$  of (112)  
1144 for  $[t_1, t_2] = [\tilde{t}, \bar{t}]$  and enddata  $p^{\tilde{t}} = \frac{1}{\bar{y}(t, x_s(\bar{w}))}$  and apply integration by parts on  $D_{1,\varepsilon}$ .  
1145 Since  $[\tilde{t}, \bar{t}] \times \Omega$  contains no rarefaction center of  $\bar{y}$ , (111) holds such that using (116)  
1146 yields that (106) is up to some  $o(\|\delta w\|_W)$  and  $\bar{\varepsilon}O(\|\delta w\|_W)$  terms equal to some terms  
1147  $I_{1,\varepsilon} + I_{2,\varepsilon}$ , which are defined in (136). The remaining steps are concerned with further  
1148 simplifying  $I_{1,\varepsilon}$  and  $I_{2,\varepsilon}$ .

1149 In **step 2**, using the explicitly known local solution near the rarefaction center  $\bar{t}_m^a$ ,  
1150 we will derive a representation of  $I_{1,\varepsilon}$  in (141), which depends on the local solution  
1151 near the rarefaction center and the adjoint state in Definition 15.

1152 **Step 3** is concerned with the simplification of  $I_{2,\varepsilon}$ . To this end, we define the  
1153 domains

$$1154 \quad (121) \quad D_{2,\varepsilon} := \{(t, x) \in [\hat{t}, \bar{t}] \times [0, b] : \max\{0, f'(\bar{u}_m^{B,a}(\bar{t}_m^a - \varepsilon))(t - \bar{t}_m^a + \varepsilon)\} \leq x \leq \xi_r(t)\},$$

$$1155 \quad (122) \quad D_2 := \{(t, x) \in [\hat{t}, \bar{t}] \times [0, b] : \max\{0, f'(\bar{u}_m^{B,a}(\bar{t}_m^a))(t - \bar{t}_m^a)\} \leq x \leq \xi_r(t)\},$$

1157 where  $\hat{t} > \bar{t}_{m-1}^a$  and  $\varepsilon$  is chosen according to (137). We note that due to the rarefaction  
1158 center of  $\bar{y}$  in  $(\bar{t}_m^a, 0)$  the OSLC (111) is violated. To solve this problem, we replace  
1159 the coefficients  $c$  and  $\tilde{c}$  on  $[\hat{t}, \bar{t}] \times \mathbb{R}$  by suitable  $\tilde{c}_{\text{loc}}, c_{\text{loc}}$  such that the OSLC (111)  
1160 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  $w \in B_\rho^W(\bar{w})$   
1161 with a sufficiently small  $\rho > 0$ . Let (108) be considered on  $[\hat{t}, \bar{t}] \times \Omega$ , multiply it with  
1162 the reversible solution  $\hat{p}_2$  of (112) for  $[t_1, t_2] = [\hat{t}, \bar{t}]$ ,  $\tilde{c}$  replaced by  $\tilde{c}_{\text{loc}}$  and enddata  
1163  $p^{\hat{t}} = p_1(\hat{t}, \cdot)$ . Then we apply integration by parts on  $D_{2,\varepsilon}$ . Observing that (109), (110)  
1164 and (111) are satisfied, we can use (116) to rewrite  $I_{2,\varepsilon}$  in terms of the boundary data,  
1165 the reversible solution  $p_2$  of (113) with  $c$  replaced by  $c_{\text{loc}}$  and the term  $I_{21}$ , which is  
1166 defined in (156).

1167 In **step 4**, we have a closer look at the term  $I_{21}$ . Choosing a time point  $s$  with  
1168  $0 < s < \varepsilon_g$ , we define the set

$$1169 \quad (123) \quad D_3 := \{(t, x) \in [s, \hat{t}] \times [0, b] : 0 \leq x \leq \xi_r(t)\}.$$



1192 by parts on the domain  $D_{1,\bar{\varepsilon}}$  yielding

$$\begin{aligned}
1193 \quad & \frac{1}{[\bar{y}(\bar{t}, x_s(\bar{w}))]} \int_{x_s(\bar{w})-\bar{\varepsilon}}^{x_s(\bar{w})+\bar{\varepsilon}} \Delta y(\bar{t}, x) dx = \int_{\zeta_l(\bar{t})}^{\zeta_r(\bar{t})} \hat{p}_1(\bar{t}, x) \Delta y(\bar{t}, x) dx \\
1194 \quad & + \int_{D_{1,\bar{\varepsilon}}} \hat{p}_1(t, x) (g(t, x, \bar{y}, u_1) - g(t, x, \bar{y}, \bar{u}_1)) dx dt \\
1195 \quad & - \int_{\bar{t}}^{\bar{t}} \hat{p}_1(t, \zeta_l(t)) \Delta y(t, \zeta_l(t)) (\tilde{c}(t, \zeta_l(t)) - c(t, \zeta_l(t))) dt \\
1196 \quad (127) \quad & + \int_{\bar{t}}^{\bar{t}} \hat{p}_1(t, \zeta_r(t)) \Delta y(t, \zeta_r(t)) (\tilde{c}(t, \zeta_r(t)) - c(t, \zeta_r(t))) dt. \\
1197
\end{aligned}$$

1198 Using (126) and the regularity of  $g$  w.r.t.  $u_1$ , we can rewrite the second term on the  
1199 right hand side of (127) by

$$\begin{aligned}
1200 \quad & \int_{D_{1,\bar{\varepsilon}}} \hat{p}_1(t, x) (g(t, x, \bar{y}, u_1) - g(t, x, \bar{y}, \bar{u}_1)) dx dt \\
1201 \quad & = \int_{D_{1,\bar{\varepsilon}}} p_1(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 dx dt + o(\|\delta w\|_W) \\
1202 \quad & = \int_{D_{1,\bar{\varepsilon}} \setminus D_1} p_1(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 dx dt \\
1203 \quad & + \int_{D_1} p_1(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 dx dt + o(\|\delta w\|_W) \\
1204
\end{aligned}$$

1205 Hereby, the term  $o(\|\delta w\|_W)$  is uniform w.r.t.  $\bar{\varepsilon} > 0$ . From [33, Lemma 3.1.15] we can  
1206 deduce that

$$1207 \quad (128) \quad \|\xi_{l/r}(\cdot) - \zeta_{l/r}(\cdot)\|_{C([\bar{t}, \bar{t}])} \leq C\bar{\varepsilon}.$$

1209 This and the boundedness of  $g_{u_1}(\cdot, \bar{y}, \bar{u}_1)$  yield

$$1210 \quad \int_{D_{1,\bar{\varepsilon}} \setminus D_1} p_1(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 dx dt = \bar{\varepsilon} O(\|\delta w\|_W)$$

1212 and hence

$$\begin{aligned}
1213 \quad & \int_{D_{1,\bar{\varepsilon}}} \hat{p}_1(t, x) (g(t, x, \bar{y}, u_1) - g(t, x, \bar{y}, \bar{u}_1)) dx dt = \int_{D_1} p_1(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 dx dt \\
1214 \quad (129) \quad & + \bar{\varepsilon} O(\|\delta w\|_W) + o(\|\delta w\|_W).
\end{aligned}$$

1216 Since  $\xi_r$  ends in a point where  $\bar{u}_0$  is smooth and  $\xi_l$  ends in the inner of a rarefaction  
1217 center we can use [33, Lemma 6.2.1] yielding smooth local solutions  $w \mapsto Y_{l/r}(\cdot, w)$   
1218 defined on some stripes  $S_{l/r}$ . Using those local solutions, the definitions of  $\tilde{c}$  and  $c$   
1219 and the uniform boundedness of  $\hat{p}_1$ , we can show that the last two integrals on the  
1220 right hand side of (127) are equal to  $O(\|\delta w\|_W^2)$ :

$$\begin{aligned}
1221 \quad & - \int_{\bar{t}}^{\bar{t}} \hat{p}_1(t, \zeta_l(t)) \Delta y(t, \zeta_l(t)) (\tilde{c}(t, \zeta_l(t)) - c(t, \zeta_l(t))) dt \\
1222 \quad (130) \quad & + \int_{\bar{t}}^{\bar{t}} \hat{p}_1(t, \zeta_r(t)) \Delta y(t, \zeta_r(t)) (\tilde{c}(t, \zeta_r(t)) - c(t, \zeta_r(t))) dt = O(\|\delta w\|_W^2) \\
1223
\end{aligned}$$

1224 In order to derive  $Y_l$ , we apply [33, Lemma 6.2.1] to IBVP on the truncated space-time  
 1225 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  
 1226 (130) is uniform w.r.t.  $\bar{\varepsilon}$ .

1227 Considering the first integral on the right hand side of (127), we can use (126)  
 1228 and Corollary 7 to show that

$$1229 \quad (131) \quad \int_{\zeta_l(\tilde{t})}^{\zeta_r(\tilde{t})} \hat{p}_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx = \int_{\zeta_l(\tilde{t})}^{\zeta_r(\tilde{t})} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx + o(\|\delta w\|_W),$$

1231 where  $o(\|\delta w\|_W)$  is uniform w.r.t.  $\bar{\varepsilon}$ . We further obtain

$$1232 \quad \int_{\zeta_l(\tilde{t})}^{\zeta_r(\tilde{t})} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx = \int_{\zeta_l(\tilde{t})}^{\xi_l(\tilde{t})} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx$$

$$1233 \quad (132) \quad + \int_{\xi_l(\tilde{t})}^{\xi_r(\tilde{t})} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx + \int_{\xi_r(\tilde{t})}^{\zeta_r(\tilde{t})} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx.$$

1235 With the help of the local solutions  $Y_{l/r}$  that we used to prove (130) and the bounded-  
 1236 ness of  $p_1$ , we can show the estimation

$$1237 \quad \left| \int_{\zeta_l(\tilde{t})}^{\xi_l(\tilde{t})} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx \right| + \left| \int_{\xi_r(\tilde{t})}^{\zeta_r(\tilde{t})} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx \right|$$

$$1238 \quad \leq (|\xi_l(\tilde{t}) - \zeta_l(\tilde{t})| + |\xi_r(\tilde{t}) - \zeta_r(\tilde{t})|) \cdot O(\|\delta w\|_W),$$

1240 where the term  $O(\|\delta w\|_W)$  is uniform w.r.t.  $\bar{\varepsilon}$ . Using this result and (128), we can  
 1241 deduce that

$$1242 \quad (133) \quad \left| \int_{\zeta_l(\tilde{t})}^{\xi_l(\tilde{t})} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx \right| + \left| \int_{\xi_r(\tilde{t})}^{\zeta_l(\tilde{t})} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx \right| \leq \bar{\varepsilon} O(\|\delta w\|_W).$$

1244 Inserting (133) in (132), we obtain that (131) can be rewritten as

$$1245 \quad (134) \quad \int_{\zeta_l(\tilde{t})}^{\zeta_r(\tilde{t})} \hat{p}_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx = \int_{\xi_l(\tilde{t})}^{\xi_r(\tilde{t})} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx + \bar{\varepsilon} O(\|\delta w\|_W) + o(\|\delta w\|_W).$$

1247 Inserting (129), (130) and (134) in (106) yields

$$1248 \quad \frac{1}{[\bar{y}(\tilde{t}, x_s(\bar{w}))]} \int_{x_s(\bar{w}) - \bar{\varepsilon}}^{x_s(\bar{w}) + \bar{\varepsilon}} \Delta y(\tilde{t}, x) \, dx = \int_{\xi_l(\tilde{t})}^{\xi_r(\tilde{t})} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx$$

$$1249 \quad (135) \quad + \int_{D_1} p_1(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 \, dx \, dt + o(\|\delta w\|_W) + \bar{\varepsilon} O(\|\delta w\|_W).$$

1251 We observe that due to [33, Lemma 6.3.5],  $p_1$  coincides on  $D_1$  with the adjoint state  $p$   
 1252 defined in Definition 15 with enddata  $p^{\tilde{t}}(\cdot) = \mathbb{1}_{x_s(\bar{w})}(\cdot) \frac{1}{[\bar{y}(\tilde{t}, x_s(\bar{w}))]}$ . We will now have a  
 1253 closer look at the first term of the right hand side of (135). For all  $\varepsilon \geq 0$  small enough  
 1254 it holds that

$$1255 \quad \int_{\xi_l(\tilde{t})}^{\xi_r(\tilde{t})} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx = \int_{\xi_l(\tilde{t})}^{f'(\bar{u}_m^{B,a}(\bar{t}_m^a - \varepsilon) \cdot (\tilde{t} - \bar{t}_m^a + \varepsilon))} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx$$

$$1256 \quad (136) \quad + \int_{f'(\bar{u}_m^{B,a}(\bar{t}_m^a - \varepsilon) \cdot (\tilde{t} - \bar{t}_m^a + \varepsilon))}^{\xi_r(\tilde{t})} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx =: I_{1,\varepsilon} + I_{2,\varepsilon}.$$

1258 In what follows, we choose  $\varepsilon$  as a function of  $\delta w$  satisfying

$$1259 \quad (137) \quad \varepsilon(\delta w) \geq 2\|\delta w\|_W \quad \text{and} \quad \varepsilon(\delta w) \rightarrow 0 \quad \text{if} \quad \|\delta w\|_W \rightarrow 0.$$

1261 **Step 2:** Considering  $I_{1,\varepsilon}$ , we note that since  $\xi_l$  ends in the inner of a rarefaction wave,  
1262 for all  $w \in B_\rho^W(\bar{w})$  with  $\rho > 0$  small enough it holds that

$$1263 \quad (138) \quad y(\tilde{t}, x, w) = \begin{cases} f'^{-1}\left(\frac{x}{\tilde{t}-t_m^a}\right) & \text{if } x \in [f'(u_{m+1}^{B,a}(t_m^a)) \cdot (\tilde{t}-t_m^a), f'(u_m^{B,a}(t_m^a)) \cdot (\tilde{t}-t_m^a)], \\ Y(\tilde{t}, x, w) & \text{if } x \in [f'(u_m^{B,a}(t_m^a)) \cdot (\tilde{t}-t_m^a), f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t}-\bar{t}_m^a) + \delta), \end{cases}$$

1265 where  $\delta > 0$  is a small constant and  $w \in B_\rho^W(\bar{w}) \ni w \rightarrow Y(\tilde{t}, \cdot, w) \in C(I)$  is a smooth  
1266 function obtained by [33, Lemma 6.2.7] with

$$1267 \quad I = (f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t} - \bar{t}_m^a) - \delta, f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t} - \bar{t}_m^a) + \delta).$$

1269 We note that  $y(\tilde{t}, x, w)$  is continuous in  $x = f'(u_m^{B,a}(t_m^a)) \cdot (\tilde{t} - t_m^a)$  for all  $w \in B_\rho^W(\bar{w})$ .

1270 Using a Taylor approximation of  $f'(u_m^{B,a}(t_m^a)) \cdot (\tilde{t} - t_m^a)$  in  $u_m^{B,a} = \bar{u}_m^{B,a}$  and  $t_m^0 = \bar{t}_m^0$ ,  
1271 we can deduce that there exists a constant  $C > 0$  such that

$$1272 \quad \xi_l(\tilde{t}) < f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t} - \bar{t}_m^a) - C\|\delta w\|_W \quad \text{and} \\ 1273 \quad (139) \quad [\xi_l(\tilde{t}), f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t} - \bar{t}_m^a) - C\|\delta w\|_W] \subset [\xi_l(\tilde{t}), f'(u_m^{B,a}(t_m^a)) \cdot (\tilde{t} - t_m^a)]$$

1275 hold for all  $w \in B_\rho^W(\bar{w})$  with  $\rho > 0$  small enough. Then (138) and (139) yield

$$1276 \quad (140) \quad \Delta y(\tilde{t}, x)|_{[\xi_l(\tilde{t}), f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t} - \bar{t}_m^a) - C\|\delta w\|_W]} = \frac{x \cdot \delta t_m^a}{f''(f'^{-1}(\frac{x}{\tilde{t}-\bar{t}_m^a})) \cdot (\tilde{t} - \bar{t}_m^a)^2} + o(\|\delta w\|_W).$$

1278 Using (140), we can rewrite the term  $I_{1,\varepsilon}$  in (136) as follows:

$$1279 \quad I_{1,\varepsilon} = \int_{\xi_l(\tilde{t})}^{f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t} - \bar{t}_m^a) - C\|\delta w\|_W} p_1(\tilde{t}, x) \frac{x \cdot \delta t_m^a}{f''(f'^{-1}(\frac{x}{\tilde{t}-\bar{t}_m^a})) \cdot (\tilde{t} - \bar{t}_m^a)^2} dx \\ 1280 \quad + \int_{f'(\bar{u}_m^{B,a}(\bar{t}_m^a - \varepsilon)) \cdot (\tilde{t} - \bar{t}_m^a + \varepsilon)}^{f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t} - \bar{t}_m^a) - C\|\delta w\|_W} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) dx + o(\|\delta w\|_W) \\ 1281 \quad = \int_{\xi_l(\tilde{t})}^{f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t} - \bar{t}_m^a)} p_1(\tilde{t}, x) \frac{x \cdot \delta t_m^a}{f''(f'^{-1}(\frac{x}{\tilde{t}-\bar{t}_m^a})) \cdot (\tilde{t} - \bar{t}_m^a)^2} dx \\ 1282 \quad - \int_{f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t} - \bar{t}_m^a) - C\|\delta w\|_W}^{f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t} - \bar{t}_m^a)} p_1(\tilde{t}, x) \frac{x \cdot \delta t_m^a}{f''(f'^{-1}(\frac{x}{\tilde{t}-\bar{t}_m^a})) \cdot (\tilde{t} - \bar{t}_m^a)^2} dx \\ 1283 \quad + \int_{f'(\bar{u}_m^{B,a}(\bar{t}_m^a - \varepsilon)) \cdot (\tilde{t} - \bar{t}_m^a + \varepsilon)}^{f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t} - \bar{t}_m^a) - C\|\delta w\|_W} p_1(\tilde{t}, x) \Delta y(\tilde{t}, x) dx + o(\|\delta w\|_W)$$

1285 Due to the regularity of  $f$  and  $p_1$ , the second integral on the right hand side is equal  
1286 to  $o(\|\delta w\|_W)$ . Using again the regularity of  $p_1$ , the choice of  $\varepsilon$  in (137) and the

1287 representation of  $y(\tilde{t}, x, w)$  in (138), which is in particular continuous, we obtain that  
 1288 the third integral is also equal to  $o(\|\delta w\|_W)$ . Hence, we can conclude

$$1289 \quad I_{1,\varepsilon} = \int_{\xi_l(\tilde{t})}^{f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t} - \bar{t}_m^a)} p_1(\tilde{t}, x) \frac{x \cdot \delta t_m^a}{f''(f'^{-1}(\frac{x}{\tilde{t} - \bar{t}_m^a}))(\tilde{t} - \bar{t}_m^a)^2} dx + o(\|\delta w\|_W).$$

1290

1292 Since  $p_1$  coincides with the adjoint  $p$  on  $D_1$ , we further obtain

$$1293 \quad I_{1,\varepsilon} = \int_{\xi_l(\tilde{t})}^{f'(\bar{u}_m^{B,a}(\bar{t}_m^a)) \cdot (\tilde{t} - \bar{t}_m^a)} p(\tilde{t}, x) \frac{x \cdot \delta t_m^a}{f''(f'^{-1}(\frac{x}{\tilde{t} - \bar{t}_m^a})) \cdot (\tilde{t} - \bar{t}_m^a)^2} dx + o(\|\delta w\|_W)$$

$$1294 \quad = \int_{\xi_l(\tilde{t})/(\tilde{t} - \bar{t}_m^a)}^{f'(\bar{u}_m^{B,a}(\bar{t}_m^a))} p(\tilde{t}, z(\tilde{t} - \bar{t}_m^a)) \frac{z \cdot \delta t_m^a}{f''(f'^{-1}(z))} dz + o(\|\delta w\|_W)$$

$$1295 \quad (141) \quad = \int_{f'(\bar{u}_{m+1}^{B,a}(\bar{t}_m^a))}^{f'(\bar{u}_m^{B,a}(\bar{t}_m^a))} \lim_{t \searrow \bar{t}_m^a} p(t, z(t - \bar{t}_m^a)) \frac{z \cdot \delta t_m^a}{f''(f'^{-1}(z))} dz + o(\|\delta w\|_W).$$

1296

1297 The last equality holds due to the fact that for all  $z \in [f'(\bar{u}_{m+1}^{B,a}(\bar{t}_m^a)), f'(\bar{u}_m^{B,a}(\bar{t}_m^a))]$   
 1298 the term  $p(\cdot, z(\cdot - \bar{t}_m^a))$  is constant on  $(\bar{t}_m^a, \tilde{t}]$  and is equal to zero for all

$$1299 \quad z \in [f'(\bar{u}_{m+1}^{B,a}(\bar{t}_m^a)), \xi_l(\tilde{t})/(\tilde{t} - \bar{t}_m^a)].$$

1301 **Step 3:** In order to compute  $I_{2,\varepsilon}$ , we define suitable coefficients  $\tilde{c}_{\text{loc}}, c_{\text{loc}}$  such  
 1302 that  $\tilde{c}_{\text{loc}}|_{D_{2,\varepsilon}} \equiv \tilde{c}$  and  $c_{\text{loc}}|_{D_{2,\varepsilon}} \equiv c$  hold for all  $w \in B_\rho^W(\bar{w})$  with a sufficiently small  
 1303  $\rho > 0$  such that (109), (110) and (111) are satisfied. To this end, we choose some  
 1304  $\hat{t} \in (\bar{t}_{m-1}^a, \bar{t}_m^a)$  and define  $\tilde{c}_{\text{loc}}$  and  $c_{\text{loc}}$  on  $[\hat{t}, \tilde{t}] \times \mathbb{R}$  by

$$1305 \quad \tilde{c}_{\text{loc}}(t, x) = M_{f'}, \text{ if } (t, x) \in [\hat{t}, \tilde{t}] \times [0, \xi_r(t)] \setminus D_{2, \frac{\varepsilon}{2}}, \quad \tilde{c}_{\text{loc}}(t, x) = \tilde{c}(t, x), \text{ else}$$

$$1306 \quad c_{\text{loc}}(t, x) = M_{f'}, \text{ if } (t, x) \in [\hat{t}, \tilde{t}] \times [0, \xi_r(t)] \setminus D_2, \quad c_{\text{loc}}(t, x) = c(t, x), \text{ else.}$$

1308 Recalling the sets  $D_{2,\varepsilon}$  and  $D_2$  defined in (121) and (122), we choose the function  
 1309  $\varepsilon(\delta w)$  in (137) such that

$$1310 \quad (142) \quad D_{2,\varepsilon} \subset D_{2, \frac{\varepsilon}{2}} \subset D_2 \quad \text{for all } w \in B_\rho^W(\bar{w})$$

1312 for sufficiently small  $\rho > 0$ .

1313 Due to the construction of  $\tilde{c}_{\text{loc}}$  and  $c_{\text{loc}}$  and the choice of  $\varepsilon$  in (137), one can show  
 1314 that (109), (110) and (111) hold on  $[\hat{t}, \tilde{t}] \times \mathbb{R}$ . Therefore,

$$1315 \quad (143) \quad p_t + \tilde{c}_{\text{loc}} p_x = -\tilde{d}p \quad \text{on } (\hat{t}, \tilde{t}) \times \mathbb{R}, \quad p(\tilde{t}, \cdot) = p_1(\tilde{t}, \cdot) \quad \text{on } \mathbb{R}$$

$$1316 \quad (144) \quad \text{and } p_t + c_{\text{loc}} p_x = -dp \quad \text{on } (\hat{t}, \tilde{t}) \times \mathbb{R}, \quad p(\tilde{t}, \cdot) = p_1(\tilde{t}, \cdot) \quad \text{on } \mathbb{R}$$

1318 admit reversible solutions  $\hat{p}_2, p_2 \in C^{0,1}([\hat{t}, \tilde{t}] \times \mathbb{R})$ , respectively, satisfying

$$1319 \quad (145) \quad \hat{p}_2 \rightarrow p_2 \text{ in } C([\hat{t}, \tilde{t}] \times [-R, R]) \text{ if } \|\delta w\|_W \rightarrow 0$$

1321 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,  
 1322 if we consider (108) on  $(\hat{t}, \tilde{t}) \times \mathbb{R}$ , multiply it with  $\hat{p}_2$  and apply integration by parts

1323 on  $D_{2,\varepsilon}$ , then we obtain

$$\begin{aligned}
1324 \quad I_{2,\varepsilon} &= \int_{D_{2,\varepsilon}} \hat{p}_2(t, x)(g(t, x, \bar{y}, u_1) - g(t, x, \bar{y}, \bar{u}_1)) \, dx \, dt + \int_0^{\xi_r(\tilde{t})} \hat{p}_2(\tilde{t}, x) \Delta y(\tilde{t}, x) \, dx \\
1325 \quad &+ \int_{\bar{t}_m^a - \varepsilon}^{\tilde{t}} \hat{p}_2(t, \gamma(t, \varepsilon)) \Delta y(t, \gamma(t, \varepsilon)) (f'(\bar{u}_m^{B,a}(\bar{t}_m^a - \varepsilon)) - \tilde{c}(t, \gamma(t, \varepsilon))) \, dt \\
1326 \quad &- \int_{\tilde{t}}^{\bar{t}_m^a - \varepsilon} \hat{p}_2(t, \xi_r(t)) \Delta y(t, \xi_r(t)) (f'(\bar{y}(t, \xi_r(t))) - \tilde{c}(t, \xi_r(t))) \, dt \\
1327 \quad (146) \quad &+ \int_{\tilde{t}}^{\bar{t}_m^a - \varepsilon} \hat{p}_2(t, 0) \tilde{c}(t, 0+) \Delta y(t, 0+) \, dt =: I_{20} + I_{21} + I_{22} + I_{23} + I_{24}, \\
1328
\end{aligned}$$

1329 where  $\gamma(t, \varepsilon) := f'(\bar{u}_m^{B,a}(\bar{t}_m^a - \varepsilon))(t - \bar{t}_m^a + \varepsilon)$ .

1330 Concerning (146), we note that  $g(\cdot, y(w), u_1) = 0$  holds on  $D_{2,0} \setminus D_{2,\varepsilon}$  for all  
1331  $w \in B_\rho^W(\bar{w})$ . From this result, the regularity of  $g$  and (145) we obtain

$$1332 \quad (147) \quad I_{20} = (p_2(\cdot), g_{u_1}(\cdot, \bar{y}, \bar{u}_1) \delta u_1)_{2, D_{2,0}} + o(\|\delta w\|_W).$$

1334 Now, we show that

$$1335 \quad (148) \quad \|I_{22}\| \leq C \|\delta w\|_W^2$$

1337 and note that

$$1338 \quad (149) \quad \|I_{23}\| \leq C \|\delta w\|_W^2$$

1340 can be proved analogously.

1341 To this end, we note that since the point  $(\tilde{t}, f'(\bar{u}_m^{B,a}(\bar{t}_m^a))(\tilde{t} - \bar{t}_m^a))$  lies on the right  
1342 boundary of the rarefaction wave emanating from  $(\bar{t}_m^a, 0)$ , [33, Lemma 6.2.7] implies  
1343 that there exists a stripe  $S$  with

$$\begin{aligned}
1344 \quad &\{(t, x) \in [\bar{t}_m^a, \tilde{t}] \times \Omega : x = f'(\bar{u}_m^{B,a}(\bar{t}_m^a))(t - \bar{t}_m^a)\} \subset S \\
1345 \quad (150) \quad &\{(t, x) \in [\bar{t}_m^a - \delta, \tilde{t}] \times \Omega : x = M_{f'}(t - \bar{t}_m^a + \delta)\} \subset S,
\end{aligned}$$

1347 and a continuously Fréchet-differentiable mapping  $B_\rho^W(\bar{w}) \ni w \mapsto Y_m(\cdot; w) \in C(S)$   
1348 such that

$$1349 \quad (151) \quad y(\cdot; w)|_{\hat{S}} = Y_m(\cdot; w) \text{ for all } w \in B_\rho^W(\bar{w})$$

1351 holds, where

$$1352 \quad (152) \quad \hat{S} := \left\{ (t, x) \in S : x > \max\left\{0, f'(u_m^{B,a}(\bar{t}_m^a)) \cdot (t - \bar{t}_m^a + \frac{\varepsilon}{2})\right\} \right\}$$

1354 provided that  $\rho > 0$  and  $\delta$  are small enough and  $\tilde{t} > \bar{t}_m^a$  is sufficiently close to  $\bar{t}_m^a$ .  
1355 Hereby, (150) implies that  $\hat{S}$  is nonempty.

1356 We observe that we can choose  $\varepsilon > 0$  in (137) such that  $(t, \gamma(t, \varepsilon)) \in \hat{S}$  holds for  
1357 all  $t \in [\bar{t}_m^a - \varepsilon, \tilde{t}]$  and all  $w \in B_\rho^W(\bar{w})$ . Therefore, (151) yields for all  $w \in B_\rho^W(\bar{w})$

$$\begin{aligned}
1358 \quad &\|\Delta y(t, \gamma(t, \varepsilon))\|_{C([\bar{t}_m^a - \varepsilon, \tilde{t}])} = \|Y_m(t, \gamma(t, \varepsilon); w) - Y_m(t, \gamma(t, \varepsilon); \bar{w})\|_{C([\bar{t}_m^a - \varepsilon, \tilde{t}])} \\
1359 \quad &\leq \|Y_m(\cdot; w) - Y_m(\cdot; \bar{w})\|_{C(\hat{S})} = \left\| \frac{\partial}{\partial w} Y_m(\cdot; \bar{w}) \right\|_{C(\hat{S})} \cdot \|\delta w\|_W + o(\|\delta w\|_W) \\
1360 \quad (153) \quad &\leq C \|\delta w\|_W.
\end{aligned}$$

1362 Next, we note that  $I_{22}$  in (146) can be estimated from above by

$$1363 \quad (154) \quad |I_{22}| \leq C \int_{\hat{t}_m^a - \varepsilon}^{\hat{t}} |\Delta y(t, \gamma(t, \varepsilon))|^2 dt,$$

1365 where we have used the uniform boundedness of  $\hat{p}_2$  and the regularity of  $f$ . From (153)  
1366 and (154) we obtain that (148) holds. Using similar arguments and (104), we obtain

$$1367 \quad (155) \quad I_{24} = (p_2(\cdot, 0), f'(\bar{y}(\cdot, 0+))\delta u_m^{B,a}(\cdot))_{2, [\hat{t}, \hat{t}_m^a]} + o(\|\delta w\|_W).$$

1369 Concerning  $I_{21}$ , from (145) and Corollary 7, we can deduce that

$$1370 \quad (156) \quad I_{21} = \int_0^{\xi_r(\hat{t})} p_2(\hat{t}, x) \Delta y(\hat{t}, x) dx + o(\|\delta w\|_W).$$

1372 Due to [33, Lemma 6.3.5],  $p_2$  and the adjoint state  $p$  coincide on  $D_{2,0}$ . Therefore, we  
1373 can replace  $p_2$  by  $p$  in the terms  $I_{20}$ ,  $I_{21}$  and  $I_{24}$ .

1374 **Step 4:** Next, we want to further simplify the term  $I_{21}$  in (156). To this end,  
1375 we choose some  $s \in (0, \varepsilon_g)$  and observe that since  $\tilde{c}$  has no rarefaction centers on the  
1376 domain  $[s, \hat{t}] \times \mathbb{R}$ , (109), (110) and (111) hold. Therefore,

$$1377 \quad (157) \quad p_t + \tilde{c}p_x = -\tilde{d}p \quad \text{on } (s, \hat{t}) \times \mathbb{R}, \quad p(\hat{t}, \cdot) = p_2(\hat{t}, \cdot) \quad \text{on } \mathbb{R}$$

$$1378 \quad (158) \quad \text{and } p_t + cp_x = -dp \quad \text{on } (s, \hat{t}) \times \mathbb{R}, \quad p(\hat{t}, \cdot) = p_2(\hat{t}, \cdot) \quad \text{on } \mathbb{R}$$

1380 admit reversible solutions  $\hat{p}_3, p_3 \in C^{0,1}([s, \hat{t}] \times \mathbb{R})$ , respectively. Moreover,

$$1381 \quad (159) \quad \hat{p}_3 \rightarrow p_3 \quad \text{in } C([s, \hat{t}] \times [-R, R]) \quad \text{if } \|\delta w\|_W \rightarrow 0$$

1383 holds for all  $R > 0$ . Now, we consider (108) on  $(s, \hat{t}) \times \Omega$ , multiply it with  $\hat{p}_3$  and  
1384 apply integration by parts on  $D_3$  defined in (123). This yields

$$1385 \quad I_{21} = \int_{D_3} \hat{p}_3(t, x) (g(t, x, \bar{y}, u_1) - g(t, x, \bar{y}, \bar{u}_1)) dx dt$$

$$1386 \quad + \int_s^{\hat{t}} \hat{p}_3(t, 0+) \tilde{c}(t, 0+) \Delta y(t, 0+) dt$$

$$1387 \quad - \int_s^{\hat{t}} \hat{p}_3(t, \xi_r(t)) \Delta y(t, \xi_r(t)) (c(t, \xi_r(t)) - \tilde{c}(t, \xi_r(t))) dt$$

$$1388 \quad + \int_0^{\xi_r(s)} \hat{p}_3(s, x) \Delta y(s, x) dx + o(\|\delta w\|_W)$$

$$1389 \quad (160) \quad =: I_{31} + I_{32} + I_{33} + I_{34} + o(\|\delta w\|_W).$$

1391 Analogously to the simplification of (146) in **Step 2**, one can also simplify the terms  
1392 on the right hand side of (160). Using the regularity of  $g$  and (159), we obtain that

$$1393 \quad (161) \quad I_{31} = \int_{D_3} p_3(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 dt dx + o(\|\delta w\|_W).$$

1395 Using that  $\xi_r$  ends in a point where the initial data is smooth and  $\bar{w}$  satisfies (ND),  
1396 we can use the same arguments as in the estimation of (148) and obtain

$$1397 \quad (162) \quad I_{33} = O(\|\delta w\|_W^2).$$

1399 Moreover, using (159) and Corollary 7, we obtain

$$1400 \quad (163) \quad I_{34} = \int_0^{\xi_r(s)} p_3(s, x) \Delta y(s, x) dx + o(\|\delta w\|_W).$$

1402 Now we want to have a closer look at the term  $I_{32}$ . Observing that  $\Delta y(\cdot, 0+) = \delta u_{B,a}$   
1403 holds on  $[s, \hat{t}]$  due to (104), Lemma 6 yields that  $\|\Delta y(\cdot, 0+)\|_{1, (s, \hat{t})} = O(\|\delta w\|_W)$ . Using  
1404 this result, (159) and the uniform boundedness of  $\tilde{c}$ , we can rewrite  $I_{32}$  by

$$1405 \quad (164) \quad I_{32} = \int_s^{\hat{t}} p_3(t, 0) \tilde{c}(t, 0+) \Delta y(t, 0+) dt + o(\|\delta w\|_W).$$

1407 Moreover, it holds that

$$1408 \quad \left\| \Delta y(\cdot, 0+) - \sum_{i=1}^{n_{t,a}+1} \delta u_i^{B,a}(\cdot) \mathbb{1}_{I_{B,a}^i(\bar{w})}(\cdot) - \sum_{i=1}^{n_{t,a}} \text{sgn}(\delta t_i^a) \mathbb{1}_{I(\bar{t}_i^a, \bar{t}_i^a + \delta t_i^a)}(\cdot) [\bar{u}_{B,a}(\bar{t}_i^a)] \right\|_{1, (s, \hat{t})}$$

$$1409 \quad = o(\|\delta w\|_W).$$

1411 Using this result and the uniform boundedness of  $\tilde{c}$  and  $p_3$ , we can rewrite the right  
1412 hand side of (164) by

$$1413 \quad \int_s^{\hat{t}} p_3(t, 0+) \tilde{c}(t, 0+) \Delta y(t, 0+) dt + o(\|\delta w\|_W)$$

$$1414 \quad = \sum_{i=1}^{n_{t,a}+1} (p_3(\cdot, 0), \tilde{c}(t, 0+) \delta u_i^{B,a}(\cdot))_{2, I_{B,a}^i \cap [s, \hat{t}]}$$

$$1415 \quad + \sum_{i \in I_{s,a}, s < \bar{t}_i^a < \hat{t}} (p_3(\cdot, 0), \tilde{c}(t, 0+) \cdot \text{sgn}(\delta t_i^a) [\bar{u}_{B,a}(\bar{t}_i^a)])_{2, I(\bar{t}_i^a, \bar{t}_i^a + \delta t_i^a)} + o(\|\delta w\|_W)$$

$$1416 \quad (165) \quad =: I_{321} + I_{322} + o(\|\delta w\|_W).$$

1418 Since  $\tilde{c}(\cdot, 0+) \rightarrow c(\cdot, 0+)$  in  $L^1([s, \hat{t}])$  due to Corollary 7 and  $p_3$  is bounded, we obtain  
1419 that

$$1420 \quad (166) \quad I_{321} = \sum_{i=1}^{n_{t,a}+1} (p_3(\cdot, 0), c(t, 0+) \delta u_i^{B,a}(\cdot))_{2, I_{B,a}^i \cap [s, \hat{t}]} + o(\|\delta w\|_W).$$

1422 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  
1423 be treated analogously and the case  $\delta t_i^a = 0$  is trivial. We observe that for all  $i \in I_{s,a}$   
1424 with  $s < \bar{t}_i^a < \hat{t}$  it holds that

$$1425 \quad \left\| [\bar{u}_{B,a}(\bar{t}_i^a)] - (u_i^{B,a}(\cdot) - \bar{u}_{i+1}^{B,a}(\cdot)) \right\|_{L^1(\bar{t}_i^a, \bar{t}_i^a + \delta t_i^a)} = o(\|\delta w\|_W).$$

1427 This together with the uniform boundedness of  $\tilde{c}$  and  $p_3$  yields

$$1428 \quad I_{322} = \sum_{i \in I_{s,a}, s < \bar{t}_i^a < \hat{t}} (p_3(\cdot, 0), \tilde{c}(\cdot, 0+) \cdot (u_i^{B,a}(\cdot) - \bar{u}_{i+1}^{B,a}(\cdot)))_{2, I(\bar{t}_i^a, \bar{t}_i^a + \delta t_i^a)} + o(\|\delta w\|_W)$$

$$1429 \quad (167) \quad = \sum_{i \in I_{s,a}, s < \bar{t}_i^a < \hat{t}} (p_3(\cdot, 0), f(u_i^{B,a}(\cdot)) - f(\bar{u}_{i+1}^{B,a}(\cdot)))_{2, I(\bar{t}_i^a, \bar{t}_i^a + \delta t_i^a)} + o(\|\delta w\|_W),$$

1430

1431 where the equality follows from the definition of  $\tilde{c}$ . We further observe that due to the  
1432 regularity of  $p_3$ ,  $f$  and  $u_j^{B,a}$  for  $j = 1, \dots, n_{t,a}$  it holds that

$$\begin{aligned}
1433 \quad I_{322} &= \sum_{i \in I_{s,a}, s < \bar{t}_i^a < \hat{t}} p_3(\bar{t}_i^a, 0)(f(\bar{u}_i^{B,a}(\bar{t}_i^a)) - f(\bar{u}_{i+1}^{B,a}(\bar{t}_i^a)))\delta t_i^a + o(\|\delta w\|_W) \\
1434 \quad (168) \quad &= \sum_{i \in I_{s,a}, s < \bar{t}_i^a < \hat{t}} p_3(\bar{t}_i^a, 0)[f(\bar{y}(\bar{t}_i^a, 0+))] \delta t_i^a + o(\|\delta w\|_W). \\
1435
\end{aligned}$$

1436 Inserting (166) and (168) in (165) and (165) in (164), we obtain

$$\begin{aligned}
1437 \quad I_{32} &= \int_s^{\hat{t}} p_3(t, 0+)c(t, 0+)\Delta y(t, 0+) dt + o(\|\delta w\|_W) \\
1438 \quad &= \sum_{i=1}^{n_{t,a}+1} (p_3(\cdot, 0), f'(\bar{u}_i^{B,a}(\cdot))\delta u_i^{B,a}(\cdot))_{2, I_{B,a}^i \cap [s, \hat{t}]} \\
1439 \quad (169) \quad &+ \sum_{i \in I_{s,a}, s < \bar{t}_i^a < \hat{t}} p_3(\bar{t}_i^a, 0)[f(\bar{y}(\bar{t}_i^a, 0+))] \delta t_i^a + o(\|\delta w\|_W). \\
1440
\end{aligned}$$

1441 Inserting (161), (162), (163) and (169) in (160) finally yields

$$\begin{aligned}
1442 \quad I_{21} &= (p_3(\cdot), g_{u_1}(\cdot, \bar{y}, \bar{u}_1)\delta u_1)_{2, D_3} \\
1443 \quad &+ (p_3(s, \cdot), \Delta y(s, \cdot))_{2, (0, \xi_r(s))} + \sum_{i=1}^{n_{t,a}+1} (p_3(\cdot, 0), f'(\bar{u}_i^{B,a})\delta u_i^{B,a})_{2, I_{B,a}^i \cap [s, \hat{t}]} \\
1444 \quad (170) \quad &+ \sum_{i \in I_{s,a}, s < \bar{t}_i^a < \hat{t}} p_3(\bar{t}_i^a, 0)[f(\bar{y}(\bar{t}_i^a, 0+))] \delta t_i^a + o(\|\delta w\|_W). \\
1445
\end{aligned}$$

1446 We note that  $p_3$  coincides with the adjoint state  $p$  on  $D_3$  such that we obtain from  
1447 (170)

$$\begin{aligned}
1448 \quad I_{21} &= (p(\cdot), g_{u_1}(\cdot, \bar{y}, \bar{u}_1)\delta u_1)_{2, D_3} \\
1449 \quad &+ (p(s, \cdot), \Delta y(s, \cdot))_{2, (0, \xi_r(s))} + \sum_{i=1}^{n_{t,a}+1} (p(\cdot, 0), f'(\bar{u}_i^{B,a})\delta u_i^{B,a})_{2, I_{B,a}^i \cap [s, \hat{t}]} \\
1450 \quad (171) \quad &+ \sum_{i \in I_{s,a}, s < \bar{t}_i^a < \hat{t}} p(\bar{t}_i^a, 0)[f(\bar{y}(\bar{t}_i^a, 0+))] \delta t_i^a + o(\|\delta w\|_W). \\
1451
\end{aligned}$$

1452 The second term of the right side of (171) can be rewritten as follows:

$$\begin{aligned}
1453 \quad &(p(s, \cdot), \Delta y(s, \cdot))_{2, (0, \xi_r(s))} \\
1454 \quad &= (p(s, \cdot), \Delta y(s, \cdot))_{2, (0, \bar{x}_l^0 - \varepsilon + f'(\bar{u}_l^0(\bar{x}_l^0 - \varepsilon)) \cdot s]} \\
1455 \quad &+ (p(s, \cdot), \Delta y(s, \cdot))_{2, (\bar{x}_l^0 - \varepsilon + f'(\bar{u}_l^0(\bar{x}_l^0 - \varepsilon)) \cdot s, \bar{x}_l^0 + \varepsilon + f'(\bar{u}_{l+1}^0(\bar{x}_l^0 + \varepsilon)) \cdot s]} \\
1456 \quad &+ (p(s, \cdot), \Delta y(s, \cdot))_{2, (\bar{x}_l^0 + \varepsilon + f'(\bar{u}_{l+1}^0(\bar{x}_l^0 + \varepsilon)) \cdot s, \xi_r(s))} \\
1457 \quad (172) \quad &= I_{1,\varepsilon}^0 + I_{2,\varepsilon}^0 + I_{3,\varepsilon}^0. \\
1458
\end{aligned}$$

1459 **Step 5:** Using similar arguments as in the estimation of  $I_{2,\varepsilon}$ , one can firstly show that

$$1460 \quad I_{1,\varepsilon}^0 = (p(0, \cdot), \delta u_0(\cdot))_{2, (0, \bar{x}_l^0 - \varepsilon)} + o(\|\delta w\|_W), \\
1461$$

1462 where the term  $o(\|\delta w\|_W)$  is uniform w.r.t.  $\varepsilon$ . Next, we observe that

$$1463 \quad \left\| \delta u_0(\cdot) - \sum_{i=1}^l \delta u_i^0(\cdot) \mathbb{1}_{I_0^i(\bar{w})}(\cdot) - \sum_{i=1}^{l-1} \operatorname{sgn}(\delta x_i^0) \mathbb{1}_{I(\bar{x}_i^0, \bar{x}_i^0 + \delta x_i^0)}(\cdot) [\bar{u}_0(\bar{x}_i^0)] \delta x_i^0 \right\|_{1, (0, \bar{x}_l^0 - \varepsilon)}$$

$$1464 \quad = o(\|\delta w\|_W),$$

1466 where the term  $o(\|\delta w\|_W)$  is uniform w.r.t.  $\varepsilon$  again. Hence, by the regularity of  $p$  and  
1467 the choice of  $\varepsilon$  in (137), we obtain

$$1468 \quad I_{1, \varepsilon}^0 = \sum_{i=1}^l (p(0, \cdot), \delta u_i^0(\cdot))_{2, I_0^i(\bar{w}) \cap [0, \xi_r(0)]} + \sum_{i=1}^{l-1} \mathbb{1}_{[0, \xi_r(0)]}(\bar{x}_i^0) p(0, \bar{x}_i^0) [\bar{u}_0(\bar{x}_i^0)] \delta x_i^0$$

$$1469 \quad + o(\|\delta w\|_W)$$

$$1470 \quad (173) \quad = \sum_{i=1}^l (p(0, \cdot), \delta u_i^0(\cdot))_{2, I_0^i(\bar{w})} + \sum_{i=1}^{l-1} p(0, \bar{x}_i^0) [\bar{u}_0(\bar{x}_i^0)] \delta x_i^0 + o(\|\delta w\|_W),$$

$$1471$$

1472 where the second equality holds due to the fact that the adjoint state  $p$  is by definition  
1473 equal to zero on  $\{(t, x) \in [0, \bar{t}] \times [0, b] : x > \xi_r(t)\}$ . Analogously, one can show that

$$1474 \quad (174) \quad I_{3, \varepsilon}^0 = \sum_{i=l+1}^{n_x+1} (p(0, \cdot), \delta u_i^0(\cdot))_{2, I_0^i(\bar{w})} + \sum_{i=l+1}^{n_x} p(0, \bar{x}_i^0) [\bar{u}_0(\bar{x}_i^0)] \delta x_i^0 + o(\|\delta w\|_W).$$

$$1475$$

1476 Similar to the estimation of  $I_{1, \varepsilon}$ , using that

$$1477 \quad y(s, x, w) \Big|_{f'(u_l^0(x_l^0)) \cdot s + x_l^0, f'(u_{l+1}^0(x_l^0)) \cdot s + x_l^0} = f'^{-1} \left( \frac{x - x_l^0}{s} \right) \quad \forall w \in B_\rho^W(\bar{w})$$

$$1478$$

1479 for sufficiently small  $\rho > 0$ , we can further show that

$$1480 \quad (175) \quad I_{2, \varepsilon}^0 = - \int_{f'(\bar{u}_l^0(\bar{x}_l^0))}^{f'(\bar{u}_{l+1}^0(\bar{x}_l^0))} \lim_{t \searrow 0} p(t, zt + \bar{x}_l^0) \frac{\delta x_l^0}{f''(f'^{-1}(z))} dz + o(\|\delta w\|_W),$$

$$1481$$

1482 where the term  $o(\|\delta w\|_W)$  is uniform w.r.t.  $\varepsilon$ . Inserting (173), (174) and (175) in  
1483 (172), we obtain

$$1484 \quad (p(s, \cdot), \Delta y(s, \cdot))_{2, (0, \xi_r(s))} = \sum_{i=1}^{n_x+1} (p(0, \cdot), \delta u_i^0(\cdot))_{2, I_0^i(\bar{w})}$$

$$1485 \quad - \int_{f'(\bar{u}_l^0(\bar{x}_l^0))}^{f'(\bar{u}_{l+1}^0(\bar{x}_l^0))} \lim_{t \searrow 0} p(t, zt + \bar{x}_l^0) \frac{\delta x_l^0}{f''(f'^{-1}(z))} dz$$

$$1486 \quad (176) \quad + \sum_{i \in I_{s, 0}(\bar{w})} p(0, \bar{x}_i^0) [\bar{u}_0(\bar{x}_i^0)] \delta x_i^0 + o(\|\delta w\|_W).$$

$$1487$$

1488 Inserting (176) in (171), we can deduce

$$\begin{aligned}
1489 \quad I_{21} &= (p(\cdot), g_{u_1}(\cdot, y, u_1)\delta u_1)_{2, D_3} + \sum_{i=1}^{n_x+1} (p(0, \cdot), \delta u_i^0(\cdot))_{2, I_i^0(\bar{w})} \\
1490 \quad &- \int_{f'(\bar{u}_i^0(\bar{x}_i^0))}^{f'(\bar{u}_{i+1}^0(\bar{x}_i^0))} \lim_{t \searrow 0} p(t, zt + \bar{x}_i^0) \frac{\delta x_i^0}{f''(f'^{-1}(z))} dz \\
1491 \quad &+ \sum_{i \in I_{s,0}(\bar{w})} p(0, \bar{x}_i^0) [\bar{u}_0(\bar{x}_i^0)] \delta x_i^0 + \sum_{i=1}^{n_{t,a}+1} (p(\cdot, 0), f'(\bar{u}_i^{B,a}(\cdot)) \delta u_i^{B,a}(\cdot))_{2, I_{B,a}^i(\bar{w}) \cap [0, \bar{t}]} \\
1492 \quad (177) \quad &+ \sum_{i \in I_{s,a}(\bar{w}), s \leq \bar{t}_i^a < \bar{t}} p(\bar{t}_i^a, 0) [f(\bar{y}(\bar{t}_i^a, 0+))] \delta t_i^a + o(\|\delta w\|_W). \\
1493
\end{aligned}$$

1494 As already mentioned at the end of step 2,  $p_2$  and the adjoint state  $p$  coincide on  $D_{2,0}$   
1495 such that we can replace  $p_2$  by  $p$  in (147) and (155) yielding

$$\begin{aligned}
1496 \quad (178) \quad &I_{20} = (p(\cdot), g_{u_1}(\cdot, \bar{y}, \bar{u}_1)\delta u_1)_{2, D_{2,0}} + o(\|\delta w\|_W) \\
1497 \quad (179) \quad &\text{and } I_{24} = (p(\cdot, 0), f'(\bar{y}(\cdot, 0+)) \delta u_m^{B,a}(\cdot))_{2, [\bar{t}, \bar{t}_m^a]} + o(\|\delta w\|_W). \\
1498
\end{aligned}$$

1499 Inserting (148), (149), (177), (178) and (179) in (146), we obtain

$$\begin{aligned}
1500 \quad I_{2,\varepsilon} &= (p(\cdot), g_{u_1}(\cdot, \bar{y}, \bar{u}_1)\delta u_1)_{2, D_{2,0} \cup D_3} \\
1501 \quad &+ \sum_{i=1}^{n_{t,a}+1} (p(\cdot, 0), f'(\bar{u}_i^{B,a}(\cdot)) \delta u_i^{B,a}(\cdot))_{2, I_{B,a}^i(\bar{w}) \cap [0, \bar{t}_m^a]} \\
1502 \quad &+ \sum_{i \in I_{s,a}(\bar{w}), s \leq \bar{t}_i^a < \bar{t}} p(\bar{t}_i^a, 0) [f(\bar{y}(\bar{t}_i^a, 0+))] \delta t_i^a + \sum_{i=0}^{n_x+1} (p(0, \cdot), \delta u_i^0(\cdot))_{2, I_i^0(\bar{w})} \cdot \\
1503 \quad &- \int_{f'(\bar{u}_i^0(\bar{x}_i^0))}^{f'(\bar{u}_{i+1}^0(\bar{x}_i^0))} \lim_{t \searrow 0} p(t, zt + \bar{x}_i^0) \frac{\delta x_i^0}{f''(f'^{-1}(z))} dz \\
1504 \quad (180) \quad &+ \sum_{i \in I_{s,0}(\bar{w})} p(0, \bar{x}_i^0) [\bar{u}_0(\bar{x}_i^0)] \delta x_i^0 + o(\|\delta w\|_W). \\
1505
\end{aligned}$$

1506 Since  $p(\cdot, 0)|_{(\bar{t}_m^a, \bar{t}]} \equiv 0$ , choosing  $s > 0$  sufficiently small, we obtain that (180) can be  
1507 simplified to

$$\begin{aligned}
1508 \quad I_{2,\varepsilon} &= (p(\cdot), g_{u_1}(\cdot, \bar{y}, \bar{u}_1)\delta u_1)_{2, D_{2,0} \cup D_3} \\
1509 \quad &+ \sum_{i=1}^{n_{t,a}+1} (p(\cdot, 0), f'(\bar{u}_i^{B,a}(\cdot)) \delta u_i^{B,a}(\cdot))_{2, I_{B,a}^i(\bar{w}) \cap [0, \bar{t}]} \\
1510 \quad &+ \sum_{i \in I_{s,a}(\bar{w}), \bar{t}_i^a \leq \bar{t}} p(\bar{t}_i^a, 0) [f(\bar{y}(\bar{t}_i^a, 0+))] \delta t_i^a + \sum_{i=0}^{n_x+1} (p(0, \cdot), \delta u_i^0(\cdot))_{2, I_i^0(\bar{w})} \cdot \\
1511 \quad &- \int_{f'(\bar{u}_i^0(\bar{x}_i^0))}^{f'(\bar{u}_{i+1}^0(\bar{x}_i^0))} \lim_{t \searrow 0} p(t, zt + \bar{x}_i^0) \frac{\delta x_i^0}{f''(f'^{-1}(z))} dz \\
1512 \quad (181) \quad &+ \sum_{i \in I_{s,0}(\bar{w})} p(0, \bar{x}_i^0) [\bar{u}_0(\bar{x}_i^0)] \delta x_i^0 + o(\|\delta w\|_W). \\
1513
\end{aligned}$$

1514 Inserting (181) and (141) in (136) and further (136) in (135) yields

$$\begin{aligned}
1515 & \frac{1}{[\bar{y}(\bar{t}, x_s(\bar{w}))]} \int_{x_s(\bar{w})-\bar{\varepsilon}}^{x_s(\bar{w})+\bar{\varepsilon}} \Delta y(\bar{t}, x) dx = \int_{D_{2,0} \cup D_3 \cup D_1} p(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 dx dt \\
1516 & + \int_{f'(\bar{u}_{m+1}^{B,a}(\bar{t}_m^a))}^{f'(\bar{u}_m^{B,a}(\bar{t}_m^a))} \lim_{t \searrow \bar{t}_m^a} p(t, z(t - \bar{t}_m^a)) \frac{\delta t_m^a \cdot z}{f''(f'^{-1}(z))} dz \\
1517 & + \sum_{i=1}^{n_{t,a}+1} (p(\cdot, 0), f'(u_i^{B,a}(\cdot))) \delta u_i^{B,a}(\cdot)_{2, I_{B,a}^i(\bar{w}) \cap [0, \bar{t}]} + \sum_{i \in I_{s,a}, \bar{t}_i^a \leq \bar{t}} p(\bar{t}_i^a, 0) [f(\bar{y}(t_i, 0+))] \delta t_i^a \\
1518 & + \sum_{i=0}^{n_x+1} (p(0, \cdot), \delta u_i^0(\cdot))_{2, I_0^i} + \sum_{i \in I_{s,0}} p(0, \bar{x}_i^0) [\bar{u}_0(\bar{x}_i^0)] \delta x_i^0 \\
1519 & - \int_{f'(\bar{u}_l^0(\bar{x}_l^0))}^{f'(\bar{u}_{l+1}^0(\bar{x}_l^0))} \lim_{t \searrow 0} p(t, zt + \bar{x}_l^0) \frac{\delta x_l^0}{f''(f'^{-1}(z))} dz + o(\|\delta w\|_W) + \bar{\varepsilon} O(\|\delta w\|_W). \\
1520 &
\end{aligned}$$

1521 Here, we have used that  $p_1$  coincides with the adjoint state  $p$  on  $D_1$ . Setting  $\bar{\varepsilon} := \bar{\varepsilon}(\delta w)$   
1522 such that  $\bar{\varepsilon}(\delta w) \rightarrow 0$  if  $\|\delta w\|_W \rightarrow 0$  and using (105), we can finally deduce Fréchet-  
1523 differentiability of the mapping (12) with derivative

$$\begin{aligned}
1524 & \frac{\partial}{\partial w} x_s(w) \cdot \delta w = \int_{D_{2,0} \cup D_3 \cup D_1} p(t, x) g_{u_1}(t, x, \bar{y}, \bar{u}_1) \delta u_1 dx dt \\
1525 & + \int_{f'(\bar{u}_{m+1}^{B,a}(\bar{t}_m^a))}^{f'(\bar{u}_m^{B,a}(\bar{t}_m^a))} \lim_{t \searrow \bar{t}_m^a} p(t, z(t - \bar{t}_m^a)) \frac{\delta t_m^a \cdot z}{f''(f'^{-1}(z))} dz \\
1526 & + \sum_{i=1}^{n_{t,a}+1} (p(\cdot, 0), f'(u_i^{B,a}(\cdot))) \delta u_i^{B,a}(\cdot)_{2, I_{B,a}^i(\bar{w}) \cap [0, \bar{t}]} \\
1527 & + \sum_{i \in I_{s,a}, \bar{t}_i^a \leq \bar{t}} p(\bar{t}_i^a, 0) [f(\bar{y}(t_i, 0+))] \delta t_i^a \\
1528 & + \sum_{i=0}^{n_x+1} (p(0, \cdot), \delta u_i^0(\cdot))_{2, I_0^i} + \sum_{i \in I_{s,0}} p(0, \bar{x}_i^0) [\bar{u}_0(\bar{x}_i^0)] \delta x_i^0 \\
1529 & (182) \quad - \int_{f'(\bar{u}_l^0(\bar{x}_l^0))}^{f'(\bar{u}_{l+1}^0(\bar{x}_l^0))} \lim_{t \searrow 0} p(t, zt + \bar{x}_l^0) \frac{\delta x_l^0}{f''(f'^{-1}(z))} dz. \\
1530 &
\end{aligned}$$

1531 Since it holds  $\text{supp}_x(p) \subset D_{2,0} \cup D_3 \cup D_1$ , we can finally conclude that the representation  
1532 in (182) coincides with the representation in Theorem 16. In the case of  $\bar{n}$  rarefaction  
1533 center  $\{i_1, \dots, i_{\bar{n}}\}$ , the term  $(p(s, \cdot), \Delta y(s, \cdot))_{2, \Omega}$  in (172) is splitted as follows:

$$\begin{aligned}
1534 & (p(s, \cdot), \Delta y(s, \cdot))_{2, \Omega} = (p(s, \cdot), \Delta y(s, \cdot))_{2, (-\infty, \bar{x}_{i_1}^0 - \varepsilon + f'(\bar{u}_{i_1}^0(\bar{x}_{i_1}^0 - \varepsilon)) \cdot s]} \\
1535 & + \sum_{j=2}^{\bar{n}-1} (p(s, \cdot), \Delta y(s, \cdot))_{2, (\bar{x}_{i_j}^0 - \varepsilon + f'(\bar{u}_{i_j}^0(\bar{x}_{i_j}^0 - \varepsilon)) \cdot s, \bar{x}_{i_j}^0 + \varepsilon + f'(\bar{u}_{i_{j+1}}^0(\bar{x}_{i_j}^0 + \varepsilon)) \cdot s]} \\
1536 & + (p(s, \cdot), \Delta y(s, \cdot))_{2, (\bar{x}_{i_{\bar{n}}}^0 + \varepsilon + f'(\bar{u}_{i_{\bar{n}+1}}^0(\bar{x}_{i_{\bar{n}}}^0 + \varepsilon)) \cdot s, \infty)} = \sum_{k=1}^{\bar{n}+1} I_{k, \varepsilon}^0 \\
1537 &
\end{aligned}$$

1538 Then we can treat all terms  $I_{k, \varepsilon}^0$  analogously to  $I_{1, \varepsilon}^0$ ,  $I_{2, \varepsilon}^0$  and  $I_{3, \varepsilon}^0$ , respectively and  
1539 obtain the desired result. In the case that also the right boundary is within in shock

1540 funnel and there are several rarefaction centers, we introduce time points  $\tilde{t}_i > \bar{t}_i^a$  for  
 1541  $i \in I_{r,a}(\bar{w})$  and  $\tilde{t}_i > \bar{t}_i^b$  for  $i \in I_{r,b}(\bar{w})$  satisfying (119), respectively. The treatment of  
 1542 the case that  $|\mathbb{T}^a|, |\mathbb{T}^b| > 1$  is basically the same as it was shown in Step 3.  $\square$

1543

## REFERENCES

- 1544 [1] C. BARDOS, A. Y. LE ROUX, AND J.-C. NÉDÉLEC, *First order quasilinear equations with*  
 1545 *boundary conditions*, Comm. Partial Differential Equations, 4 (1979), pp. 1017–1034,  
 1546 doi:10.1080/03605307908820117.
- 1547 [2] C. BARDOS AND O. PIRONNEAU, *A formalism for the differentiation of con-*  
 1548 *servation laws*, Comptes Rendus Mathématique, 335 (2002), pp. 839 – 845,  
 1549 doi:https://doi.org/10.1016/S1631-073X(02)02574-8.
- 1550 [3] F. BOUCHUT AND F. JAMES, *One-dimensional transport equations with discontinuous coeffi-*  
 1551 *cients*, Nonlinear Anal., 32 (1998), pp. 891–933, doi:10.1016/S0362-546X(97)00536-1.
- 1552 [4] A. BRESSAN, *One dimensional hyperbolic systems of conservation laws*, in Current developments  
 1553 in mathematics, 2002, Int. Press, Somerville, MA, 2003, pp. 1–37.
- 1554 [5] A. BRESSAN AND A. MARSON, *A maximum principle for optimally controlled systems of*  
 1555 *conservation laws*, Rend. Sem. Mat. Univ. Padova, 94 (1995), pp. 79–94, [http://www.](http://www.numdam.org/item?id=RSMUP_1995__94__79_0)  
 1556 [numdam.org/item?id=RSMUP\\_1995\\_\\_94\\_\\_79\\_0](http://www.numdam.org/item?id=RSMUP_1995__94__79_0).
- 1557 [6] A. BRESSAN AND W. SHEN, *Optimality conditions for solutions to hyperbolic balance laws*, in  
 1558 Control methods in PDE-dynamical systems, vol. 426 of Contemp. Math., Amer. Math.  
 1559 Soc., Providence, RI, 2007, pp. 129–152, doi:10.1090/conm/426/08187.
- 1560 [7] E. CASAS, *Control of an elliptic problem with pointwise state constraints*, SIAM J. Control  
 1561 Optim., 24 (1986), pp. 1309–1318, doi:10.1137/0324078.
- 1562 [8] C. CASTRO, F. PALACIOS, AND E. ZUAZUA, *An alternating descent method for the optimal*  
 1563 *control of the inviscid Burgers equation in the presence of shocks*, Math. Models Methods  
 1564 Appl. Sci., 18 (2008), pp. 369–416, doi:10.1142/S0218202508002723.
- 1565 [9] G. M. COCLITE, K. H. KARLSEN, AND Y.-S. KWON, *Initial-boundary value problems for*  
 1566 *conservation laws with source terms and the Degasperis-Procesi equation*, J. Funct. Anal.,  
 1567 257 (2009), pp. 3823–3857, doi:10.1016/j.jfa.2009.09.022.
- 1568 [10] C. M. DAFERMOS, *Generalized characteristics and the structure of solutions of hyperbolic*  
 1569 *conservation laws*, Indiana Univ. Math. J., 26 (1977), pp. 1097–1119.
- 1570 [11] J. C. DE LOS REYES AND I. YOUSEPT, *Regularized state-constrained boundary optimal*  
 1571 *control of the Navier-Stokes equations*, J. Math. Anal. Appl., 356 (2009), pp. 257–279,  
 1572 doi:10.1016/j.jmaa.2009.02.042.
- 1573 [12] J. ELSTRODT, *Maß- und Integrationstheorie*, Springer-Lehrbuch., Springer, Berlin, fourth ed.,  
 1574 2005. Grundlehren Mathematik.
- 1575 [13] M. GILES AND S. ULBRICH, *Convergence of linearized and adjoint approximations for discon-*  
 1576 *tinuous solutions of conservation laws. Part 1: Linearized approximations and linearized*  
 1577 *output functionals*, SIAM J. Numer. Anal., 48 (2010), pp. 882–904, doi:10.1137/080727464.
- 1578 [14] M. GILES AND S. ULBRICH, *Convergence of linearized and adjoint approximations for discon-*  
 1579 *tinuous solutions of conservation laws. Part 2: Adjoint approximations and extensions*,  
 1580 SIAM J. Numer. Anal., 48 (2010), pp. 905–921, doi:10.1137/09078078X.
- 1581 [15] L. GOSSE AND F. JAMES, *Numerical approximations of one-dimensional linear conserva-*  
 1582 *tion equations with discontinuous coefficients*, Math. Comp., 69 (2000), pp. 987–1015,  
 1583 doi:10.1090/S0025-5718-00-01185-6.
- 1584 [16] S. HAJIAN, M. HINTERMÜLLER, AND S. ULBRICH, *Total variation diminishing schemes in*  
 1585 *optimal control of scalar conservation laws*, IMA J. Numer. Anal., 39 (2019), pp. 105–140,  
 1586 doi:10.1093/imanum/drx073.
- 1587 [17] M. HERTY, A. KURGANOV, AND D. KUROCHKIN, *Numerical method for optimal control*  
 1588 *problems governed by nonlinear hyperbolic systems of PDEs*, Commun. Math. Sci., 13  
 1589 (2015), pp. 15–48, doi:10.4310/CMS.2015.v13.n1.a2.
- 1590 [18] M. HERTY, A. KURGANOV, AND D. KUROCHKIN, *On convergence of numerical methods for*  
 1591 *optimization problems governed by scalar hyperbolic conservation laws*, in Theory, numerics  
 1592 and applications of hyperbolic problems. I, vol. 236 of Springer Proc. Math. Stat., Springer,  
 1593 Cham, 2018, pp. 691–706, doi:10.1007/978-3-319-91545-6\_5.
- 1594 [19] M. HINTERMÜLLER AND K. KUNISCH, *Path-following methods for a class of constrained*  
 1595 *minimization problems in function space*, SIAM J. Optim., 17 (2006), pp. 159–187,  
 1596 doi:10.1137/040611598.
- 1597 [20] M. HINZE, R. PINNAU, M. ULBRICH, AND S. ULBRICH, *Optimization with PDE constraints*,  
 1598 vol. 23 of Mathematical Modelling: Theory and Applications, Springer, New York, 2009.

- 1599 [21] K. ITO AND K. KUNISCH, *Semi-smooth Newton methods for state-constrained optimal control*  
 1600 *problems*, Systems Control Lett., 50 (2003), pp. 221–228, doi:10.1016/S0167-6911(03)00156-  
 1601 7.
- 1602 [22] K. KRUMBIEGEL, I. NEITZEL, AND A. RÖSCH, *Sufficient optimality conditions for the Moreau-*  
 1603 *Yosida-type regularization concept applied to semilinear elliptic optimal control problems*  
 1604 *with pointwise state constraints*, Ann. Acad. Rom. Sci. Ser. Math. Appl., 2 (2010), pp. 222–  
 1605 246.
- 1606 [23] K. KRUMBIEGEL AND A. RÖSCH, *A virtual control concept for state constrained optimal control*  
 1607 *problems*, Comput. Optim. Appl., 43 (2009), pp. 213–233, doi:10.1007/s10589-007-9130-0.
- 1608 [24] S. N. KRUŽKOV, *First order quasilinear equations in several independent variables*, Math. USSR  
 1609 Sb., 10 (1970), pp. 217–243.
- 1610 [25] R. LECAROS AND E. ZUAZUA, *Tracking control of 1D scalar conservation laws in the presence*  
 1611 *of shocks*, in Trends in contemporary mathematics, vol. 8 of Springer INdAM Ser., Springer,  
 1612 Cham, 2014, pp. 195–219.
- 1613 [26] R. LECAROS AND E. ZUAZUA, *Control of 2D scalar conservation laws in the presence of shocks*,  
 1614 Math. Comp., 85 (2016), pp. 1183–1224, <https://doi.org/10.1090/mcom/3015>.
- 1615 [27] M. J. LIGHTHILL AND G. B. WHITHAM, *On kinematic waves. II. A theory of traffic*  
 1616 *flow on long crowded roads*, Proc. Roy. Soc. London. Ser. A., 229 (1955), pp. 317–345,  
 1617 doi:10.1098/rspa.1955.0089.
- 1618 [28] J. MÁLEK, J. NEČAS, M. ROKYTA, AND M. RUŽIČKA, *Weak and measure-valued solutions*  
 1619 *to evolutionary PDEs*, vol. 13 of Applied Mathematics and Mathematical Computation,  
 1620 Chapman & Hall, London, 1996, doi:10.1007/978-1-4899-6824-1.
- 1621 [29] S. MARTIN, *First order quasilinear equations with boundary conditions in the  $L^\infty$  framework*,  
 1622 J. Differential Equations, 236 (2007), pp. 375–406, doi:10.1016/j.jde.2007.02.007.
- 1623 [30] C. MEYER AND I. YOUSEPT, *Regularization of state-constrained elliptic optimal control problems*  
 1624 *with nonlocal radiation interface conditions*, Comput. Optim. Appl., 44 (2009), pp. 183–212,  
 1625 doi:10.1007/s10589-007-9151-8.
- 1626 [31] I. NEITZEL AND F. TRÖLTZSCH, *On convergence of regularization methods for nonlinear*  
 1627 *parabolic optimal control problems with control and state constraints*, Control Cybernet.,  
 1628 37 (2008), pp. 1013–1043.
- 1629 [32] F. OTTO, *Initial-boundary value problem for a scalar conservation law*, C. R. Acad. Sci. Paris  
 1630 Sér. I Math., 322 (1996), pp. 729–734.
- 1631 [33] S. PFAFF, *Optimal control of a hyperbolic conservation law on a bounded domain with switch-*  
 1632 *ing controls*, Dissertation, Fachbereich Mathematik, Technische Universität Darmstadt,  
 1633 Germany, 2015.
- 1634 [34] S. PFAFF AND S. ULBRICH, *Optimal boundary control of nonlinear hyperbolic conservation*  
 1635 *laws with switched boundary data*, SIAM J. Control Optim., 53 (2015), pp. 1250–1277,  
 1636 doi:10.1137/140995799.
- 1637 [35] S. PFAFF AND S. ULBRICH, *Optimal control of scalar conservation laws by on/off-switching*,  
 1638 Optim. Methods Softw., 32 (2017), pp. 904–939.
- 1639 [36] P. I. RICHARDS, *Shock waves on the highway*, Operations Res., 4 (1956), pp. 42–51.
- 1640 [37] P. SCHÄFER AGUILAR, J. M. SCHMITT, S. ULBRICH, AND M. MOOS, *On the numerical*  
 1641 *discretization of optimal control problems for conservation laws*, Control and Cybernetics,  
 1642 48 (2019).
- 1643 [38] S. ULBRICH, *Optimal Control of Nonlinear Hyperbolic Conservation Laws with Source Terms*,  
 1644 Department of Mathematics, TU München, 2002. Habilitation.
- 1645 [39] S. ULBRICH, *A sensitivity and adjoint calculus for discontinuous solutions of hyperbolic*  
 1646 *conservation laws with source terms*, SIAM J. Control Optim., 41 (2002), pp. 740–797,  
 1647 doi:10.1137/S0363012900370764.
- 1648 [40] S. ULBRICH, *Adjoint-based derivative computations for the optimal control of discontinuous*  
 1649 *solutions of hyperbolic conservation laws*, Systems Control Lett., 48 (2003), pp. 313–328,  
 1650 doi:10.1016/S0167-6911(02)00275-X.
- 1651 [41] J. VOVELLE, *Convergence of finite volume monotone schemes for scalar conservation laws on*  
 1652 *bounded domains*, Numer. Math., 90 (2002), pp. 563–596, doi:10.1007/s002110100307.
- 1653 [42] J. ZOWE AND S. KURCZYUSZ, *Regularity and stability for the mathematical programming problem*  
 1654 *in Banach spaces*, Appl. Math. Optim., 5 (1979), pp. 49–62, doi:10.1007/BF01442543.