PRIMAL-DUAL INTERIOR POINT METHODS FOR PDE-CONSTRAINED OPTIMIZATION

MICHAEL ULBRICH* AND STEFAN ULBRICH†

Abstract. This paper provides a detailed analysis of a primal-dual interior-point method for PDE-constrained optimization. Considered are optimal control problems with control constraints in L^p. It is shown that the developed primal-dual interior-point method converges globally and locally superlinearly. Not only the easier L^∞-setting is analyzed, but also a more involved L^q-analysis, $q < \infty$, is presented. In L^∞, the set of feasible controls contains interior points and the Fréchet differentiability of the perturbed optimality system can be shown. In the L^q-setting, which is highly relevant for PDE-constrained optimization, these nice properties are no longer available. Nevertheless, a convergence analysis is developed using refined techniques. In particular, two-norm techniques and a smoothing step are required.

1. Introduction. This paper is concerned with the analysis of primal-dual interior point methods for optimization problems with PDE- and pointwise inequality constraints. We assume that the problem has optimal control structure and that the inequality constraints are posed on the controls only. In contrast to state constraints, this situation allows for a rigorous analysis. Related investigations of other Newton-based algorithms were conducted in, e.g., [3, 7, 10, 11, 13, 17, 15, 16, 18, 19, 20] for comparable problem settings. For primal-dual interior point methods, although intensively investigated in finite dimensional mathematical programming, see, e.g., [6] and the references therein, only little rigorous theory is available in the function space framework of optimal control problems. Earlier investigations of modern optimization methods in function space have resulted in valuable deep understanding of algorithms for PDE constrained optimization. In particular, in all analyses, a certain problem structure is required for a successful local convergence analysis. A common theme is that an L^p-setting for the inequalities is required and that a smoothing property or smoothing step must be available. Furthermore, the usual backtracking in interior point methods to keep iterates strictly positive has to be augmented by suitable projection techniques, at least if the primal-dual Newton step for the control is not in L^∞. Finally, integrated barriers are the appropriate choice, which result in a weighting of the pointwise barriers after discretization. All of these crucial ingredients are not visible in the finite dimensional analysis. A further important benefit of an abstract analysis in function space is that it is the prerequisite for proving mesh independence results, see, e.g., [1, 2, 8].

The purpose of this paper is to give a rigorous analysis of the global and fast local convergence of a primal-dual interior point method for PDE-constrained optimization. The analysis covers not only the (easier) L^∞ setting but also the quite involved but in practice highly relevant L^q-setting, $q < \infty$. The crucial point is that for the analysis in the L^∞-setting one needs that the corresponding adjoint state (i.e., the Lagrange multiplier for the state equation) is also in L^∞, which is usually not the case for complex systems like, e.g., the Navier-Stokes equations [5, 9, 14]. One of the difficulties in the L^q-setting, $q < \infty$, is that the set of feasible controls does not contain interior points with respect to the L^q-topology. As a consequence, the barrier function is not Fréchet-differentiable in L^q. This requires elaborate techniques, including a suitable scaling of the primal-dual Newton system, a two-norm approach, and a smoothing step.

The paper covers both, global and superlinear local convergence. It is organized as follows: In section 2 the considered problem class is described and it is illustrated that elliptic optimal control problems fit into this class. Then, first order optimality (KKT) conditions are derived. As a first step towards interior point methods, a barrier problem is formulated, its unique solvability is proved, and optimality conditions are stated that result in perturbed KKT conditions that form the basis for

*TU München, Chair of Mathematical Optimization, Zentrum Mathematik, M1: Optimization, Boltzmannstr. 3, D-85747 Garching b. München, Germany (mulbrich@ma.tum.de).

†TU Darmstadt, Fachbereich Mathematik, AG10: Nonlinear Optimization and Optimal Control, Schlossgartenstr. 7, D-64289 Darmstadt, Germany (ulbrich@mathematik.tu-darmstadt.de).

the primal dual Newton step. Section 3 presents and illustrates a functional analytic setting that is used in the rest of the paper. In section 4, properties of the central path are derived, in particular the boundedness of the dual variables in L^q and the boundedness of the central path. Section 5 is devoted to the analysis of the primal-dual Newton system on a suitable neighborhood of the central path. A key result is the uniformly bounded invertibility of the suitably scaled linear operator in the primal dual Newton system on bounded subsets of the neighborhood. As a simple consequence, the norm of the inverse of the unscaled operator is uniformly bounded by $O(1/\sqrt{\mu})$ on bounded subsets of the neighborhood. The Hölder continuity of the central path is proved in section 6. The conceptual primal-dual interior-point method is formulated in section 7. It includes a projection onto the neighborhood of the central path that replaces the usual backtracking. In section 8, the method is analyzed in the L^∞-setting. Quadratic local convergence towards the central path and global linear convergence are proved. Finally, in section 9, the more involved analysis of the method in L^q, $q < \infty$, is carried out. As for other approaches, an inevitable norm gap occurs that has to be closed by a smoothing step. Such a smoothing step is derived and incorporated in the algorithm. For the resulting method, global linear and local superlinear convergence is proved.

Notations. We denote the L^p-norm by $\|\cdot\|_p$, $1 \le p \le \infty$. For Banach spaces X, Y we denote by $\mathcal{L}(X,Y)$ the space of bounded linear operators from X to Y equipped with the operator norm $\|\cdot\|_{X,Y}$. X^* is the dual space of a Banach space X and $\langle\cdot,\cdot\rangle_{X^*,X}$ is the corresponding dual pairing. By $\mathrm{leb}\,(\cdot)$ we denote the Lebesgue measure on $\mathbb{R}^n$. Throughout the paper equalities and inequalities between L^p-functions are meant almost everywhere. If $X \subset Y$ is a continuous embedding, we write $I_{X,Y}$, $I_{X,Y}x = x$ for the embedding operator. Sometimes, if no confusion is to be expected, we save space by writing I instead of $I_{X,Y}$.

2. Control constrained optimal control problem. Let $\Omega \subset \mathbb{R}^n$ be a bounded open domain with sufficiently smooth boundary. We consider the optimal control problem with control constraints

$$\min_{y\in Y, u\in U} J(y,u) \quad \text{s.t.} \quad c(y,u) = 0, \quad a \le u \le b, \tag{2.1}$$

where $U = L^p(\Omega)$, $p \in [2,\infty)$, $a, b \in L^\infty$, $b - a \ge \nu > 0$, and Y is a Banach space. We set

$$\mathcal{B} := \{u \in U \ : \ a \le u \le b\}, \quad x = (y,u), \quad X := Y \times U,$$

and assume that there exists an open set $\mathcal{D} \subset L^p(\Omega)$, $\mathcal{D} \supset \mathcal{B}$ such that

(A1) $J : Y \times U \to \mathbb{R}$, $c : Y \times U \to \Lambda$ are twice locally Lipschitz-continuously differentiable and there exist uniform Lipschitz constants on bounded subsets of $Y \times \mathcal{B}$.

(A2) $c_y(y,u) \in \mathcal{L}(Y,\Lambda)$ has a bounded inverse for all $(y,u) \in Y \times \mathcal{D}$ and $\|c_y(y,u)^{-1}\|_{\Lambda,Y}$ is uniformly bounded on bounded subsets of $Y \times \mathcal{B}$.

(A3) For all $u \in \mathcal{D}$ there exists a unique solution $y = y(u) \in Y$ of

$$c(y(u),u) = 0$$

and there exists $M_y > 0$ with

$$\|y(u)\|_Y \le M_y \quad \forall\, u \in \mathcal{B}.$$

(A4) The reduced objective functional

$$u \in (\mathcal{B}, \|\cdot\|_\infty) \mapsto J(y(u),u) =: \hat{J}(u)$$

is lower semicontinuous w.r.t. sequential L^∞-weak* convergence.

REMARK 2.1. By the implicit function theorem (A1)–(A3) ensure that $u \in \mathcal{D} \mapsto y(u) \in Y$ and $u \in \mathcal{D} \mapsto J(y(u),u)$ are twice locally Lipschitz-continuously differentiable and in addition Lipschitz continuous on $\mathcal{B}$.

For convenience we identify $U^* = L^p(\Omega)^*$ with $L^{p'}(\Omega)$, $\frac{1}{p} + \frac{1}{p'} = 1$, via the dual pairing

$$\langle v, u, \rangle_{U^*,U} = \langle v, u \rangle := \int_\Omega vu \, d\xi.$$

We recall that in function spaces of distributions it is common practice to extend $\langle \cdot, \cdot \rangle$ to the distributional dual pairing. In our examples, we typically work with the Sobolev spaces $H_0^1(\Omega)$ and $\tilde{Y} = H_0^1(\Omega) \cap H^2(\Omega)$. The dual spaces with respect to the dual pairing $\langle \cdot, \cdot \rangle$ then result in the following continuous and dense embeddings:

$$\tilde{Y} \subset H_0^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega) = H_0^1(\Omega)^* \subset \tilde{Y}^*.$$

In this sense, we can (and will) interpret H_0^1-functions as (nice) L^2-functions and L^2-functions as (nice) H^{-1}-functions (the latter are generalized functions). Furthermore, we will omit operators of the form $I_{Y_1,Y_2} : y \in Y_1 \mapsto y \in Y_2$ if $Y_1 \subset Y_2$ is continuously embedded. In this case, we also have $I_{Y_1,Y_2}^* : y' \in Y_2^* \mapsto y' \in Y_1^*$, thus $I_{Y_1,Y_2}^* = I_{Y_2^*,Y_1^*}$, i.e., the adjoint acts like the identity.

PROPOSITION 2.2. *Under assumptions* (A1)–(A4) *problem (2.1) has a solution.*

Proof. Take a minimizing sequence $(y(u_k), u_k)$. Since $(u_k) \subset \mathcal{B}$, it is bounded in L^∞ and has a weak*-convergent subsequence, which we denote again by (u_k) for simplicity, with limit $\bar{u} \in \mathcal{B}$. But by (A4) we have

$$\limsup_{k \to \infty} J(y(u_k), u_k) \geq J(y(\bar{u}), \bar{u})$$

and thus $(y(\bar{u}), \bar{u})$ solves (2.1), since $(y(u_k), u_k)$ is a minimizing sequence. □

2.1. An Example. As a standard example we consider the following elliptic control problem

$$\begin{aligned} \min_{y,u} \quad & \frac{1}{2}\|y - y_d\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 \\ \text{s.t.} \quad & -\Delta y = u \quad \text{in } \Omega, \\ & y = 0 \quad \text{in } \partial\Omega, \\ & a \leq u \leq b \quad \text{in } \Omega. \end{aligned}$$

Here, $\Omega \subset \mathbb{R}^n$ and $a, b \in L^\infty(\Omega)$ are as specified above, $u \in U := L^2(\Omega)$ is the control, $y \in H^1(\Omega)$ is the state, $y_d \in L^2(\Omega)$ is the desired state, and $\alpha > 0$ is a regularization parameter.

There are at least two reasonable ways to choose the functional analytic setting. We choose $p = 2$, $\mathcal{D} = U$ and have $U^* = U = L^2$.

2.1.1. First setting. The first setting is to consider the usual weak solution of the state equation. Here, the state space is $Y = H_0^1(\Omega)$ and the PDE is considered in the weak form

$$\int_\Omega \nabla y(\xi)^T \nabla v(\xi) \, d\xi = \int_\Omega u(\xi) v(\xi) \, d\xi \quad \forall \, v \in H_0^1(\Omega).$$

This results in the abstract state equation

$$c(y, u) := Ay - u = 0 \quad \text{in } \Lambda$$

with $\Lambda = Y^* = H^{-1}(\Omega)$. Note that, as mentioned earlier, we have omitted the embedding operator $I_{L^2,H^{-1}} \in \mathcal{L}(L^2(\Omega), \Lambda)$, $I_{L^2,H^{-1}} u = u$. The operator $A \in \mathcal{L}(Y, \Lambda)$ is defined by

$$\langle Ay, v \rangle_{H^{-1},H_0^1} = \langle Ay, v \rangle = \int_\Omega \nabla y(\xi)^T \nabla v(\xi) \, d\xi \qquad \forall \, y, v \in H_0^1(\Omega).$$

It is well known that $A \in \mathcal{L}(Y, \Lambda)$ is invertible. Clearly,

$$J(y,u) = \frac{1}{2}\|y - y_d\|_2^2 + \frac{\alpha}{2}\|u\|_2^2$$

is twice continuously differentiable with

$$\begin{aligned}
\langle J_y(y,u), v\rangle_{Y^*,Y} &= \langle y - y_d, v\rangle = (y - y_d, v)_2, \\
\langle J_u(y,u), w\rangle_{U,U^*} &= \alpha\langle u, w\rangle = \alpha(u,w)_2, \\
\langle J_{yy}(y,u)v_2, v_1\rangle_{Y^*,Y} &= (v_1, v_2)_2, \\
\langle J_{uu}(y,u)w_2, w_1\rangle_{U^*,U} &= \alpha(w_1, w_2)_2.
\end{aligned}$$

In short notation:

$$J_y(y,u) = y - y_d, \quad J_u(y,u) = \alpha u, \quad J_{yy}(y,u) = I_{Y^*,Y}, \quad J_{uu}(y,u) = \alpha I_U.$$

Furthermore, $c(y,u) = Ay - u$ is twice continuously differentiable with

$$c_y(y,u) = A, \quad c_u(y,u) = -I_{L^2,H^{-1}}, \quad c''(y,u) = 0.$$

The uniform Lipschitz constants on bounded sets for c, J' and their derivatives are clear due to bounded linearity. The uniform Lipschitz continuity of J on bounded sets follows from the boundedness of J' on bounded sets. Hence, (A1) is shown. (A2) follows since

$$c_y(y,u) = A$$

is constant and $A \in \mathcal{L}(Y, \Lambda)$ is invertible. (A3) follows from $y(u) = A^{-1}Bu$. Finally (A4) is satisfied since $u \in L^2 \mapsto J(A^{-1}u, u) \in \mathbb{R}$ is convex and continuous, hence sequentially weakly lower semicontinuous. As a consequence, $\hat{J}$ is also lower semicontinuous w.r.t. sequential L^∞-weak* convergence.

2.1.2. Second setting. From the assumptions on Ω and standard regularity results for elliptic equations it follows that the solution of the state equation enjoys more regularity, namely, $y \in \tilde{Y} := H_0^1 \cap H^2$ (we use a "˜" to distinguish from the first setting). Hence, we can write the state equation also as follows:

$$\tilde{c}(y,u) := \tilde{A}y - u = 0 \quad \text{in } \tilde{\Lambda} := L^2$$

with $\tilde{A} \in \mathcal{L}(\tilde{Y}, \tilde{\Lambda})$. Again $\tilde{A}$ is invertible in $\mathcal{L}(\tilde{Y}, \tilde{\Lambda})$.

Just as before, we can calculate derivatives and verify the assumptions (A1)–(A4).

2.2. Optimality conditions. If we define the closed convex set

$$\mathcal{C} := \{0\} \times \mathcal{B} \subset \Lambda \times U$$

and the constraint function

$$h(y,u) := \begin{pmatrix} c(y,u) \\ u \end{pmatrix}$$

then the constraint in (2.1) can be written as

$$h(y,u) \in \mathcal{C}.$$

Denote for $(\lambda, z) \in \Lambda^* \times U^*$ the Lagrangian function for the abstract problem

$$\min_{y \in Y, u \in U} J(y,u) \quad \text{s.t.} \quad h(y,u) \in \mathcal{C}$$

by

$$L(y,u,\lambda,z) = J(y,u) + \langle \lambda, c(y,u)\rangle_{\Lambda^*,\Lambda} + (z,u)_2.$$

Let $\bar{x} = (\bar{y},\bar{u}) \in Y \times \mathcal{B}$ be a local solution of (2.1). Since $c_y(\bar{x})$ is surjective by (A2), the operator

$$h'(\bar{x}) = \begin{pmatrix} c_y(\bar{x}) & c_u(\bar{x}) \\ 0 & I \end{pmatrix} \in \mathcal{L}(X, \Lambda \times U)$$

is surjective and therefore *Robinson's constraint qualification* [12]

$$0 \in \text{int}(h(\bar{x}) + h_x(\bar{x}) \cdot X - \mathcal{C})$$

is satisfied. By standard optimality theory, see [4, Prop. 3.2], there exist $(\bar{\lambda}, \bar{z}) \in \Lambda^* \times U^*$ with

$$L_x(\bar{x},\lambda,z) = (J_y + c_y^*\bar{\lambda}, J_u + c_u^*\bar{\lambda})(\bar{x}) + (0,\bar{z}) = 0, \quad h(\bar{x}) \in \mathcal{C}, \quad (\bar{\lambda},\bar{z}) \in N_{\mathcal{C}}(h(\bar{x})) \tag{2.2}$$

with the normal cone

$$\begin{aligned} N_{\mathcal{C}}(h(\bar{x})) &:= \{(\lambda,z) \in \Lambda^* \times U^* \;:\; \langle \lambda, w - c(\bar{x})\rangle_{\Lambda^*,\Lambda} + (z, v-\bar{u}) \le 0 \;\; \forall\, (w,v) \in \mathcal{C}\} \\ &= \{(\lambda,z) \in \Lambda^* \times U^* \;:\; (z, v - \bar{u}) \le 0 \;\; \forall\, v \in \mathcal{B}\} \\ &= \left\{(\lambda,z) \in \Lambda^* \times U^* \;:\; z|_{\{\bar{u}=a\}} \le 0, \quad z|_{\{\bar{u}=b\}} \ge 0, \; z|_{\{a<\bar{u}<b\}} = 0\right\}. \end{aligned}$$

Hence, using the splitting $\bar{z} = \bar{z}_b - \bar{z}_a$, $\bar{z}_b, \bar{z}_a \ge 0$, we can write (2.2) in the following form: there exist $\bar{\lambda} \in \Lambda^*$ and $\bar{z}_a, \bar{z}_b \in U^*$ such that with the Lagrangian

$$\ell(y,u,\lambda,z_a,z_b) = J(y,u) + \langle \lambda, c(y,u)\rangle_{\Lambda^*,\Lambda} - (z_a, u-a)_2 - (z_b, b-u)_2$$

the first order optimality conditions hold

$$\begin{cases} \ell_y(\bar{x},\bar{\lambda},\bar{z}_b,\bar{z}_a) = J_y(\bar{x}) + c_y(\bar{x})^*\bar{\lambda} = 0, \\ \ell_u(\bar{x},\bar{\lambda},\bar{z}_b,\bar{z}_a) = J_u(\bar{x}) + c_u(\bar{x})^*\bar{\lambda} + \bar{z}_b - \bar{z}_a = 0, \\ c(\bar{x}) = 0, \quad a \le \bar{u} \le b, \\ (\bar{u}-a)\,\bar{z}_a = 0, \quad \bar{z}_a \ge 0, \\ (b-\bar{u})\,\bar{z}_b = 0, \quad \bar{z}_b \ge 0. \end{cases} \tag{2.3}$$

2.3. Barrier problem. Now consider the associated barrier-problem

$$\begin{aligned} &\min\; J_\mu(y,u) := J(y,u) - \mu\int_\Omega \ln(u-a)\,dx - \mu\int_\Omega \ln(b-u)\,dx \\ &\text{s.t.} \quad c(y,u) = 0, \quad a \le u \le b. \end{aligned} \tag{2.4}$$

PROPOSITION 2.3. *Under assumptions* (A1)–(A4) *problem (2.4) has for any* $\mu > 0$ *a solution.*

Proof. Take a minimizing sequence $(y(u_k), u_k)$. Since $(u_k) \subset \mathcal{B}$, it is bounded in L^∞ and has a weak*-convergent subsequence, for simplicity again denoted by (u_k), with limit $\bar{u} \in \mathcal{B}$. But by (A4) we have

$$J(y(u_k),u_k) \ge J(y(\bar{u}),\bar{u}). \tag{2.5}$$

Moreover, the barrier term satisfies

$$-\mu\int_\Omega \ln(u_k - a)\,dx - \mu\int_\Omega \ln(b-u_k)\,dx \ge -2\mu\text{leb}\,(\Omega)\ln(\|b-a\|_\infty).$$

This shows that

$$J_\mu(y(u_k), u_k) \geq J(y(\bar{u}), \bar{u}) - 2\mu \text{leb}\,(\Omega) \ln(\|b-a\|_\infty) \geq -M_1$$

with a constant $M_1 > 0$.

Moreover, we have $u_n \to \bar{u}$ in L^∞-weak* and thus also in L^2-weak. It remains to show that

$$u \in (\mathcal{B}, \|\cdot\|_2) \mapsto -\mu \int_\Omega \ln(u-a)\, dx - \mu \int_\Omega \ln(b-u)\, dx =: f_a(u) + f_b(u)$$

is lower semicontinuous w.r.t. weak convergence.

We consider only $f_a : u \mapsto -\mu \int_\Omega \ln(u-a)\, dx$, since $f_b : u \mapsto -\mu \int_\Omega \ln(b-u)\, dx$ can be treated analogously. Since $\mathcal{B} \subset L^2(\Omega)$ is convex and $f_a : \mathcal{B} \mapsto \mathbb{R} \cup \{\infty\}$ is convex, it is sufficient to show that the mapping is lower semicontinuous w.r.t. strong convergence, see Jost, Lemma 4.2.2. To this end, let $\mathcal{B} \ni v_k \to \bar{v}$ in L^2 where without restriction $f_a(v_k) \leq C$ with a constant $C > 0$. We observe that

$$\begin{aligned} -\mu \ln(v_k - a) &= -\mu \ln(\max(v_k - a, 1)) - \mu \ln(\min(v_k - a, 1)) =: g_k + h_k, \\ -\mu \ln(\bar{v} - a) &= -\mu \ln(\max(\bar{v} - a, 1)) - \mu \ln(\min(\bar{v} - a, 1)) =: \bar{g} + \bar{h}. \end{aligned}$$

We have the estimate

$$|g_k - \bar{g}| \leq \mu |\max(v_k - a, 1) - \max(\bar{v} - a, 1)| \leq \mu |v_k - \bar{v}|$$

and thus

$$\liminf_{k\to\infty} f_k(v_k) = \liminf_{k\to\infty} \int_\Omega (g_k + h_k)\, dx = \int_\Omega \bar{g}\, dx + \liminf_{k\to\infty} \int_\Omega h_k\, dx.$$

Moreover, $h_k \geq 0$, $|g_k| \leq \mu |v_k - a - 1|$ and

$$0 \leq \int_\Omega h_k\, dx \leq C - \int_\Omega g_k\, dx \leq C + \mu \|v_k - a - 1\|_1 \leq C + C'.$$

For a subsequence (again denoted by (v_k)) we have $v_k \to \bar{v}$ a.e. and thus $h_k \to \bar{h}$ a.e. Now the Lemma of Fatou yields that $\bar{h} \in L^1(\Omega)$ and

$$0 \leq \int_\Omega \bar{h}\, dx = \int_\Omega \liminf_{k\to\infty} h_k\, dx \leq \liminf_{k\to\infty} \int_\Omega h_k\, dx \leq C + C'.$$

This concludes the proof that

$$\liminf_{k\to\infty} f_a(v_k) = \liminf_{k\to\infty} \int_\Omega (g_k + h_k)\, dx \geq \int_\Omega (\bar{g} + \bar{h})\, dx = f_a(\bar{v}).$$

As mentioned before, the same holds for $v_k \to \bar{v}$ weakly, see Jost, Lemma 4.2.2.

Applying this to the minimizing sequence, we obtain together with (2.5)

$$\begin{aligned} \liminf_{k\to\infty} J_\mu(y(u_k), u_k) &= \liminf_{k\to\infty} J(y(u_k), u_k) + f_a(u_k) + f_b(u_k) \\ &\geq J(y(\bar{u}), \bar{u}) + f_a(\bar{u}) + f_b(\bar{u}) = J_\mu(y(\bar{u}), \bar{u}). \end{aligned}$$

Hence, $(y(\bar{u}), \bar{u})$ solves (2.4). □

It is obvious that $\{\bar{u} = a\}$ and $\{\bar{u} = b\}$ have measure zero, since $J_\mu(\bar{y}, \bar{u}) < \infty$. Therefore, it is easy to derive the following necessary optimality conditions for (2.4).

LEMMA 2.4. *Let assumptions* (A1)–(A4) *hold and let* $(\bar{y}, \bar{u})$ *be a local solution of (2.4). Then there is* $\bar{\lambda} \in \Lambda^*$ *such that*

$$
\left\{
\begin{aligned}
& J_y(\bar{x}) + c_y(\bar{x})^* \bar{\lambda} = 0, \\
& J_u(\bar{x}) + c_u(\bar{x})^* \bar{\lambda} + \frac{\mu}{b - \bar{u}} - \frac{\mu}{\bar{u} - a} = 0, \\
& c(\bar{x}) = 0, \quad a < \bar{u} < b.
\end{aligned}
\right. \tag{2.6}
$$

REMARK 2.5. By introducing the artificial variables $\bar{z}_a = \frac{\mu}{\bar{u}-a}$ and $\bar{z}_b = \frac{\mu}{b-\bar{u}}$ we can write (2.6) as the perturbed KKT-conditions

$$
\left\{
\begin{aligned}
& \ell_y(\bar{x}, \bar{\lambda}, \bar{z}_b, \bar{z}_a) = J_y(\bar{x}) + c_y(\bar{x})^* \bar{\lambda} = 0, \\
& \ell_u(\bar{x}, \bar{\lambda}, \bar{z}_b, \bar{z}_a) = J_u(\bar{x}) + c_u(\bar{x})^* \bar{\lambda} + \bar{z}_b - \bar{z}_a = 0, \\
& c(\bar{x}) = 0, \quad a \le \bar{u} \le b, \\
& (\bar{u} - a)\, \bar{z}_a = \mu, \quad \bar{z}_a \ge 0, \\
& (b - \bar{u})\, \bar{z}_b = \mu, \quad \bar{z}_b \ge 0.
\end{aligned}
\right. \tag{2.7}
$$

We call the solution set (2.7) parameterized by $\mu > 0$ *central path*. We will see that under appropriate assumptions the central path is actually a Hölder-continuous curve that converges for $\mu \to 0$ to a solution of (2.1).□

Proof. By (A3) there exists for any $u \in \mathcal{D}$ a unique solution $y = y(u) \in Y$ of $c(y, u) = 0$. By (A2) and the implicit function theorem the mapping $u \in (\mathcal{D}, \|\cdot\|_U) \mapsto y(u) \in Y$ is continuously differentiable with

$$c_y y_u = -c_u.$$

Thus the reduced objective functional $u \in (\mathcal{D}, \|\cdot\|_U) \mapsto \hat{J}(u)$ is continuously differentiable with derivative

$$
\begin{aligned}
\hat{J}_u(u) &= -\langle J_y(x), c_y(x)^{-1} c_u(x) \;\cdot\; \rangle_{Y^*,Y} + J_u(x) \\
&= -c_u(x)^* (c_y(x)^{-1})^* J_y(x) + J_u(x) \\
&= -c_u(x)^* (c_y(x)^*)^{-1} J_y(x) + J_u(x),
\end{aligned}
$$

where $x = (y(u), u)$. Let $(y(\bar{u}), \bar{u})$ be the solution of (2.4). With the unique solution $\bar{\lambda} \in \Lambda^*$ of

$$J_y(\bar{x}) + c_y(\bar{x})^* \bar{\lambda} = 0$$

we have

$$\hat{J}_u(\bar{u}) = c_u(\bar{x})^* \bar{\lambda} + J_u(\bar{x}). \tag{2.8}$$

We show that

$$w := \hat{J}_u(\bar{u}) - \frac{\mu}{\bar{u} - a} + \frac{\mu}{b - \bar{u}} = 0 \quad \text{a.e.}$$

We know that $a < \bar{u} < b$ almost everywhere. The sets

$$M_k := \{a + 1/k \le \bar{u} \le b - 1/k\}$$

are monotone increasing with $\bigcup_{k=1}^{\infty} M_k = \Omega \setminus N$ with a set N of measure zero. Let $v \in L^\infty(\Omega)$ be arbitrary, then $v_k := v 1_{M_k} \to v$ in $U = L^p(\Omega)$, since $p < \infty$. For all $t \in (-\rho, \rho)$, $\rho > 0$ small enough, we have $a + 1/(2k) \le \bar{u} + t v_k \le b - 1/(2k)$ and therefore the function

$$
\begin{aligned}
h_k : t \in (-\rho, \rho) \mapsto & J_\mu(y(\bar{u} + t v_k), \bar{u} + t v_k) \\
& = \hat{J}(\bar{u} + t v_k) - \mu \int_\Omega \ln(b - (\bar{u} + t v_k))\, dx - \mu \int_\Omega \ln((\bar{u} + t v_k) - a)
\end{aligned}
$$

is continuously differentiable with

$$h_k'(t) = \left\langle \hat{J}_u(\bar{u} + tv_k) + \frac{\mu}{b - (\bar{u} + tv_k)} - \frac{\mu}{(\bar{u} + tv_k) - a}, v_k \right\rangle_2 .$$

Since $(y(\bar{u}), \bar{u})$ is optimal for (2.4) and $\bar{u} + tv_k \in \mathcal{B}$ for $t \in (-\rho, \rho)$, the function h_k has a minimum at $t = 0$ and thus

$$0 = h_k'(0) = \left\langle \hat{J}_u(\bar{u}) + \frac{\mu}{b - \bar{u}} - \frac{\mu}{\bar{u} - a}, v_k \right\rangle_2 .$$

Taking the limit $k \to \infty$ we obtain

$$\left\langle \hat{J}_u(\bar{u}) + \frac{\mu}{b - \bar{u}} - \frac{\mu}{\bar{u} - a}, v \right\rangle_2 = 0.$$

This holds for all $v \in L^\infty(\Omega)$ and by density for all $v \in U$. We deduce with (2.8) that

$$c_u(\bar{x})^* \bar{\lambda} + J_u(\bar{x}) + \frac{\mu}{b - \bar{u}} - \frac{\mu}{\bar{u} - a} = 0.$$

□

REMARK 2.6. For the special case of linear elliptic control problems, the previous results were shown in a different way in [11]. The control problem in [11] satisfies our assumption $(A5)_q$ below with $q = \infty$. In this particular case the solution of the barrier problem (2.4) lies in the interior of $\mathcal{B}$, see Corollary 4.4 below, and $\bar{z}_a, \bar{z}_b$ are bounded in L^∞. The analysis in [11] makes essential use of this fact.

In this paper we cover the much more general setting that $\bar{z}_a, \bar{z}_b$ are only bounded in L^q for some $q > p$. This is of essential interest to cover state equations, where the state or adjoint equation does not allow a priori estimates in L^∞. In the latter case solutions of (2.4) can touch the boundary of $\mathcal{B}$ on a zero set and are thus no interior points in the classical sense. Nevertheless, we will see that also in this setting interior point methods with a projection are convergent, since – roughly speaking – the measure of the set where the solution of (2.4) has distance $\leq \varepsilon$ to the boundary of $\mathcal{B}$ tends to zero as $\varepsilon \searrow 0$. The analysis is considerably more involved than for the case $(A5)_\infty$. □

3. A function space setting. Unfortunately, it is not possible to work with soft analysis only. Rather, we need a carefully adjusted function space setting, where a typical requirement will be that a continuous (or differentiable) mapping $h : X_1 \to Y_1$ also defines a mapping $h : X_2 \to Y_2$ from a stronger space $X_2 \subset X_1$ to a stronger space $Y_2 \subset Y_1$. For instance, as a trivial example, the identity mapping $X_1 \ni x \mapsto x \in X_1$ induces the identity mapping $X_2 \ni x \mapsto x \in X_2$ for any stronger space $X_2 \subset X_1$.

We make the following assumptions, which are satisfied for many elliptic and parabolic optimal control problems, see [].

$(A5)_q$ There are $q \in (p, \infty]$ and Banach spaces $\Sigma \subset \Lambda^*$, $V \subset Y^*$ such that the following holds:

1. The mapping

$$(y, u, \lambda) \in Y \times L^q(\Omega) \times \Sigma \mapsto \ell_u(y, u, \lambda, 0, 0) \in L^q(\Omega)$$

is differentiable and its derivative is Lipschitz continuous on bounded subsets of $Y \times \mathcal{B} \times \Sigma$.

2. The mapping

$$(y, u) \in Y \times U \mapsto J_y(y, u) \in V$$

is differentiable and its derivative is Lipschitz continuous on bounded subsets of $Y \times \mathcal{B}$.

3. The operator

$$(y,u) \in Y \times U \mapsto c_y^*(y,u) \in \mathcal{L}(\Sigma, V)$$

is differentiable and its derivative is Lipschitz continuous on bounded subsets of $Y \times \mathcal{B}$.

4. The following mappings are continuous and uniformly bounded on bounded sets:

$$\begin{aligned}
&(y,u) \in Y \times (\mathcal{B}, \|\cdot\|_q) \mapsto J_u(y,u) \in L^q(\Omega),\\
&(y,u) \in Y \times U \mapsto c_u^*(y,u) \in \mathcal{L}(\Sigma, L^q(\Omega)),\\
&(y,u) \in Y \times U \mapsto c_y^{-*}(y,u) \in \mathcal{L}(V, \Sigma),\\
&(y,u,\lambda) \in Y \times U \times \Sigma \mapsto \ell_{uy}(y,u,\lambda,0,0) \in \mathcal{L}(Y, L^q(\Omega)),\\
&(y,u,\lambda) \in Y \times U \times \Sigma \mapsto \ell_{yu}(y,u) \in \mathcal{L}(L^2(\Omega), V),\\
&(y,u,\lambda) \in Y \times U \times \Sigma \mapsto \ell_{uu}(y,u) \in \mathcal{L}(L^t(\Omega), L^t(\Omega)), \quad t \in [2,q].
\end{aligned}$$

5. The reduced gradient has the structure

$$\ell_u(y,u,\lambda,0,0) = \beta(u) + \hat{g}_s(y,u,\lambda), \quad \beta \in C^1(\mathbb{R}),\ \beta' \geq \alpha_0 > 0,$$

where

$$(y,u,\lambda) \in Y \times U \times \Sigma \mapsto \hat{g}_s(y,u,\lambda) \in L^q(\Omega)$$

is Lipschitz continuous on bounded sets.

6. The reduced Hessian

$$\hat{H}(y,u,\lambda) := \ell_{uu} + c_u^* c_y^{-*} \ell_{yy} c_y^{-1} c_u - c_u^* c_y^{-*} \ell_{yu} - \ell_{uy} c_y^{-1} c_u \tag{3.1}$$

has the structure

$$\hat{H}(y,u,\lambda) = \beta'(u) I + \hat{H}_s(y,u,\lambda),$$

where $(y,u,\lambda) \in Y \times U \times \Sigma \mapsto \hat{H}_s(y,u,\lambda) \in \mathcal{L}(L^2, L^q)$ is uniformly bounded on bounded subsets.

3.1. Verification for the elliptic control problem.

3.1.1. First setting. In the first setting we can verify (A5$_q$) for any $q > 2$ such that $Y = H_0^1(\Omega)$ is continuously embedded in $L^q(\Omega)$. We do not need the additional spaces V and Σ, since we just can choose $V = Y^* = H^{-1}(\Omega)$ and $\Sigma = \Lambda^* = H_0^1(\Omega) = Y$.

We have

$$\begin{aligned}
\ell(y,u,\lambda,z_a,z_b) &= \frac{1}{2}\|y-y_d\|_2^2 + \frac{\alpha}{2}\|u\|_2^2\\
&\quad + \langle \lambda, Ay - u\rangle - (z_a, u-a)_2 + (z_b, u-b)_2,\\
J_y(y,u) &= y - y_d, \quad J_u(y,u) = \alpha u,\\
\ell_y(y,u,\lambda,z_a,z_b) &= J_y(y,u) + A^*\lambda = y - y_d + A\lambda,\\
\ell_u(y,u,\lambda,z_a,z_b) &= \alpha u - \lambda - z_a + z_b,\\
\ell_{yu}(y,u,\lambda,z_a,z_b) &= 0,\\
\ell_{uy}(y,u,\lambda,z_a,z_b) &= 0,\\
\ell_{yy}(y,u,\lambda,z_a,z_b) &= I_{Y,Y^*},\\
\ell_{uu}(y,u,\lambda,z_a,z_b) &= \alpha I.
\end{aligned}$$

Therefore, (A5$_q$) is a direct consequence of the following observations:

1. $\ell_u(y,u,\lambda,0,0) = \alpha u - \lambda \in L^q(\Omega) + \Sigma = L^q(\Omega)$.
2. $J_y(y,u) = y - y_d \in L^2 \subset H^{-1} = V$.
3. $c_y^*(y,u) = A \in \mathcal{L}(H_0^1, H^{-1}) = \mathcal{L}(\Sigma, V)$.
4. $J_u(y,u) = \alpha u$,
   $c_u^*(y,u) = I_{\Lambda^*,U^*} = I_{H_0^1,L^2} \in \mathcal{L}(\Sigma, L^q(\Omega))$,
   $c_y^{-*}(y,u) = A^{-1} \in \mathcal{L}(H^{-1}, H_0^1) = \mathcal{L}(V,\Sigma)$,
   $\ell_{uy}(y,u,\lambda,0,0) = 0$,
   $\ell_{yu}(y,u) = 0$,
   $\ell_{uu}(y,u) = \alpha I$.
5. $\ell_u(y,u,\lambda,0,0) = \alpha u - \lambda = \beta(u) + \hat{g}_s(y,u,\lambda)$
   with $\beta(t) = \alpha t,\ \beta'(t) = \alpha,\ \hat{g}_s(y,u,\lambda) = -\lambda$.
   Hence, $\beta \in C^1(\mathbb{R})$, $\beta'(t) = \alpha \geq \alpha_0$ for any $\alpha_0 \in (0,\alpha]$, and $\hat{g}_s(y,u,\lambda) = -\lambda \in H_0^1 \subset L^q(\Omega)$.
6. $\hat{H}(y,u,\lambda) = \alpha I + I_{H_0^1,L^2}A^{-1}I_{H_0^1,H^{-1}}A^{-1}I_{L^2,H^{-1}} = \alpha I + A^{-2} = \beta'(u)I + \hat{H}_s(y,u,\lambda)$,
   where $\hat{H}_s(y,u,\lambda) = A^{-2} \in \mathcal{L}(H^{-1}, H_0^1) \subset \mathcal{L}(L^2, L^q)$ is constant and thus uniformly bounded on bounded subsets.

3.1.2. Second setting. Here, we choose $V = L^2$, $\Sigma = \tilde{Y}$, and verify (A5$_q$) for any $q \in (2,\infty]$ satisfying $\tilde{Y} = H_0^1 \cap H^2 \subset L^q$ (continuous embedding). In particular, for $n \leq 3$, we have $\tilde{Y} \subset L^\infty$.

It is quite obvious that the operator $A \in \mathcal{L}(H_0^1, H^{-1})$ from the first setting is the unique extension of $\tilde{A} \in \mathcal{L}(\tilde{Y}, L^2) \subset \mathcal{L}(\tilde{Y}, H^{-1})$ to $H_0^1 \supset \tilde{Y}$. Therefore, $\tilde{A}^* \in \mathcal{L}(L^2, \tilde{Y}^*)$ is the unique extension of $A^* = A \in \mathcal{L}(H_0^1, H^{-1}) \subset \mathcal{L}(H_0^1, \tilde{Y}^*)$ to $L^2 \supset H_0^1$. Since, as said, $A = A^*$ is the unique extension of $\tilde{A}$, this shows that $\tilde{A}^* \in \mathcal{L}(L^2, \tilde{Y}^*)$ is the unique extension of $\tilde{A} \in \mathcal{L}(\tilde{Y}, L^2) \subset \mathcal{L}(\tilde{Y}, \tilde{Y}^*)$ to L^2. Hence,

(3.2) $\quad y \in \tilde{Y} \mapsto \tilde{A}^* y \in L^2$ is continuous linear and boundedly invertible in $\mathcal{L}(\tilde{Y}, L^2)$.

The derivatives can be computed similar to the first setting. The validity of (A5$_q$) follows from

1. $\ell_u(y,u,\lambda,0,0) = \alpha u - \lambda \in L^q(\Omega) + \Sigma = L^q(\Omega) + \tilde{Y} = L^q(\Omega)$.
2. $J_y(y,u) = y - y_d \in L^2 = V$.
3. $c_y^*(y,u) = \tilde{A}^* \in \mathcal{L}(\tilde{Y}, L^2) = \mathcal{L}(\Sigma, V)$ (see (3.2)).
4. $J_u(y,u) = \alpha u$,
   $c_u^*(y,u) = I_{\Lambda^*,U^*} = I_{L^2} \in \mathcal{L}(\Sigma, L^q(\Omega))$,
   $c_y^{-*}(y,u) = \tilde{A}^{-*} \in \mathcal{L}(L^2, \tilde{Y}) = \mathcal{L}(V,\Sigma)$ (see (3.2)),
   $\ell_{uy}(y,u,\lambda,0,0) = 0$,
   $\ell_{yu}(y,u) = 0$,
   $\ell_{uu}(y,u) = \alpha I$.
5. $\ell_u(y,u,\lambda,0,0) = \alpha u - \lambda = \beta(u) + \hat{g}_s(y,u,\lambda)$
   with $\beta(u) = \alpha t,\ \beta'(t) = \alpha,\ \hat{g}_s(y,u,\lambda) = -\lambda$.
   Hence, $\beta \in C^1(\mathbb{R})$, $\beta'(t) = \alpha \geq \alpha_0$ for any $\alpha_0 \in (0,\alpha]$, and
   $$\hat{g}_s(y,u,\lambda) = -\lambda \in \Sigma = \tilde{Y} \subset L^q(\Omega) \quad \text{for } \lambda \in \Sigma.$$
6. By (3.2), $\hat{H}(y,u,\lambda) = \alpha I + \tilde{A}^{-*}I_{\tilde{Y},\tilde{Y}^*}\tilde{A}^{-1} = \alpha I + \tilde{A}^{-2} = \beta'(u)I + \hat{H}_s(y,u,\lambda)$,
   where $\hat{H}_s(y,u,\lambda) = \tilde{A}^{-2} \in \mathcal{L}(L^2, \tilde{Y}) \subset \mathcal{L}(L^2, L^q)$ is constant and thus uniformly bounded on bounded subsets.

4. Properties of the central path. We study next the regularity of the dual variables $\bar{z}_a, \bar{z}_b$ on the central path.

LEMMA 4.1. *Let* (A1)–(A5)$_q$ *hold and let* $(y_\mu, u_\mu, \lambda_\mu, z_{a,\mu}, z_{b,\mu})$ *be a solution of (2.7). Then there holds* $\lambda_\mu \in \Sigma$,

$$0 \le z_{a,\mu}, z_{b,\mu} \le \max(3\mu/\nu, 2|J_u(x_\mu) + c_u(x_\mu)^*\lambda_\mu|), \tag{4.1}$$

and thus with (A5)$_q$

$$\|z_{a,\mu}\|_q, \|z_{b,\mu}\|_q \le \|\max(3\mu/\nu, 2|J_u(x_\mu) + c_u(x_\mu)^*\lambda_\mu|)\|_q < \infty. \tag{4.2}$$

Proof. From the first equation in (2.7) we see that

$$\lambda_\mu = -c_y(x_\mu)^{-*}J_y(x_\mu) \in c_y(x_\mu)^{-*}V \subset \Sigma.$$

We have

$$z_{a,\mu} = \frac{\mu}{u_\mu - a}, \quad z_{b,\mu} = \frac{\mu}{b - u_\mu}.$$

This yields on the set $M = \{u_\mu - a \le (b - u_\mu)/2\}$ the estimate $z_{a,\mu}|_M \ge 2z_{b,\mu}|_M$ and thus by (2.7)

$$0 \le \frac{1}{2} z_{a,\mu}|_M \le z_{a,\mu}|_M - z_{b,\mu}|_M = (J_u(x_\mu) + c_u(x_\mu)^*\lambda_\mu)|_M.$$

On the complement $M^c := \Omega \setminus M$ we have

$$\frac{3}{2}(u_\mu - a)|_{M^c} \ge \frac{1}{2}(b - a)|_{M^c} \ge \frac{1}{2}\nu$$

and thus $(u_\mu - a)|_{M^c} \ge \nu/3$. Both cases together prove (4.1) for $z_{a,\mu}$. The estimate for $z_{b,\mu}$ is obtained in the same way. Since $\lambda_\mu \in \Sigma$, the right hand side of (4.1) is in L^q by (A5)$_q$ and thus (4.2) is obvious. □

We introduce for $s \in [1, \infty]$ the function spaces

$$W_s := Y \times L^s \times \Sigma \times L^s \times L^s,$$
$$W_s' := V \times L^s \times \Lambda \times L^s \times L^s,$$

equipped with the norms

$$\|(y, u, \lambda, z_a, z_b)\|_{W_s} = \|y\|_Y + \|\lambda\|_\Sigma + \left\|\sqrt{u^2 + z_a^2 + z_b^2}\right\|_s,$$
$$\|(y, u, v, z_a, z_b)\|_{W_s'} = \|y\|_V + \|v\|_\Lambda + \left\|\sqrt{u^2 + z_a^2 + z_b^2}\right\|_s.$$

REMARK 4.2. The choice of the Euclidean norm for $(u(\xi), z_a(\xi), z_b(\xi)) \in \mathbb{R}^3$ will be convenient, since we will later use a pointwise orthogonal projection of these components with respect to the Euclidean inner product on $\mathbb{R}^3$.□

As a direct consequence of the previous lemma all solutions of the perturbed optimality conditions (2.7) are contained in a bounded set of W_q.

COROLLARY 4.3. *Let* (A1)–(A5)$_q$ *hold. Then for any* $\mu_0 > 0$ *there exists a constant* $C_{\mu_0} > 0$ *such that for all* $0 < \mu \le \mu_0$ *any solution* $w_\mu = (y_\mu, u_\mu, \lambda_\mu, z_{a,\mu}, z_{b,\mu})$ *of (2.7) is in* W_q *and*

$$\|w_\mu\|_{W_q} \le C_{\mu_0}.$$

Proof. Since $u_\mu \in \mathcal{B}$ we have $\|u_\mu\|_q \le \|a\|_q + \|b\|_q =: C_u$. By Remark 2.1 $u \in U \to y(u) \in Y$ is Lipschitz continuous on $\mathcal{B}$ and therefore $\|y_\mu\|_Y = \|y(u_\mu)\|_Y \le C_y$ with a constant C_y. Now (2.7) yields

$$\lambda_\mu = -c_y(x_\mu)^{-*} J_y(x_\mu)$$

and by (A5)$_q$ the right hand side is uniformly bounded in V, since x_μ lies in a bounded subset of $Y \times \mathcal{B}$. Finally, this implies with (A5)$_q$ that the right hand side of (4.2) is uniformly bounded. The proof is complete. ☐

If (A5)$_\infty$ is satisfied then we can deduce immediately that solutions of the barrier problem are true interior points. In fact, we have the simple

COROLLARY 4.4. *Let* (A1)–(A5)$_\infty$ *hold. Then for any* $\mu_0 > 0$ *there exists a constant* $C_{\mu_0} > 0$ *such that for all* $0 < \mu \le \mu_0$ *any solution* $w_\mu = (y_\mu, u_\mu, \lambda_\mu, z_{a,\mu}, z_{b,\mu})$ *of (2.7) satisfies*

$$u_\mu - a \ge \frac{\mu}{C_{\mu_0}}, \quad b - u_\mu \ge \frac{\mu}{C_{\mu_0}}.$$

Proof. Corollary (4.3) yields a constant $C_{\mu_0} > 0$ with

$$\|z_{a,\mu}\|_\infty, \|z_{b,\mu}\|_\infty \le \|w_\mu\|_{W_\infty} \le C_{\mu_0}.$$

Now the last two equations in (2.7) yield

$$u_\mu - a = \frac{\mu}{z_{a,\mu}} \ge \frac{\mu}{C_{\mu_0}}, \quad b - u_\mu = \frac{\mu}{z_{b,\mu}} \ge \frac{\mu}{C_{\mu_0}}.$$

☐

REMARK 4.5. For linear elliptic control problems, which satisfy (A1)–(A5)$_\infty$, this result was shown in [11], where it is used to prove the existence of solutions for (2.4). We used a different proof to cover also the more difficult case that (A5)$_q$ holds only for some $q < \infty$. ☐

We show next, that the dual variables $\lambda_\mu, z_{a,\mu}, z_{b,\mu}$ depend continuously on the primal variables y_μ, u_μ.

LEMMA 4.6. *Let* (A1)–(A5)$_q$ *hold and let* $(y_\mu, u_\mu, \lambda_\mu, z_{a,\mu}, z_{b,\mu})$ *be a solution of (2.7). Then*

$$\lambda_\mu = -c_y(x_\mu)^{-*} J_y(x_\mu)$$

and for any measurable sets $M, N \subset \Omega$ *one has*

$$\begin{aligned} z_{a,\mu}|_M &= \left.\frac{\mu}{u_\mu - a}\right|_M, \\ z_{b,\mu}|_N &= \left.\frac{\mu}{b - u_\mu}\right|_N, \\ z_{a,\mu}|_N &= (z_{b,\mu} + J_u(x_\mu) + c_u(x_\mu)^* \lambda_\mu)|_N \\ z_{b,\mu}|_M &= (z_{a,\mu} - J_u(x_\mu) - c_u(x_\mu)^* \lambda_\mu)|_M. \end{aligned} \tag{4.3}$$

Moreover, if $u_\eta \to u_\mu$ *in* $L^q(\Omega)$ *as* $\eta \to \mu$ *then*

$$(u_\eta, y_\eta, \lambda_\eta, z_{a,\eta}, z_{b,\eta}) \to (u_\mu, y_\mu, \lambda_\mu, z_{a,\mu}, z_{b,\mu}) \quad \textit{in } W_q.$$

Proof. The equations for λ_μ, $z_{a,\mu}$, and $z_{b,\mu}$ follow directly from (2.7).

Now assume that $u_\eta \to u_\mu$ in $L^q(\Omega)$ as $\eta \to \mu$. Then $y_\eta = y(u_\eta) \to y_\mu = y(u_\mu)$ in Y follows from (A1)–(A3), see Remark 2.1. Moreover,

$$\lambda_\eta = -c_y(x_\eta)^{-*} J_y(x_\eta) \to \lambda_\mu = -c_y(x_\mu)^{-*} J_y(x_\mu) \quad \text{in } \Sigma$$

is a direct consequence of (A5)$_q$.

Finally, we know by Lemma 4.1 that $z_{a,\eta}, z_{b,\eta} \in L^q(\Omega)$ for all $\eta > 0$.

Now we use the formulas (4.3) with

$$\begin{aligned} M = M_\eta &:= \{u_\eta \geq (b+a)/2,\ u_\mu \geq (b+a)/2\}\,, \\ N = N_\eta &:= \{u_\eta < (b+a)/2,\ u_\mu < (b+a)/2\}\,. \end{aligned}$$

Then be obtain on M_η by (4.3)

$$\begin{aligned} |(z_{a,\eta} - z_{a,\mu})_{M_\eta}| &= \left| \left(\frac{\eta}{u_\eta - a} - \frac{\mu}{u_\mu - a} \right)_{M_\eta} \right| \leq \frac{|\eta - \mu|}{u_\eta - a}\bigg|_{M_\eta} + \frac{\mu |u_\eta - u_\mu|}{(u_\eta - a)(u_\mu - a)}\bigg|_{M_\eta} \\ &\leq \frac{2|\eta - \mu|}{\nu} + \frac{4\mu}{\nu^2} |(u_\eta - u_\mu)_{M_\eta}| \end{aligned}$$

and thus

$$\left\|(z_{a,\eta} - z_{a,\mu})|_{M_\eta}\right\|_q \leq C(|\eta - \mu| + \left\|(u_\eta - u_\mu)|_{M_\eta}\right\|_q) \to 0 \quad \text{as } \eta \to \mu.$$

Now (4.3) yields with (A5)$_q$

$$\begin{aligned} \left\|(z_{b,\eta} - z_{b,\mu})|_{M_\eta}\right\|_q &\leq \left\|(z_{a,\eta} - z_{a,\mu})|_{M_\eta}\right\|_q + \left\|J_u(x_\eta) - J_u(x_\mu)\right\|_q \\ &\quad + \left\|c_u(x_\eta)^* \lambda_\eta - c_u(x_\mu)^* \lambda_\mu)\right\|_q \to 0 \quad \text{as } \eta \to \mu. \end{aligned}$$

In the same way we obtain

$$\left\|(z_{a,\eta} - z_{a,\mu})|_{N_\eta}\right\|_q + \left\|(z_{b,\eta} - z_{b,\mu})|_{N_\eta}\right\|_q \to 0 \quad \text{as } \eta \to \mu.$$

Finally, (4.3) yields on $J_\eta := \Omega \setminus (M_\eta \cup N_\eta)$

$$\text{sgn}(z_{a,\eta} - z_{a,\mu})|_{J_\eta} \neq -\,\text{sgn}(z_{b,\eta} - z_{b,\mu})|_{J_\eta}$$

and thus the difference of the second equation in (2.7) for η and μ, respectively, yields

$$\begin{aligned} &\left\|(z_{b,\eta} - z_{b,\mu})|_{J_\eta}\right\|_q + \left\|(z_{a,\eta} - z_{a,\mu})|_{J_\eta}\right\|_q = \left\|(z_{b,\eta} - z_{b,\mu} + z_{a,\mu} - z_{a,\eta})|_{J_\eta}\right\|_q \\ &\leq \left\|J_u(x_\eta) - J_u(x_\mu)\right\|_q + \left\|c_u(x_\eta)^* \lambda_\eta - c_u(x_\mu)^* \lambda_\mu)\right\|_q \to 0 \quad \text{as } \eta \to \mu. \end{aligned}$$

Since $M_\eta \cup N_\eta \cup J_\eta = \Omega$, we have shown that

$$\left\|z_{a,\eta} - z_{a,\mu}\right\|_q + \left\|z_{b,\eta} - z_{b,\mu}\right\|_q \to 0 \quad \text{as } \eta \to \mu$$

which concludes the proof. □

5. Analysis of the primal-dual Newton system. Throughout this section we assume that (A1)–(A5)$_q$ with some $q > p$ hold.

5.1. Primal-dual Newton system. The formal application of Newton's method to the perturbed KKT-system (2.7) yields with the multiplication operators

$$Z_a := z_a \cdot I, \quad Z_b := z_b \cdot I, \quad U_a := (u-a) \cdot I, \quad U_b := (b-u) \cdot I$$

the primal-dual Newton system

$$\begin{pmatrix} \ell_{yy} & \ell_{yu} & c_y^* & 0 & 0 \\ \ell_{uy} & \ell_{uu} & c_u^* & -I & I \\ c_y & c_u & 0 & 0 & 0 \\ 0 & Z_a & 0 & U_a & 0 \\ 0 & -Z_b & 0 & 0 & U_b \end{pmatrix} \begin{pmatrix} s_y \\ s_u \\ s_\lambda \\ s_a \\ s_b \end{pmatrix} = - \begin{pmatrix} \ell_y(y,u,\lambda,z_a,z_b) \\ \ell_u(y,u,\lambda,z_a,z_b) \\ c(y,u) \\ z_a(u-a)-\mu \\ z_b(b-u)-\mu \end{pmatrix}.$$

We write this briefly as

$$DF_\mu(w)\, s = -F_\mu(w).$$

For convenience, we will also use the abbreviations

$$u_a = u - a, \quad u_b = b - u.$$

To ensure a certain quality of the primal-dual Newton step, we will keep the iteration in the following wide neighborhood of the central path

$$\begin{aligned} N_{-\infty,q}(\mu) := \Big\{ & (y,u,\lambda,z_a,z_b) \in Y \times \mathcal{B} \times \Sigma \times L^q \times L^q \; : \\ & (u-a)z_a \geq \gamma\mu, \;\; (b-u)z_b \geq \gamma\mu, \\ & z_a|_{\{u>(b+a)/2\}} \leq \frac{2\max(\mu_{-\infty},\mu)}{\nu}, \;\; z_b|_{\{u<(b+a)/2\}} \leq \frac{2\max(\mu_{-\infty},\mu)}{\nu}, \\ & \min(z_a|_{\{u=(b+a)/2\}}, z_b|_{\{u=(b+a)/2\}}) \leq \frac{2\max(\mu_{-\infty},\mu)}{\nu} \Big\} \end{aligned}$$

with constants $\gamma \in (0,1)$, $\mu_{-\infty} > 0$. By Corollary 4.3 any solution of (2.7) is contained in $N_{-\infty,q}(\mu)$.

Notation. For multiplication operators S, T associated with measurable functions $s, t : \Omega \to \mathbb{R}$ we write $S \geq T$ if and only if $s \geq t$ almost everywhere. $S > T$, $S \leq T$ and $S < T$ are defined analogously.

5.2. Regularity of the primal-dual Newton system. We will now show that under a coercivity condition for the reduced Hessian $\hat{H}$ and under assumptions (A1)-(A5)$_q$ for all $w \in N_{-\infty,q}(\mu)$ the operator $DF_\mu(w) \in \mathcal{L}(W_t, W_t')$, $t \in [p,q]$ has an inverse with

$$\|DF_\mu(w)^{-1}\|_{W_t', W_t} \leq \frac{C}{\min(1, \sqrt{\mu})}.$$

Moreover, if we premultiply $DF_\mu(w)$ by the scaling operator

(5.1) $$S(w) = \begin{pmatrix} I & & & & \\ & I & & & \\ & & I & & \\ & & & (U_a+Z_a)^{-1} & \\ & & & & (U_b+Z_b)^{-1} \end{pmatrix}$$

we will show that even

$$\|(S(w)DF_\mu(w))^{-1}\|_{W_t',W_t} \le C.$$

Setting

$$\hat{U}_a = (U_a + Z_a)^{-1}U_a, \quad \hat{U}_b = (U_b + Z_b)^{-1}U_b, \\ \hat{Z}_a = (U_a + Z_a)^{-1}Z_a, \quad \hat{Z}_b = (U_b + Z_b)^{-1}Z_b, \tag{5.2}$$

we have $\hat{U}_a + \hat{Z}_a = I$, $\hat{U}_b + \hat{Z}_b = I$ and

$$S(w)DF_\mu(w) = \begin{pmatrix} \ell_{yy} & \ell_{yu} & c_y^* & 0 & 0 \\ \ell_{uy} & \ell_{uu} & c_u^* & -I & I \\ c_y & c_u & 0 & 0 & 0 \\ 0 & \hat{Z}_a & 0 & \hat{U}_a & 0 \\ 0 & -\hat{Z}_b & 0 & 0 & \hat{U}_b \end{pmatrix}$$

where we omit the arguments. For convenience, we use also the abbreviations

$$\hat{u}_a = \frac{u_a}{u_a + z_a}, \quad \hat{z}_a = \frac{z_a}{u_a + z_a}, \quad \hat{u}_b = \frac{u_b}{u_b + z_b}, \quad \hat{z}_b = \frac{z_b}{u_b + z_b}.$$

We show first that $S(w)DF_\mu(w)$ has a bounded inverse. This fact will play an essential role in this paper.

LEMMA 5.1. *Let (A1)–(A5)$_q$ hold for some $q \in]2,\infty]$ and let $w = (y, u, \lambda, z_l, z_r) \in N_{-\infty,q}(\mu)$ for some $\gamma \in (0, 1)$, $\mu_{-\infty} > 0$.*

If the reduced Hessian $\hat{H}(y, u, \lambda)$ in (3.1) satisfies

$$(v, \hat{H}(y, u, \lambda)v) \ge \alpha\|v\|_2^2 \quad \forall\, v \in L^2(\Omega) \tag{5.3}$$

with some $\alpha > 0$ then $S(w)DF_\mu(w)$ has a bounded inverse in $\mathcal{L}(W_t', W_t)$ for all $t \in [p, q]$ and

$$\|(S(w)DF_\mu(w))^{-1}\|_{W_t',W_t} \le C$$

with a constant $C > 0$. The constant C can be chosen uniformly on bounded subsets of $\{(\mu, w) \in (0,\infty) \times N_{-\infty,q}(\mu)\}$ on which (5.3) holds uniformly.

Moreover, also $DF_\mu(w)$ has a bounded inverse in $\mathcal{L}(W_t', W_t)$ for all $t \in [p, q]$ and

$$\|DF_\mu(w)^{-1}\|_{W_t',W_t} \le \frac{C'}{\min(1, \sqrt{\mu})},$$

where $C' = \frac{C}{\min(1,2\sqrt{\gamma})}$ with the above constant C.

Proof. We consider the equation

$$\begin{pmatrix} \ell_{yy} & \ell_{yu} & c_y^* & 0 & 0 \\ \ell_{uy} & \ell_{uu} & c_u^* & -I & I \\ c_y & c_u & 0 & 0 & 0 \\ 0 & \hat{Z}_a & 0 & \hat{U}_a & 0 \\ 0 & -\hat{Z}_b & 0 & 0 & \hat{U}_b \end{pmatrix} \begin{pmatrix} s_y \\ s_u \\ s_\lambda \\ s_a \\ s_b \end{pmatrix} = \begin{pmatrix} r_y \\ r_u \\ r_\lambda \\ r_a \\ r_b \end{pmatrix} \tag{5.4}$$

with $r = (r_y, r_u, r_\lambda, r_a, r_b) \in W_t'$, $t \in [p, q]$, where we omit the arguments.

We note that by (A5)$_q$ the operator $S(w)DF_\mu(w)$ on the left hand side of (5.4) is in $\mathcal{L}(W_t, W_t')$, $p \le t \le q$.

Elimination with the last two block rows and subsequently with the first and third block row yields as above with $\hat{H}(y,u,\lambda)$ in (3.1) the reduced system

$$(\hat{H}(y,u,\lambda) + \hat{U}_a^{-1}\hat{Z}_a + \hat{U}_b^{-1}\hat{Z}_b)s_u = \\ = r_u + \hat{U}_a^{-1}r_a - \hat{U}_b^{-1}r_b + B_1 r_\lambda - B_2 r_y =: r'_u. \tag{5.5}$$

with the abbreviations

$$B_1 = c_u^* c_y^{-*}\ell_{yy}c_y^{-1} - \ell_{uy}c_y^{-1}, \quad B_2 = c_u^* c_y^{-*}.$$

Note that $B_1 \in \mathcal{L}(\Lambda, L^q(\Omega))$, $B_2 \in \mathcal{L}(V, L^q(\Omega))$ by (A5)$_q$.

Let

$$0 < \varepsilon \le 1/2$$

and define $I_1 = \{\hat{U}_a \le \varepsilon, \hat{U}_b > \varepsilon\}$, $I_2 = \{\hat{U}_b \le \varepsilon, \hat{U}_a > \varepsilon\}$, $I_3 = \{\hat{U}_a \le \varepsilon, \hat{U}_b \le \varepsilon\}$ and $I_4 = \Omega \setminus (I_1 \cup I_2 \cup I_3)$. Multiply by

$$D = \varepsilon I|_{I_4} + \hat{U}_a|_{I_1} + \hat{U}_b|_{I_2} + \min(\hat{U}_a, \hat{U}_b)|_{I_3},$$

where $I|_{I_4} = 1_{I_4}I$, $\hat{U}_a|_{I_1} = \hat{u}_a 1_{I_1}, \ldots$.

Then we obtain on I_1

$$(\hat{U}_a|_{I_1}\hat{H}(y,u,\lambda) + (\hat{Z}_a + \hat{U}_a\hat{U}_b^{-1}\hat{Z}_b)|_{I_1})s_u = (\hat{U}_a r_u + r_a - \hat{U}_a\hat{U}_b^{-1}r_b + \hat{U}_a B_1 r_\lambda - \hat{U}_a B_2 r_y)|_{I_1}.$$

We have

$$\hat{U}_a\hat{U}_b^{-1}|_{I_1} \le I, \quad \hat{Z}_a|_{I_1} \ge 1/2.$$

Thus, the right hand side is pointwise $\le |r_u| + |r_a| + |r_b| + |B_1 r_\lambda| + |B_2 r_y|$ and the operator on the left has the form $(\delta|_{I_1}I + D|_{I_1}\hat{H})$ with a function $\delta \in L^\infty(\Omega)$, $\delta \ge 1/2$.

On I_2 the situation is analogous. Similarly, we have on I_3 with $\hat{U}_{ab} := \min(\hat{u}_a, \hat{u}_b) \cdot I$

$$(\hat{U}_{ab}|_{I_3}\hat{H}(y,u,\lambda) + (\hat{U}_{ab}\hat{U}_a^{-1}\hat{Z}_a + \hat{U}_{ab}\hat{U}_b^{-1}\hat{Z}_b)|_{I_3})s_u \\ = (\hat{U}_{ab}r_u + \hat{U}_{ab}\hat{U}_a^{-1}r_a - \hat{U}_{ab}\hat{U}_b^{-1}r_b + \hat{U}_{ab}B_1 r_\lambda - \hat{U}_{ab}B_2 r_y)|_{I_3}.$$

Again, the right hand side is pointwise $\le |r_u| + |r_a| + |r_b| + |B_1 r_\lambda| + |B_2 r_y|$ and the operator on the left has the form $\delta|_{I_3}I + D|_{I_3}\hat{H})$ with $\delta \in L^\infty(\Omega)$, $\delta \ge 1/2$.

On I_4 we obtain

$$(\varepsilon I|_{I_4}\hat{H}(y,u,\lambda) + \varepsilon(\hat{U}_a^{-1}\hat{Z}_a + \hat{U}_b^{-1}\hat{Z}_b)|_{I_4})s_u = \varepsilon(r_u + \hat{U}_a^{-1}r_a - \hat{U}_b^{-1}r_b + B_1 r_\lambda - B_2 r_y)|_{I_4}.$$

Since $\hat{U}_a|_{I_4} > \varepsilon$ and $\hat{U}_b|_{I_4} > \varepsilon$ the right hand side is $\le \varepsilon|r_u| + |r_a| + |r_b| + \varepsilon|B_1 r_\lambda| + \varepsilon|B_2 r_y|$. The operator has on I_4 the form $\delta_{I_4}I + \varepsilon I_{I_4}\hat{H}$ with $\delta \ge 0$.

Thus, after multiplication with D the operator on the left hand side has the form $\delta I + D\hat{H}$ with $\delta|_{I_1 \cup I_2 \cup I_3} \ge 1/2$ and $\delta|_{I_4} \ge 0$ and the right hand side is pointwise $\le |r_u| + |r_a| + |r_b| + |B_1 r_\lambda| + |B_2 r_y|$. Moreover, we have

$$\| |r_u| + |r_a| + |r_b| + |B_1 r_\lambda| + |B_2 r_y| \|_t \le (3 + \|B_1\|_{\Lambda, L^t} + \|B_2\|_{\Sigma, L^t})\|r\|_{W'_t}. \tag{5.6}$$

Let without restriction $\alpha \le 1/2$ in (5.3). Then we have for all $s \in L^2(\Omega)$ with the abbreviations $s_i = s_{I_i}$, $i = 1, \ldots, 4$,

$$\begin{aligned}(s, (\delta I + D\hat{H})s) \ge\ & \alpha(s_1, s_1) + \alpha(s_2, s_2) + \alpha(s_3, s_3) + \varepsilon\alpha(s_4, s_4) \\ & + \varepsilon(s_4, \hat{H}(s_1 + s_2 + s_3)) + (s_1, \hat{U}_a\hat{H}s) + (s_2, \hat{U}_b\hat{H}s) \\ & + (s_3, \hat{U}_{ab}\hat{H}s).\end{aligned}$$

Using that $(\beta u - v/\beta, \beta u - v/\beta) \geq 0$ and thus

$$2(u,v) \leq \beta^2 (u,u) + \frac{1}{\beta^2}(v,v)$$

we obtain

$$(s_1, \hat{U}_a \hat{H} s) \geq -\frac{\alpha}{4}\|s_1\|_2^2 - \frac{1}{\alpha}\|\hat{U}_a|_{I_1}\hat{H}s\|_2^2 \geq -\frac{\alpha}{4}\|s_1\|_2^2 - \frac{\varepsilon^2}{\alpha}\|\hat{H}s\|_2^2$$

and analogously

$$(s_2, \hat{U}_b \hat{H} s) \geq -\frac{\alpha}{4}\|s_2\|_2^2 - \frac{\varepsilon^2}{\alpha}\|\hat{H}s\|_2^2,$$
$$(s_3, \min(\hat{U}_a, \hat{U}_b)\hat{H} s) \geq -\frac{\alpha}{4}\|s_3\|_2^2 - \frac{\varepsilon^2}{\alpha}\|\hat{H}s\|_2^2.$$

Finally,

$$\varepsilon(s_4, \hat{H}(s_1+s_2+s_3)) \geq -\frac{\varepsilon\alpha}{4}\|s_4\|_2^2 - \frac{\varepsilon}{\alpha}\|\hat{H}(s_1+s_2+s_3)\|_2^2.$$

Now set

$$\varepsilon = \min(\alpha, \alpha^2)/(1 + 16\|\hat{H}\|_{L^2,L^2}^2).$$

Since $\|s\|_2^2 = \|s_1\|_2^2 + \|s_2\|_2^2 + \|s_3\|_2^2 + \|s_4\|_2^2$, inserting these estimates yields

$$(s, (\delta I + D\hat{H})s) \geq \frac{\alpha}{2}(s_1,s_1) + \frac{\alpha}{2}(s_2,s_2) + \frac{\alpha}{2}(s_3,s_3) + \frac{\varepsilon\alpha}{2}(s_4,s_4).$$

This shows together with (5.6) that

$$\|s_u\|_2 \leq \frac{C}{\varepsilon\alpha}\|r\|_{W_2'} \tag{5.7}$$

where C depends only on α and $\hat{H}$ but not on μ.

To obtain a bound in L^t-topology we multiply (5.5) by s_u. By the structure of $\hat{H}$ according to (A5)$_q$ this yields the pointwise estimate

$$\begin{aligned} 0 &\leq (\alpha_0 + \hat{u}_a^{-1}\hat{z}_a + \hat{u}_b^{-1}\hat{z}_b)s_u^2 \leq s_u(r_u - \hat{H}_s s_u) + s_u\hat{u}_a^{-1}r_a - s_u\hat{u}_b^{-1}r_b \\ &\quad + s_u(B_1 r_\lambda) - s_u(B_2 r_y). \end{aligned} \tag{5.8}$$

Since $\hat{u}_a + \hat{z}_a = \hat{u}_b + \hat{z}_b = 1$, we have

$$\frac{\hat{u}_a^{-1}}{\alpha_0 + \hat{u}_a^{-1}\hat{z}_a + \hat{u}_b^{-1}\hat{z}_b} \leq \frac{\hat{u}_a^{-1}}{\alpha_0 + \hat{u}_a^{-1}\hat{z}_a} = \frac{1}{\alpha_0\hat{u}_a + \hat{z}_a} \leq \frac{1}{\min(\alpha_0, 1)}$$

and analogously

$$\frac{\hat{u}_b^{-1}}{\alpha_0 + \hat{u}_a^{-1}\hat{z}_a + \hat{u}_b^{-1}\hat{z}_b} = \frac{1}{\alpha_0\hat{u}_b + \hat{z}_b} \leq \frac{1}{\min(\alpha_0, 1)}.$$

Division of (5.8) by $\left(\alpha_0 + \hat{u}_a^{-1}\hat{z}_a + \hat{u}_b^{-1}\hat{z}_b\right)|s_u|$ yields

$$|s_u| \leq \frac{1}{\alpha_0}|r_u - \hat{H}_s s_u + B_1 r_\lambda - B_2 r_y| + \frac{1}{\min(\alpha_0,1)}(|r_a| + |r_b|).$$

This yields together with (5.7) for all $t \in [p,q]$

$$
\begin{aligned}
\|s_u\|_t &\le \frac{1}{\alpha_0}\Big(\|r_u\|_t + \|\hat{H}_s\|_{L^2,L^t}\|s_u\|_2 + \|B_1\|_{\Lambda,L^t}\|r_\lambda\|_\Lambda + \|B_2\|_{V,L^t}\|r_y\|_V\Big) \\
&\quad + \frac{1}{\min(\alpha_0,1)}\left(\|r_a\|_t + \|r_b\|_t\right) \\
&\le C'\|r\|_{W_t'}.
\end{aligned} \tag{5.9}
$$

We derive now also bounds for s_y, s_λ, s_a, s_b. We have

$$
s_y = c_y^{-1}(-c_u s_u + r_\lambda)
$$

and thus $L^t \subset L^p$ for $t \in [p,q]$ yields

$$
\|s_y\|_Y \le \|c_y^{-1}c_u\|_{L^t,Y}\|s_u\|_t + \|c_y^{-1}\|_{\Lambda,Y}\|r_\lambda\|_\Lambda. \tag{5.10}
$$

Next, we obtain

$$
s_\lambda = c_y^{-*}(r_y - \ell_{yy}s_y - \ell_{yu}s_u),
$$

which yields by (A5)$_q$

$$
\|s_\lambda\|_\Sigma \le \|c_y^{-*}\|_{V,\Sigma}(\|r_y\|_V + \|\ell_{yy}\|_{Y,V}\|s_y\|_Y + \|\ell_{yu}\|_{L^t,V}\|s_u\|_t). \tag{5.11}
$$

To estimate s_a, s_b we partition Ω into the sets

$$
\Omega_a = \{u > (b+a)/2\} \cup \left\{u = (b+a)/2,\ \ z_a \le \frac{2\max(\mu_{-\infty},\mu)}{\nu}\right\}, \quad \Omega_a^c = \Omega \setminus \Omega_a.
$$

Now (5.4) yields

$$
\begin{aligned}
s_a|_{\Omega_a} &= \hat{U}_a^{-1}(r_a - \hat{Z}_a s_u)|_{\Omega_a}, \\
s_b|_{\Omega_a^c} &= \hat{U}_b^{-1}(r_b + \hat{Z}_b s_u)|_{\Omega_a^c}, \\
s_a|_{\Omega_a^c} &= (\ell_{uy}s_y + \ell_{uu}s_u + c_u^* s_\lambda + s_b - r_u)|_{\Omega_a^c}, \\
s_b|_{\Omega_a} &= (-\ell_{uy}s_y - \ell_{uu}s_u - c_u^* s_\lambda + s_a + r_u)|_{\Omega_a}.
\end{aligned}
$$

By the definition of the neighborhood $N_{-\infty,q}(\mu)$ we have

$$
z_a|_{\Omega_a} \le \frac{2\max(\mu_{-\infty},\mu)}{\nu} =: C_\mu, \quad z_b|_{\Omega_a^c} \le \frac{2\max(\mu_{-\infty},\mu)}{\nu} = C_\mu \tag{5.12}
$$

and thus

$$
\hat{u}_a|_{\Omega_a} = \left.\frac{u_a}{u_a + z_a}\right|_{\Omega_a} \ge \frac{1}{1+2C_\mu\nu^{-1}}, \quad \hat{u}_b|_{\Omega_a^c} = \left.\frac{u_b}{u_b + z_b}\right|_{\Omega_a^c} \ge \frac{1}{1+2C_\mu\nu^{-1}}.
$$

Using that $0 \le \hat{z}_a \le 1$, $0 \le \hat{z}_b \le 1$ this yields for all $t \in [p,q]$

$$
\begin{aligned}
\|s_a|_{\Omega_a}\|_t &\le (1+2C_\mu\nu^{-1})(\|r_a\|_t + \|s_u\|_t) \\
\|s_b|_{\Omega_a^c}\|_t &\le (1+2C_\mu\nu^{-1})(\|r_b\|_t + \|s_u\|_t) \\
\|s_a|_{\Omega_a^c}\|_t &\le \|l_{uy}\|_{Y,L^t}\|s_y\|_Y + \|l_{uu}\|_{L^t,L^t}\|s_u\|_t \\
&\quad + \|c_u^*\|_{\Sigma,L^t}\|s_\lambda\|_\Sigma + \|s_b|_{\Omega_a^c}\|_t + \|r_u\|_t \\
\|s_b|_{\Omega_a}\|_t &\le \|l_{uy}\|_{Y,L^t}\|s_y\|_Y + \|l_{uu}\|_{L^t,L^t}\|s_u\|_t \\
&\quad + \|c_u^*\|_{\Sigma,L^t}\|s_\lambda\|_\Sigma + \|s_a|_{\Omega_a}\|_t + \|r_u\|_t.
\end{aligned}
$$

We conclude that

$$\|s_a\|_t \le (1+2C_\mu\nu^{-1})(\|r_a\|_t + \|s_u\|_t) + C'(\|s_y\|_Y + \|s_u\|_t + \|s_\lambda\|_\Sigma + \|r_u\|_t), \tag{5.13}$$

$$\|s_b\|_t \le (1+2C_\mu\nu^{-1})(\|r_b\|_t + \|s_u\|_t) + C'(\|s_y\|_Y + \|s_u\|_t + \|s_\lambda\|_\Sigma + \|r_u\|_t). \tag{5.14}$$

Now (5.9), (5.10), (5.11), (5.13), (5.14) yield

$$\|(S(w)DF_\mu(w))^{-1}\|_{W_t',W_t} \le C''.$$

It is easy to check that C'' can be chosen uniformly on bounded subsets of $\{(\mu,w) \in (0,\infty) \times N_{-\infty,q}(\mu)\}$ on which (5.3) holds uniformly.

Finally, the definition of the neighborhood $N_{-\infty,q}(\mu)$ yields

$$u_a + z_a \ge 2\sqrt{u_a z_a} \ge 2\sqrt{\gamma\mu}, \quad u_b + z_b \ge 2\sqrt{u_b z_b} \ge 2\sqrt{\gamma\mu}.$$

Therefore, the scaling matrix $S(w)$ in (5.1) satisfies

$$\|S(w)\|_{W_t',W_t'} \le \frac{1}{\min(1, 2\sqrt{\gamma\mu})}$$

and is invertible. Thus, $DF_\mu(w)^{-1} = (S(w)DF_\mu(w))^{-1}S(w)$ and

$$\begin{aligned}\|DF_\mu(w)^{-1}\|_{W_t',W_t} &= \|(S(w)DF_\mu(w))^{-1}S(w)\|_{W_t',W_t} \\ &\le \|(S(w)DF_\mu(w))^{-1}\|_{W_t',W_t}\|S(w)\|_{W_t',W_t'} \le \frac{C''}{\min(1, 2\sqrt{\gamma\mu})}.\end{aligned}$$

□

We have the following variant of Lemma 5.1 that will be useful for showing the Hölder-continuity of the central path.

LEMMA 5.2. *Let the assumptions of Lemma 5.1 hold, but assume only that*

$$w = (y, u, \lambda, z_a, z_a) \in \frac{1}{2}(N_{-\infty,q}(\mu_1) + N_{-\infty,q}(\mu_2)) \tag{5.15}$$

with $\mu_1, \mu_2 \in (0,\infty)$ and set $\mu = \min(\mu_1, \mu_2)$.

Then $DF_\mu(w)$ has a bounded inverse in $\mathcal{L}(W_t', W_t)$ for all $t \in [p,q]$ and

$$\|DF_\mu(w)^{-1}\|_{W_t',W_t} \le \frac{C}{\min(1, \sqrt{\mu})} \tag{5.16}$$

with a constant $C > 0$. The constant C can be chosen uniformly on bounded subsets of

$$\left\{(\mu_1, \mu_2, w) \in (0,\infty)^2 \times \frac{1}{2}(N_{-\infty,q}(\mu_1) + N_{-\infty,q}(\mu_2))\right\}$$

on which (5.3) holds uniformly.

REMARK 5.3. (5.15) is weaker than $(y, u, \lambda, z_a, z_a) \in N_{-\infty,q}(\mu)$, since the constraints

$$z_a|_{\{u>(b+a)/2\}} \le \frac{2\max(\mu_{-\infty},\mu)}{\nu}, \quad z_b|_{\{u<(b+a)/2\}} \le \frac{2\max(\mu_{-\infty},\mu)}{\nu},$$
$$\min(z_a|_{\{u=(b+a)/2\}}, z_b|_{\{u=(b+a)/2\}}) \le \frac{2\max(\mu_{-\infty},\mu)}{\nu}$$

are nonconvex and can be violated by points satisfying (5.15).□

Proof. Let $\tilde{w} = (\tilde{y}, \tilde{u}, \tilde{\lambda}, \tilde{z}_a, \tilde{z}_b) \in N_{-\infty,q}(\mu_1)$, $\bar{w} = (\bar{y}, \bar{u}, \bar{\lambda}, \bar{z}_a, \bar{z}_b) \in N_{-\infty,q}(\mu_2)$, and $w = (y, u, \lambda, z_a, z_b) = \frac{1}{2}(\tilde{w} + \bar{w})$ according to (5.15). Without restriction we assume that $\mu_1 \leq \mu_2$ then $\mu = \min(\mu_1, \mu_2) = \mu_1$.

We modify the proof of Lemma 5.1, but consider this time the system

$$DF_\mu(w)s = \hat{r} =: \begin{pmatrix} r_y \\ r_u \\ r_\lambda \\ \hat{r}_a \\ \hat{r}_b \end{pmatrix}, \text{ i.e., } S(w)DF_\mu(w)s = \begin{pmatrix} r_y \\ r_u \\ r_\lambda \\ (U_a+Z_a)^{-1}\hat{r}_a \\ (U_b+Z_b)^{-1}\hat{r}_b \end{pmatrix} =: \begin{pmatrix} r_y \\ r_u \\ r_\lambda \\ r_a \\ r_b \end{pmatrix},$$

which yields

$$\begin{pmatrix} \ell_{yy} & \ell_{yu} & c_y^* & 0 & 0 \\ \ell_{uy} & \ell_{uu} & c_u^* & -I & I \\ c_y & c_u & 0 & 0 & 0 \\ 0 & \hat{Z}_a & 0 & \hat{U}_a & 0 \\ 0 & -\hat{Z}_b & 0 & 0 & \hat{U}_b \end{pmatrix} \begin{pmatrix} s_y \\ s_u \\ s_\lambda \\ s_a \\ s_b \end{pmatrix} = \begin{pmatrix} r_y \\ r_u \\ r_\lambda \\ (U_a+Z_a)^{-1}\hat{r}_a \\ (U_b+Z_b)^{-1}\hat{r}_b \end{pmatrix} = \begin{pmatrix} r_y \\ r_u \\ r_\lambda \\ r_a \\ r_b \end{pmatrix}, \tag{5.17}$$

with $\hat{U}_a, \hat{U}_b, \hat{Z}_a, \hat{Z}_b$ in (5.2).

$(y, u, \lambda, z_a, z_a)$ according to (5.15) satisfies all constraints of $N_{-\infty,q}(\mu)$ with the possible exception of the nonconvex constraints

$$\begin{aligned} &z_a|_{\{u>(b+a)/2\}} \leq \frac{2\max(\mu_{-\infty}, \mu)}{\nu}, \quad z_b|_{\{u<(b+a)/2\}} \leq \frac{2\max(\mu_{-\infty}, \mu)}{\nu}, \\ &\min(z_a|_{\{u=(b+a)/2\}}, z_b|_{\{u=(b+a)/2\}}) \leq \frac{2\max(\mu_{-\infty}, \mu)}{\nu}. \end{aligned} \tag{5.18}$$

The only point, where the latter property is used in the proof of Lemma 5.1, is (5.12) for the derivation of (5.13), (5.14). Therefore, we still obtain the estimates (5.7), (5.9), (5.10), (5.11), which yield a constant $C > 0$ with

$$\|s_u\|_t + \|s_y\|_Y + \|s_\lambda\|_\Sigma \leq C\|r\|_{W_t'} \quad \forall\, t \in [p, q]. \tag{5.19}$$

We derive now bounds for $\|s_a\|_t$ and $\|s_b\|_t$. Since (5.18) does not necessarily hold, we have to modify the proof of Lemma 5.1. Consider the subsets

$$\begin{aligned} \Omega_a = &\left(\{\tilde{u} > (b+a)/2\} \cup \left\{\tilde{u} = (b+a)/2,\ \tilde{z}_a \leq \frac{2\max(\mu_{-\infty}, \mu_1)}{\nu}\right\}\right) \\ &\cap \left(\{\bar{u} > (b+a)/2\} \cup \left\{\bar{u} = (b+a)/2,\ \bar{z}_a \leq \frac{2\max(\mu_{-\infty}, \mu_2)}{\nu}\right\}\right), \\ \Omega_b = &\left(\{\tilde{u} < (b+a)/2\} \cup \left\{\tilde{u} = (b+a)/2,\ \tilde{z}_a > \frac{2\max(\mu_{-\infty}, \mu_1)}{\nu}\right\}\right) \\ &\cap \left(\{\bar{u} < (b+a)/2\} \cup \left\{\bar{u} = (b+a)/2,\ \bar{z}_a > \frac{2\max(\mu_{-\infty}, \mu_2)}{\nu}\right\}\right). \end{aligned}$$

This yields by using $\mu_1 \leq \mu_2$

$$u_a|_{\Omega_a} \geq \frac{b-a}{2}, \quad z_a|_{\Omega_a} \leq \frac{1}{2}\left(\frac{2\max(\mu_{-\infty}, \mu_1)}{\nu} + \frac{2\max(\mu_{-\infty}, \mu_2)}{\nu}\right) \leq \frac{2\max(\mu_{-\infty}, \mu_2)}{\nu}$$

and similarly

$$u_b|_{\Omega_b} \geq \frac{b-a}{2}, \quad z_b|_{\Omega_b} \leq \frac{1}{2}\left(\frac{2\max(\mu_{-\infty},\mu_1)}{\nu} + \frac{2\max(\mu_{-\infty},\mu_2)}{\nu}\right) \leq \frac{2\max(\mu_{-\infty},\mu_2)}{\nu}.$$

Hence, we have with $\Omega' = \Omega_a \cup \Omega_b$

$$z_a|_{\Omega_a} \leq \frac{2\max(\mu_{-\infty},\mu_2)}{\nu}., \quad z_b|_{\Omega'\setminus\Omega_a} \leq \frac{2\max(\mu_{-\infty},\mu_2)}{\nu}.$$

Thus, (5.12) holds on Ω' instead of Ω with μ_2 instead of μ and we obtain exactly as in the proof of Lemma 5.1 the following analogs of (5.13), (5.14) on Ω'

$$\|s_a|_{\Omega'}\|_t \leq (1+2C_{\mu_2}\nu^{-1})(\|r_a\|_t + \|s_u\|_t) + C'(\|s_y\|_Y + \|s_u\|_t + \|s_\lambda\|_\Sigma + \|r_u\|_t), \tag{5.20}$$

$$\|s_b|_{\Omega'}\|_t \leq (1+2C_{\mu_2}\nu^{-1})(\|r_b\|_t + \|s_u\|_t) + C'(\|s_y\|_Y + \|s_u\|_t + \|s_\lambda\|_\Sigma + \|r_u\|_t). \tag{5.21}$$

It remains to estimate $\|s_a|_{\Omega''}\|_t$, $\|s_b|_{\Omega''}\|_t$ for the set

$$\Omega'' = \Omega \setminus \Omega'.$$

By the definition of Ω' we have

$$u_a|_{\Omega''} = \left.\frac{\tilde{u}_a + \bar{u}_a}{2}\right|_{\Omega''} \geq \frac{b-a}{4} \geq \frac{\nu}{4}, \quad u_b|_{\Omega''} = \left.\frac{\tilde{u}_b + \bar{u}_b}{2}\right|_{\Omega''} \geq \frac{b-a}{4} \geq \frac{\nu}{4}.$$

We now split Ω'' into the sets

$$\Omega''_a = \left\{|\hat{Z}_a s_u| \leq |r_a|\right\}, \quad \Omega''_b = \left\{|\hat{Z}_b s_u| \leq |r_b|\right\}, \quad \Omega''_r = \Omega'' \setminus (\Omega''_a \cup \Omega''_b).$$

This yields by using (5.2)

$$\begin{aligned}
|s_a|_{\Omega''_a}| &= |\hat{U}_a^{-1}(r_a - \hat{Z}_a s_u)|_{\Omega''_a}| \leq \left.\frac{2|u_a+z_a||r_a|}{u_a}\right|_{\Omega''_a} \leq 8\nu^{-1}|\hat{r}_a|_{\Omega''_a}|,\\
|s_b|_{\Omega''_b}| &= |\hat{U}_b^{-1}(r_b - \hat{Z}_b s_u)|_{\Omega''_b}| \leq \left.\frac{2|u_b+z_b||r_b|}{u_b}\right|_{\Omega''_b} \leq 8\nu^{-1}|\hat{r}_b|_{\Omega''_b}|,\\
s_a|_{\Omega''_b} &= (\ell_{uy}s_y + \ell_{uu}s_u + c_u^* s_\lambda + s_b - r_u)|_{\Omega''_b},\\
s_b|_{\Omega''_a} &= (-\ell_{uy}s_y - \ell_{uu}s_u - c_u^* s_\lambda + s_a + r_u)|_{\Omega''_a}.
\end{aligned}$$

Hence, we obtain on $\Omega''_a \cup \Omega''_b$.

$$\|s_a|_{\Omega''_a\cup\Omega''_b}\|_t \leq 8\nu^{-1}\|\hat{r}_a\|_t + C'(\|s_y\|_Y + \|s_u\|_t + \|s_\lambda\|_\Sigma + \|r_u\|_t), \tag{5.22}$$

$$\|s_b|_{\Omega''_a\cup\Omega''_b}\|_t \leq 8\nu^{-1}\|\hat{r}_b\|_t + C'(\|s_y\|_Y + \|s_u\|_t + \|s_\lambda\|_\Sigma + \|r_u\|_t). \tag{5.23}$$

Finally, we have on Ω''_r

$$s_a|_{\Omega''_r} = \hat{U}_a^{-1}(r_a - \hat{Z}_a s_u)|_{\Omega''_r}, \quad s_b|_{\Omega''_r} = \hat{U}_b^{-1}(r_b + \hat{Z}_b s_u)|_{\Omega''_r}$$

and thus by the definition of Ω''_r

$$\begin{aligned}
\operatorname{sgn}(s_a|_{\Omega''_r}) &= -\operatorname{sgn}(s_u|_{\Omega''_r}),\\
\operatorname{sgn}(s_b|_{\Omega''_r}) &= \operatorname{sgn}(s_u|_{\Omega''_r}).
\end{aligned}$$

Hence, the second line in (5.17) yields

$$|s_a|_{\Omega_r''}| + |s_b|_{\Omega_r''}| = |(s_b - s_a)|_{\Omega_r''}| = |r_u - \ell_{uy}s_y - \ell_{uu}s_u - c_u^* s_\lambda|_{\Omega_r''}.$$

Therefore, (5.20), (5.21) hold also on Ω_r'' and we have shown that

$$\begin{aligned}\|s_a\|_t &\le 8\nu^{-1}\|\hat{r}_a\|_t + 2(1+2C_{\mu_2}\nu^{-1})(\|r_a\|_t + \|s_u\|_t)\\ &\quad + 2C'(\|s_y\|_Y + \|s_u\|_t + \|s_\lambda\|_\Sigma + \|r_u\|_t),\\ \|s_b\|_t &\le 8\nu^{-1}\|\hat{r}_b\|_t + 2(1+2C_{\mu_2}\nu^{-1})(\|r_b\|_t + \|s_u\|_t)\\ &\quad + 2C'(\|s_y\|_Y + \|s_u\|_t + \|s_\lambda\|_\Sigma + \|r_u\|_t).\end{aligned}$$

Together with (5.19) we conclude that the solution of (5.17) satisfies for all $t \in [p,q]$

$$\|s\|_{W_t} \le C''(\|r\|_{W_t'} + \|\hat{r}\|_{W_t}),$$

where C'' can be chosen uniformly for all $\mu_1, \mu_2 \in (0, \mu_0]$ only depending on μ_0.

Since $\mu = \min(\mu_1, \mu_2)$, we obtain as at the end of the proof of Lemma 5.1

$$u_a + z_a \ge 2\sqrt{\gamma\mu}, \quad u_b + z_b \ge 2\sqrt{\gamma\mu}$$

and thus by the definition of r in (5.17)

$$\|r\|_{W_t'} \le \max\left(1, \frac{1}{2\sqrt{\gamma\mu}}\right)\|\hat{r}\|_{W_t'}.$$

Therefore, (5.16) is proven, where the constant C can be chosen uniformly on bounded subsets of $\left\{(\mu_1,\mu_2,w) \in (0,\infty)^2 \times \frac{1}{2}(N_{-\infty,q}(\mu_1) + N_{-\infty,q}(\mu_2))\right\}$ on which (5.3) holds uniformly. □

6. Hölder continuity of the central path. We will now state conditions under which the central path defines a Hölder continuous curve that converges for $\mu \searrow 0$ to a solution of (2.1).

The analysis of the central path is quite obvious if (A5)$_q$ holds for $q = \infty$ and more involved in the case $q < \infty$. This is caused by the fact that

$$(u,z) \in (\mathcal{B}, \|\cdot\|_q) \times L^q \mapsto uz \in L^q$$

is only differentiable in the case $q = \infty$. Otherwise we have to weaken the image space to achieve differentiability. More precisely we have the following result.

LEMMA 6.1. *Let Z be an open bounded set in L^∞. Then for any $p < q \le \infty$ the mapping*

$$u,z \in (Z, \|\cdot\|_q) \times L^q \mapsto uz \in L^p$$

is continuously differentiable and

$$\begin{aligned}\|(u+u')(z+z') - uz - (uz' + zu')\|_p &= \|u'z'\|_p \le \|(u')^p\|_{q/(q-p)}^{1/p}\|(z')^p\|_{q/p}^{1/p}\\ &= \|u'\|_{pq/(q-p)}\|z'\|_q.\end{aligned}$$

If $p/(q-p) > 1$ then

$$\|u'\|_{pq/(q-p)}\|z'\|_q \le \|u'\|_q^{(q-p)/p}\|u'\|_\infty^{1-(q-p)/p}\|z'\|_q.$$

Moreover,

$$u,z \in L^\infty \times L^\infty \mapsto uz \in L^\infty$$

is continuously differentiable and

$$\|(u+u')(z+z') - uz - (uz' + zu')\|_\infty = \|u'z'\|_\infty \le \|u'\|_\infty \|z'\|_\infty.$$

Proof. The proof is already obvious from the statement of the lemma, where Hölders inequality was used. □

We consider now first the case that $q = \infty$.

LEMMA 6.2. *Let (A1)–(A4) and (A5)$_q$ with $q = \infty$ hold. If $u \in \mathcal{B} \to J(y(u), u)$ is convex then for any $\mu > 0$ the central path $\mu \in (0, \infty) \to w(\mu) \in W_\infty$ according to (2.7) is well defined.*

If for $\mu_0 > 0$ the reduced Hessian satisfies

$$(v, \hat{H}(y(\mu), u(\mu), \lambda(\mu))v) \ge \alpha \|v\|_2^2 \quad \forall\, v \in L^2(\Omega), \;\; \forall\, \mu \in (0, \mu_0]$$

with some $\alpha > 0$ then the central path $\mu \in (0, \mu_0] \to w(\mu) \in W_\infty$ is continuously differentiable, satisfies

$$\|\dot{w}(\mu)\|_{W_\infty} \le \frac{C}{\sqrt{\mu}} \quad \forall\, \mu \in (0, \mu_0]$$

with a constant $C > 0$ and is thus Hölder-continuous with index $1/2$. More precisely, we have with $L = 2C$

$$\|w(\mu_1) - w(\mu_2)\|_{W_\infty} \le L|\sqrt{\mu_1} - \sqrt{\mu_2}| \le L\sqrt{|\mu_1 - \mu_2|} \quad \forall\, \mu_1, \mu_2 \in (0, \mu_0]. \tag{6.1}$$

Moreover, $\bar{w} = \lim_{\mu \searrow 0} w(\mu)$ exists in W_∞, $\bar{w} \in W_\infty$ satisfies the KKT-conditions (2.7) and $(\bar{y}, \bar{u})$ is global solution of (2.1).

Proof. Then the barrier problem (2.4) has the strictly convex reduced objective function $u \mapsto J_\mu(y(u), u)$, and thus the solution $(\bar{y}, \bar{u}) = (y(\bar{u}), \bar{u})$ provided by Proposition 2.3 is the unique solution of (2.4). Now also $\bar{\lambda}, \bar{z}_a, \bar{z}_b$ are uniquely determined by the first and the last two equations in (2.7). Thus, together with Corollary 4.3 the central path $\mu \in (0, \infty) \to w(\mu) \in W_\infty$ is well defined and bounded on bounded subsets $(0, \mu_0]$.

The mapping

$$F_\mu : W_\infty \to W'_\infty$$

is by (A1)–(A5)$_\infty$ and by Lemma 6.1 continuously differentiable. For $\mu > 0$ the primal-dual central path $\mu \mapsto w(\mu) := (y, u, \lambda, z_a, z_b)(\mu)$ given by (2.7) is the unique solution of

$$F_\mu((y, u, \lambda, z_a, z_b)(\mu)) = 0, \quad a \le u(\mu) \le b.$$

Since $w(\mu) \in N_{-\infty,\infty}(\mu)$, $DF_\mu(w(\mu)) \in \mathcal{L}(W_\infty, W'_\infty)$ has by Lemma 5.1 a bounded inverse. Thus, the implicit function theorem shows that $\mu \to w(\mu)$ is continuously differentiable and that the derivative w.r.t. μ satisfies

$$DF_\mu(w(\mu))\dot{w}(\mu) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}. \tag{6.2}$$

For fixed $\mu_0 > 0$ Lemma 5.1 yields a constant $C > 0$ with

$$\|DF_\mu(w(\mu))^{-1}\|_{W'_\infty, W_\infty} \le \frac{C}{\sqrt{\mu}} \quad \forall\, \mu \in (0, \mu_0].$$

Hence, we conclude that

$$\|\dot{w}(\mu)\|_{W_\infty} \le \frac{C}{\sqrt{\mu}} \quad \forall\, \mu \in (0, \mu_0].$$

But this gives for all $0 < \mu_1 < \mu_2 \le \mu_0$

$$\begin{aligned}\|w(\mu_2) - w(\mu_1)\|_{W_\infty} &\le \int_{\mu_1}^{\mu_2} \|\dot{w}(\mu)\|_{W_\infty}\, d\mu \le \int_{\mu_1}^{\mu_2} \frac{C}{\sqrt{\mu}}\, d\mu \\ &= 2C(\sqrt{\mu_2} - \sqrt{\mu_1}) = 2C\frac{\mu_2 - \mu_1}{\sqrt{\mu_2} + \sqrt{\mu_1}} \le 2C\sqrt{\mu_2 - \mu_1}.\end{aligned}$$

Thus, we have shown that $w(\cdot) \in C^{1/2}((0, \mu_0]; W_\infty)$ for any $\mu_0 > 0$. Hence, the central path is Hölder-continuous in W_∞ and admits a continuation until $\mu = 0$, i.e., $\bar{w} = \lim_{\mu \searrow 0} w(\mu)$ exists in W_∞. By continuity, $\bar{w}$ satisfies $F_0(\bar{w}) = 0$, which are just the KKT-conditions (2.3). Consequently, $(\bar{y}, \bar{u})$ is a global solution of (2.1), since the reduced objective functional and $\mathcal{B}$ are convex. □

For the general case we use the following auxiliary result.

LEMMA 6.3. *Let (A1)–(A4) and (A5)$_q$ with some $p < q < \infty$ hold. If $u \in \mathcal{B} \to J(y(u), u)$ is convex then for any $\mu > 0$ the central path $\mu \in (0, \infty) \to w(\mu) \in W_q$ according to (2.7) is well defined.*

If for $\mu_0 > 0$ the reduced Hessian satisfies

$$(v, \hat{H}(y(\mu), u(\mu), \lambda(\mu))v) \ge \alpha \|v\|_2^2 \quad \forall\, v \in L^2(\Omega), \ \ \forall\, \mu \in (0, \mu_0]$$

with some $\alpha > 0$ then the central path $\mu \in (0, \mu_0] \to w(\mu) \in W_q$ is continuous.

Proof. Then the barrier problem (2.4) has the strictly convex reduced objective function $u \mapsto J_\mu(y(u), u)$, and thus the solution $(\bar{y}, \bar{u}) = (y(\bar{u}), \bar{u})$ provided by Proposition 2.3 is the unique solution of (2.4). Now also $\bar{\lambda}, \bar{z}_a, \bar{z}_b$ are uniquely determined by the first and the last two equations in (2.7). Thus, together with Corollary 4.3 the central path $\mu \in (0, \infty) \to w(\mu) \in W_q$ is well defined and bounded on bounded subsets $(0, \mu_0]$.

We show next the continuity of the central path. Let $0 < \bar{\mu}$ be given and let $\bar{\mu} \le \mu < \eta$ be arbitrary. We write

$$J_\mu(y, u) = J(y, u) + \mu B(u)$$

with the log-barrier term B. We know that $B(u) \ge -c_B$ on $\mathcal{B}$ with some $c_B \ge 0$. Now we have for the solutions u_μ, u_η of the barrier problems (2.4)

$$\begin{aligned}J_\mu(y(u_\mu), u_\mu) &\le J_\mu(y(u_\eta), u_\eta) = J_\eta(y(u_\eta), u_\eta) + (\mu - \eta)B(u_\eta) \\ &\le J_\eta(y(u_\mu), u_\mu) + (\mu - \eta)B(u_\eta) \\ &= J_\mu(y(u_\mu), u_\mu) + (\mu - \eta)(B(u_\eta) - B(u_\mu)) \\ &\le J_\mu(y(u_\mu), u_\mu) + (\eta - \mu)(c_B + B(u_\mu)).\end{aligned}$$

The first and third row shows that

$$B(u_\mu) - B(u_\eta) \ge 0$$

and thus in particular $B(u_\mu) \le B(u_{\bar{\mu}})$. This yields

$$J_\mu(y(u_\mu), u_\mu) \le J_\mu(y(u_\eta), u_\eta) \le J_\mu(y(u_\mu), u_\mu) + (\eta - \mu)(c_B + B(u_{\bar{\mu}}))$$

and thus

$$|J_\mu(y(u_\eta), u_\eta) - J_\mu(y(u_\mu), u_\mu)| \le (\eta - \mu)|c_B + B(u_{\bar{\mu}})| \to 0 \quad \text{for } \eta - \mu \to 0, \ \ \bar{\mu} \le \mu < \eta.$$

As we have already observed, (A1)–(A4) imply that the reduced objective functional is twice continuously differentiable. Since the barrier terms are convex and $J_\mu'(y(u_\mu), u_\mu) = 0$, we have

$$J_\mu(y(u_\eta), u_\eta) - J_\mu(y(u_\mu), u_\mu) \geq (u_\eta - u_\mu, \hat{H}(u(\tau))(u_\eta - u_\mu)) \geq \alpha \|u_\eta - u_\mu\|_2^2$$

with $u(\tau) = u_\mu + \tau(u_\eta - u_\mu)$ and appropriate $\tau \in [0, 1]$. This yields

$$u_\eta - u_\mu \to 0 \text{ in } L^2 \text{ for } \eta \to \mu, \quad \eta, \mu \geq \bar{\mu}$$

and thus in all L^s, $s < \infty$ by interpolation with the uniform L^∞-bound. Now Lemma 4.6 yields

$$(\bar{u}_\eta, \bar{y}_\eta, \bar{\lambda}_\eta, \bar{z}_{a,\eta}, \bar{z}_{b,\eta}) \to (\bar{u}_\mu, \bar{y}_\mu, \bar{\lambda}_\mu, \bar{z}_{a,\mu}, \bar{z}_{b,\mu}) \quad \text{in } W_q \text{ for } \eta \to \mu, \quad \eta, \mu \geq \bar{\mu}.$$

□

LEMMA 6.4. *Let the assumptions of Lemma 6.3 hold. Then the central path $\mu \in (0, \infty) \to w(\mu) \in W_q$ is well defined. Moreover, $\mu \in (0, \mu_0] \to w(\mu) \in W_q$ is continuous, satisfies*

$$\limsup_{\eta \to \mu} \frac{\|w(\eta) - w(\mu)\|_{W_q}}{\eta - \mu} \leq \frac{C}{\sqrt{\mu}} \quad \forall\, \mu \in (0, \mu_0] \tag{6.3}$$

with a constant $C > 0$ and is thus Hölder-continuous with index $1/2$ in all spaces W_t, $t \in [p, q]$. More precisely, we have with $L = 2C$

$$\|w(\mu_1) - w(\mu_2)\|_{W_t} \leq L|\sqrt{\mu_1} - \sqrt{\mu_2}| \leq L\sqrt{|\mu_1 - \mu_2|} \quad \forall\, \mu_1, \mu_2 \in (0, \mu_0]. \tag{6.4}$$

Moreover, $\bar{w} = \lim_{\mu \searrow 0} w(\mu)$ exists in W_q, $\bar{w} \in W_q$ satisfies the KKT-conditions (2.7) and $(\bar{y}, \bar{u})$ is global solution of (2.1).

Proof. By assumptions (A1), (A5$_q$) the first three components of the mapping

$$F_\mu : W_q \to W_q'$$

are continuously differentiable. Moreover, we have

$$u_a(\eta) z_a(\eta) - u_a(\mu) z_a(\mu) = \frac{u_a(\eta) + u_a(\mu)}{2}(z_a(\eta) - z_a(\mu)) + \frac{z_a(\eta) + z_a(\mu)}{2}(u_a(\eta) - u_a(\mu)).$$

Therefore, we obtain with

$$\tilde{w}(\mu) := \left(y(\mu), \frac{u(\eta) + u(\mu)}{2}, \lambda(\mu), \frac{z_a(\eta) + z_a(\mu)}{2}, \frac{z_b(\eta) + z_b(\mu)}{2} \right)$$

the identity

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ \eta - \mu \\ \eta - \mu \end{pmatrix} = F_\mu(w(\eta)) - F_\mu(w(\mu)) \tag{6.5}$$

$$= DF_\mu(\tilde{w}(\mu))(w(\eta) - w(\mu)) + o_{W_q'}(\|w(\eta) - w(\mu)\|_{W_q}),$$

where

$$\frac{\|o_{W_q'}(\|w(\eta) - w(\mu)\|_{W_q})\|_{W_q'}}{\|w(\eta) - w(\mu)\|_{W_q}} \to 0 \quad \text{for } \eta \to \mu,$$

since $\|w(\eta)-w(\mu)\|_{W_q} \to 0$ for $\eta \to \mu$ by Lemma 6.3.

Since

$$\tilde{w}(\eta) \in \frac{1}{2}(N_{-\infty,q}(\mu)+N_{-\infty,q}(\eta)),$$

Lemma 5.2 yields

$$\|DF_\mu(\tilde{w}(\eta))^{-1}\|_{W_q',W_q} \le \frac{C}{\sqrt{\min(\mu,\eta)}}.$$

For $|\eta-\mu| \le \varepsilon$, $\varepsilon > 0$ small enough we have with the remainder term in (6.5) clearly

$$\left\| w(\eta)-w(\mu)+DF_\mu(\tilde{w}(\eta))^{-1} o_{W_q'}(\|w(\eta)-w(\mu)\|_{W_q}) \right\|_{W_q} \ge \frac{1}{2}\|w(\eta)-w(\mu)\|_{W_q}$$

for all $|\eta-\mu| \le \varepsilon$. We conclude with (6.5) that

$$\frac{\|w(\eta)-w(\mu)\|_{W_q}}{\eta-\mu} \le \frac{2C}{\sqrt{\min(\mu,\eta)}} \left\| \begin{pmatrix} 0\\0\\0\\1\\1 \end{pmatrix} \right\|_{W_q'} \le \frac{2CC'}{\sqrt{\min(\mu,\eta)}} \quad \text{for all } |\eta-\mu| \le \varepsilon.$$

This shows (6.3) and the Hölder continuity with index $1/2$ follows immediately. By writing the integral $\int_{\mu_1}^{\mu_2} \frac{C}{\sqrt{\mu}}\, d\mu$ as a limit of Riemann sums and by using (6.3), we see that again

$$\begin{aligned} \|w(\mu_1)-w(\mu_2)\|_{W_t} &\le \left| \int_{\mu_1}^{\mu_2} \frac{C}{\sqrt{\mu}}\, d\mu \right| = 2C|\sqrt{\mu_1}-\sqrt{\mu_2}| \\ &\le 2C\sqrt{|\mu_1-\mu_2|} \quad \forall\, \mu_1,\mu_2 \in (0,\mu_0]. \end{aligned}$$

The fact that $\bar{w} = \lim_{\mu \searrow 0} w(\mu)$ exists in W_q, satisfies (2.3) and is global solution of (2.1) follows now exactly as at the end of the proof of Lemma 6.2. □

7. A primal-dual interior point method. Let (A1)–(A5)$_q$ hold. The previous considerations show that for $w \in N_{-\infty,q}(\mu)$ the solution s of the primal-dual Newton system

$$DF_\mu(w)s = -F_\mu(w)$$

is only contained in W_q. Therefore, in the case $q < \infty$ we cannot ensure $w + \alpha s \in N_{-\infty,q}(\mu)$ by choosing an appropriate stepsize $\alpha \in (0,1]$. Instead, we use in addition a projection onto the neighborhood $N_{-\infty,q}(\mu)$.

DEFINITION 7.1. *We denote by P_μ a projection onto $N_{-\infty,q}(\mu)$ in W_q, i.e.*

$$P_\mu(w) \in N_{-\infty,q}(\mu), \quad \|P_\mu(w)-w\|_{W_q} = \min\left\{ \|\tilde{w}-w\|_{W_q} \ : \ \tilde{w} \in N_{-\infty,q}(\mu) \right\}.$$

If more than one projection point exists, P_μ selects one of them.

Obviously,

$$P_\mu(y,u,\lambda,z_a,z_b) = (y,*,\lambda,*,*),$$

i.e. P_μ does not change the y- and λ-component. Furthermore, the projection does not depend on q, since it reduces to a pointwise projection in $\mathbb{R}^3$ with respect to the Euclidean norm of the (u,z_a,z_b)-part.

REMARK 7.2. The form of our neighborhood $N_{-\infty,q}$ allows an easy computation of the projection P_μ. In fact, for almost all $\xi \in \Omega$ we have to project the point $(u(\xi), z_a(\xi), z_b(\xi)) \in \mathbb{R}^3$ onto the set

$$\begin{aligned} N_\xi = &\{(v, s_a, s_b) \ : \ v, s_a, s_b > 0, \ (v - a(\xi))s_a \geq \gamma\mu, \ (b(\xi) - v)s_b \geq \gamma\mu\} \\ &\cap\Big(([\frac{a(\xi)+b(\xi)}{2}, b(\xi)] \times [0, \frac{2\max(\mu_{-\infty}, \mu}{\nu}] \times [0,\infty)) \\ &\quad \cup ([a(\xi), \frac{a(\xi)+b(\xi)}{2}] \times [0,\infty) \times [0, \frac{2\max(\mu_{-\infty}, \mu}{\nu}]\Big). \end{aligned}$$

The first set is convex and it is easy to compute the projection onto it. The second set is the union of two cuboids of infinite length and again it is easy to project onto them. Moreover, if a point is not contained in the first set, one has only to project it onto it. Otherwise, if it is not contained in the second set, it is sufficient to project onto it.□

We show now that P_μ has a Lipschitz constant ≤ 2.

LEMMA 7.3. *Let (A5)$_q$ hold and let $w_\mu = (y_\mu, u_\mu, \lambda_\mu, z_{a,\mu}, z_{b,\mu}) \in W_q$ be a point on the central path. Then we have*

$$\|P_\mu(w) - w_\mu\|_{W_q} \leq 2\|w - w_\mu\|_{W_q} \quad \forall\, w \in W_q.$$

Proof. Since $w_\mu \in N_{-\infty,q}(\mu)$, we have

$$\|P_\mu(w) - w\|_{W_q} \leq \|w_\mu - w\|_{W_q}.$$

Hence,

$$\|P_\mu(w) - w_\mu\|_{W_q} \leq \|P_\mu(w) - w\|_{W_q} + \|w - w_\mu\|_{W_q} \leq 2\|w - w_\mu\|_{W_q}.$$

□

We consider now the following conceptual algorithm.

Algorithm PDPF: Projected Primal-Dual Interior-Point Method.

1. Choose $\nu \in (0,1)$, $C_0 > 0$, $0 < \sigma_{min} < 1$ and $\mu_0 > 0$. Select the constants $\gamma \in (0,1)$ and $\mu_{-\infty} > \mu_0$ for the neighborhood and choose $w_0 := (y_0, u_0, \lambda_0, z_{a,0}, z_{b,0}) \in N_{-\infty,q}(\mu_0)$ such that
$$\|F_{\mu_0}(w_0)\|_{V\times L^q\times\Lambda\times L^{q/2}(\Omega)^2} \leq C_0\sqrt[3]{\mu_0}.$$
Set $\bar\mu_0 = \mu_0$ and $k = 0$.
2. Solve the Newton-System
$$DF_{\mu_k}(w_k)\, s_k = -F_{\mu_k}(w_k)$$
and choose the maximal stepsize $\alpha_k \in 2^{-j}$, $j \in \mathbb{N}_0$, such that
$$\|F_{\mu_k}(P_{\mu_k}(w_k + \alpha_k s_k))\|_{V\times L^q\times\Lambda\times L^{q/2}(\Omega)^2} \leq C_0(1 - \nu(1 - \sqrt[3]{\sigma_k})\alpha_k)\sqrt[3]{\bar\mu_k}.$$
Set $w_{k+1} = P_{\mu_k}(w_k + \alpha_k s_k)$ and $\bar\mu_{k+1} = (1 - \nu(1 - \sqrt[3]{\sigma_k})\alpha_k)^3\bar\mu_k$.
3. If
$$\|F_{\mu_k}(w_{k+1})\|_{V\times L^q\times\Lambda\times L^{q/2}(\Omega)^2} \leq \frac{1}{2}C_0\sqrt[3]{\mu_k}$$
then choose the smallest $\sigma_k \in [\sigma_{min}, 1)$ with
$$\|F_{\sigma_k\mu_k}(w_{k+1})\|_{V\times L^q\times\Lambda\times L^{q/2}(\Omega)^2} \leq C_0\sqrt[3]{\mu_k}$$
and set $\mu_{k+1} = \sigma_k\mu_k$, $\bar\mu_{k+1} = \mu_k$.
Else set $\mu_{k+1} = \mu_k$

4. Set $k := k + 1$ and goto 2.

REMARK 7.4. This is only an exemplary globalization mechanism and is not the topic of this paper. We will see that under assumptions (A1)–(A5)$_\infty$ it accepts the choice $\alpha_k = 1$ and $\sigma_k = \sigma_{min}$ if σ_{min} is close enough to 1 and if w_0 is close enough to the central path.

We will analyze Algorithm PDPF under assumption (A5)$_\infty$. If merely (A5)$_q$ for $q < \infty$ holds, we will have to modify the algorithm by introducing a smoothing step, see Algorithm PDPFS in section 9. Appropriate implementations of Algorithm PDPFS will even yield superlinear convergence.

For $q < \infty$ the norm in Algorithms PDPF and PDPFS for measuring the residual is weaker than $\|\cdot\|_{W_q'}$. This is to ensure that it depends locally Lipschitz continuously on $w \in W_q$. One can show that $\|F_\mu(w)\|_{V\times L^q\times\Lambda\times L^{q/2}(\Omega)^2} \le C_1, \quad w \in N_{-\infty,q}(\mu)$ implies $\|F_\mu(w)\|_{W_q'} \le C_2$ and therefore $\|s_k\|_{W_q} \le CC_2/\sqrt{\mu_k}$. □

8. Global linear convergence for the L^∞-setting. We assume throughout this section that the assumptions (A1)–(A4) and (A5)$_\infty$ hold.

8.1. Quadratic local convergence towards the central path. We show first that the primal-dual iteration

$$DF_\mu(w)\, s = -F_\mu(w), \quad w_+ = w + s. \tag{8.1}$$

yields quadratic local convergence towards the central path.

LEMMA 8.1. *Let $\mu_0 > 0$ and $\rho_0 > 0$ be fixed. Assume that (A1)–(A4) and (A5)$_\infty$ hold and that*

$$\begin{aligned}(v, \hat{H}(y,u,\lambda)v) \ge \alpha\|v\|_2^2 \quad & \forall\, v \in L^2(\Omega),\\ & \forall\, w \in N_{-\infty,\infty}(\mu),\ \|w - w(\mu)\|_{W_\infty} \le \rho_0,\ \mu \in (0,\mu_0].\end{aligned} \tag{8.2}$$

Then there exists a constant $C > 0$ such that for any $0 < \mu \le \mu_0$ and for any $w \in N_{-\infty,\infty}(\mu)$ with $\|w - w(\mu)\|_\infty \le \rho_0$ the solution w_+ of the primal dual Newton step (8.1) satisfies

$$\|w_+ - w(\mu)\|_{W_\infty} \le \frac{C}{\sqrt{\mu}}\|w - w(\mu)\|_{W_\infty}^2.$$

For the projected iterate $P_\mu(w_+) \in N_{-\infty,\infty}(\mu)$ the estimate holds

$$\|P_\mu(w_+) - w(\mu)\|_{W_\infty} \le \frac{2C}{\sqrt{\mu}}\|w - w(\mu)\|_{W_\infty}^2.$$

and thus the projected iteration converges locally with quadratic rate.

Proof. We know by Corollary 4.3 that $w(\mu)$ is uniformly bounded in W_∞ for all $\mu \in (0, \mu_0]$. Thus, $\|w - w(\mu)\|_{W_\infty} \le \rho_0$ implies $\|w\|_{W_\infty} \le M$ for some constant $M > 0$. Hence, all μ, w in (8.2) are contained in a bounded subset of $\{(\mu, w) \in (0,\infty) \times N_{-\infty,\infty}(\mu)\}$ and Lemma 5.1 yields a constant $C > 0$ with

$$\|DF_\mu(w)^{-1}\|_{W_\infty', W_\infty} \le \frac{C}{\sqrt{\mu}}.$$

We have

$$DF_\mu(w)(w_+ - w) = -F_\mu(w), \quad DF_\mu(w)(w(\mu) - w(\mu)) = -F_\mu(w(\mu))$$

and thus

$$DF_\mu(w)(w_+ - w(\mu)) = F_\mu(w(\mu)) - F_\mu(w) - DF_\mu(w)(w(\mu) - w). \tag{8.3}$$

Since the first three components of $F_\mu : W_\infty \to W'_\infty$ are by (A1), (A5)$_q$ Lipschitz continuously differentiable on bounded subsets, this gives

$$DF_\mu(w)(w_+ - w(\mu)) = \begin{pmatrix} R_1(w(\mu)-w) \\ R_2(w(\mu)-w) \\ R_3(w(\mu)-w) \\ (u(\mu)-u)(z_a(\mu)-z_a) \\ (u(\mu)-u)(z_b(\mu)-z_b) \end{pmatrix},$$

where with a Lipschitz constant $L > 0$

$$\|R_1(w(\mu)-w)\|_V + \|R_2(w(\mu)-w)\|_\infty + \|R_3(w(\mu)-w)\|_\Lambda \le L\|w(\mu)-w\|_{W_\infty}^2.$$

Therefore,

$$\|w_+ - w(\mu)\|_{W_\infty} \le \|DF_\mu(w)^{-1}\|_{W'_\infty, W_\infty} \big(L\|w(\mu)-w\|_{W_\infty}^2 + \|(u(\mu)-u)(z_a(\mu)-z_a)\|_\infty + \|(u(\mu)-u)(z_b(\mu)-z_b)\|_\infty\big).$$

This yields

$$\|w_+ - w(\mu)\|_{W_\infty} \le \frac{C}{\sqrt{\mu}}(2+L)\|w-w(\mu)\|_{W_\infty}^2.$$

Finally, the estimate for $P_\mu(w_+)$ follows from Lemma 7.3. □

8.2. Global linear convergence of the interior point method. The previous result yields linear convergence for a short step method.

THEOREM 8.2. *Let $\mu_0 > 0$ and $\rho_0 > 0$ be fixed. Assume that (A1)–(A4) and (A5)$_\infty$ hold and that*

$$(v, \hat{H}(y,u,\lambda)v) \ge \alpha\|v\|_2^2 \quad \forall\, v \in L^2(\Omega), \quad \forall\, w \in N_{-\infty,\infty}(\mu),\ \|w-w(\mu)\|_{W_\infty} \le \rho_0,\ \mu \in (0,\mu_0]. \tag{8.2}$$

Then there are constants $\bar{\rho} \in (0, \rho_0]$ and $\bar{\sigma}_{min} \in (0,1)$ such that Algorithm PFPF has the following convergence property:

For any starting point $w \in N_{-\infty,\infty}(\mu_0)$ with $\|w - w(\mu_0)\|_{W_\infty} \le \bar{\rho}$, Algorithm PDPF with $\sigma_{min} \in (\bar{\sigma}_{min}, 1)$ chooses

$$\alpha_k = 1, \quad \mu_{k+1} = \sigma_k\mu_k = \sigma_{min}\mu_k$$

and generates a sequence with

$$\|w_k - w(\mu_k)\|_{W_\infty} \le C\sqrt{\mu_k} \tag{8.4}$$

$$\|w_k - \bar{w}\|_{W_\infty} \le (C+L)\sqrt{\mu_k} \tag{8.5}$$

$$\mu_k = \sigma_0 \cdots \sigma_{k-1}\mu_0 = \sigma_{min}^k \mu_0 \tag{8.6}$$

with constants $C, L > 0$. Here, $\bar{w} = \lim_{\mu\searrow 0} w(\mu)$ is the solution of (2.1).

Proof. Consider an arbitrary $\mu \in (0, \mu_0]$. Then there exists by Lemma 8.1 a constant $C > 0$ such that for any $w \in N_{-\infty,\infty}(\mu)$ with $\|w - w(\mu)\|_{W_\infty} \le \rho_0$ the estimate holds

$$\|P_\mu(w_+) - w(\mu)\|_{W_\infty} \le 2\|w_+ - w(\mu)\|_{W_\infty}^2 \le \frac{2C}{\sqrt{\mu}}\|w - w(\mu)\|_{W_\infty}^2,$$

where w_+ is the result of the primal-dual Newton step (8.1).

Now fix $\tau \in (0,1)$ such that

$$\bar{\rho} := \frac{\tau\sqrt{\mu_0}}{2C} \le \rho_0. \tag{8.7}$$

Then for any $w \in N_{-\infty,\infty}(\mu)$ with

$$\|w - w(\mu)\|_{W_\infty} \le \frac{\tau\sqrt{\mu}}{2C},$$

we have

$$\|P_\mu(w_+) - w(\mu)\|_{W_\infty} \le \frac{2C}{\sqrt{\mu}}\|w - w(\mu)\|^2_{W_\infty} \le \tau\|w - w(\mu)\|_{W_\infty} \le \frac{\tau^2\sqrt{\mu}}{2C}. \tag{8.8}$$

Moreover, we have with the Hölder constant L of the central path in (6.1) for $0 < \sigma < 1$

$$\|w(\mu) - w(\sigma\mu)\|_{W_\infty} \le L(1 - \sqrt{\sigma})\sqrt{\mu}.$$

and thus

$$\|P_\mu(w_+) - w(\sigma\mu)\|_{W_\infty} \le \frac{\tau^2\sqrt{\mu}}{2C} + L(1 - \sqrt{\sigma})\sqrt{\mu}.$$

Therefore, we can ensure that the new iterate satisfies

$$\|P_\mu(w_+) - w(\sigma\mu)\|_{W_\infty} \le \frac{\tau\sqrt{\sigma\mu}}{2C}, \tag{8.9}$$

if $\sigma \in (0,1)$ is chosen such that

$$\frac{\tau^2}{2C} + L(1 - \sqrt{\sigma}) \le \frac{\tau\sqrt{\sigma}}{2C}.$$

Since $\tau \in (0,1)$, this holds for $\sigma \in (0,1)$ sufficiently close to 1, more precisely for

$$1 > \sigma \ge \bar{\sigma}_{min} := \left(\frac{\tau^2 + 2LC}{\tau + 2LC}\right)^2. \tag{8.10}$$

Thus we obtain by induction: If $\bar{\rho}$ is chosen according to (8.7) and $\bar{\sigma}_{min}$ is given by (8.10) then Algorithm PDPF with $\sigma_{min} \in [\bar{\sigma}_{min}, 1)$ and with the fixed choices $\alpha_k = 1$ and $\mu_{k+1} = \sigma_{min}\mu_k$ generates a sequence $w_{k+1} = P_{\mu_k}(w_k + s_k)$ with (see (8.9))

$$\|w_{k+1} - w(\mu_{k+1})\|_{W_\infty} \le \frac{\tau\sqrt{\mu_{k+1}}}{2C} \le \bar{\rho}, \quad \mu_{k+1} = \sigma_{min}\,\mu_k \tag{8.11}$$

and (see (8.8))

$$\|w_{k+1} - w(\mu_k)\|_{W_\infty} \le \frac{\tau^2\sqrt{\mu_k}}{2C} \le \bar{\rho}. \tag{8.12}$$

With the solution $\bar{w} = \lim_{\mu \searrow 0} w(\mu)$ of (2.1) we have by (8.11) in addition

$$\|w_k - \bar{w}\|_{W_\infty} \le \|w_k - w(\mu_k)\|_{W_\infty} + \|w(\mu_k) - \bar{w}\|_{W_\infty} \le \frac{\tau\sqrt{\mu_k}}{2C} + L\sqrt{\mu_k}.$$

This proves (8.4), (8.5), (8.6).

We still have to show that after a possible reduction of $\rho > 0$ the globalization strategy of Algorithm PDPF admits the choice $\alpha_k = 1$ and $\mu_{k+1} = \sigma_{min}\mu_k$. To this purpose we observe that Algorithm PDPF chooses $\alpha_k = 1$, $\sigma_k = \sigma_{min}$, and $\mu_{k+1} = \sigma_{min}\mu_k$ if

$$\|F_{\mu_k}(P_{\mu_k}(w_k + s_k))\|_{V\times L^q\times\Lambda\times L^{q/2}(\Omega)^2} \le \frac{C_0}{2}\sqrt[3]{\mu_k}. \tag{8.13}$$

and if

$$\|F_{\sigma_{min}\mu_k}(P_{\mu_k}(w_k + s_k))\|_{V\times L^q\times\Lambda\times L^{q/2}(\Omega)^2} \le C_0\sqrt[3]{\mu_k}. \tag{8.14}$$

But the operators $F_{\mu_k} : W_\infty \to V \times L^q \times \Lambda \times L^{q/2}(\Omega)^2 = W'_\infty$ are obviously Lipschitz-continuous on the bounded set

$$\big\{w \in N_{-\infty,\infty}(\mu) \;:\; \|w - w(\mu)\|_{W_\infty} \le \rho, \quad 0 < \mu \le \mu_0\big\}$$

with some Lipschitz constant L_F. Hence, possibly after reducing $\tau > 0$ (and thus $\bar\rho > 0$), (8.13) follows from (8.8) and (8.14) follows from (8.9). □

9. Global linear and superlinear local convergence for the general L^q-setting. If $(A5)_q$ holds only for some $p < q < \infty$ the convergence analysis is more delicate. Under a strict complementarity assumption and by using an additional smoothing step we will prove global linear convergence in the general L^q-setting. Moreover, we will also show that superlinear local convergence is achieved if μ_k is reduced fast enough.

We refine our analysis as follows. Since under the assumptions of Lemma 5.1

$$\|DF_\mu(w(\mu))^{-1}\|_{W'_t,W_t} = O(\mu^{-1/2}), \text{ but } \|(S(w)DF_\mu(w))^{-1}\|_{W'_t,W_t} \le C, \quad t \in [p,q],$$

with the scaling operator

$$S(w) = \begin{pmatrix} I & & & & \\ & I & & & \\ & & I & & \\ & & & (U_a + Z_a)^{-1} & \\ & & & & (U_b + Z_b)^{-1} \end{pmatrix}, \tag{5.1}$$

we use now for the analysis of the primal-dual Newton iteration instead of (8.3) the scaled equation

$$S(w)DF_\mu(w)(w_+ - w(\mu)) = S(w)\left(F_\mu(w(\mu)) - F_\mu(w) - DF_\mu(w)(w(\mu) - w)\right). \tag{9.1}$$

Then we use a two-norm technique based on the $L^p - L^q$ norm gap to estimate the right hand side. Independently of the size of $\mu > 0$ this will yield an estimate of the form

$$\|w_+ - w(\mu)\|_{W_p} = o(\|w - w(\mu)\|_{W_q}).$$

The norm gap will then be closed be using a smoothing step. We recall that with the notations

$$\hat U_a = (U_a + Z_a)^{-1}U_a, \quad \hat U_b = (U_b + Z_b)^{-1}U_b,$$
$$\hat Z_a = (U_a + Z_a)^{-1}Z_a, \quad \hat Z_b = (U_b + Z_b)^{-1}Z_b,$$

we have $\hat U_a + \hat Z_a = I$, $\hat U_b + \hat Z_b = I$ and

$$S(w)DF_\mu(w) = \begin{pmatrix} \ell_{yy} & \ell_{yu} & c_y^* & 0 & 0 \\ \ell_{uy} & \ell_{uu} & c_u^* & -I & I \\ c_y & c_u & 0 & 0 & 0 \\ 0 & \hat Z_a & 0 & \hat U_a & 0 \\ 0 & -\hat Z_b & 0 & 0 & \hat U_b \end{pmatrix}$$

where we omit the arguments.

9.1. Refined analysis of the primal-dual Newton step. We first prove a similar result as in Lemma 8.1 where we avoid the μ-dependent convergence factor by using a two-norm technique. We will need a strict complementarity assumption.

DEFINITION 9.1. *Let $\bar{w} = (\bar{y}, \bar{u}, \bar{\lambda}, \bar{z}_a, \bar{z}_b) \in Y \times U \times \Lambda^* \times U^* \times U^*$ satisfy the KKT-conditions (2.3). Then* strict complementarity *holds at $\bar{w}$ if*

$$\text{leb}\left(\{\min(\bar{u}_a + \bar{z}_a, \bar{u}_b + \bar{z}_b) = 0\}\right) = 0.$$

where leb () *is the Lebesgue measure on Ω.*

We define the function

$$\omega(t) = \text{leb}\left(\{\min(\bar{u}_a + \bar{z}_a, \bar{u}_b + \bar{z}_b) \le t\}\right). \tag{9.2}$$

Under a strict complementarity assumption we then have

$$\lim_{t \searrow 0} \omega(t) = 0. \tag{9.3}$$

If $\omega(t) = O(t^\kappa)$ as $t \searrow 0$, we say that strong strict complementarity holds.

DEFINITION 9.2. *Let $\bar{w} = (\bar{y}, \bar{u}, \bar{\lambda}, \bar{z}_a, \bar{z}_b) \in Y \times U \times \Lambda^* \times U^* \times U^*$ satisfy the KKT-conditions (2.3). Then* strong strict complementarity *holds at $\bar{w}$ if there exist constants $C_c > 0$, $\kappa > 0$ such that*

$$\omega(t) \le C_c t^\kappa \quad \forall\, t \ge 0. \tag{9.4}$$

We start with the following technical result.

LEMMA 9.3. *Let the assumptions of Lemma 6.4 hold and let $\bar{w} = \lim_{\mu \searrow 0} w(\mu)$ be the solution of (2.3). Define the function*

$$\omega_\mu(t) = \text{leb}\left(\{\min(u_a(\mu) + z_a(\mu), u_b(\mu) + z_b(\mu)) \le t\}\right). \tag{9.5}$$

Then there exists a constant $C > 0$ with

$$\omega_\mu(t) \begin{cases} = 0 & \text{for } t < 2\sqrt{\mu}, \\ \le \omega(2\max(t, t^{\frac{q}{q+1}})) + C\max(t, t^{\frac{q}{q+1}}) & \text{for } t \ge 2\sqrt{\mu} \end{cases}$$

with ω according to (9.2).

Proof. Since $u_a(\mu) z_a(\mu) = \mu$, we have $u_a(\mu) + z_a(\mu) \ge 2\sqrt{\mu}$ and similarly $u_b(\mu) + z_b(\mu) \ge 2\sqrt{\mu}$. This shows that $\omega_\mu(t) = 0$ for $t < 2\sqrt{\mu}$. For brevity, we set now

$$u_a = u_a(\mu), \quad u_b = u_b(\mu), \quad z_a = z_a(\mu), \quad z_b = z_b(\mu).$$

Now let $t \ge 2\sqrt{\mu}$. Then we have for any $s \ge t$

$$\begin{aligned}
\omega_\mu(t) &\le \text{leb}\left(\{\min(u_a + z_a, u_b + z_b) \le s\}\right) \\
&\le \text{leb}\left(\{\min(\bar{u}_a + \bar{z}_a, \bar{u}_b + \bar{z}_b) - \max(|u_a - \bar{u}_a| + |z_a - \bar{z}_a|, |u_b - \bar{u}_b| + |z_b - \bar{z}_b|) \le s\}\right) \\
&\le \omega(2s) + \text{leb}\left(\{\max(|u_a - \bar{u}_a| + |z_a - \bar{z}_a|, |u_b - \bar{u}_b| + |z_b - \bar{z}_b|) \ge s\}\right) \\
&\le \omega(2s) + \text{leb}\left(\{|u_a - \bar{u}_a| \ge s/2\}\right) + \text{leb}\left(\{|z_a - \bar{z}_a| \ge s/2\}\right) \\
&\quad + \text{leb}\left(\{|u_b - \bar{u}_b| \ge s/2\}\right) + \text{leb}\left(\{|z_b - \bar{z}_b| \ge s/2\}\right) \\
&\le \omega(2s) + (s/2)^{-q}(\|u_a - \bar{u}_a\|_q^q + \|z_a - \bar{z}_a\|_q^q + \|u_b - \bar{u}_b\|_q^q + \|z_b - \bar{z}_b\|_q^q).
\end{aligned}$$

By Lemma 6.4 there exists a Hölder constant $L > 0$ with $\|w - w(\mu)\|_{W_q} \le L\sqrt{\mu}$. This yields

$$\omega_\mu(t) \le \omega(2s) + 4(s/2)^{-q} L^q \mu^{q/2}.$$

Now let $t \geq 2\sqrt{\mu}$. If $t \geq 1$ then we obtain with the choice $s = t$

$$\omega_\mu(t) \leq \omega(2t) + 4(t/2)^{-q} L^q \mu^{q/2} \leq \omega(2t) + 4L^q \leq \omega(2t) + 4L^q t.$$

If $t \leq 1$ then the choice $s = t^{q/(q+1)} = t^{1-1/(q+1)}$ yields

$$\omega_\mu(t) \leq \omega(2t^{q/(q+1)}) + 4t^{-q+q/(q+1)}(2L)^q \mu^{q/2} \leq \omega(2t^{q/(q+1)}) + 4L^q t^{q/(q+1)}.$$

□

We show now that for $\|w - w(\mu)\|_{W_q}$ small enough the result w_+ of the primal-dual Newton step satisfies $\|w_+ - w(\mu)\|_{W_p} = o(\|w - w(\mu)\|_{W_q})$, where the estimate is uniform in $\mu \in (0, \mu_0]$. Thus, in contrast to Lemma 8.1 we avoid the μ-dependent convergence factor but obtain a $W_p - W_q$ norm gap. This norm gap will be closed by using a smoothing step.

LEMMA 9.4. *Let $\mu_0 > 0$ and $\rho_0 > 0$ be fixed. Assume that (A1)–(A4) and (A5)$_q$ hold with some $q \in (p, \infty)$ and that*

$$(9.6) \qquad \begin{aligned} &(v, \hat{H}(y,u,\lambda)v) \geq \alpha \|v\|_2^2 \quad \forall\, v \in L^2(\Omega), \\ &\qquad\qquad \forall\, w \in N_{-\infty,q}(\mu),\ \|w - w(\mu)\|_{W_q} \leq \rho_0,\ \mu \in (0, \mu_0]. \end{aligned}$$

If in addition $\bar{w} = \lim_{\mu \searrow 0} w(\mu)$ satisfies strict complementarity then there exists a constant $C > 0$ such that for any $0 < \mu \leq \mu_0$ and for any $w \in N_{-\infty,q}(\mu)$ with $\|w - w(\mu)\|_{W_q} \leq \rho_0$ the solution w_+ of the primal dual Newton step (8.1) satisfies

$$(9.7) \qquad \begin{aligned} \|P_\mu(w_+) - w(\mu)\|_{W_p} &\leq 2\|w_+ - w(\mu)\|_{W_p} \\ &\leq 2C \left(\omega(4\|w - w(\mu)\|_{W_q}^{\eta q'})^{1/q'} + \|w - w(\mu)\|_{W_q}^{\eta} \right) \|w - w(\mu)\|_{W_q} \\ &= o(\|w - w(\mu)\|_{W_q}) \end{aligned}$$

with

$$(9.8) \qquad \eta = \frac{(q-p)\min(p, q-p)}{pq(p+1)}, \quad q' = \frac{qp}{q-p}.$$

Moreover, if $\bar{w}$ satisfies strong strict complementarity (9.4) then

$$(9.9) \qquad \|P_\mu(w_+) - w(\mu)\|_{W_p} \leq 2\|w_+ - w(\mu)\|_{W_p} \leq 2C\|w - w(\mu)\|_{W_q}^{1+\eta}$$

with

$$(9.10) \qquad \eta = \frac{\min(1,\kappa)(q-p)\min(p, q-p)}{p^2(q+1) + \min(1,\kappa)p(q-p)}.$$

Proof. We know by Corollary 4.3 that $w(\mu)$ is uniformly bounded in W_q for all $\mu \in (0, \mu_0]$. Thus, $\|w - w(\mu)\|_{W_q} \leq \rho_0$ implies $\|w\|_{W_q} \leq M$ for some constant $M > 0$. Hence, all μ, w in (8.2) are contained in a bounded subset of $\{(\mu, w) \in (0, \infty) \times N_{-\infty,q}(\mu)\}$ and Lemma 5.1 yields a constant $C > 0$ with

$$\|(S(w)DF_\mu(w))^{-1}\|_{W_t', W_t} \leq C, \quad p \leq t \leq q,$$

where $S(w)$ is the scaling operator (5.1). We have

$$DF_\mu(w)(w_+ - w) = -F_\mu(w), \quad DF_\mu(w)(w(\mu) - w(\mu)) = -F_\mu(w(\mu))$$

and thus

$$DF_\mu(w)(w_+ - w(\mu)) = F_\mu(w(\mu)) - F_\mu(w) - DF_\mu(w)(w(\mu) - w). \tag{9.11}$$

Since the first three components of $F : W_q \to W_p'$ are by (A1), (A5)$_q$ Lipschitz continuously differentiable on bounded subsets, this gives

$$DF_\mu(w)(w_+ - w(\mu)) = \begin{pmatrix} R_1(w(\mu)-w) \\ R_2(w(\mu)-w) \\ R_3(w(\mu)-w) \\ (u(\mu)-u)(z_a(\mu)-z_a) \\ (u(\mu)-u)(z_b(\mu)-z_b) \end{pmatrix},$$

where with a Lipschitz constant $L > 0$

$$\|R_1(w(\mu)-w)\|_\Sigma + \|R_2(w(\mu)-w)\|_p + \|R_3(w(\mu)-w)\|_\Lambda \le L\|w(\mu)-w\|_{W_q}^2.$$

To obtain an operator with uniformly bounded inverse on the left hand side we multiply with the scaling operator $S(w)$ in (5.1) and obtain

$$S(w)DF_\mu(w)(w_+ - w(\mu)) = \begin{pmatrix} R_1(w(\mu)-w) \\ R_2(w(\mu)-w) \\ R_3(w(\mu)-w) \\ \frac{(u(\mu)-u)(z_a(\mu)-z_a)}{u_a+z_a} \\ \frac{(u(\mu)-u)(z_b(\mu)-z_b)}{u_b+z_b} \end{pmatrix}.$$

Therefore,

$$\|w_+ - w(\mu)\|_{W_p} \le \|(S(w)DF_\mu(w))^{-1}\|_{W_p',W_p}\Bigg(L\|w(\mu)-w\|_{W_q}^2 + \left\|\frac{(u(\mu)-u)(z_a(\mu)-z_a)}{u_a+z_a}\right\|_p + \left\|\frac{(u(\mu)-u)(z_b(\mu)-z_b)}{u_b+z_b}\right\|_p\Bigg).$$

This yields

$$\|w_+ - w(\mu)\|_{W_p} \le C(L\|w-w(\mu)\|_{W_q}^2 + \|R_a\|_p + \|R_b\|_p) \tag{9.12}$$

with

$$R_a := \frac{(u(\mu)-u)(z_a(\mu)-z_a)}{u_a+z_a}, \quad R_b := \frac{(u(\mu)-u)(z_b(\mu)-z_b)}{u_b+z_b}.$$

It remains to estimate $\|R_a\|_p + \|R_b\|_p$. We show first that for $w \in N_{-\infty,q}(\mu)$

$$|R_a| \le \left(1 + \frac{1}{2\sqrt{\gamma}}\right)\max(|u-u(\mu)|, |z_a - z_a(\mu)|),$$
$$|R_b| \le \left(1 + \frac{1}{2\sqrt{\gamma}}\right)\max(|u-u(\mu)|, |z_b - z_b(\mu)|).$$

To this end we note that $(u(\mu)-a)z_a(\mu) = \mu$, $(u-a)z_a \ge \gamma\mu$. This yields with $u_a, z_a \ge 0$

$$u_a + z_a \ge 2\sqrt{u_a z_a} \ge 2\sqrt{\gamma\mu}.$$

Now we have

$$R_a = \left(\frac{-u_a}{u_a + z_a} + \frac{u_a(\mu)}{u_a + z_a}\right)(z_a(\mu) - z_a) = \left(\frac{-z_a}{u_a + z_a} + \frac{z_a(\mu)}{u_a + z_a}\right)(u_a(\mu) - u_a).$$

Since $u_a(\mu)z_a(\mu) = \mu$ we have $\min(u_a(\mu), z_a(\mu)) \leq \sqrt{\mu}$. On $\{u_a(\mu) \leq \sqrt{\mu}\}$ we have

$$|R_a| \leq \left(1 + \frac{1}{2\sqrt{\gamma}}\right)|z_a - z_a(\mu)|$$

and on $\{z_a(\mu) \leq \sqrt{\mu}\}$ we have

$$|R_a| \leq \left(1 + \frac{1}{2\sqrt{\gamma}}\right)|u_a - u_a(\mu)|.$$

The estimate for $|R_b|$ is obtained in the same way.

To estimate $\|R_a\|_p + \|R_b\|_p$ we split Ω for an arbitrary $\beta \in (0, \min(1, (q-p)/p))$ into the sets

$$J = \{u_a + z_a \geq \|w - \bar{w}\|_{W_q}^{\beta}\}, \quad J^c = \Omega \setminus J.$$

We have with ω_μ in (9.5)

$$\begin{aligned}
\text{leb}\,(\{u_a + z_a \leq t\}) &\leq \text{leb}\,(\{u_a(\mu) + z_a(\mu) - |u_a - u_a(\mu)| - |z_a - z_a(\mu)| \leq t\}) \\
&\leq \omega_\mu(2t) + \text{leb}\,(\{|u_a - u_a(\mu)| + |z_a - z_a(\mu)| \geq t\}) \\
&\leq \omega_\mu(2t) + \text{leb}\,(\{|u_a - u_a(\mu)| \geq t/2\}) + \text{leb}\,(|z_a - z_a(\mu)| \geq t/2\}) \\
&\leq \omega_\mu(2t) + (t/2)^{-q}(\|u_a - u_a(\mu)\|_q^q + \|z_a - z_a(\mu)\|_q^q) \\
&\leq \omega_\mu(2t) + 2(t/2)^{-q}\|w - w(\mu)\|_{W_q}^q
\end{aligned}$$

This yields the upper bound for the measure of J^c

$$\text{leb}\,(J^c) \leq \omega_\mu(2\|w - \bar{w}\|_{W_q}^{\beta}) + 2^{q+1}\|w - w(\mu)\|_{W_q}^{(1-\beta)q}.$$

Using $\|uv\|_p \leq \|u\|_{q'}\|v\|_q$, $q' = pq/(q-p)$, we have

$$\|R_a\|_{p,J} \leq \|u - u(\mu)\|_{q',J}\|z - z(\mu)\|_{q,J}^{1-\beta}.$$

If $q' \leq q$ this yields

$$\|R_a\|_{p,J} \leq c_{q,q'}\|u - u(\mu)\|_{q,J}\|z - z(\mu)\|_{q,J}^{1-\beta}$$

otherwise $\|v\|_{q'} \leq \|v\|_q^{q/q'}\|b - a\|_\infty^{1-q/q'}$ and thus

$$\|R_a\|_{p,J} \leq \|b - a\|_\infty^{1-(q-p)/p}\|u - u(\mu)\|_{q,J}^{(q-p)/p}\|z - z(\mu)\|_{q,J}^{1-\beta}.$$

Combining both cases we obtain

$$\|R_a\|_{p,J} \leq C_1\|w - w(\mu)\|_{W_q}^{1+\min(1,(q-p)/p)-\beta}.$$

On the complement set J^c we obtain by using that $\|1\|_{q',J^c} = \text{leb}\,(J^c)^{1/q'}$

$$\begin{aligned}
\|R_a\|_{p,J^c} &\leq \left(1 + \frac{1}{2\sqrt{\gamma}}\right)\|1\|_{q',J^c}\left(\|u - u(\mu)\|_{q,J^c} + \|z - z(\mu)\|_{q,J^c}\right) \\
&\leq \left(2 + \frac{1}{\sqrt{\gamma}}\right)\text{leb}\,(J^c)^{\frac{1}{q'}}\,\|w - w(\mu)\|_{W_q} \\
&\leq \left(2 + \frac{1}{\sqrt{\gamma}}\right)\left(\omega_\mu(2\|w - w(\mu)\|_{W_q}^{\beta}) + 2^{q+1}\|w - w(\mu)\|_{W_q}^{(1-\beta)q}\right)^{\frac{1}{q'}}\|w - w(\mu)\|_{W_q} \\
&\leq \left(2 + \frac{1}{\sqrt{\gamma}}\right)\left(\omega_\mu(2\|w - w(\mu)\|_{W_q}^{\beta})^{\frac{1}{q'}} + 2^{\frac{q+1}{q'}}\|w - w(\mu)\|_{W_q}^{\frac{(1-\beta)q}{q'}}\right)\|w - w(\mu)\|_{W_q}.
\end{aligned}$$

Using Lemma 9.3 and $q/q' = (q-p)/p$, we finally obtain constants $C_2, C_3 > 0$ with

$$\begin{aligned}\|R_a\|_p &\le C_1 \|w - w(\mu)\|_{W_q}^{\min(1,(q-p)/p)-\beta} \|w - w(\mu)\|_{W_q} \\ &\quad + C_2 \left(\omega(4\|w - w(\mu)\|_{W_q}^{\beta q/(q+1)})^{(q-p)/(qp)} + \|w - w(\mu)\|_{W_q}^{\beta(q-p)/(p(q+1))} \right) \|w - w(\mu)\|_{W_q} \\ &\quad + C_3 \|w - w(\mu)\|_{W_q}^{(1-\beta)(q-p)/p} \|w - w(\mu)\|_{W_q}.\end{aligned}$$

The last term has at least the order of the first term. Balancing the orders of the first and second term leads to

$$\beta = \frac{(q+1)\min(p, q-p)}{q(p+1)}$$

and results in the estimate

$$\|R_a\|_p \le C' \left(\omega(4\|w - w(\mu)\|_{W_q}^{\eta q'})^{1/q'} + \|w - w(\mu)\|_{W_q}^{\eta} \right) \|w - w(\mu)\|_{W_q}$$

with

$$\eta = \frac{(q-p)\min(p, q-p)}{pq(p+1)}, \quad q' = \frac{qp}{q-p}.$$

The same estimate is valid for $\|R_b\|_p$. Inserting this in (9.12) yields (9.7).

If in addition strong strict complementarity (9.4) holds, i.e., $\omega(t) \le C_c t^\kappa$, then the middle term has order $O(\|w - w(\mu)\|_{W_q}^{\min(1,\kappa)\beta(q-p)/(p(q+1))})$ and balancing with the first term gives

$$\beta = \frac{(q+1)\min(p, q-p)}{p(q+1) + \min(1,\kappa)(q-p)}.$$

Inserting this choice of β leads to the asserted estimate (9.9). □

We see that a norm gap occurs in the estimates (9.7), (9.9). To close the norm gap we will use a smoothing step.

9.2. Smoothing steps. We construct now an operator

$$Q_\mu : (y, u, \lambda, z_a, z_b) \in W_p \mapsto (y, \tilde{u}, \lambda, \tilde{z}_a, \tilde{z}_b) \in W_q$$

such that there exists a constant $L_S > 0$ with

$$\|Q_\mu(w) - w(\mu)\|_{W_q} \le L_S \|w - w(\mu)\|_{W_p} \tag{9.13}$$

to close the norm gap in (9.7). Let (A1)-(A5)$_q$ and the assumptions of Lemma 6.4 hold. Then $w(\mu)$ is the unique solution of (2.7) and satisfies with the notations of (A5)$_q$, 5. in particular

$$\begin{aligned}0 = \ell_u(w(\mu)) &= \beta(u(\mu)) + \hat{g}_s(y(\mu), u(\mu), \lambda(\mu)) + z_b(\mu) - z_a(\mu) \\ &= \beta(u(\mu)) + \frac{\mu}{b - u(\mu)} - \frac{\mu}{u(\mu) - a} + \hat{g}_s(y(\mu), u(\mu), \lambda(\mu)) \\ &=: \beta_\mu(u(\mu)) + \hat{g}_s(y(\mu), u(\mu), \lambda(\mu)),\end{aligned}$$

where $\beta \in C^1(\mathbb{R}),\ \beta' \ge \alpha_0 > 0$ and where

$$\hat{g}_s : (y, u, \lambda) \in Y \times U \times \Sigma \mapsto J_u(y, u) - \beta(u) + c_u(y, u)^* \lambda \in L^q(\Omega)$$

is by (A5)$_q$, 5. well defined and Lipschitz on bounded sets. Since the mappings

$$\beta_{\mu;\xi} : t \in (a(\xi), b(\xi)) \mapsto \beta(t) + \frac{\mu}{b(\xi) - t} - \frac{\mu}{t - a(\xi)}$$

satisfy $\beta'_{\mu;\xi} \geq \alpha_0 > 0$ and $\beta_{\mu;\xi}((a(\xi), b(\xi))) = \mathbb{R}$, the inverse mappings $\beta_{\mu;\xi}^{-1} : \mathbb{R} \to (a(\xi), b(\xi))$ exist and are Lipschitz continuous with Lipschitz constant $\leq 1/\alpha_0$. Thus, also the mapping

$$\beta_\mu : u \in \{v \ : \ a < v < b\} \mapsto \beta_\mu(u) = \beta(u) + \frac{\mu}{b-u} - \frac{\mu}{u-a}$$

has a Lipschitz continuous inverse

$$\beta_\mu^{-1} : L^q(\Omega) \to (\{v \ : \ a < v < b\}, \|\cdot\|_q)$$

and we have

$$u(\mu) = \beta_\mu^{-1}(-\hat{g}_s(y(\mu), u(\mu), \lambda(\mu))).$$

Thus, given any $w \in W_q$ the “smoothed” control

$$u_+ := \beta_\mu^{-1}(-\hat{g}_s(y, u, \lambda)) \tag{9.14}$$

satisfies

$$\|u_+ - u(\mu)\|_q \leq \frac{1}{\alpha_0}\|\hat{g}_s(y, u, \lambda) - \hat{g}_s(y(\mu), u(\mu), \lambda(\mu))\|_q \leq \frac{L_g}{\alpha_0}\|w - w(\mu)\|_{W_p} \tag{9.15}$$

with a Lipschitz constant L_g of $\hat{g}_s$ according to (A5)$_q$, 5..

Smoothing of the z_a, z_b-components can now be obtained by using the identities

$$z_a(\mu) = \frac{\mu}{u(\mu) - a}, \quad z_b(\mu) = \frac{\mu}{b - u(\mu)}, \tag{9.16}$$

In fact, after computing u_+ we choose $z_{a,+}$, $z_{b,+}$ according to

$$z_{a,+} = \frac{\mu}{u_+ - a}, \qquad z_{b,+} = \frac{\mu}{b - u_+}. \tag{9.17}$$

We will see that this leads in fact to an operator that has the smoothing property (9.13).

DEFINITION 9.5. *The smoothing operator $Q_\mu : W_p \to W_q$ is defined by*

$$Q_\mu(y, u, \lambda, z_a, z_b) = (y, u_+, \lambda, z_{a,+}, z_{b,+})$$

with u_+ according to (9.14) and $z_{a,+}, z_{b,+}$ according to (9.17), where $\hat{g}_s = (y, u, \lambda) = J_u(y, u) - \beta(u) + c_u(y, u)^*\lambda$ according to (A5)$_q$, 5.

THEOREM 9.6. *Let the assumptions of Lemma 6.3 hold. Then for any $\mu_0 > 0$, $\rho_0 > 0$ there is a constant $L_S > 0$ such that the smoothing operator $Q_\mu : W_p \to W_q$ of Definition 9.5 is well defined and satisfies*

$$\|Q_\mu(w) - w(\mu)\|_{W_q} \leq L_S\|w - w(\mu)\|_{W_p} \quad \forall\, w \in W_p,\ \|w - w(\mu)\|_{W_p} \leq \rho_0, \quad \forall\, \mu \in (0, \mu_0]. \tag{9.18}$$

Moreover, $(y, u_+, \lambda, z_{a,+}, z_{b,+}) = Q_\mu(w)$ satisfies

$$a < u_+ < b, \quad (u_+ - a)z_{a,+} = \mu, \quad (b - u_+)z_{b,+} = \mu$$

and thus $Q_\mu(w) \in N_{-\infty,q}(\mu)$.

Proof. We have already shown that (9.15) holds, where L_g is the Lipschitz constant of $\hat{g}_s$ on the bounded set of all $w \in W_p$ considered in (9.18). Moreover, we have by (9.17) and the choice (9.14) of u_+

$$z_{b,+} - z_{a,+} = -\beta(u_+) - \hat{g}_s(y, u, \lambda). \tag{9.19}$$

On the other hand

$$z_b(\mu) - z_a(\mu) = -\beta(u(\mu)) - \hat{g}_s(y(\mu), u(\mu), \lambda(\mu)). \tag{9.20}$$

Now consider the following partition of Ω

$$\begin{aligned} \Omega_a &= \{u_+ \geq (b+a)/2,\ u(\mu) \geq (b+a)/2\}, \quad \Omega_b = \{u_+ < (b+a)/2,\ u(\mu) < (b+a)/2\}, \\ \Omega_a' &= \{u_+ \geq (b+a)/2,\ u(\mu) < (b+a)/2\}, \quad \Omega_b' = \{u_+ < (b+a)/2,\ u(\mu) \geq (b+a)/2\}. \end{aligned}$$

Then be obtain on Ω_a by (9.17) and (9.20)

$$|(z_{a,+} - z_a(\mu))|_{\Omega_a}| = \left| \frac{\mu}{u_+ - a} - \frac{\mu}{u(\mu) - a} \right|_{\Omega_a} = \left. \frac{\mu |u^+ - u(\mu)|}{(u^+ - a)(u(\mu) - a)} \right|_{\Omega_a} \leq \frac{4\mu}{\nu^2} |(u^+ - u(\mu))_{\Omega_a}|$$

Now (9.19), (9.20) yield

$$\begin{aligned} |(z_{b,+} - z_b(\mu))_{\Omega_a}| &\leq |(z_{a,+} - z_a(\mu))_{\Omega_a}| + |(\beta(u_+) - \beta(u(\mu)))_{\Omega_a}| \\ &\quad + |(\hat{g}_s(y,u,\lambda) - \hat{g}_s(y(\mu), u(\mu), \lambda(\mu)))_{\Omega_a}|. \end{aligned}$$

Completely similar we obtain on Ω_b by (9.17) and (9.20)

$$|(z_{b,+} - z_b(\mu))_{\Omega_b}| = \left| \frac{\mu}{b - u_+} - \frac{\mu}{b - u(\mu)} \right|_{\Omega_b} \leq \frac{4\mu}{\nu^2} |(u^+ - u(\mu))_{\Omega_b}|$$

Now again (9.19), (9.20) yield

$$\begin{aligned} |(z_{a,+} - z_a(\mu))_{\Omega_b}| &\leq |(z_{b,+} - z_b(\mu))_{\Omega_b}| + |(\beta(u_+) - \beta(u(\mu)))_{\Omega_b}| \\ &\quad + |(\hat{g}_s(y,u,\lambda) - \hat{g}_s(y(\mu), u(\mu), \lambda(\mu)))_{\Omega_b}|. \end{aligned}$$

Finally, (9.17), (9.16) yield on Ω_a'

$$(z_{a,+} - z_a(\mu))_{\Omega_a'} < 0, \quad (z_{b,+} - z_b(\mu))_{\Omega_a'} > 0$$

and thus the difference of (9.19), (9.20) yields

$$\begin{aligned} &|(z_{b,+} - z_b(\mu))_{\Omega_a'}| + |(z_{a,+} - z_a(\mu))_{\Omega_a'}| = |(z_{b,+} - z_b(\mu) + z_a(\mu) - z_{a,+})_{\Omega_a'}| \\ &\leq |(\beta(u_+) - \beta(u(\mu)))_{\Omega_a'}| + |(\hat{g}_s(y,u,\lambda) - \hat{g}_s(y(\mu), u(\mu), \lambda(\mu)))_{\Omega_a'}|. \end{aligned}$$

Analogously we obtain on Ω_b'

$$\begin{aligned} &|(z_{b,+} - z_b(\mu))_{\Omega_b'}| + |(z_{a,+} - z_a(\mu))_{\Omega_b'}| = |(z_{b,+} - z_b(\mu) + z_a(\mu) - z_{a,+})_{\Omega_b'}| \\ &\leq |(\beta(u_+) - \beta(u(\mu)))_{\Omega_b'}| + |(\hat{g}_s(y,u,\lambda) - \hat{g}_s(y(\mu), u(\mu), \lambda(\mu)))_{\Omega_b'}|. \end{aligned}$$

Taking all together, we have shown that

$$\begin{aligned} &|z_{b,+} - z_b(\mu)| + |z_{a,+} - z_a(\mu)| \\ &\leq |\beta(u_+) - \beta(u(\mu))| + |\hat{g}_s(y,u,\lambda) - \hat{g}_s(y(\mu), u(\mu), \lambda(\mu))| + \frac{8\mu_0}{\nu^2} |u^+ - u(\mu)| \\ &\leq |\hat{g}_s(y,u,\lambda) - \hat{g}_s(y(\mu), u(\mu), \lambda(\mu))| + \left(\frac{8\mu_0}{\nu^2} + \sup_{[\inf a, \sup b]} \beta' \right) |u^+ - u(\mu)|. \end{aligned}$$

Hence, (9.15) yields

$$\|z_{b,+} - z_b(\mu)\|_q + \|z_{a,+} - z_a(\mu)\|_q \leq \left(L_g + \frac{L_g}{\alpha_0} \left(\frac{8\mu_0}{\nu^2} + \sup_{[\inf a, \sup b]} \beta' \right) \right) \|w - w(\mu)\|_{W_p}.$$

Together with (9.15), (9.18) is proven. The last statement is obvious by the definition of Q_μ. □

If we replace the projection P_μ in Lemma 9.4 by the smoothing operator Q_μ (which yields a point in the neighborhood) then we obtain directly the following corollary.

COROLLARY 9.7. *Let the assumptions of Lemma 9.4 hold. If $\bar{w} = \lim_{\mu \searrow 0} w(\mu)$ satisfies strict complementarity then there exist a constants $L_S, C > 0$ such that for any $0 < \mu \leq \mu_0$ and for any $w \in N_{-\infty,q}(\mu)$ with $\|w - w(\mu)\|_{W_q} \leq \rho_0$ the solution w_+ of the primal dual Newton step (8.1) satisfies*

(9.21)

$$\begin{aligned}\|Q_\mu(w_+) - w(\mu)\|_{W_q} &\leq L_S \|w_+ - w(\mu)\|_{W_p} \\ &\leq 2L_S C \left(\omega(4\|w - w(\mu)\|_{W_q}^{\eta q'})^{1/q'} + \|w - w(\mu)\|_{W_q}^{\eta}\right) \|w - w(\mu)\|_{W_q} \\ &= o(\|w - w(\mu)\|_{W_q})\end{aligned}$$

with

$$\eta = \frac{(q-p)\min(p, q-p)}{pq(p+1)}, \quad q' = \frac{qp}{q-p}.$$

Moreover, if $\bar{w}$ satisfies strong strict complementarity (9.4) then

$$\|Q_\mu(w_+) - w(\mu)\|_{W_q} \leq L_S \|w_+ - w(\mu)\|_{W_p} \leq 2L_S C \|w - w(\mu)\|_{W_q}^{1+\eta} \tag{9.22}$$

with

$$\eta = \frac{\min(1,\kappa)(q-p)\min(p, q-p)}{p^2(q+1) + \min(1,\kappa)p(q-p)}.$$

Note that no norm gap appears in (9.21) and (9.22).

9.3. A modified interior point method with smoothing step. We consider now the following modification of Algorithm PDPF.

Algorithm PDPFS: Projected Primal-Dual Interior-Point Method with Smoothing.

1. Choose $\nu \in (0,1)$, $C_0 > 0$, $0 < \sigma_{min} < 1$ and $\mu_0 > 0$. Select the constants $\gamma \in (0,1)$ and $\mu_{-\infty} > \mu_0$ for the neighborhood and choose $w_0 := (y_0, u_0, \lambda_0, z_{a,0}, z_{b,0}) \in N_{-\infty,q}(\mu_0)$ such that

   $$\|F_{\mu_0}(w_0)\|_{V \times L^q \times \Lambda \times L^{q/2}(\Omega)^2} \leq C_0 \sqrt[3]{\mu_0}.$$

   Choose $j_{max} \in \mathbb{N}_0$. Set $\bar{\mu}_0 = \mu_0$ and $k = 0$.
2. Solve the Newton-System

   $$DF_{\mu_k}(w_k)\, s_k = -F_{\mu_k}(w_k).$$

   Choose – if possible – the maximal stepsize $\alpha_k \in \left\{1, 2^{-1}, \dots, 2^{-j_{max}}\right\}$ such that

   $$\|F_{\mu_k}(Q_{\mu_k}(w_k + \alpha_k s_k))\|_{V \times L^q \times \Lambda \times L^{q/2}(\Omega)^2} \leq C_0(1 - \nu(1 - \sqrt[3]{\sigma_k})\alpha_k) \sqrt[3]{\bar{\mu}_k}.$$

   and set $w_{k+1} = Q_{\mu_k}(w_k + \alpha_k s_k)$ and $\bar{\mu}_{k+1} = (1 - \nu(1 - \sqrt[3]{\sigma_k})\alpha_k)^3 \bar{\mu}_k$.
   Otherwise choose the maximal stepsize $\alpha_k \in 2^{-j}$, $j \in \mathbb{N}_0$, such that

   $$\|F_{\mu_k}(P_{\mu_k}(w_k + \alpha_k s_k))\|_{V \times L^q \times \Lambda \times L^{q/2}(\Omega)^2} \leq C_0(1 - \nu(1 - \sqrt[3]{\sigma_k})\alpha_k) \sqrt[3]{\bar{\mu}_k}.$$

   and set $w_{k+1} = P_{\mu_k}(w_k + \alpha_k s_k)$ and $\bar{\mu}_{k+1} = (1 - \nu(1 - \sqrt[3]{\sigma_k})\alpha_k)^3 \bar{\mu}_k$.

3. If

$$\|F_{\mu_k}(w_{k+1})\|_{V\times L^q\times\Lambda\times L^{q/2}(\Omega)^2} \le \frac{1}{2}C_0\sqrt[3]{\mu_k}$$

then choose $\sigma_{min,k}\in(0,1)$ and choose the smallest $\sigma_k\in[\sigma_{min,k},1)$ with

$$\|F_{\sigma_k\mu_k}(w_{k+1})\|_{V\times L^q\times\Lambda\times L^{q/2}(\Omega)^2} \le C_0\sqrt[3]{\mu_k}$$

and set $\mu_{k+1}=\sigma_k\mu_k$, $\bar\mu_{k+1}=\mu_k$.
Else set $\mu_{k+1}=\mu_k$

4. Set $k:=k+1$ and goto 2.

We are now in the position to prove global linear and local superlinear convergence of algorithm PDPFS.

THEOREM 9.8. *Let $\mu_0>0$ and $\rho_0>0$ be fixed. Assume that (A1)–(A4) and (A5)$_q$ with some $q\in(p,\infty)$ hold, that $\bar w=\lim_{\mu\searrow 0}w(\mu)$ satisfies strict complementarity and that*

$$(v,\hat H(y,u,\lambda)v)\ge\alpha\|v\|_2^2 \quad \forall\, v\in L^2(\Omega), \qquad \forall\, w\in N_{-\infty,q}(\mu),\ \|w-w(\mu)\|_{W_q}\le\rho_0,\ \mu\in(0,\mu_0]. \tag{9.6}$$

Then there exist constants $\bar\rho\in(0,\rho_0]$, $\bar\sigma_{min}\in(0,1)$ and a sequence $\bar\sigma_{min}\ge\bar\sigma_{min,k}\searrow 0$ such that Algorithm PFPFS has the following convergence property:

For any starting point $w\in N_{-\infty,q}(\mu_0)$ with $\|w-w(\mu_0)\|_{W_q}\le\bar\rho$, Algorithm PDPFS with $\sigma_{min,k}\in[\bar\sigma_{min,k},\sigma_{max}]$, $\sigma_{max}\in[\bar\sigma_{min},1)$ chooses

$$\alpha_k=1,\quad \mu_{k+1}=\sigma_k\mu_k=\sigma_{min,k}\mu_k$$

and generates a sequence with

$$\|w_k-w(\mu_k)\|_{W_q}\le C\sqrt{\mu_k} \tag{9.23}$$

$$\|w_k-\bar w\|_{W_q}\le (C+L)\sqrt{\mu_k} \tag{9.24}$$

$$\mu_k=\sigma_0\cdots\sigma_{k-1}\mu_0=\sigma_{min,0}\cdots\sigma_{min,k-1}\le\sigma_{max}^k\mu_0\to 0 \tag{9.25}$$

with constants $C,L>0$. This show global R-linear convergence with rate σ_{max}.

The choice $\sigma_{min,k}=O(\bar\sigma_{min,k})$ yields $\sigma_{min,k}\searrow 0$ and thus R-superlinear convergence.

If strong strict complementarity (9.4) holds at $\bar w$, then

$$\bar\sigma_{min,k}=O(\mu_k^\eta)$$

with $\eta>0$ according to (9.10) and the choice $\sigma_{min,k}=O(\bar\sigma_{min,k})$ yields R-superlinear convergence with rate $1+\eta$.

Proof. Consider an arbitrary $\mu\in(0,\mu_0]$. Then there exists by Corollary 9.7 a constant $C>0$ such that for any $w\in N_{-\infty,q}(\mu)$ with $\|w-w(\mu)\|_{W_q}\le\rho_0$ the estimate holds

$$\|Q_\mu(w_+)-w(\mu)\|_{W_q}\le\psi(\|w_+-w(\mu)\|_{W_q})\|w_+-w(\mu)\|_{W_q}$$

with

$$\psi(t)=2L_SC(\omega(4t^{\eta q'})^{1/q'}+t^\eta)$$

and $\eta>0$, q' according to (9.8), where w_+ is the result of the primal-dual Newton step (8.1).

Choose $\tau > 0$ such that

$$\psi(\tau) \leq \frac{1}{4}$$

and set

$$\bar{\rho} = \min(\tau, 2L\sqrt{\mu_0}) \tag{9.26}$$

with the Hölder constant L of the central path. Assume – which we will show later – that Algorithm PDPFS uses the iterates

$$w_{k+1} = Q_{\mu_k}(w_k + s_k), \quad \mu_{k+1} = \sigma_k \mu_k, \tag{9.27}$$

where σ_k satisfies

$$\sigma_k \geq \bar{\sigma}_{min,k} := \begin{cases} \max\left\{\left(1 - \frac{\tau(1-\psi(\tau))}{L\sqrt{\mu_k}}\right)_+^2, 4\psi(\tau)^2\right\} & \text{if } \tau < 2L\sqrt{\mu_k}, \\ 4\psi(2L\sqrt{\mu_k})^2 & \text{if } \tau \geq 2L\sqrt{\mu_k}. \end{cases} \tag{9.28}$$

We note that $0 < \bar{\sigma}_{min,k} < 1$, more precisely,

$$\bar{\sigma}_{min,k} \leq \max\left\{\left(1 - \frac{\tau(1-\psi(\tau))}{L\sqrt{\mu_0}}\right)_+^2, 4\psi(\tau)^2\right\} =: \bar{\sigma}_{min} < 1.$$

We show next that the choice (9.28) yields

$$\|w_k - w(\mu_k)\|_{W_q} \leq \min(\tau, 2L\sqrt{\mu_{k-1}}), \tag{9.29}$$

$$\|w_k - w(\mu_{k-1})\|_{W_q} \leq \min(\tau, 2L\sqrt{\mu_{k-1}}), \tag{9.30}$$

where we set $\mu_{-1} = \mu_0$. In fact, this is true for $k = 0$ if $\|w_0 - w(\mu_0)\|_{W_q} \leq \bar{\rho}$. Moreover, to proceed by induction we observe that

$$\begin{aligned} \|w_{k+1} - w(\mu_{k+1})\|_{W_q} &\leq \|w(\mu_{k+1}) - w(\mu_k)\|_{W_q} + \|w_{k+1} - w(\mu_k)\|_{W_q} \\ &\leq L(\sqrt{\mu_k} - \sqrt{\mu_{k+1}}) + \psi(\|w_k - w(\mu_k)\|_{W_q})\|w_k - w(\mu_k)\|_{W_q} \\ &\leq L(1 - \sqrt{\sigma_k})\sqrt{\mu_k} + \psi(\min(\tau, 2L\sqrt{\mu_{k-1}}))\min(\tau, 2L\sqrt{\mu_{k-1}}). \end{aligned}$$

We consider now three cases.

Case 1: $\tau < 2L\sqrt{\mu_k}$. Then we obtain by using (9.28)

$$\begin{aligned} \|w_{k+1} - w(\mu_{k+1})\|_{W_q} &\leq \|w(\mu_{k+1}) - w(\mu_k)\|_{W_q} + \|w_{k+1} - w(\mu_k)\|_{W_q} \\ &\leq L(1 - \sqrt{\sigma_k})\sqrt{\mu_k} + \psi(\tau)\tau \leq \tau = \min(\tau, 2L\sqrt{\mu_k}). \end{aligned}$$

Case 2: $2L\sqrt{\mu_k} \leq \tau < 2L\sqrt{\mu_{k-1}}$. Then (9.28) yields $\mu_k \geq 4\psi(\tau)^2\mu_{k-1}$ and thus

$$\begin{aligned} \|w_{k+1} - w(\mu_{k+1})\|_{W_q} &\leq \|w(\mu_{k+1}) - w(\mu_k)\|_{W_q} + \|w_{k+1} - w(\mu_k)\|_{W_q} \\ &\leq L(1 - \sqrt{\sigma_k})\sqrt{\mu_k} + \psi(\tau)\tau \leq L\sqrt{\mu_k} + \psi(\tau)2L\sqrt{\mu_{k-1}} \\ &\leq 2L\sqrt{\mu_k} = \min(\tau, 2L\sqrt{\mu_k}). \end{aligned}$$

Case 3: $2L\sqrt{\mu_{k-1}} \le \tau$. Then (9.28) yields $\mu_k \ge 4\psi(2L\sqrt{\mu_{k-1}})^2\mu_{k-1}$ and thus

$$\begin{aligned}\|w_{k+1} - w(\mu_{k+1})\|_{W_q} &\le \|w(\mu_{k+1}) - w(\mu_k)\|_{W_q} + \|w_{k+1} - w(\mu_k)\|_{W_q} \\ &\le L(1-\sqrt{\sigma_k})\sqrt{\mu_k} + \psi(2L\sqrt{\mu_{k-1}})2L\sqrt{\mu_{k-1}} \\ &\le L\sqrt{\mu_k} + L\sqrt{\mu_k} \le 2L\sqrt{\mu_k} = \min(\tau, 2L\sqrt{\mu_k}).\end{aligned}$$

Hence, in all three cases (9.29), (9.30) hold.

We still have to show that after a possible reduction of $\tau > 0$ Algorithm PDPFS actually generates w_k, μ_k satisfying (9.27), as long as σ_k satisfies (9.28). To this end we observe that Algorithm PDPFS chooses

$$w_{k+1} = Q_{\mu_k}(w_k + s_k), \quad \mu_{k+1} = \sigma_k\mu_k,$$

if

$$\|F_{\mu_k}(Q_{\mu_k}(w_k + s_k))\|_{V\times L^q\times\Lambda\times L^{q/2}(\Omega)^2} \le \frac{C_0}{2}\sqrt[3]{\mu_k}. \tag{9.31}$$

and if

$$\|F_{\sigma_k\mu_k}(Q_{\mu_k}(w_k + s_k))\|_{V\times L^q\times\Lambda\times L^{q/2}(\Omega)^2} \le C_0\sqrt[3]{\mu_k}. \tag{9.32}$$

But the operators $F_{\mu_k} : W_q \to V \times L^q \times \Lambda \times L^{q/2}(\Omega)^2$ are obviously Lipschitz-continuous on the bounded set

$$\Big\{w \in N_{-\infty,q}(\mu)\ :\ \|w - w(\mu)\|_{W_q} \le \rho, \quad 0 < \mu \le \mu_0\Big\}$$

with some Lipschitz constant L_F. Hence, (9.27), (9.30) yield

$$\|F_{\mu_k}(Q_{\mu_k}(w_k + s_k))\|_{V\times L^q\times\Lambda\times L^{q/2}(\Omega)^2} \le L_F \min(\tau, 2L\sqrt{\mu_k})$$

which implies (9.31) for $\tau > 0$ small enough. Then also (9.32) holds, since by (9.27), (9.29)

$$\|F_{\sigma_k\mu_k}(Q_{\mu_k}(w_k + s_k))\|_{V\times L^q\times\Lambda\times L^{q/2}(\Omega)^2} \le L_F \min(\tau, 2L\sqrt{\mu_k}).$$

□

REFERENCES

[1] E. ALLGOWER, K. BÖHMER, F. POTRA, AND W. RHEINBOLDT, *A mesh-independence principle for operator equations and their discretizations.*, SIAM J. Numer. Anal., 23 (1986), pp. 160–169.

[2] W. ALT, *Discretization and mesh-independence of Newton's method for generalized equations.*, in Fiacco, A. (ed.), Mathematical programming with data perturbations, Lecture Notes in Pure and Appl. Math. 195, Dekker, New York, 1998, pp. 1–30.

[3] M. BERGOUNIOUX, K. ITO, AND K. KUNISCH, *Primal-dual strategy for constrained optimal control problems.*, SIAM J. Control Optim., 37 (1999), pp. 1176–1194.

[4] J. F. BONNANS AND A. SHAPIRO, *Optimization problems with perturbations: a guided tour*, SIAM Rev., 40 (1998), pp. 228–264.

[5] J. C. DE LOS REYES AND K. KUNISCH, *A semi-smooth Newton method for control constrained boundary optimal control of the Navier-Stokes equations*, Nonlinear Anal., 62 (2005), pp. 1289–1316.

[6] A. FORSGREN, P. E. GILL, AND M. H. WRIGHT, *Interior methods for nonlinear optimization*, SIAM Review, 44 (2002), pp. 525–597.

[7] M. HINTERMÜLLER, K. ITO, AND K. KUNISCH, *The primal-dual active set strategy as a semismooth newton method.*, SIAM J. Optim., 13 (2003), pp. 865–888.

[8] M. HINTERMÜLLER AND M. ULBRICH, *A mesh-independence result for semismooth Newton methods*, Math. Program., 101 (2004), pp. 151–184.

[9] M. HINZE AND K. KUNISCH, *Second order methods for optimal control of time-dependent fluid flow*, SIAM J. Control Optim., 40 (2001), pp. 925–946 (electronic).

[10] C. T. KELLEY AND E. W. SACHS, *Multilevel algorithms for constrained compact fixed point problems*, SIAM J. Sci. Comput., 15 (1994), pp. 645–667.

[11] U. PRÜFERT, F. TRÖLTSCH, AND M. WEISER, *The convergence of an interior point method for an elliptic control problem with mixed control-state constraints.*, Tech. Rep. 36-2004, Inst. of Math., TU Berlin, 2005.

[12] S. M. ROBINSON, *Stability theory for systems of inequalities. II: Differentiable nonlinear systems.*, SIAM J. Numer. Anal., 13 (1976), pp. 497–513.

[13] A. SCHIELA AND M. WEISER, *Superlinear convergence of the control reduced interior point method for PDE constrained optimization.*, Tech. Rep. ZIB 05-15, ZIB, Berlin, 2005.

[14] M. ULBRICH, *Constrained optimal control of Navier-Stokes flow by semismooth Newton methods*, Systems Control Lett., 48 (2003), pp. 297–311. Optimization and control of distributed systems.

[15] ———, *Semismooth Newton methods for operator equations in function spaces.*, SIAM J. Optim., 13 (2003), pp. 805–841.

[16] M. ULBRICH AND S. ULBRICH, *Global convergence of trust-region interior-point algorithms for infinite-dimensional nonconvex minimization subject to pointwise bounds.*, SIAM J. Control Optim., 37 (1999), pp. 731–764.

[17] ———, *Superlinear convergence of affine-scaling interior-point Newton methods for infinite-dimensional nonlinear problems with pointwise bounds.*, SIAM J. Control Optim., 38 (2000), pp. 1938–1984.

[18] M. WEISER, *Interior point methods in function space.*, SIAM J. Control Optim., 44 (2005), pp. 1766–1786.

[19] M. WEISER, T. GÄNZLER, AND A. SCHIELA, *A control reduced primal interior point method for PDE constrained optimization.*, Tech. Rep. ZIB 04-38, ZIB, Berlin, 2004.

[20] M. WEISER AND A. SCHIELA, *Function space interior point methods for PDE constrained optimization.*, PAMM, 4 (2004), pp. 43–46.