A HIGHER DIMENSIONAL PARTIAL LEGENDRE TRANSFORM, AND REGULARITY OF DEGENERATE MONGE-AMPÈRE EQUATIONS

CRISTIAN RIOS, ERIC T. SAWYER, AND RICHARD L. WHEEDEN

ABSTRACT. In dimension $n \geq 3$, we define a generalization of the classical two dimensional partial Legendre transform, that reduces interior regularity of the generalized Monge-Ampère equation $\det D^2 u = k(x, u, Du)$ to regularity of a divergence form quasilinear system of special form. This is then used to obtain smoothness of $C^{2,1}$ solutions, having $n-1$ nonvanishing principal curvatures, to certain subelliptic Monge-Ampère equations in dimension $n \geq 3$. A corollary is that if $k \geq 0$ vanishes only at nondegenerate critical points, then a $C^{2,1}$ convex solution u is smooth if and only if the symmetric function of degree $n-1$ of the principal curvatures of u is positive, and moreover, u fails to be $C^{3,1-\frac{2}{n}+\varepsilon}$ when not smooth.

1. Introduction

We consider regularity of the generalized Monge-Ampère equation,

$$\det D^2 u = k(x, u, Du), \qquad x \in \Omega, \tag{1.1}$$

where k is smooth and nonnegative in $\Omega \times \mathbb{R} \times \mathbb{R}^n$, and Ω is a convex domain in $\mathbb{R}^n$. We first introduce a higher dimensional partial Legendre transform corresponding to a convex solution u of (1.1),

$$\begin{cases} s &= x_1 \\ t_2 &= \frac{\partial u}{\partial x_2}(x) \\ \vdots & \\ t_n &= \frac{\partial u}{\partial x_n}(x) \end{cases},$$

and show that the vector-valued function $\mathbf{v} = (v_\ell)_{\ell=2}^n = (x_\ell(s, \mathbf{t}))_{\ell=2}^n$ is a weak solution of the divergence form quasilinear system (elliptic when $k > 0$),

$$\mathcal{L} v_\ell \equiv \left\{ \frac{\partial^2}{\partial s^2} + \frac{\partial}{\partial \mathbf{t}'} k \left(co \left[\frac{\partial \mathbf{v}}{\partial \mathbf{t}'} \right] \right)' \frac{\partial}{\partial \mathbf{t}} \right\} v_\ell = 0, \qquad 2 \leq \ell \leq n, \tag{1.2}$$

where $\left(co \left[\frac{\partial \mathbf{v}}{\partial \mathbf{t}'} \right] \right)'$ denotes the transposed cofactor matrix of $\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}$. See subsection 2.1.3 below for the derivation. The significant feature of the system (1.2) is that the degeneracy of the operator is incorporated (at least when we assume that $\det \frac{\partial \mathbf{t}}{\partial \mathbf{x}'} = \det \left[\frac{\partial^2 u}{\partial x_i \partial x_j} \right]_{i,j=2}^n > 0$) into the function k appearing in the coefficient matrix, thus permitting the use of subelliptic De Giorgi - Nash - Moser theory as in [5] and [21].

Date: October 1, 2002.

1991 *Mathematics Subject Classification.* 35B65, 35B45.

Key words and phrases. Monge-Ampère, partial Legendre transform, regularity.

This is in contrast to the usual quasilinear system obtained by differentiating (1.1) with respect to $\frac{\partial}{\partial x_\sigma}$:

$$trace\left\{\left(co\left[D^2u\right]\right)'\left(\nabla\nabla'\frac{\partial u}{\partial x_\sigma}\right)\right\}=k_{x_\sigma}+k_u\frac{\partial u}{\partial x_\sigma}+k_{Du}\left(D\frac{\partial u}{\partial x_\sigma}\right),\qquad 1\leq\sigma\leq n.$$

This can be put into divergence form using $\nabla'\left(co\left[D^2u\right]\right)'=\mathbf{0}'$ (see section 4.3 of the appendix below):

$$\nabla'\left(co\left[D^2u\right]\right)'\nabla\left(\frac{\partial u}{\partial x_\sigma}\right)=k_{x_\sigma}+k_u\frac{\partial u}{\partial x_\sigma}+k_{Du}\left(D\frac{\partial u}{\partial x_\sigma}\right),\qquad 1\leq\sigma\leq n,$$

but the degeneracy when $k=0$ remains embedded in the matrix $\left(co\left[D^2u\right]\right)'$. In dimension $n=2$, the transposed cofactor matrix in (1.2) defaults to 1, yielding the classical equation (see (1.3) below) for the two dimensional partial Legendre transform. For a discussion of other partial Legendre transforms, see the appendix.

We then apply this partial Legendre transform to generalize the two dimensional regularity theorem of Guan [9] to higher dimensions: namely that a $C^{2,1}(\Omega)$ convex solution u to (1.1) is smooth if k vanishes to finite order in a certain sense, and if $n-1$ of the principal curvatures of the solution u are bounded away from zero (fewer than $n-1$ nonvanishing principal curvatures do not suffice). Before continuing further with the development of (1.2) and a discussion of the regularity application, we briefly review some history.

In the case $k>0$, the equation (1.1) is elliptic and the theory is well developed. For example, if $k=k(x)$, Caffarelli, Nirenberg and Spruck have shown in [1] that there is a unique smooth convex solution u to the Dirichlet problem for (1.1) in $\overline{\Omega}$ with smooth data when $\partial\Omega$ has positive Gaussian curvature. However, if k is permitted to vanish in Ω, regularity may fail spectacularly. For example, if $u(x)=|x|^{2+\frac{2}{n}}$, then by rotation invariance and homogeneity, (1.1) holds with $k=c_n|x|^2$, and thus u is a $C^{2,\frac{2}{n}}$ solution, and no better, of the Monge-Ampère equation with analytic k that vanishes to the least order possible. The best possible regularity for the degenerate Dirichlet problem is given by Guan [10], and Guan, Trudinger and Wang [12]; for k nonnegative and smooth, there is a unique $C^{1,1}\left(\overline{\Omega}\right)$ convex solution u to the Dirichlet problem for (1.1) in the generalized sense of Alexandrov.

In two dimensions, Guan [9] has shown that a $C^{1,1}(\Omega)$ solution u to (1.1) is smooth if k vanishes to finite order in a certain sense, and if one principal curvature of the solution u is bounded away from zero (see also earlier work in two dimensions in Xu [24]). This identifies the rank of the Hessian of the solution as an obstacle to regularity even in the subelliptic case. Three main ingredients were used in the proof:

1. the two dimensional partial Legendre transform (called semispherical mapping in [17] and Legendre-like transform in [22]), $\left\{\begin{array}{rcl}s&=&x\\t&=&u_y(x,y)\end{array}\right.$, associated to a convex solution u of $u_{xx}u_{yy}-u_{xy}^2=k(x,y)$, which results in a divergence form quasilinear equation for $v=y(s,t)$;

$$\mathcal{L}v=\left\{\partial_s^2+\partial_t k(s,v)\partial_t\right\}v=0,\tag{1.3}$$

2. Franchi's extension [5] of the De Giorgi - Nash - Moser theory to certain subelliptic linear divergence form equations in two dimensions, including (1.3),
3. a regularity theorem for divergence form quasilinear equations with elliptic extendible operator, including (1.3).

In [21], two of the authors have extended the two dimensional case of Franchi's subelliptic theorem in [5] to higher dimensions. We would now like to use the higher dimensional partial Legendre transform introduced above. Unfortuately, while the system (1.2) is elliptic when $k > 0$, it is not diagonal in the principal terms, and is never strongly elliptic (see the appendix). However, if we use the divergence-free property of the matrix $M = \left(co\left[\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right]\right)'$, namely $\partial_{\mathbf{t}'} M = \mathbf{0}'$, and differentiate (1.2), we obtain that the vector-valued function $\mathbf{p} = D\mathbf{v} = \left(\frac{\partial v_i}{\partial t_j}\right)_{2 \le i \le n, 1 \le j \le n}$ $(s = t_1)$ satisfies the divergence form quasilinear system (with $M = M(\mathbf{p})$ a function of $\mathbf{p}$)

$$\mathcal{L}\mathbf{p} \equiv \left\{ \frac{\partial^2}{\partial s^2} + \frac{\partial}{\partial \mathbf{t}'} k M(\mathbf{p}) \frac{\partial}{\partial \mathbf{t}} \right\} \mathbf{p} = \mathbf{f}\left((s, \mathbf{t}), \mathbf{v}, \mathbf{p}, D\mathbf{p}\right), \tag{1.4}$$

that is diagonal in the principal terms, strongly elliptic when $k > 0$, and has inhomogeneous term $\mathbf{f}$ that is quadratic in $D\mathbf{p}$. Here we view the scalar operator in braces as acting on each component of $\mathbf{p}$ separately. See the beginning of section 3 below for a precise statement.

Notation: We use boldface characters to denote column vectors, or $m \times 1$ matrices, and a prime to denote the transpose of a matrix. Square matrices will typically be denoted with square brackets. Juxtaposition of matrices indicates matrix multiplication. Note however that we do not view the matrix $\frac{\partial \mathbf{v}}{\partial \mathbf{t}'} = \left[\frac{\partial v_i}{\partial t_j}\right]_{2 \le i,j \le n}$ as the juxtaposition of $\frac{\partial}{\partial \mathbf{t}'}$ and $\mathbf{v}$. Moreover, when a column vector such as $\mathbf{v}$ appears as an argument of a function, we will often think of it as a row vector while continuing to write $\mathbf{v}$ rather than $\mathbf{v}'$.

We now generalize the theorem of Guan in [9] by giving subelliptic conditions on k which yield smoothness for convex solutions $u \in C^{2,1}(\Omega)$ to the generalized Monge-Ampère equation (1.1) provided u has $n-1$ nonvanishing principal curvatures.

Theorem 1.1. *Let $u \in C^{2,1}(\Omega)$ be a convex solution to (1.1) with k smooth on $\Omega \times \mathbb{R} \times \mathbb{R}^n$ satisfying*

$$k(x, u, Du) \approx \left(|x_1|^{2m} + \psi(x_1, \mathbf{x})\right) K(x, u, Du), \qquad x = (x_1, \mathbf{x}) \in \Omega, \tag{1.5}$$

where K is smooth and positive on $\Omega \times \mathbb{R} \times \mathbb{R}^n$, ψ is smooth and nonnegative on Ω, m is a positive integer and $\psi(x_1, \mathbf{x})^{\frac{1}{2m}}$ is Lipschitz. If

$$d = \det \begin{bmatrix} \frac{\partial^2 u}{\partial x_2^2} & \cdots & \frac{\partial^2 u}{\partial x_2 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 u}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 u}{\partial x_n^2} \end{bmatrix} > 0 \tag{1.6}$$

everywhere in Ω, then $u \in C^\infty(\Omega)$.

Remark 1.1. *The theorem fails if merely a minor of size $(n-2) \times (n-2)$ is assumed nondegenerate in (1.6). For example, $u(x) = |(x_1, x_2)|^3 + \sum_{j=3}^n \frac{1}{2} x_j^2$ lies* *in* $C^{2,1}(\Omega)$*, satisfies (1.1) with* $k = 18\left(|x_1|^2 + |x_2|^2\right)$*, and verifies*

$$\det \begin{bmatrix} \frac{\partial^2 u}{\partial x_3^2} & \cdots & \frac{\partial^2 u}{\partial x_3 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 u}{\partial x_n \partial x_3} & \cdots & \frac{\partial^2 u}{\partial x_n^2} \end{bmatrix} = 1,$$

yet u *fails to be in* $C^3(\Omega)$*. Of course* d *in (1.6) vanishes when* $k = 0$*.*

We point out that the theorem applies in particular to the equation of prescribed Gaussian curvature,

$$\frac{\det D^2 u}{\left(1 + |Du|^2\right)^{\frac{n+2}{2}}} = k_n(x), \qquad x \in \Omega. \tag{1.7}$$

One geometric consequence is that if a $C^{2,1}$ convex function u has graph with smooth Gaussian curvature $k_n(x) \approx |x|^{2m}$, $m \in \mathbb{N}$, so that u solves (1.7), then u is smooth provided $k_{n-1}(0) > 0$. Here k_j denotes the j^{th} elementary symmetric function of the principal curvatures of u, $1 \leq j \leq n$, often referred to as the j^{th} elementary symmetric curvature (k_n is the Gaussian curvature and k_1 the mean curvature). For j even, k_j is an isometry invariant of the surface. To apply Theorem 1.1, rotate coordinates so that $d(0) = k_{n-1}(0)$ in (1.6). The same result holds for a $C^{2,1}$ convex solution u to (1.1), and in the case $m = 1$ leads to the following characterization of regularity for solutions to (1.1).

Corollary 1.2. *Suppose* k *is smooth and nonnegative in* $\Omega \times \mathbb{R} \times \mathbb{R}^n$*,* $\Omega \subset \mathbb{R}^n$*, vanishes at the origin in* Ω*, and has a nondegenerate critical point there, i.e.* $k(x,z,p) \approx |x|^2$ *for* (x,z,p) *in compact subsets of* $\Omega \times \mathbb{R} \times \mathbb{R}^n$*. If* u *is a* $C^{2,1}$ *convex solution to (1.1), then* $u \in C^\infty(\Omega)$ *if and only if* $k_{n-1}(0) > 0$*. In the event that* $k_{n-1}(0) = 0$*,* $u \notin C^{3,\beta}(\Omega)$ *for any* $\beta > 1 - \frac{2}{n}$*. Thus when* $k \approx |x|^2$*, all* $C^{3,\beta}$ *convex solutions to (1.1) with* $\beta > 1 - \frac{2}{n}$ *are smooth.*

It suffices to prove that $u \in C^{3,\beta}(\Omega)$, $\beta > 1 - \frac{2}{n}$, implies $k_{n-1}(0) > 0$. However, if $k_{n-1}(0) = 0$, then $k_{n-1} \in C^{1,\beta}$ and $k_{n-1} \geq 0$ imply $k_{n-1}(x) \leq C|x|^{1+\beta}$. If $\{\lambda_j(x)\}_{j=1}^n$ are the principal curvatures of u at x, then

$$k_{n-1}(x)^n = \left(\sum_{j=1}^n \frac{k_n(x)}{\lambda_j(x)}\right)^n \geq \prod_{j=1}^n \left(\frac{k_n(x)}{\lambda_j(x)}\right) = k_n(x)^{n-1} \approx |x|^{2n-2},$$

which yields a contradiction if $\beta > 1 - \frac{2}{n}$. Note that in $n = 2$ dimensions, Iaia [13] has sharpened this argument to show that if $u \in C^3$ solves (1.1) with smooth $k \approx |x|^2$, then $k_1(0) > 0$, and so by Guan's theorem [9], $u \in C^\infty$; the example $u = |x|^3$ shows that $u \in C^{2,1}$ is not enough.

The proof of Theorem 1.1 proceeds along the lines of Guan's two dimensional proof outlined above, using the higher dimensional partial Legendre transform and the system (1.4), the extension to higher dimensions of Franchi's subelliptic result in [21], and an extension of Guan's quasilinear regularity theorem to equations of the form (1.4). To the best of the authors' knowledge, this represents the first C^∞ regularity result for the degenerate Monge-Ampère equation in higher dimensions. In the appendix, we comment briefly on the possibility of using the methods of Morrey and Campanato on the system (1.4).

2. Quasilinear divergence form systems

We begin by recalling the partial Legendre transformation corresponding to a convex solution u of the two dimensional generalized Monge-Ampère equation in a planar domain Ω,

$$u_{xx}u_{yy} - (u_{xy})^2 = k(x, y, u, u_x, u_y), \qquad (x, y) \in \Omega, \tag{2.1}$$

where $k(x, y, r, z, t)$ is smooth and nonnegative in $\Omega \times \mathbb{R}^3$. The partial Legendre transformation $(s, t) = T(x, y)$ (as in [9], [17] and [22]) is given by

$$\left\{ \begin{array}{ccc} s & = & x \\ t & = & u_y(x, y) \end{array} \right. . \tag{2.2}$$

We note that if $u_{yy} > 0$ (in particular if $k > 0$), then u_y is strictly increasing in y, making T one-to-one on the set where k is nonvanishing. If we assume that $u \in C^{1,1}$ with $u_{yy} \geq c > 0$ and set

$$\left\{ \begin{array}{ccccc} v & = & y & = & y(s, t) \\ z & = & u_x(x, y) & = & u_x(s, y(s, t)) \\ r & = & u(x, y) & = & u(s, y(s, t)) \end{array} \right. ,$$

where $(x, y) = (s, y(s, t))$ is the inverse partial Legendre transform, then the Lipschitz functions v, z and r satisfy the quasilinear divergence form equation

$$\partial_s^2 v + \partial_t k(s, v(s, t), r(s, t), z(s, t), t) \partial_t v = 0, \qquad (s, t) \in T(\Omega), \tag{2.3}$$

in the weak sense. Indeed, as in [22] and [9], the Jacobian of T is $\begin{bmatrix} 1 & 0 \\ u_{xy} & u_{yy} \end{bmatrix}$, and that of $S = T^{-1}$ is $\begin{bmatrix} 1 & 0 \\ -\frac{u_{xy}}{u_{yy}} & \frac{1}{u_{yy}} \end{bmatrix}$. Thus

$$\left\{ \begin{array}{ccccc} \partial_s & = & x_s\partial_x + y_s\partial_y & = & \partial_x - \frac{u_{xy}}{u_{yy}}\partial_y, \\ \partial_t & = & x_t\partial_x + y_t\partial_y & = & \frac{1}{u_{yy}}\partial_y, \end{array} \right.$$

and for $\eta \in C_c^\infty(T(\Omega))$ we have by (2.1),

$$\begin{aligned}
(2.4) \qquad & \int_{T(\Omega)} (y_s\eta_s + ky_t\eta_t)\, dsdt \\
= \; & \int_\Omega \left\{ \left(y_x - \frac{u_{xy}}{u_{yy}} y_y \right) \left(\eta_x - \frac{u_{xy}}{u_{yy}} \eta_y \right) + k \left(\frac{1}{u_{yy}} y_y \right) \left(\frac{1}{u_{yy}} \eta_y \right) \right\} u_{yy} dxdy \\
= \; & \int_\Omega \left\{ -\frac{u_{xy}}{u_{yy}} \eta_x + \left(\frac{u_{xy}}{u_{yy}} \right)^2 \eta_y + k \left(\frac{1}{u_{yy}} \right)^2 \eta_y \right\} u_{yy} dxdy \\
= \; & \int_\Omega \left\{ -u_{xy}\eta_x + \frac{(u_{xy})^2 + k}{u_{yy}} \eta_y \right\} dxdy \\
= \; & \int_\Omega \left\{ -u_{xy}\eta_x + u_{xx}\eta_y \right\} dxdy = \int_\Omega \left\{ -u_{xy}\eta_x + u_{xy}\eta_x \right\} dxdy = 0,
\end{aligned}$$

by approximation, since both η and u_x are Lipschitz functions of (x, y), hence in $W^{1,2}$.

2.1. **Higher dimensional partial Legendre transform.** In dimension $n \geq 3$, we define a generalization of the two dimensional partial Legendre transform (2.2) above, that results in a divergence form quasilinear system of $n-1$ equations in $n-1$ unknowns that is diagonal in the principal terms, with inhomogeneous term that is quadratic in the first order derivatives of the unknown functions. Elliptic systems of this nature in two dimensions have been studied in Schulz [22], and in higher dimensions with Hörmander vector fields in Xu and Zuily [25].

2.1.1. *A Cauchy-Riemann system.* The key to generalizing the transform above to higher dimensions is to rewrite (2.3) as a first order Cauchy-Riemann system. For this we need to calculate the s and t derivatives of $r = u$ and $z = u_x$; recall that $\partial_s = \partial_x - \frac{u_{xy}}{u_{yy}}\partial_y$ and $\partial_t = \frac{1}{u_{yy}}\partial_y$ so that

$$\begin{cases} v_s &= -\frac{u_{xy}}{u_{yy}} \\ v_t &= \frac{1}{u_{yy}} \end{cases} .$$

We thus have

$$\begin{cases} r_s &= \left(\partial_x - \frac{u_{xy}}{u_{yy}}\partial_y\right) u &= z + t v_s \\ r_t &= \frac{1}{u_{yy}}\partial_y u &= t v_t \\ z_s &= \left(\partial_x - \frac{u_{xy}}{u_{yy}}\partial_y\right) u_x &= k v_t \\ z_t &= \frac{1}{u_{yy}}\partial_y u_x &= -v_s \end{cases} , \tag{2.5}$$

where k is evaluated at $(s, v(s,t), r(s,t), z(s,t), t)$. We may view these equations as a Cauchy-Riemann system,

$$\begin{cases} \frac{\partial z}{\partial s} &= k(s,v,r,z,t)\frac{\partial v}{\partial t} \\ \frac{\partial z}{\partial t} &= -\frac{\partial v}{\partial s} \end{cases} , \tag{2.6}$$

together with the compatibility conditions

$$\begin{cases} \frac{\partial r}{\partial s} &= z + t\frac{\partial v}{\partial s} \\ \frac{\partial r}{\partial t} &= t\frac{\partial v}{\partial t} \end{cases} .$$

Note that the divergence form equation (2.3) is obtained from the Cauchy-Riemann equations (2.6) using $z_{st} = z_{ts}$.

2.1.2. *Generalized Cauchy-Riemann equations.* In higher dimensions, we consider the generalized Monge-Ampère equation,

$$\det D^2 u(x) = k(x, u, Du), \qquad x \in \Omega, \tag{2.7}$$

where Ω is a convex domain in $\mathbb{R}^n$. Our partial Legendre transform is the following combination of the identity map and the gradient map of u. Keeping in mind our desire to obtain an analogue of the Cauchy-Riemann system (2.6), we introduce variables s and $\mathbf{t} = (t_2, ..., t_n)'$ by

$$\begin{cases} s &= x_1 \\ t_2 &= \frac{\partial u}{\partial x_2}(x) \\ \vdots & \\ t_n &= \frac{\partial u}{\partial x_n}(x) \end{cases} , \tag{2.8}$$

and consider the functions z and $\mathbf{v} = (v_2, ..., v_n)'$ given by

$$\begin{cases} z &= \frac{\partial u}{\partial x_1}(x) \\ v_2 &= x_2 \\ \vdots & \\ v_n &= x_n \end{cases},$$

along with $r = u$. It is also convenient to write $\mathbf{x} = (x_2, ..., x_n)'$ for the variables complementary to x_1. We now assume that $u \in C^{1,1}$ and $d = \det [u_{ij}]_{i,j=2}^n \geq c > 0$, so that both the partial Legendre transformation and its inverse are Lipschitz (note that $k > 0$ implies $d > 0$), and view $\mathbf{v}$, r and z as functions of s and $\mathbf{t}$.

We claim that z and $\mathbf{v} = (v_2, ..., v_n)'$ satisfy the Cauchy-Riemann system

$$\begin{cases} \frac{\partial z}{\partial s} &= k\left((s, \mathbf{v}), r, (z, \mathbf{t})\right) \det \frac{\partial \mathbf{v}}{\partial \mathbf{t}'} \\ \frac{\partial z}{\partial \mathbf{t}} &= -\frac{\partial \mathbf{v}}{\partial s} \end{cases}, \tag{2.9}$$

where $\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}$ is the Jacobian matrix $\left[\frac{\partial v_i}{\partial t_j}\right]_{2 \leq i,j \leq n}$, $\frac{\partial z}{\partial \mathbf{t}}$ is the column vector $\left(\frac{\partial z}{\partial t_2}, ..., \frac{\partial z}{\partial t_n}\right)'$, and $\frac{\partial \mathbf{v}}{\partial s}$ is the column vector $\left(\frac{\partial v_2}{\partial s}, ..., \frac{\partial v_n}{\partial s}\right)'$; and that the function r satisfies the compatibility condition

$$\frac{\partial r}{\partial (s, \mathbf{t}')} = (z, 0, ..., 0) + \mathbf{t}' \frac{\partial \mathbf{v}}{\partial (s, \mathbf{t}')}. \tag{2.10}$$

To see this, we compute the Jacobian matrix $J = \frac{\partial (s, \mathbf{t})}{\partial (x_1, \mathbf{x}')}$ (here and in similar situations below we are viewing $(s, \mathbf{t})$ as a column vector) of the transformation (2.8):

$$J = \frac{\partial (s, \mathbf{t})}{\partial (x_1, \mathbf{x}')} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{bmatrix},$$

and its inverse,

$$J^{-1} = \frac{\partial (x_1, \mathbf{x})}{\partial (s, \mathbf{t}')} = \frac{1}{d} \begin{bmatrix} d & 0 & \cdots & 0 \\ b_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix},$$

where

$$d = \det \frac{\partial \mathbf{t}}{\partial \mathbf{x}'} = \det \frac{\partial \mathbf{t}}{\partial \mathbf{v}'} = \det \begin{bmatrix} u_{22} & \cdots & u_{2n} \\ \vdots & \ddots & \vdots \\ u_{n2} & \cdots & u_{nn} \end{bmatrix} > 0,$$

$\begin{bmatrix} c_{22} & \cdots & c_{2n} \\ \vdots & \ddots & \vdots \\ c_{n2} & \cdots & c_{nn} \end{bmatrix}$ is the transposed cofactor matrix of $\frac{\partial \mathbf{t}}{\partial \mathbf{v}'} = \begin{bmatrix} u_{22} & \cdots & u_{2n} \\ \vdots & \ddots & \vdots \\ u_{n2} & \cdots & u_{nn} \end{bmatrix}$,
and

$$b_{i1} = -\det \begin{bmatrix} u_{22} & \cdots & u_{2(i-1)} & u_{21} & u_{2(i+1)} & \cdots & u_{2n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ u_{n2} & \cdots & u_{n(i-1)} & u_{n1} & u_{n(i+1)} & \cdots & u_{nn} \end{bmatrix}, \tag{2.11}$$

for $2 \leq i \leq n$. Note that the final determinant above is computed for the matrix $\frac{\partial \mathbf{t}}{\partial \mathbf{v}'}$ with its i^{th} column replaced by the column $(u_{21}, ..., u_{n1})'$. Note that we refer to the columns (rows) of a matrix that is indexed by $2 \leq i, j \leq n$ as the second column (row) through to the n^{th} column (row).

Our goal now is to show that the Jacobian matrix $\frac{\partial(z,\mathbf{v})}{\partial(s,\mathbf{t}')}$ satisfies

$$\frac{\partial (z, \mathbf{v})}{\partial (s, \mathbf{t}')} = \begin{bmatrix} \frac{\partial z}{\partial s} & \frac{\partial z}{\partial \mathbf{t}'} \\ \frac{\partial \mathbf{v}}{\partial s} & \frac{\partial \mathbf{v}}{\partial \mathbf{t}'} \end{bmatrix} = \begin{bmatrix} k \det \frac{\partial \mathbf{v}}{\partial \mathbf{t}'} & -\frac{\partial \mathbf{v}'}{\partial s} \\ \frac{\partial \mathbf{v}}{\partial s} & \frac{\partial \mathbf{v}}{\partial \mathbf{t}'} \end{bmatrix},$$

which is the Cauchy-Riemann system (2.9). For this we note that the derivatives $\frac{\partial}{\partial s}, \frac{\partial}{\partial t_2}, ..., \frac{\partial}{\partial t_n}$ can be expressed in terms of x-derivatives via the chain rule and the columns of J^{-1}. In particular,

$$\frac{\partial}{\partial s} = \sum_{i=1}^{n} \frac{\partial x_i}{\partial s} \frac{\partial}{\partial x_i} = \frac{\partial}{\partial x_1} + \frac{1}{d} \sum_{i=2}^{n} b_{i1} \frac{\partial}{\partial x_i},$$

and so

$$\begin{aligned} \frac{\partial z}{\partial s} &= \frac{1}{d} \left(d \frac{\partial}{\partial x_1} + \sum_{i=2}^{n} b_{i1} \frac{\partial}{\partial x_i} \right) \frac{\partial}{\partial_{x_1}} u(x) \\ &= \frac{1}{d} \left\{ u_{11} d + \sum_{i=2}^{n} u_{i1} b_{i1} \right\} \\ &= \frac{1}{d} \{k\} = \frac{k}{\det \frac{\partial \mathbf{t}}{\partial \mathbf{v}'}} = k \det \frac{\partial \mathbf{v}}{\partial \mathbf{t}'}, \end{aligned}$$

which is the first of the Cauchy-Riemann equations in (2.9). To obtain the rest, we read off from J^{-1} that

$$\frac{\partial}{\partial t_j} = \frac{1}{d} \sum_{i=2}^{n} c_{ij} \frac{\partial}{\partial x_i},$$

and so

$$\begin{aligned} \frac{\partial z}{\partial t_j} &= \frac{1}{d} \left(\sum_{i=2}^{n} c_{ij} \frac{\partial}{\partial x_i} \right) \frac{\partial}{\partial_{x_1}} u(x) \\ &= \frac{1}{d} \sum_{i=2}^{n} u_{1i} c_{ij} \\ &= -\frac{b_{j1}}{d} = -\frac{\partial x_j}{\partial s} = -\frac{\partial v_j}{\partial s}, \end{aligned}$$

by (2.11) for $2 \leq j \leq n$. This establishes the remaining equations in (2.9). Finally, to establish the compatibility condition (2.10), we compute

$$\begin{aligned}
\frac{\partial r}{\partial s} &= \frac{1}{d}\left(d\frac{\partial}{\partial x_1}+\sum_{i=2}^{n} b_{i1}\frac{\partial}{\partial x_i}\right)u\left(x\right) \\
&= \frac{\partial u}{\partial x_1}+\sum_{i=2}^{n}\frac{\partial u}{\partial x_i}\frac{b_{i1}}{d} \\
&= z+\sum_{i=2}^{n} t_i\frac{\partial v_i}{\partial s}=z+\mathbf{t}'\frac{\partial \mathbf{v}}{\partial s},
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial r}{\partial t_j} &= \frac{1}{d}\left(\sum_{i=2}^{n} c_{ij}\frac{\partial}{\partial x_i}\right)u\left(x\right) \\
&= \sum_{i=2}^{n}\frac{\partial u}{\partial x_i}\frac{c_{ij}}{d} \\
&= \sum_{i=2}^{n} t_i\frac{\partial v_i}{\partial t_j}=\mathbf{t}'\frac{\partial \mathbf{v}}{\partial t_j}.
\end{aligned}$$

2.1.3. *A divergence form quasilinear system.* We now use the equality of mixed second order partial derivatives, $\frac{\partial}{\partial \mathbf{t}}\left(\frac{\partial z}{\partial s}\right)=\frac{\partial}{\partial s}\left(\frac{\partial z}{\partial \mathbf{t}}\right)$, along with (2.9) to obtain that the Lipschitz functions $\mathbf{v}=\left(v_2,...,v_n\right)'$ satisfy, in the weak sense, the divergence form quasilinear system

$$\frac{\partial}{\partial \mathbf{t}}\left(k\det\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right)=\frac{\partial}{\partial \mathbf{t}}\left(\frac{\partial z}{\partial s}\right)=\frac{\partial}{\partial s}\left(\frac{\partial z}{\partial \mathbf{t}}\right)=\frac{\partial}{\partial s}\left(-\frac{\partial \mathbf{v}}{\partial s}\right)=-\frac{\partial^2 \mathbf{v}}{\partial s^2},$$

which we can rewrite explicitly as

$$\frac{\partial^2 v_\ell}{\partial s^2}+\frac{\partial}{\partial t_\ell}\left(k\det\frac{\partial\left(v_2,...,v_n\right)}{\partial\left(t_2,...,t_n\right)}\right)=0,\qquad 2\leq \ell\leq n. \tag{2.12}$$

This is of divergence form,

$$div_{(s,\mathbf{t}')}\mathbf{F}^\ell\left(\left(s,\mathbf{t}\right),\mathbf{v},D\mathbf{v}\right)=0,\qquad 2\leq\ell\leq n, \tag{2.13}$$

where the n-vector $\mathbf{F}^\ell$ is given by

$$\mathbf{F}^\ell\left(\left(s,\mathbf{t}\right),\mathbf{v},\mathbf{p}^2,...,\mathbf{p}^n\right)=\left(p_1^\ell,0,...,0,k\det\left[p_\tau^\sigma\right]_{2\leq\sigma,\tau\leq n},0,...,0\right)'. \tag{2.14}$$

Here $k\det\left[p_\tau^\sigma\right]_{2\leq\sigma,\tau\leq n}$ occurs in the ℓ^{th} position for $2\leq\ell\leq n$, where we are writing $p_\tau^\sigma=\frac{\partial v_\sigma}{\partial t_\tau}$ for $2\leq\sigma,\tau\leq n$, and $p_1^\sigma=\frac{\partial v_\sigma}{\partial s}$ for $2\leq\sigma\leq n$. Note that the superscript σ indexes the rows and the subscript τ indexes the columns of the matrix $\left[p_\tau^\sigma\right]_{2\leq\sigma,\tau\leq n}$.

Unfortunately, in the preliminary form (2.13) - (2.14), the system fails to be elliptic when $k>0$ (see the appendix). To rectify this lack of ellipticity when $k>0$, we exploit the symmetry of $\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}$ (which follows from the symmetry of $\frac{\partial \mathbf{t}}{\partial \mathbf{v}'}=\frac{\partial \mathbf{t}}{\partial \mathbf{x}'}$). If $M=\left(\det\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right)\left[\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right]^{-1}$ denotes the transposed cofactor matrix of $\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}$, then

$$M\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}=\left(\det\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right)\left[\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right]^{-1}\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}=\left(\det\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right)I_{n-1}, \tag{2.15}$$

where I_{n-1} denotes the $(n-1)\times(n-1)$ identity matrix. Equating columns and using the symmetry of $\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}$, we obtain that

$$M\frac{\partial v_\ell}{\partial \mathbf{t}} = M\frac{\partial \mathbf{v}}{\partial t_\ell} = \left(\det \frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right)\mathbf{e}_\ell, \tag{2.16}$$

and so we can rewrite (2.12) as

$$\left\{\frac{\partial^2}{\partial s^2} + \frac{\partial}{\partial \mathbf{t}'}kM\frac{\partial}{\partial \mathbf{t}}\right\} v_\ell = 0, \qquad 2\le \ell \le n. \tag{2.17}$$

This exhibits a divergence form quasilinear system satisfied in the weak sense by $\mathbf{v} = (v_2, ..., v_n)'$, which is now elliptic when $k>0$ is independent of u and Du (see the appendix). This system is not however diagonal in the principal terms since the coefficients in the matrix M involve first order derivatives of the unknowns v_ℓ. Moreover, the system (2.17) is never strongly elliptic (see the appendix). We will denote the nonlinear operator in braces in (2.17) by $\mathcal{L}$ when there is no possibility of confusion.

Remark 2.1. *When $k = k(x)$ in (1.1) is independent of u and Du, the function z satisfies the equation*

$$\mathcal{L}z \equiv \left\{\frac{\partial^2}{\partial s^2} + \frac{\partial}{\partial \mathbf{t}'}kM\frac{\partial}{\partial \mathbf{t}}\right\} z = \frac{\partial k}{\partial x_1}\det \frac{\partial \mathbf{v}}{\partial \mathbf{t}'}.$$

See the appendix below for a proof (we will not use this equation in our application).

2.1.4. *A diagonal strongly elliptic system for first derivatives.* We will obtain a quasilinear system that is diagonal in the principal terms, and strongly elliptic when $k>0$, by differentiating the system (2.17). Thus we now assume as well that $u\in C^{1,1}\cap W^{3,2}$, so that $\mathbf{v}\in C^{0,1}\cap W^{2,2}$. However, if we simply proceed directly with applying a derivative $\partial_\sigma = \frac{\partial}{\partial t_\sigma}$, $t_1 = s$, to (2.17), we obtain

$$\mathcal{L}(\partial_\sigma v_\ell) = -\frac{\partial}{\partial \mathbf{t}'}(\partial_\sigma(kM))\frac{\partial v_\ell}{\partial \mathbf{t}}, \qquad 2\le \ell\le n, \tag{2.18}$$

and since the entries of M are homogeneous polynomials of degree $n-2$ in the derivatives of the v_ℓ, the right side appears to contain second order derivatives of the $\partial_\sigma v_\ell$ that arise when $\frac{\partial}{\partial \mathbf{t}'}$ hits $\partial_\sigma M$. However, we can make use of a further crucial property of the matrix M, namely that its columns are divergence free in the $\mathbf{t}$ variables:

$$\frac{\partial}{\partial \mathbf{t}'}M = (0,...,0). \tag{2.19}$$

This is a consequence of the equality of mixed second order partial derivatives of $\mathbf{v}$ - see the appendix.

Armed with (2.19), we continue the calculation in (2.18) to obtain

$$\begin{aligned}
\mathcal{L}(\partial_\sigma v_\ell) &= -\frac{\partial}{\partial \mathbf{t}'}(\partial_\sigma(kM))\frac{\partial v_\ell}{\partial \mathbf{t}} \\
&= -trace\left[\{k(\partial_\sigma M) + (\partial_\sigma k)M\}\frac{\partial}{\partial \mathbf{t}}\frac{\partial v_\ell}{\partial \mathbf{t}'}\right] \\
&\quad -\left\{\left(\frac{\partial}{\partial \mathbf{t}'}k\right)(\partial_\sigma M) + \left(\frac{\partial}{\partial \mathbf{t}'}\partial_\sigma k\right)M\right\}\frac{\partial v_\ell}{\partial \mathbf{t}} \\
&= f_\sigma^\ell,
\end{aligned}$$

for $2 \leq \ell \leq n$, since both of the row vectors $\frac{\partial}{\partial \mathbf{t}'} M$ and $\frac{\partial}{\partial \mathbf{t}'}(\partial_\sigma M) = \partial_\sigma\left(\frac{\partial}{\partial \mathbf{t}'} M\right)$ vanish by (2.19). If we write

$$\mathbf{p} = \left(p^i_j\right)_{2\leq i\leq n, 1\leq j\leq n} = \left(\frac{\partial}{\partial t_j} v_i\right)_{2\leq i\leq n, 1\leq j\leq n},$$

we note that (using (2.9) and (2.10))

$$\mathbf{f} = \left(f^\ell_\sigma\right)_{2\leq \ell\leq n, 1\leq\sigma\leq n} = k\mathcal{A}\left(\mathbf{p}, D\mathbf{p}\right) + (\nabla k)' \mathcal{B}\left(\mathbf{p}, D\mathbf{p}\right) + \mathcal{C}\left(\mathbf{p}\right) \tag{2.20}$$

where ∇k denotes the gradient of k with respect to its original variables; $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ have entries that are polynomials in $\mathbf{p}$ and $D\mathbf{p}$ of degree depending on n, with coefficients that are smooth functions of $(s, \mathbf{t})$, $\mathbf{v}$, r and z; $\mathcal{A}(\mathbf{p}, D\mathbf{p})$ is homogeneous of degree *two* in $D\mathbf{p}$, and $\mathcal{B}(\mathbf{p}, D\mathbf{p})$ is homogeneous of degree *one* in $D\mathbf{p}$.

Thus for $u \in C^{1,1} \cap W^{3,2}$ we obtain that $\left(p^i_j\right)_{2\leq i\leq n, 1\leq j\leq n}$ is a weak solution of the *diagonal* (in the sense that the principal terms act diagonally) divergence form quasilinear system,

$$\mathcal{L}p^i_j = \left\{\frac{\partial^2}{\partial s^2} + \frac{\partial}{\partial \mathbf{t}'} kM \frac{\partial}{\partial \mathbf{t}}\right\} p^i_j = f^i_j, \quad 2 \leq i \leq n, 1 \leq j \leq n, \tag{2.21}$$

where M is the transposed cofactor matrix of $\left[p^i_j\right]_{2\leq i,j\leq n}$ and $\mathbf{f} = \left(f^i_j\right)_{2\leq i\leq n, 1\leq j\leq n} = \mathbf{f}(x, \mathbf{p}, D\mathbf{p})$ is the polynomial of degree two in $D\mathbf{p}$ as given in (2.20) above. We also have that $k = k((s, \mathbf{v}), r, (z, \mathbf{t}))$ where r and z satisfy the compatibility conditions,

$$\begin{aligned} \frac{\partial r}{\partial (s, \mathbf{t}')} &= (z, 0, ..., 0) + \mathbf{t}' \frac{\partial \mathbf{v}}{\partial (s, \mathbf{t}')}, \\ \frac{\partial z}{\partial (s, \mathbf{t}')} &= \left(k \det \frac{\partial \mathbf{v}}{\partial \mathbf{t}'}, -\frac{\partial \mathbf{v}'}{\partial s}\right). \end{aligned} \tag{2.22}$$

Furthermore, the system (2.21) is strongly elliptic when $k > 0$ and independent of u and Du (see the appendix).

3. Regularity of solutions to degenerate Monge-Ampère equations

We will investigate regularity of solutions $u \in C^{1,1} \cap W^{3,2}$ to the n-dimensional Monge-Ampère equation (2.7) via the partial Legendre transform (2.8), and the resulting system (2.21), which we abbreviate

$$\mathcal{L}\mathbf{p} = \mathbf{f}, \tag{3.1}$$

where

$$\begin{aligned} \mathcal{L}\mathbf{p} &= \left\{\frac{\partial^2}{\partial s^2} + \frac{\partial}{\partial \mathbf{t}'} kM(\mathbf{p}) \frac{\partial}{\partial \mathbf{t}}\right\} \mathbf{p}, \\ \mathbf{p} &= \left(p^i_j\right)_{2\leq i\leq n, 1\leq j\leq n}, \\ \mathbf{f} &= \left(f^i_j\right)_{2\leq i\leq n, 1\leq j\leq n}. \end{aligned}$$

Here $M = M(\mathbf{p})$ is the transposed cofactor matrix of $\left[p^i_j\right]_{2\leq i,j\leq n}$, also given by

$$M = \left(\det \frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right)\left[\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right]^{-1} = \left(\det \frac{\partial \mathbf{t}}{\partial \mathbf{v}'}\right)^{-1}\left[\frac{\partial \mathbf{t}}{\partial \mathbf{v}'}\right] = \frac{1}{d}\begin{bmatrix} \frac{\partial^2 u}{\partial x_2^2} & \cdots & \frac{\partial^2 u}{\partial x_2 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 u}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 u}{\partial x_n^2} \end{bmatrix},$$

and

$$k = k\left(\left(s, \mathbf{v}\right), r, \left(z, \mathbf{t}\right)\right),$$

where $\mathbf{v}, r, z$ satisfy the compatibility conditions (see (2.22))

$$\begin{aligned} \frac{\partial \mathbf{v}}{\partial\left(s, \mathbf{t}'\right)} &= \left(p_j^i\right)_{2 \leq i \leq n, 1 \leq j \leq n}, \\ \frac{\partial r}{\partial\left(s, \mathbf{t}'\right)} &= \left(z, 0, ..., 0\right) + \mathbf{t}'\left(p_j^i\right)_{2 \leq i \leq n, 1 \leq j \leq n}, \\ \frac{\partial z}{\partial\left(s, \mathbf{t}'\right)} &= \left(k \det\left[p_j^i\right]_{2 \leq i, j \leq n}, -p_1^2, ..., -p_1^n\right). \end{aligned} \tag{3.2}$$

Finally, we recall from (2.20) that

$$\mathbf{f} = k\mathcal{A}\left(\mathbf{p}, D\mathbf{p}\right) + \left(\nabla k\right)' \mathcal{B}\left(\mathbf{p}, D\mathbf{p}\right) + \mathcal{C}\left(\mathbf{p}\right),$$

where the polynomial $\mathcal{A}\left(\mathbf{p}, D\mathbf{p}\right)$ (respectively $\mathcal{B}\left(\mathbf{p}, D\mathbf{p}\right)$) is homogeneous of degree two (respectively one) in $D\mathbf{p}$, with coefficients that are smooth functions of $\left(s, \mathbf{t}\right), \mathbf{v}, r$ and z.

In dimension $n = 2$, Guan [9] has proved Theorem 1.1 for $\psi\left(x_1, x_2\right) = B\left|x_2\right|^{2\ell}$, $\ell \geq m$, $B \geq 0$, under the weaker regularity hypothesis $u \in C^{1,1}$ (his proof also works when $\psi\left(x_1, x_2\right)^{\frac{1}{2m}}$ is Lipschitz). In this case, the 1×1 matrix M is simply 1, and the scalar equation $\mathcal{L}v = \left\{\frac{\partial^2}{\partial s^2} + \frac{\partial}{\partial t} k \frac{\partial}{\partial t}\right\} v = 0$ can be mollified by convolution, prior to being differentiated, by the Commutator Lemma in [11], which requires only v Lipschitz, i.e. $u \in C^{1,1}$. This permits Moser-type arguments to be applied to the equation for $\mathbf{p}$ as in [5] since the right side $\mathbf{f}$ of (3.1) is now linear, and no longer quadratic, in $D\mathbf{p}$. However, in view of the higher dimensional $C^{1,1}$ *a priori* estimates in [10] and [12], it would be desirable to extend Theorem 1.1 to $u \in C^{1,1}\left(\Omega\right)$ as well, but this remains an open question.

3.1. **Subelliptic preliminaries.** In proving Theorem 1.1, we will follow the approach in Guan [9] with three differences. First, we use the higher dimensional partial Legendre transform described above in place of the classical two dimensional transform. The necessity of using the differentiated system (3.1) in place of (2.17) accounts for the extra degree of regularity in our hypothesis $u \in C^{2,1}$. Second, we will use the generalization to higher dimensions in [21] of Franchi's two dimensional result [5] on Hölder continuity of subelliptic equations with rough coefficients. Third, we will prove an extension to more general equations of Guan's hypoellipticity theorem for subelliptic quasilinear divergence form equations [9]. In order to state these results we need some definitions.

Definition 3.1. *Let $A\left(x\right) = \left[a_{ij}\left(x\right)\right]_{i,j=1}^n$ be a symmetric nonnegative matrix with bounded measurable coefficients defined in a domain $\Omega \subset \mathbb{R}^n$. We say that a vector field $T = \sum_{i=1}^n \alpha_i\left(x\right) \frac{\partial}{\partial x_i}$, with bounded coefficients α_i, is* subunit *with respect to $A\left(x\right)$ in Ω if*

$$\left(\sum_{i=1}^n \alpha_i\left(x\right) \xi_i\right)^2 \leq \xi' A\left(x\right) \xi, \qquad x \in \Omega, \xi \in \mathbb{R}^n.$$

Definition 3.2. *Let* $A(y) = [a_{ij}(y)]_{i,j=1}^n$ *be a symmetric nonnegative Lipschitz matrix defined in a domain* $\Omega \subset \mathbb{R}^N$*. We say that* $A(y)$ *is* subordinate *in* Ω *if*

$$\sum_{j=1}^n \left(\sum_{i=1}^n \frac{\partial}{\partial y_\ell} a_{ij}(y) \xi_i \right)^2 \leq C \xi' A(y) \xi, \qquad y \in \Omega, \xi \in \mathbb{R}^n, 1 \leq \ell \leq N. \tag{3.3}$$

If we denote the (symmetric) matrix $\left[\frac{\partial}{\partial y_\ell} a_{ij}(y)\right]_{i,j=1}^n$ by $\partial_\ell A(y)$, then the left side of (3.3) is

$$|\partial_\ell A(y) \xi|^2 = \xi' \partial_\ell A(y)' \partial_\ell A(y) \xi,$$

and so (3.3) can be rephrased as

$$(\partial_\ell A(y))^2 \preceq CA(y), \qquad y \in \Omega, \tag{3.4}$$

where $B \preceq A$ means $A - B$ is nonnegative semidefinite. The scalar case $n = 1$ of (3.4) is well-known to hold with $C = \|A\|_{C^2}$ (see e.g. the section on interpolation inequalities in the appendix to [20]), and persists for diagonal matrices, but fails in the general matrix-valued case for $n \geq 2$, as evidenced by $A(y) = \begin{bmatrix} 1 & y \\ y & y^2 \end{bmatrix}$, $-1 < y < 1$. We will use (3.3) mainly when $N = n$, in which case $A(x)$ is subordinate in Ω if and only if there is $c > 0$ such that the vector fields asssociated to the rows of $\frac{\partial}{\partial x_\ell} A(x)$, namely $c \sum_{i=1}^n \frac{\partial}{\partial x_\ell} a_{ij}(x) \frac{\partial}{\partial x_i}$, are subunit with respect to $A(x)$ in Ω for $1 \leq j \leq n, 1 \leq \ell \leq n$. Another fact used below is that for any symmetric nonnegative matrix $A(y)$, we have

$$(A(y))^2 \preceq CA(y), \qquad y \in \Omega, \tag{3.5}$$

where C is the supremum over Ω of the maximum eigenvalue of A (this is easily seen by diagonalizing A). In the case $N = n$, (3.5) is equivalent to the vector fields $c \sum_{i=1}^n a_{ij}(x) \frac{\partial}{\partial x_i}$ being subunit with respect to $A(x)$ in Ω for some constant $c > 0$, $1 \leq j \leq n$. In [9], condition (3.3) is referred to as subunit, but we prefer subordinate so as not to conflict with subunit vector field.

Definition 3.3. *Let* $A(x) = [a_{ij}(x)]_{i,j=1}^n$ *be a symmetric nonnegative semidefinite matrix with bounded measurable coefficients defined in a domain* $\Omega \subset \mathbb{R}^n$*. We say that*

$$L = \nabla' A(x) \nabla = \sum_{i,j=1}^n \frac{\partial}{\partial x_i} a_{ij}(x) \frac{\partial}{\partial x_j}$$

is α-subelliptic *in* Ω *for* $\alpha > 0$*, if there is a positive function* $C(\cdot,\cdot,\cdot,\cdot,\cdot)$ *defined on* $\mathcal{P}(\Omega) \times [0,\infty)^4$*, increasing in each variable separately, such that for all* m*-tuples* $\mathbf{T} = (T_1, ..., T_m)$ *of bounded subunit (with respect to* $A(x)$*) vector fields, all bounded functions* $f, \mathbf{g}$*, and all compact subsets* K *of* Ω*,* every *weak solution* $u \in W^{1,2}(\Omega)$ *to the divergence form equation*

$$Lu = f + \mathbf{T}'\mathbf{g},$$

satisfies

$$\|u\|_{C^{0,\alpha}(K)} \leq C(K, \|u\|_2, \|f\|_\infty, \|\mathbf{g}\|_\infty, m).$$

Here $\mathbf{T}'$ *denotes the transpose of* $\mathbf{T}$.

Definition 3.4. *We say that $L = \nabla' A(x)\nabla$ is α-*elliptic extendible *in Ω for $\alpha > 0$ if for every x_0 and Ω_1 with $x_0 \in \Omega_1 \Subset \Omega$, there is a symmetric smooth nonnegative subordinate matrix $B(x)$ in Ω that vanishes in a neighbourhood $\mathcal{N} \Subset \Omega_1$ of x_0, is elliptic in $\Omega - \Omega_1$, and such that*

$$L_\varepsilon = \nabla' (A(x) + B(x) + \varepsilon I)\nabla$$

is α-subelliptic in Ω, uniformly in $0 < \varepsilon < 1$.

We will need the following extension of a theorem in [9].

Theorem 3.1. *Suppose that $\mathbf{p} = (p_\ell)_{1\le\ell\le N}, \mathbf{v} = (v_\ell)_{1\le\ell\le N_0} \in C^{0,1}(\Omega)$ and that $\mathbf{p}$ is a weak solution of the system*

$$\left\{\sum_{i,j=1}^{n} \frac{\partial}{\partial x_i} a_{ij}(x,\mathbf{v},\mathbf{p}) \frac{\partial}{\partial x_j}\right\} p_\ell = h_\ell(x,\mathbf{v},\mathbf{p},D\mathbf{p}), \qquad 1 \le \ell \le N, \tag{3.6}$$

where $a_{ij} \in C^\infty(\Gamma)$, where Γ is a subdomain of $\Omega \times \mathbb{R}^{N_0} \times \mathbb{R}^N$, $A(x,\mathbf{v},\mathbf{p}) = [a_{ij}(x,\mathbf{v},\mathbf{p})]_{ij=1}^n$ is symmetric, nonnegative semidefinite, and subordinate in relatively compact subdomains of Γ, $\mathbf{h} = (h_\ell)_{1\le\ell\le N} \in C^\infty\left(\Gamma \times \mathbb{R}^{nN}\right)$ and where

$$D\mathbf{v} = \Psi(x,\mathbf{v},\mathbf{p}),$$

for $\Psi \in C^\infty(\Gamma)$. Let $\widetilde{L} = \nabla' \widetilde{A}(x)\nabla = \sum_{i,j=1}^n \frac{\partial}{\partial x_i}\widetilde{a_{ij}}(x)\frac{\partial}{\partial x_j}$ be the scalar linear operator with $\widetilde{a_{ij}}(x) = a_{ij}(x,\mathbf{v}(x),\mathbf{p}(x))$. Suppose that $\widetilde{L}$ is α-elliptic extendible in Ω for some $\alpha > 0$, that

$$trace\,[\widetilde{a_{ij}}]_{ij=1}^n \ge c > 0 \quad in\ \Omega, \tag{3.7}$$

and that $\mathbf{h}$ has the product decomposition

$$h_\ell(x,\mathbf{v},\mathbf{p},D\mathbf{p}) = H_{\ell,0}(x,\mathbf{v},\mathbf{p}) + \sum_{\mu=1}^{M} H_{\ell,\mu}(x,\mathbf{v},\mathbf{p})\,\Phi_{\ell,\mu}(x,\mathbf{v},\mathbf{p},D\mathbf{p}), \qquad 1 \le \ell \le N,$$

with $H_{\ell,\mu}$ and $\Phi_{\ell,\mu}$ smooth functions of their arguments, and where the vector fields

$$H_{\ell,\mu}(x,\mathbf{v}(x),\mathbf{p}(x))\frac{\partial}{\partial x_k} \tag{3.8}$$

are subunit with respect to $\widetilde{A}$ for $1 \le \mu \le M, 1 \le \ell \le N, 1 \le k \le n$. Then $\mathbf{p}$ and $\mathbf{v}$ are both smooth in Ω.

In [9], this result was stated and proved only for scalar equations with smooth right side $h(x)$ in (3.6), but the extension of the proof to systems of the above form is relatively straightforward, though technical, and we give the details in a subsection below. We also note that in [11], Guan eliminated the problematic hypothesis of elliptic extendibility. However, this does not seem to apply to our situation where $D\mathbf{p}$ can enter the right side nonlinearly.

The theorem from [21] that we will need extends the two dimensional result of Franchi in [5] as follows. The reverse Hölder norm $\|a\|_{RH_\infty}$ of a nonnegative function a on the real line is given by the least constant C such that

$$\sup_{s\in I} a(s) \le C\frac{1}{|I|}\int_I a(s)\,ds,$$

for all intervals I.

Theorem 3.2. *([21]) Suppose $a_{ij}(x) \in L^{\infty}(\Omega)$, and $A(x) = [a_{ij}(x)]_{ij=1}^{n}$ satisfies*

$$c\left(\xi_1^2 + a(x)^2 \|\boldsymbol{\xi}\|^2\right) \leq \xi' A(x)\, \xi \leq C\left(\xi_1^2 + a(x)^2 \|\boldsymbol{\xi}\|^2\right),$$

for $x \in \Omega$ and $(\xi_1, \boldsymbol{\xi}) \in \mathbb{R}^n$, where a satisfies $\|a\|_{C^{0,1}} \leq C$, $\|a(\cdot, \mathbf{x})\|_{RH_\infty} \leq C$ and the nondegeneracy condition $\|a(\cdot, \mathbf{x})\|_{L^\infty} \geq c > 0$, for $\mathbf{x} = (x_2, ..., x_n) \in \mathbb{R}^{n-1}$. Then $L = \nabla' A(x) \nabla$ is α-subelliptic in Ω for some $\alpha > 0$ depending only on the constants c, C above.

Remark 3.1. *The special case of Theorem 3.2 that is needed here is when*

$$a(x) = \sqrt{|x_1|^{2m} + \psi(x)},$$

where $\psi \geq 0$ and $\psi(x)^{\frac{1}{2m}}$ is Lipschitz (see (3.10) below). In [6], Franchi and Lanconelli have obtained the Hölder continuity of weak solutions u to $\nabla' A(x) \nabla u = 0$ for coefficients such as $a(x) = |x_1|^m$. Of course more general right-hand sides are also required by the definition of α-subelliptic.

3.2. **Proof of regularity for the Monge-Ampère equation.** We now prove Theorem 1.1. Let $(s, \mathbf{v}) = (x_1, \mathbf{x})$, $(z, \mathbf{t}) = \left(\frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial \mathbf{x}}\right)$ and $r = u$ be the variables in the partial Legendre transform T discussed in the sections above, and let

$$A = A((s, \mathbf{t}), \mathbf{v}, r, z, \mathbf{p}) = \begin{bmatrix} 1 & 0 \\ 0 & k((s, \mathbf{v}), r, (z, \mathbf{t}))\, M(\mathbf{p}) \end{bmatrix}.$$

Then with $t_1 = s$,

$$\mathbf{p} = \left(p_j^i\right)_{2 \leq i \leq n, 1 \leq j \leq n} = \left(\frac{\partial v_i}{\partial t_j}\right)_{2 \leq i \leq n, 1 \leq j \leq n}$$

is a $C^{0,1}(\Omega)$ weak solution of the system (3.1), which can be written $\nabla' A \nabla \mathbf{p} = \mathbf{f}$, in $\Omega' \equiv T\Omega$. In order to apply Theorem 3.1 with x there replaced by $(s, \mathbf{t})$, we consider the linear operator

$$\nabla' \widetilde{A} \nabla = \frac{\partial^2}{\partial s^2} + \frac{\partial}{\partial \mathbf{t}'} k((s, \mathbf{v}(s, \mathbf{t})), r(s, \mathbf{t}), (z(s, \mathbf{t}), \mathbf{t}))\, M(\mathbf{p}(s, \mathbf{t})) \frac{\partial}{\partial \mathbf{t}}$$

where k, M, $\mathbf{v}$, r and z are as in (3.1). Here $\widetilde{A}$ is given by

$$\widetilde{A}(s, \mathbf{t}) = \begin{bmatrix} 1 & 0 \\ 0 & k((s, \mathbf{v}(s, \mathbf{t})), r(s, \mathbf{t}), (z(s, \mathbf{t}), \mathbf{t}))\, M(\mathbf{p}(s, \mathbf{t})) \end{bmatrix},$$

and the function k satisfies

$$k = k((s, \mathbf{v}), r, (z, \mathbf{t})) \approx \left(|s|^{2m} + \psi(s, \mathbf{v})\right) K((s, \mathbf{v}), r, (z, \mathbf{t})). \tag{3.9}$$

We now establish the hypotheses of Theorem 3.1. The lower bound $c = 1$ on the trace (3.7) is obvious. The compatibility conditions (3.2) imply that

$$D(\mathbf{v}, r, z) = \Psi((s, \mathbf{t}), \mathbf{v}, r, z, \mathbf{p}),$$

with Ψ smooth (note that we replace $\mathbf{v}$ in Theorem 3.1 with the vector $(\mathbf{v}, r, z)$). Note that since M is positive definite, the quadratic form of A has the lower bound

$$(\zeta, \boldsymbol{\xi})' A((s, \mathbf{t}), \mathbf{v}, r, z, \mathbf{p})(\zeta, \boldsymbol{\xi}) \geq \zeta^2 + ck((s, \mathbf{v}), r, (z, \mathbf{t})) |\boldsymbol{\xi}|^2.$$

The standard inequality $|\nabla k| \leq C\sqrt{k}$ (see e.g. the appendix in [20]) now shows that $A((s, \mathbf{t}), \mathbf{v}, r, z, \mathbf{p})$ is subordinate in relatively compact subregions of its domain. Formula (2.20) yields the desired product decomposition for $\mathbf{h}$, and together with the inequality $|\nabla k| \leq C\sqrt{k}$, (2.20) yields as well the subunit property of the vector fields in (3.8). Thus in order to apply Theorem 3.1, it only remains to prove that $\nabla' \widetilde{A} \nabla$ is α-elliptic extendible in Ω'.

So fix a point $(s_0, \mathbf{t}_0) \in \Omega'$. We follow the argument of Guan [9]. Without loss of generality, we may suppose that $k = 0$ at $(s_0, \mathbf{t}_0)$ and that in fact $(s_0, \mathbf{t}_0) = (0, \mathbf{0})$. We choose $\delta > 0$ sufficiently small and a smooth nonnegative function $\eta(\mathbf{t})$, such that $\eta = 0$ for $|\mathbf{t}| < \delta$, $\eta > 0$ for $|\mathbf{t}| > 2\delta$ and $\eta(\mathbf{t})^{\frac{1}{2m}}$ is Lipschitz (i.e. all zeroes of η vanish to order at least $2m$). Then we define

$$B = \begin{bmatrix} 0 & 0 \\ 0 & \eta(\mathbf{t}) I_{n-1} \end{bmatrix},$$

where I_{n-1} denotes the $(n-1) \times (n-1)$ identity matrix. Clearly the operator $\nabla' \left(\widetilde{A} + B \right) \nabla$ is elliptic in $\Omega' - \overline{B_{3\delta}}$ since $|s|^{2m} + \eta(\mathbf{t})$ is positive there, and $\nabla' \left(\widetilde{A} + B \right) \nabla = \nabla' \widetilde{A} \nabla = \widetilde{\mathcal{L}}$ in B_δ. Here $B_\delta = B((0, \mathbf{0}), \delta)$ is the ball centred at $(0, \mathbf{0})$ with radius δ. The inequality $|\nabla \eta| \leq C\sqrt{\eta}$ shows that B is subordinate in Ω'. We further observe using (3.9) that

$$c\left(\xi_1^2 + a_\varepsilon(s,\mathbf{t})^2 \|\boldsymbol{\xi}\|^2\right) \leq \xi' \left(\widetilde{A} + B + \varepsilon I \right) \xi \leq C\left(\xi_1^2 + a_\varepsilon(s,\mathbf{t})^2 \|\boldsymbol{\xi}\|^2\right)$$

for $0 \leq \varepsilon < 1$, where

$$a_\varepsilon(s,\mathbf{t}) = \sqrt{|s|^{2m} + \psi(s, \mathbf{v}(s,\mathbf{t})) + \eta(\mathbf{t}) + \varepsilon} \tag{3.10}$$

since M is positive definite and K is positive in Ω'. We now claim that $a_\varepsilon(s,\mathbf{t})$ satisfies the hypotheses of Theorem 3.2 uniformly in $0 \leq \varepsilon < 1$, namely that

$$\begin{aligned} \|a_\varepsilon\|_{Lip1} &\leq C, \\ \|a_\varepsilon(\cdot,\mathbf{t})\|_{RH_\infty} &\leq C, \qquad \mathbf{t} \in \mathbb{R}^{n-1}, \\ \|a_\varepsilon(\cdot,\mathbf{t})\|_{L^\infty} &\geq c > 0, \qquad \mathbf{t} \in \mathbb{R}^{n-1}. \end{aligned} \tag{3.11}$$

With this established, Theorem 3.2 completes the proof that $\nabla' \widetilde{A} \nabla$ is α-elliptic extendible in Ω'. Then Theorem 3.1 (with $\mathbf{v}$ replaced by $(\mathbf{v}, r, z)$) shows that $\mathbf{p}$, $\mathbf{v}$, r and z are smooth in Ω'. Since $\det \frac{\partial(s,\mathbf{t})}{\partial(x_1,\mathbf{x}')} = \frac{\partial \mathbf{t}}{\partial \mathbf{x}'} = d > 0$ by (1.6), we conclude that $u = r$ is smooth in Ω, and this completes the proof of Theorem 1.1.

So it remains to prove (3.11). It is enough to prove the case $\varepsilon = 0$ since we may replace ψ by $\psi + \varepsilon$. We now write $a(s,\mathbf{t})$ for $a_\varepsilon(s,\mathbf{t})$. The first inequality in (3.11) follows immediately from the fact that $\mathbf{v}$ is Lipschitz, since then so also are the functions $|s|$, $\psi(s, \mathbf{v}(s,\mathbf{t}))^{\frac{1}{2m}}$ and $\eta(\mathbf{t})^{\frac{1}{2m}}$, and hence their ℓ^{2m} length as a vector in $\mathbb{R}^3$; a is the m^{th} power of this length. The RH_∞ inequality,

$$\sup_{s \in I} a(s,\mathbf{t}) \leq C \frac{1}{|I|} \int_I a(s,\mathbf{t})\, ds,$$

for all intervals I and points $\mathbf{t}$, is easier to check separately in the two cases

$$\begin{aligned} \sup_{s \in I} |s|^m &\geq \sup_{s \in I} \sqrt{\widetilde{\psi}(s,\mathbf{t})}, \\ \sup_{s \in I} |s|^m &\leq \sup_{s \in I} \sqrt{\widetilde{\psi}(s,\mathbf{t})}, \end{aligned}$$

where we have set $\widetilde{\psi}(s,\mathbf{t})=\psi(s,\mathbf{v})+\eta(\mathbf{t})$. Indeed, in the first case

$$\sup_{s\in I} a(s,\mathbf{t})\leq C\sup_{s\in I}|s|^m\leq C\frac{1}{|I|}\int_I |s|^m\,ds\leq C\frac{1}{|I|}\int_I a(s,\mathbf{t})\,ds.$$

In the second case,

$$\sup_{s\in I} a(s,\mathbf{t})\leq C\sup_{s\in I}\sqrt{\widetilde{\psi}(s,\mathbf{t})}.$$

Let $s_1\in I$ be such that $\sqrt{\widetilde{\psi}(s_1,\mathbf{t})}=\sup_{s\in I}\sqrt{\widetilde{\psi}(s,\mathbf{t})}$. Then we observe that

$$|I|^m\leq C\sup_{s\in I}|s|^m\leq C\sqrt{\widetilde{\psi}(s_1,\mathbf{t})}$$

implies

$$\widetilde{\psi}(s_1,\mathbf{t})^{\frac{1}{2m}}\geq c\,|I|\,.$$

Since $\widetilde{\psi}(s,\mathbf{t})^{\frac{1}{2m}}=[\psi(s,\mathbf{v}(s,\mathbf{t}))+\eta(\mathbf{t})]^{\frac{1}{2m}}$ is Lipschitz (this is the ℓ^{2m} length of the Lipschitz vector $\left(\psi^{\frac{1}{2m}},\eta^{\frac{1}{2m}}\right)$), we have

$$\widetilde{\psi}(s,\mathbf{t})^{\frac{1}{2m}}\geq\frac{1}{2}\widetilde{\psi}(s_1,\mathbf{t})^{\frac{1}{2m}}$$

for s in an interval J of length at least $c\,|I|$ that contains s_1 and is contained in I. Then we conclude,

$$\begin{aligned}\frac{1}{|I|}\int_I a(s,\mathbf{t})\,ds &\geq \frac{1}{|I|}\int_J\sqrt{\widetilde{\psi}(s,\mathbf{t})}ds\\ &\geq \frac{|J|}{|I|}\sqrt{\frac{1}{2^{2m}}\widetilde{\psi}(s_1,\mathbf{t})}\\ &\geq c\frac{|J|}{|I|}\sup_{s\in I} a(s,\mathbf{t})\\ &\geq c\sup_{s\in I} a(s,\mathbf{t})\,.\end{aligned}$$

The nondegeneracy inequality in (3.11) follows from $a(s,\mathbf{t})\geq|s|^m$.

3.3. **Hypoellipticity of the quasilinear system.** We now give the proof of Theorem 3.1. We will treat each equation in (3.6) as a scalar equation in the unknown $p_\ell\in C^{0,1}$ with $\widetilde{a_{ij}}(x)\in C^{0,1}$, and as in [9], approximate p_ℓ by solutions to elliptic linear Dirichlet problems which have better regularity properties. Then we use the argument in [9] to show that $p_\ell\in C^{1,\alpha-\delta}$, $1\leq\ell\leq N$, $\delta>0$, and thus that $\widetilde{a_{ij}}(x)\in C^{1,\alpha-\delta}$. Finally we use the commutator lemma in [9], which we reproduce below, to show that if $\widetilde{a_{ij}}(x)\in C^{m,\beta}$, then $p_\ell\in C^{m,\beta+\alpha-\delta}$, $1\leq\ell\leq N$, $\delta>0$, and thus that $\widetilde{a_{ij}}(x)\in C^{m,\beta+\alpha-\delta}$. Iterating this argument and differentiating the equation as necessary will complete the proof. The details follow the statement of Guan's commutator lemma.

Definition 3.5. *We let $\Lambda^t(\mathbb{R}^n)$ denote the Hölder-Zygmund spaces for $t\in\mathbb{R}$ (see [23] where these spaces are denoted C_*^t), and set*

$$C_*^t=\begin{cases}\Lambda^t(\mathbb{R}^n)=C^{m,\alpha}(\mathbb{R}^n) & \text{for } t=m+\alpha>0,\quad m\in\mathbb{N}\cup\{0\},\quad 0<\alpha<1\\ C^{m-1,1}(\mathbb{R}^n) & \text{for } t=m,\quad m\in\mathbb{N}\end{cases}.$$

We denote by O_t^λ *the collection of linear operators bounded from* $\Lambda^{s+\lambda}_{compact}(\mathbb{R}^n)$ *to* $\Lambda^s_{loc}(\mathbb{R}^n)$ *for all* $0 < s \le t$.

Lemma 3.3. *(Guan [9]) Let* $a(x,\mathbf{u}) \in C^\infty(\mathbb{R}^n \times \mathbb{R}^N)$ *vanish for* x *outside a compact subset of* $\mathbb{R}^n$*, and suppose* $\mathbf{u}(x) \in C^t_*$ *for some* $t \ge 1$*. Then for every* $0 < s < t$ *and* $\delta > 0$*, there are operators* $B_0 \in O^{s-2+\delta}_{t-s}$*,* $\mathbf{B} \in O^{s-1+\delta}_{t-s}$ *and* $B_j \in O^{s-1+\delta}_{t-s}$*,* $1 \le j \le n$*, such that for* $|D|^s = (-\triangle)^{\frac{s}{2}}$*,*

$$[|D|^s, a(x,\mathbf{u}(x))] = \sum_{j=1}^n a_{x_j}(x,\mathbf{u}(x)) B_j + a_{\mathbf{u}}(x,\mathbf{u}(x)) \cdot \mathbf{B} + B_0.$$

This lemma was proved in [9] for u a scalar function, but the proof persists for $\mathbb{R}^N$-valued $\mathbf{u} = (u_j)_{j=1}^N$ with the vector-valued decomposition $\mathbf{u} = \mathbf{u}^\sharp + \mathbf{u}^\flat$, $\mathbf{u}^\sharp = \left(u_j^\sharp\right)_{j=1}^N$ where $u_j = u_j^\sharp + u_j^\flat$ is the usual symbol splitting as in [9].

Proof. (of Theorem 3.1) Fix $x_0 \in \Omega_0 \Subset \Omega$ where Ω_0 will be chosen sufficiently small below, and let $B(x)$ and $\mathcal{N} \Subset \Omega_1 \Subset \Omega_0$ be as in the definition of α-elliptic extendibility for $\widetilde{A}(x) = [a_{ij}(x,\mathbf{v}(x),\mathbf{p}(x))]_{i,j=1}^n$. For $\varepsilon \ge 0$, define

$$\begin{aligned} A^\varepsilon(x,\mathbf{v},\mathbf{p}) &= A(x,\mathbf{v},\mathbf{p}) + B(x) + \varepsilon I, \\ \widetilde{A^\varepsilon}(x) &= A^\varepsilon(x,\mathbf{v}(x),\mathbf{p}(x)). \end{aligned}$$

Fix $\varphi \in C_c^\infty(\mathcal{N})$ and note by (3.6) and the fact that $B(x)$ vanishes in $\mathcal{N}$, that $\mathbf{p}^0 = \varphi\mathbf{p}$ is a weak solution to the *scalar* linear Dirichlet problems

$$\left\{ \begin{array}{rcll} \nabla' \widetilde{A^0}(x) \nabla \mathbf{p}^0 &=& \mathbf{h}^0(x,\mathbf{v},\mathbf{p},D\mathbf{p}), & x \in \Omega_0 \\ \mathbf{p}^0 &=& 0, & x \in \partial\Omega_0 \end{array} \right. \tag{3.12}$$

where $\mathbf{h}^0 = \left(h_\ell^0\right)_{1 \le \ell \le N}$ is given by

$$h_\ell^0(x,\mathbf{v},\mathbf{p},D\mathbf{p}) = \varphi h_\ell(x,\mathbf{v},\mathbf{p},D\mathbf{p}) + (\nabla\varphi)' A(x,\mathbf{v},\mathbf{p}) \nabla p_\ell + \nabla' [p_\ell A(x,\mathbf{v},\mathbf{p}) \nabla\varphi].$$

By weak solution, we mean that both sides of the first equation in (3.12) can be multiplied by any test function in $W_0^{1,2}(\Omega_0)$, and then integrated over Ω_0 to yield equality. For $\varepsilon > 0$ let $\mathbf{p}^\varepsilon$ be the unique weak solution to the scalar *elliptic* linear Dirichlet problems

$$\left\{ \begin{array}{rcll} \nabla' \widetilde{A^\varepsilon}(x) \nabla \mathbf{p}^\varepsilon &=& \mathbf{h}^0(x,\mathbf{v},\mathbf{p},D\mathbf{p}), & x \in \Omega_0 \\ \mathbf{p}^\varepsilon &=& 0, & x \in \partial\Omega_0 \end{array} \right. . \tag{3.13}$$

Since $\widetilde{A^\varepsilon}(x)$ is Lipschitz and $\mathbf{h}^0(x,\mathbf{v},\mathbf{p},D\mathbf{p})$ is bounded, standard elliptic theory for scalar equations (e.g. Corollary 9.18 and Theorem 8.34 in [8]) shows that

$$\begin{aligned} \mathbf{p}^\varepsilon &\in W^{2,p}_{loc}(\Omega_0), \quad 1 < p < \infty, \\ \mathbf{p}^\varepsilon &\in C^{1,\beta}\left(\overline{\Omega_0}\right), \quad 0 < \beta < 1, \end{aligned}$$

with norms depending on $\varepsilon > 0$. Since *trace* $\left(\widetilde{A^\varepsilon}\right) \ge c > 0$ independent of $\varepsilon > 0$ in Ω by (3.7), the maximum principle yields

$$\|\mathbf{p}^\varepsilon\|_{L^\infty\left(\overline{\Omega_0}\right)} \le C, \tag{3.14}$$

independent of $\varepsilon > 0$. Indeed, if $w_\ell^\varepsilon = p_\ell^\varepsilon + C_1 |x - x_0|^2 - C_2$, we have

$$\nabla' \widetilde{A^\varepsilon} \nabla w_\ell^\varepsilon = h_\ell^0 + 2C_1 trace\left(\widetilde{A^\varepsilon}\right) + 2C_1\left(\nabla' \widetilde{A^0}\right)(\mathbf{x} - \mathbf{x}_0) \geq 1$$

in Ω_0 (note $\nabla' \widetilde{A^\varepsilon} = \nabla' \widetilde{A^0}$), provided we choose Ω_0 sufficiently small and C_1 sufficiently large, independent of $\varepsilon > 0$. We also have $w_\ell^\varepsilon \leq 0$ on $\partial\Omega_0$ by the boundary condition in (3.13) if we choose C_2 sufficiently large, independent of $\varepsilon > 0$. Theorem 8.1 in [8] now shows that $w_\ell^\varepsilon \leq 0$ in Ω_0. Arguing as above with $-p_\ell^\varepsilon$ in place of p_ℓ^ε, we obtain (3.14).

Lemma 3.4. *Let $0 < \alpha < 1$ be as in Theorem 3.1. Then for any $0 < \beta < \alpha$, we have*

$$\|\mathbf{p}^\varepsilon\|_{C^{1,\beta}\left(\overline{\Omega_0}\right)} \leq C_\beta,$$

independent of $\varepsilon > 0$, and

$$\mathbf{p} \in C^{1,\beta}_{loc}(\Omega_0).$$

Proof. Fix k and differentiate the ℓ^{th} equation in (3.13) with respect to $\frac{\partial}{\partial x_k}$ to obtain that $q_\ell^\varepsilon \equiv \frac{\partial p_\ell^\varepsilon}{\partial x_k} \in W^{1,p}_{loc}(\Omega_0)$ is a weak solution in Ω_0 of the divergence form equation

$$\begin{aligned}(3.15)\qquad \nabla' \widetilde{A^\varepsilon}(x) \nabla q_\ell^\varepsilon &= \frac{\partial}{\partial x_k}\varphi H_{\ell,0}(x, \mathbf{v}, \mathbf{p}) \\ &+ \sum_{\mu=1}^M \frac{\partial}{\partial x_k}\varphi H_{\ell,\mu}(x, \mathbf{v}, \mathbf{p})\, \Phi_{\ell,\mu}(x, \mathbf{v}, \mathbf{p}, D\mathbf{p}) \\ &+ \frac{\partial}{\partial x_k}\left[(\nabla\varphi)' \widetilde{A} \nabla p_\ell\right] + \nabla'\left\{\widetilde{A}\frac{\partial}{\partial x_k}(p_\ell \nabla\varphi) + \left(\frac{\partial}{\partial x_k}\widetilde{A}\right) p_\ell \nabla\varphi\right\} \\ &- \nabla'\left[\left(\frac{\partial}{\partial x_k}\widetilde{A^\varepsilon}\right)\nabla p_\ell^\varepsilon\right].\end{aligned}$$

Now suppose, in order to derive a contradiction, that for some $0 < \beta < \alpha$ and $1 \leq \ell \leq N$,

$$\sup_{0<\varepsilon<1} \|p_\ell^\varepsilon\|_{C^{1,\beta}\left(\overline{\Omega_0}\right)} = \infty.$$

Then there is a sequence $\{\varepsilon_j\}_{j=1}^\infty$ with $\lim_{j\to\infty} \varepsilon_j = 0$ and

$$c_j \equiv \left\|p_\ell^{\varepsilon_j}\right\|_{C^{1,\beta}\left(\overline{\Omega_0}\right)} > j, \qquad j \geq 1.$$

Now we take $\varepsilon = \varepsilon_j$ in (3.15), multiply by c_j^{-1} and rewrite the equation as

$$(3.16)\qquad \nabla' \widetilde{A^{\varepsilon_j}} \nabla\left(c_j^{-1} q_\ell^{\varepsilon_j}\right) = c_j^{-1} f_\ell^{\varepsilon_j} + \left(\mathbf{T}_\ell^{\varepsilon_j}\right)' c_j^{-1} \mathbf{g}_\ell^{\varepsilon_j},$$

where, we claim, $\mathbf{T}_\ell^{\varepsilon_j}$ is a collection of subunit vector fields with the supremum norms of $c_j^{-1} f_\ell^{\varepsilon_j}$ and $c_j^{-1}\mathbf{g}_\ell^{\varepsilon_j}$ bounded independent of j.

Indeed, the first term on the right side of (3.15) is bounded, and by (3.8), the second term is a sum of transposed subunit (with respect to $\widetilde{A^\varepsilon}$) vector fields applied to bounded functions. Since $A(x, \mathbf{v}, \mathbf{p})$ is subordinate in Γ, and $\mathbf{v}, \mathbf{p}$ are Lipschitz, the chain rule shows that $\widetilde{A}$ is subordinate in Ω_0. Thus the dot product of any row of $\frac{\partial}{\partial x_k}\widetilde{A}$ with ∇ is a multiple of a subunit (with respect to $\widetilde{A^\varepsilon}$) vector field (see the comments following (3.3)). Thus the components $\sum_{j=1}^n \left(\frac{\partial}{\partial x_k}\widetilde{a_{ij}}\right)\frac{\partial}{\partial x_j}$, $1 \leq i \leq n$, of the vector $\widetilde{A}_{x_k}\nabla$ are multiples of subunit vector fields. As a consequence, the term $c_j^{-1}\nabla'\left\{\left(\frac{\partial}{\partial x_k}\widetilde{A}\right)p_\ell\nabla\varphi\right\}$ is a sum of transposed subunit (with respect to $\widetilde{A^\varepsilon}$) vector fields applied to bounded functions. The term $c_j^{-1}\nabla'\left\{\widetilde{A}\frac{\partial}{\partial x_k}\left(p_\ell\nabla\varphi\right)\right\}$ is also of this form using (3.5). Now each of $\widetilde{A}$, B and εI is subordinate in Ω_0, hence also $\widetilde{A^\varepsilon}$, and thus each component of $\nabla'\left(\frac{\partial}{\partial x_k}\widetilde{A^{\varepsilon_j}}\right)$ is a multiple of a transposed subunit vector field. The function $c_j^{-1}\nabla p_\ell^{\varepsilon_j}$ is bounded since the definition of c_j yields

$$\left\|c_j^{-1}\nabla p_\ell^{\varepsilon_j}\right\|_{L^\infty\left(\overline{\Omega_0}\right)} \leq Cc_j^{-1}\left\|p_\ell^{\varepsilon_j}\right\|_{C^{1,\beta}\left(\overline{\Omega_0}\right)} = C.$$

Thus the term $c_j^{-1}\nabla'\left[\left(\frac{\partial}{\partial x_k}\widetilde{A^{\varepsilon_j}}\right)\nabla p_\ell^{\varepsilon_j}\right]$ is also a sum of transposed subunit vector fields applied to bounded functions. We expand the remaining term as c_j^{-1} times

$$\frac{\partial}{\partial x_k}\left[(\nabla\varphi)'\,\widetilde{A}\nabla p_\ell\right] = \left(\frac{\partial}{\partial x_k}\nabla\varphi\right)'\widetilde{A}\nabla p_\ell + (\nabla\varphi)'\left(\frac{\partial}{\partial x_k}\widetilde{A}\right)\nabla p_\ell + (\nabla\varphi)'\,\widetilde{A}\nabla\frac{\partial}{\partial x_k}p_\ell,$$

and note that the first two terms on the right are bounded. By (3.5), the third term on the right is a sum $\sum_\sigma T_\sigma g_\sigma$ of subunit (not transposed) vector fields T_σ applied to bounded functions g_σ, but where the coefficients α_j^σ of $T_\sigma = \sum_{j=1}^n \alpha_j^\sigma \frac{\partial}{\partial x_j}$ are Lipschitz functions. Thus $T_\sigma g_\sigma = -T'_\sigma g_\sigma - \sum_{j=1}^n \frac{\partial\alpha_j^\sigma}{\partial x_j}g_\sigma$ has the required form.

Now by our hypotheses, the linear operators $\nabla'\widetilde{A^\varepsilon}\nabla$ are α-subelliptic in Ω_0 (even in Ω) for $\alpha > 0$ as in Theorem 3.1, uniformly in $0 < \varepsilon < 1$, and we thus obtain from (3.16) that

$$\left\|c_j^{-1}q_\ell^{\varepsilon_j}\right\|_{C^{0,\alpha}(K)} \leq C_K,$$

independent of $j \geq 1$, for K compact in Ω_0. Together with (3.14), this yields

$$\left\|c_j^{-1}p_\ell^{\varepsilon_j}\right\|_{C^{1,\alpha}(K)} = \left\|c_j^{-1}p_\ell^{\varepsilon_j}\right\|_{L^\infty(K)} + \left\|c_j^{-1}q_\ell^{\varepsilon_j}\right\|_{C^{0,\alpha}(K)} \leq C_K,$$

independent of $j \geq 1$, for K compact in Ω_0. However, since $\widetilde{A^\varepsilon}(x)$ is elliptic in $\overline{\Omega_0} - \Omega_1$ independent of $\varepsilon > 0$, standard elliptic theory (Corollary 8.36 in [8]) applied to (3.13) in $\overline{\Omega_0} - \Omega_1$ yields

$$\|p_\ell^\varepsilon\|_{C^{1,\sigma}\left(\overline{\Omega_0}-\Omega_1\right)} \leq C_\sigma,$$

independent of $\varepsilon > 0$, for all $0 < \sigma < 1$. Combining the last two inequalities yields

$$\left\|c_j^{-1}p_\ell^{\varepsilon_j}\right\|_{C^{1,\alpha}\left(\overline{\Omega_0}\right)} \leq C$$

with a constant C that is independent of $j \geq 1$. Now $0 < \beta < \alpha$ and so there is a subsequence, which we continue to write as $\left\{c_j^{-1}p_\ell^{\varepsilon_j}\right\}_{j=1}^\infty$, with $\lim_{j\to\infty}\varepsilon_j = 0$ and which converges in $C^{1,\beta}\left(\overline{\Omega_0}\right)$ to a solution u_ℓ of the Dirichlet problem,

$$\begin{cases} \nabla' A^0(x,\mathbf{v},\mathbf{p})\nabla u_\ell &= 0, \quad x \in \Omega_0 \\ u_\ell &= 0, \quad x \in \partial\Omega_0 \end{cases},$$

since $\lim_{j\to\infty} c_j^{-1}h_\ell^0(x,\mathbf{v},\mathbf{p},D\mathbf{p}) = 0$. By uniqueness (Theorem 8.1 of [8]), $u_\ell = 0$ in $\overline{\Omega_0}$ and this contradicts

$$\|u_\ell\|_{C^{1,\beta}\left(\overline{\Omega_0}\right)} = \lim_{j\to\infty}\left\|c_j^{-1}p_\ell^{\varepsilon_j}\right\|_{C^{1,\beta}\left(\overline{\Omega_0}\right)} = 1,$$

which completes the proof that

$$\sup_{0<\varepsilon<1} \|p_\ell^\varepsilon\|_{C^{1,\beta}\left(\overline{\Omega_0}\right)} < \infty, \qquad 1 \le \ell \le N,\ 0 < \beta < \alpha.$$

Now if $0 < \gamma < \beta < \alpha$, there is a sequence $\{\varepsilon_j\}_{j=1}^\infty$ with $\lim_{j\to\infty} \varepsilon_j = 0$ such that $p_\ell^{\varepsilon_j}$ converges in $C^{1,\gamma}\left(\overline{\Omega_0}\right)$ to a solution $\widetilde{p_\ell^0}$ of the Dirichlet problem,

$$\left\{ \begin{array}{rcll} \nabla' A^0(x,\mathbf{v},\mathbf{p}) \nabla \widetilde{p_\ell^0} & = & h_\ell^0(x,\mathbf{v},\mathbf{p},D\mathbf{p}), & x \in \Omega_0 \\ \widetilde{p_\ell^0} & = & 0, & x \in \partial\Omega_0 \end{array} \right., \qquad 1 \le \ell \le N.$$

Since $p_\ell^0 = \varphi p_\ell$ is also a solution to this Dirichlet problem by (3.12), uniqueness yields $\varphi p_\ell = \widetilde{p_\ell^0} \in C^{1,\gamma}\left(\overline{\Omega_0}\right)$. This completes the proof of Lemma 3.4.

Thus $\mathbf{p}$ is now more regular in Ω_0, hence in Ω, and using the elliptic equation $D\mathbf{v} = \Psi(x,\mathbf{v},\mathbf{p})$, then so also are the coefficient matrix $A(x,\mathbf{v},\mathbf{p})$ and the data $\mathbf{h}(x,\mathbf{v},\mathbf{p},D\mathbf{p})$ in (3.6), namely

(3.17)
$$\mathbf{p} \in C^{1,\beta}_{loc}(\Omega), \mathbf{v} \in C^{2,\beta}_{loc}(\Omega), A(x,\mathbf{v},\mathbf{p}) \in C^{1,\beta}_{loc}(\Omega), \mathbf{h}(x,\mathbf{v},\mathbf{p},D\mathbf{p}) \in C^{0,\beta}_{loc}(\Omega),$$

for all $0 < \beta < \alpha$. Indeed, if $\mathbf{v}$ is Lipschitz and $\mathbf{p} \in C^{j,\beta}$, $j \ge 0$, then a simple induction using $D\mathbf{v} = \Psi(x,\mathbf{v},\mathbf{p})$ shows that $\mathbf{v} \in C^{j+1,\beta}$. The increased smoothness of the data in (3.17) suggests we should differentiate equation (3.6) by a fractional amount less than α. This will use the commutator lemma 3.3 of Guan stated at the beginning of this subsection.

Lemma 3.5. *Let α be as in Theorem 3.1. Suppose that for some $0 < \beta < 1$, we have $\mathbf{p} \in C^{1,\beta}_{loc}(\Omega)$ and $\sup_{0<\varepsilon<1} \|\mathbf{p}^\varepsilon\|_{C^{1,\beta}\left(\overline{\Omega_0}\right)} < \infty$. Then for $0 < \gamma < \beta$ and $0 < \delta < 1$, we have*

$$\begin{array}{rcl} \sup_{0<\varepsilon<1} \|\mathbf{p}^\varepsilon\|_{C_*^{1+\gamma+\alpha}\left(\overline{\Omega_0}\right)} & < & \infty, \\ \mathbf{p} & \in & C_*^{1+\gamma+\alpha-\delta}(\Omega_0). \end{array}$$

Proof. In order to apply the fractional differentiation operator $|D|^\gamma = (-\triangle)^{\frac{\gamma}{2}}$ to (3.15), we must first multiply the function q_ℓ^ε, which is only defined in $\overline{\Omega_0}$, by a smooth cutoff function supported in Ω_0. So let $\psi \in C_c^\infty(\Omega_0)$ and set $w_\ell^\varepsilon = |D|^\gamma \psi q_\ell^\varepsilon = |D|^\gamma \psi \frac{\partial}{\partial x_k} p_\ell^\varepsilon$. Since $\beta > 0$, (3.17) and Schauder theory applied to equation (3.13) yield that $\mathbf{p}^\varepsilon \in C^{2,\beta}\left(\overline{\Omega_0}\right)$ for $\varepsilon > 0$, and so $w_\ell^\varepsilon \in C^{1,\beta-\gamma}\left(\overline{\Omega_0}\right) \subset W^{1,2}(\Omega_0)$. In order to exploit the special form of $\widetilde{A^\varepsilon}$, we first need to write (3.15) out in full using the chain rule,

$$\begin{aligned}(3.18)\qquad & \nabla' A^{\varepsilon}(x,\mathbf{v},\mathbf{p})\nabla q_{\ell}^{\varepsilon} \equiv \mathcal{E} \\ = & \frac{\partial}{\partial x_k}\varphi H_{\ell,0}(x,\mathbf{v},\mathbf{p}) \\ & +\sum_{\mu=1}^{M}\frac{\partial}{\partial x_k}\varphi H_{\ell,\mu}(x,\mathbf{v},\mathbf{p})\Phi_{\ell,\mu}(x,\mathbf{v},\mathbf{p},D\mathbf{p}) \\ & +\frac{\partial}{\partial x_k}\left[(\nabla\varphi)' A(x,\mathbf{v},\mathbf{p})\nabla p_{\ell}\right] \\ & +\nabla'\left\{A(x,\mathbf{v},\mathbf{p})\frac{\partial}{\partial x_k}(p_{\ell}\nabla\varphi)+A_{x_k}(x,\mathbf{v},\mathbf{p})p_{\ell}\nabla\varphi\right\} \\ & +\nabla'\left\{(\nabla_{\mathbf{v}}A)(x,\mathbf{v},\mathbf{p})\frac{\partial\mathbf{v}}{\partial x_k}p_{\ell}\nabla\varphi+(\nabla_{\mathbf{p}}A)(x,\mathbf{v},\mathbf{p})\frac{\partial\mathbf{p}}{\partial x_k}p_{\ell}\nabla\varphi\right\} \\ & -\nabla'\left[\left\{A_{x_k}^{\varepsilon}(x,\mathbf{v},\mathbf{p})+(\nabla_{\mathbf{v}}A^{\varepsilon})(x,\mathbf{v},\mathbf{p})\frac{\partial\mathbf{v}}{\partial x_k}\right\}\nabla p_{\ell}^{\varepsilon}\right] \\ & -\nabla'\left[\left\{(\nabla_{\mathbf{p}}A^{\varepsilon})(x,\mathbf{v},\mathbf{p})\frac{\partial\mathbf{p}}{\partial x_k}\right\}\nabla p_{\ell}^{\varepsilon}\right].\end{aligned}$$

We will apply $|D|^{\gamma}$ to the equation

$$\begin{aligned}\nabla' A^{\varepsilon}(x,\mathbf{v},\mathbf{p})\nabla(\psi q_{\ell}^{\varepsilon}) &= \psi\nabla' A^{\varepsilon}(x,\mathbf{v},\mathbf{p})\nabla q_{\ell}^{\varepsilon} \\ &\quad +\nabla' A^{\varepsilon}(x,\mathbf{v},\mathbf{p})(\nabla\psi)q_{\ell}^{\varepsilon} \\ &\quad +\left(\nabla'\psi\right)A^{\varepsilon}(x,\mathbf{v},\mathbf{p})\nabla q_{\ell}^{\varepsilon} \\ &= \psi\mathcal{E}+\mathcal{F},\end{aligned}$$

with the term $\psi\mathcal{E}$ written as

$$\begin{aligned}&\sum_{\mu=1}^{M}\frac{\partial}{\partial x_k}\varphi H_{\ell,\mu}(x,\mathbf{v},\mathbf{p})\psi\Phi_{\ell,\mu}(x,\mathbf{v},\mathbf{p},D\mathbf{p}) \\ &+\frac{\partial}{\partial x_k}\left[\psi(\nabla\varphi)' A(x,\mathbf{v},\mathbf{p})\nabla p_{\ell}\right] \\ &+\nabla'\left\{A(x,\mathbf{v},\mathbf{p})\psi\frac{\partial}{\partial x_k}(p_{\ell}\nabla\varphi)+A_{x_k}(x,\mathbf{v},\mathbf{p})\psi p_{\ell}\nabla\varphi\right\} \\ &+\nabla'\left\{(\nabla_{\mathbf{v}}A)(x,\mathbf{v},\mathbf{p})\frac{\partial\mathbf{v}}{\partial x_k}\psi p_{\ell}\nabla\varphi+(\nabla_{\mathbf{p}}A)(x,\mathbf{v},\mathbf{p})\frac{\partial\mathbf{p}}{\partial x_k}\psi p_{\ell}\nabla\varphi\right\} \\ &-\nabla'\left[\left\{A_{x_k}^{\varepsilon}(x,\mathbf{v},\mathbf{p})+(\nabla_{\mathbf{v}}A^{\varepsilon})(x,\mathbf{v},\mathbf{p})\frac{\partial\mathbf{v}}{\partial x_k}\right\}\psi\nabla p_{\ell}^{\varepsilon}\right] \\ &-\nabla'\left[\left\{(\nabla_{\mathbf{p}}A^{\varepsilon})(x,\mathbf{v},\mathbf{p})\frac{\partial\mathbf{p}}{\partial x_k}\right\}\psi\nabla p_{\ell}^{\varepsilon}\right]+\mathcal{G},\end{aligned}$$

where $\mathcal{G}\in C^{0,\beta}(\Omega_0)$ upon using (3.17). Note that all of the above terms vanish outside the support of ψ, and thus we can multiply any of the functions there by a cutoff function $\omega\in C_c^{\infty}(\Omega_0)$ satisfying $\omega=1$ on the support of ψ. Thus we may assume everything is compactly supported and so the pseudodifferential calculus used below is justified. We obtain that for $\varepsilon > 0$, w_ℓ^ε is a weak solution in Ω_0 of

$$
\begin{aligned}
(3.19)\qquad & \nabla' \widetilde{A^\varepsilon}(x) \nabla w_\ell^\varepsilon \\
= \quad & |D|^\gamma \mathcal{F} + |D|^\gamma \mathcal{G} + \sum_{\mu=1}^{M} \frac{\partial}{\partial x_k} \varphi H_{\ell,\mu}(x, \mathbf{v}, \mathbf{p}) |D|^\gamma \psi \Phi_{\ell,\mu}(x, \mathbf{v}, \mathbf{p}, D\mathbf{p}) \\
& + \frac{\partial}{\partial x_k} trace \left\{ \widetilde{A}(x) |D|^\gamma \left(\psi \nabla p_\ell (\nabla \varphi)' \right) \right\} \\
& + \nabla' \left\{ \widetilde{A}(x) |D|^\gamma \psi \frac{\partial}{\partial x_k}(p_\ell \nabla \varphi) + A_{x_k}(x, \mathbf{v}(x), \mathbf{p}(x)) |D|^\gamma (\psi p_\ell \nabla \varphi) \right\} \\
& + \nabla' \left\{ \nabla_{\mathbf{v}} A(x, \mathbf{v}(x), \mathbf{p}(x)) |D|^\gamma \left(\frac{\partial \mathbf{v}}{\partial x_k} \psi p_\ell \nabla \varphi \right) \right\} \\
& + \nabla' \left\{ \nabla_{\mathbf{p}} A(x, \mathbf{v}(x), \mathbf{p}(x)) |D|^\gamma \left(\frac{\partial \mathbf{p}}{\partial x_k} \psi p_\ell \nabla \varphi \right) \right\} \\
& - \nabla' \left\{ A_{x_k}^\varepsilon(x, \mathbf{v}(x), \mathbf{p}(x)) |D|^\gamma \psi \nabla p_\ell^\varepsilon \right\} \\
& - \nabla' \left\{ \nabla_{\mathbf{v}} A^\varepsilon(x, \mathbf{v}(x), \mathbf{p}(x)) |D|^\gamma \left(\frac{\partial \mathbf{v}}{\partial x_k} \psi \nabla p_\ell^\varepsilon \right) \right\} \\
& - \nabla' \left\{ \nabla_{\mathbf{p}} A^\varepsilon |D|^\gamma \left(\frac{\partial \mathbf{p}}{\partial x_k} \psi \nabla p_\ell^\varepsilon \right) \right\} + \mathcal{H},
\end{aligned}
$$

where $\mathcal{H}$ is the sum of the following commutator terms (recall that everything is compactly supported in Ω_0),

$$
\begin{aligned}
(3.20)\qquad & \sum_{\mu=1}^{M} \frac{\partial}{\partial x_k} \left[|D|^\gamma, \varphi H_{\ell,\mu}(x, \mathbf{v}(x), \mathbf{p}(x)) \right] \psi \Phi_{\ell,\mu}(x, \mathbf{v}, \mathbf{p}, D\mathbf{p}) \\
& + \frac{\partial}{\partial x_k} trace \left\{ \left[|D|^\gamma, \widetilde{A} \right] \left(\psi \nabla p_\ell (\nabla \varphi)' \right) \right\} + \nabla' \left[|D|^\gamma, \widetilde{A} \right] \psi \frac{\partial}{\partial x_k}(p_\ell \nabla \varphi) \\
& + \nabla' \left[|D|^\gamma, \widetilde{A_{x_k}} \right] (\psi p_\ell \nabla \varphi) \\
& + \nabla' \left\{ \left[|D|^\gamma, \widetilde{\nabla_{\mathbf{v}} A} \right] \left(\frac{\partial \mathbf{v}}{\partial x_k} \psi p_\ell \nabla \varphi \right) + \left[|D|^\gamma, \widetilde{\nabla_{\mathbf{p}} A} \right] \left(\frac{\partial \mathbf{p}}{\partial x_k} \psi p_\ell \nabla \varphi \right) \right\} \\
& - \nabla' \left\{ \left[|D|^\gamma, \widetilde{A_{x_k}^\varepsilon} \right] \psi \nabla p_\ell^\varepsilon \right\} - \nabla' \left\{ \left[|D|^\gamma, \widetilde{\nabla_{\mathbf{v}} A^\varepsilon} \right] \left(\frac{\partial \mathbf{v}}{\partial x_k} \psi \nabla p_\ell^\varepsilon \right) \right\} \\
& - \nabla' \left\{ \left[|D|^\gamma, \widetilde{\nabla_{\mathbf{p}} A^\varepsilon} \right] \left(\frac{\partial \mathbf{p}}{\partial x_k} \psi \nabla p_\ell^\varepsilon \right) \right\} \\
& - \nabla' \left\{ \left[|D|^\gamma, A^\varepsilon(x, \mathbf{v}(x), \mathbf{p}(x)) \right] \nabla \psi \frac{\partial}{\partial x_k} p_\ell^\varepsilon \right\}.
\end{aligned}
$$

Note that we write $\widetilde{A_{x_k}}(x) = A_{x_k}(x, \mathbf{v}(x), \mathbf{p}(x))$ and similarly for $\widetilde{\nabla_{\mathbf{v}} A}$, etc. Now $\mathcal{G} \in C^{0,\beta}(\Omega_0)$, so that the second term on the right side of (3.19) lies in $C^{0,\beta-\gamma}$, and in particular is bounded. Since $\widetilde{A}$ and $\widetilde{A^\varepsilon}$ are subordinate in Ω_0, we see as in the proof of lemma 3.4, that all of the component vector fields in

$$
\varphi H_{\ell,\mu}(x, \mathbf{v}(x), \mathbf{p}(x)) \frac{\partial}{\partial x_k}, \widetilde{A} \nabla, \widetilde{A_{x_k}} \nabla, \widetilde{\nabla_{\mathbf{v}} A} \nabla, \widetilde{\nabla_{\mathbf{p}} A} \nabla, \widetilde{A_{x_k}^\varepsilon} \nabla, \widetilde{\nabla_{\mathbf{v}} A^\varepsilon} \nabla, \widetilde{\nabla_{\mathbf{p}} A^\varepsilon} \nabla,
$$

$\mu \geq 1$, are bounded and multiples of subunit vector fields with respect to $\widetilde{A^{\varepsilon}}$ (use (3.8) for the first vector field). These vector fields act on functions such as $|D|^{\gamma} \Phi_{\ell,\mu}(x, \mathbf{v}, \mathbf{p}, D\mathbf{p})$ and $|D|^{\gamma} \nabla p_{\ell}^{\varepsilon}$ which by hypothesis lie in $C^{0,\beta-\gamma}$ uniformly in $\varepsilon > 0$, hence are bounded independent of $\varepsilon > 0$. Thus the remaining terms in (3.19), apart from $|D|^{\gamma} \mathcal{F}$ and $\mathcal{H}$, have the form $\mathbf{T}'\mathbf{g}$ for $\mathbf{g}$ bounded independent of $\varepsilon > 0$, and $\mathbf{T}$ subunit.

Consider $\mathcal{H}$ first. By the commutator lemma 3.3 above with $s = \gamma$ and $t = 1+\beta$, each of the commutators

$$\begin{gathered}
\left[|D|^{\gamma}, \varphi H_{\ell,\mu}(x, \mathbf{v}(x), \mathbf{p}(x))\right], \left[|D|^{\gamma}, \widetilde{A}(x)\right], \\
\left[|D|^{\gamma}, \widetilde{A_{x_k}}(x)\right], \left[|D|^{\gamma}, \widetilde{\nabla_{\mathbf{v}} A}(x)\right], \left[|D|^{\gamma}, \widetilde{\nabla_{\mathbf{p}} A}(x)\right], \\
\left[|D|^{\gamma}, \widetilde{A^{\varepsilon}_{x_k}}(x)\right], \left[|D|^{\gamma}, \widetilde{\nabla_{\mathbf{v}} A^{\varepsilon}}(x)\right], \left[|D|^{\gamma}, \widetilde{\nabla_{\mathbf{p}} A^{\varepsilon}}(x)\right]
\end{gathered}$$

in particular lies in $O^{\gamma-1+\delta}_{1+\beta-\gamma}$ for all $\delta > 0$, and since they all act on functions whose $C^{0,\beta}$ norms are uniformly bounded in $\varepsilon > 0$, all but the last term in (3.20) lie in $C^{0,\beta-\gamma-\delta}$ independent of ε, and so are also bounded independent of $0 < \varepsilon < 1$ if we take δ small enough. We use the full force of the commutator lemma 3.3 to obtain as in [9] that the final term in (3.20) has the form

$$-\nabla' \left(\sum_{j=1}^{n} A^{\varepsilon}_{x_j}(x, \mathbf{v}(x), \mathbf{p}(x)) B_j + A^{\varepsilon}_{(\mathbf{v},\mathbf{p})}(x, \mathbf{v}(x), \mathbf{p}(x)) \cdot \mathbf{B} + B_0 \right) \nabla \psi \frac{\partial}{\partial x_k} p_{\ell}^{\varepsilon}, \tag{3.21}$$

where $\mathbf{B}, B_j \in O^{\gamma-1+\delta}_{1+\beta-\gamma}$ and $B_0 \in O^{\gamma-2+\delta}_{1+\beta-\gamma}$. Thus $\mathbf{B}\nabla, B_j\nabla \in O^{\gamma+\delta}_{1+\beta-\gamma}$ and $B_0\nabla \in O^{\gamma-1+\delta}_{1+\beta-\gamma}$. In particular, $\nabla' B_0 \nabla \in O^{\gamma+\delta}_{\beta-\gamma}$ and since $\psi \frac{\partial}{\partial x_k} p_{\ell}^{\varepsilon} \in C^{0,\beta}_{loc}$, we have for $0 < \delta < \beta - \gamma$,

$$\nabla' B_0 \nabla \psi \frac{\partial}{\partial x_k} p_{\ell}^{\varepsilon} \in C^{0,\beta-\gamma-\delta}_{loc} \subset L^{\infty},$$

independent of $0 < \varepsilon < 1$. Since A^{ε} is subordinate in relatively compact subregions of Γ, we have that the remaining terms in (3.21) have the form $\mathbf{T}'\mathbf{g}$ where $\mathbf{g}$ is bounded and $\mathbf{T}$ is subunit with respect to $\widetilde{A^{\varepsilon}}$, all independent of $0 < \varepsilon < 1$. Finally we consider the remaining term $|D|^{\gamma} \mathcal{F}$. We have

$$\begin{aligned}
|D|^{\gamma} \mathcal{F} &= |D|^{\gamma} \nabla' A^{\varepsilon}(x, \mathbf{v}(x), \mathbf{p}(x)) (\nabla \psi) q_{\ell}^{\varepsilon} + |D|^{\gamma} \left(\nabla' \psi\right) \widetilde{A^{\varepsilon}}(x) \nabla q_{\ell}^{\varepsilon} \\
&= \nabla' A^{\varepsilon}(x, \mathbf{v}(x), \mathbf{p}(x)) |D|^{\gamma} ((\nabla \psi) q_{\ell}^{\varepsilon}) \\
&\quad + \nabla' \left[|D|^{\gamma}, A^{\varepsilon}(x, \mathbf{v}(x), \mathbf{p}(x))\right] ((\nabla \psi) q_{\ell}^{\varepsilon})
\end{aligned}$$

plus a bounded function. Now the first term on the right is a sum of transposed subunit vector fields applied to bounded functions by (3.5), and the second term on the right is handled using the commutator lemma as above.

Altogether we have

$$\nabla' \widetilde{A^{\varepsilon}} \nabla w_{\ell}^{\varepsilon} = f + \mathbf{T}'\mathbf{g},$$

where $f, \mathbf{g}$ are bounded and $\mathbf{T}$ is a collection of bounded subunit vector fields with respect to $\widetilde{A^{\varepsilon}}$, all independent of $0 < \varepsilon < 1$. Since the linear operators $\nabla' \widetilde{A^{\varepsilon}} \nabla$ are

α-subelliptic in Ω_0 uniformly in $\varepsilon > 0$, we conclude that $\sup_{\varepsilon>0} \|\mathbf{w}^\varepsilon\|_{C^{0,\alpha}(K)} \leq C_K$, and thus that

$$\begin{aligned}\|\mathbf{p}^\varepsilon\|_{C_*^{1+\gamma+\alpha}(K)} &\approx \|\mathbf{p}^\varepsilon\|_{L^\infty(K)} + \sum_k \left\| |D|^\gamma \frac{\partial}{\partial x_k} \mathbf{p}^\varepsilon \right\|_{C^{0,\alpha}(K)} \\ &\leq C_K + \|\mathbf{w}^\varepsilon\|_{C^{0,\alpha}(K)} \leq C_K,\end{aligned}$$

independent of $\varepsilon > 0$ for every compact subset K in Ω_0. Since $\widetilde{A^\varepsilon}(x)$ is $C_*^{1+\beta}$ and elliptic in $\overline{\Omega_0} - \Omega_1$ independent of $\varepsilon > 0$, and $\mathbf{h}^0$ is in C_*^β, we have by Schauder theory (Lemma 6.18 in [8]) applied to (3.13) that,

$$\|\mathbf{p}^\varepsilon\|_{C_*^{2+\beta}\left(\overline{\Omega_0}-\Omega_1\right)} \leq C_\beta,$$

independent of $\varepsilon > 0$. Combining these inequalities yields

$$\|\mathbf{p}^\varepsilon\|_{C_*^{1+\gamma+\alpha}\left(\overline{\Omega_0}\right)} \leq C$$

with a constant C that is independent of $\varepsilon > 0$. The rest of Lemma 3.5 now follows easily upon considering a sequence of ε_j's for which $\{\mathbf{p}^{\varepsilon_j}\}_{j=1}^\infty$ converges in $C_*^{1+\gamma+\alpha-\delta}\left(\overline{\Omega_0}\right)$, $\delta > 0$, as in Lemma 3.4.

We now iterate Lemma 3.5 until we reach $\mathbf{p} \in C_*^{2+\delta}(\Omega)$, for some $\delta > 0$, differentiate equation (3.6), and then apply the above procedure again to get to $\mathbf{p} \in C_*^{3+\delta}(\Omega)$, for some $\delta > 0$. Continuing in this way we obtain $\mathbf{p}, \mathbf{v} \in C^\infty(\Omega)$, and this completes the proof of Theorem 3.1.

4. Appendix

Throughout this appendix we will assume, for the sake of simplicity, that the partial Legendre transform arises from the Monge-Ampère equation (1.1) where $k = k(x)$ is independent of u and Du. In the general case, the systems in the next subsection are complicated in that they involve the unknowns $r = u$ and $z = \frac{\partial u}{\partial x_1}$ as well as $\mathbf{v}$, and the formula for $\mathcal{L}z$ in subsection 4.3 below is not as simple.

4.1. **Failure of ellipticity of system (2.12).** Here we show that the system (2.12),

$$\frac{\partial^2 v_\ell}{\partial s^2} + \frac{\partial}{\partial t_\ell}\left(k \det \frac{\partial(v_2, ..., v_n)}{\partial(t_2, ..., t_n)}\right) = 0, \quad 2 \leq \ell \leq n,$$

is not elliptic. Indeed, with $\mathbf{F}^\ell$ as in (2.14), we note that for each fixed ℓ and σ, the matrix

$$\frac{\partial \mathbf{F}^\ell}{\partial p^\sigma} = \left[\frac{\partial F_i^\ell}{\partial p_j^\sigma}\right]_{1\leq i,j\leq n} = \begin{bmatrix} \delta_\sigma^\ell & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & kq_2^\sigma & \cdots & kq_n^\sigma \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

has $\begin{pmatrix} 0 & kq_2^\sigma & \cdots & kq_n^\sigma \end{pmatrix}$ in the ℓ^{th} row where $\left[q_j^\sigma\right]_{2\leq\sigma,j\leq n}$ is the transposed cofactor matrix of $\left[p_j^\sigma\right]_{2\leq\sigma,j\leq n}$, and all other rows vanish except for the Dirac delta function δ_σ^ℓ at the far left in the first row. Thus $\frac{\partial \mathbf{F}^\ell}{\partial p^\sigma}$ has the n-dimensional quadratic form

$$Q_{\ell\sigma}(\zeta,\boldsymbol{\xi}) = \left(\zeta,\boldsymbol{\xi}'\right)\frac{\partial \mathbf{F}^\ell}{\partial p^\sigma}(\zeta,\boldsymbol{\xi}) = \delta_\sigma^\ell\zeta^2 + k\sum_{j=2}^n q_j^\sigma \xi_\ell \xi_j .$$

Now the system (2.12) is elliptic if $\det\left[Q_{\ell\sigma}(\zeta,\boldsymbol{\xi})\right]_{2\leq \ell,\sigma\leq n} \neq 0$ for $(\zeta,\boldsymbol{\xi})$ away from the origin in $\mathbb{R}^n$. However, when $\zeta = 0$, the determinant $\det\left[Q_{\ell\sigma}(0,\boldsymbol{\xi})\right]_{2\leq \ell,\sigma\leq n}$ vanishes since the matrix $\left[Q_{\ell\sigma}(0,\boldsymbol{\xi})\right]_{2\leq \ell,\sigma\leq n}$ has rank one - indeed, the ℓ^{th} row is $k\xi_\ell$ times the fixed vector

$$\left(\sum_{j=2}^n q_j^2 \xi_j, ..., \sum_{j=2}^n q_j^n \xi_j\right).$$

4.1.1. *Ellipticity of systems (2.17) and (2.21).* Let $P = \left[p_\beta^\alpha\right]_{2\leq\alpha,\beta\leq n}$, $coP = \left[c_\beta^\alpha\right]_{2\leq\alpha,\beta\leq n}$ be the cofactor matrix of P, and $(coP)' = \left[q_\beta^\alpha\right]_{2\leq\alpha,\beta\leq n}$ denote the transpose of coP. We write $p^\alpha = (p_2^\alpha, ..., p_n^\alpha)$ for the α^{th} row of P, and $p_\beta = \left(p_\beta^2, ..., p_\beta^n\right)'$ for the β^{th} column of P, so that for example $(q^\alpha)' = c_\alpha$. Now $M = (coP)'$ and $\frac{\partial v_\ell}{\partial \mathbf{t}} = \left(p^\ell\right)'$ in the system (2.17),

$$\left\{\frac{\partial^2}{\partial s^2} + \frac{\partial}{\partial \mathbf{t}'}kM\frac{\partial}{\partial \mathbf{t}}\right\}v_\ell = 0, \qquad 2\leq \ell\leq n,$$

and so in divergence form (2.17) becomes,

$$div_{(s,\mathbf{t}')}\mathbf{F}^\ell\left((s,\mathbf{t}),\mathbf{v},D\mathbf{v}\right) = 0, \qquad 2\leq \ell\leq n,$$

where the column vector $\mathbf{F}^\ell$ is given by

$$\mathbf{F}^\ell\left((s,\mathbf{t}),\mathbf{v},\mathbf{p}\right) = \left(p_1^\ell, k\left((s,\mathbf{t}),\mathbf{v}\right)M\frac{\partial v_\ell}{\partial \mathbf{t}}\right) = \left(p_1^\ell, k\,(coP)'\left(p^\ell\right)'\right).$$

The ellipticity of the system (2.17) for $k > 0$ follows from the following claim, together with the fact that the matrix M is bounded and positive definite when $k > 0$.

Claim: $Q_{\ell\sigma}(\zeta,\boldsymbol{\xi}) \equiv \left(\zeta,\boldsymbol{\xi}'\right)\frac{\partial \mathbf{F}^\ell}{\partial p^\sigma}(\zeta,\boldsymbol{\xi}) = \delta_\sigma^\ell\left(\zeta^2 + k\boldsymbol{\xi}'M\boldsymbol{\xi}\right), \qquad 2\leq \ell,\sigma\leq n.$

Since the quadratic form associated to an antisymmetric matrix vanishes, it is enough to show that the $(n-1)\times(n-1)$ Jacobian matrix $\frac{\partial}{\partial p^\sigma}\left[(coP)'\left(p^\ell\right)'\right]$ satisfies

$$\frac{\partial}{\partial p^\sigma}\left[(coP)'\left(p^\ell\right)'\right] = \delta_\sigma^\ell\,(coP)' \pmod{\mathcal{A}_{n-1}}, \qquad 2\leq \ell,\sigma\leq n,$$

where $\mathcal{A}_{n-1}$ denotes the space of antisymmetric $(n-1)\times(n-1)$ matrices, and $\frac{\partial}{\partial p^\sigma}=\left(\frac{\partial}{\partial p_2^\sigma},...,\frac{\partial}{\partial p_n^\sigma}\right)$. We compute

$$\begin{aligned}
\frac{\partial}{\partial p^\sigma}\left[(coP)'\left(p^\ell\right)'\right] &= \frac{\partial}{\partial p^\sigma}\begin{pmatrix} q^2\left(p^\ell\right)' \\ \vdots \\ q^n\left(p^\ell\right)'\end{pmatrix} = \begin{pmatrix} q^2\frac{\partial}{\partial p^\sigma}\left(p^\ell\right)' \\ \vdots \\ q^n\frac{\partial}{\partial p^\sigma}\left(p^\ell\right)'\end{pmatrix} + \begin{pmatrix} p^\ell\frac{\partial}{\partial p^\sigma}\left(q^2\right)' \\ \vdots \\ p^\ell\frac{\partial}{\partial p^\sigma}\left(q^n\right)'\end{pmatrix} \\
&= \begin{pmatrix} q^2\delta^\ell_\sigma I_{n-1} \\ \vdots \\ q^n\delta^\ell_\sigma I_{n-1}\end{pmatrix} + \begin{pmatrix} p^\ell\frac{\partial}{\partial p^\sigma}c_2 \\ \vdots \\ p^\ell\frac{\partial}{\partial p^\sigma}c_n\end{pmatrix} = \delta^\ell_\sigma\left(coP\right)' + \sum_{\tau=2}^n p^\ell_\tau\begin{pmatrix} \frac{\partial}{\partial p^\sigma}c^\tau_2 \\ \vdots \\ \frac{\partial}{\partial p^\sigma}c^\tau_n\end{pmatrix} \\
&= \delta^\ell_\sigma\left(coP\right)' + \sum_{\tau=2}^n p^\ell_\tau\frac{\partial}{\partial p^\sigma}\left(c^\tau\right)'.
\end{aligned}$$

It remains to show that each of the matrices $\frac{\partial}{\partial p^\sigma}\left(c^\tau\right)'=\left[\frac{\partial}{\partial p^\sigma_\nu}c^\tau_\mu\right]_{2\leq\mu,\nu\leq n}$ is antisymmetric for $2\leq\sigma,\tau\leq n$. However, upon examining the matrix $P=\left[p^\alpha_\beta\right]_{2\leq\alpha,\beta\leq n}$, and using the definition of the cofactor matrix $coP=\left[c^\alpha_\beta\right]_{2\leq\alpha,\beta\leq n}$, we easily see that

$$\frac{\partial}{\partial p^\sigma_\nu}c^\tau_\mu=-\frac{\partial}{\partial p^\sigma_\mu}c^\tau_\nu,$$

and this completes the proof that (2.17) is elliptic.

We now show that the system (2.17) fails to be strongly elliptic. Such a system is *strongly elliptic* if the larger quadratic form

$$\mathcal{Q}\left(\boldsymbol{\eta}\right)=\sum_{\ell,\sigma=2}^n\left(\boldsymbol{\eta}^\ell\right)'\frac{\partial\mathbf{F}^\ell}{\partial p^\sigma}\boldsymbol{\eta}^\sigma \tag{4.1}$$

is positive definite for $\boldsymbol{\eta}=\left(\boldsymbol{\eta}^2,...,\boldsymbol{\eta}^n\right)\in\mathbb{R}^{n(n-1)}$. This notion of ellipticity is used when applying divergence structure techniques to derivatives of solutions. For the system (2.17), the quadratic form

$$\mathcal{Q}\left(\boldsymbol{\eta}\right)=\sum_{\ell,\sigma=2}^n\left(\zeta^\ell,\boldsymbol{\xi}^{\ell\prime}\right)\frac{\partial\mathbf{F}^\ell}{\partial p^\sigma}\left(\zeta^\sigma,\boldsymbol{\xi}^\sigma\right)$$

may vanish for $\boldsymbol{\eta}=\left(\left(\zeta^\ell,\boldsymbol{\xi}^\ell\right)\right)_{\ell=2}^n$ away from the origin, as a lengthy computation reveals. For example when $n=3$, $M=\begin{bmatrix} p^3_3 & -p^2_3 \\ -p^3_2 & p^2_2\end{bmatrix}$ and

$$\begin{aligned}
\mathbf{F}^2\left((s,\mathbf{t}),\mathbf{v},\mathbf{p}\right) &= \left(p^2_1,k\left[p^3_3p^2_2-\left(p^2_3\right)^2\right],kp^2_2\left(p^2_3-p^3_2\right)\right), \\
\mathbf{F}^3\left((s,\mathbf{t}),\mathbf{v},\mathbf{p}\right) &= \left(p^3_1,kp^3_3\left(p^3_2-p^2_3\right),k\left[p^2_2p^3_3-\left(p^3_2\right)^2\right]\right).
\end{aligned}$$

Thus the 6×6 block matrix $\left[\left[\frac{\partial \mathbf{F}_i^\ell}{\partial p_j^\sigma}\right]_{i,j=1}^3\right]_{\ell,\sigma=2}^3$ is given by

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & kp_3^3 & -2kp_3^2 & 0 & 0 & kp_2^2 \\ 0 & k\left(p_3^2-p_2^3\right) & kp_2^2 & 0 & -kp_2^2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -kp_3^3 & 0 & kp_3^3 & k\left(p_2^3-p_3^2\right) \\ 0 & kp_3^3 & 0 & 0 & -2kp_2^3 & kp_2^2 \end{bmatrix}$$

and with $\boldsymbol{\eta}=\left(\boldsymbol{\eta}^2,\boldsymbol{\eta}^3\right)=(0,0,1,0,1,0)$, we obtain

$$\mathcal{Q}\left(\boldsymbol{\eta}\right)=(0,0,1,0,1,0)\left[\left[\frac{\partial \mathbf{F}_i^\ell}{\partial p_j^\sigma}\right]_{i,j=1}^3\right]_{\ell,\sigma=2}^3 (0,0,1,0,1,0)'=0.$$

For the solutions we consider in (2.17), we have the additional symmetry $p_j^\sigma = p_\sigma^j$ for $2\leq \sigma,j\leq n$, and thus the application of divergence structure techniques only requires the positivity (when $k>0$) of $\mathcal{Q}\left(\boldsymbol{\eta}\right)$ on the $\frac{(n-1)(n+2)}{2}$-dimensional subspace

$$\mathcal{S}_n=\left\{\boldsymbol{\eta}=\left(\boldsymbol{\eta}^2,...,\boldsymbol{\eta}^n\right)\in\mathbb{R}^{n(n-1)}:\eta_j^i=\eta_i^j \textit{ for } 2\leq i,j\leq n\right\}.$$

Even this fails, as the above calculation demonstrates when $n=3$.

Finally we note that the system (2.21),

$$\mathcal{L}p_j^i=\left\{\frac{\partial^2}{\partial s^2}+\frac{\partial}{\partial \mathbf{t}'}kM\frac{\partial}{\partial \mathbf{t}}\right\}p_j^i=f_j^i, \quad 2\leq i\leq n, 1\leq j\leq n,$$

is strongly elliptic when $k>0$; since M is independent of the Dp_j^i, it follows that (2.21) is diagonal in the principal terms, and since M is positive definite and bounded, (2.21) is then strongly elliptic. Indeed, the matrix $\left[Q_{\ell,\sigma}\right]_{\ell,\sigma=2}^n$ is diagonal with $Q_{\ell,\ell}\left(\zeta^\ell,\boldsymbol{\xi}^\ell\right)>0$ for $\left(\zeta^\ell,\boldsymbol{\xi}^\ell\right)\neq(0,\mathbf{0})$ and $2\leq\ell\leq n$.

4.2. **The equation for $\mathcal{L}z$.** Here we establish the formula $\mathcal{L}z=\frac{\partial k}{\partial x_1}\det\frac{\partial \mathbf{v}}{\partial \mathbf{t}}$. Indeed, from (2.9) we obtain

$$\mathcal{L}z=\frac{\partial}{\partial s}\left(k\det\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right)-\frac{\partial}{\partial \mathbf{t}'}\left(kM\frac{\partial \mathbf{v}}{\partial s}\right)=I-II.$$

From the formula $\frac{\partial}{\partial s}\det A=trace\left\{(coA)'\frac{\partial}{\partial s}A\right\}$ with $A=k^{\frac{1}{n-1}}\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}$, we have

$$I=\frac{\partial}{\partial s}\left(\det k^{\frac{1}{n-1}}\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right)=trace\left\{k^{\frac{n-2}{n-1}}M\frac{\partial}{\partial s}k^{\frac{1}{n-1}}\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right\},$$

and so using (2.15),

$$\begin{aligned} I-trace\left\{kM\frac{\partial}{\partial s}\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right\} &= trace\left\{M\frac{1}{n-1}\left(\frac{\partial}{\partial s}k\right)\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right\} \\ &= \frac{1}{n-1}\left(\frac{\partial}{\partial s}k\right)trace\left\{M\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}\right\} \\ &= \left(\frac{\partial}{\partial s}k\right)\det\frac{\partial \mathbf{v}}{\partial \mathbf{t}'}. \end{aligned}$$

Using (2.19) we have

$$II=\frac{\partial}{\partial\mathbf{t}'}\left(Mk\frac{\partial\mathbf{v}}{\partial s}\right)=trace\left\{M\frac{\partial}{\partial\mathbf{t}'}k\frac{\partial\mathbf{v}}{\partial s}\right\},$$

and so by the symmetry of M and (2.16), we have $II-trace\left\{kM\frac{\partial}{\partial\mathbf{t}'}\frac{\partial\mathbf{v}}{\partial s}\right\}$

$$\begin{aligned}
&= trace\left\{M\left(\frac{\partial\mathbf{v}}{\partial s}\right)\left(\frac{\partial}{\partial\mathbf{t}'}k\right)\right\}=trace\left\{M\left[\frac{\partial v_i}{\partial s}\frac{\partial}{\partial t_j}k\right]_{2\le i,j\le n}\right\}\\
&= trace\left\{M\left[\frac{\partial v_i}{\partial s}\sum_{\ell=2}^{n}k_\ell\frac{\partial v_\ell}{\partial t_j}\right]_{2\le i,j\le n}\right\}=\sum_{\ell=2}^{n}k_\ell\; trace\left\{M\left[\frac{\partial v_i}{\partial s}\frac{\partial v_\ell}{\partial t_j}\right]_{2\le i,j\le n}\right\}\\
&= \sum_{\ell=2}^{n}k_\ell\left(\frac{\partial\mathbf{v}}{\partial s}\right)'M\frac{\partial v_\ell}{\partial\mathbf{t}}=\sum_{\ell=2}^{n}k_\ell\left(\frac{\partial\mathbf{v}}{\partial s}\right)'\mathbf{e}_\ell\left(\det\frac{\partial\mathbf{v}}{\partial\mathbf{t}'}\right)\\
&= \left(\sum_{\ell=2}^{n}k_\ell\frac{\partial v_\ell}{\partial s}\right)\det\frac{\partial\mathbf{v}}{\partial\mathbf{t}'},
\end{aligned}$$

where $k_\ell=\frac{\partial k}{\partial x_\ell}$. Combining these equalities, we obtain

$$\begin{aligned}
\mathcal{L}z &= I-II=\left(\frac{\partial}{\partial s}k-\sum_{\ell=2}^{n}k_\ell\frac{\partial v_\ell}{\partial s}\right)\det\frac{\partial\mathbf{v}}{\partial\mathbf{t}'}\\
&= \left(k_1+\sum_{\ell=2}^{n}k_\ell\frac{\partial v_\ell}{\partial s}-\sum_{\ell=2}^{n}k_\ell\frac{\partial v_\ell}{\partial s}\right)\det\frac{\partial\mathbf{v}}{\partial\mathbf{t}'}\\
&= k_1\det\frac{\partial\mathbf{v}}{\partial\mathbf{t}'}.
\end{aligned}$$

4.3. **The divergence-free property of M.** Here we establish the divergence-free property (2.19) of the matrix M. In fact, if M is the transposed cofactor matrix of any $n\times n$ Jacobian matrix $\frac{\partial\mathbf{v}}{\partial\mathbf{t}'}=\left[\frac{\partial v_i}{\partial t_j}\right]_{1\le i,j\le n}$, then $\nabla'M=\mathbf{0}'$. To see this we observe that the $\mathbf{t}$-divergence of the σ^{th} column of M can be written

$$\det\begin{bmatrix}
\frac{\partial v_1}{\partial t_1} & \cdots & \frac{\partial v_1}{\partial t_n}\\
\vdots & & \vdots\\
\frac{\partial v_{\sigma-1}}{\partial t_1} & \cdots & \frac{\partial v_{\sigma-1}}{\partial t_n}\\
\frac{\partial}{\partial t_1} & \cdots & \frac{\partial}{\partial t_n}\\
\frac{\partial v_{\sigma+1}}{\partial t_1} & \cdots & \frac{\partial v_{\sigma+1}}{\partial t_n}\\
\vdots & & \vdots\\
\frac{\partial v_n}{\partial t_1} & \cdots & \frac{\partial v_n}{\partial t_n}
\end{bmatrix},$$

where it is understood that the determinant must be expanded along the σ^{th} row. However, if we do just that, and use the formula

$$\frac{\partial}{\partial t_j}\det\left(A_1,...,A_n\right)=\sum_{i=1}^{n}\det\left(A_1,...,\frac{\partial A_i}{\partial t_j},...,A_n\right),$$

where the A_i are the columns of a square matrix, we obtain that in the case $\sigma = 1$,

$$\det \begin{bmatrix} \frac{\partial}{\partial t_1} & \cdots & \frac{\partial}{\partial t_n} \\ \frac{\partial v_2}{\partial t_1} & \cdots & \frac{\partial v_2}{\partial t_n} \\ \vdots & & \vdots \\ \frac{\partial v_n}{\partial t_1} & \cdots & \frac{\partial v_n}{\partial t_n} \end{bmatrix}$$

expands as

$$\begin{aligned} &\det \begin{bmatrix} \frac{\partial}{\partial t_1}\frac{\partial v_2}{\partial t_2} & \cdots & \frac{\partial v_2}{\partial t_n} \\ \vdots & & \vdots \\ \frac{\partial}{\partial t_1}\frac{\partial v_n}{\partial t_2} & \cdots & \frac{\partial v_n}{\partial t_n} \end{bmatrix} + ... + \det \begin{bmatrix} \frac{\partial v_2}{\partial t_2} & \cdots & \frac{\partial}{\partial t_1}\frac{\partial v_2}{\partial t_n} \\ \vdots & & \vdots \\ \frac{\partial v_n}{\partial t_2} & \cdots & \frac{\partial}{\partial t_1}\frac{\partial v_n}{\partial t_n} \end{bmatrix} \\ &- \det \begin{bmatrix} \frac{\partial}{\partial t_2}\frac{\partial v_2}{\partial t_1} & \frac{\partial v_2}{\partial t_3} & \cdots & \frac{\partial v_2}{\partial t_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial}{\partial t_2}\frac{\partial v_n}{\partial t_1} & \frac{\partial v_n}{\partial t_3} & \cdots & \frac{\partial v_n}{\partial t_n} \end{bmatrix} - ... - \det \begin{bmatrix} \frac{\partial v_2}{\partial t_1} & \frac{\partial v_2}{\partial t_3} & \cdots & \frac{\partial}{\partial t_2}\frac{\partial v_2}{\partial t_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial v_n}{\partial t_1} & \frac{\partial v_n}{\partial t_3} & \cdots & \frac{\partial}{\partial t_2}\frac{\partial v_n}{\partial t_n} \end{bmatrix} \\ &\vdots \\ &+ (-1)^{n+1} \det \begin{bmatrix} \frac{\partial}{\partial t_n}\frac{\partial v_2}{\partial t_1} & \cdots & \frac{\partial v_2}{\partial t_{n-1}} \\ \vdots & & \vdots \\ \frac{\partial}{\partial t_n}\frac{\partial v_n}{\partial t_1} & \cdots & \frac{\partial v_n}{\partial t_{n-1}} \end{bmatrix} + ... + (-1)^{n+1} \det \begin{bmatrix} \frac{\partial v_2}{\partial t_1} & \cdots & \frac{\partial}{\partial t_n}\frac{\partial v_2}{\partial t_{n-1}} \\ \vdots & & \vdots \\ \frac{\partial v_n}{\partial t_1} & \cdots & \frac{\partial}{\partial t_n}\frac{\partial v_n}{\partial t_{n-1}} \end{bmatrix}. \end{aligned}$$

The above expression vanishes identically by the equality of mixed second order partial derivatives. For example, the columns $\begin{pmatrix} \frac{\partial}{\partial t_1}\frac{\partial v_2}{\partial t_2} \\ \vdots \\ \frac{\partial}{\partial t_1}\frac{\partial v_n}{\partial t_2} \end{pmatrix}$ and $\begin{pmatrix} \frac{\partial}{\partial t_2}\frac{\partial v_2}{\partial t_1} \\ \vdots \\ \frac{\partial}{\partial t_2}\frac{\partial v_n}{\partial t_1} \end{pmatrix}$ are equal, and appear as the first columns of two otherwise identical matrices above, whose determinants appear with opposite sign. A similar reasoning, combined with interchanging appropriate columns, shows that the last term on the first line cancels with the first term on the last line.

4.4. **The other partial Legendre transforms.** Given $1 \leq \kappa \leq n$, we modify the transform (2.8) by defining variables $\mathbf{s} = (s_1, ..., s_\kappa)$ and $\mathbf{t} = (t_{\kappa+1}, ..., t_n)$ by

$$\begin{cases} s_1 &= x_1 \\ \vdots & \\ s_\kappa &= x_\kappa \\ t_{\kappa+1} &= \frac{\partial u}{\partial x_{\kappa+1}}(x) \\ \vdots & \\ t_n &= \frac{\partial u}{\partial x_n}(x) \end{cases},$$

and consider the functions $\mathbf{z} = (z_1, ..., z_\kappa)$ and $\mathbf{v} = (v_{\kappa+1}, ..., v_n)$ given by

$$\begin{cases} z_1 & = & \frac{\partial u}{\partial x_1}(x) \\ \vdots & & \\ z_\kappa & = & \frac{\partial u}{\partial x_\kappa}(x) \\ v_{\kappa+1} & = & x_{\kappa+1} \\ \vdots & & \\ v_n & = & x_n \end{cases}.$$

Calculations as in subsection 2.1 above (but more elaborate), yield in place of (2.9) the Cauchy-Riemann equations

$$\begin{cases} \det \frac{\partial \mathbf{z}}{\partial \mathbf{s}'} & = & k\left((\mathbf{s}, \mathbf{v}), r, (\mathbf{z}, \mathbf{t})\right) \det \frac{\partial \mathbf{v}}{\partial \mathbf{t}'} \\ \frac{\partial \mathbf{z}}{\partial \mathbf{t}'} & = & -\frac{\partial \mathbf{v}'}{\partial \mathbf{s}} \end{cases}.$$

Indeed, the chain rule yields

$$\frac{\partial z_i}{\partial s_j} = \frac{1}{d} \det \begin{pmatrix} u_{ij} & u_{i,\kappa+1} & \cdots & u_{in} \\ u_{\kappa+1,j} & u_{\kappa+1,\kappa+1} & & u_{\kappa+1,n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n,j} & u_{n,\kappa+1} & \cdots & u_{nn} \end{pmatrix}, \quad 1 \leq i, j \leq \kappa,$$

where $d = \det [a_{ij}]_{\kappa+1 \leq i,j \leq n}$, as well as

$$\frac{\partial \mathbf{v}}{\partial \mathbf{t}'} = \begin{pmatrix} u_{\kappa+1,\kappa+1} & \cdots & u_{\kappa+1,n} \\ \vdots & \ddots & \vdots \\ u_{n,\kappa+1} & \cdots & u_{nn} \end{pmatrix}^{-1}.$$

Now apply the determinant formula

$$\det [\widetilde{a}_{ij}]_{1 \leq i,j \leq \kappa} = \det [a_{ij}]_{1 \leq i,j \leq n} \left(\det [a_{ij}]_{\kappa+1 \leq i,j \leq n}\right)^{\kappa-1},$$

where

$$\widetilde{a}_{ij} = \det \begin{pmatrix} a_{ij} & a_{i,\kappa+1} & \cdots & a_{in} \\ a_{\kappa+1,j} & a_{\kappa+1,\kappa+1} & \cdots & a_{\kappa+1,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,j} & a_{n,\kappa+1} & \cdots & a_{nn} \end{pmatrix}, \quad 1 \leq i, j \leq \kappa,$$

which can be proved by writing $\det [\widetilde{a}_{ij}]_{1 \leq i,j \leq \kappa}$ as a sum of signed products $sgn(\sigma)\, \widetilde{a}_{i,\sigma(i)}$ over all permutations σ of $\{1, ..., \kappa\}$, and then applying Laplace's expansion to each determinant $\widetilde{a}_{i,\sigma(i)}$.

4.5. **The difficulties with alternate methods.** We begin with a discussion of the difficulties encountered in applying the Campanato method to the system (3.1), as used by Xu and Zuily in [25] to treat equations of the form

$$\left\{ \sum_{i,j=1}^{m} X_i^* M(x, \mathbf{p}) X_j \right\} p_\ell = f_\ell(x, \mathbf{p}, D\mathbf{p}), \quad 1 \leq \ell \leq N, \tag{4.2}$$

where $\mathbf{p} = (p_1, ..., p_N)$ is assumed continuous and $D\mathbf{p}$ locally square integrable, M is smooth and elliptic, $\mathbf{f}$ is smooth and has at most quadratic growth in $D\mathbf{p}$, and

$\{X_j\}_{j=1}^m$ is a collection of smooth linear vector fields satisfying Hörmander's condition (see also [15], [2], [3] and [7]). Note that our equation (3.1) fails to be of this form with x replaced by $(s, \mathbf{t})$ in (4.2) since our vector fields $X_j = \sqrt{k(s, \mathbf{v}(s, \mathbf{t}))}\frac{\partial}{\partial t_j}$ are nonlinear. A key step in the Campanato method is to freeze coefficients at a point in the elliptic part $M(\mathbf{p}(s, \mathbf{t}))$ of the operator, and then solve a Dirichlet problem for this frozen operator $\mathcal{L}_0$. While the solution $\mathbf{w}$ to the quasilinear system with operator $\mathcal{L}_0$ verifies the needed estimates by a generalization of Guan's theorem in [11], the degeneracies of the solution occur when $k(s, \mathbf{w}(s, \mathbf{t})) = 0$, and do not match the degeneracies of $\mathbf{p}$ which occur when $k(s, \mathbf{v}(s, \mathbf{t})) = 0$. If instead, we freeze the function $\mathbf{v}$ in $k(s, \mathbf{v}(s, \mathbf{t}))$, and freeze $M(\mathbf{p}(s, \mathbf{t}))$ at a point, then the solution to the Dirichlet problem with the linear operator $\mathcal{L}_0$ fails to have sufficient smoothness for applying Sobolev-type inequalities as in [25].

On the other hand, in the special case when $\psi(x_1, \mathbf{x})$ vanishes in Theorem 1.1, then of course $k(s, \mathbf{w}(s, \mathbf{t})) \approx k(s, \mathbf{v}(s, \mathbf{t}))$, and so the above difficulties with degeneracies disappears. It is likely that the Campanato method in [25] can be adapted to prove Theorem 1.1 when $\psi(x_1, \mathbf{x}) \equiv 0$ under the weaker regularity hypothesis $u \in C^2 \cap W^{3,2}$. This will be pursued elsewhere.

There is also a difficulty in applying the method of Guan in [11], as convolution with a smooth approximate identity does not behave well on the term that is quadratic in $D\mathbf{p}$ on the right side of (3.1).

Finally, any attempt to apply linear elliptic regularization to the simpler equation (2.17), $\mathcal{L}\mathbf{v} = \mathbf{0}$, such as adding εI_n to the coefficient matrix, is doomed since the matrix coefficients involve $\mathbf{p}$, and are thus no more regular than derivatives of $\mathbf{v}$. The classical elliptic theory for $\nabla' A \nabla v = 0$ requires $\nabla' A \in L^q(\mathbb{R}^n)$, $q > n$, in order to conclude that $v \in W^{2,2}$ (see e.g. Theorem 10.1 in chapter 3 of [14]), and applied to the above situation would yield only that $\nabla \mathbf{p} \in L^q$ implies $\nabla^2 \mathbf{v} \in L^2$, a trivial conclusion.

References

[1] L. Caffarelli, L. Nirenberg and J. Spruck, The Dirichlet problem for nonlinear second order elliptic equations, I. Monge-Ampère equations, *Comm. Pure Appl. Math.* **37** (1984), 369-402.

[2] S. Campanato, Equazioni ellittiche del II ordine e spazi $\mathcal{L}^{(2,\lambda)}$, *Ann. Mat. Pura Appl. (4)* **69** (1965), 321-381.

[3] S. Campanato, *Sistemi ellittici in forma divergenza. Regolarità all' interno*, Quaderni. Pisa: Scuola Normale Superiore 1980.

[4] A. Douglis and L. Nirenberg, Interior estimates for elliptic systems of partial differential equations, *Comm. Pure Appl. Math.* **8** (1955), 503-538.

[5] B. Franchi, Weighted Sobolev-Poincaré inequalities and pointwise estimates for a class of degenerate elliptic equations, *Trans. Amer. Math. Soc.* **327** (1991), 125-158.

[6] B. Franchi and E. Lanconelli, Hölder regularity theorem for a class of linear nonuniformly elliptic operators with measurable coefficients, *Annali Scuola Normale Sup. Pisa* **4, 10** (1983), 523-541.

[7] M. Giaquinta, Multiple integrals in the calculus of variations and nonlinear elliptic systems, *Annals of Math. Studies* **105**, Princeton Univ. Press, New Jersey 1983.

[8] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, revised 3rd printing, 1998.

[9] P. Guan, Regularity of a class of quasilinear degenerate elliptic equations, *Advances in Mathematics* **132** (1997), 24-45.

[10] P. Guan, C^2 *a priori* estimates for degenerate Monge-Ampère equations, *Duke Math. J.* **86** (1997), 323-346.

[11] P. Guan, Quasilinear degenerate elliptic equations in divergence form, *Contemporary Mathematics* **205** (1997), 93-99.

[12] P. Guan, N. S. Trudinger and X.-J. Wang, On the Dirichlet problem for degenerate Monge-Ampère equations, *Acta Math.* **182** (1999), 87-104.

[13] J. Iaia, Isometric embedding of surfaces with non negative curvature in $\mathbb{R}^3$, *Duke Math. J.* **67** (1992), 423-459.

[14] O. A. Ladyzhenskaya and N. N. Ural'tseva, *Linear and Quasilinear Elliptic Equations*, Moscow: Izdat. "Nauka" 1964 [Russian]. English Translation: New York: Academic Press 1968. 2nd Russian ed. 1973.

[15] C. B. Morrey Jr., Second order elliptic systems of differential equations, In: *Annals of Math. Studies* **33**, 101-159, Princeton Univ. Press, New Jersey 1954.

[16] J. Moser, On Harnack's theorem for elliptic differential equations, *Comm. Pure Appl. Math.* **14** (1961), 577-591.

[17] A. V. Pogorelov, *Monge-Ampère Equations of Elliptic Type*, Groningen, Noordhoff, 1964.

[18] C. Rios, E. Sawyer and R. L. Wheeden, Regularity of degenerate quasilinear equations, *in preparation.*

[19] E. Sawyer, A symbolic calculus for rough pseudodifferential operators, *Indiana U. Math. J.* **45** (1996), 289-332.

[20] E. Sawyer and R. L. Wheeden, Regularity of degenerate Monge-Ampère and prescribed Gaussian curvature equations in two dimensions, *preprint.*

[21] E. Sawyer and R. L. Wheeden, Hölder continuity of subelliptic equations with rough coefficients, *in preparation.*

[22] F. Schulz, Regularity Theory for Quasilinear Elliptic Systems and Monge-Ampère Equations in Two Dimensions, *Lecture Notes in Math.* **1445** (1990), Springer-Verlag.

[23] M. Taylor, *Pseudodifferential Operators and Nonlinear PDE*, Birkhauser, Basel, 1991.

[24] C.-J. Xu, Opérateurs sou-elliptiques et régularité des solitions aux dérivées partielles nonlinéaires du second ordre en deux variables, *Comm. Partial Differential Equations* **11** (1986), 1575-1603.

[25] C.-J. Xu and C. Zuily, Higher interior regularity for quasilinear subelliptic systems, *Prépublications Univ. Paris-Sud Math.* **425** (1995), 1-24.

McMaster University, Hamilton, Ontario
E-mail address: riosc@math.mcmaster.ca

McMaster University, Hamilton, Ontario
E-mail address: sawyer@mcmaster.ca

Rutgers University, New Brunswick, N.J.
E-mail address: wheeden@math.rutgers.edu