Lecture 26: Consensus Equilibrium for Two Models

Charles A. Bouman

Purdue University

Monday, October 26, 2026

Outline

1. The consensus equilibrium (CE) equations
2. CE and the MAP estimate: proximal maps and general agents
3. Serial solution: Douglas-Rachford, Mann iterations, and the convergence of ADMM
4. Parallel solution: stacked states and the averaging operator

Reading: Sections 10.3 to 10.3.4, and Appendix B.

Today's Big Ideas

1. PnP-ADMM's limit solves the CE equations, not an optimization.
2. For proximal maps, the CE solution is exactly the MAP estimate.
3. General agents give a balance of actions, with no cost minimized.
4. Douglas-Rachford finds the fixed point; this proves ADMM converges.
5. A parallel form gives the same solution; it has two advantages.

What is PnP doing?

A trained denoiser H is usually not a proximal map. Then PnP-ADMM does not minimize a cost function.

1. What exactly is PnP-ADMM computing?
2. Under what conditions does it converge?
3. If it converges, what is the limit?

The answer

Define **consensus equilibrium (CE)**: the solution is defined by an equilibrium equation instead of an optimization. CE fuses models that are actions with models that are objectives.

The CE equations

PnP-ADMM iterates

$$x \leftarrow F(v-u)\ , \qquad v \leftarrow H(x+u)\ , \qquad u \leftarrow u+(x-v)\ .$$

At convergence the u update gives $x = v$. So two equations hold.

Definition: CE equations

The **CE equations** are

$$x^* = F(x^* - u^*)\ , \tag{10.5}$$
$$x^* = H(x^* + u^*)\ . \tag{10.6}$$

The CE equations replace $x^* = \arg\min_x\{f(x)+h(x)\}$ as the definition of the solution x^*.

Subgradients

Definition: Subgradient

The **subgradient** of f at x_o, denoted $\nabla f(x)|_{x=x_o}$, is the set of vectors c such that for x in an open ball about x_o,

$$f(x) - f(x_o) \geq c^t(x - x_o) \ .$$

- If f is differentiable, the subgradient is $\{\nabla f(x_o)\}$.
- Example: $\nabla|x| = [-1, 1]$ at $x = 0$.

Theorem: Minimum of a Convex Function

Let $f : \mathbb{R}^N \to \mathbb{R} \cup \{\infty\}$ be convex. Then x_o is a global minimum of f if and only if $\mathbf{0} \in \nabla f(x)|_{x=x_o}$.

Theorem: Equivalence of CE and MAP Estimates

Theorem: Equivalence of CE and MAP Estimates

Let F and H be proximal maps of proper closed convex functions f and h. Then x^* is a solution to the CE equations (10.5) and (10.6) if and only if x^* is a global minimum of $f(x) + h(x)$.

- The equilibrium of F and H is the minimum of f and h.
- The proof uses one fact: the minimum theorem applied to F:

$$x = F(w) \iff \mathbf{0} \in \nabla f(x) + \frac{1}{\sigma^2}(x - w) \ .$$

Proof: CE solution implies global minimum

Assume (x^*, u^*) solves the CE equations.

Apply the proximal map condition to $x^* = F(x^* - u^*)$ and to $x^* = H(x^* + u^*)$:

$$\mathbf{0} \in \nabla f(x^*) + \frac{1}{\sigma^2} u^* \ ,$$
$$\mathbf{0} \in \nabla h(x^*) - \frac{1}{\sigma^2} u^* \ .$$

Add the two lines. The u^* terms cancel:

$$\mathbf{0} \in \nabla f(x^*) + \nabla h(x^*) \subset \nabla \left[f(x) + h(x) \right]|_{x=x^*} \ .$$

By the minimum theorem, x^* is a global minimum of $f(x) + h(x)$.

Proof: global minimum implies CE solution

Assume x^* is a global minimum of $f(x) + h(x)$.

Then $\mathbf{0} \in \nabla f(x^*) + \nabla h(x^*)$, so there are $c_f \in \nabla f(x^*)$ and $c_h \in \nabla h(x^*)$ with $c_f + c_h = \mathbf{0}$.

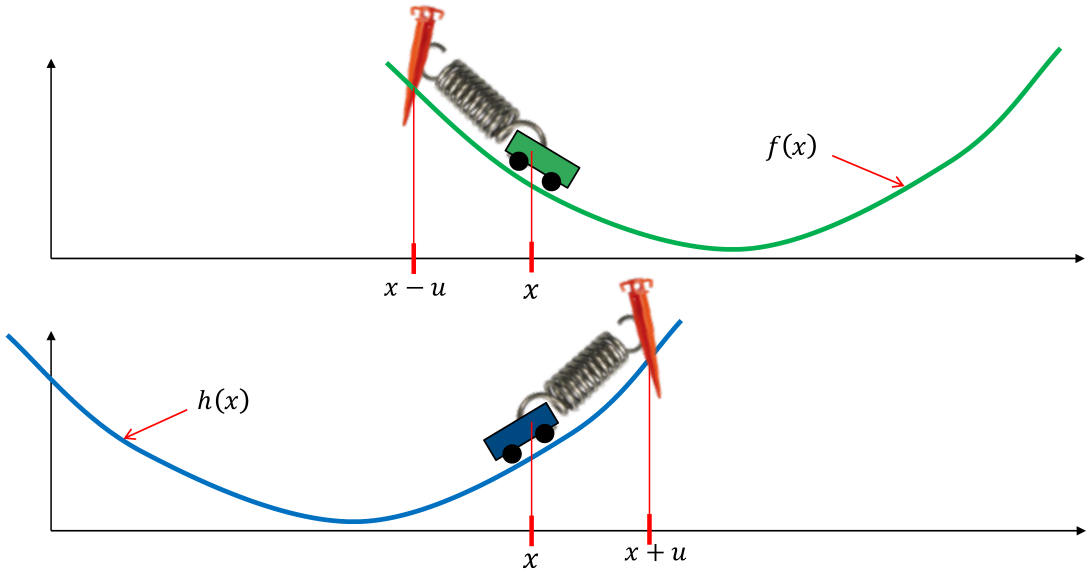
Choose $u^* = -\sigma^2 c_f = \sigma^2 c_h$. Then

$$\mathbf{0} = c_f + \frac{1}{\sigma^2} u^* \in \nabla f(x^*) + \frac{1}{\sigma^2} u^* \ ,$$
$$\mathbf{0} = c_h - \frac{1}{\sigma^2} u^* \in \nabla h(x^*) - \frac{1}{\sigma^2} u^* \ .$$

By the proximal map condition, $x^* = F(x^* - u^*)$ and $x^* = H(x^* + u^*)$. $\square$

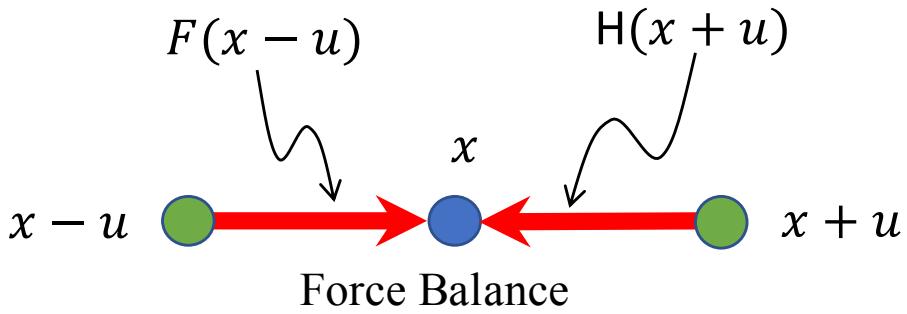
The first step uses $\nabla[f + h](x^*) = \nabla f(x^*) + \nabla h(x^*)$, which holds, for example, when f is differentiable.

CE solution when the agents are proximal maps



- Each cart is the proximal map solution for one function.
- The two springs are stretched by $\pm u$, with stakes at $x^* - u$ and $x^* + u$.
- The minimum of $f + h$ is where the two carts reach the same location x^*.

CE solution when the agents are general



- The CE equations are a **force balance** of equal, opposing actions.
- This equilibrium does **not** minimize any cost function.
- PnP mixes actions with objectives: a DNN prior and a sensor model.

Serial agents: a change of variables

Do the CE equations have a solution, and how do we compute it?

Define $w_1 = x - u$ and $w_2 = x + u$. The CE equations become

$$F(w_1^*) = \frac{w_1^* + w_2^*}{2} \,, \tag{10.7}$$

$$H(w_2^*) = \frac{w_1^* + w_2^*}{2} \,. \tag{10.8}$$

Convert back to the original variables with

$$x^* = \frac{w_1^* + w_2^*}{2} \,, \qquad u^* = \frac{w_2^* - w_1^*}{2} \,.$$

The fixed-point operator T

Rearrange (10.7) and (10.8), writing Fw for $F(w)$:

$$(2F - I)\, w_1^* = w_2^* \ ,$$
$$(2H - I)\, w_2^* = w_1^* \ .$$

Substitute the first equation into the second:

$$(2H - I)\,(2F - I)\, w_1^* = w_1^* \ .$$

Result

w_1^* solves the CE equations if and only if it is a **fixed point** of

$$T = (2H - I)\,(2F - I) \ ,$$

that is, $Tw_1^* = w_1^*$.

Douglas-Rachford with Mann iterations

The **Douglas-Rachford** (DR) algorithm finds the fixed point with **Mann iterations**:

$$w_1 \leftarrow (1-\rho)w_1 + \rho T w_1 \ ,$$

where $\rho \in (0,1)$ is a user-selectable parameter.

Theorem: Mann Iteration

Consider $w^{(k)} = (1-\rho)w^{(k-1)} + \rho T w^{(k-1)}$.

- If T is nonexpansive with a fixed point and $\rho \in (0,1)$, then $w^{(k)} \to w^*$ with $Tw^* = w^*$.
- Contractive T, $\rho \in (0,1]$: $w^{(k)} \to w^*$ with $Tw^* = w^*$.

T is **contractive** if $\|T(x) - T(y)\| \leq \lambda \|x - y\|$ for some $\lambda \in [0,1)$.

The PnP-DR-Serial algorithm

PnP-DR-Serial Algorithm

Initialize Reconstruction: $w_1 \in \mathbb{R}^N$

Until Converged $\{$

$$\begin{aligned} w_1' &\leftarrow (2H - I)(2F - I)w_1 \\ w_1 &\leftarrow (1-\rho)w_1 + \rho w_1' \end{aligned}$$

$\}$

$w_2^* \leftarrow (2F - I)w_1^*$

$x^* \leftarrow \frac{w_1^* + w_2^*}{2}$

$\text{Return}(x^*)$

- F and H can be plug-in forward and prior agents.
- Each iteration applies F and then H.

Theorem: Serial PnP Convergence

Theorem: Serial PnP Convergence

If $T = (2H - I)(2F - I)$ is nonexpansive and T has a fixed point, then the PnP-DR-Serial algorithm converges to a solution of the CE equations.

Proof. This follows directly from the Mann Iteration theorem. □

- With a plug-in H, nonexpansive T is sufficient for convergence.
- PnP often converges when T is not nonexpansive; proof is harder.
- PnP-DR-Serial allows other ρ; larger ρ converges faster.

This proves that ADMM converges

Theorem: Serial ADMM Convergence (Theorem B.3)

Let F and H be proximal maps of proper closed convex functions f and h, and let $f(x) + h(x)$ have a global minimum. Then the PnP-DR-Serial algorithm

- is equivalent to ADMM for $f(x) + h(x)$ when $\rho = 1/2$,
- converges to a solution of the CE equations, and
- converges to $x^* = \arg\min_x \{f(x) + h(x)\}$.

- At $\rho = 1/2$, it **is** ADMM with split $x = v$, $\sigma^2 = 1/a$.
- So this proves Lecture 24's ADMM convergence for $x = v$, $B = I$.

Proof of Serial ADMM Convergence

Define H to be **firmly nonexpansive** if $\forall x, y \in \mathbb{R}^N$,

$$\|H(x) - H(y)\|^2 \leq (x - y)^t (H(x) - H(y)) \ .$$

1. A proximal map (proper, closed, convex) is firmly nonexpansive.
2. If H is firmly nonexpansive, then $2H - I$ is nonexpansive.
3. Compositions of nonexpansive maps are nonexpansive; so T is.
4. A global minimum exists, so CE has a solution; T has a fixed point.
5. By the Mann theorem, w_1 converges to a fixed point of T, the minimum.

The equivalence with ADMM at $\rho = 1/2$ follows from algebra on the update equations. Items 1 and 2 are stated here without proof.

Parallel agents: stacked states and the averaging operator

Define the stacked state and the stacked operator:

$$\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}, \qquad \mathbf{F}(\mathbf{w}) = \begin{bmatrix} F(w_1) \\ H(w_2) \end{bmatrix}.$$

Define the **averaging operator**

$$\mathbf{G}(\mathbf{w}) = \begin{bmatrix} \bar{w} \\ \bar{w} \end{bmatrix}, \qquad \text{where} \qquad \bar{w} = \frac{w_1 + w_2}{2}.$$

CE equations in parallel form

Equations (10.7) and (10.8) become

$$\mathbf{F}(\mathbf{w}^*) = \mathbf{G}(\mathbf{w}^*).$$

The averaging operator applied twice

The average of the average is the average:

$$\mathbf{G}\mathbf{G} = \mathbf{G} \ .$$

$\mathbf{G}$ is linear, so

$$(2\mathbf{G} - I)^2 = 4\mathbf{G}\mathbf{G} - 4\mathbf{G} + I = I \ ,$$

and therefore

$$(2\mathbf{G} - I)^{-1} = (2\mathbf{G} - I) \ .$$

The parallel fixed-point operator $\mathbf{T}$

Using $(2\mathbf{G} - I)^{-1} = (2\mathbf{G} - I)$:

$$\begin{aligned} \mathbf{F}\mathbf{w}^* &= \mathbf{G}\mathbf{w}^* \ , \\ (2\mathbf{F} - I)\mathbf{w}^* &= (2\mathbf{G} - I)\mathbf{w}^* \ , \\ (2\mathbf{G} - I)(2\mathbf{F} - I)\mathbf{w}^* &= \mathbf{w}^* \ . \end{aligned}$$

Result

$\mathbf{w}^*$ solves the CE equations if and only if $\mathbf{T}\mathbf{w}^* = \mathbf{w}^*$, where $\mathbf{T} = (2\mathbf{G} - I)(2\mathbf{F} - I)$.

- Mann iterations $\mathbf{w} \leftarrow \rho\mathbf{T}\mathbf{w} + (1 - \rho)\mathbf{w}$ converge when $\mathbf{T}$ is nonexpansive and has a fixed point.
- If F and H are firmly nonexpansive, $\mathbf{T}$ is nonexpansive.

The PnP-DR-Parallel algorithm

PnP-DR-Parallel Algorithm

Initialize Reconstruction: $x^{(0)} \in \mathbb{R}^N$

$$\mathbf{w} \leftarrow \begin{bmatrix} x^{(0)} \\ x^{(0)} \end{bmatrix}$$

Until Converged $\{$

$$\begin{aligned} \mathbf{w}' &\leftarrow (2\mathbf{G} - I)(2\mathbf{F} - I)\mathbf{w} \\ \mathbf{w} &\leftarrow \rho\mathbf{w}' + (1-\rho)\mathbf{w} \end{aligned}$$

$\}$

$x^* \leftarrow \frac{w_1 + w_2}{2}$

$\mathrm{Return}(x^*)$

- $\mathbf{F}$ applies F and H to separate states in parallel.

Two advantages over consensus ADMM

Consensus ADMM is ADMM for the split in which each agent has its own copy of x and the copies are constrained to be equal. When $\rho = 1/2$ and F and H are proximal maps, PnP-DR-Parallel converges to the same MAP estimate.

First advantage

PnP-DR has a well-defined solution even when F and H are not proximal maps. ADMM minimizes a cost function, and without proximal maps there is no cost function to minimize.

Second advantage

Larger values of ρ typically converge much faster. PnP-DR requires only that $\mathbf{T}$ be nonexpansive, a far weaker condition than being a proximal map.

Summary

1. The CE equations $x^* = F(x^* - u^*)$ and $x^* = H(x^* + u^*)$ define the solution in place of an optimization.
2. For proximal maps of proper closed convex f and h, the CE solution is $x^* = \arg\min_x \{f(x) + h(x)\}$.
3. For general agents, CE is a force balance; it minimizes no cost.
4. The CE solution satisfies $Tw_1^* = w_1^*$ with $T = (2H - I)(2F - I)$, and Mann iterations $w_1 \leftarrow (1 - \rho)w_1 + \rho Tw_1$ compute it. At $\rho = 1/2$ this is ADMM, so ADMM converges.
5. In parallel form $\mathbf{F}(\mathbf{w}^*) = \mathbf{G}(\mathbf{w}^*)$, and the solution is the fixed point of $\mathbf{T} = (2\mathbf{G} - I)(2\mathbf{F} - I)$. This form has two advantages over consensus ADMM.