Finding Fixed Points by Averaging with Well-Behaved Maps

Marina Epelman* Thomas L. Magnanti† Georgia Perakis‡

October 1997, Revised September 25, 2001§

Abstract

We introduce and examine a general framework for parametrically solving fixed point problems through a family of problems that are "better behaved" than our original problem. One class of examples parametrically combines a well-behaved map with the fixed point map into a composite map. Special cases include (outside) averaging the underlying map with a contractive map, or alternatively, with a nonexpansive map whose fixed points include those of the original map, averaging with the identity map, and inside averaging (i.e. averaging before applying the underlying map). We establish conditions under which the trajectory of fixed points of the parametric map converge to a fixed point of the original problem as the parameter approaches zero. To implement this solution framework, we also consider an iterative scheme that approximates the parametric fixed point trajectory. We establish a convergence result and characterize the limit points of this approximation scheme.

Key Words:
Fixed Point Problems, Averaging Schemes, Nonexpansive Maps.

1 Introduction

Algorithms for solving problems in many applied settings establish a mapping T whose iterative application leads to a fixed point solution that solves the original problem. The convergence of this iterative procedure to a fixed point solution often requires strong assumptions on the algorithmic map T that restrict the algorithm's domain of applicability. For example, the classical function iteration $x_{k+1} = T(x_k)$ requires the map T to be contractive. In many instances, various forms of averaging extend the range of applicability of the algorithmic (fixed point) map T.

In this paper, we introduce a general averaging framework for solving the finite dimensional fixed point problem

$$\text{Find } \; x^* \in K \; \text{ satisfying } \; x^* = T(x^*)$$

*Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan.
†School of Engineering and Sloan School of Management, MIT, Cambridge, MA.
‡Sloan School of Management and Operations Research Center, MIT, Cambridge, MA.
§This research has been supported by NSF grant DMI-9634736.

defined over a given ground set $K \subseteq R^n$ by a given map $T : K \rightarrow K$. This problem formulation is closely related to several problem classes in mathematical programming, such as linear and nonlinear optimization, complementarity problems, and variational inequality problems (see [12] for more details). For example, we obtain an equivalent fixed point reformulation of variational inequality problems, that is,

$$\text{Find } x^* \in K : \quad f(x^*)^t(x - x^*) \geq 0, \quad \forall x \in K,$$

through the map $T = \Pr_K(I - \rho f)$ defined by the projection operator $\Pr_K$ onto the set K and some positive constant ρ.

To solve the fixed point problem, we introduce a general averaging framework that considers at each step a parameterization of the original fixed point problem. The averaging framework thus creates a family of problems which are easier to solve than the original problem. A special case of this framework is averaging the fixed point map T with maps g that are "well-behaved" to counteract the "bad" properties of the map T. In particular, the framework includes as special cases averaging with the identity map, with nonexpansive maps whose fixed points include those of the original problem, as well as with general contractive maps.

The framework includes as special cases "outside averaging" $\lambda g(x) + (1 - \lambda)T(x)$ with an averaging parameter $0 \leq \lambda \leq 1$. Halpern [8] and Browder [3] introduced this type of averaging and Bauschke [1] and Wittmann [15] studied it further in the special case when $g(x) =$ constant. Dunn [5] introduced the special case of $g(x) = x$, which Magnanti and Perakis further studied for variational inequality problems (see [10] and [11]). The framework also includes "inside averaging" $T(\lambda g(x) + (1 - \lambda)x)$. It also permits averaging with the identity map (line search procedures), or outside and inside averaging with contractive maps as well as with the proximal point map. When applied to fixed point and variational inequality problems, this framework produces certain known methods as well as several new ones.

Our goals in this paper are to

(i) study averaging schemes based upon parameterizations of (that is, averaging) the original fixed point map with "well-behaved" maps such as contractive maps, the identity map or, more generally, nonexpansive maps whose fixed points include those of the original problem,

(ii) develop a general averaging framework that unifies several averaging approaches,

(iii) solve a large class of fixed point problems by considering fixed point maps T satisfying properties weaker than contractiveness (including forms of nonexpansiveness that permit fixed point problems with multiple solutions), and

(iv) create as a special case averaging schemes that are more efficient than the classical function iteration even when the underlying map is contractive.

1.1 Notation and preliminaries

Throughout this paper we will be working with a convex set $K \subseteq \Re^n$ and a map $T : K \to K$. We refer to the problem

$$\text{Find } \; x^* \in K \; \text{ satisfying } \; x^* = T(x^*) \tag{1}$$

as the *fixed point problem* associated with the map T and the set K. We let FP(T), which we assume to be a nonempty set, denote the set of *fixed point* solutions of problem (1). We permit the ground set K to be unbounded and the fixed point problem to have multiple solutions. That is, $FP(T)$ can contain more than one point.

Definition 1 *A map $G : K \to K$ is called contractive (or a contraction) with a contraction constant $a \in [0, 1)$ if for any $x, y \in K$,*

$$\|G(x) - G(y)\| \leq a\|x - y\|.$$

A map $G : K \to K$ is called nonexpansive if for any $x, y \in K$,

$$\|G(x) - G(y)\| \leq \|x - y\|.$$

A map $G : K \to K$ is called firmly nonexpansive if for any $x, y \in K$,

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

Equivalently, G is firmly nonexpansive if for any $x, y \in K$,

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

We use the following result repeatedly in our analysis (see, for example, Dunn[5]):

Proposition 1 *Suppose $G : K \to K$ is nonexpansive, with FP(G) $\neq \emptyset$, and the sequence $\{x_k\}_{k=0}^{\infty} \subset K$ satisfies the condition $\lim_{k\to\infty} \|x_k - G(x_k)\| = 0$. Then every limit point of the sequence is a fixed point of G.*

We denote the identity map by I. Throughout our discussion, we use the following well-known results concerning the relationship between a given map $G : K \to K$ and the map $(I - G) : K \to K$:

Proposition 2 *[4] A map G is contractive with contraction constant $a \in [0, 1)$ if and only if the map $I - G$ is a strongly monotone map with monotonicity constant $\frac{1-a^2}{2}$. Even more strongly, for any x and y,*

$$(x - G(x) - y + G(y))^t(x - y) \geq \frac{1 - a^2}{2}\|x - y\|^2 + \frac{1}{2}\|x - G(x) - y + G(y)\|^2.$$

A map G is nonexpansive if and only if the map $I - G$ is a strongly-f-monotone map with monotonicity constant $\frac{1}{2}$. That is, for any x and y,

$$(x - G(x) - y + G(y))^t(x - y) \geq \frac{1}{2}\|x - G(x) - y + G(y)\|^2.$$

For any convex closed set $S \subseteq \Re^n$ and any point $y \in \Re^n$, we denote the Euclidean projection of y onto S by $\Pr_S(y)$.

2 Averaging trajectories

2.1 Generalized Averaging Map

As a first step in our analysis, we approximate the original fixed point problem with a parameterized family of fixed point problems

$$x_\lambda = F_\lambda(x_\lambda) \tag{2}$$

with parameterization parameter $\lambda \in [0, 1]$. Our intent is to choose the family of functions $F_\lambda(x) : K \to K$ to be a parameterization of the map T and be "better behaved" than the map T. In particular, we make the following assumptions on this family of functions:

A1 For all $\lambda \in (0, 1]$, the parameterized problem $x_\lambda = F_\lambda(x_\lambda)$ has at least one solution.

A2 The function $F_\lambda(x) = F(\lambda, x) : [0, 1] \times K \to K$ is a continuous function of (λ, x).

A3 $\mathrm{FP}(T) \supseteq \mathrm{FP}(F_0) \neq \emptyset$.

We refer to any map $F_\lambda(\cdot)$ satisfying conditions **A1**–**A3** as a *generalized averaging map* for T.

For parameterizations satisfying these three assumptions, the following result provides a characterization of the limit points of the set $\{x_\lambda\}_{\lambda \in (0,1]}$ induced by the parameterized fixed point problems (2) as $\lambda \to 0^+$.

Proposition 3 *Suppose $T : K \to K$ is a map with $\mathrm{FP}(T) \neq \emptyset$. Suppose further that the family of maps $F_\lambda : K \to K$ satisfies conditions **A1**–**A3**. Then as $\lambda \to 0^+$, every limit point of the sequence (or trajectory) $\{x_\lambda\}_{\lambda \in (0,1]}$ is a fixed point of T.*

Proof: Condition **A1** states that the iterates induced by (2) exist. Let $\bar{x} \in K$ be a limit point of the sequence (or trajectory) $\{x_\lambda\}$ as $\lambda \to 0^+$, i.e., for some sequence $\{\lambda_k\} \subset (0, 1]$ with $\lim_{k\to\infty} \lambda_k = 0$, $\bar{x} = \lim_{k\to\infty} x_{\lambda_k}$. Then as $k \to \infty$

$$x_{\lambda_k} - F_0(x_{\lambda_k}) = F_{\lambda_k}(x_{\lambda_k}) - F_0(x_{\lambda_k}) = F(\lambda_k, x_{\lambda_k}) - F(0, x_{\lambda_k}) \to 0$$

since $\lambda_k \to 0$, x_{λ_k} is a convergent subsequence, and, by assumption **A2**, the function $F(\lambda, x)$ is continuous. Consequently, the limit point $\bar{x}$ is a fixed point of F_0, and therefore, from condition **A3**, a fixed point of T. ∎

The previous analysis has characterized the limit points of the trajectory of solutions of the parametric family of fixed point subproblems as solutions of the original problem. To further understand the nature of the limit points, we need to impose some additional conditions on the generalized averaging map F_λ.

A4 $F_\lambda = F(G_\lambda)$, for maps $F : K \to K$ and $G_\lambda : K \to K$ satisfying the conditions

- F is nonexpansive and $\mathrm{FP}(T) \subseteq \mathrm{FP}(F)$,
- G_0 is nonexpansive and $\mathrm{FP}(T) \subseteq \mathrm{FP}(G_0)$,

- $\lim_{\lambda\to 0^+} \frac{\|G_\lambda(x_\lambda)-x_\lambda\|^2}{\lambda} = 0.$

A5 For all $\lambda \in [0,1]$, all $x \in K$ and all $x^* \in \mathrm{FP}(T)$,

$$G_\lambda(x)^t(x-x^*) \leq \lambda G_1(x)^t(x-x^*) + (1-\lambda)G_0(x)^t(x-x^*).$$

A6a $G_1(x)$ is a nonexpansive map.

A6b $G_1(x)$ is a contractive map with contraction constant $a \in [0,1)$.

Note that condition **A4** implies that $\mathrm{FP}(T) \subseteq \mathrm{FP}(F_0)$ since if $x^* \in \mathrm{FP}(T)$ then $F_0(x^*) = F(G_0(x^*)) = F(x^*) = x^*$ from the set inclusions in condition **A4**. Taken together, conditions **A3** and **A4**, imply that $\mathrm{FP}(T) = \mathrm{FP}(F(G_0))$.

It might initially appear unclear how to choose the maps F and G_0 satisfying conditions **A3** and **A4** without finding the fixed points of T beforehand. Examples in Section 2.2 show that we need not already know the fixed points of T, for example, by letting $F = T$ and $G_0 = I$, or $F = I$ and $G_0 = T$. In both cases, these maps trivially satisfy the set containment conditions imposed by assumptions **A3** and **A4**.

By imposing these additional assumptions, we can further characterize the limit points of the trajectory induced by the parametric fixed point problem (1).

Theorem 1
*I) Suppose the generalized averaging map F_λ satisfies conditions **A1–A6a**. Then as $\lambda \to 0^+$, every limit point $\bar{x}$ of the sequence $\{x_\lambda\}$ is a fixed point of T and satisfies the condition*

$$\bar{x} = \mathrm{Pr}_{\mathrm{FP}(T)} G_1(\bar{x}).$$

*II) Suppose, alternatively, the generalized averaging map F_λ satisfies conditions **A1–A6b**. Then as $\lambda \to 0^+$, the sequence $\{x_\lambda\}$ has at most one limit point x^* which is the unique fixed point of T defined as*

$$x^* = \mathrm{Pr}_{\mathrm{FP}(T)} G_1(x^*).$$

Proof: Suppose assumptions **A1–A5** and either **A6a** or **A6b** (which implies **A6a**) are valid for the family F_λ. Let $x \in \mathrm{FP}(T)$ be an arbitrary fixed point of T.

Condition **A4** implies that $x \in \mathrm{FP}(F)$ and F is nonexpansive. Proposition 2 shows that

$$0 \leq ((G_\lambda(x_\lambda) - F(G_\lambda(x_\lambda))) - (x - F(x)))^t(G_\lambda(x_\lambda) - x)$$

$$= (G_\lambda(x_\lambda) - x_\lambda)^t(G_\lambda(x_\lambda) - x) = \|G_\lambda(x_\lambda) - x_\lambda\|^2 + (G_\lambda(x_\lambda) - x_\lambda)^t(x_\lambda - x) \tag{3}$$

$$= \|G_\lambda(x_\lambda) - x_\lambda\|^2 - (x_\lambda - G_0(x_\lambda))^t(x_\lambda - x) + (G_\lambda(x_\lambda) - G_0(x_\lambda))^t(x_\lambda - x).$$

Condition **A5** implies that for all $\lambda \in [0,1]$,

$$G_\lambda(x_\lambda)^t(x_\lambda - x) \leq \lambda G_1(x_\lambda)^t(x_\lambda - x) + (1-\lambda)G_0(x_\lambda)^t(x_\lambda - x),$$

or, upon rearranging terms,

$$(G_\lambda(x_\lambda) - G_0(x_\lambda))^t (x_\lambda - x) \le \lambda (G_1(x_\lambda) - G_0(x_\lambda))^t (x_\lambda - x).$$

Substituting this inequality into (3), we conclude that for all $\lambda \in [0, 1]$,

$$0 \le \|G_\lambda(x_\lambda) - x_\lambda\|^2 - (x_\lambda - G_0(x_\lambda))^t(x_\lambda - x) + \lambda(G_1(x_\lambda) - G_0(x_\lambda))^t(x_\lambda - x)$$

$$= \|G_\lambda(x_\lambda) - x_\lambda\|^2 - (1-\lambda)(x_\lambda - G_0(x_\lambda))^t(x_\lambda - x) + \lambda(G_1(x_\lambda) - x_\lambda)^t(x_\lambda - x).$$

Since $x \in \mathrm{FP}(T) \subseteq \mathrm{FP}(G_0)$ and G_0 is a nonexpansive map (from condition **A4**),

$$(x_\lambda - G_0(x_\lambda))^t(x_\lambda - x) \ge 0.$$

Therefore, for all $\lambda \in (0, 1]$,

$$\frac{\|G_\lambda(x_\lambda) - x_\lambda\|^2}{\lambda} \ge (x_\lambda - G_1(x_\lambda))^t(x_\lambda - x).$$

Condition **A4** implies that $\limsup_{\lambda \to 0^+}(x_\lambda - G_1(x_\lambda))^t(x_\lambda - x) \le 0$. Therefore as $\lambda \to 0^+$, any limit point $\bar{x}$ (which, if it exists, is also a fixed point solution of the map T as we have shown in Theorem 3) satisfies the condition

$$(\bar{x} - G_1(\bar{x}))^t(\bar{x} - x) \le 0 \quad \forall x \in \mathrm{FP}(T). \tag{4}$$

Condition (4) can alternatively be interpreted as $\bar{x} = \mathrm{Pr}_{\mathrm{FP}(T)} G_1(\bar{x})$, thus establishing the first claim of the theorem.

To establish the second claim of the theorem, suppose assumption **A6b** is valid. Since the set $\mathrm{FP}(T)$ is nonempty, the map $\mathrm{Pr}_{\mathrm{FP}(T)} G_1 : \mathrm{FP}(T) \to \mathrm{FP}(T)$ is well-defined. Under assumption **A6b**, this map is a contraction, and therefore has a unique fixed point $x^* \in \mathrm{FP}(T)$. Therefore, $\bar{x} = x^*$, i.e., as $\lambda \to 0^+$, the sequence $\{x_\lambda\}$ has at most one limit point x^*, establishing the second part of the theorem. ∎

Note that if the generalized averaging map F_λ satisfies conditions **A1**–**A6a** or **A1**–**A6b**, and the induced trajectory $\{x_\lambda\}$ is bounded, then the trajectory possesses limit points as $\lambda \to 0$, each of them a fixed point of T. Moreover, when G_1 is a contractive mapping (i.e., condition **A6b** applies), the entire trajectory will converge.

Example: To illustrate the limiting conditions, consider the fixed point problem with $K = \Re^2$ and the map $T(x_1, x_2) = (2 - x_1, x_2)$. The set of fixed point solutions is $\mathrm{FP}(T) = \{(1, x_2) : x_2 \in \Re\}$.

The classical function iteration cycles by reflecting about the line of fixed points $FP(T)$. To remedy this cycling behavior we consider special cases of the parameterization scheme (1). The first parameterization considers the map

$$F^1_\lambda(x) = \lambda\frac{x}{2} + (1-\lambda)T(x).$$

In this case, $F^1_\lambda(x) = F^1(G^1_\lambda(x))$ with $F^1 = I$ and $G^1_\lambda = \frac{\lambda}{2}I + (1-\lambda)T$. This averaging map satisfies conditions **A1**–**A6b** and the point $x^* = (1, 0)$ satisfies the condition

$x^* = \mathrm{Pr}_{\mathrm{FP}(T)} G_1^1(x^*)$. The map F_λ^1 creates the averaging trajectory $x_\lambda^1 = \lambda\frac{x_\lambda^1}{2} + (1-\lambda)T(x_\lambda^1)$. Algebraic manipulations yield the explicit expression for x_λ^1 to be $x_\lambda^1 = \left(\frac{4-4\lambda}{4-3\lambda}, 0\right)$, converging to x^* as $\lambda \to 0$.

As another example, let $F_\lambda^2(x) = T(\lambda\frac{x}{2} + (1-\lambda)x) = T((1-\frac{1}{2}\lambda)x)$. In this case, $F_\lambda^2(x) = F^2(G_\lambda^2(x))$ with $F^2 = T$ and $G_\lambda^2 = (1-\frac{1}{2}\lambda)I$. Again, the point $x^* = (1,0)$ is the unique point satisfying the condition $x^* = Pr_{\mathrm{FP}(T)} G_1^2(x^*)$. This map produces the trajectory $x_\lambda^2 = T((1-\frac{1}{2}\lambda)x_\lambda^2)$ that can be expressed explicitly as $x_\lambda^2 = (\frac{4}{4-\lambda}, 0)$. This trajectory also converges to x^* as $\lambda \to 0$. Figure 1 illustrates the averaging trajectories for this example.

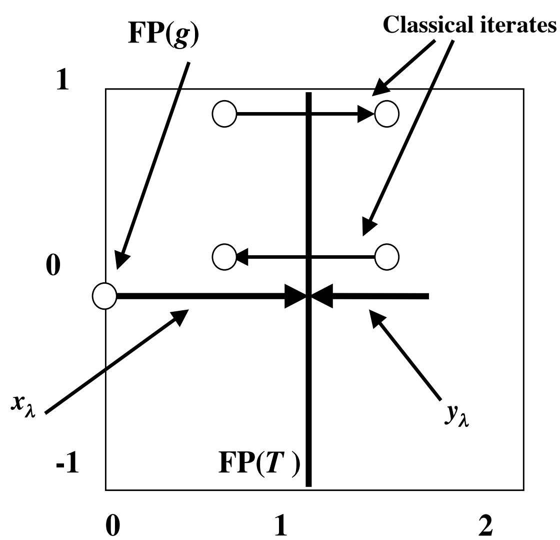

Figure 1 : Averaging Trajectories

Observe that these averaging schemes satisfy conditions **A1–A6b**. Moreover, although the trajectories $\{x_\lambda^1\}$ and $\{x_\lambda^2\}$ are quite different, they both approach the same limit point. As Theorem 1 illustrates, this result is not a coincidence, since the point $x^* = Pr_{\mathrm{FP}(T)} G_1(x^*)$ is the same in both cases.

As yet another example, suppose $F = I$ and

$$G_\lambda^3 = \lambda x + (1-\lambda)(I + c(I-T))^{-1}(x), \text{ with } c > 0.$$

This averaging map satisfies conditions **A1–A6a**, and therefore, the first (weaker) part of Theorem 1 applies[1].

In the next section we show that these types of averaging trajectories are special cases of the averaging framework we have considered in this section.

2.2 Examples of generalized averaging maps

We next introduce several examples of generalized averaging maps $F_\lambda(\cdot)$ and show that they are special cases of the averaging framework we have introduced (i.e., satisfy conditions **A1 – A6**). We assume in these examples that the ground set K is a closed convex set, the map $T : K \to K$ is nonexpansive, and the map $g : K \to K$ is either (i) contractive, or (ii) nonexpansive with fixed points containing those of the original problem (e.g., $g = I$). The special cases to follow typically use various forms of inside and outside averaging.

The fact that the fixed point map T is nonexpansive allows us to apply the proximal point map $J_{c(I-T)} \triangleq (I + c(I - T))^{-1}$ for some constant $c > 0$ to the map $(I - T)$ in some of these examples[2]. For a discussion of the proximal point map and its relationship to fixed point problems, see Eckstein and Bertsekas [6] and Rockafellar [14].

1. **Outside averaging** Let

$$F_\lambda(x) = F_\lambda^{\text{out}}(x) \triangleq \lambda g(x) + (1 - \lambda)T(x).$$

We refer to this map as "outside averaging" since the values it generates are convex combinations of the values of the original map T and values of the map g. The prior literature has examined special cases of outside averaging. Halpern [8] and Browder [3] introduced, and Bauschke [1] and Wittmann [15] further studied, a special case of this map with the constant map $g(x) \equiv x_0$ for some constant $x_0 \in K$. Dunn [5] and Magnanti and Perakis [10, 11] study this type of averaging for the choice of map $g(x) = x$. The developments in this paper allow us to extend the results to a richer class of maps $g(x)$.

Observe that $F_\lambda = F(G_\lambda)$ with $F = I$ and $G_\lambda = \lambda g + (1 - \lambda)T$. As demonstrated next, this parametric scheme satisfies conditions **A1 – A6.**

A1 When g is contractive, $F_\lambda(x)$ is also contractive for all $\lambda \in (0, 1]$ and, therefore, has a fixed point. When g is nonexpansive with nonempty set of fixed points containing those of T, $F_\lambda(x)$ is also nonexpansive with fixed points containing those of T.

A2 $F(\lambda, x) = F_\lambda(x)$ is continuous in (λ, x) since $T(x)$ and $g(x)$ are nonexpansive, and hence continuous.

A3 $F_0 = T$, therefore, condition holds trivially.

A4 Holds trivially

[1] In fact, it is not hard to see that $\text{FP}(F_\lambda) = \text{FP}(T)$ for any $\lambda \in [0, 1]$.

[2] For example, whenever $\text{Image}(2I - T) = \Re^n$, the map $J_{c(I-T)}$ is defined over the entire space $\Re^n$.

A5 Holds trivially, since

$$G_\lambda \equiv \lambda G_1 + (1-\lambda)G_0.$$

A6 Condition **A6a** holds whenever $g(x)$ is nonexpansive; condition **A6b** holds whenever $g(x)$ is a contraction.

It is not hard to see that when g is a contraction, the fixed point x_λ is unique for each $\lambda > 0$, and the trajectory $\{x_\lambda\}$ is bounded for $\lambda \in (0,1]$. Indeed, let $x^* \in \mathrm{FP}(T)$ be an arbitrary fixed point, and suppose the contraction constant of g is $\alpha \in [0,1)$. Then

$$\|x_\lambda - x^*\| \leq \lambda\|g(x_\lambda) - x^*\| + (1-\lambda)\|T(x_\lambda) - T(x^*)\|$$

$$\leq \lambda\|g(x_\lambda)-g(x^*)\|+\lambda\|g(x^*)-x^*\|+(1-\lambda)\|x_\lambda-x^*\| \leq (\alpha\lambda+1-\lambda)\|x_\lambda-x^*\|+\lambda\|g(x^*)-x^*\|,$$

or, after rearranging terms and cancelling $\lambda > 0$,

$$\|x_\lambda - x^*\| \leq \frac{\|g(x^*) - x^*\|}{1-\alpha}.$$

We conclude that the trajectory $\{x_\lambda\}$ is bounded and, therefore, converges to the unique limit point x^* as described in Theorem 1.

2. **Inside averaging:**

$$F_\lambda(x) = F_\lambda^{\mathrm{in}}(x) \triangleq T(\lambda g(x) + (1-\lambda)x).$$

We refer to this map as "inside averaging" since it generates values by applying the map T to the convex combination of the identity map and map $g(x)$, which are, therefore, averaged "inside," or within the argument, of T.

Observe that $F_\lambda = F(G_\lambda)$ for $F = T$ and $G_\lambda = \lambda g + (1-\lambda)I$. As demonstrated next, with one provision, this parametric scheme satisfies conditions **A1** – **A6**. The provision is that when $g(x)$ is nonexpansive, we need to select the sequence $\{x_\lambda\}$ from a bounded set. In Section 3, we discuss a procedure that permits us to eliminate the boundedness condition.

A1 When g is contractive, $F_\lambda(x)$ is also contractive for all $\lambda \in (0,1]$ and, therefore, has a fixed point. When g is nonexpansive with a nonempty set of fixed points containing those of T, $F_\lambda(x)$ is also nonexpansive with fixed points containing those of T.

A2 $F(\lambda, x) = F_\lambda(x)$ is continuous in (λ, x) since $T(x)$ and $g(x)$ are nonexpansive, and hence continuous.

A3 $F_0 = T$, therefore, the condition holds trivially.

A4 The first two statements hold trivially. The last statement of condition **A4** is equivalent to

$$\lim_{\lambda \to 0} \lambda\|g(x_\lambda) - x_\lambda\|^2 = 0. \tag{5}$$

When $g(x)$ is a contraction, the fixed point x_λ is unique for each $\lambda > 0$, and the trajectory $\{x_\lambda\}$ is bounded for $\lambda \in (0,1]$, which implies (5). When $g(x)$ is nonexpansive with $\mathrm{FP}(T) \subseteq \mathrm{FP}(g)$, each map $F_\lambda(x)$ might have more than one fixed point (e.g., any point $x^* \in \mathrm{FP}(T)$ is a fixed point of F_λ). However, as long as we select the sequence $\{x_\lambda\}$ from a bounded set, it would satisfy condition (5).

A5 Holds trivially, since

$$G_\lambda \equiv \lambda G_1 + (1-\lambda)G_0.$$

A6 Condition **A6a** holds whenever $g(x)$ is nonexpansive; condition **A6b** holds whenever $g(x)$ is a contraction.

When g is a contractive map, then an argument similar to the one used in our discussion of outside averaging, shows that the trajectory $\{x_\lambda\}$ is bounded for $\lambda \in (0,1]$ and, therefore, converges to the unique fixed point x^* described in Theorem 1.

3. **Outside averaging with the proximal point map:** Whenever $0 \in K$ and the proximal point map $J_{c(I-T)}(x) = (I + c(I-T))^{-1}(x)$ is defined on K, we can consider the map

$$F_\lambda(x) = (1-\lambda)J_{c(I-T)}(x) \equiv (1-\lambda)(I + c(I-T))^{-1}(x).$$

In this case, $F = I$ and $G_\lambda = \lambda(1-\lambda)J_{c(I-T)}$, and we can establish the validity of conditions **A1** –**A6b** as in our discussion in Example 1, since the map $J_{c(I-T)}$ is nonexpansive with fixed points coinciding with those of T. Moreover, we can replace the constant map $g(x) \equiv 0$ by an arbitrary map $g : K \to K$ satisfying the properties as discussed in Example 1.

4. **Inside averaging with the proximal point map:** Again, assuming $0 \in K$ and the proximal point map $J_{c(I-T)}(x)$ is well defined on K, we consider the map

$$F_\lambda(x) = J_{c(I-T)}((1-\lambda)x) \equiv (I + c(I-T))^{-1}((1-\lambda)x).$$

In this case, $F = J_{c(I-T)}$ and $G_\lambda = (1-\lambda)I$.

Remarks:

1. Note that Examples 3 and 4 are analogous to outside and inside averaging of the map $J_{c(I-T)}$ with $g \equiv 0$. Furthermore, as in Examples 1 and 2, we can apply averaging for a more general contractive map g (or nonexpansive map g whose fixed points contain those of T), or a family of maps g_λ as in Example 6 to follow.

2. It is interesting to compare the trajectories generated by the averaging schemes of Example 1 (with $g(x) \equiv 0$) and Example 4. Let us denote the points on these trajectories by x^1_λ and x^4_λ, respectively. We compare these trajectories by noting that

$$(1-\lambda)x^4_\lambda = (I + c(I-T))(x^4_\lambda) = (c+1)x^4_\lambda - cT(x^4_\lambda).$$

Therefore, $x^4_\lambda = \frac{c}{c+\lambda}T(x^4_\lambda)$, implying that

$$x^4_\lambda = x^1_{\frac{\lambda}{c+\lambda}}.$$

Notice that in this case the trajectory of the fixed points generated by inside averaging with the proximal point mapping (that is, the trajectory $\{x_\lambda^4\}$) is the same as the trajectory of the fixed points generated by outside averaging (that is, the sequence $\{x_\lambda^1\}$). Observe that whenever $c \geq 1$, the trajectory $\{x_\lambda^4\}$ converges faster than the trajectory $\{x_\lambda^1\}$.

5. **Convex combinations of any F_λ maps that satisfy the conditions A1-A6a or A1-A6b**

6. **Averaging with a family of maps g_λ:** In Examples 1–5, we can replace the map g with a family of maps $g_\lambda : K \to K$ if for some constants $\alpha \in [0, 1)$ and $\gamma \geq 0$:

$$\|g_\lambda(x) - g_\lambda(y)\| \leq \alpha \|x - y\| \ \forall x, y \in K, \ \forall \lambda \in [0, 1]$$

and

$$\|g_{\lambda^1}(x) - g_{\lambda^2}(x)\| \leq \gamma |\lambda^1 - \lambda^2|, \ \forall x \in K, \ \forall \lambda^1, \lambda^2 \in [0, 1].$$

The resulting family F_λ satisfies conditions **A1**–**A3**, and therefore Proposition 3 applies. The resulting parametric scheme might violate condition **A5** and, therefore, Theorem 1 might not be directly applicable to the limit points of the trajectory of fixed points induced by this family. However, it is not hard to show that, as λ approaches 0, this trajectory approaches the trajectory of fixed points induced by (inside or outside) averaging with the map g_0, and consequently converges to $x^* = \Pr_{\mathrm{FP}(T)}(g_0(x^*))$. This extension allows us to incorporate inexact computations on the early stages of the averaging (i.e., for larger values of λ), as long as the magnitude of errors decreases as λ approaches 0.

3 Approximate averaging trajectories — a first approach

The averaging framework described in the previous section provides an intuitive way of approximating fixed point solutions of a map T with the trajectory of fixed points solutions induced from a class of parameterized subproblems. Following this trajectory exactly and computing the fixed points of the parameterized subproblems typically will be very expensive computationally. As an extreme example, note that the trivial averaging scheme $F_\lambda \equiv T$ satisfies conditions **A1**–**A6**. Therefore, we would like to solve the original fixed point problem by choosing the parametric families F_λ more judiciously, i.e., by imposing stronger assumptions. These stronger assumptions will also allow us to compute points that lie close to the trajectory of fixed points for the parameterized subproblems, rather than solving the subproblems exactly, while still guaranteeing convergence to the desired fixed points of T.

To motivate the general approach in this section, consider the iterative scheme that approximates the trajectory of fixed points induced by outside averaging of Example 1:

$$x_0 \in K, \ x_{k+1} = \lambda_{k+1} g(x_k) + (1 - \lambda_{k+1}) T(x_k), k > 0.$$

Halpern [8] first considered this iterative scheme and Wittmann [15] and Bauschke [?] further analyzed it, all for the case of constant map g. Averaging with the identity map (see [5], [10], [11]) is another special case of this scheme. The framework we introduce in this paper, however, extends well beyond these very special cases and includes as special cases, among many others, inside and outside averaging with contractive maps g. As a first approach, in this section we consider the following solution scheme that approximates the trajectory of fixed points for the parameterized subproblems:

$$x_0 \in K, \ x_{k+1} = F_{\lambda_{k+1}}(x_k), \ k > 0, \tag{6}$$

with the sequence of parameter values $\lambda_k \in (0,1]$ and the family of functions $F_\lambda : K \to K$ satisfying the following conditions:

B1 $\mathrm{FP}(T) \supseteq \mathrm{FP}(F_0) \neq \emptyset$

B2 For some constant $\alpha \in (0,1]$,

$$\|F_\lambda(x) - F_\lambda(y)\| \leq (1 - \alpha\lambda)\|x - y\|,$$

for any $\lambda \in [0,1]$ and any $x, y \in K$.

B3 For some constant $L > 0$,

$$\|F_\lambda(x) - F_\mu(x)\| \leq L|\lambda - \mu| \cdot \|x\|,$$

for any $\lambda, \mu \in [0,1]$ and any $x \in K$.

B4 $\lambda_k \to 0^+, \ \sum_{k=1}^{\infty} \lambda_k = +\infty, \ \sum_{k=1}^{\infty} |\lambda_k - \lambda_{k+1}| < +\infty.$

Observe that conditions **B1–B3** imply conditions **A1–A3**.

Before analyzing the behavior of the general averaging map $F_\lambda(x)$, we examine several special cases. In particular, we show that several iterative schemes which are the algorithmic extensions of the examples of the previous section satisfy properties **B1–B4**.

In the following examples, we assume that the map $T : K \to K$ is nonexpansive with $\mathrm{FP}(T) \neq \emptyset$ and the "step sizes" $\{\lambda_k\}$ satisfy condition **B4**.

Examples of the iterative scheme:

1. **Outside averaging:**

$$x_{k+1} = \lambda_{k+1} g(x_k) + (1 - \lambda_{k+1}) T(x_k),$$

for a contractive map $g : K \to K$ with bounded norm and a contraction constant $a \in [0,1)$. In this case,

$$F_\lambda(x) = \lambda g(x) + (1 - \lambda) T(x),$$

and

B1 $F_0(x) = T(x)$, therefore, $\mathrm{FP}(T) = \mathrm{FP}(F_0)$.

B2 For any λ, x, y,

$$\|F_\lambda(x) - F_\lambda(y\| \leq \lambda a\|x - y\| + (1 - \lambda)\|x - y\| = (1 - \alpha\lambda)\|x - y\|$$

with $\alpha = (1 - a) \in (0, 1]$.

B3 For any λ, μ, x,

$$\|F_\lambda(x) - F_\mu(x)\| \leq |\lambda - \mu|\|g(x)\| + |\lambda - \mu|\|T(x)\| = |\lambda - \mu|(\|g(x)\| + \|T(x)\|),$$

which implies the condition

2. **Inside averaging:**

$$x_{k+1} = T(\lambda_{k+1}g(x_k) + (1 - \lambda_{k+1})x_k),$$

for a contractive map $g : K \to K$ with bounded norm and a contraction constant $a \in [0, 1)$. In this case,

$$F_\lambda(x) = T(\lambda g(x) + (1 - \lambda)x),$$

and

B1 $F_0(x) = T(x)$, therefore, $\mathrm{FP}(T) = \mathrm{FP}(F_0)$.

B2 Since T is nonexpansive, for any λ, x, y,

$$\|F_\lambda(x) - F_\lambda(y)\| \leq \|\lambda g(x) + (1 - \lambda)x - \lambda g(y) - (1 - \lambda)y\|$$

$$\leq \lambda a\|x - y\| + (1 - \lambda)\|x - y\| = (1 - \alpha\lambda)\|x - y\|$$

with $\alpha = 1 - a \in (0, 1]$.

B3 For any λ, μ, x,

$$\|F_\lambda(x) - F_\mu(x)\| \leq \|\lambda g(x) + (1 - \lambda)x - \mu g(x) - (1 - \mu)x\|$$

$$\leq |\lambda - \mu|\|g(x)\| + |\lambda - \mu|\|x\| = |\lambda - \mu|(\|g(x)\| + \|x\|),$$

which implies the condition.

3. **Outside averaging with the proximal point mapping**: Assume $0 \in K$ and the map $J_{c(I-T)}$ is defined on K. Then we can define

$$x_{k+1} = (1 - \lambda_{k+1})J_{c(I-T)}(x_k) \equiv (1 - \lambda_{k+1})(I + c(I - T))^{-1}(x_k).$$

In this case,

$$F_\lambda(x) = (1 - \lambda)J_{c(I-T)}(x),$$

and

B1 $F_0(x) = (I + c(I - T))^{-1}(x)$, therefore, $FP(F_0) = FP(T)$.

B2 Since $J_{c(I-T)}$ is a nonexpansive map, for any λ, x, y,

$$\|F_\lambda(x) - F_\lambda(y)\| \leq (1-\lambda)\|x-y\|,$$

satisfying the condition with $\alpha = 1$.

B3 For any λ, μ, x,

$$\|F_\lambda(x) - F_\mu(x)\| = \|(1-\lambda)J_{c(I-T)}(x) - (1-\mu)J_{c(I-T)}(x)\| = |\lambda - \mu| \cdot \|J_{c(I-T)}(x)\|,$$

which implies the condition.

4. **Inside averaging with the proximal point mapping**: Again, assume that $0 \in K$ and $J_{c(I-T)}$ is well-defined on K. Let

$$x_{k+1} = (I + c(I-T))^{-1}((1-\lambda_{k+1})x_k).$$

In this case,

$$F_\lambda(x) = J_{c(I-T)}((1-\lambda)x),$$

and

B1 $F_0(x) = (I + c(I-T))^{-1}(x)$ and, therefore, $FP(F_0) = FP(T)$.

B2 Since $J_{c(I-T)}$ is a nonexpansive map, for any λ, x, y,

$$\|F_\lambda(x) - F_\lambda(y)\| \leq (1-\lambda)\|x-y\|,$$

satisfying the condition with $\alpha = 1$.

B3 Since $J_{c(I-T)}$ is nonexpansive, for any λ, μ, x,

$$\|F_\lambda(x) - F_\mu(x)\| \leq |\lambda - \mu| \cdot \|x\|,$$

which implies the condition.

5. **Convex combinations**: Similar to the previous cases.

The following results (Theorems 2 and 3) extend to the general iterative scheme (6) the results of Wittmann [15] and Bauschke [1], which, essentially, concerned the case of outside averaging with a constant map. Some of the proof techniques used in establishing these theorems are similar to those used by Bauschke.

Theorem 2 *Suppose $T : K \to K$ is a nonexpansive map with* $\mathrm{FP}(T) \neq \emptyset$. *Suppose further that the sequence $\{\lambda_k\}$ and the family of functions F_λ satisfy conditions **B1 — B4**. Then every limit point of the sequence $\{x_k\}$ induced by the averaging scheme (6) is a fixed point of the map T.*

Proof: We begin by showing that the sequence $\{x_k\}$ induced by the averaging scheme (6) is bounded. Let x^* be an arbitrary point in $\mathrm{FP}(T)$. Define

$$D \equiv \max\left\{\|x_0 - x^*\|,\ \frac{\|x^*\|L}{\alpha}\right\},$$

with the constants $L > 0$ and $\alpha \in (0,1]$ specified as in conditions **B2** and **B3**. We will show by induction that $\|x_k - x^*\| \le D$ for any $k \ge 0$. This inequality is valid by definition for $k = 0$. Suppose the inequality is valid for some $k \ge 0$. Then

$$\|x_{k+1} - x^*\| = \|F_{\lambda_{k+1}}(x_k) - F_0(x^*)\| \le \|F_{\lambda_{k+1}}(x_k) - F_{\lambda_{k+1}}(x^*)\| + \|F_{\lambda_{k+1}}(x^*) - F_0(x^*)\|$$

$$\le (1 - \alpha\lambda_{k+1})\|x_k - x^*\| + \lambda_{k+1}L\|x^*\|$$

$$\le (1 - \alpha\lambda_{k+1})D + \alpha D\lambda_{k+1} = D,$$

establishing the desired relationship.

Since the sequence $\{x_k\}$ is bounded, for some constant $C > 0$, $L\|x_k\| \le C$ and $\|x_k - x_{k+1}\| \le C$ for all $k \ge 0$. Then for any $k \ge 0$,

$$\|x_{k+1} - x_k\| = \|F_{\lambda_{k+1}}(x_k) - F_{\lambda_k}(x_{k-1})\|$$

$$\le \|F_{\lambda_{k+1}}(x_k) - F_{\lambda_{k+1}}(x_{k-1})\| + \|F_{\lambda_{k+1}}(x_{k-1}) - F_{\lambda_k}(x_{k-1})\|$$

$$\le (1 - \alpha\lambda_{k+1})\|x_k - x_{k-1}\| + |\lambda_{k+1} - \lambda_k| \cdot L\|x_{k-1}\|$$

$$\le (1 - \alpha\lambda_{k+1})\|x_k - x_{k-1}\| + |\lambda_{k+1} - \lambda_k|C.$$

Applying this relationship inductively, and using the fact that $1 - \alpha\lambda_k \le 1$ for all k, shows that for all $k \ge m > 0$

$$\|x_{k+1} - x_k\| \le \|x_m - x_{m-1}\| \prod_{i=m}^{k}(1 - \alpha\lambda_{i+1}) + C\sum_{i=m}^{k}|\lambda_{i+1} - \lambda_i|.$$

Thus,

$$\overline{\lim_{k\to\infty}}\|x_{k+1} - x_k\| \le C\prod_{i=m}^{\infty}(1 - \alpha\lambda_{i+1}) + C\sum_{i=m}^{\infty}|\lambda_{i+1} - \lambda_i|. \tag{7}$$

The assumptions $\lim_{k\to\infty}\lambda_k = 0$ and $\sum_{k=1}^{\infty}\lambda_k = +\infty$ from **B4** imply that $\prod_{i=m}^{\infty}(1 - \alpha\lambda_{i+1}) = 0$, (see, for example, Proposition 2.1 of [1]). Therefore, $\lim_{m\to\infty}\prod_{i=m}^{\infty}(1 - \alpha\lambda_{i+1}) = 0$. Additionally, from assumption **B4**, $\lim_{m\to\infty}\sum_{i=m}^{\infty}|\lambda_{i+1} - \lambda_i| = 0$. Letting $m \to \infty$ in (7), we obtain $\lim_{k\to\infty}\|x_{k+1} - x_k\| = 0$. Also,

$$\|x_k - F_0(x_k)\| \le \|x_k - x_{k+1}\| + \|x_{k+1} - F_0(x_k)\|$$

$$= \|x_k - x_{k+1}\| + \|F_{\lambda_{k+1}}(x_k) - F_0(x_k)\|$$

$$\le \|x_k - x_{k+1}\| + \lambda_{k+1}L\|x_k\| \le \|x_k - x_{k+1}\| + \lambda_{k+1}C \longrightarrow_{k\to\infty} 0.$$

Therefore, every limit point of the sequence $\{x_k\}$ is a fixed point solution of the map F_0. Property **B1** implies that any limit point is also a fixed point solution of T. In addition, since the sequence $\{x_k\}$ is bounded, it has at least one limit point. ∎

The previous theorem states that the limit points of the sequence induced by the iterative scheme (6) are, indeed, fixed points of T. Can we, as before, characterize these limiting fixed points? To do so, we consider a stronger version of assumptions **B1**–**B4**, obtained by replacing conditions **B2** and **B3** with the following stronger conditions.

B2′ $F_\lambda = F(G_\lambda)$, with $F : K \to K$ and $G_\lambda : K \to K$ satisfying the assumptions:

- F is nonexpansive and $\mathrm{FP}(T) \subseteq \mathrm{FP}(F)$.
- G_0 is nonexpansive, $\mathrm{FP}(T) \subseteq \mathrm{FP}(G_0)$, and there exists $\alpha \in (0, 1]$ such that

$$\|G_\lambda(x) - G_\lambda(y)\| \le (1 - \alpha\lambda)\|x - y\|, \ \forall x, y \in K$$

- For all $\lambda \in [0, 1]$, all $x \in K$ and all $x^* \in \mathrm{FP}(T)$,

$$G_\lambda(x^*)^t(x - x^*) \le \lambda G_1(x^*)^t(x - x^*) + (1 - \lambda)G_0(x^*)^t(x - x^*).$$

B3′ For some constant $L > 0$, $\|G_\lambda(x) - G_\mu(x)\| \le L|\lambda - \mu| \cdot \|x\|$ for any $\lambda, \mu \in [0, 1]$ and any $x \in K$.

As demonstrated in the following proposition, the last assumption of condition **B2′** is essentially a weaker version of condition **A5** of Section 2.

Proposition 4 *Suppose the family of maps $G_\lambda : K \to K$, $\lambda \in [0, 1]$ satisfies condition **A5**, i.e., for all $\lambda \in [0, 1]$, all $x \in K$ and all $x^* \in \mathrm{FP}(T)$,*

$$G_\lambda(x)^t(x - x^*) \le \lambda G_1(x)^t(x - x^*) + (1 - \lambda)G_0(x)^t(x - x^*).$$

Suppose further that for all $\lambda \in [0, 1]$, the map G_λ is continuous. Then

$$G_\lambda(x^*)^t(x - x^*) \le \lambda G_1(x^*)^t(x - x^*) + (1 - \lambda)G_0(x^*)^t(x - x^*)$$

for all $\lambda \in [0, 1]$, all $x \in K$ and all $x^* \in \mathrm{FP}(T)$.

Proof: Let $\lambda \in [0, 1]$, $x^* \in \mathrm{FP}(T)$ and $y \in K$ be arbitrary. For $t \in (0, 1]$, let $x = ty + (1 - t)x^* \in K$. Then by supposition,

$$tG_\lambda(ty + (1 - t)x^*)^t(y - x^*) \le t[\lambda G_1(ty + (1 - t)x^*) + (1 - \lambda)G_0(ty + (1 - t)x^*)]^t(y - x^*),$$

and since $t > 0$,

$$G_\lambda(ty + (1 - t)x^*)^t(y - x^*) \le [\lambda G_1(ty + (1 - t)x^*) + (1 - \lambda)G_0(ty + (1 - t)x^*)]^t(y - x^*)$$

Letting $t \to 0^+$, this inequality implies that

$$G_\lambda(x^*)^t(y - x^*) \le \lambda G_1(x^*)^t(y - x^*) + (1 - \lambda)G_0(x^*)^t(y - x^*). \quad ∎$$

Note that the Examples 1 through 4 of generalized averaging mappings satisfy these new conditions with F and G_λ defined as follows:

1. $F = I$ and $G_\lambda = \lambda g(x) + (1-\lambda)T(x)$.
2. $F = T$ and $G_\lambda = \lambda g(x) + (1-\lambda)x$.
3. $F = I$ and $G_\lambda = (1-\lambda)J_{c(I-T)}$.
4. $F = J_{c(I-T)}$ and $G_\lambda = (1-\lambda)x$.

The next result establishes convergence of the sequence $\{x_k\}$ induced by the iterative scheme (6) if it satisfies conditions **B1**, **B2′**, **B3′** and **B4**.

Theorem 3 *Suppose $T : K \to K$ is a nonexpansive map with $\mathrm{FP}(T) \neq \emptyset$. Suppose further that the sequence $\{\lambda_k\}$ and the family of functions F_λ satisfy conditions **B1**, **B2′**, **B3′** and **B4**. Then the sequence $\{x_k\}$ induced by the iterative scheme (6) converges to the fixed point solution $x^* = \mathrm{Pr}_{\mathrm{FP}(T)} G_1(x^*) \in \mathrm{FP}(T)$.*

Proof: First observe that condition **B2′** implies that the map $G_1(x)$ is a contraction. Therefore, the map $\mathrm{Pr}_{\mathrm{FP}(T)} G_1 : \mathrm{FP}(T) \to \mathrm{FP}(T)$ is also a contraction, and so it has a unique fixed point $x^* \in \mathrm{FP}(T) \subseteq \mathrm{FP}(G_0)$. The definition of the Euclidean projection implies that

$$(\bar{x} - x^*)^t (G_1(x^*) - x^*) \leq 0 \text{ for all } \bar{x} \in \mathrm{FP}(T). \tag{8}$$

From Theorem 2, any limit point of the sequence $\{x_k\}$ induced by the iterates (6) is a fixed point of T and therefore, of G_0. Consequently,

$$\limsup_{k\to\infty} (G_0(x_k) - x^*)^t (G_1(x^*) - x^*) = \limsup_{k\to\infty} (x_k - x^*)^t (G_1(x^*) - x^*) \leq 0,$$

with the inequality following from expression (8). Therefore, for any $\epsilon > 0$,

$$(G_0(x_k) - x^*)^t (G_1(x^*) - x^*) \leq \epsilon \tag{9}$$

for k sufficiently large.

Define $C = \max\{L\|x^*\|,\ \sup_{k\geq 0} L\|x_k\|\}$ (recall from the proof of Theorem 2 that the supremum is finite). For any $\epsilon > 0$,

$$2C^2 \lambda_k \leq \epsilon \tag{10}$$

for k sufficiently large.

Let $\epsilon > 0$ be arbitrary, and suppose k is sufficiently large as required by conditions (9) and (10). Then if we let $\lambda = \lambda_{k+1}$,

$$\|x_{k+1} - x^*\|^2 = \|F(G_\lambda(x_k)) - F(G_0(x^*))\|^2 \leq \|G_\lambda(x_k) - G_0(x^*)\|^2$$

$$= \|G_\lambda(x_k) - G_\lambda(x^*)\|^2 + \|G_\lambda(x^*) - G_0(x^*)\|^2 + 2(G_\lambda(x_k) - G_\lambda(x^*))^t (G_\lambda(x^*) - G_0(x^*))$$

$$= \|G_\lambda(x_k) - G_\lambda(x^*)\|^2 + \|G_\lambda(x^*) - G_0(x^*)\|^2 + 2(G_\lambda(x_k) - G_0(x_k))^t (G_\lambda(x^*) - G_0(x^*))$$

$$+2(G_0(x_k) - G_0(x^*))^t (G_\lambda(x^*) - G_0(x^*)) - 2\|G_\lambda(x^*) - G_0(x^*)\|^2$$

$$\leq (1-\alpha\lambda)^2\|x_k - x^*\|^2 + 2C^2\lambda^2 + 2(G_0(x_k) - G_0(x^*))^t(G_\lambda(x^*) - G_0(x^*)).$$

To obtain the previous inequality, we have made use of the Cauchy-Schwartz inequality and conditions **B2′** and **B3′**. Applying condition **B2′** with $x = G_0(x_k)$ to the last term in this expression and using the fact that $G_0(x^*) = x^*$ gives

$$\|x_{k+1} - x^*\|^2 \leq (1-\alpha\lambda)^2\|x_k - x^*\|^2 + 2C^2\lambda^2 + 2\lambda(G_0(x_k) - x^*)^t(G_1(x^*) - G_0(x^*))$$

$$\leq (1-\alpha\lambda)^2\|x_k - x^*\|^2 + 2C^2\lambda^2 + 2\lambda\epsilon \leq (1-\alpha\lambda)^2\|x_k - x^*\|^2 + 3\lambda\epsilon.$$

Recall that $\lambda = \lambda_{k+1}$. Since $1 - \alpha\lambda_{k+1} \leq 1$,

$$\|x_{k+1} - x^*\|^2 \leq (1-\alpha\lambda_{k+1})\|x_k - x^*\|^2 + 3\epsilon\lambda_{k+1}. \tag{11}$$

Let

$$\alpha_m = 1 - \prod_{i=m}^{k}(1-\alpha\lambda_{i+1}),\ m \leq k.$$

Observe that $\alpha_m \in [0,1]$ for any m. We will show that

$$\|x_{k+1} - x^*\|^2 \leq \frac{3}{\alpha}\epsilon\alpha_m + (1-\alpha_m)\|x_m - x^*\|^2 \tag{12}$$

for any $m \leq k$, with m chosen sufficiently large to satisfy conditions (9) and (10).

For $m = k$, $\alpha_k = \alpha\lambda_{k+1}$, and we rewrite relationship (11) as

$$\|x_{k+1} - x^*\|^2 \leq (1-\alpha_k)\|x_k - x^*\| + \frac{3\epsilon}{\alpha}\alpha_k,$$

establishing (12) for $m = k$.

Suppose (12) holds for some $m \leq k$. By definition,

$$\alpha_{m-1} = 1 - (1-\alpha\lambda_m)(1-\alpha_m) = \alpha_m + \alpha\lambda_m(1-\alpha_m).$$

Applying (11) to the term $\|x_m - x^*\|^2$, we obtain

$$\|x_{k+1} - x^*\|^2 \leq (1-\alpha_m)\|x_m - x^*\|^2 + \frac{3}{\alpha}\epsilon\alpha_m$$

$$\leq (1-\alpha_m)\left((1-\alpha\lambda_m)\|x_{m-1} - x^*\|^2 + 3\epsilon\lambda_m\right) + \frac{3}{\alpha}\epsilon\alpha_m$$

$$= (1-\alpha_{m-1})\|x_{m-1} - x^*\|^2 + \frac{3}{\alpha}\epsilon(\alpha_m + (1-\alpha_m)\alpha\lambda_m)$$

$$= (1-\alpha_{m-1})\|x_{m-1} - x^*\|^2 + \frac{3}{\alpha}\epsilon\alpha_{m-1},$$

establishing (12).

We conclude that for m sufficiently large,

$$\|x_{k+1} - x^*\|^2 \leq \frac{3}{\alpha}\epsilon\alpha_m + \|x_m - x^*\|^2\prod_{i=m}^{k}(1-\alpha\lambda_{i+1})$$

for any $k \geq m$. Note that the sequence $\{\alpha\lambda_i\} \in [0,1)$ converges to 0 and satisfies

$$\sum_{i=1}^{\infty} (\alpha\lambda_i) = +\infty.$$

Therefore, $\prod_{i=1}^{\infty}(1-\alpha\lambda_i) = 0$ (see, for example, Proposition 2.1 of Bauschke [1]), and so

$$\overline{\lim_k} \|x_{k+1} - x^*\|^2 \leq \frac{3}{\alpha}\epsilon.$$

Therefore, since ϵ is arbitrarily small,

$$\|x_{k+1} - x^*\| \longrightarrow_{k\to\infty} 0.$$

We conclude that the sequence $\{x_k\}$ converges to the fixed point solution x^* of T satisfying the property $x^* = \Pr_{\mathrm{FP}(T)}(G_1(x^*))$, completing the proof. ∎

Remark:

Following Halpern [8], as an alternative of the approximate averaging framework we have introduced, we could consider multiple applications of the averaging map $F_{\lambda_{k+1}}$ rather than one. As a result we can impose less restrictive assumptions on the sequence of step sizes $\{\lambda_k\}$. By applying the averaging map $F_{\lambda_{k+1}}$ several times, we approximate the trajectory of the parameterized fixed point subproblems more accurately.

To provide some motivation, consider the sequence x_λ induced by the parameterized fixed point subproblem $x_\lambda = F_\lambda(x_\lambda)$ as in (2). If conditions **B1**–**B3** are valid, then the trajectory $\{x_\lambda\}$ is bounded, and as $\lambda \to 0^+$, the limit points of x_λ are fixed point solutions of the original problem. If, in addition, F_λ is a contractive map for every $\lambda > 0$, then we can compute the fixed point x_λ as

$$x_\lambda = \lim_{k\to\infty} F_\lambda^k(X_0) \tag{13}$$

for any starting point $X_0 \in K$. We therefore consider an iterative scheme that "follows" the path of fixed points $\{x_\lambda\}$. That is, given a sequence $\lambda_k \to 0$, $\lambda_k \in (0,1]$ and a sequence of nonnegative integers $\{m_k\}$, for an arbitrary $X_0 \in K$, we can approximate the original fixed point problem through

$$X_{k+1} = F_{\lambda_{k+1}}^{m_{k+1}}(X_k) \tag{14}$$

by ensuring that the iterates X_k stay "close" to the fixed point solutions of the corresponding subproblems $x_{\lambda_k} = F_{\lambda_k}(x_{\lambda_k})$.

The following result formalizes this iterative scheme.

Proposition 5 *Suppose $F_\lambda : K \to K$ satisfies conditions **B1**–**B3**. Assume the sequences $\{\lambda_k\}$ and $\{m_k\}$ satisfy the conditions $\lambda_k \to 0$ and $(1-\alpha\lambda_k)^{m_k} \to 0$ as $k \to \infty$. Then the limit points of the sequence $\{X_k\}$ induced by relation (14) are fixed points of T.*

Proof: It is not hard to see that the trajectory $\{x_\lambda\}_{\lambda\in(0,1]}$ is bounded. Moreover, following an argument similar to the one employed in the proof of Theorem 2, we can show that the sequence $\{X_k\}$ is also bounded. Consequently, for some constant L, $\|X_k - x_\lambda\| \leq L$ for all $k \geq 0$ and all $\lambda \in (0,1]$. Furthermore, since x_{λ_k} is a fixed point solution of the map F_{λ_k} and, therefore, of the map $F_{\lambda_k}^{m_k}$,

$$\|X_{k+1} - x_{\lambda_{k+1}}\| = \|F_{\lambda_{k+1}}^{m_{k+1}}(X_k) - F_{\lambda_{k+1}}^{m_{k+1}}(x_{\lambda_{k+1}},)\|$$
$$\leq (1-\alpha\lambda_{k+1})^{m_{k+1}}\|X_k - x_{\lambda_{k+1}}\| \leq (1-\alpha\lambda_{k+1})^{m_{k+1}} L \longrightarrow_{k\to\infty} 0.$$

Therefore, the set of limit points of the sequence $\{X_k\}$ are contained in the set of limit points of the trajectory $\{x_\lambda\}_{\lambda\in(0,1]}$, which, by Theorem 2, are fixed points of T. ∎

4 Approximate averaging trajectories — a second approach

In the previous section, we introduced a general averaging framework that induced points lying close to the trajectory of fixed points of a class of parameterized subproblems. The averaging maps F_λ for this analysis were contractive for each fixed value of $\lambda \in (0,1)$. We next study a similar general averaging framework when the map F_λ is nonexpansive rather than contractive. To compensate for this weaker property on map F_λ, we need the additional assumption that its fixed points contain those of the original map T. The points generated in this case approximate the trajectory of fixed points of the parameterized subproblems FP(F_λ) satisfying assumptions **A1-A6b**. Averaging with the identity map (line search procedures) is a type of averaging that is a special case. This averaging scheme has received much attention in the literature (see for example among others [5], [10], [11]). For this type of averaging, the averaging map F_λ is nonexpansive rather than contractive and its fixed point solutions include those of the original map T. Our averaging framework in this section incorporates this type of averaging.

In particular, as before, we will consider the sequence of iterates

$$x_0 \in K,\ x_{k+1} = F_{\lambda_{k+1}}(x_k),\ k > 0, \tag{15}$$

with the sequence of parameter values $\lambda_k \in (0,1]$ and the family of functions $F_\lambda : K \to K$ satisfying the following conditions:

C1 FP(F_0) = FP(T) $\neq \emptyset$.

C2 $F_\lambda = F(G_\lambda)$, with $F : K \to K$ and $G_\lambda : K \to K$ satisfying the assumptions

- F is nonexpansive and FP(T) $\subseteq$ FP(F).
- G_0 is nonexpansive and FP(T) = FP(G_0).
- G_1 is nonexpansive and FP(T) $\subseteq$ FP(G_1).

C3 For all $\lambda \in [0,1]$, all $x^* \in$ FP(T) and all $x \in K$,

$$G_\lambda(x^*)^t(x-x^*) \leq \lambda G_1(x^*)^t(x-x^*) + (1-\lambda)G_0(x^*)^t(x-x^*),$$

which in light of **C2** implies

$$(G_\lambda(x^*) - G_0(x^*))^t(x-x^*) \leq 0.$$

C4 For all $\lambda \in [0,1]$, all $x^* \in \mathrm{FP}(T)$ and all $x \in K$,

$$\|G_\lambda(x) - G_\lambda(x^*)\|^2 \le \|x - x^*\|^2 - \gamma(\lambda)\|x - G_0(x)\|^2,$$

for some function $\gamma : [0,1] \to \mathbb{R}_+$.

C5 The step sizes λ_k satisfy the conditions $\sum_{k=1}^{\infty} \gamma(\lambda_k) = +\infty$.

Once again, we begin by examining several special cases:

1. **Averaging with the identity map:** We begin by considering a classical averaging scheme and describing how it fits into this generalized averaging framework:

$$F_\lambda(x) = \lambda x + (1-\lambda)T(x).$$

This averaging scheme (see Dunn [5] and Magnanti and Perakis [10], [11]) is applicable for a map T that is nonexpansive.

C1 $F_0 = T$, so the condition holds trivially.

C2 $F = I,\ G_\lambda = \lambda I + (1-\lambda)T$, i.e., $G_0 = T$ and $G_1 = I$, so the condition is satisfied.

C3 Holds trivially since $G_\lambda = \lambda G_1 + (1-\lambda)G_0$.

C4 For all $\lambda \in [0,1]$, all $x^* \in \mathrm{FP}(T)$ and all $x \in K$, since $x^* = G_\lambda(x^*)$, we have (see Dunn [5]):

$$\|G_\lambda(x) - G_\lambda(x^*)\|^2 + \lambda(1-\lambda)\|T(x) - x\|^2 = \|\lambda x + (1-\lambda)T(x) - x^*\|^2 + \lambda(1-\lambda)\|T(x) - x\|^2$$
$$= \|(1-\lambda)(T(x) - x) + (x - x^*)\|^2 + \lambda(1-\lambda)\|T(x) - x\|^2$$
$$= 2(1-\lambda)(T(x) - x)^t(x - x^*) + \|x - x^*\|^2 + (1-\lambda)\|T(x) - x\|^2$$
$$= (1-\lambda)\left(2(T(x) - x)^t(x - x^*) + \|T(x) - x\|^2 + \|x - x^*\|^2\right) + \lambda\|x - x^*\|^2$$
$$= (1-\lambda)\|T(x) - x^*\|^2 + \lambda\|x - x^*\|^2 \le \|x - x^*\|^2,$$

i.e.,

$$\|G_\lambda(x) - G_\lambda(x^*)\|^2 \le \|x - x^*\|^2 - \lambda(1-\lambda)\|T(x) - x\|^2 = \|x - x^*\|^2 - \lambda(1-\lambda)\|G_0(x) - x\|^2.$$

Therefore, the condition is satisfied with $\gamma(\lambda) = \lambda(1-\lambda)$.

C5 The condition can be satisfied by selecting the sequence $\{\lambda_k\} \subset [0,1]$ to satisfy $\sum_{k=1}^{\infty} \lambda_k(1-\lambda_k) = \infty$, just as in the cited papers.

2. **Outside averaging:**

$$F_\lambda^{\mathrm{out}}(x) = \lambda g(x) + (1-\lambda)T(x)$$

We consider the case when the map T is firmly nonexpansive and the map g is nonexpansive with $\mathrm{FP}(g) \supseteq \mathrm{FP}(T)$, (an example of a map g is the identity map).

C1 $F_0 = T$, so the condition holds trivially

C2 $F = I$, $G_\lambda = \lambda g + (1-\lambda)T$, i.e., $G_0 = T$ and $G_1 = g$ — by the above assumptions, the condition is satisfied.

C3 Holds at equality, since $G_\lambda = \lambda G_1 + (1-\lambda)G_0$.

C4 Observe that $G_\lambda(x^*) = x^*$. Therefore,

$$\|G_\lambda(x) - G_\lambda(x^*)\|^2 = \|\lambda g(x) + (1-\lambda)T(x) - x^*\|^2 =$$

$$= \lambda^2\|g(x) - x^*\|^2 + (1-\lambda)^2\|T(x) - x^*\|^2 + 2\lambda(1-\lambda)(g(x) - x^*)^t(T(x) - x^*) \leq$$

$$\lambda^2\|g(x)-x^*\|^2+(1-\lambda)^2\|T(x)-x^*\|^2+\lambda(1-\lambda)\|g(x)-x^*\|^2+\lambda(1-\lambda)\|T(x)-x^*\|^2 =$$

$$\lambda\|g(x)-x^*\|^2+(1-\lambda)\|T(x)-x^*\|^2 \leq \lambda\|g(x)-x^*\|^2+(1-\lambda)\|x-x^*\|^2-(1-\lambda)\|T(x)-x\|^2$$

$$\leq \|x - x^*\|^2 - (1-\lambda)\|G_0(x) - x\|^2.$$

The second inequality follows since T is firmly nonexpansive. Therefore, the condition is satisfied with $\gamma(\lambda) = (1-\lambda)$.

C5 This condition can be satisfied by selecting the sequence $\{\lambda_k\} \subseteq (0,1]$ to satisfy $\sum_{k=1}^{\infty}(1-\lambda_k) = \infty$.

Observe that in this method, g could be the identity map. In this case, the method reduces to averaging with the identity map. In fact, for the case of averaging with the identity map, it is sufficient to require that that map T is nonexpansive rather than firmly nonexpansive (see for example [5], [10], [11]).

3. **Outside averaging with the proximal point mapping:**

$$F_\lambda(x) = \lambda g(x) + (1-\lambda)J_{c(I-T)}(x) = \lambda g(x) + (1-\lambda)(I + c(I-T))^{-1}(x).$$

We consider the case when the map T is nonexpansive and the map g is nonexpansive with $\mathrm{FP}(g) \supseteq \mathrm{FP}(T)$.

This is essentially an extension of the previous case, since the proximal point operator $J_{c(I-T)}$ is firmly nonexpansive whenever T is nonexpansive.

Theorem 4 *Suppose $T : K \to K$ is a nonexpansive map with $\mathrm{FP}(T) \neq \emptyset$. Suppose further that the sequence $\{\lambda_k\}$ and the family of functions F_λ satisfy conditions **C1–C6**. Then the sequence $\{x_k\}$ induced by the averaging scheme (15) converges to a fixed point of the map T.*

Proof: Let $x^* \in \mathrm{FP}(T)$. Then

$$\|G_{\lambda_{k+1}}(x_k) - x^*\|^2 = \|G_{\lambda_{k+1}}(x_k) - G_0(x^*)\|^2$$

$$= \|G_{\lambda_{k+1}}(x_k)-G_{\lambda_{k+1}}(x^*)\|^2-\|G_{\lambda_{k+1}}(x^*)-G_0(x^*)\|^2+2(G_{\lambda_{k+1}}(x_k)-G_0(x^*))^t(G_{\lambda_{k+1}}(x^*)-G_0(x^*))$$

$$\leq \|G_{\lambda_{k+1}}(x_k) - G_{\lambda_{k+1}}(x^*)\|^2,$$

with the inequality obtained by applying condition **C3** with $x = G_{\lambda_{k+1}}(x_k)$.

Conditions **C2** and **C4** and the fact that $x^* = F(G_{\lambda_{k+1}}(x^*))$ imply that

$$\|x_{k+1} - x^*\|^2 = \|F(G_{\lambda_{k+1}}(x_k)) - x^*\|^2 \leq \|G_{\lambda_{k+1}}(x_k) - G_{\lambda_{k+1}}(x^*)\|^2$$

$$\leq \|x_k - x^*\|^2 - \gamma(\lambda_{k+1})\|x_k - G_0(x_k)\|^2 \tag{16}$$

Relation (16) implies that the sequence $\{\|x_k - x^*\|^2\}$ is nonincreasing, and hence converges for any fixed point $x^* \in \mathrm{FP}(T)$. Moreover, it implies that the sequence $\{x_k\}$ is bounded and, therefore, has at least one limit point.

We claim that for some limit point $\bar{x}$ of the sequence $\{x_k\}$, $\|\bar{x} - G_0(\bar{x})\|^2 = 0$. Suppose this is not the case. Then for some sufficiently large K and some constant $B > 0$, $\|x_k - G_0(x_k)\|^2 \geq B$ for all $k \geq K$. Since the sequence $\{\lambda_k\}$ satisfies condition **C5**,

$$0 \leq \lim_{k\to\infty} \|x_k - x^*\|^2 \leq \|x_K - x^*\|^2 - B\sum_{i=K}^{+\infty}\gamma(\lambda_{i+1}) = -\infty$$

which is a contradiction.

We conclude that the sequence $\{x_k\}$ has a limit point $\bar{x}$ satisfying the conditions $\|\bar{x} - G_0(\bar{x})\|^2 = 0$, implying that $\bar{x} \in \mathrm{FP}(G_0) = \mathrm{FP}(T)$. Therefore, the sequence $\{\|x_k - \bar{x}\|^2\}$ is convergent. Since $\bar{x}$ is at limit point of $\{x_k\}$, we conclude that $\lim_{k\to\infty}\|x_k - \bar{x}\|^2 = 0$. ∎

Future research related to this work might fruitfully examine a number of issues, for example:

1. Test these methods computationally to determine how they perform in practice.
2. Examine how this approach applies to the solution of variational inequality, complementarity, linear and nonlinear optimization problems.
3. By making necessary refinements, improve the convergence behavior of averaging schemes. Moreover, develop results by imposing weaker assumptions on the original problem map T.
4. Develop results on the rate of convergence of averaging schemes, perhaps by considering additional refinements on the schemes.

Acknowledgments: We are grateful to the Associate Editor and the referees for their constructive comments on the previous version of this paper that has led us to weaken some of our original assumptions.

References

[1] H.H. Bauschke. The approximation of fixed points of compositions of nonexpansive mappings in Hilbert space. *Preprint.* Center for Experimental and Constructive Mathematics, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6, 1994.

[2] F.E. Browder. Fixed-point theorems for noncompact mappings in Hilbert space. *Proc. Nat. Acad. Sci. U.S.A.*, 53: 1272-1276, 1965.

[3] F.E. Browder. Convergence of approximants to fixed points of nonexpansive nonlinear mappings in Banach spaces. *Archive for Rational Mechanics and Analysis*, 24:82–90, 1967.

[4] F. E. Browder and W.V. Petryshn. Construction of fixed points of nonlinear mappings in a Hilbert space. *Journal of Analysis and Applications*, 20:197-228.

[5] J.C. Dunn. On the recursive averaging processes and Hilbert space extensions of the contraction mapping principle. *Journal of the Franklin Institute*, 295:117-133, 1973.

[6] J. Eckstein and D.P. Bertsekas. On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55:293–318, 1992.

[7] A. Genel and J. Lindenstrauss. An example concerning fixed points. *Israel Journal of Mathematics*, 22:81–86, 1975.

[8] B. Halpern. Fixed points of nonexpanding maps. *Bulletin of the American Mathematical Society*, 73:957–961, 1967.

[9] S. Kaniel. Construction of a fixed point for contractions in Banach space. *Israel Journal of Mathematics*, 9:535–540, 1971.

[10] T. L. Magnanti and G. Perakis. Averaging schemes for variational inequalities and systems of equations. *Mathematics of Operations Research*, 22(3):568-587, 1997.

[11] T. L. Magnanti and G. Perakis. Solving variational inequalities and fixed point problems by averaging and optimizing potentials. *preprint*, 1997.

[12] T. L. Magnanti and G. Perakis. Computing fixed points by averaging. *Transportation and network analysis: Current trends* , Miscellenea in honor of Michael Florian, Eds. M. Gendreau and P. Marcotte, 2001.

[13] S. Reich. Weak convergence theorems for nonexpansive mappings in Banach spaces. *Journal of Mathematical Analysis and Applications*, 67:274–276, 1979.

[14] R.T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14(5):877–898, 1976.

[15] R. Wittmann. Approximation of fixed points of nonexpansive mappings. *Archiv der Mathematik*, 58:486–491, 1992.