# Using Randomization to Break the Curse of Dimensionality

John Rust
*Yale University*
March, 2000

**Goal:** Solve *Bellman equation* for *value function* $V$:

$$V = \Gamma(V)$$

where $\Gamma : B \to B$ is a *contraction mapping* defined by

$$\Gamma(V)(s) = \max_{a \in A} \left[ u(s,a) + \beta \int V(s')p(s'|s,a)\lambda(ds') \right]$$

where $\beta \in (0,1)$, $\lambda$ is Lebesgue measure, and $B$ is the Banach space of bounded, continuous functions on the $d$-dimensional hypercube under supremum norm, $\|f\| = sup_{s \in [0,1]^d} |f(s)|$.

**Problem:** Bellman's *curse of dimensionality:* the amount of time required to compute an $\epsilon$-approximation to fixed point $V$ using *deterministic algorithms* increases exponentially in $d$:

$$comp^{\textit{wor-det}}(\epsilon, \beta, d) = \Theta\left( \frac{1}{((1-\beta)^2\epsilon)^{2d}} \right)$$

**Idea:** Use randomization to break the curse of dimensionality. Consider successive approximations or *multigrid* algorithms to compute fixed point, but replace $\Gamma$ by *random Bellman operator* $\tilde{\Gamma}_N : B \to B$ defined by:

$$\tilde{\Gamma}_N(V)(s) = \max_{a \in A} \left[ u(s,a) + \frac{\beta}{N} \sum_{i=1}^{N} V(\tilde{s}_i)p(\tilde{s}_i|s,a) \right]$$

where $\{\tilde{s}_1, \ldots, \tilde{s}_N\}$ are *random grid points: $N$ IID* uniform draws from $[0,1]^d$.

**Outline of talk:**

1. Introduction/Motivation
2. Review of Markovian Decision Processes (MDP's)
3. Review of Information-Based Complexity Theory
4. Complexity of Continuous Decision Processes (CDP's): Randomization Can't Break the Curse of Dimensionality
5. Complexity of Discrete Decision Processes (DDP's): Randomization Can Break the Curse of Dimensionality
6. Concluding Comments: Are algorithms based on psuedo-random generators deterministic or random?

**Definition 2.1:** A *Markovian Decision Process* consists of the following objects:

- A *time index* $t \in \{0, 1, 2, \ldots, T\},\ T \leq \infty$
- A *state space* $S$,
- An *action space* $A$,
- A family of *constraint sets* $s \to A(s) \subseteq A$,
- A *utility function* $u(s, a)$,
- A Markov *transition probability* $p(ds'|s, a)$,
- A *discount factor* $\beta \in (0, 1)$.

**Definition 2.2:** A *Discrete Decision Process* is an MDP with the following property:

- There is a finite set $A$ such that $A(s) \subset A$ for each $s \in S$.

**Finite Horizon Dynamic Programming:** Terminal Period decision rule and value function are given by:

$$\alpha_T(s_T) = \underset{a_T \in A(s_T)}{argmax}\, u(s_T, a_T), \tag{2.2}$$

$$V_T(s_T) = \max_{a_T \in A(s_T)} u(s_T, a_T). \tag{2.3}$$

Backward induction (successive approximation) steps for periods $t = T-1, \ldots, 0$: yields:

$$\alpha_t(s_t) = \underset{a_t \in A(s_t)}{argmax} \left[ u(s_t, a_t) + \beta \int V_{t+1}(s_{t+1}) p(ds_{t+1}|s_t, a_t) \right], \tag{2.4}$$

$$V_t(s_t) = \max_{a_t \in A(s_t)} \left[ u(s_t, a_t) + \beta \int V_{t+1}(s_{t+1}) p(ds_{t+1}|s_t, a_t) \right]. \tag{2.5}$$

If we define the *Bellman operator* $\Gamma$ by:

$$\Gamma(W)(s) \equiv \max_{d \in A(s)} \left[ u(s, a) + \beta \int W(s') p(ds'|s, a) \right], \tag{2.6}$$

then the backward induction steps can be written more compactly as:

$$V_{T-t} = \Gamma^t(V_T), \quad t = 0, \ldots, T. \tag{2.7}$$

**Infinite horizon case:** Optimal stationary decision rule $\alpha$ is given by:

$$\alpha(s) = \underset{d \in A(s)}{argmax} \left[ u(s, a) + \beta \int V(s') p(ds'|s, a) \right], \tag{2.8}$$

where $V$ is the solution to *Bellman's equation:*

$$V(s) = \max_{d \in A(s)} \left[ u(s, a) + \beta \int V(s') p(ds'|s, a) \right]. \tag{2.9}$$

Bellman's equation can be re-written more compactly in operator notation as:

$$V = \Gamma(V). \tag{2.10}$$

The method of *successive approximations* is essentially equivalent to using the solution to a long, but finite horizon MDP problem to approximate the solution to the infinite horizon problem.

**Lemma 2.1:** *Suppose* $\{\Gamma_N\}$ *are a family of contraction mappings on a Banach space* $B$ *indexed by* $N$ *that that converge pointwise: i.e.* $\forall W \in B$ *we have*

$$\underset{N\to\infty}{lim} \Gamma_N(W) = \Gamma(W), \tag{2.11}$$

*and the approximate fixed point* $V_N = \Gamma_N(V_N)$ *satisfies the error bound:*

$$\|V_N - V\| \leq \frac{\|\Gamma_N(V) - \Gamma(V)\|}{(1-\beta)}. \tag{2.12}$$

**Lemma 2.2:** *Let* $\Gamma$ *be a contraction mapping on a Banach space* $B$ *with fixed point* $V = \Gamma(V)$. *If* $W$ *is an arbitrary element of* $B$ *the following inequalities hold:*

$$\|W - \Gamma(W)\| \leq (1+\beta)\|V - W\| \tag{2.14}$$

$$\|\Gamma^t(W) - V\| \leq \beta^t \|\Gamma(W) - W\|/(1-\beta) \tag{2.15}$$

$$\|\Gamma^{t+1}(W) - V\| \leq \beta\|\Gamma^t(W) - V\|. \tag{2.16}$$

**Lemma 2.3:** *Let* $\Gamma$ *be a Bellman operator. Then we have:*

$$\Gamma : B\left(0, \frac{K}{(1-\beta)}\right) \rightarrow B\left(0, \frac{K}{(1-\beta)}\right), \tag{2.17}$$

*where* $B(0,r) = \{V \in B | \ \|V\| \leq r\}$ *and the constant* $K$ *is given by:*

$$K \equiv \underset{s\in S}{sup} \ \underset{d\in A(s)}{sup} |u(s,a)|. \tag{2.18}$$

**Lemma 2.4:** *Suppose there exists an integer* $\overline{N}(\epsilon, \beta)$ *such that for all* $N \geq \overline{N}(\epsilon, \beta)$ *we have:*

$$\left\| \hat{\Gamma}_N(W) - \Gamma(W) \right\| \leq (1 - \beta)\epsilon, \tag{4.4}$$

*uniformly for all* $W \in B(0, K/(1-\beta))$ *where the constant* $K$ *is given by:*

$$K \equiv \sup_{s \in S} \sup_{d \in A(s)} |u(s, a)|. \tag{4.5}$$

*If we begin the backward induction using any estimate of the terminal value function* $\hat{V}_T$ *satisfying:*

$$\begin{aligned} &\|\hat{V}_T - V_T\| \leq \epsilon, \\ &\|\hat{V}_T\| \leq K/(1 - \beta), \end{aligned} \tag{4.6}$$

*then* $\hat{V}_t$ *will be uniformly within* $\epsilon$ *of* $V_t$ *for all* $t$*:*

$$\max_{t \in \{1,\ldots,T\}} \|\hat{V}_t - V_t\| \leq \epsilon. \tag{4.7}$$

**Information-Based Complexity Theory**

Traub, Wasilikowski, and Woźniakowski, 1988.

1. Define a mathematical problem via a *solution operator* $\Lambda : F \rightarrow B$. For example, in multivariate integration $\Lambda$ is given by:
$$\Lambda(f) = \int_{[0,1]^d} f(s)\lambda(ds), \tag{3.1}$$

2. Approximate $\Lambda$ using a "computable" mapping $U : F \rightarrow B$ represented by:
$$U(f) = \phi_N\left(I_N(f)\right), \tag{3.4}$$
where $I_N(f) : F \rightarrow R^N$ represents an *information operator* (*oracle*) and $\phi_N : R^N \rightarrow B$ is an *algorithm* that maps the information about $f$ into an approximate solution. We consider algorithms using *standard information:*
$$I_N(f) = (f(s_1), \ldots, f(s_N)), \tag{3.5}$$
where $\{s_1, \ldots, s_N\}$ can be thought of as defining "grid points" in the domain of the function $f$. An example of an algorithm in the case of multivariate integration is the simple sample average operator:
$$\phi_N\left(I_N(f)\right) = \frac{1}{N}\sum_{i=1}^{N} f(s_i). \tag{3.7}$$

3. The information operator $I_N$, the algorithm $\phi_N$, and the number of grid points $N$ are all treated as variables to be chosen in order to obtain approximations of a specified degree of accuracy. Thus, we can think of an arbitrary algorithm as consisting of the pair $U_N = (I_N, \phi_N)$. We say that $U_N(f)$ is an *$\epsilon$-approximation* of $f \in F$ if:
$$\|\Lambda(f) - U_N(f)\| \leq \epsilon$$

4. The *cost* of computing an approximate solution $U_N(f)$ is given by:
$$c(U_N, f) = c_1(I_N, f) + c_2(\phi_N, I_N(f)), \tag{3.6}$$
where $c_1(I_N, f)$ is the cost of computing the $N$ information points $I_N(f)$ (i.e. of evaluating the function $f$ at the $N$ grid points $\{s_1, \ldots, s_N\}$), and $c_2(\phi_N, I_N(f))$ denotes the combinatorial cost of using the information $I_N(f) \in R^N$ to compute the approximate solution $U_N(f) = \phi_N(I_N(f))$. In many problems such as the multivariate integration problem it is reasonable to presume that $c(I_N, f)$ and $c(\phi_N, I_N(f))$ are both proportional to $N$.

**Definition 3.1:** The *worst case deterministic complexity* of a mathematical problem $\Lambda$ is given by:

$$comp^{\textit{wor-det}}(\epsilon) = \underset{U_N}{inf} \left\{ c(U_N) \mid e(U_N) \leq \epsilon \right\}, \tag{3.8}$$

*where the functions* $e(U_N)$ *and* $c(U_N)$ *are defined by:*

$$\begin{aligned} e(U_N) &= \underset{f \in F}{sup} \left\| \Lambda(f) - U_N(f) \right\|, \\ c(U_N) &= \underset{f \in F}{sup}\, c(U_N, f). \end{aligned} \tag{3.9}$$

For example the worst case complexity of integration of functions that are differentiable of order $r$ is given by:

$$comp^{\textit{wor-det}}(\epsilon, d, r) = \Theta\left(\frac{1}{\epsilon^{d/r}}\right). \tag{3.10}$$

Thus, multivariate integration is subject to the *curse of dimensionality:* i.e. complexity increases exponentially in the problem dimension $d$.

**Complexity of random algorithms**

A random algorithm $\tilde{U}_N$ can be decomposed as a composition of a random information operator $\tilde{I}_N$ and a random combinatory algorithm $\tilde{\phi}_N$:

$$\tilde{U}_N(f) = \tilde{\phi}_N\left(\tilde{I}_N(f)\right). \tag{3.11}$$

Monte carlo integration is a simple example of a random algorithm $\tilde{U}_N$. It is based on the random information

$$\tilde{I}_N(f) = \{f(\tilde{s}_1), \ldots, f(\tilde{s}_N)\}, \tag{3.12}$$

where $\{\tilde{s}_1, \ldots, \tilde{s}_N\}$ are *IID* uniform random draws from $[0,1]^d$, and the deterministic linear algorithm $\phi_N : R^N \to R$ given in (3.7).

In order to ensure that a random algorithm $\tilde{U}_N$ is well defined we need to define an underlying probability space $(\Omega, \Im, \mu)$ and measurable mappings $\tilde{I}_N : \Omega \to R^N$ and $\tilde{\phi}_N : R^N \times \Omega \to B$ such that $\tilde{U}_N(f)$ is a well defined random element of $B$ for each $f \in F$. We can think of a random algorithm as a kind of *mixed strategy* $\tilde{U}_N = (\tilde{\phi}_N, \tilde{I}_N, \Omega, \Im, \mu)$.

**Definition 3.2:** The *worst case randomized complexity* of a mathematical problem $\Lambda$ is given by:

$$comp^{\textit{wor-ran}}(\epsilon) \equiv \inf_{\tilde{U}_N} \left\{ c(\tilde{U}_N) \mid e(\tilde{U}_N) \leq \epsilon \right\}, \tag{3.13}$$

*where $e(\tilde{U}_N)$ and $c(\tilde{U}_N)$ are defined by:*

$$\begin{aligned} e(\tilde{U}_N) &= \sup_{f \in F} \int_\Omega \left\| \Lambda(f) - \tilde{U}_N(f) \right\| \mu(d\omega), \\ c(\tilde{U}_N) &= \sup_{f \in F} \int_\Omega c\left(\tilde{U}_N(\omega, f), f\right) \mu(d\omega). \end{aligned} \tag{3.14}$$

**Example:** The simple monte carlo integration algorithm puts an upper bound on the complexity of the problem given by:

$$comp^{wor\text{-}ran}(\epsilon, d) = O\left(\frac{1}{\epsilon^2}\right). \tag{3.15}$$

To see this, note that Hölder's inequality implies:

$$e(\tilde{U}_N) = \sup_{f \in F} E\left\{\left|\frac{1}{N}\sum_{i=1}^{N} f(\tilde{s}_i) - \int f(s)\lambda(ds)\right|\right\} \leq \frac{\sup_{f \in F}\sigma(f)}{\sqrt{N}}. \tag{3.18}$$

Bakhvalov 1959 and Novak 1987 have established lower bounds on the randomized worst case complexity of multivariate integration.

**Theorem 3.1** *The worst case randomized complexity of the multivariate integration problem is given by:*

$$comp^{wor\text{-}ran}(\epsilon, d, r) = \Theta\left(\frac{1}{\epsilon^m}\right), \tag{3.20}$$

*where $m$ is given by:*

$$m = \frac{2}{\left(\frac{2r}{d}\right) + 1}. \tag{3.21}$$

It follows that in high dimensional problems simple monte carlo integration is an approximately *optimal algorithm.*

## Worst Case Complexity of CDP's

General solution strategy: discretize state space into a finite number of points $|S|$ and actions $|A|$. Backward recursion becomes:

$$V_t(s_j) = \max_{a \in A(s_j)} \left[ u(s_j, a) + \beta \sum_{i=1}^{|S|} V_{t+1}(s_i) p(s_i | s_j, a) \right], \ j = 1, \ldots, |S| \tag{4.1}$$

Notice amount of computer time/space required per backward induction step is $O(|A||S|^2)$.

**Definition 4.1** Let $\Lambda : F \to B$ be the solution operator of the general MDP problem of Definition 2.1. Then we have $V = \Lambda(u, p)$, where $V$ is the value function defined by recursions $(2.2), \ldots, (2.5)$ in the finite horizon case or recursion $(2.10)$ in the infinite horizon case, $B = C[0,1]^{d_s}$, and $F = \{(u, p)\}$ is the set of problem elements satisfying regularity conditions $(A1), \ldots, (A5)$ below:

**(A1)** *$S$ is a Borel subset of $[0,1]^{d_s}$; $A(s)$ is a compact subset of $[0,1]^{d_a}$ $\forall s \in S$.*

**(A2)** *The transition probability is absolutely continuous with respect to Lebesgue measure $\lambda(ds)$ on $S$, with density $p(s'|s, a)$ existing for each $a \in A(s)$ and $s \in S$.*

**(A3)** *The utility function $u(s, a)$ is Lipschitz continuous Lipschitz bound $K_u$.*

**(A4)** *The transition density $p(s'|s, a)$ is jointly Lipschitz continuous in $(s', s, a)$ with Lipschitz bound $K_p$.*

**(A5)** *The mapping $s \to A(s)$ is a Lipschitz continuous correspondence.*

**Chow-Tsitsiklis 1991 Multigrid Algorithm**

Discretize state space $S = [0,1]^{d_s}$ into $N_s = (1/h)^{d_s}$ subcubes and the action space $A = [0,1]^{d_a}$ into $N_a = (1/h)^{d_a}$ subcubes. Let $k(s)$ denote subcube that contains state $s \in S$. Then define discretized utility and transition probabilities by:

$$\begin{aligned} u_h(s,a) &= u(s_{k(s)}, a) \\ p_h(s'|s,a) &= \frac{p(s_{k(s')}|s_{k(s)}, a)}{\int p(s_{k(s')}|s_{k(s)}, a)\lambda(ds')} \end{aligned} \tag{4.10}$$

and corresponding discretized Bellman operator by:

$$\begin{aligned} \hat{\Gamma}_h(V)(s) &= \max_{a \in A_h(s)} \left[ u_h(s,a) + \beta \int V(s') p_h(s'|s,a) \lambda(ds') \right] \\ &= \max_{a \in A_h(s)} \left[ u_h(s,a) + \frac{\beta}{N} \sum_{k=1}^{N} V(s_k) p_h(s_k|s_{k(s)}, a) \right]. \end{aligned} \tag{4.11}$$

**Theorem 4.2:** *There exist constants $K_1$ and $K_2$ such that for all $h$ sufficiently small and all $V \in B$ we have:*

$$\|\Gamma(V) - \hat{\Gamma}_h(V)\| \leq (K_1 + \beta K_2 \|V\|) h. \tag{4.12}$$

To guarantee that $\|\hat{V}_t - V_t\| \leq \epsilon$ at each $t$ we must choose $h$ sufficiently small so that $\|\Gamma(V) - \hat{\Gamma}_h(V)\| \leq (1-\beta)\epsilon$ uniformly for all $V \leq K/(1-\beta)$. Theorem 4.2 shows that we can guarantee this by setting the grid size $h \leq \overline{h}$, where $\overline{h}$ is given by:

$$\overline{h} = \frac{(1-\beta)^2 \epsilon}{K_1(1-\beta) + \beta K_2 K}. \tag{4.13}$$

The resulting discretized MDP has $|S| = O(1/[(1-\beta)^2\epsilon)]^{d_s})$ states and $|A| = O(1/[(1-\beta)\epsilon]^{d_a})$ actions. Recalling that $O(T|A||S|^2)$ operations are required to solve the discrete finite horizon MDP, it follows that an upper bound on complexity is given by the following result of Chow and Tsitsiklis 1989:

**Theorem 4.3:** *The worst case deterministic complexity of the infinite horizon MDP problem* $\Lambda : F \to B$ *is given by:*

$$comp^{\textit{wor-det}}(\epsilon, d_s, d_a, \beta) = \Theta\left(\frac{T}{((1-\beta)^2\epsilon)^{2d_s+d_a}}\right). \tag{4.14}$$

*where* $T$ *is a constant independent of* $\beta$ *and* $\epsilon$*.*

**Are there algorithms that attain this lower bound?**

Let $T(\epsilon, \beta)$ be the number of iterations to get within $\epsilon$ of the fixed point:

- **successive approximations**

$$T(\beta, \epsilon) = \Theta\left(\frac{log\,(1/(1-\beta)\epsilon)}{|\,log(\beta)|}\right). \tag{4.15}$$

- **multigrid algorithm** (to be described)

$$T(\beta, \epsilon) = \Theta\left(\frac{1}{|\,log(\beta)|}\right), \tag{4.16}$$

## Worst Case Complexity of DDP's

**Observation:** In DDP's there are only a finite number of possible actions in each state. Essentially all of the work involved in computing Bellman operator is the evaluation of the conditional expectation $E\{V|s, a\}$ for $s \in S$, $a \in A$. Since monte carlo integration breaks curse of dimensionality for integration, it seems reasonable to conjecture that it breaks the curse of dimensionality for the DDP problem as well.

**Problem:** Evaluation of Bellman operator requires computation of infinitely many integrals, one for each possible state $s \in S$. In addition, Bellman operator is nonlinear due to $max$ operator: even though monte carlo estimate of $E\{V|s, a\}$ may be unbiased for each $(s, a)$ pair, random Bellman operator $\tilde{\Gamma}_N(V)(s)$ will generally be upward biased due to Jensen's inequality.

**Intuition:** If transition probability $p(s'|s, a)$ satisfies a Lipschitz condition, there is an implied "rigidity" in the function $E\{V|s, a\}$ as $s$ varies. Thus, if monte carlo integral is close to $E\{V|s, a\}$ for any particular $s \in S$ it will also be close uniformly for $s \in S$. For sufficiently large $N$, problem of upward bias in $\tilde{\Gamma}_N(V)$ disappears since it converges uniformly to $\Gamma(V)$ as $N \to \infty$.

**Technical problems:** Need *Banach Space Central Limit Theorem* to show that $\|\tilde{\Gamma}_N(V) - \Gamma(V)\| = O_P(1/\sqrt{N})$. Must also show that this convergence is uniform for all $V$ in a ball of finite radius, and is uniform for all $(u, p)$ satisfying certain regularity conditions in order to show that randomization breaks curse of dimensionality on a worst case basis. Finally, must show there is a finite bounding constant $K(d)$ such that *for all* $N \geq 1$ we have:

$$\underset{V}{sup}\,\underset{(u,p)}{sup}\, E\left\{\|\tilde{\Gamma}_N(V) - \Gamma(V)\|\right\} \leq \frac{K(d)}{\sqrt{N}}$$

where $K(d)$ grows at most polynomially with dimension $d$. Showing this requires compactness arguments and *maximal inequalities for Gaussian processes.*

**Definition 5.1** A *random Bellman operator* $\tilde{\Gamma}_N : B \to B$ (where $B = C[0,1]^d$) is defined by:

$$\tilde{\Gamma}_N(V)(s) = \underset{a \in A}{max}\left[u(s,a) + \frac{\beta}{N}\sum_{k=1}^{N} V(\tilde{s}_k)p(\tilde{s}_k|s,a)\right], \tag{5.1}$$

*where* $\{\tilde{s}_1, \ldots, \tilde{s}_N\}$ *are IID draws with respect to Lebesgue measure* $\lambda$ *from the unit hypercube* $[0,1]^d$.

The proof of the convergence of $\tilde{\Gamma}_N(V)$ to $\Gamma(V)$ is based on the linear operators $\tilde{\Gamma}_{a,N}$ and $\Gamma_a$ defined by:

$$\begin{aligned}\tilde{\Gamma}_{a,N}(V)(s) &= u(s,a) + \frac{\beta}{N}\sum_{k=1}^{N} V(\tilde{s}_k)p(\tilde{s}_k|s,a)\\ \Gamma_a(V)(s) &= u(s,a) + \beta\int V(s')p(s'|s,a)\lambda(ds').\end{aligned} \tag{5.2}$$

Note that we have

$$\begin{aligned}\tilde{\Gamma}_N(V) &= \underset{a \in A}{max}\,\tilde{\Gamma}_{a,N}(V)\\ \Gamma(V) &= \underset{a \in A}{max}\,\Gamma_a(V),\end{aligned} \tag{5.3}$$

where the maximization on the right hand side of (5.3) is performed pointwise for each $s \in S$.

Note $\tilde{\Gamma}_N$ is not guaranteed to be a contraction with probability 1 for finite $N$. However following normalized operator $\hat{\Gamma}_N$ is a contraction with probability 1 for all $N \geq 1$:

$$\hat{\Gamma}_N(V)(s) = \max_{a \in A} \left[ u(s,a) + \beta \sum_{k=1}^{N} V(s_k) p_N(s_k|s,a) \right], \tag{5.4}$$

where $p_N$ is defined by:

$$p_N(s_k|s,a) = \frac{p(s_k|s,a)}{\sum_{i=1}^{N} p(s_i|s,a)}. \tag{5.5}$$

**Strategy:** Establish asymptotic properties of linear operators $\tilde{\Gamma}_{a,N}$ first since they are simple sample averages of *IID* random elements of $B$. Then show that asymptotic properties of the normalized operators $\hat{\Gamma}_{a,N}$ and $\hat{\Gamma}_N$ are first order equivalent to $\tilde{\Gamma}_{a,N}$ and $\tilde{\Gamma}_N$.

**Lemma 5.1** $\forall N \geq 1$ we have:

$$\begin{aligned} \|\tilde{\Gamma}_N(V) - \Gamma(V)\| &= \| \max_{a \in A} \tilde{\Gamma}_{a,N}(V) - \max_{a \in A} \Gamma_a(V) \| \\ &\leq \max_{a \in A} \|\tilde{\Gamma}_{a,N}(V) - \Gamma_a(V)\| \\ &\leq \sum_{a \in A} \|\hat{\Gamma}_{a,N}(V) - \Gamma_a(V)\|. \end{aligned} \tag{5.6}$$

**Definition 5.2** Let $\Lambda : F \to B$ be the solution operator corresponding to the DDP problem of Definition 2.2. Then we have $V = \Lambda(u,p)$, where $V$ is the value function defined by recursions $(2.2), \ldots, (2.5)$ in the finite horizon case or recursion $(2.10)$ in the infinite horizon case, $B = C[0,1]^d$, and $F = \{(u,p)\}$ is the set of problem elements satisfying regularity conditions $(A1), \ldots, (A4)$ below:

**(A1)** $S = [0,1]^d$.

**(A2)** $\exists K_u < \infty, \forall a \in A, \forall s, s' \in S \; |u(s,a) - u(s',a)| \le K_u |s - s'|$.

**(A3)** The transition probability has a density $p(s'|s,a)$ with respect to Lebesgue measure $\lambda$ on $[0,1]^d$.

**(A4)** $\exists K_p < \infty, \forall a \in A, \forall (s',s), (t',t) \in S \times S \; |p(s'|s,a) - p(t'|t,a)| \le K_p |(s',s) - (t',t)|$.

**Definition:** *Let $BL(K) \subset B$ denote the set of uniformly Lipschitz functions with Lipschitz bound $K$, i.e.*

$$BL(K) = \{ f \in B | \; |f(s) - f(s')| \le K|s - s'|, \; \|f\| \le K \} . \tag{5.15}$$

**Lemma 5.2** *$\forall K > 0$, $\forall N \ge 1$ the operators $\tilde{\Gamma}_N$, $\tilde{\Gamma}_{a,N}$, $\hat{\Gamma}_N$, $\hat{\Gamma}_{a,N}$, $\Gamma$, and $\Gamma_a$ map $B(0,K)$ into $BL(K_u + \beta K K_p)$, where $B(0,K)$ is the ball of radius $K$ in $B$.*

**Lemma 5.3** *For each $a \in A$ and $V \in B$ the following limits hold with probability 1:*

$$\begin{aligned}
\underset{N\to\infty}{lim} \tilde{\Gamma}_N(V) &= \Gamma(V) \\
\underset{N\to\infty}{lim} \tilde{\Gamma}_{a,N}(V) &= \Gamma_a(V) \\
\underset{N\to\infty}{lim} \hat{\Gamma}_N(V) &= \Gamma(V) \\
\underset{N\to\infty}{lim} \hat{\Gamma}_{a,N}(V) &= \Gamma_a(V).
\end{aligned} \tag{5.17}$$

**Theorem 5.1:** *The random operators* $(\tilde{\Gamma}_N, \hat{\Gamma}_N, \tilde{\Gamma}_{a,N}, \hat{\Gamma}_{a,N})$ *are uniformly consistent: i.e.* $\forall V \in BL(K)$ *we have with probability 1:*

$$\begin{aligned} &\lim_{N\to\infty} \sup_{V\in BL(K)} |\Gamma(V) - \tilde{\Gamma}_N(V)| = 0, \\ &\lim_{N\to\infty} \sup_{V\in BL(K)} \|\Gamma(V) - \hat{\Gamma}_N(V)\| = 0 \\ &\lim_{N\to\infty} \sup_{V\in BL(K)} \|\Gamma_a(V) - \tilde{\Gamma}_{a,N}(V)\| = 0 \\ &\lim_{N\to\infty} \sup_{V\in BL(K)} \|\Gamma_a(V) - \hat{\Gamma}_{a,N}(V)\| = 0. \end{aligned} \tag{5.20}$$

**Corollary** *The normalization factors converge uniformly to 1 as* $N \to \infty$ *with probability 1:*

$$\lim_{N\to\infty} \max_{a\in A} \sup_{s\in[0,1]^d} \left| 1 - \frac{1}{N} \sum_{i=1}^{N} p(s_i|s,a) \right| = 0. \tag{5.23}$$

**Theorem 5.2** $\forall a \in A$, *and* $\forall V \in B = C[0,1]^d$ *we have:*

$$\sqrt{N}[\tilde{\Gamma}_{a,N}(V) - \Gamma_a(V)] \Longrightarrow \tilde{Z}_a(V) \sim N(0, \Sigma_a(V)), \tag{5.24}$$

*where* $N(0,\Sigma_a(V))$ *denotes a Gaussian random element with mean* $0 \in B$ *and covariance operator* $\Sigma_a(V) \in L(B', B')$, *where* $B'$ *is the dual space to* $B$ *and* $L(B', B')$ *is the Banach space of bounded bilinear operators on* $B'$. *In this case,* $B'$ *is congruent to the space* $P(S)$ *of finite, regular signed Borel measures on S, and given* $\mu_1, \mu_2 \in P(S)$ *the covariance operator* $\Sigma_a(V)$ *has the representation:*

$$\begin{aligned} \mu_1 \Sigma_a(V) \mu_2 = \beta^2 \Big[ &\int_{s'} V^2(s') \int_s p(s'|s,a)\mu_1(ds) \times \\ &\int_s p(s'|s,a)\mu_2(ds)\lambda(ds') - \\ &\int_{s'} \int_s V(s')p(s'|s,a)\mu_1(ds)\lambda(ds') \times \\ &\int_{s'} \int_s V(s')p(s'|s,a)\mu_2(ds)\lambda(ds') \Big]. \end{aligned} \tag{5.25}$$

**Theorem 5.3** *For each* $V \in B$ *we have:* $\sqrt{N}\|\tilde{\Gamma}_N(V) - \Gamma(V)\| = O_p(1)$.

**Theorem 5.4** *Let* $\tilde{Z}_a(V)$ *denote the Gaussian random element in Theorem 5.2. Then* $\tilde{Z}_a(V)$ *can be represented as a stochastic integral with respect to the Brownian Bridge process* $\tilde{B}$ *on* $[0,1]^d$*:*

$$\tilde{Z}_a(V)(s) \simeq \beta \int V(s')p(s'|s,a)\tilde{B}(ds'). \tag{5.37}$$

**Proof** (sketch) Let $\lambda_N(s)$ denote the empirical CDF of the *IID* uniform realizations $\{s_1,\dots,s_N\}$,

$$\lambda_N(s) = \frac{1}{N}\sum_{i=1}^{N} I\{s_i \le s\}. \tag{5.38}$$

Then it is straightforward to verify that the following representation obtains:

$$\sqrt{N}\left[\tilde{\Gamma}_{a,N}(V) - \Gamma_a(V)\right](s) =$$
$$\int V(s')p(s'|s,a)\sqrt{N}\left[\lambda_N(ds') - \lambda(ds')\right],$$

where $\sqrt{N}[\lambda_N - \lambda]$ is an *empirical process* indexed by $s \in [0,1]^d$. Technically the process is indexed by a class $\Delta = \{D_s | s \in [0,1]^d\}$ of $d$-dimensional "intervals" of the form:

$$I_s = I\left\{s' \in [0,1]^d |\ s' \le s\right\}.$$

By the multivariate version of Donsker's Theorem (Gänßler, 1983) $\sqrt{N}[\lambda_N - \lambda]$ converges weakly to a $d$-dimensional Brownian Bridge process, which has the following representation in terms of a standard $d$-dimensional Wiener process:

$$\tilde{B}(t) = \tilde{W}(t) - t\tilde{W}(\mathbf{1}), \tag{5.41}$$

where $\mathbf{1}$ denotes a vector in $[0,1]^d$ consisting of all 1's. The stochastic integral with respect to the $d$-dimensional Brownian Bridge process can be defined in terms of the usual Ito integral with respect to the $d$-dimensional Wiener process as follows:

$$\begin{aligned} \int h(s)\tilde{B}(ds) &= \int h(s)\tilde{W}(ds) - \tilde{W}(\mathbf{1}) \int h(s)\lambda(ds), \\ &\sim N\left(0, h\Sigma h\right), \end{aligned} \tag{5.43}$$

where the covariance operator $\Sigma$ is given by:

$$f\Sigma g = \int f(s)g(s)\lambda(ds) - \int f(s)\lambda(ds) \int g(s)\lambda(ds). \tag{3.16}$$

Using equations (5.43) and (3.16) it is straightfoward to show that the Gaussian random element $\beta \int V(s')p(s'|s,a)\tilde{B}(ds')$ has the same covariance operator as the Gaussian random element in equation (5.25) of Theorem 5.2. Since the characteristic functional uniquely determines the distribution of a random element, and since the characteristic functional of a centered Gaussian random element is uniquely determined by its covariance operator, the result follows.

**Corollary** *The difference between the normalized Bellman operator $\hat{\Gamma}_N(V)$ and the non-normalized random Bellman operator $\tilde{\Gamma}_N(V)$ converges weakly to a Gaussian random element which has the following stochastic integral representation:*

$$\sqrt{N}\left(\hat{\Gamma}_{a,N}(V)(s) - \tilde{\Gamma}_{a,N}(V)(s)\right) \Longrightarrow \beta E\{V\}(s) \int p(s'|s,a)\tilde{B}(ds'). \tag{5.47}$$

*where $E\{V\}(s) = \int V(s')p(s'|s,a)\lambda(ds')$ and $\tilde{B}(ds)$ is a Brownian Bridge process on $[0,1]^d$.*

**Corollary** *For each $V \in B$ we have:*

$$\sqrt{N}\|\hat{\Gamma}_N(V) - \Gamma(V)\| = O_p(1) \tag{0.1}$$

**Using Maximal Inequalities to Bound Expected Error**

- To show that random Bellman operator succeeds in breaking the curse of dimensionality, we need an error bound in the expected error that holds for all $N \geq 1$, i.e. we need a bound of the form

$$\underset{p \in BL(K_p)}{sup} \; \underset{\|V\| \leq K_v}{sup} E\{\|\tilde{\Gamma}_N(V) - \Gamma(V)\|\} \leq \frac{\gamma(d)|A|K_p K_v}{\sqrt{N}}, \quad (5.54)$$

  where the constant $\gamma(d)$ grows at most polynomially in the dimension $d$.

- We do this using a *maximal inequality for empirical processes* due to Pollard, 1989. The idea behind Pollard's bound can be explained heuristically as follows. The random linear operators $\tilde{Z}_{a,N}(V) \equiv \sqrt{N}[\tilde{\Gamma}_{a,N}(V) - \Gamma_a(V)]$ are stochastic processes indexed by a parameter $t \in [0,1]^d$. These stochastic processes can be represented as stochastic integrals with respect to the empirical process $\tilde{B}_N$:

$$\tilde{Z}_{a,N}(V)(t) = \int V(s) p(s|t,a) \tilde{B}_N(ds). \quad (5.55)$$

  Theorem 5.4 establishes that the stochastic processes $\tilde{Z}_{a,N}(V)$ converge weakly to the Gaussian stochastic process $\tilde{Z}_a(V)$ defined by:

$$\tilde{Z}_a(V)(t) = \int V(s) p(s|t,a) \tilde{B}(ds), \quad (5.56)$$

  where $\tilde{B}$ is the Brownian bridge process on $[0,1]^d$. Using "chaining arguments" Pollard derived the following maximal inequality for Gaussian processes in terms of an integral involving the "$\epsilon$-capacity" $D(\epsilon, S, \rho)$.

**Theorem 5.5** *Let* $(S, \rho)$ *be a pseudometric space and* $\{\tilde{Z}(t)|t \in S\}$ *be a Gaussian process with* $\rho$*-continuous sample paths for which:*

$$E\left\{|\tilde{Z}(s) - \tilde{Z}(t)|^2\right\} \leq \rho(s,t)^2 \tag{5.57}$$

*for all* $s$ *and* $t$ *in* $S$. *Then there exists a universal constant* $C$ *such that for each* $s_0 \in S$ *we have:*

$$\begin{aligned} E\left\{\|\tilde{Z}\|\right\} &\equiv E\left\{\underset{s\in S}{sup}\,|\tilde{Z}(s)|\right\} \\ &\leq E\left\{|\tilde{Z}(s_0)|\right\} + C\int_0^\delta \sqrt{log\,(D(x,S,\rho))}dx, \end{aligned} \tag{5.58}$$

*where* $\delta = sup_{s\in S}\,\rho(s,s_0)$ *and* $D(\epsilon, S, \rho)$ *is the* $\epsilon$*-capacity of* $S$*:*

$$sup\,\{k|\,\exists s_1,\ldots,s_k \in S\,:\,\rho(s_i,s_j) > \epsilon \quad \text{for}\, i \neq j,\; i,j = 1,\ldots,k\}\,.$$

We find it slightly easier to work with the metric entropy:

$$H(\epsilon, S, \rho) \equiv inf\left\{k|\exists s_1,\ldots,s_k \in S\,:\,\underset{s\in S}{sup}\,\underset{i\leq k}{min}\,\rho(s_i,s) \leq \epsilon\right\}. \tag{5.60}$$

Dudley 1984 showed that $H$ and $D$ satisfy:

$$H(\epsilon,S,\rho) \leq D(\epsilon,S,\rho) \leq H(\epsilon/2,S,\rho). \tag{5.61}$$

Using (5.61) and the maximal inequality in Theorem 5.5 we can derive a uniform bound on $E\{\|\tilde{Z}_a(V)\|\}$.

**Theorem 5.6** *Let* $\tilde{Z}_a(V)$ *be the limiting Gaussian process derived in Theorems 5.2 and 5.4. Then* $E\{\|\tilde{Z}_a(V)\|\}$ *satisfies the following bound:*

$$E\{\|\tilde{Z}_a(V)\|\} \leq \beta\left[1 + d\sqrt{\pi}C\right]K_p\|V\|, \tag{5.62}$$

*where* $C$ *is the universal constant from equation (5.58) of Theorem 5.5.*

**Corollary:** *The bound in Theorem 5.6 holds uniformly for all $p \in BL(K_p)$ and all $\|V\| \in B(0, K_v)$:*

$$\sup_{p \in BL(K_p)} \sup_{V \in B(0,K_v)} E\{\|\tilde{Z}_a(V)\|\} \leq \beta \left[1 + d\sqrt{\pi}C\right] K_p K_v. \tag{5.70}$$

The next step is to show that a version of the inequality (5.70) not only holds in the limit as $N \to \infty$, but also for all $N \geq 1$. The "cost" of having the inequality apply for all $N \geq 1$ is a slight increase in the bounding constant — by a factor of $\sqrt{\pi/2}$. The proof is based on the following "symmetrization inequality" of Pollard 1989:

**Lemma 5.4** *For each $N \geq 1$ we have:*

$$\begin{aligned} E\{\sqrt{N}\|\tilde{\Gamma}_{a,N}(V) - \Gamma_a(V)\|\} &= E\left\{\sup_{s \in [0,1]^d} \left|\int V(t)p(t|s,a)\tilde{B}_N(dt)\right|\right\} \\ &\leq \sqrt{\frac{\pi}{2}} E\left\{\sup_{s \in [0,1]^d} |\tilde{Z}_{a,N}(s,\xi)|\right\}, \end{aligned} \tag{5.73}$$

*where $\xi$ denotes the realized IID sample of of points from $[0,1]^d$: i.e. $\xi \equiv (s_1, \ldots, s_N)$ and conditional on $\xi$ $\tilde{Z}_{a,N}(s,\xi)$ is a Gaussian process on $[0,1]^d$ defined by:*

$$\tilde{Z}_{a,N}(s,\xi) \equiv \frac{\beta}{\sqrt{N}} \sum_{i=1}^{N} \tilde{z}_i V(s_i) p(s_i|s,a), \tag{5.74}$$

*where the $\{\tilde{z}_i\}$ are IID $N(0,1)$ variates.*

Since $\tilde{Z}_{a,N}(t,\xi)$ a Gaussian process conditional on the realized value of $\xi$, we can apply the maximal inequality for Gaussian processes (5.58) to obtain the following inequality:

$$\begin{aligned} &E_\xi \left\{\sup_{s \in [0,1]} \left|\tilde{Z}_{a,N}(s,\xi)\right|\right\} \\ &\leq E_\xi \left\{\left|\tilde{Z}_{a,N}(s_0,\xi)\right|\right\} + C\int_0^\delta log\left(D(x,S,\rho)\right)^{1/2} dx \end{aligned} \tag{5.75}$$

where $D(\epsilon, S, \rho)$ denotes the $\epsilon$-capacity of $S$ with respect to the pseudometric $\rho$ defined in the proof of theorem 5.6. Taking expectations over $\tilde{\xi} = (\tilde{s}_1, \ldots, \tilde{s}_N)$ yields the following bound:

**Theorem 5.7** *Let $\tilde{Z}_{a,N}(V)$ be the sub-Gaussian process given by:*

$$\tilde{Z}_{a,N}(V) = \sqrt{N}\left[\tilde{\Gamma}_{a,N}(V) - \Gamma_a(V)\right], \tag{5.76}$$

*then for each $N \geq 1$ we have:*

$$E\left\{\sqrt{N}\|\tilde{\Gamma}_{a,N}(V) - \Gamma_a(V)\|\right\} \leq \sqrt{\frac{\pi}{2}}\beta\left[1 + d\sqrt{\pi}C\right]K_p\|V\|. \tag{5.77}$$

**Corollary** *The following uniform bound holds:*

$$\underset{p\in BL(K_p)}{sup}\ \underset{V\in B(0,K_v)}{sup}\ E\{\|\tilde{\Gamma}_{a,N}(V) - \Gamma_a(V)\|\} \leq \frac{\gamma(d)K_pK_v}{\sqrt{N}}, \tag{5.81}$$

*where the constant $\gamma(d)$ is given by:*

$$\gamma(d) = \sqrt{\frac{\pi}{2}}\beta\left[1 + d\sqrt{\pi}C\right]. \tag{5.82}$$

Using inequality (5.81) and inequality (5.6) from Lemma 5.1, we obtain the main result of this section:

**Theorem 5.8** *For each $N \geq 1$ the expected error in the random Bellman operator satisfies the uniform bound:*

$$\underset{p\in BL(K_p)}{sup}\ \underset{V\in B(0,K_v)}{sup}\ E\left\{\|\tilde{\Gamma}_N(V) - \Gamma(V)\|\right\} \leq \frac{\gamma(d)|A|K_pK_v}{\sqrt{N}}, \tag{5.83}$$

*where the bounding constant $\gamma(d)$ is given in equation (5.82).*

**Corollary** *The expected error in the normalized Bellman operator $\hat{\Gamma}_N$ satisfies:*

$$\underset{p\in BL(K_p)}{sup}\ \underset{V\in B(0,K_v)}{sup}\ E\left\{\|\hat{\Gamma}_N(V) - \Gamma(V)\|\right\} \leq 2\frac{\gamma(d)|A|K_pK_v}{\sqrt{N}}, \tag{0.2}$$

*where the bounding constant $\gamma(d)$ is given in equation (5.82).*

**Random Successive Approximations Algorithm**

$$\hat{V}_T(s_i) = \underset{a \in A}{argmax}\, u(s_i, a) \quad i = 1, \ldots, N. \tag{5.89}$$

Subsequent value functions $V_{T-t}$, $t = 0, \ldots, T$ are generated by successive application of the $\hat{\Gamma}_N$ operator:

$$\hat{V}_{T-t}(s_i) = \hat{\Gamma}_N^t(V_T)(s_i), \tag{5.90}$$

for $i = 1, \ldots, N$, $t = 0, \ldots, T$.

**Theorem 5.9:** *Randomization breaks the curse of dimensionality of solving finite horizon DDP's: i.e. the worst case complexity of the class of randomized algorithms for solving a $T$-period DDP problem with $|A|$ possible actions, discount factor $\beta \in (0,1)$ and utility function and transition probability $(u, p)$ satisfying $(A1), \ldots, (A4)$ is given by:*

$$comp^{\textit{wor-ran}}(\epsilon, d) = O\left(\frac{T d^4 |A|^3 K_u^4 K_p^4}{(1-\beta)^8 \epsilon^4}\right). \tag{5.91}$$

**Corollary** *Randomization breaks the curse of dimensionality of solving infinite horizon DDP problems. An upper bound on the worst case randomized complexity of the infinite horizon DDP problem is given by:*

$$comp^{\textit{wor-ran}}(2\epsilon, d) = O\left(\frac{log\,(1/(1-\beta)\epsilon) d^4 |A|^3 K_u^4 K_p^4}{|\,log(\beta)|(1-\beta)^8 \epsilon^4}\right). \tag{5.94}$$

**Random Multigrid Algorithm**

The multigrid algorithm consists of a number of "outer" iterations $k = 1, 2, \ldots$, where a number $N_k$ of uniform random grid points $\{s_1, \ldots, s_{N_k}\}$ is drawn at each iteration $k$ independently of the grid points drawn at previous iterations $k-1, k-2, \ldots$ of the multigrid algorithm. The basic idea is to start at iteration $k = 0$ with a relatively small number of grid points $N_0$ and successively increase the number of grid points drawn at each iteration according to:

$$N_k = 2^{2k} N_0. \tag{5.96}$$

Within each outer iteration $k$, a number $T(k)$ of inner successive approximation steps are taken using the random Bellman operator $\hat{\Gamma}_{N_k}$. Let $\hat{V}_k$ denote the value function produced at the termination of the $T(k)$ successive approximation steps at outer iteration $k$. The starting point for successive approximations in outer iteration $k$ is the value function $\hat{V}_{k-1}$ produced at outer iteration $k-1$. Thus we have the recursion:

$$\hat{V}_k = \hat{\Gamma}_{N_k}^{T(k)}(\hat{V}_{k-1}), \tag{5.97}$$

where the starting point for iteration $0$ of the multigrid algorithm is given by $max_{a \in A}\, u(s, a)$. The number of inner successive approximation steps $T(k)$ is determined as the smallest integer $t$ satisfying:

$$\left| \hat{\Gamma}_{N_k}^{t}(\hat{V}_{k-1}) - \hat{\Gamma}_{N_k}^{t-1}(\hat{V}_{k-1}) \right| \leq \frac{K}{\sqrt{N_k}\beta(1-\beta)}. \tag{5.99}$$

**Lemma 5.5** *There exists a constant $c$ independent of $\beta$ and $\epsilon$ such that:*

$$T(k) \leq \frac{c}{|log(\beta)|}, \quad k = 1, \ldots, k^*. \tag{5.104}$$

**Theorem 5.10** *An upper bound on the worst case complexity of infinite horizon DDP problems is given by:*

$$comp^{\textit{wor-ran}}(2\epsilon, d) = O\left(\frac{|A|^3 d^4 K_u^4 K_p^4}{|\, log(\beta)|(1-\beta)^8 \epsilon^4}\right). \tag{5.114}$$

*and can be attained by the random multigrid algorithm described above.*