Mean Field Limits By Population Acceleration

Harsha Honnappa\textsuperscript{1} and Rahul Jain\textsuperscript{2} and Amy R. Ward\textsuperscript{3}

Abstract—Mean field or large population limits are useful analytical tools that provide well justified approximations to finite systems. Motivated by problems in Game Theory and Queueing Theory we provide a framework to develop mean field limits that emerge from finite population systems in a very general setting. Our results prove new generalizations of the standard Glimenko–Cantelli and Donsker's Theorems. We prove these limits by making use of an analytical technique called population acceleration first introduced in [1] to study transitory queueing models.

I. INTRODUCTION

Asymptotic approximations help simplify analyses of mathematically complicated systems. For instance, in queueing theory non-Markovian queueing systems are typically studied in a heavy-traffic setting, where the load in the system is close to one, and time and space are scaled inversely in proportion to the load in the system. Here, it can be rigorously shown that the workload and queue length processes can be approximated by a reflected Brownian Motion ([2], [3]). Typically, it is not natural to think of the heavy-traffic setting in a mean field sense. However, in [4], [1] we studied transitory queueing systems by increasing the population size to infinity without rescaling time, similar to a mean field limit. We call this technique population acceleration. It is shown that the workload and queue length processes can be justifiably approximated by a non-stationary reflected diffusion process. Mean field asymptotics have also been used in queueing theory to study large, complex, queueing networks in the limit of a large network size [5], [6]. Typically, large network sizes justify a standard ansatz in queueing theory that nodes in a large queueing network are statistically independent.

In Game Theory too, beginning with Schmeidler [7], there is a large literature of identifying so-called non-atomic or mean field equilibria; see [7], [8] on one-shot games, and [9] for dynamic games. Here, mean field refers to the fact that the equilibria emerge as a result of the collective effect of a large number of infinitely divisible players (i.e., each player is non-atomic or of measure zero). Mean field equilibria can be easier to compute than the atomic equilibria, while still providing significant utility as an approximation to the finite game. This is usually the case when the strategy/action space is infinite or high dimensional. As an example, consider a finite population of customers arriving at a lunch cafeteria. It is natural to expect that customers choose a time to arrive at the cafeteria such that their expected waiting time is minimized. This type of strategic arrival behavior was studied in [19], who introduced the Concert Queueing Game and identified the mean field (or non-atomic) Nash equilibrium arrival profile. This work was subsequently extended by [20] to feedforward queueing networks, where the Nash equilibrium (joint) arrival and routing profile were identified in a non-atomic setting. On the other hand, computing the equilibrium profile in the finite population game is prohibitively hard as shown by [21]; indeed the authors can only solve for the equilibrium numerically. However, it is shown that in the limit of a large population size the finite population equilibrium does converge to the mean field equilibrium, implying that the latter is a reasonable facsimile for the former. More generally, mean field equilibria have been used as approximations to a variety of finite population games in the literature [22], [23]; see Section 4 of [22] for an excellent survey of relevant results in Game Theory and Mechanism Design.

In any mean field system, what is interesting is the collective effect of the infinitesimal participants in the system (which could be a game, or a stochastic system like a queue). Mathematically speaking, this collective effect is most naturally represented as empirical sums in the finite population (or pre-limit) setting, and the mean field limits emerge as empirical process limits. Standard empirical process limit theory assumes that the pre-limit variables are statistically identical. This assumption is with a significant loss of generality - for instance, from a game theoretic perspective, this implies that only symmetric mixed strategies can be considered. Our goal, in this paper, is to provide a framework to generalize the development of empirical process limits where the pre-limit variables are not statistically independent, and to do so in a problem agnostic manner. In particular, we study the generic situation of each participant in the system sampling a random variable, and the collective effect presenting as a summation over the random samples. This situation is manifest in games with mixed strategies, and in queueing theory where users sample a time to arrive at a queue from a distribution function (see [1]).

A. Model

Let $\mathcal{K} \coloneqq [0, 1]$ represent the 'universe' of all possible participants in the system, and $\mathcal{S}$ a compact sample/strategy space; in this paper we assume this is an interval of the real line. Associated with user $s \in \mathcal{K}$ is a sampling distribution

\begin{itemize}
  \item [1] Electrical Engineering, University of Southern California, Los Angeles CA 90089 honnappa@usc.edu
  \item [2] Electrical Engineering and Industrial Systems Engineering, University of Southern California, Los Angeles CA 90090 rahul.jain@usc.edu
  \item [3] Data Sciences Operations, Marshall School of Business, University of Southern California, Los Angeles CA 90089 amyward@marshall.usc.edu
\end{itemize}
{G_s(t); \forall t \in S}. The mean field average profile \( F(t) := \int_0^1 G_s(t) \, ds \) is a Lebesgue-Stiltjes average over the collection of sampling distributions.

Consider a system with \( N \) users. Given a collection of sampling distributions, \( \{G_1(t), G_2(t), \ldots, G_N(t)\} \), let \( T_i \) be a random sample from distribution \( G_i \), and define the population average profile as \( \frac{1}{N} \sum_{i=1}^N G_i(t) \). Under what circumstances will this population average scale to a mean field profile? Further, is it possible to quantify the difference between these? In this paper we provide affirmative answers to these questions. We draw upon and expand results in empirical process theory, by proving new generalized Glivenko-Cantelli and generalized Donsker’s Theorems. As noted before, we develop these limits as the population size scales to infinity. To achieve this, we consider a non-decreasing sequence of population sizes \( N_n \) (formalized in Section III below), that tends to infinity as the index increases. For each population size, the random samples and sampling distributions form a row of an infinite triangular array. As the population size increases the number of samples obtained in any compact subset of the support of the sampling distributions increases. Put another way, the “sampling rate” is accelerated by the population size. We call this population acceleration (PA).

II. RELATED WORK

The work in this paper is related to many different areas of applied mathematics. In statistics and probability theory, in particular, the standard Glivenko-Cantelli and Donsker’s Theorems (see [11], [12]) are used to justify the convergence of empirical processes of independent and identically distributed random variables. For non-identically distributed random variables, it is necessary to look for ‘generalized’ Glivenko-Cantelli and Donsker’s Theorems. There has been some work on this in the empirical process theory literature: in [13] and [14] the authors prove the existence of limits and demonstrate convergence of empirical sums of non-identically distributed random variables. However, they do not explicitly identify the limits. For applications, this can be a serious lacunae. Our results, on the other hand, explicitly identify the limit, and provide an alternative approach for proving the limits.

We also note that the techniques developed in this paper were used to prove fluid and diffusion approximations to a class of finite population queueing models in [4], [1]. Typical analyses of queueing models implicitly assume the existence of an infinite population of users. However, in analyzing time-varying, non-stationary and (possibly) non-ergodic queueing systems it is useful to study the model on a finite time-horizon (hence implying a finite population). The results in [4], [1] provide a framework for studying such queueing models, and appropriate limits were identified.

III. MEAN FIELD LIMITS

Assume that \( N_n \) is a deterministic, non-decreasing sequence of natural numbers and that \( N_n = O(n) \) and \( n \geq 1 \). Consider a sequence of systems indexed by \( n \), with population size \( N_n \) in the \( n \)th system, and with sample space \( A = [0, 1] \). Let \( (\Omega, F, \mathbb{P}) \) represent some probability space. Let \( F_{n,i} : A \rightarrow [0, 1] \), for each \( i \in \{1, \ldots, N_n\} \), represent the sampling distribution of user \( i \) in the \( n \)th system. Notice that \( \{F_{n,i}, i \in \{1, \ldots, N_n\}\} \) \( \forall n \in \mathbb{N} \) forms a triangular array of distribution functions. Let \( T_n := (T_{n,1}, \ldots, T_{n,N_n}) \) represent a tuple of the random samples in the finite population system, where \( T_{n,i} : \Omega \rightarrow A \). Assume that the joint distribution of \( T_n \) is of product form with respect to \( \mathbb{P} \). \( \mathbb{P}(T_n \in \Pi_{n=1}^{N_n}[0, t_i]) = \Pi_{i=1}^{N_n} F_{n,i}(t_i) \), for any Borel set \( \Pi_{n=1}^{N_n}[0, t_i] \subset [0, 1]^{N_n} \).

The population average profile in the \( n \)th system is \( F_n := \frac{1}{N_n} \sum_{i=1}^{N_n} F_{n,i} \). However, what is really observed as a result of sampling is \( T_n \), and the empirical average profile is the empirical sum \( A^n(t) := \frac{1}{N_n} F_n(t), \forall t \in [0, 1] \), where \( A^n(t) := \sum_{i=1}^{N_n} \mathbf{1}(T_{n,i} \leq t) \). As noted in Section I, our goal is to identify the limit, as \( n \rightarrow \infty \), to which the population average strategy profile scales to. Intuitively, one might expect that any limit should be an integral over the individual sampling distributions.

We formalize this notion, by placing the following restriction on the cumulative distribution functions \( F_{n,i} \). Recall that \( K \) represents the set of user indices (in the ‘universe’ of possible users). Let \( (K, \mathcal{B}(K), m) \) represent the sample space of the user indices, where \( \mathcal{B}(K) \) is the Borel \( \sigma \)-algebra on \( K \) and \( m \) is the Lebesgue measure. Let \( \mathcal{A} \) be the space of all distribution functions with support \( A \). Following [10], we define a random distribution function as a mapping \( \Upsilon : K \rightarrow \mathcal{A} \). Thus, \( (F_n(t) := \Upsilon(s)(t), \forall t \in A) \), is the distribution function of customer \( s \in K \). Clearly, \( \Upsilon \) induces a sample space \( (\mathcal{L}_A, \mathcal{B}(\mathcal{L}_A), \mathbb{P}) \) where \( \mathcal{B}(\mathcal{L}_A) \) is the Borel \( \sigma \)-algebra containing the weak-* topology on \( \mathcal{L}_A \) and \( \mathbb{P} = m \circ \Upsilon^{-1} \) is the measure induced on the space \( \mathcal{L}_A \).

The average distribution function \( \bar{F} \) is now well defined in relation to \( \mathbb{P} \) as \( \bar{F} := \int_{F \in \mathcal{L}_A} F \, d\mathbb{P}(F) = \int_{K} \Upsilon(s)(t) \, m(ds) = \int_{\mathcal{L}_A} F_n(t) \, d\mathbb{P} \forall t \in A \). Notice that \( \Upsilon \) is a measure-valued stochastic process with domain \([0, 1]\) and range \( \mathcal{L}_{[0,1]} \). It is useful to view \( \Upsilon \) in the following sense: it represents a summary of the “beliefs” of all possible users \( s \in [0, 1] \). While the total order property of \( K \) plays no role in our description of the population of customers, it is not unusual to expect that customers “close” to each other, in the sense of the Euclidean norm on \( K \), should have similar beliefs. Thus, we impose the condition that \( \Upsilon \) satisfies \( \| \Upsilon(\omega_1) - \Upsilon(\omega_2) \| \leq K|\omega_1 - \omega_2| \), for any \( \omega_1, \omega_2 \in K \), \( K < \infty \) is some given constant and \( \| \cdot \| := \sup_{\omega \in A} \| \cdot \| \).

In particular, in the simplest case where \( \Upsilon(\omega) = \delta_F(\omega) \) for all \( \omega \in K \) and some \( F \in \mathcal{L}_A \) (i.e., in an i.i.d. model), this condition is satisfied automatically.

Without loss of generality, assume user \( i \in \{1, \ldots, N_n\} \) corresponds to the point \( i/N_n \in (0, 1) \). We start by proving that the population average profile has a well defined limit. The following lemma shows that \( \bar{F}_n \) converges to a limit \( \bar{F} \) as \( n \rightarrow \infty \).

**Lemma 1:** There exists a distribution function \( \bar{F} \) such
that
\[ \bar{F}_n(t) \rightarrow \bar{F}(t) := \int_K F_p(t) m(dp), \tag{1} \]
uniformly on \([0,1]\) as \(n \to \infty\).

The proof is relegated to the appendix.

The limit for the empirical profile is a generalization of the Glivenko-Cantelli Theorem [15] to triangular arrays of non-identically distributed random variables. In Theorem 1 we show that the normalized process \(A^n\) converges uniformly on compact sets of \([0,1]\) to \(F\) in \(D := D[0,1]\), the space of all functions that are right continuous with left limits with domain \([0,1]\). We prove this result by demonstrating the uniform convergence of the sample paths of the empirical profile. We make the reasonable assumption that none of the distribution functions \(F_p\) share discontinuity points in the support. This implies that the limit \(\bar{F}\) is (almost surely) continuous, allowing us to prove convergence in the uniform metric. Here, \(U\) represents convergence uniformly on compact sets of \([0,\infty)\), and \(a.s.\) represents almost sure convergence with respect to \(P\).

**Theorem 1 (Glivenko-Cantelli for Triangular Arrays):**

The process \(A^n = \frac{1}{N_n} (A^n - F)\) satisfies a functional strong law of large numbers,
\[ \bar{A}_n \xrightarrow{a.s.} \bar{F} \text{ in } (D,U), \]
as \(n \to \infty\).

The proof is presented in the appendix.

**Remarks.** Versions of this theorem have been proved in the literature and we draw attention, in particular, to Theorem 1 of [13] that proves an existence result and does not explicitly identify the mean field limit. Our construction of the empirical distribution via random distribution functions allows us to do so.

As a consequence of our construction of the empirical profile space and Lemma 1, we can also quantify the error in scaling the empirical profile to the mean field profile in a functional Central Limit Theorem (fCLT) result. This result generalizes the well known Donsker’s Theorem. Let \(A^n := \sqrt{N_n} (A^n - F)\) represent this “error” process. Here \(\Rightarrow\) represents weak convergence.

A straightforward calculation shows that the covariance function of the error process is \(K_n(s,t) := E[A^n(s)A^n(t)] = \frac{1}{N^2_n} \sum_{i=1}^{N_n} F_{n,i}(s \wedge t) - F_{n,i}(s) F_{n,i}(t)\). The following lemma shows that \(K_n\) has a well defined limit as \(n \to \infty\).

**Lemma 2:** There exists a function \(K(s,t)\) such that,
\[ K_n(s,t) \rightarrow K(s,t) := \int_K (F_p(s \wedge t) - F_p(s)F_p(t)) m(dp), \tag{2} \]
as \(n \to \infty\), uniformly for all \(s,t \in [0,1]\).

The proof is available in the appendix.

**Theorem 2 (Donsker’s for Triangular Arrays):** The mean field error \(A^n\) satisfies a functional central limit theorem,
\[ \bar{A}_n \Rightarrow \bar{W} \text{ in } (D,U), \]
as \(n \to \infty\), where \(\bar{W}\) is a mean zero Gaussian process with covariance function \(K(s,t)\) defined in (2) and continuous sample paths.

The proof can be found in the appendix.

**Remarks.** Our result is a generalization of Hahn’s Central Limit Theorem (Theorem 2 in [18]) to nonidentically distributed random elements of \(D\). We also draw attention to Theorem 1.1 of [14] that proves the existence of an empirical process limit for triangular arrays, under the sufficient condition that an appropriate covariance function exists. However, it does not specifically identify the Gaussian process limit and our identification of the empirical distribution function space with random distribution functions enables this identification.

The covariance structure of the process \(\bar{W}\) is interesting in itself, and we make the following observations. First, notice that the covariance function is an average of the covariance functions of the Brownian Bridge processes \(W^0 \circ F_p\) (with \(p \in K\)) where \(W^0\) is a standard Brownian Bridge. In a sense, these are Brownian Bridge processes associated with empirical processes of random samples from the function \(F_p\). Second, differentiating the expression for \(K(t,t)\) with respect to \(t\) we have \(\frac{dK(t,t)}{dt} = \int_K (f_p(t) - 2f_p(t)F_p(t)) m(dp)\), where \(f_p\) is the density (or at least the right-derivative) of the distribution function \(F_p\). This is the average of the infinitesimal variance of the Brownian Bridges \(W^0 \circ F_p\).

Recall that the infinitesimal mean and variance of a diffusion process are defined as \(\mu_t = \int_0^h \frac{d}{dh} E[X(t+h)-X(t)|X(t)=x] \, dx\) and \(\sigma_t^2 = \int_0^h \frac{d}{dh} \left( E[X(t+h)-X(t)]^2 |X(t)=x]\right) \, dx\) as \(h \to 0\) (resp.). For the Brownian Bridge process \(W^0 \circ F_p\), it is well known that the infinitesimal mean and variance are (for a fixed \(p \in K\))
\[ \mu_p(t,y) = -\frac{y f_p(t)}{1 - F_p(t)} \]
\[ \sigma_p^2(t,y) = f_p(t). \]

Further, it can be shown that the mean and variance of the Brownian Bridge satisfies the following o.d.e.’s:
\[ \frac{d}{dt} \mathbb{E} [W^0 \circ F_p(t)] = \mathbb{E} \left[ \mu_p(t, W^0 \circ F_p(t)) \right] \]
\[ = \frac{-f_p(t)}{1 - F_p(t)} \mathbb{E} [W^0 \circ F_p(t)] = 0 \]
\[ \frac{d}{dt} \text{Var} (W^0 \circ F_p(t)) = + 2 \mathbb{E} [W^0 \circ F_p(t) \times \mu_p(t, W^0 \circ F_p(t))] \]
\[ = f_p(t) - 2f_p(t)F_p(t). \]

Comparing the variance derivative above with \(\frac{dK(t,t)}{dt}\), we conjecture that the \(\bar{W}\) is a Gaussian diffusion process with infinitesimal generator equal to the average of the infinitesimal generators of the Brownian Bridges \(W^0 \circ F_p\). However, we have not been able to verify that the process is Markov with respect to its natural filtration to make a definitive conclusion.

A particular case is that of symmetric sampling distributions, implying that the random samples are independent.
and identically distributed. This result, of course, is the standard fCLT for the empirical process (see [16] for a deeper exposition).

**Corollary 1:** For each \( n \geq 1 \), let \( \{T_{n,i}, i = 1, \ldots, N_n\} \) be a triangular array of i.i.d. random samples drawn from a distribution \( F \). Then, as \( n \to \infty \)

\[
\hat{A}^n \Rightarrow W^0 \circ F \quad \text{in} (D, U).
\]

Here \( W^0 \) is the standard Brownian Bridge process defined on the common sample space.

A formal proof of this result is standard and omitted (see Chapter 13 of [16]). However, it is also straightforward to see this from Theorem 2 by setting \( F_p = F \) for all \( p \in [0,1] \). It can be readily verified that the Gaussian process \( \tilde{W} \) is equal in distribution to a Brownian Bridge process.

**IV. Conclusions**

In this paper, we provide a framework for studying the convergence of finite population empirical average profiles to mean field profiles. We do so by making use of the notion of a random distribution function as introduced by Dubins and Freedman in [10], and proving new generalizations of the well known Glivenko-Cantelli Theorem and Donsker’s Theorem of empirical process theory. To be precise, in the standard theory samples are assumed to be independent and identically distributed (i.e. one assumes symmetric mixed strategies), however, in general, sampling distributions need not be identical. While there are some results in the empirical process theory literature that imply the existence of limits to the strategy profiles in this generalized setting, none of these are constructive and the limit is never identified. Our framework allows the identification of the limit and to precisely quantify the error between the empirical average profile and the mean field profile in a functional Central Limit Theorem, where the limit is a Gaussian process whose covariance function is an average over an infinite collection of Brownian Bridge processes. While our main contribution is the proof of generalized Glivenko-Cantelli and Donsker’s Theorems, our framework appears to be useful for studying when equilibrium strategy profiles in the finite population game converge to the non-atomic equilibrium.

The reason for studying mean field equilibria is that they are tractable approximations to finite population equilibria. However, it is not clear what happens if mean field equilibrium strategies are followed in the finite population game. For instance, in the Concert Queueing Game, the symmetric Nash Equilibrium Profile is a uniform distribution function. However, this strategy can be strictly improved upon in any finite population Concert Queueing Game. On the other hand, is it likely that a mean field equilibrium will imply an \( \epsilon \)-Nash equilibrium in the finite population game? Is it possible to quantify this relationship in a meaningful manner? We believe that the random distribution function \( \Upsilon \) provides the right tool to answer these questions in a very general manner. To make this precise, suppose \( \Upsilon^* \) is a random distribution function such that \( \bar{F}^*(t) := \int_{[0,1]} \Upsilon^*(s)(t) m(ds) = \int_0^1 F_p^*(t) ds \) is an equilibrium strategy profile of the non-atomic limit game. Now, consider the \( n \)th game in sequence such that the population size is \( N_n \). Then, does \( \{\Upsilon^*(1/N_n), \ldots, \Upsilon^*(N_n/N_n)\} \) form an \( \epsilon \)-Nash Equilibrium? Does \( \frac{1}{N_n} \sum_{i=1}^{N_n} \Upsilon^*(i/N_n) \) form an \( \epsilon \)-Nash Equilibrium Profile? How far away are we from the “equilibrium path?” We leave a complete analysis of these questions for future work.

**APPENDIX**

**A. Proof of Lemma 1**

For each \( n \in \mathbb{N} \) we have \( F_{n,i} = \Upsilon \left( \frac{i}{N_n} \right) \), \( i = 1, \ldots, N_n \). Therefore, \( (1) \) can be rewritten as

\[
\bar{F}_{n}(t) = \frac{1}{N_n} \sum_{i=1}^{N_n} \Upsilon \left( \frac{i}{N_n} \right)(t),
\]

and we prove that

\[
\left\| \frac{1}{\sqrt{N_n}} \sum_{i=1}^{N_n} \Upsilon \left( \frac{i}{N_n} \right)(t) - \int_{[0,1]} \Upsilon(s)(t) m(ds) \right\|_{[0,1]} \to 0
\]

as \( n \to \infty \). Notice that \( \bar{F}(t) \) is a Riemann-Stieltjes integral with respect to the Lebesgue measure. Therefore, it is natural to view \( F_n \) as a Riemann-Stieltjes (pre-limit) sum. For a fixed \( t \in [0,1] \), therefore, we show that the Riemann sums converge to the Riemann-Stieltjes integral.

Let \( M_i(t) := \sup \Upsilon(x)(t) \) and \( m_i(t) := \inf \Upsilon(x)(t) \) for all \( x \in \left[ \frac{i-1}{N_n}, \frac{i}{N_n} \right] \) and \( i = 1, \ldots, N_n \). We define the “upper” and “lower” Riemann sums as (respectively) \( U_n(t) := \sum_{i=1}^{N_n} M_i \left( \frac{i}{N_n} - \frac{i-1}{N_n} \right) \) and \( N_n(t) := \sum_{i=1}^{N_n} m_i \left( \frac{i}{N_n} - \frac{i-1}{N_n} \right) \). Clearly, for every \( \epsilon > 0 \) and large enough \( n \), \( U_n(t) - N_n(t) < \epsilon \) due to the Lipschitz continuity property assumed for \( \Upsilon \). This is tantamount to showing that the bound holds for at least one possible partition of \([0,1]\). Then, by Theorem 6.6 of [24], it follows that the limit exists and is equal to \( \bar{F}(t) \). The Lipschitz continuity property implies that the limit clearly holds for all \( t \in S \), implying uniform convergence.

**B. Proof of Lemma 2**

The first summation in the definition of \( K_n(s,t) = \frac{1}{N_n} \sum_{i=1}^{N_n} F_{n,i}(s \land t) - \frac{1}{N_n} \sum_{i=1}^{N_n} F_{n,i}(s) F_{n,i}(t) \) converges to \( \int_K F_p(s \land t) m(dp) \) as \( n \to \infty \) by Lemma 1. The second summation converges as well by using the same Riemann-Stieltjes summation argument used in the proof of Lemma 1, and the limit is \( \int_K F_p(t) F_p(s)m(dp) \).
C. Proof of Theorem 1

First, fix \( t \in [0,1] \) and \( \epsilon > 0 \). Consider,

\[
\Pr \left( \left| \bar{A}_n(t) - \bar{Y}(t) \right| > \epsilon \right) = \left( \frac{1}{N} \right)^2 \sum_{i=1}^{N_n} \Pr \left( \left| \bar{Y}(i/N_n(t)) \right| > \epsilon \right)
\]

\[
= \frac{1}{N^2} \sum_{n=1}^{N_n} \mathbb{E} \left| \bar{Y}(i/N_n(t)) \right|^4
\]

\[
= \frac{1}{N^2} \sum_{n=1}^{N_n} \left( \sum_{i=1}^{N_n} \left( \bar{Y}(i/N_n(t)) \right)^4 \right)
\]

\[
\leq \frac{1}{N^2} \sum_{n=1}^{N_n} \left( \sum_{i=1}^{N_n} \left( \bar{Y}(i/N_n(t)) \right)^2 \right)
\]

\[
\leq \frac{1}{N^2} \sum_{n=1}^{N_n} \left( \sum_{i=1}^{N_n} \left( \bar{Y}(i/N_n(t)) \right)^2 \right)
\]

where the first inequality is due to Chebyshev’s inequality and the second equality follows by expanding the fourth power of the summation and taking an expectation. The last inequality follows by noting that \( \bar{Y}(i/N_n(t)) \) is Bernoulli with mean \( \bar{Y}(i/N_n(t)) \). Since the fourth central moment of a Bernoulli random variable with expectation \( p(1-p) \) is bounded by \( 4p^2(1-p) \) for all \( p \in (0,1) \), the second equality follows by expanding the fourth moment of a Bernoulli random variable.

By Lemma 1, \( \bar{A}_n \) converges to \( \bar{F} \) almost surely pointwise.

Next, consider a uniform partition of the support \([0,1]\), and suppose \( \frac{1}{2M} \leq t \leq \frac{1}{M} \), where \( j = 1, \ldots, M \) and \( M \) is the size of the partition. Then, for fixed \( n \), \( \bar{A}_n \left( \frac{1}{M} \right) \leq \bar{A}_n(t) \leq \bar{A}_n(t) \), implying that \( \frac{1}{N} \sum_{i=1}^{N_n} \left( \bar{Y}(i/N_n(t)) \right) = \bar{F}_n \).

For each \( M \), there exists \( n_M \) such that for all \( n \geq n_M \), \( |F_{n,i}(j/M) - F_{n,i}(j-1/M)| \leq \frac{1}{M} \). Further, for \( \epsilon > 0 \), there exists \( n_M' \) such that for all \( n \geq n_M' \),

\[
\left| \frac{1}{N_n} \sum_{i=1}^{N_n} \left( \bar{Y}(i/N_n(t)) \right) - \left( \frac{1}{N_n} \sum_{i=1}^{N_n} \left( \bar{Y}(i/N_n(t)) \right) \right) \right| < \frac{\epsilon}{2},
\]

where \( k = j - 1/M \). It follows that \( \sup_{t \in [0,1]} \left| \frac{1}{N_n} \sum_{i=1}^{N_n} \left( \bar{Y}(i/N_n(t)) \right) - \left( \frac{1}{N_n} \sum_{i=1}^{N_n} \left( \bar{Y}(i/N_n(t)) \right) \right) \right| < \epsilon + \frac{1}{M} \).

Since \( \epsilon \) is arbitrary, letting \( M \to \infty \) the desired result follows.

D. Proof of Theorem 2

We first prove pointwise convergence by verifying the sufficiency of the Lyapunov Central Limit Theorem (Theorem 7.3 [16]). Fix \( t \in [0,1] \) and \( \delta > 0 \), and consider

\[
\frac{1}{N_n} \sum_{i=1}^{N_n} E \left| F_{n,i}(t) - 1 \right|^{2+\delta} \leq \frac{1}{N_n} \sum_{i=1}^{N_n} \left( F_{n,i}(t) - 1 \right)^{2+\delta} \leq \frac{1}{N_n} \sum_{i=1}^{N_n} \left( F_{n,i}(t) - 1 \right)^{2+\delta} \Rightarrow \mathcal{N}(0, 1).
\]

By Lemma 2 it follows that

\[
\frac{1}{\sqrt{N_n}} \sum_{i=1}^{N_n} \left( F_{n,i}(t) - 1 \right) \Rightarrow \mathcal{N}(0, 1).
\]

Finally, we verify the sufficiency of Theorem 15.6 of [16] to show that \( A_n \Rightarrow \bar{W} \) in \( (D, J) \). To ease the notation, let \( X_{n,i} := (1 \bar{Y}(i/N_n(t)) - F_{n,i}(t)) \). It can be shown, by Chebyshev’s inequality and some algebra, that for any \( \lambda > 0 \) and \( t_1 \leq t \leq t_2 \), \( \lambda^4 \text{Pr}(A_n(t) - A_n(t_1)) \geq \lambda, |A_n(t) - A_n(t_2)| \geq \lambda) \leq C \), where \( C \geq 8 \) and the bound is true for all \( t_2 > t_1 \).

Theorem 15.6 of [16] shows that if \( \text{Pr}(A_n(t) - A_n(t_1)) \geq \lambda, |A_n(t) - A_n(t_2)| \geq \lambda) \leq (G(t_2) - G(t_1))^{2\alpha} \), where \( G \) is a non-decreasing function on \([0,1]\) and \( \alpha > 1/2 \), then \( A_n \) converges weakly to a limit in \( (D, J) \). Therefore, \( A_n \Rightarrow \bar{W} \) as \( n \to \infty \). Finally, by part (ii) of Theorem 1.1 of [14] we know that \( \bar{W} \) has continuous sample paths, implying that \( A_n \Rightarrow \bar{W} \) in \( (D_{lim}, U) \), thus completing the proof.
REFERENCES


