Lecture 15
Discrete Time Random Processes

The goal of this lecture is to prepare the theoretical grounds for the applications in the next lecture. We will derive a fundamental result on discrete time random processes which shows that, under certain mild conditions, the values of the random variables in such a process are concentrated around their expectations.

15.1. Preliminaries

Let Ω denote a probability space. In our applications, Ω will be derived from the uniform distribution on bipartite graphs with a given number of nodes, and given fractions of nodes of various degrees. A *discrete time random process* $Q = (Q_0, Q_1, \ldots)$ over Ω with state space S is a sequence of random variables $Q\colon \Omega \to S$, where S is a measurable set. We will call the indices of the random variables involved the *time indices*. The *history* of a discrete time random process up to time t is the vector $H_t = (Q_0, Q_1, \ldots, Q_t)$.

Suppose that $X = (X_0, X_1, \ldots, X_N)$ and $H = (H_0, H_1, \ldots, H_N)$ are sequences of real valued random variables. We call X a *supermartingale* with respect to H if $\mathrm{E}[X_i] < \infty$ for all $0 \leq i \leq N$, $\mathrm{E}[X_i | H_0, \ldots, H_i] = X_i$ (X_i is determined if $H_0, \ldots, H_i$ are known) and

$$\mathrm{E}[X_{i+1} \mid H_0, \ldots, H_i] \leq X_i.$$

Similarly, we call X a *submartingale* with respect to H if

$$\mathrm{E}[X_{i+1} \mid H_0, \ldots, H_i] \geq X_i.$$

X is called a *martingale* with respect to H if it is both a super- and a submartingale with respect to H. It is called a martingale (without reference to H) if it is a martingale with respect to itself.

Theorem 15.1 (Azuma-Hoeffding Inequality). *Suppose that $X = (X_0, \ldots, X_N)$ is a supermartingale with respect to $H = (H_0, \ldots, H_N)$ and that $|X_{k+1} - X_k| \leq c$ for some $c > 0$ and $0 \leq k < N$. Then for all $\alpha > 0$ we have*

$$\Pr[X_k \geq X_0 + \alpha c] \leq \mathrm{e}^{-\alpha^2/2k}.$$

Proof. By Markov's inequality we have for all $a \geq 0$ and all $\lambda > 0$:

$$\Pr[X_k - X_0 \geq \lambda] = \Pr[\mathrm{e}^{a(X_k - X_0)} \geq \mathrm{e}^{a\lambda}] \leq \frac{\mathrm{E}[\mathrm{e}^{a(X_k - X_0)}]}{\mathrm{e}^{a\lambda}}. \tag{15.1}$$

Lecture by Amin Shokrollahi
Scribes by Amin Shokrollahi, Satish Korada

Let $Y = (Y_0, \dots, Y_N)$ be defined by $Y_0 = X_0$ and $Y_k := X_k - X_{k-1}$ for $k \geq 1$. Since the function $\exp(ax) = \mathrm{e}^{ax}$ is convex, we have

$$\begin{aligned}\exp(aY_i) &= \exp\left(ac\frac{1+Y_i/c}{2} - ac\frac{1-Y_i/c}{2}\right)\\ &\leq \frac{1+Y_i/c}{2}\mathrm{e}^{ac} + \frac{1-Y_i/c}{2}\mathrm{e}^{-ac}\\ &= \frac{\mathrm{e}^{ac}+\mathrm{e}^{-ac}}{2} + \frac{Y_i}{2c}(\mathrm{e}^{ac} - \mathrm{e}^{-ac}).\end{aligned}$$

We also have

$$\begin{aligned}\frac{\mathrm{e}^{ac}+\mathrm{e}^{-ac}}{2} &= \sum_{j=0}^{\infty}\frac{a^{2j}c^{2j}}{(2j)!}\\ &\leq \sum_{j=0}^{\infty}\frac{a^{2j}c^{2j}}{2^j j!}\\ &= \mathrm{e}^{a^2c^2/2}.\end{aligned}$$

Further,

$$\mathrm{E}[Y_i|H_0,\dots,H_{i-1}] = \mathrm{E}[X_i - X_{i-1}|H_0,\dots,H_{i-1}] = \mathrm{E}[X_i|H_0,\dots,H_{i-1}] - X_{i-1} \leq 0,$$

by assumption. This implies that

$$\mathrm{E}[\mathrm{e}^{aY_i}|H_0,\dots,H_{i-1}] \leq \mathrm{e}^{a^2c^2/2}.$$

For random variables X and Z we have $\mathrm{E}[X] = \mathrm{E}[\mathrm{E}[X|Z]]$. Therefore, we have

$$\begin{aligned}\mathrm{E}[\mathrm{e}^{a(X_k-X_0)}] &= \mathrm{E}[\mathrm{e}^{a(Y_1+\cdots+Y_k)}]\\ &= \mathrm{E}\left[\mathrm{E}[\mathrm{e}^{a(Y_1+\cdots+Y_k)}|H_0,\dots,H_{k-1}]\right]\\ &= \mathrm{E}\left[\mathrm{e}^{a(Y_1+\cdots+Y_{k-1})}\mathrm{E}[\mathrm{e}^{aY_k|H_0,\dots,H_{k-1}}]\right]\\ &\leq \mathrm{E}\left[\mathrm{e}^{a(Y_1+\cdots+Y_{k-1})}\mathrm{e}^{a^2c^2/2}\right]\\ &\vdots \quad \vdots\\ &\leq \mathrm{e}^{a^2c^2k/2}.\end{aligned}$$

From (15.1) we obtain

$$\Pr[X_k - X_0 \geq \lambda] \leq \mathrm{e}^{a^2c^2k/2-\lambda a}.$$

Choosing $a = \lambda/(kc^2)$, this gives

$$\Pr[X_k - X_0 \geq \lambda] \leq \mathrm{e}^{-\frac{\lambda^2}{2kc^2}},$$

which finishes the proof (set $\lambda = \alpha c$). □

We immediately obtain the following corollaries.

Corollary 15.2. *Suppose that $X = (X_0, \dots, X_N)$ is a submartingale with respect to $H = (H_0, \dots, H_N)$ and that $|X_{k+1} - X_k| \leq c$ for some $c > 0$ and $0 \leq k < N$. Then for all $\alpha > 0$ we have*

$$\Pr[X_k \leq X_0 - \alpha c] \leq \mathrm{e}^{-\alpha^2/2k}.$$

Proof. Set $Z_k := -X_k$. Then $Z = (Z_0, \ldots, Z_N)$ forms a supermartingale with respect to $(H_0, \ldots, H_N)$, and $|Z_{k+1} - Z_k| \leq c$ for $0 \leq k < N$. Therefore, by Theorem 15.1 we have

$$\mathrm{e}^{-\alpha^2/2k} \geq \Pr[Z_k \geq Z_0 + \alpha c] = \Pr[X_k \leq X_0 - \alpha c],$$

which finishes the proof. □

Corollary 15.3. *Suppose that $X = (X_0, \ldots, X_N)$ is a martingale with respect to $H = (H_0, \ldots, H_N)$ and that $|X_{k+1} - X_k| \leq c$ for some $c > 0$ and $0 \leq k < N$. Then for all $\alpha > 0$ we have*

$$\Pr[|X_k - X_0| \geq \alpha c] \leq 2\mathrm{e}^{-\alpha^2/2k}.$$

Proof. Since X is both a sub- and a supermartingale, we obtain from Theorem 15.1 and corollary 15.2

$$\begin{aligned} \Pr[X_k - X_0 \geq \alpha c] &\leq \mathrm{e}^{-\alpha^2/2k}, \\ \Pr[X_k - X_0 \leq -\alpha c] &\leq \mathrm{e}^{-\alpha^2/2k}. \end{aligned}$$

Combining the two assertions yields the result. □

We say that a function $f\colon \mathbb{R}^n \to \mathbb{R}$ satisfies the *Lipschitz condition* on $D \subseteq \mathbb{R}^n$ if there exists $L > 0$ such that for all $u, v \in D$ we have

$$|f(u) - f(v)| \leq L||u - v||_1,$$

where $||x||_1 = \sum_{i=1}^n |x_i|$ is the 1-norm of x. Note that if f satisfies the Lipschitz condition on D, then f is continuous on D. Moreover, if D is bounded, then f is bounded on D (because $||u - v||_1$ is bounded for $u, v \in D$).

15.2. Assumptions

In the following we will list a fairly long list of assumptions under which we can prove sharp concentration results for values of certain random variables that appear in a discrete time random process.

1. $Q = (Q_0, Q_1, \ldots)$ is a discrete time random process over a probability space Ω with state space S.
2. d, E are integers, and we have measurable functions $y^{(i)}\colon S^+ := \cup_{i \geq 1} S^i \to \mathbb{R}$, $i = 1, \ldots, d$, such that $y^{(i)}(H_t) \leq E$ for all $t \leq E$, where $H_t = (Q_0, \ldots, Q_t)$ is the history of Q up to time t.
3. We set $Y_t^{(i)} := y^{(i)}(H_t)$ for $1 \leq i \leq d$.
4. We further set for $1 \leq i \leq d$
$$\zeta_i := \mathrm{E}\left[Y_0^{(i)}\right].$$
5. $\mathbb{D}$ is an open and bounded subset of $\mathbb{R}^{d+1}$ and there exists $0 \leq \varphi \leq 1$ such that for all $0 \leq t \leq \varphi E$ we have
$$\frac{1}{E}\left(t, Y_t^{(1)}, \ldots, Y_t^{(d)}\right) \in \mathbb{D}.$$
6. We have $(0, \zeta_1, \ldots, \zeta_d) \in \mathbb{D}$.
7. $f_1, \ldots, f_d$ are functions from $\mathbb{R}^{d+1}$ to $\mathbb{R}$ satisfying the Lipschitz condition on $\mathbb{D}$.
8. N is an integer such that for all $1 \leq i \leq d$ we have $\sup_{x \in \mathbb{D}} |f_i(x)| \leq N$ for all i.
9. There exists $c > 0$ such that for all $0 \leq t < \varphi E$ and all $1 \leq i \leq d$ we have
$$|Y_{t+1}^{(i)} - Y_t^{(i)}| \leq c.$$

10. We have for all $1 \leq i \leq d$ and all $0 \leq t < \varphi E$:

$$\mathrm{E}\left[Y_{t+1}^{(i)} - Y_t^{(i)} \;\middle|\; H_t\right] = f_i\left(\frac{t}{E}, \frac{Y_t^{(1)}}{E}, \ldots, \frac{Y_t^{(d)}}{E}\right).$$

11. There is $a \in \mathbb{R}$ such that for all $1 \leq i \leq d$:

$$\Pr\left[\left| Y_0^{(i)} - \mathrm{E}\big[Y_0^{(i)}\big] \right| \geq aE^{5/6}\right] \leq \mathrm{e}^{-\sqrt[3]{E}}.$$

15.3. Sharp Concentration

Under the assumptions laid out in the previous section, we can formulate and prove a general concentration result for the random variables $Y_t^{(i)}$.

Theorem 15.4. *Under the assumptions of Section 15.2. we have:*

(1) *There are unique continuously differentiable functions $z_1, \ldots, z_d$ defined on $[0, \varphi]$ such that*

$$\frac{\mathrm{d}z_i}{\mathrm{d}\tau} = f_i(\tau, z_1, \ldots, z_d),$$

and

$$z_i(0) = \frac{\zeta_i}{E}.$$

(2) *There exists $\eta \in \mathbb{R}$ such that for all $1 \leq i \leq d$ and for all $0 \leq t \leq \varphi E$ we have*

$$\Pr\left[\left| Y_t^{(i)} - Ez_i(t/E) \right| \geq \eta E^{5/6}\right] \leq bE^{1/3}\mathrm{e}^{-\sqrt[3]{E}}.$$

The proof of this theorem will proceed in several steps. We will outline the two results we need for the proof, and prove the theorem using those results in this section. The next two sections are then devoted to the proofs of these three intermediate results.

Proposition 15.5. *Assumptions being as in Theorem 15.4, let w be an integer smaller than φE and larger than 0. Then there exist $\gamma > 0$ and $C_1 > 0$ such that for all $\alpha, \beta > 0$, all $1 \leq i \leq d$, and all $0 \leq t \leq \varphi E - w$ we have*

$$\Pr\left[\left| Y_{t+w} - Y_t - wf\left(\frac{t}{E}, \frac{Y_t^{(1)}}{E}, \ldots, \frac{Y_t^{(d)}}{E}\right) \right| \geq (\gamma + \beta)\frac{w^2}{E} + \alpha C_1\right] \leq 2\mathrm{e}^{-\frac{\alpha^2}{2w}}.$$

Proposition 15.6. *Assumptions being as in Proposition 15.5, let $H(\beta) := (\beta w/E + 1)$ and $K(\beta) = (\gamma + N\beta)w^2/E + \alpha C_1$, and*

$$T_\ell(\beta) := aE^{5/6}H(\beta)^\ell + K(\beta)\frac{H(\beta)^\ell - 1}{H(\beta) - 1},$$

for $\ell = 0, 1, \ldots, \lfloor \varphi E/w \rfloor$. Moreover, suppose that $\alpha^2 = 2wE^{1/3}$. Then there exists some $\beta > 0$ depending on d and the Lipschitz constant L of the f_i (Assumption 7) such that for all such ℓ and all $1 \leq i \leq d$ we have

$$\Pr\left[\left| Y_{\ell w}^{(i)} - Ez_i(\ell w/E) \right| \geq T_\ell(\beta)\right] \leq 2(\ell + 1)\mathrm{e}^{-\frac{\alpha^2}{2w}} = 2(\ell + 1)\mathrm{e}^{-\sqrt[3]{E}}.$$

Proposition 15.7. *Assumptions being as in Proposition 15.6, we have for all* $0 \le t \le \varphi E$ *and all* $1 \le i \le d$*:*

$$\Pr\left[\left| Y_t^{(i)} - E z_i(t/E) \right| \ge cw + T_M(\beta) + Nw\right] \le 2d\varphi \frac{E}{w} \mathrm{e}^{-\frac{\alpha^2}{2w}} = 2d\varphi \frac{E}{w} \mathrm{e}^{-\sqrt[3]{E}},$$

where $M = \lceil \varphi E/w \rceil$.

We can now prove our main concentration result.

Proof. (Of Theorem 15.4.) (1) The assertion on the existence of z_i is a standard result in the theory of ordinary differential equations [**?**].

(2) We use Proposition 15.7 with $w = \lceil E^{2/3} \rceil$ and $\alpha = \sqrt{2E}$. Note that $\alpha^2 = 2E = 2wE^{1/3}$, so that the assumptions of Proposition 15.6 and hence those of Proposition 15.7 are satisfied. We have

$$T_M = aE^{5/6} H^M + K \frac{H^M - 1}{H - 1},$$

where we have suppressed the dependency on β since we are choosing the value for β guaranteed by Proposition 15.6. Note that $H^M \le \mathrm{e}^{\varphi\beta}$. Since $M \le \varphi E/w$, we have

$$\begin{aligned} T_M &\le aE^{5/6}\mathrm{e}^{\varphi\beta} + (\mathrm{e}^{\varphi\beta} - 1)\left[\frac{\gamma + N\beta}{\beta} w + \frac{\alpha E}{w} \frac{C_1}{\beta}\right] \\ &\le aE^{5/6}\mathrm{e}^{\varphi\beta} + (\mathrm{e}^{\varphi\beta} - 1)\left[\frac{\gamma + N\beta}{\beta} E^{2/3} + E^{5/6} \frac{C_1}{\beta}\right], \end{aligned}$$

so that there exists an absolute constant η such that $Cw + T_M + Nw \le \eta E^{5/6}$.Since $\varphi E/w \le \varphi E^{1/3} \le E^{1/3}$, we obtain the statement of the theorem by applying Proposition 15.7. □

We now embark on the proofs of the three propositions above.

15.4. Proof of Proposition 15.5

Let the parameter i be fixed. To ease up the notation, we will suppress the dependency on i, and will write Y_t and f for $Y_t^{(i)}$ and f_i.

Let $0 \le k < w$. By Assumption 10 in Section 15.2., we have

$$\mathrm{E}\left[Y_{t+k+1} - Y_{t+k} \;\middle|\; H_{t+k}\right] = f\left(\frac{t+k}{E}, \frac{Y_{t+k}^{(1)}}{E}, \dots, \frac{Y_{t+k}^{(d)}}{E}\right).$$

By the Lipschitz condition of f on $\mathbb{D}$ (Assumption 7) and the fact that $\mathbb{D}$ includes the points $(t, Y_t^{(1)}, \dots, Y_t^{(d)})/E$ for $t \le \varphi E$ (Assumption 5), we have for $t \le \varphi E - w$

$$\begin{aligned} f\left(\frac{t+k}{E}, \frac{Y_{t+k}^{(1)}}{E}, \dots, \frac{Y_{t+k}^{(d)}}{E}\right) - f\left(\frac{t}{E}, \frac{Y_t^{(1)}}{E}, \dots, \frac{Y_t^{(d)}}{E}\right) &\le L\left(\frac{k}{E} + \sum_{i=1}^{d} \frac{Y_{t+k}^{(i)} - Y_k^{(i)}}{E}\right) \\ &\le \frac{k}{E} L(cd+1), \end{aligned}$$

the latter inequality being a consequence of $Y_{t+k}^{(i)} - Y_t^{(i)} \le |Y_{t+k}^{(i)} - Y_t^{(i)}| \le kc$ (iteration of Assumption 8). Setting $\theta := L(cd+1)$, we obtain

$$\mathrm{E}\left[Y_{t+k+1} - Y_{t+k} \;\middle|\; H_{t+k}\right] \le f\left(\frac{t}{E}, \frac{Y_t^{(1)}}{E}, \dots, \frac{Y_t^{(d)}}{E}\right) + \theta\frac{k}{E}. \tag{15.2}$$

Let $X = (X_0, X_1, \ldots)$ be a sequence of random variables with $X_0 = 0$, and

$$X_{k+1} := X_k + Y_{t+k+1} - Y_{t+k} - f\left(\frac{t}{E}, \frac{Y_t^{(1)}}{E}, \ldots, \frac{Y_t^{(d)}}{E}\right) - \theta\frac{k}{E}.$$

By (15.2) we have $\mathrm{E}[X_{k+1} - X_k | H_{t+k}] \leq 0$. Since X_k is completely determined by H_{t+k}, $\mathrm{E}[X_k|H_{t+k}] = X_k$, and we see that X forms a supermartingale with respect to $(H_t, \ldots, H_{t+k}, \ldots)$. Moreover,

$$|X_{k+1} - X_k| \leq |Y_{t+k+1} - Y_{t+k}| + \left| f\left(\frac{t}{E}, \frac{Y_t^{(1)}}{E}, \ldots, \frac{Y_t^{(d)}}{E}\right) \right| + \left| \theta\frac{k}{E} \right| \leq C_1,$$

where $C_1 := c + N + |\theta|$. (See Assumption 8 for a definition of N.) Since $X_0 = 0$, we have by Theorem 15.1 we have for all $\alpha \geq 0$:

$$\Pr[X_w \geq \alpha C_1] \leq \mathrm{e}^{-\frac{\alpha^2}{2w}}.$$

But

$$X_w = Y_{t+w} - Y_t - wf\left(\frac{t}{E}, \frac{Y_t^{(1)}}{E}, \ldots, \frac{Y_t^{(d)}}{E}\right) - \frac{\theta}{E}\frac{w(w-1)}{2}.$$

Hence, setting $\gamma := \theta/2$, we obtain

$$\Pr\left[Y_{t+w} - Y_t - wf\left(\frac{t}{E}, \frac{Y_t^{(1)}}{E}, \ldots, \frac{Y_t^{(d)}}{E}\right) \geq \alpha C_1 + \gamma\frac{w^2}{E}\right] \leq \mathrm{e}^{-\frac{\alpha^2}{2w}}.$$

To get the other bound we construct a submartingale

$$\mathrm{E}\left[Y_{t+k+1} - Y_{t+k} \;\middle|\; H_{t+k}\right] \geq f\left(\frac{t}{E}, \frac{Y_t^{(1)}}{E}, \ldots, \frac{Y_t^{(d)}}{E}\right) - \theta\frac{k}{E}. \tag{15.3}$$

using the lower bound

$$f\left(\frac{t+k}{E}, \frac{Y_{t+k}^{(1)}}{E}, \ldots, \frac{Y_{t+k}^{(d)}}{E}\right) - f\left(\frac{t}{E}, \frac{Y_t^{(1)}}{E}, \ldots, \frac{Y_t^{(d)}}{E}\right) \quad \geq \quad -\frac{k}{E}L(cd+1),$$

Now, following the same argument as before and using Corollary 15.2, we have

$$\Pr\left[Y_{t+w} - Y_t - wf\left(\frac{t}{E}, \frac{Y_t^{(1)}}{E}, \ldots, \frac{Y_t^{(d)}}{E}\right) \leq -\alpha C_1 - \gamma\frac{w^2}{E}\right] \leq \mathrm{e}^{-\frac{\alpha^2}{2w}}.$$

15.5. Proof of Proposition 15.6

As in the proof of Proposition 15.5 we will suppress the dependency on i, and will write Y_t and f for $Y_t^{(i)}$ and f_i.

For $\ell = 0$, the assertion follows from Assumption 11. Suppose now that the assertion is true for ℓ. Let

$$\begin{aligned}
A_1 &= Y_{(\ell+1)w} - Y_{\ell w} - wf\left(\frac{\ell w}{E}, \frac{Y_{\ell w}^{(1)}}{E}, \ldots, \frac{Y_{\ell w}^{(d)}}{E}\right) \\
A_2 &= Y_{\ell w} - Ez\left(\frac{\ell}{w}\right) \\
A_3 &= Ez\left(\frac{(\ell+1)w}{E}\right) - Ez\left(\frac{\ell w}{E}\right) - wf\left(\frac{\ell w}{E}, \frac{Y_{\ell w}^{(1)}}{E}, \ldots, \frac{Y_{\ell w}^{(d)}}{E}\right).
\end{aligned}$$

Then

$$\left| Y_{(\ell+1)w} - Ez\left(\frac{(\ell+1)w}{E}\right) \right| = |A_1 + A_2 - A_3| \le |A_1| + |A_2| + |A_3|.$$

We will establish upper bounds for A_1, A_2, A_3. First, note that by the induction hypothesis

$$\forall\, i = 1, \ldots, d: \quad \Pr\left[\left| \frac{Y_{\ell w}^{(i)}}{E} - z_i\left(\frac{\ell w}{E}\right) \right| \ge T_\ell \right] \le 2(\ell+1)\mathrm{e}^{-\frac{\alpha^2}{2w}},$$

and hence with probability at least $1 - 2(\ell+1)d\mathrm{e}^{-\alpha^2/2w}$ we have

$$\sum_{i=1}^{d} \left| \frac{Y_{\ell w}^{(i)}}{E} - z_i\left(\frac{\ell w}{E}\right) \right| \le d\frac{T_\ell}{E} \tag{15.4}$$

and

$$|A_2| \le T_\ell. \tag{15.5}$$

In addition, by Proposition 15.5 we have

$$\Pr\left[|A_1| \le \gamma\frac{w^2}{E} + \alpha C_1 \right] \ge 1 - 2\mathrm{e}^{-\frac{\alpha^2}{2w}}. \tag{15.6}$$

Since z is continuously differentiable in $\mathbb{D}$, by the mean value theorem there exists ξ with $\ell w/E \le \xi \le (\ell+1)w/E$ such that $z((\ell+1)w/E) - z(\ell w/E) = wz'(\xi)/E$, where $z'(\tau)$ is the derivative of z with respect to τ. Because $\mathrm{d}z/\mathrm{d}\tau = f(\tau, z_1(\tau), \ldots, z_d(\tau))$, we see that

$$A_3 = w\left(f(\xi, z_1(\xi), \ldots, z_d(\xi)) - f\left(\frac{\ell w}{E}, \frac{Y_{\ell w}^{(1)}}{E}, \ldots, \frac{Y_{\ell w}^{(d)}}{E}\right)\right).$$

Using the Lipschitz condition on f (Assumption 7), we have

$$\begin{aligned} |A_3| &\le wL\left(\left| \xi - \frac{\ell w}{E} \right| + \left| \frac{Y_{\ell w}}{E} - z(\xi) \right| \right) \\ &\le Lw\left(\frac{w}{E} + \sum_{i=1}^{d} \left| \frac{Y_{\ell w}^{(i)}}{E} - z_i\left(\frac{\ell w}{E}\right) \right| + \sum_{i=1}^{d} \left| z_i\left(\frac{\ell w}{E}\right) - z_i(\xi) \right| \right) \end{aligned}$$

By the mean value theorem $|z_i(\ell w/E) - z_i(\xi)| = (\xi - \ell w/E)|f_i(\psi)| \le w|f_i(\psi)|/E$ for some ψ with $\ell w/E \le \psi \le \xi$. Therefore, by the definition of N in Assumption 8,

$$\sum_{i=1}^{d} \left| z_i\left(\frac{\ell w}{E}\right) - z_i(\xi) \right| \le \frac{w}{E}dN.$$

Using (15.4)- (15.6), we see that with probability at least $1 - 2(\ell+2)d\mathrm{e}^{-\alpha^2/2w}$ we have

$$\begin{aligned} |A_1| &\le \gamma\frac{w^2}{E} + \alpha C_1 \\ |A_2| &\le T_\ell \\ |A_3| &\le \le Lw\left(\frac{w}{E} + d\frac{T_\ell}{E} + \frac{w}{E}dN \right). \end{aligned}$$

In total, with the above probability we have

$$\begin{aligned}\left| Y_{(\ell+1)w} - mz\left(\frac{(\ell+1)w}{E}\right) \right| &= |A_1 + A_2 - A_3| \\ &\leq |A_1| + |A_2| + |A_3| \\ &\leq T_\ell\left(\frac{\beta w}{E} + 1\right) + (\gamma + \beta N)\frac{w^2}{E} + \alpha C_1 \\ &= T_{\ell+1},\end{aligned}$$

where $\beta = Ld$.

15.6. Proof of Proposition 15.7

Consider any $0 \leq t \leq \varphi E$, and let $l = \lfloor t/w \rfloor$.

$$\begin{aligned}|Y_t^{(i)} - Ez_i(t/E)| &\leq |Y_t^{(i)} - Y_{lw}^{(i)}| + |Y_{lw}^{(i)} - Ez_i(l)| + |Ez_i(l) - Ez_i(t/E)| \\ &\leq cw + T_M(\beta) + Nw\end{aligned}$$

where, the second inequality holds with probability atleast $1 - 2d\varphi\frac{E}{w}\mathrm{e}^{-\frac{\alpha^2}{2w}}$. This proves the Proposition 15.7