The greedy independent set in a random graph with given degrees

Malwina Luczak[1] [2]

School of Mathematical Sciences
Queen Mary University of London
e-mail: m.luczak@qmul.ac.uk

January 2016
Monash University

[1] Joint work with Graham Brightwell and Svante Janson

[2] Supported by EPSRC Leadership Fellowship, grant reference EP/J004022/2

Random greedy algorithm for independent sets

We study the most naive randomised algorithm for finding a maximal independent set $\mathcal{S}$ in a (multi)graph G.

- We start with $\mathcal{S}$ empty, and we consider the vertices one by one, in a uniformly random order.
- Choose a vertex v uniformly at random from those not already chosen.
- If v has no neighbours that are already in $\mathcal{S}$, put v in $\mathcal{S}$.
- Repeat until all vertices have been chosen.

Motivation

- Easy to analyse.
- May give reasonable bounds on the size of a *largest* independent set.
- Simplicity may be a crucial advantage for applications in distributed computing.
- Connected to various application areas.

Motivation: car parking

In application areas, there is a "continuum" version, typically involving a greedy process for packing d-dimensional unit cubes into $[0, M]^d$ (M large). Unit cubes arrive one at a time, choose a location for their bottom corner uniformly at random in $[0, M-1]^d$, and occupy the space if no already-placed cube overlaps it.

The one-dimensional version is Rényi's car-parking process, see Rényi (1958). See Penrose (2001) for rigorous results on this car-parking process in higher dimensions.

There is also a "discrete" version, taking place on a regular lattice, where an object arrives and selects a location on the lattice uniformly at random, and then inhibits later objects from occupying neighbouring points.

Motivation: applications

- In chemistry and physics, this process is called *random sequential adsorption*. It models some physical processes, such as the deposition of a thin film of liquid onto a crystal.
- In statistics, the greedy process is known as *simple sequential inhibition*; see for instance Diggle (2014).
- There are other potential applications to areas as diverse as linguistics and sociology.
- See Evans (1993); Cadilhe, Araújo and Privman (2007); Bermolen, Jonckheere and Moyal (2013); Finch (2003).

Random graphs with a given degree sequence

- For $n \in \mathbb{N}$ and a sequence $(d_i)_1^n$ of non-negative integers, let $G(n, (d_i)_1^n)$ be a simple graph (no loops or multiple edges) on n vertices chosen uniformly at random from among all graphs with degree sequence $(d_i)_1^n$.
- Must have $\sum_{i=1}^n d_i$ even, at least.

Configuration model

- We let $G^*(n, (d_i)_1^n)$ be the random multigraph with given degree sequence $(d_i)_1^n$ defined by the configuration model: take a set of d_i half-edges for each vertex i and combine the half-edges into pairs by a uniformly random matching of the set of all half-edges.
- In general, this produces a multigraph, so there can be loops and multiple edges.
- Conditioned on the multigraph being a (simple) graph, we obtain $G(n, (d_i)_1^n)$, the uniformly distributed random graph with the given degree sequence.

Configuration model

- Let $n_k = n_k(n) = \#\{i : d_i = k\}$, the number of vertices of degree k in $G(n, (d_i)_1^n)$ (or $G^*(n, (d_i)_1^n)$).
- Then $\sum_k n_k = n$, and we need $\sum_k kn_k$ even.
- E.g., if $n_k = n$ for some k, we get a random k-regular graph.
- We assume that $n_k/n \to p_k$ as $n \to \infty$ for each k, for a probability distribution $(p_k)_0^\infty$.
- We assume that $(p_k)_0^\infty$ has mean $\lambda = \sum_k kp_k \in (0, \infty)$, and that the average vertex degree $\sum_k kn_k/n$ converges to λ.
- We also assume $\sum_k k^2 n_k = O(n)$.

We now consider generating an independent set $\mathcal{S}$ in the random multigraph $G^*(n,(d_i)_1^n)$ via the greedy independent set process.

Let $S_\infty = S_\infty^{(n)}$ be the size of $\mathcal{S}$ at the end of the process; the expected value of S_∞/n is sometimes called the *jamming constant* of the (multi)graph.

Note: a multigraph may have loops, and it might be thought natural to exclude a looped vertex from the independent set, but as a matter of convenience we do not do this, and we allow looped vertices into our independent set. Ultimately, the main interest is in the case of graphs.

- We prove our results for the greedy independent set process on G^*, and, by conditioning on G^* being simple, we deduce that these results also hold for the greedy independent set process on G.
- For this, we use a standard argument that relies on the probability that G^* is simple being bounded away from zero as $n \to \infty$.
- By the main theorem of Janson (2009), this occurs if and only if $\sum_k k^2 n_k(n) = O(n)$. Equivalently, the second moment of the degree distribution of a random vertex is uniformly bounded. (When considering the jamming constant of the multigraph, we can relax this.)

Our main result

Theorem

Assume that $n_k/n \to p_k$ for each k and that $\sum_k kn_k/n \to \lambda = \sum_k kp_k$. Let $S_\infty^{(n)}$ be the size of a random greedy independent set in the random multigraph $G^*(n,(d_i)_1^n)$. Let τ_∞ be the unique value in $(0,\infty]$ such that

$$\lambda \int_0^{\tau_\infty} \frac{e^{-2\sigma}}{\sum_k kp_k e^{-k\sigma}}\, d\sigma = 1.$$

Then

$$\frac{S_\infty^{(n)}}{n} \to \lambda \int_0^{\tau_\infty} e^{-2\sigma} \frac{\sum_k p_k e^{-k\sigma}}{\sum_k kp_k e^{-k\sigma}}\, d\sigma \quad \textit{in probability.}$$

The same holds if $S_\infty^{(n)}$ is the size of a random greedy independent set in the random graph $G(n,(d_i)_1^n)$, if we assume also that $\sum_k k^2 n_k = O(n)$ as $n \to \infty$.

Since $S_\infty^{(n)}/n$ is bounded by 1, it follows that the expectation $\mathbb{E}S_\infty^{(n)}/n$ also tends to the same limit under the hypotheses in the theorem.

Theorem

Under the assumptions of the previous theorem, let $S_\infty^{(n)}(k)$ denote the number of vertices of degree k in the random greedy independent set in the random multigraph $G^*(n,(d_i)_1^n)$. Then, for each $k = 0, 1, \dots,$

$$\frac{S_\infty^{(n)}(k)}{n} \to \lambda \int_0^{\tau_\infty} e^{-2\sigma} \frac{p_k e^{-k\sigma}}{\sum_j j p_j e^{-j\sigma}} \, d\sigma \quad \textit{in probability.}$$

The same holds in the random graph $G(n,(d_i)_1^n)$, if we assume additionally that $\sum_k k^2 n_k = O(n)$ as $n \to \infty$.

We do not know whether the theorems hold also for the simple random graph $G(n, (d_i)_1^n)$ without the additional hypothesis that $\sum_k k^2 n_k = O(n)$.

We leave this as an open problem.

Special case: random regular graphs ($p_d = 1$ for some d)

The jamming constant of a random 2-regular graph, in the limit as $n \to \infty$, is the same as that of a single cycle (or path), again in the limit as the number of vertices tends to infinity.

An equivalent version of the greedy process in this case is for "cars" to arrive sequentially, choose some pair of adjacent vertices on the cycle, and occupy both if they are both currently empty. (Discrete variant of the Rényi parking problem).

The limiting density of occupied vertices was first calculated by Flory in 1939 to be $\frac{1}{2}(1 - e^{-2})$.

Earlier results

- The process, as described above, was analysed by Wormald (1995) for d-regular graphs with $d \geq 3$, and independently by Frieze and Suen for $d = 3$. The independent set has size approximately $\frac{1}{2}\left(1 - \left(\frac{1}{d-1}\right)^{2/(d-2)}\right)n$.
- In the case of an Erdős-Rényi random graph with $p = c/n$, the independent set has size approximately $\frac{\log(c+1)}{c}n$. (McDiarmid 1984)

In the cases mentioned on the previous slide (random d-regular graphs, Erdős-Rényi random graphs with $p = c/n$), we will show how to recover the known results from our theorem by evaluating the integrals.

Suppose $p_0 + p_1 = 1$.

Then we find from

$$\int_0^{\tau_\infty} e^{-2\sigma} \frac{\sum_k k p_k}{\sum_k k p_k e^{-k\sigma}} \, d\sigma = 1$$

that $\tau_\infty = \infty$, and the formula for the limit of S_∞/n evaluates to $p_0 + \frac{1}{2}p_1$, as expected.

If a multigraph has n_0 isolated vertices, $n_1 = n - n_0 - o(n)$ vertices of degree 1, and $\frac{1}{2}n_1 + o(n)$ edges in total, then any maximal independent set has size $n_0 + \frac{1}{2}n_1 + o(n)$.

Suppose there is some $\ell \geq 2$ such that $p_\ell > 0$.

Then

$$\frac{\lambda e^{-2\sigma}}{\sum_k k p_k e^{-k\sigma}} = \frac{\sum_k k p_k e^{-2\sigma}}{\sum_k k p_k e^{-k\sigma}} > e^{-\sigma}$$

for all $\sigma > 0$, and hence

$$\int_0^\infty \frac{\lambda e^{-2\sigma}}{\sum_k k p_k e^{-k\sigma}}\, d\sigma > \int_0^\infty e^{-\sigma}\, d\sigma = 1.$$

As the integrand is positive and bounded on finite intervals, this implies that there is a unique finite value τ_∞ satisfying the equation

$$\lambda \int_0^{\tau_\infty} \frac{e^{-2\sigma}}{\sum_k k p_k e^{-k\sigma}}\, d\sigma = 1.$$

Random d-regular graphs

Suppose $p_d = 1$ for some $d \geq 2$.

Evaluating the integral in the formula and setting it equal to 1:

$$1 = \int_{\sigma=0}^{\tau_\infty} e^{(d-2)\sigma}\, d\sigma = \frac{1}{d-2}(e^{(d-2)\tau_\infty} - 1),$$

for $d \geq 3$, and so $\tau_\infty = \frac{\log(d-1)}{d-2}$.
For $d = 2$ we obtain $\tau_\infty = 1$.

Random d-regular graphs

Our formula for the jamming constant reduces to

$$\int_0^{\tau_\infty} e^{-2\sigma}\, d\sigma = \frac{1}{2}(1 - e^{-2\tau_\infty}) = \frac{1}{2}\left(1 - \frac{1}{(d-1)^{2/(d-2)}}\right),$$

for $d \geq 3$, and $\frac{1}{2}(1 - e^{-2})$ for $d = 2$.

This indeed agrees with Wormald's formula for $d = 3$, as well as Flory's formula for $d = 2$.

$G(n,p)$ for $p = c/n$

Vertex degrees are random, but by conditioning on the vertex degrees we can apply the results above, with the asymptotic Poisson degree distribution $p_k = \frac{c^k e^{-c}}{k!}$, for each $k \geq 0$.

Here $\lambda = \sum_k k p_k = c$, $\sum_k p_k e^{-k\sigma} = e^{-c} e^{ce^{-\sigma}}$,
$\sum_k k p_k e^{-k\sigma} = c e^{-c} e^{-\sigma} e^{ce^{-\sigma}}$.

We find that

$$1 = e^c \int_0^{\tau_\infty} e^{-\sigma} e^{-ce^{-\sigma}} \, d\sigma = \frac{e^c}{c} \left[e^{-ce^{-\tau_\infty}} - e^{-c} \right],$$

and rearranging gives

$$e^{-\tau_\infty} = 1 - \frac{\log(c+1)}{c}.$$

Then

$$\frac{S_\infty}{n} \to \int_0^{\tau_\infty} e^{-\sigma}\, d\sigma = 1 - e^{-\tau_\infty} = \frac{\log(c+1)}{c},$$

which agrees with the known value, which can be found from first principles by a short calculation; see McDiarmid (1984).

Furthermore

$$\frac{S_\infty(k)}{n} \to \int_0^{\tau_\infty} \frac{c^k}{k!} e^{-(k+1)\sigma} e^{-ce^{-\sigma}}\, d\sigma = \frac{1}{c} \int_{c-\log(c+1)}^{c} \frac{x^k}{k!} e^{-x}\, dx.$$

Hence the asymptotic degree distribution in the random greedy independent set can be described as a mixture of $Po(\mu)$, with parameter μ uniformly distributed in $[c - \log(c+1), c]$.

Earlier general result

A recent preprint of Bermolen, Jonckheere and Moyal contains a study of the general case, under a 6th moment condition.

They prove that their process is approximated by the unique solution of an infinite-dimensional differential equation. The paper gives no explicit form for the solution (except in the case of a random 2-regular graph, and for the Poisson distribution; in the latter case, the authors substitute a Poisson distribution for the number of empty vertices of degree k and show that this satisfies their equations), and the differential equation itself involves the second moment of the degree sequence.

The authors evaluate the solution numerically in several explicit instances, and extract the jamming constant.

Our random process

To study a random process on a random graph with given degrees, it often pays to think of the random process as running in parallel with the generation of the random graph.

So here, we analyse a continuous-time Markovian process, which generates a random multigraph on a fixed set $\mathcal{V} = \{1, \ldots, n\}$ of n vertices, vertex i with degree d_i, along with an independent set $\mathcal{S}$ in the multigraph.

At each time $t \geq 0$, the vertex set $\mathcal{V}$ is partitioned into three classes:

(a) a set $\mathcal{S}_t$ of vertices already placed into the independent set, with all half-edges out of $\mathcal{S}_t$ paired,

(b) a set $\mathcal{B}_t$ of *blocked* vertices, where at least one half-edge has been paired with a half-edge from $\mathcal{S}_t$,

(c) a set $\mathcal{E}_t$ of *empty* vertices, from which no half-edge has yet been paired.

At all times, the only paired edges are those with at least one endpoint in $\mathcal{S}_t$. For $j = 1, 2, \ldots,$ we set $\mathcal{E}_t(j)$ to be the set of vertices in $\mathcal{E}_t$ of degree j.

Initially all vertices are empty, i.e., $\mathcal{E}_0 = \mathcal{V}$.

Each vertex v has an independent exponential clock, with rate 1. When the clock of vertex $v \in \mathcal{E}_t$ goes off, the vertex is placed into the independent set and all its half-edges are paired. This results in the following changes:

(a) v is moved from $\mathcal{E}_t$ to $\mathcal{S}_t$,

(b) each half-edge incident to v is paired in turn with some other uniformly randomly chosen currently unpaired half-edge,

(c) all the vertices in $\mathcal{E}_t$ where some half-edge has been paired with a half-edge from v are moved to $\mathcal{B}_t$.

So we only generate neighbours of a vertex as it is chosen.

- Some half-edges from v may be paired with half-edges from $\mathcal{B}_t$, or indeed with other half-edges from v: no change in the status of a vertex results from such pairings.
- The clocks of vertices in $\mathcal{B}_t$ are ignored.
- The process terminates when $\mathcal{E}_t$ is empty. At this point, there may still be some unpaired half-edges attached to blocked vertices: these may be paired off uniformly at random to complete the creation of the random multigraph.

The pairing generated is a uniform random pairing of all the half-edges.

The independent set generated in the random multigraph can also be described as follows: vertices have clocks that go off in a random order, and when the clock at any vertex goes off, it is placed in the independent set if possible.

Thus our process does generate a random greedy independent set in the random multigraph.

Key variables

The variables we track in our analysis of the process are:

- $E_t(j) = |\mathcal{E}_t(j)|$, the number of empty vertices of degree j at time t, for each $j \geq 0$,
- the total number U_t of unpaired half-edges,
- the number $S_t = |\mathcal{S}_t|$ of vertices that have so far been placed in the independent set.

Vector $(U_t, E_t(0), E_t(1), \ldots, S_t)$ is Markovian.

- At each time t, there are $E_t(j)$ clocks associated with empty vertices of degree j.
- When the clock at one such vertex v goes off, its j half-edges are paired uniformly at random within the pool of U_t available half-edges, so U_t goes down by exactly $2j - 2\ell$, where ℓ is the number of loops generated at v, which has a distribution that can be derived from $|U_t|$ and the degree of v.
- The distribution of the numbers of vertices of each $\mathcal{E}_t(k)$ that are paired with one of the half-edges out of v is a straightforward function of the vector given. Meanwhile S_t increases by one each time a clock at an empty vertex goes off.

Drifts: S_t

Next, we calculate the drifts in each of our variables, as a function of the current state.

S_t has drift $\sum_{k=0}^{\infty} E_t(k)$, since S_t increases by 1 each time the clock at an empty vertex goes off, and they all go off with rate 1.

Drifts: U_t

U_t has drift $-\sum_k kE_t(k)\Big(2 - \frac{k-1}{U_t-1}\Big)$.

When the clock at a vertex of degree k goes off, then the number of free half-edges decreases by $k + (k - 2L) = 2k - 2L$, where L is the number of loops created at the vertex.
We have, conditionally on U_t, $\mathbb{E}L = \binom{k}{2}\frac{1}{U_t-1}$, and thus the expected number of removed half-edges is

$$2k - 2\mathbb{E}L = k\Big(2 - \frac{k-1}{U_t - 1}\Big).$$

Now multiply by $E_t(k)$ and sum over k.

Drifts: $E_t(k)$

$E_t(k)$ has drift

$$-E_t(k) - \sum_j p_{jk}(U_t)E_t(j)\Big(E_t(k) - \delta_{jk}\Big),$$

where $p_{jk}(u)$ is the probability that vertices v and w of degrees j and k, respectively, are connected by at least one edge in a configuration model with u half-edges.

When the clock at an empty vertex of degree j goes off, this reduces $E_t(j)$ by 1.

Also, each empty vertex of degree k is joined to it with probability $p_{jk}(U_t)$ (since we may ignore the half-edges already paired).

Hence the expected total decrease of $E_t(k)$ when a vertex in $\mathcal{E}_t(j)$ goes off is $p_{jk}(U_t)E_t(k)$ when $j \neq k$ and $1 + p_{jk}(U_t)(E_t(k) - 1)$ when $j = k$.

Now multiply by $E_t(j)$ and sum over j.

We have, with the sums really finite and extending only to $j \wedge k$,

$$p_{jk}(u) = \sum_{m=1}^{\infty} \frac{(-1)^{m-1}}{m!} \frac{(j)_m (k)_m}{2^m((u-1)/2)_m}.$$

Hence one can show that

$$\frac{jk}{u-1} \geq p_{jk}(u) \geq \frac{jk}{u-1} - \frac{j(j-1)k(k-1)}{2(u-1)(u-3)}.$$

Limiting differential equations

From the above, the scaled random variables $U_t/n, E_t(k)/n, S_t/n$ should converge to deterministic functions $u_t, e_t(k), s_t$ solving the following differential equations.

- $$\frac{du_t}{dt} = -2\sum_k ke_t(k).$$
- $$\frac{de_t(j)}{dt} = -e_t(j) - \frac{je_t(j)\sum_k ke_t(k)}{u_t} \quad \text{for each } j.$$
- $$\frac{ds_t}{dt} = \sum_k e_t(k).$$

Change of variable

Setting $h_t(j) = e^t e_t(j)$ for each j results in:

$$\frac{dh_t(j)}{dt} = -\frac{jh_t(j)\sum_k ke_t(k)}{u_t} \quad \text{for each } j.$$

Now we make a time-change $t \to \tau$ where $\dfrac{d\tau}{dt} = \dfrac{\sum_k ke_t(k)}{u_t}$.

Our equations then become:

- $\dfrac{du_\tau}{d\tau} = -2u_\tau;$
- $\dfrac{dh_\tau(j)}{d\tau} = -jh_\tau(j) \quad \text{for each } j.$

Solution of equations

These equations are now decoupled, and can be solved:

$$u_\tau = \sum_k k n_k e^{-2\tau}; \quad h_\tau(j) = n_j e^{-j\tau}.$$

Substituting into the time-change equation gives:

$$\frac{d\tau}{dt} = e^{-t} \frac{\sum_k k n_k e^{-k\tau}}{\sum_k k n_k e^{-2\tau}}.$$

The equation separates, so:

$$1 - e^{-t} = \int_0^t e^{-s}\, ds = \int_0^{\tau_t} e^{-2\sigma} \frac{\sum_k k n_k}{\sum_k k n_k e^{-k\sigma}}\, d\sigma.$$

Since the integrand is positive, this determines τ_t uniquely for every $t \in [0, \infty)$.

Thus u_t, $h_t(j)$ and $e_t(j)$ are determined uniquely.

This is at least assuming $\sum_k k e_t(k) > 0$ and $u_t > 0$, which we show any subsequential limit of our process must satisfy. If these conditions are met, then the time change is well defined.

Expression for s_∞

The original process runs for all time, so the time-changed process runs up to some time τ_∞, where

$$1 = \int_0^{\tau_\infty} e^{-2\sigma} \frac{\sum_k k n_k}{\sum_k k n_k e^{-k\sigma}} \, d\sigma.$$

This time τ_∞ is finite except in a trivial case.

Now we have, working in the time-changed process,

$$\begin{aligned} s_\infty &= \int_0^\infty \sum_k e_t(k) \, dt = \int_0^{\tau_\infty} \frac{u_\tau \sum_k e_\tau(k)}{\sum_k k e_\tau(k)} \, d\tau \\ &= \frac{\sum_k k n_k}{n} \int_0^{\tau_\infty} e^{-2\tau} \frac{\sum_k n_k e^{-k\tau}}{\sum_k k n_k e^{-k\tau}} \, d\tau. \end{aligned}$$

Proof - with martingales

We can write

$$U_t = U_0 - \int_0^t \sum_k kE_s(k)\Big(2 - \frac{k-1}{U_s - 1}\Big)\, ds + M_t,$$

and, for each $k \in \mathbb{Z}^+$,

$$\begin{aligned} E_t(k) &= E_0(k) - \int_0^t E_s(k)\, ds \\ &\quad - \int_0^t \sum_j p_{jk}(U_s)E_s(j)\Big(E_s(k) - \delta_{jk}\Big)\, ds + M_t(k), \end{aligned}$$

where M_t and each $M_t(k)$ is a martingale.

The martingale M_t has quadratic variation $[M]_t$ given by

$$\begin{aligned}[M]_t &= \sum_{0\leq s\leq t}(\Delta M_s)^2 = \sum_{0\leq s\leq t}(\Delta U_s)^2 \\ &\leq \sum_{s\geq 0}(\Delta U_s)^2 \leq \sum_j (2j)^2 n_j = o(n^2),\end{aligned}$$

since $\sum_k kn_k/n$ is uniformly summable.
(In fact, $[M]_t = O(n)$ if the second moment of the degree distribution is uniformly bounded.)

Likewise, for each k,

$$[M(k)]_t \leq \sum_{s\geq 0} (\Delta E_s(k))^2 \leq \sum_j (j+1)^2 n_j = o(n^2),$$

with the bound being $O(n)$ if the second moment of the degree distribution is uniformly bounded.

Doob's inequality then gives $\sup_{t\geq 0} |M_t| = o_P(n)$ and, for each k, $\sup_{t\geq 0} |M_t(k)| = o_P(n)$.

The integrand in the definition of M_t is at most $4\lambda n$. This implies that $(U_t - M_t)/n$, $n \geq 1$, is what is known as a "uniformly Lipschitz family", and it is also uniformly bounded on each finite interval $[0, t_0]$. For each k, the same is true for $(E_t(k) - M_t(k))/n$, $n \geq 1$.

Using the Arzela-Ascoli Theorem, it can then be shown that there is a subsequence along which we have

$$\frac{U_t - M_t}{n} \to u_t \qquad \text{and each} \qquad \frac{E_t(k) - M_t(k)}{n} \to e_t(k)$$

in distribution, in $C[0, \infty)$, for some random continuous functions u_t and $e_t(k)$.

By the Skorokhod coupling lemma, we may assume that these limits hold a.s. (in $C[0,\infty)$, i.e., uniformly on compact sets), and also that $\sup_{t\geq 0}|M_t|/n \xrightarrow{\text{a.s.}} 0$ and $\sup_{t\geq 0}|M_t(k)|/n \xrightarrow{\text{a.s.}} 0$.

Hence, along the subsequence, there are (random) continuous functions u_t and $e_t(k)$, $k = 0, 1, 2, \ldots$, such that

$$\frac{U_t}{n} \to u_t \qquad \text{and each} \qquad \frac{E_t(k)}{n} \to e_t(k)$$

a.s., uniformly on compact sets.

It is then possible to show that these limiting functions u_t, $e_t(k)$ $(k = 0, 1, \dots)$, are differentiable, and satisfy the differential equations we saw before, as well as the conditions $\sum_k ke_t(k) \geq 0$ and $u_t > 0$.

Hence they are equal to the unique solutions we found earlier.

This means that u_t and $e_t(k)$, which a priori are random, in fact are deterministic.

Moreover, since the limits u_t and $e_t(k)$ are continuous, we have convergence in the Skorokhod space $D[0, \infty)$.

Consequently, each subsequence of $(U_t/n, E_t(k)/n : k = 0, 1, \ldots)$ has a subsequence which converges in distribution in the Skorokhod topology, and each convergent subsequence converges to the same $(u_t, e_t(k) : k = 0, 1, \ldots)$.

This implies that the whole sequence $(U_t/n, E_t(k)/n : k = 0, 1, \ldots)$ must in fact converge to $(u_t, e_t(k) : k = 0, 1, \ldots)$ in distribution in the Skorokhod topology. Since the limit is deterministic, the convergence holds in probability.

We complete the proof by showing that, furthermore, $S_t/n \to s_t$ in probability.