The Embedding Capacity of Information Flows Under Renewal Traffic

Stefano Marano, Vincenzo Matta, Ting He, *Member, IEEE*, and Lang Tong, *Fellow, IEEE*

***Abstract*—Given two independent point processes and a certain rule for matching points between them, what is the fraction of matched points over infinitely long streams? In many application contexts, e.g., secure networking, a meaningful matching rule is that of a maximum causal delay, and the problem is related to embedding a flow of packets in cover traffic such that no timing analysis can detect it. We study the best undetectable embedding policy and the corresponding maximum flow rate—that we call the embedding capacity—under the assumption that the cover traffic can be modeled as an arbitrary renewal process. We find that computing the embedding capacity requires the inversion of a very structured linear system that, for a broad range of renewal models encountered in practice, admits a fully analytical expression in terms of the renewal function of the processes. This result enables us to explore the properties of the embedding capacity, obtaining closed-form solutions for selected distribution families and a suite of sufficient conditions on the capacity ordering. We test our solution on real network traces, which shows a remarkable match for tight delay constraints. A gap between the predicted and the actual embedding capacities appears for looser constraints, and further investigation reveals that it is caused by inaccuracy of the renewal traffic model rather than of the solution itself.**

***Index Terms*—Embedding capacity, information flow, intrusion detection and security, point processes and inference, Riemann–Hilbert problem.**

I. Introduction

CONSIDER the pair of timing sequences represented by the point processes $\mathcal{S}$ and $\mathcal{T}$ in Fig. 1, where points are matched according to some prescribed rule. What is the maximum achievable fraction of matched points (embedding capacity) given the two processes and the matching rule? How do statistical properties of the point processes affect the maximum fraction of matching?

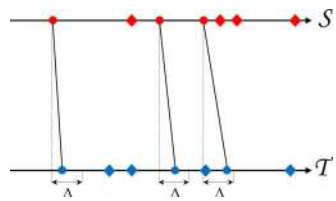

Fig. 1. Notional sketch of the addressed problem, with arrival epochs of the processes $\mathcal{S}$ and $\mathcal{T}$ matched according to a delay constraint Δ. Matched points are marked by circles, and unmatched by diamonds.

The point processes $\mathcal{S}$ and $\mathcal{T}$ represent sequences of events generated by two different entities that are externally observable. The matching rule describes the expected relationship between the events if an event at one entity is caused by another event at the other entity, and this nexus of cause and effect can be inferred by the fraction of events satisfying the matching rule over the entire sequences. For instance, the aforementioned problem arises in intelligence applications aimed at tracing relationships among individuals in social networks (see [1] for a recent survey), where the point processes model specific observable activity patterns (e.g., tags, shared links, favorite games, preferences or attitudes, and so on) of network members; or in the problem of discovering neuron connections by measurements of firing sequences [2], [3], where the point processes model the spike trains produced by neural activities.

Closer to the communication area is the network security application concerning the detection of clandestine information flows, see, e.g., [4], where nodes relaying packets for each other try to hide the fact of relaying, which is often an indicator of network attacks (such as stepping-stone attacks [5], [6]). In the last decades, a prominent role against clandestine communication has been played by traffic analysis aimed at discovering source–relay pairs by analyzing timing information in the network traffic [7], [8]. In this context, the two processes represent the sequences of time epochs (traffic patterns) at which successive packets leave two nodes of the network and, for security requirements, packets are encrypted so that they do not reveal special characteristics. Still, the act of transmission itself cannot be kept secret, and timing analysis can be performed.

Given that nodes are unable to hide the act of transmission, they must hide the information flow packets into their normal transmission scheduling, which provide *cover traffic* for the desired flow. The nodes can mask the timing relationships by properly delaying the transmission of information packets and/or multiplexing information packets with dummy packets

Manuscript received March 10, 2011; revised April 02, 2012; accepted September 24, 2012. Date of publication November 15, 2012; date of current version February 12, 2013. L. Tong was supported in part by the Army Research Office MURI Program under Award W911NF-08-1-0238 and in part by the National Science Foundation under Award CCF 1018115. This paper was presented in part at the Annual Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, in 2010, and in part at the 2011 IEEE Information Theory Workshop.

S. Marano and V. Matta are with the Department of Information and Electrical Engineering, and Applied Mathematics, University of Salerno, I-84084 Fisciano (SA), Italy (e-mail: marano@unisa.it; vmatta@unisa.it).

T. He is with IBM T. J. Watson Research Center, Yorktown, NY 10598 USA (e-mail: the@us.ibm.com).

L. Tong is with the Department of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: ltong@ece.cornell.edu).

Communicated by M. Franceschetti, Associate Editor for Communication Networks.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2012.2227672

or packets from other flows. With a sufficient amount of perturbation, an information flow can be disguised as traffic of arbitrary patterns.

However, network protocols must be faced with physical constraints that pose some limitations on the admissible scheduling patterns. A sensible constraint is that of causal bounded delay [5], meaning that the relaying of (information flow) packets must occur within a maximum allowed latency (see Fig. 1). Thus, every transmission schedule (or cover traffic) has a certain capacity of being utilized to transmit information flows covertly. The matching capability of a particular schedule takes the operational meaning of an *embedding capacity*, that is, the maximum fraction of information packets that can be embedded in the cover traffic following this schedule, leaving no chances of discovering the presence of the flow itself. The embedding capacity establishes a fundamental limit on the sensitivity of information flow detection, i.e., a smart attacker can send information flows with a normalized rate up to the embedding capacity without being detected by any flow detector [9], [10].

In this respect, the intruder might embed information packets into traffic streams corresponding to noninteracting nodes, i.e., *statistically independent* point processes. Thus, while general processes containing a common information flow will *not* be statistically independent, this specific embedding procedure would guarantee that an information flow can be embedded into realizations from strictly independent processes, and the embedding capacity quantifies the maximum delivering rate.

A closely related application is in the area of anonymous networking [7], [11], [12]. As the dual of information flow detection, the goal of an anonymity-supporting relay is to maximize the rate of information flow without revealing the act of relaying. Again, although encryption can hide the correlation in traffic content, hiding the correlation in traffic timing generally incurs loss in efficiency (dropped packets, dummy packets, see, e.g., [13]). Here, the embedding capacity characterizes the maximum efficiency of the relay in hiding the correlation between the timing of the incoming and the outgoing traffic.

The main theme of this paper is that of providing analytical tools for computing the embedding capacity of two independent and identically distributed (i.i.d.) processes, when the coupling rule is formulated in terms of a causal delay constraint, and the network traffic is modeled as an arbitrary renewal process.

It is important to stress that the network traffic characteristics depend upon the specific protocol used. Since protocols are part of the network design, a relevant question is what kind of traffic pattern offers better anonymity. In this connection, availability of simple formulas for the embedding capacity allows comparison of different traffic models in terms of their "vocation" to anonymity, and is key to the design of anonymity-compliant protocols.

A. Summary of Results

The embedding capacity for a Poisson process under causal delay constraint is known, see [4]. The Poisson assumption, however, rarely fits real traffic and, to date, analytical formulas for arbitrary renewal traffic are still missing. In the following, this gap is filled.

Indeed, we find that the embedding capacity for renewal cover traffic is related to the invariant distribution of a suitable Markov chain. First, we prove the existence of such distribution, so that capacity evaluation requires the solution of an integral equation. We attack this problem by exploiting the powerful tools offered by the Riemann–Hilbert theory, which allows us to derive the following approximation for the embedding capacity:

$$C^* \approx \frac{\lambda\Delta}{1+\dfrac{2}{\lambda\Delta}\displaystyle\int_0^{\lambda\Delta} m(t)dt}$$

where λ is the rate of the processes, Δ is the delay constraint, and $m(t)$ is the renewal function of the (scaled to unit rate) underlying process. The accuracy of this formula is excellent for a very broad range of renewal processes of interest for the applications, see Section VI-A. We also show how C^* can be computed to any degree of approximation by inverting a very structured linear system.

It is important to stress that the previous formula for C^* depends only on the renewal function which is the key quantity in renewal theory and, as such, is well studied and understood. Therefore, our formulas for C^* provide a simple way to compute the maximum fraction of information flow packets that can be anonymously embedded. In many cases of practical interest, the integral involved can be also evaluated explicitly, from which physical insights can be gained even easier.

The aforementioned expression is then used to relate the characteristics of the renewals to the embedding performance. This highlights how the maximum amount of embedded flow depends upon the cover traffic parameters, and gives insight to compare and design different anonymous-oriented traffic protocols. In particular, when the rate and/or the maximum delay is large (loose delay bound), the distribution of cover traffic affects the embedding capacity only through the dispersion index γ, namely, as $\lambda\Delta \to \infty$, the minimum fraction of unmatched points decays as $\gamma/(\lambda\Delta)$. Stochastic variability is instead the key (for any $\lambda\Delta$) to compare different interarrival distributions: less variable interarrivals yield a larger embedding capacity, and in particular it is seen that the Poisson case represents a waterfall between the classes of new better than used in expectation (NBUE) and new worse than used in expectation (NWUE) cover traffic (see Corollary 2).

B. Related Work and Organization

The roots of packet embedding into cover traffic can be traced back to the early 1980s. The problem of avoiding traffic analysis using special relay policies was first considered in [11], with the adoption of the so-called MIX relays that perform multiplexing, scrambling, and encryption of the incoming traffic in order to eliminate the correlation with the outgoing traffic. Since then, several studies have been made in order to improve relay performances, see, e.g., [13] and [14]. More recently, it has been shown how statistically independent transmission schedules can achieve perfectly anonymous relaying, with emphasis on the maximization of the carried information capacity [12].

Also related to our problem is the network security issue referred to as stepping-stone attack [5], [6], in which an adversary

launches an attack through a sequence of compromised servers, and one would like to trace the sequence to the origin of the attack. For wireless networks, an ad hoc network may be subject to the worm-hole attack [15], where the attacker hijacks the packets of a node and channels them through a covert tunnel. In such scenarios, the maximum information rate sustainable by the attackers is related to the embedding capacity of the node traffic patterns.

From an information-theoretic perspective, the problem of secure communications, in terms of maximizing the reliable rate to a legitimate receiver with secrecy constraints with respect to an eavesdropper, has been extensively studied, since the pioneering works [16]–[18], up to recent extensions, including multiaccess [19], fading [20], feedback [21], and broadcast [22] channels, among many others. We stress that the specific scenario of interest for this paper is instead secure networking with focus on anonymous relaying of information, according to the model proposed in [4] and [12].

Formal studies of the embedding properties of renewals have been carried out in [4] and [12], with extensions to distributed detection with communication constraints [23], [24]. In [4], the problem is settled up from the traffic analyzer's perspective, where the role of the embedding capacity is replaced by that of undetectable flow, and a closed formula for the capacity under the Poisson regime is found. In many applications (inside the communication area as well as outside that), general renewal traffic models are far from being approximated as Poisson, such that several extensions of the aforementioned studies in this direction have been proposed, see [25] and [26]. However, a tractable analytical formula for the embedding capacity under arbitrary renewal traffic is still missing.

The remainder of this paper is organized as follows: Section II formalizes the problem; the main results of the paper are presented in Section III, and Section IV is devoted to the main mathematical derivations. Section V addresses the problem of classification and ordering of renewal processes in terms of their embedding capacity, while Section VI concerns the application of the main theoretical findings to simulated data and experiments on real network traces. Conclusions follow in Section VII.

II. Problem Formulation

Capital letters denote random variables, and the corresponding lowercase the associate realizations, while $\mathbb{P}$ and $\mathbb{E}$ denote probability and expectation operators, respectively.

A point process on the positive real axis $\mathbb{R}_+$ is a collection of nonnegative random variables $\mathcal{S} = \{S_i, i = 1, 2, \ldots\}$ such that, for $i = 1, 2, \ldots$, $S_i < S_{i+1}$ almost surely (a.s.), and $\lim_{i\to\infty} S_i = \infty$ a.s.

Consider two point processes $\mathcal{S} = \{S_i, i = 1, 2, \ldots\}$ and $\mathcal{T} = \{T_i, i = 1, 2, \ldots\}$ defined over $\mathbb{R}_+$. Points that are matched over the two processes form an *information flow* in the sense that one point in a matched pair can be thought of as a relayed copy of the other. We are interested in delay-sensitive directional flows, for which matched points obey a causal bounded delay constraint as follows [4].

Definition 1 (Information Flow): Point processes $\mathcal{W}$ and $\mathcal{Z}$ form a Δ-bounded-delay information flow, in the direction $\mathcal{W} \to \mathcal{Z}$, if for every realization, there is a one-to-one mapping $\mathcal{M} : \{w_i\} \to \{z_i\}$ between sets $\{w_i\}$ and $\{z_i\}$, satisfying the causal bounded delay constraint $0 \le \mathcal{M}(w_i) - w_i \le \Delta, \forall i$. $\diamond$

Here, $\Delta > 0$ is a known constant representing the maximum tolerable delay during relaying.

Given point processes $\mathcal{S}$ and $\mathcal{T}$, an information flow can be selected by finding, for each realization of the processes, subsequences that admit a valid one-to-one mapping. This is controlled by an embedding policy.

Definition 2 (Embedding Policy): An embedding policy ϵ selects two (possibly path dependent and random) subsequences $\{i_k, k = 1, 2, \ldots\}$ and $\{j_k, k = 1, 2, \ldots\}$ such that the thinned point processes $\mathcal{W}^\epsilon = \{S_{i_k}, k = 1, 2, \ldots\}$ and $\mathcal{Z}^\epsilon = \{T_{j_k}, k = 1, 2, \ldots\}$ form a Δ-bounded-delay information flow in the direction $\mathcal{W}^\epsilon \to \mathcal{Z}^\epsilon$. $\diamond$

The term "embedding" is due to the fact that to an outsider who cannot observe the selection, it is not known which points belong to an information flow or even if there is a flow, and thus the flow is embedded in the overall processes $(\mathcal{S}, \mathcal{T})$. For the same reason, $(\mathcal{S}, \mathcal{T})$ is called *cover traffic*.

Let $\mathcal{E}$ be the set of the embedding policies. Given $\epsilon \in \mathcal{E}$, the cover traffic $(\mathcal{S}, \mathcal{T})$ is decomposed into

$$\mathcal{S} = \mathcal{W}^\epsilon \oplus \mathcal{U}^\epsilon, \quad \mathcal{T} = \mathcal{Z}^\epsilon \oplus \mathcal{V}^\epsilon$$

where $(\mathcal{W}^\epsilon, \mathcal{Z}^\epsilon)$ forms an information flow (in the direction $\mathcal{W}^\epsilon \to \mathcal{Z}^\epsilon$). Here, $\oplus$ is the superposition operator for point processes: $\mathcal{S} = \mathcal{W} \oplus \mathcal{U}$ means that s_i is the ith element in the sequence made of all elements of $\mathcal{W}$ and $\mathcal{U}$, arranged in increasing order, namely, $\{s_i\} = \{w_i\} \cup \{u_i\}$ with $s_1 \le s_2 \le \cdots$.

Definition 3 (Efficiency): Given cover traffic $(\mathcal{S}, \mathcal{T})$, the efficiency of an embedding policy $\epsilon \in \mathcal{E}$ is measured by

$$\eta(\epsilon) := \liminf_{t\to\infty} \frac{N_{\mathcal{W}^\epsilon}(t) + N_{\mathcal{Z}^\epsilon}(t)}{N_{\mathcal{S}}(t) + N_{\mathcal{T}}(t)}$$

where $N_{\mathcal{W}^\epsilon}(t)$ and $N_{\mathcal{Z}^\epsilon}(t)$ are the counting processes (up to t) for the embedded information flow, so are $N_{\mathcal{S}}(t)$ and $N_{\mathcal{T}}(t)$ for the cover traffic. $\diamond$

In words, the efficiency is the asymptotic fraction of matched points in the cover traffic, and we are clearly interested in the highest efficiency, which we call the *embedding capacity*.

Definition 4 (Embedding Capacity):

$$C^* := \sup\{r \in [0, 1] : \exists \epsilon \in \mathcal{E} \text{ with } \eta(\epsilon) \ge r \quad \text{a.s.}\}. \qquad \diamond$$

The embedding capacity C^* is a function of the cover traffic and of the flow constraint (i.e., Δ), omitted in the notation for simplicity. We shall focus on the case that the cover traffic processes $\mathcal{S}$ and $\mathcal{T}$ are i.i.d. renewal processes, with interarrivals modeled as absolutely continuous (with respect to the usual Lebesgue measure) random variables $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$, respectively,[1] whose common probability density function (PDF) is

[1]For simplicity, we consider nondelayed renewal processes, even though most results obviously extend to delayed renewals.

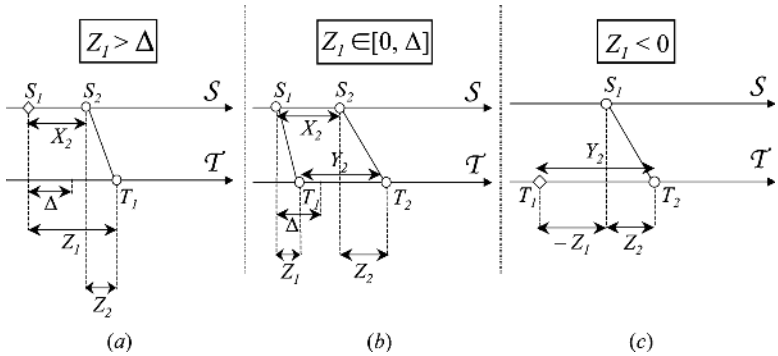

Fig. 2. Three situations arising from applying the BGM procedure to point processes $\mathcal{S}$ and $\mathcal{T}$. Chaff points are denoted by "$\diamond$." (Left) Point at S_1 is unmatched, and it is a chaff point in the process $\mathcal{S}$. (Center) All points are matched (no chaff). (Right) Chaff point is present in the process $\mathcal{T}$.

denoted by $f(t)$, and whose cumulative distribution function (CDF) by $F(t)$. Throughout the paper, it is assumed that the rate of the processes, denoted by λ, is finite and nonzero, i.e., $0 < \lambda = 1/\mathbb{E}[X] = 1/\mathbb{E}[Y] < \infty$; when the second moment is finite, we define the dispersion index as

$$\gamma := \lambda^2 \,\mathrm{VAR}[X] = \lambda^2 \,\mathrm{VAR}[Y] < \infty. \tag{1}$$

III. Characterization of the Embedding Capacity

In order to characterize the embedding capacity, we proceed by first finding an *optimal* embedding algorithm, given any realization of the two point processes $\mathcal{S}$ and $\mathcal{T}$. Then, associated with this algorithm, we define a Markov chain whose steady-state probability of staying in a certain region is directly related to the asymptotic fraction of matched points, which is the sought capacity C^*. Finally, we elaborate to compute analytically the mentioned steady-state probability.

A. Optimal Embedding Policy

As a first step toward capacity evaluation, we now show that an optimal embedding policy exists, which maximizes the number of matched points for any given cover traffic. This is achieved by an algorithm called the bounded greedy match (BGM) [27], which works as follows.

For a given realization of the two point processes $\mathcal{S}$ and $\mathcal{T}$ (all the points initially marked as "undetermined"), the BGM algorithm repeats the following steps (see Fig. 1).

1) Consider the first (in the direction of increasing time) undetermined point in the process $\mathcal{S}$, say $p^{(1)}$.
2) Find the first undetermined point in the process $\mathcal{T}$ in the interval $[p^{(1)}, p^{(1)} + \Delta]$, if any, denoted by $p^{(2)}$.
3) If such a point exists, mark both $p^{(1)}$ and $p^{(2)}$ as "matched"; otherwise, mark $p^{(1)}$ as "unmatched"; in either cases, mark all undetermined points in the process $\mathcal{T}$ before $p^{(1)}$ as "unmatched."

Matched and unmatched points are also referred to as "flow" and "chaff," respectively.

The BGM algorithm is optimal in the sense that, given two arbitrary realizations of point processes and an arbitrary value of Δ, the algorithm finds the maximum number of matched points satisfying the delay bound [27]. This result is contained in Theorem 8 of [27], which is actually given in terms of coin flips generated by two independent binomial processes. It can be seen, however, that the proof holds for any realizations of point processes,[2] as already noticed in [4] and [12]. By Definition 4, this implies the key fact that the embedding capacity is always achieved by BGM, regardless of the characteristics of the cover traffic.

B. Embedding Capacity in Terms of a Markov Chain

Our second step in deriving the embedding capacity C^* consists of modeling the behavior of BGM by a Markov chain, whose stationary distribution is directly related to C^*. Let Z_n be the nth sample of such chain: Z_n is the time difference between the "candidate" matching points at the nth iteration of BGM, as we now detail.

With reference to Fig. 2, let us consider the time difference between the first points in the two point processes $\mathcal{S}$ and $\mathcal{T}$, that is, $Z_1 = T_1 - S_1 = Y_1 - X_1$. According to the BGM algorithm, we have the following three possibilities.

1) If $Z_1 > \Delta$, the points cannot be matched, and the one in $\mathcal{S}$ is labeled as chaff. To decide the nature (chaff/nonchaff) of the point in $\mathcal{T}$, we must check whether it can be matched to the next arrival in $\mathcal{S}$, thus computing [see Fig. 2(a)]
$$Z_2 = T_1 - S_2 = Z_1 - X_2$$
where X_2 is the second interarrival of the process $\mathcal{S}$.
2) If $0 \le Z_1 \le \Delta$, the points match. To check the nature of the next incoming points, we update the process as [see Fig. 2(b)]
$$Z_2 = T_2 - S_2 = Z_1 + Y_2 - X_2$$
where Y_2 is the second interarrival of the process $\mathcal{T}$.
3) If $Z_1 < 0$, the points cannot be matched, and the one in $\mathcal{T}$ is labeled as chaff. To decide the nature of the point in $\mathcal{S}$, we must check whether it can be matched to the next arrival in $\mathcal{T}$, thus computing [see Fig. 2(c)]
$$Z_2 = T_2 - S_1 = Z_1 + Y_2.$$

[2]Specifically, the proof shows that for any pair of realizations of point processes and any embedding policy $\epsilon \in \mathcal{E}$, every unmatched point under BGM must have a corresponding unmatched point (looking backward in time) under ϵ. Note that the reference policy ϵ does not have to be sequential, i.e., it is allowed to make decisions based on the entire realizations.

By repeating for the successive points in the two streams, we see that a Markov process can be compactly defined in terms of the original renewals by the following rule

$$
\begin{aligned}
&\bullet\underline{\text{Initialization:}}\\
&\text{set:}\quad n=0;\ i_1=1;\ j_1=1;\ Z_0=0.\\
&\\
&\bullet\underline{\text{Recursion:}}\\
&n \leftarrow n+1\\
&\text{(a) Markov Chain}\\
&Z_n=\begin{cases} Z_{n-1}-X_{i_n}, & \text{if } Z_{n-1}>\Delta\\ Z_{n-1}+Y_{j_n}-X_{i_n}, & \text{if } 0\le Z_{n-1}\le\Delta\\ Z_{n-1}+Y_{j_n}, & \text{if } Z_{n-1}<0. \end{cases}\\
&\text{(b) Packet Classification \& Index Update}\\
&-\text{If } Z_n>\Delta, \text{then}: S_{i_n} \text{ is chaff}, \begin{matrix} i_{n+1}=i_n+1\\ j_{n+1}=j_n. \end{matrix}\\
&-\text{If } 0\le Z_n\le\Delta, \text{ then}\\
&\quad S_{i_n} \text{ and } T_{j_n} \text{ are flow}, \begin{matrix} i_{n+1}=i_n+1\\ j_{n+1}=j_n+1. \end{matrix}\\
&-\text{If } Z_n<0, \text{ then } T_{j_n} \text{ is chaff}, \begin{matrix} i_{n+1}=i_n\\ j_{n+1}=j_n+1. \end{matrix}
\end{aligned}
\tag{2}
$$

In the above, $X_{i_n}=S_{i_n}-S_{i_n-1}$ and $Y_{j_n}=T_{j_n}-T_{j_n-1}$ are the interarrivals of the processes $\mathcal{S}$ and $\mathcal{T}$, respectively, whose common PDF is $f(t)$. Given the sequential nature of the aforementioned recursion, the independence of the interarrivals and the independence between $\mathcal{S}$ and $\mathcal{T}$ ensure Markovianity of $\{Z_n, n=0,1,\ldots\}$.

The chain $\{Z_n, n=0,1,\ldots\}$ defined in (2) by running the BGM algorithm over a realization of the two point processes is schematically illustrated in Fig. 3, and the physical interpretation of the variables Z_n's is easily understood. From (2), a little thought reveals that the step of the chain at time n simply measures the time difference between the two points S_{i_n} and T_{j_n}, i.e., $Z_n=T_{j_n}-S_{i_n}$. Now, when $Z_n>\Delta$ (say, Z_7 in the figure), this is the distance between T_{j_n} (T_7), which is not classified yet at step n, and the chaff point S_{i_n} (S_5); similarly, when $Z_n<0$, this negative value measures how T_{j_n}, which is a chaff point, is far from S_{i_n}, which is not classified yet (for instance, $Z_3=S_3-T_3$ or $Z_4=S_3-T_4$). When Z_n stays inside the barriers, it measures the distance between the two flow points T_{j_n} and S_{i_n} (e.g., $Z_5=T_5-S_3$).

What is key for the forthcoming arguments is the operational meaning of the chain $\{Z_n, n=0,1,\ldots\}$ in terms of counting the matched points in the two streams $\mathcal{S}$ and $\mathcal{T}$. From the aforementioned discussion, note that to each step $n\ge 1$ such that $Z_n\in[0,\Delta]$, there corresponds a pair of flow points (one belonging to $\mathcal{S}$ and one to $\mathcal{T}$) matched by the BGM. Instead, to each $n\ge 1$ such that $Z_n\notin[0,\Delta]$, there corresponds a single chaff point (belonging to $\mathcal{S}$ if $Z_n>\Delta$, and belonging to $\mathcal{T}$ if $Z_n<0$). Therefore, the number of steps of the Markov chain lying inside (resp. outside) the barriers 0 and Δ defines the number of flow (resp. chaff) points marked by the BGM algorithm. Thus, a simple relationship exists between the asymptotic distribution of the chain and the fraction of flow points determined by the BGM algorithm, which is the embedding capacity. This is exploited in Section IV.

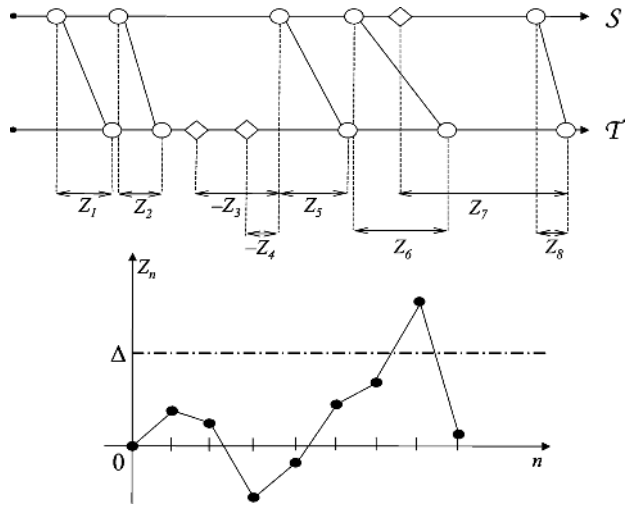

Fig. 3. Construction of a sample path of the Markov process (lower panel) from a realization of the two point processes (upper). In the upper panel, $\mathcal{S}=\{S_1,\ldots,S_6\}$, and $\mathcal{T}=\{T_1,\ldots,T_7\}$, while the corresponding interarrivals are $X_1\ldots,X_6$ and $Y_1,\ldots,Y_7$; The points marked with "$\diamond$" are those classified as chaff by the BGM algorithm.

C. Main Results

The first theorem we present, whose proof is deferred to Appendix A, establishes a connection between the embedding capacity and the invariant density of the BGM Markov chain $\{Z_n, n=0,1,\ldots\}$, expressed as the solution of an integral equation.

Theorem 1 (C^* by Markov Chain): Let $\mathcal{S}$ and $\mathcal{T}$ be two i.i.d. renewal processes, with interarrival PDF $f(t)$. Let Δ be the delay constraint, and define a Markov chain as in (2).

1) The invariant PDF $h(t)$ of the Markov chain exists and solves the following homogeneous Fredholm integral equation of the second kind [28]

$$
\begin{aligned}
h(t)=&\int_{-\infty}^{0} h(\tau)f(t-\tau)d\tau+\int_{\Delta}^{+\infty} h(\tau)f(\tau-t)d\tau\\
&+\int_{0}^{\Delta} h(\tau)f_0(t-\tau)d\tau
\end{aligned}
\tag{3}
$$

where $f_0(t)=\int_0^{+\infty} f(\tau)f(\tau-t)d\tau$ is the convolution between $f(t)$ and $f(-t)$.

2) The embedding capacity can be written as

$$
C^*=\frac{2\int_0^{\Delta} h(t)dt}{1+\int_0^{\Delta} h(t)dt}. \tag{4}
$$

$\diamondsuit$

It is useful to note that the packet-matching problem possesses a *scale-free* property: For a given distribution of the interarrivals, doubling the arrival rate "speeds up" the system so that the sample paths can be redrawn on a time axis scaled by a factor of 2, and halving Δ leaves unchanged the number of matches. We accordingly introduce a new Markov chain $\{\hat{Z}_n, n=0,1,\ldots\}$ defined by $\hat{Z}_n=\lambda(Z_n-\Delta/2)$. The

interarrivals corresponding to this modified chain are λX and λY, motivating the following definition.

Definition 5 (u-PDF): The PDF $k(t)$ of the interarrivals scaled to unit mean will be called u-PDF. $\diamond$

It is clear that the steady-state probability $\int_0^{\Delta} h(t)\,dt$ that the original chain belongs to $(0,\Delta)$ is exactly $\int_{-\lambda\Delta/2}^{\lambda\Delta/2} u(t)\,dt$, where $u(t)$ denotes the stationary density of the modified chain $\{\hat{Z}_n, n=0,1,\ldots\}$. Therefore, the piece of $u(t)$ relevant to capacity computation is

$$\omega(t) := \begin{cases} u(t), & -\frac{\lambda\Delta}{2} \le t \le \frac{\lambda\Delta}{2} \\ 0, & \text{otherwise} \end{cases}$$

whose Fourier transform is

$$\Omega(f) := \int_{-\infty}^{+\infty} \omega(t) e^{i2\pi f t} dt = \int_{-\frac{\lambda\Delta}{2}}^{\frac{\lambda\Delta}{2}} u(t) e^{i2\pi f t} dt.$$

Let $K(f)$ be the Fourier transform of the u-PDF $k(t)$, and define the "kernel"

$$\mathcal{K}(\nu, f) := 2\,\Re\left\{\frac{K(\nu)}{1-K(\nu)}\right\} \lambda\Delta \mathrm{sinc}[\lambda\Delta(f-\nu)] + \frac{(\lambda\Delta)^2}{2} \mathrm{sinc}(\lambda\Delta\nu)\mathrm{sinc}(\lambda\Delta f) \tag{5}$$

where $\mathrm{sinc}(f) := \sin(\pi f)/(\pi f)$ for $f \neq 0$, and $\mathrm{sinc}(0) := 1$. We are now in the position of stating the next theorem (this and the forthcoming results of this section are proved in Section IV).

Theorem 2 (Exact Value of C^*): Assume that the interarrivals have finite second moment. The embedding capacity of two i.i.d. renewal processes with rate λ, under delay constraint Δ, is

$$C^* = \frac{2\Omega(0)}{1+\Omega(0)} \tag{6}$$

where $\Omega(f)$ is the solution of

$$\Omega(f) + \int_{-\infty}^{\infty} \Omega(\nu)\mathcal{K}(\nu, f)\,d\nu = \frac{\lambda\Delta}{2}\,\mathrm{sinc}(\lambda\Delta f). \tag{7}$$

$\diamondsuit$

As a check, let us specialize the aforementioned equation to the case of exponential interarrivals, for which the embedding capacity is available in closed form [4]. In the exponential case, it is easily seen that $\Re\left\{\frac{K(f)}{1-K(f)}\right\} = 0$, allowing direct solution of (7), and computation of $\Omega(0) = \lambda\Delta/(2+\lambda\Delta)$. Substituting into (6), this yields

$$C^* = \frac{\lambda\Delta}{1+\lambda\Delta} \qquad \text{(exponential)}$$

that matches the known result from [4].

Note that Theorem 2 still gives an implicit solution to the problem in terms of an integral equation, which in general does not admit a closed form. On the other hand, (7) is amenable to approximate solutions, as described next and detailed later in Section IV. First of all, it is possible to obtain a "discrete" version of the integral (7), in the form

$$\sum_{k=-\infty}^{\infty} A_{hk}\Omega\left(\frac{k}{\lambda\Delta}\right) = \frac{\lambda\Delta}{2} I_h, \quad h = \ldots, -1, 0, 1, \ldots \tag{8}$$

where $I_h = 1$ for $h = 0$, and $I_h = 0$ otherwise. The coefficients A_{hk}—that define a matrix of infinite size— can be expressed in terms of the renewal function of the (unit-mean) interarrivals.

Definition 6 (u-RF): Let $N(t)$ be the number of arrivals in $(0,t)$ of a renewal process with interarrivals scaled to unit mean, i.e., distributed according to the u-PDF $k(t)$. The renewal function

$$m(t) := \mathbb{E}[N(t)]$$

will be called u-RF. $\diamond$

Indeed, using this definition of $m(t)$, the coefficients A_{hk} can be written as

$$A_{00} = 1 - \frac{\lambda\Delta}{2} + \frac{2}{\lambda\Delta}\int_0^{\lambda\Delta} m(t)dt \tag{9}$$

$$A_{kk} = 1 + \frac{2}{\lambda\Delta}\int_0^{\lambda\Delta} m(t)\left[\cos\left(\frac{2\pi kt}{\lambda\Delta}\right) + 2\pi k\left(1-\frac{t}{\lambda\Delta}\right)\sin\left(\frac{2\pi kt}{\lambda\Delta}\right)\right]dt, \quad k \neq 0 \tag{10}$$

$$A_{0k} = \frac{2\,(-1)^k}{\lambda\Delta}\int_0^{\lambda\Delta} m(t)\,\cos\left(\frac{2\pi kt}{\lambda\Delta}\right)\,dt, \quad k \neq 0 \tag{11}$$

$$A_{hk} = \frac{(-1)^{h-k}}{(h-k)}\left[h(-1)^h A_{0h} - k(-1)^k A_{0k}\right], \quad h \neq k. \tag{12}$$

To obtain suitable approximations of C^*, let $N \ge 0$ be some integer, and let us set to zero the cross terms A_{hk}, $h \neq k$, for $|h| > N$ and $|k| > N$. In this way, the matrix involved in (8) takes the following form

$$\begin{array}{ccccccc} \bullet & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \bullet & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & & & & 0 & 0 \\ 0 & 0 & & \boldsymbol{A}_N & & 0 & 0 \\ 0 & 0 & & & & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \bullet & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \bullet \end{array} \tag{13}$$

that is an infinite matrix whose nonzero entries are the central block $\boldsymbol{A}_N := \{A_{hk}\}_{h,k=-N,\ldots,N}$, of size $2N+1$ by $2N+1$, and the diagonal entries A_{kk}, $|k| > N$, shown as dots in (13).

Using (13) in expression (8) is tantamount to focus on the linear system

$$\sum_{k=-N}^{N} A_{hk}\,\Omega_k = \frac{\lambda\Delta}{2} I_h, \qquad h = -N, \ldots, N \tag{14}$$

with coefficient matrix $\boldsymbol{A}_N$. The following result holds.

Theorem 3 (Linear System Approximation): Assume that the second moment of the interarrivals is finite. Then, for any size, the matrix $\boldsymbol{A}_N$ is positive definite, such that solving the linear system in (14) gives

$$\Omega_0 = \frac{\lambda\Delta}{2}\{\boldsymbol{A}_N^{-1}\}_{00}. \tag{15}$$

$\diamondsuit$

Accordingly, one introduces the following sequence of approximants for C^*

$$C_N := \frac{\lambda\Delta\{\boldsymbol{A}_N^{-1}\}_{00}}{1+\frac{\lambda\Delta}{2}\{\boldsymbol{A}_N^{-1}\}_{00}}, \quad N=0,1,\ldots \tag{16}$$

that correspond to neglect the A_{hk}, $h \neq k$ beyond a certain layer.

From (9)–(12), we see that $\boldsymbol{A}_N$ is very structured and its degrees of freedom grow only linearly with N; in fact, $\boldsymbol{A}_N$ is completely specified by assigning one row and the main diagonal, which is very convenient for numerical tractability; also, it is expected that the approximations (16) become more and more accurate as N increases.

There is more. By exploring the regimes $\lambda\Delta \ll 1$ and $\lambda\Delta \gg 1$, in Section IV-C, we shall offer plausibility arguments for neglecting *all* the off-diagonal terms A_{hk}, $h \neq k$, that is to say, for treating the matrix in (13) as exactly diagonal. This motivates our main approximation for computing C^*, which turns out to be very accurate in many practical cases.

Main Approximation of C^*: Under the assumption of finite second moment for the interarrivals, the embedding capacity of two i.i.d. renewal processes with rate λ, under delay constraint Δ, can be approximated as

$$C^* \approx C_0 := \frac{\lambda\Delta}{1+\frac{2}{\lambda\Delta}\int_0^{\lambda\Delta} m(t)dt} \tag{17}$$

with $m(t)$ being the u-RF. $\diamondsuit$

Again, let us apply (17) in the Poisson regime. The u-RF of an exponential random variable is $m(t)=t$ that inserted in (17) gives

$$C_0 = \frac{\lambda\Delta}{1+\lambda\Delta} \qquad \text{(exponential)}$$

implying that, in this particular case, formula (17) is exact, i.e., $C^* = C_0$. This can be understood by considering that the cross terms A_{0k} in (11) are zero in the exponential case.

As mentioned, the relevance of the aforementioned claim stems from the fact that, for the typical interarrival distributions encountered in many applications, the accuracy of the fully analytical approximation (17) seems to be excellent (see Section VI-A). Accordingly, C_0 represents an accurate and mathematically tractable expression for the embedding capacity under arbitrary renewal traffic.

We would like to emphasize that the characterization (17) relates the sought capacity to the u-RF of the underlying process. This highlights the role of the renewal function $m(t)$, and reveals that its average $\frac{1}{\lambda\Delta}\int_0^{\lambda\Delta} m(t)dt$ is the key quantity in determining C_0. Thus, different traffic models can be classified with respect to their embedding capabilities just in terms of that average.

Finally, we state a corollary characterizing the asymptotic behavior of the approximate capacity C_0 in the limit of $\Delta \gg 1/\lambda$. From a known property of the renewal function [29, Corollary 3.4.7], $m(t)-t \to (\gamma-1)/2$ in the limit of $t \to \infty$, where γ is the dispersion index defined in (1). Simply plugging that expression in (17) would give $1-C_0 \sim \gamma/(\lambda\Delta)$. Indeed, we have the following result.

Corollary 1 (Scaling Law for C_0): Under the assumption of finite second moment for the interarrivals, we have

$$\lim_{\lambda\Delta\to\infty}(1-C_0)(\lambda\Delta) = \gamma.$$

$\diamondsuit$

The corollary reveals that, for large values of the product $\lambda\Delta$, the key quantity in determining the capacity is the dispersion index: given $\lambda\Delta \gg 1$, the ability for a type of (renewal) traffic to hide information flows in independent realizations only depends on the value of the dispersion index γ, and different traffic models sharing the same dispersion index must behave similarly.

IV. Proofs by Riemann–Hilbert Theory

Let us define the *normalized delay* $\delta := \lambda\Delta$, and let us work in terms of the unit-mean random variables with u-PDF $k(t)$. Accordingly, in place of the original integral (3) involving the stationary density $h(t)$, we consider the following:

$$\begin{aligned} u(t) = &\int_{-\infty}^{-\delta/2} u(\tau)k(t-\tau)d\tau + \int_{\delta/2}^{+\infty} u(\tau)k(\tau-t)d\tau \\ &+ \int_{-\delta/2}^{\delta/2} u(\tau)k_0(t-\tau)d\tau \end{aligned} \tag{18}$$

where $k_0(t)$ is the convolution between $k(t)$ and $k(-t)$.

The integral (18) involves convolutions, which suggests working in the Fourier transform domain, following a classical approach to such integral equations pioneered by Carleman, Wiener, and Hopf [30], see also [31] and [32].

After transformation, the problem can be cast in the form of a Riemann–Hilbert[3] boundary value problem, which, in a nutshell, consists in finding two functions, analytic in the upper and lower half planes, respectively, whose limiting values on the real axis must obey an assigned boundary condition [30], [33]. Accordingly, the basic tool used in the next proof is that of complex analysis.

Before, we need some basic notation and concepts about one-sided functions and their analytic Fourier transforms, which will be useful in the following. Let $g(t) : \mathbb{R} \to \mathbb{R}$ be a function belonging to $L^1(\mathbb{R})$, and define

$$g^+(t) := \begin{cases} 0, & t<0, \\ \frac{1}{2}g(0), & t=0, \\ g(t), & t>0, \end{cases} \quad g^-(t) := \begin{cases} -g(t), & t<0 \\ -\frac{1}{2}g(0), & t=0 \\ 0, & t>0 \end{cases}$$

[3]Actually, we use the name "Riemann–Hilbert", even though, in the topical literature, different terminologies are found. According to Muskhelishvili [33] "The problem formulated above is often called the Riemann problem, but the author considers this name to be incorrect [...], because it was first considered by D. Hilbert essentially in the form in which it is stated."

implying that $g(t) = g^+(t) - g^-(t)$, $t \in \mathbb{R}$. Since $g(t) \in L^1(\mathbb{R})$, the Fourier transforms of $g(t)$, $g^+(t)$ and $g^-(t)$ exist for all f. In particular, $G(f) = \int_{-\infty}^{+\infty} g(t)e^{i2\pi ft}dt$, $G^+(f) = \int_0^{+\infty} g(t)e^{i2\pi ft}dt$, and $G^-(f) = -\int_{-\infty}^0 g(t)e^{i2\pi ft}dt$.

By replacing the real parameter f by a complex variable $z = f + iy$, the aforementioned integrals become $G^+(z) = \int_0^{+\infty} g(t)e^{i2\pi zt}dt$ and $G^-(z) = -\int_{-\infty}^0 g(t)e^{i2\pi zt}dt$, which are analytic in those regions of the complex plane of the variable z in which they are absolutely convergent [30]: in particular, $G^+(z)$ is analytic for $\Im(z) > 0$, and $G^-(z)$ for $\Im(z) < 0$.

A. Proof of Theorem 2

Consider the unknown function $u(t)$ in (18) and let

$$v^+(t) := \begin{cases} 0, & t \le 0, \\ u(t+\frac{\delta}{2}), & t > 0, \end{cases} \quad v^-(t) := \begin{cases} -u\left(t-\frac{\delta}{2}\right), & t < 0 \\ 0, & t \ge 0 \end{cases}$$

$$\omega(t) = \begin{cases} u(t), & \delta/2 \le t \le \delta/2 \\ 0, & \text{otherwise} \end{cases}$$

such that

$$u(t) = v^+(t - \delta/2) - v^-(t + \delta/2) + \omega(t). \tag{19}$$

The corresponding Fourier transforms will be accordingly denoted by $V^+(f)$, $V^-(f)$, and $\Omega(f)$. Note that, from (4), we are just interested in $\int_0^{\Delta} h(t)dt = \int_{-\delta/2}^{\delta/2} \omega(t)dt = \Omega(0)$.

Transforming both sides of the integral (18) into the Fourier domain gives

$$\begin{aligned} &V^+(f)e^{i\pi\delta f} - V^-(f)e^{-i\pi\delta f} + \Omega(f) \\ &\quad = V^+(f)e^{i\pi\delta f}\bar{K}(f) - V^-(f)e^{-i\pi\delta f}K(f) \\ &\quad + \Omega(f)|K(f)|^2 \end{aligned}$$

where $\bar{a}$ is the conjugate of a. The aforementioned equation can be recast as

$$\frac{V^+(f)e^{i\pi\delta f}}{1 - K(f)} = \frac{V^-(f)e^{-i\pi\delta f}}{1 - \bar{K}(f)} - W(f) \tag{20}$$

where we define[4]

$$W(f) := \Omega(f)\frac{1 - |K(f)|^2}{|1 - K(f)|^2} = \Omega(f)\left[1 + 2\,\Re\left\{\frac{K(f)}{1 - K(f)}\right\}\right]. \tag{21}$$

For notational simplicity, let

$$W_\ell(f) := W(f)e^{-i\pi\delta f}, \qquad W_r(f) := W(f)e^{i\pi\delta f}$$

where ℓ and r refer to left-/right-shift operations (in the time domain).

Multiplying both sides of (20) by $e^{-i\pi\delta f}$, and using the factorization $W_\ell(f) = W_\ell^+(f) - W_\ell^-(f)$ yields

$$\underbrace{\frac{V^+(f)}{1 - K(f)} + W_\ell^+(f)}_{X^+(z)|_{z=f}} = \underbrace{\frac{V^-(f)e^{-i2\pi\delta f}}{1 - \bar{K}(f)} + W_\ell^-(f)}_{X^-(z)|_{z=f}}. \tag{22}$$

[4]Note that $W(f)$ is well behaved at the origin. Indeed, given the assumption of finite second moment: $\lim_{f\to 0} \Re\left\{\frac{K(f)}{1-K(f)}\right\} = \frac{\gamma-1}{2}$, having used $K'(0) = i2\pi$ and $K''(0) = -4\pi^2(1+\gamma)$.

Recalling now the properties of one-sided Fourier integrals summarized just before Section IV-A, it is easy to see that the function $X^+(z)$ (resp. $X^-(z)$) is analytic in the upper (resp. lower) half plane $\Im\{z\} > 0$ (resp. $\Im\{z\} < 0$), continuous on the real axis, with a single pole located at $z = 0$.

The asymptotic behavior of the involved functions is essentially determined by Fourier transforms, such that we assume boundedness at infinity.

Summarizing, the left-hand side (LHS) and right-hand side (RHS) of (22) are boundary values of functions that are analytic in the upper half and lower half planes, respectively. They are further bounded at infinity, and coincide on the real axis $z = f$, where there is a single pole of order one located at $z = 0$.

An application of the analytic continuation theorem [34, Ch. XVI] will allow to *glue together* the two functions in the upper and lower half planes, obtaining a function which is analytic in the whole plane, except for the single pole of order one at the origin. The (generalized) Liouville theorem [34, Th. 10.23] defines the only admissible form that such a function can assume: c/z, where c is a constant to be determined.[5] Restricting to the real-axis only, we finally get

$$\frac{V^+(f)}{1 - K(f)} + W_\ell^+(f) = \frac{V^-(f)e^{-i2\pi\delta f}}{1 - \bar{K}(f)} + W_\ell^-(f) = \frac{c}{f}. \tag{23}$$

The value of the constant c is fixed by enforcing the condition that $u(t)$ is a PDF, which, in view of (19), is equivalent to $V^+(0) - V^-(0) = 1 - \Omega(0)$. Evaluating (23) at $f = 0$ yields $V^+(0) - V^-(0) = -i\,4\pi\,c$ (recall that $K'(0) = i2\pi$), whence

$$c = i\,\frac{1 - \Omega(0)}{4\pi}. \tag{24}$$

If we repeat the aforementioned development by multiplying (20) by the complex exponential $e^{i\pi\delta f}$, we get a similar result, finally obtaining the following system of equations:

$$\begin{aligned} \frac{V^+(f)}{1 - K(f)} + W_\ell^+(f) &= c/f \quad (i) \\ \frac{V^+(f)e^{i2\pi\delta f}}{1 - K(f)} + W_r^+(f) &= c/f \quad (ii) \\ \frac{V^-(f)}{1 - \bar{K}(f)} + W_r^-(f) &= c/f \quad (iii) \\ \frac{V^-(f)e^{-i2\pi\delta f}}{1 - \bar{K}(f)} + W_\ell^-(f) &= c/f. \quad (iv) \end{aligned}$$

Solving for $V^+(f)/[1 - K(f)]$ in (i) and (ii) gives

$$W_r^+(f)e^{-i\pi\delta f} - W_\ell^+(f)e^{i\pi\delta f} = \delta\mathrm{sinc}(\delta f)\,\frac{1 - \Omega(0)}{2}. \tag{25}$$

Using (iii) and (iv) gives identical results.

Now, observe that $W_r^+(f)e^{-i\pi\delta f}$ corresponds to the following chain of operations applied to $W(f)$, regarded in the time domain. First, a right shift of $\delta/2$ (suffix r); then, setting to zero the function on the negative time-axis ($^+$ operator); finally, a left shift of $\delta/2$ (multiplication by the complex exponential). These three steps clearly correspond to the single

[5]Actually, according to the generalized Liouville theorem, the overall function should be equal to $c_0 + c_1/z$. On the other hand, we are looking for a solution $U(f)$ in the class of the functions which vanish at infinity, implying $c_0 = 0$.

operation of setting to zero the values of the original function in the (time) region $(-\infty, -\delta/2)$. Similarly, $W_\ell^+(f)e^{i\pi\delta f}$ corresponds to setting to zero the values of the original function in the region $(-\infty, \delta/2)$. As a result, the LHS of (25) reduces to select, in the time domain, the original function only in the region $(-\delta/2, \delta/2)$, i.e., a low-pass filtering of $W(f)$. Accordingly, such LHS can be rewritten as the convolution $\int_{-\infty}^{\infty} W(\nu)\delta\,\mathrm{sinc}[\delta(f-\nu)]d\nu$. Using that and recalling definition (21), straightforward algebra gives the desired claim. •

B. Proof of Theorem 3

First, note that at LHS and RHS of (7) appear Fourier transforms of functions that vanish outside the range $[-\delta/2, \delta/2]$, such that we can resort to the (pointwise) sampling theorem [35, Th. 8.4.5]. We accordingly "sample" the equation as

$$\Omega\left(\frac{h}{\delta}\right) + \int_{-\infty}^{\infty} \Omega(\nu)\mathcal{K}\left(\nu, \frac{h}{\delta}\right) d\nu = \frac{\delta}{2} I_h \tag{26}$$

and further use $\Omega(f) = \sum_k \Omega(k/\delta)\mathrm{sinc}(\delta f - k)$. Substituting into the aforementioned equation, we get the set of equations in (8), i.e., $\sum_k A_{hk}\Omega(k/\delta) = \frac{\delta}{2} I_h$, where

$$A_{hk} = 2\int_{-\infty}^{\infty} \Re\left\{\frac{K(\nu)}{1-K(\nu)}\right\} \delta\mathrm{sinc}(\delta\nu - h)\mathrm{sinc}(\delta\nu - k)d\nu + I_{h-k} + \frac{\delta}{2} I_{|h|+|k|} \tag{27}$$

that, thanks to the results in Appendix B, can be expressed in the time domain as shown in (9)–(12).

It remains to prove that the matrix $\boldsymbol{A}_N = \{A_{hk}\}_{h,k=-N,\ldots,N}$, appearing in (13), is positive definite. To show this, let us consider $\boldsymbol{x} \in \mathbb{R}^{2N+1}$. For any $h = -N, \ldots, N$, using the expression (27) and observing that $1 + 2\,\Re\left\{\frac{K(\nu)}{1-K(\nu)}\right\} = \frac{1-|K(\nu)|^2}{|1-K(\nu)|^2}$, straightforward algebra gives

$$\sum_{h,k=-N}^{N} A_{hk}x_h x_k = \frac{\delta}{2}x_0^2 + \int_{-\infty}^{\infty} \frac{1-|K(\nu)|^2}{|1-K(\nu)|^2}\left(\sum_{k=-N}^{N} x_k\,\mathrm{sinc}(\delta\nu - k)\right)^2 d\nu.$$

It is now easy to see that the aforementioned (nonnegative) quantity can be zero only if $x_k = 0, \forall k$, whence positive definiteness of $\boldsymbol{A}_N$ follows. This implies that $\boldsymbol{A}_N$ is invertible, yielding (15). •

C. Main Approximation $C^* \approx C_0$

Recall that, should the infinite matrix in (13) be diagonal, then (15) would reduce to $\Omega_0 = \delta/(2A_{00})$, thus yielding, in view of (9), the approximation C_0 in (17). By using the analytical expressions in (11) and (12), we now show that the off-diagonal terms A_{hk}, $h \neq k$ can be in fact neglected, at least in the two regimes $\delta \ll 1$ and $\delta \gg 1$. For $\delta \ll 1$, this immediately follows by triangle inequality

$$|A_{0k}| \leq \frac{2}{\delta}\int_0^{\delta} m(t)\,dt \approx 0$$

where the approximation exploits $m(0) = 0$.

As to $\delta \gg 1$, from a renewal theorem for interarrivals with finite second moment [29, Corollary 3.4.7], we know that

$$\lim_{t\to\infty}[m(t) - t] = \frac{\gamma - 1}{2}. \tag{28}$$

Thus, from (11), we write for $k \neq 0$

$$A_{0k} = (-1)^k \frac{2}{\delta}\int_0^{\delta}\left[m(t) - t - \frac{\gamma-1}{2}\right]\cos(2\pi kt/\delta)\,dt$$

which follows by $\int_0^{\delta} t\cos(2\pi kt/\delta) = 0$. Then, again by triangle inequality

$$|A_{0k}| \leq \frac{2}{\delta}\int_0^{\delta}\left|m(t) - t - \frac{\gamma-1}{2}\right| dt \approx 0$$

where the last approximation is a consequence of the Cesáro mean theorem and (28).

D. Proof of Corollary 1

We consider the limiting behavior of

$$1 - C_0 = \frac{1 + \frac{2}{\delta}\int_0^{\delta} m(t)\,dt - \delta}{1 + \frac{2}{\delta}\int_0^{\delta} m(t)\,dt}. \tag{29}$$

Now

$$\frac{2}{\delta}\int_0^{\delta} m(t)\,dt = \frac{2}{\delta}\int_0^{\delta}[m(t) - t]\,dt + \delta \sim \gamma - 1 + \delta$$

by simple application of the Cesáro mean theorem and of the renewal theorem used earlier, see (28). From (29), we get the desired result

$$\lim_{\delta\to\infty}(1 - C_0)\delta = \gamma.$$

•

V. Ordering of Embedding Capacities

In this section, we show how the approximate embedding capacity C_0 can be used for comparing different renewal processes in terms of their embedding capabilities. Let X_1 and X_2 be two nonnegative random variables with the same average value $\mathbb{E}[X_1] = \mathbb{E}[X_2] = 1/\lambda$, and with CDFs denoted by $F_1(\cdot)$ and $F_2(\cdot)$, respectively. The following definitions and results are classical in the stochastic ordering literature, and can be found in, e.g., [29] and [36].

Definition 7 (Variability or Convex Ordering): The random variable X_1 is less variable than X_2, written $X_1 \leq_v X_2$, if

$$\mathbb{E}[\phi(X_1)] \leq \mathbb{E}[\phi(X_2)] \text{ for all convex functions } \phi : \mathbb{R} \to \mathbb{R} \tag{30}$$

provided that the expectations exist. ⋄

Known Results [36, p. 110] (Sufficient and necessary conditions for convex ordering): For nonnegative random variables X_1 and X_2, with $\mathbb{E}[X_1] = \mathbb{E}[X_2] = 1/\lambda$, the condition $X_1 \leq_v X_2$ is equivalent to

$$\int_0^x [1 - F_1(t)]dt \geq \int_0^x [1 - F_2(t)]dt, \quad \text{for all } x. \tag{31}$$

Intuitively, $X_1 \leq_v X_2$ if X_1 gives less weight to the extreme values with respect to X_2. One way to get this is just to ensure that $\mathbb{E}[\phi(X_1)] \leq \mathbb{E}[\phi(X_2)]$ for convex ϕ, as stated in (30). That is why this kind of stochastic ordering is also known as convex ordering. It is also obvious that $X_1 \leq_v X_2 \Rightarrow \mathrm{VAR}[X_1] \leq \mathrm{VAR}[X_2]$, and hence, X_1 has a dispersion index smaller than or equal to that of X_2, a fact that plays a major role for the capacity in the regime of $\Delta \gg 1/\lambda$, as seen in Corollary 1.

The following theorem formally relates the classical concept of variability ordering to the embedding capacity in a straightforward and intuitive way: less variable interarrivals yield larger embedding capacities.

Theorem 4: Let C_{01} and C_{02} be the approximate embedding capacities for i.i.d. renewal processes with interarrivals distributed as X_1 and X_2, respectively. Then

$$X_1 \leq_v X_2 \Rightarrow C_{01} \geq C_{02}. \tag{32}$$

♢

Proof: The u-RF's of X_1 and X_2 can be represented as [29, Proposition 3.2.1]

$$m_1(t) = \sum_{i=1}^{\infty} \mathbb{P}\left\{\lambda S_i^{(1)} \leq t\right\}, \quad m_2(t) = \sum_{i=1}^{\infty} \mathbb{P}\left\{\lambda S_i^{(2)} \leq t\right\} \tag{33}$$

where $S_i^{(j)}$ is the ith epoch of the jth process, $j = 1, 2$. Let us focus on the single terms of the series. Since $X_1 \leq_v X_2$, then $S_i^{(1)} \leq_v S_i^{(1)}$ for each $i = 1, 2, \ldots$ (see, e.g., [36]), so that in view of (31)

$$\int_0^{\lambda\Delta} \sum_{i=1}^{n} \mathbb{P}\left\{\lambda S_i^{(1)} \leq t\right\} dt \leq \int_0^{\lambda\Delta} \sum_{i=1}^{n} \mathbb{P}\left\{\lambda S_i^{(2)} \leq t\right\} dt.$$

Applying Beppo Levi's monotone convergence theorem [37, Prob. 16.9], it is legitimate to exchange integration and limit, yielding

$$\int_0^{\lambda\Delta} m_1(t)\, dt \leq \int_0^{\lambda\Delta} m_2(t)\, dt$$

which, in the light of (33), gives $C_{01} \geq C_{02}$. •

As a consequence of the aforementioned theorem, we have the following results. First, it is of special interest to compare a given renewal process to Poisson traffic. To do so, let us define two special categories of interarrival distributions.

Definition 8 (NBUE/NWUE Classes): A nonnegative random variable X is called NBUE or NWUE if [29, p. 436]

$$\begin{aligned} &\text{NBUE} \quad \mathbb{E}[X - s | X > s] \leq \mathbb{E}[X] \qquad \forall s \geq 0 \\ &\text{NWUE} \quad \mathbb{E}[X - s | X > s] \geq \mathbb{E}[X] \qquad \forall s \geq 0. \end{aligned}$$

♢

Due to the absence of memory, the exponential distribution is such that $\mathbb{E}[X - s | X > s] = \mathbb{E}[X]$, and it belongs to both classes. The following corollary offers a comparison between the NBUE and NWUE classes, in terms of the approximate embedding capacity C_0. Recall that, in the case of exponential interarrival times, the exact capacity is $C^* = \frac{\lambda\Delta}{1+\lambda\Delta}$.

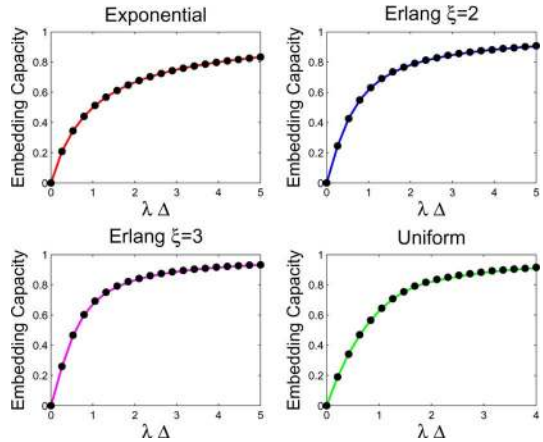

Fig. 4. Examples of traffic models for which the renewal function admits simple closed form. Dots refer to computer simulations of the embedding capacity and lines refer to analytical formulas.

Corollary 2 (Capacity ordering in NBUE/NWUE classes): Consider a pair of independent renewal processes with interarrival times distributed according to X, with $\mathbb{E}[X] = 1/\lambda$. Let Δ be the delay constraint, and C_0 the approximate embedding capacity. Then

$$\begin{aligned} X \text{ is NBUE} &\Rightarrow C_0 \geq \frac{\lambda\Delta}{1+\lambda\Delta} \\ X \text{ is NWUE} &\Rightarrow C_0 \leq \frac{\lambda\Delta}{1+\lambda\Delta}. \end{aligned} \tag{34}$$

♢

Proof: Thanks to Proposition 9.6.1 in [29], the NBUE (resp. NWUE) distributions can be shown to be less (resp. more) variable than the exponential, implying the claimed result as a direct consequence of Theorem 4. •

VI. Numerical Experiments

A. Simulations

The analytical expression of C_0 provided by (17) turns out to be quite accurate for virtually all the interarrival distributions used in our simulation studies, many of which are typical of network applications. A part of these extensive computer investigations is now summarized. In addition, we show an example where the refinements C_N in (16), with $N > 0$, provide meaningful improvements over C_0.

We start by considering some well-known interarrival distributions for which the renewal function is available in closed form, so that C_0 can be easily computed. In particular, we refer to the Erlang and the uniform random variables, whose u-PDFs are reported in Table I. The u-RF for the Erlang distribution can be found in [38, p. 57], while the u-RF for the uniform can be found in [39, Prob. 2, p. 385]. Then, C_0 follows by straightforward integration, and the final expressions are reported in Table I. Comparison to numerical simulations is depicted in Fig. 4, showing an excellent agreement.

Even when the renewal function is not known explicitly, there exist many numerical ways to compute that. Some methods exploit the definition of the renewal function in terms of interarrival distribution [29], other approaches are based on the interarrival density, and even others exploit the Fourier domain. To give an example, let us consider the Gamma family, whose

TABLE I
EMBEDDING CAPACITY (CLOSED-FORM APPROXIMATION C_0) FOR TYPICAL DISTRIBUTIONS. IN THE LAST COLUMN, THE RELATIONSHIPS BETWEEN CLASSICAL CONVEX ORDERING AND EMBEDDING CAPACITY ORDERING ARE REPORTED

	PDF $k(t)$, $t>0$	$C_0/(\lambda\Delta)$	Ordering relationship
ERLANG	$\xi\frac{(\xi t)^{\xi-1}}{\Gamma(\xi)}e^{-\xi t}$, $\xi\in\mathbb{N}$	$\left[1+\lambda\Delta+2\sum_{h=1}^{\xi-1}\frac{e^{i\frac{2\pi h}{\xi}}}{\xi(1-e^{i\frac{2\pi h}{\xi}})}\left(1-\frac{1-e^{-\xi(1-e^{i\frac{2\pi h}{\xi}})\lambda\Delta}}{\lambda\Delta\,\xi(1-e^{i\frac{2\pi h}{\xi}})}\right)\right]^{-1}$	$\xi_1\geq\xi_2\Rightarrow C_{01}\geq C_{02}$
UNIFORM	$\begin{cases}\frac{1}{2}, & t\in[0,2]\\ 0, & \text{otherwise}\end{cases}$	$\begin{cases}\frac{\lambda\Delta}{4e^{\frac{\lambda\Delta}{2}}-\lambda\Delta-4}, & \lambda\Delta\in[0,2]\\ \frac{\lambda\Delta}{4e^{\frac{\lambda\Delta}{2}}+2e^{\frac{\lambda\Delta}{2}-1}(4-\lambda\Delta)-\lambda\Delta-8}, & \lambda\Delta\in[2,4]\\ \dots & \dots\end{cases}$	meaningless
PARETO	$\frac{b/(b-1)}{\left(1+\frac{t}{b-1}\right)^{b+1}}$, $b>1$	Numerical: Use eq. (27) for computing A_{00}.	$b_1\geq b_2\Rightarrow C_{01}\geq C_{02}$
WEIBULL	$[\Gamma(1+1/b)]^b\,b\,t^{b-1}\,e^{-[\Gamma(1+1/b)\,t]^b}$ $b>0$	$\left[1+2\sum_{n=1}^{\infty}\frac{(-1)^{n-1}\,\beta_n\,[\Gamma(1+1/b)\,\lambda\Delta]^{n\,b}}{\Gamma(1+n\,b)\,(n\,b+1)}\right]^{-1}$ $\alpha_n=\frac{\Gamma(1+n\,b)}{n!}$, $\beta_1=\alpha_1$, $\beta_n=\alpha_n-\sum_{j=1}^{n-1}\alpha_j\,\beta_{n-j}$	$b_1\geq b_2\Rightarrow C_{01}\geq C_{02}$

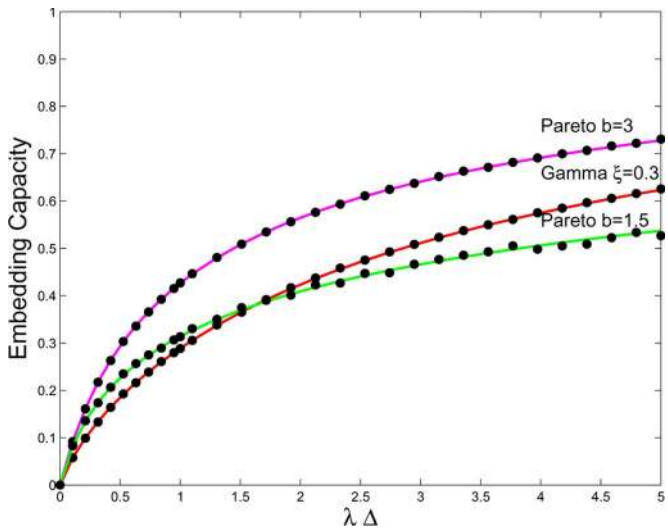

Fig. 5. Examples of different traffic models. Continuous curves refer to the approximation C_0 in Theorem 3 (17), while dots are obtained by computer simulations.

u-PDF is equivalent to that of an Erlang distribution, but with a shape parameter not necessarily integer, $\xi\in\mathbb{R}_+$. In this case, it is particularly convenient to use the expression for A_{00} obtained by (27). Computing numerically the involved integral, we get the capacity plotted in Fig. 5, for the case $\xi=0.3$. Again, the match with the results of computer simulation is excellent.

A case of special interest for network applications due to its tail behavior is the Pareto interarrival distribution, whose u-PDF is also given in Table I. Fig. 5 shows the embedding capacity, still obtained by deriving A_{00} in (27) via numerical integration, for the Pareto distribution. This distribution exhibits finite second moment whenever the shape parameter $b>2$. We first test the case $b=3$, which hence falls in the assumptions of our theorems (see Fig. 5). Then, we explore by simulation a case with infinite second moment, i.e., $b=1.5$, and Fig. 5 reveals that the accuracy of the formula is maintained.

We note in passing that the ordering results derived in Section V allow one to compare different cover traffic sharing the same interarrival PDF, but with different shape parameters. Indeed, for the mentioned traffic models, convex ordering is directly induced by ordering of the shape parameters (see, e.g., [40]), which in turn induces an ordering of the approximate embedding capacities, as stated in Theorem 4. The results of such comparisons are summarized in the last column of Table I.

In all the cases examined so far, there is no doubt that the expression C_0 is quite accurate for any practical purposes. We would like to present an example in which the analytical formula (17) is less accurate. Let us consider the following (shifted exponential) u-PDF for the interarrivals: $k(t)=\frac{1}{1-a}\,e^{-\frac{t-a}{1-a}}$, for $t\geq a$, with $0<a<1$.

The approximate embedding capacity C_0 is displayed, along with the simulated data in Fig. 6. As a first remark, note that the agreement is perfect in the range $\lambda\Delta<a$, where a linear shape is observed. This can be explained by observing that, for *any* random variable with u-PDF $k(t)$ which is zero in the range $t<a$, we have $m(t)=0$ for $t<a$. This implies that, in the range $\lambda\Delta<a$, the terms A_{0k} in (11) vanish, so that the approximation (17) is exact, and gives the linear relationship $C^*=\lambda\Delta$ in the considered range.[6] This is also consistent with earlier approximations and simulation results in [26].

Let us come back to the analysis of Fig. 6. It can be seen that the accuracy of C_0 is also very good for large $\lambda\Delta$, while, for intermediate values of the product $\lambda\Delta$, it is not satisfying. Thus, using (16), we compute the refined approximations C_1 and C_2. In particular, C_1 is still conveniently expressed in simple closed form

$$C_1=\frac{\lambda\Delta}{\frac{\lambda\Delta}{2}+A_{00}+2\,\frac{A_{01}^2}{A_{01}-A_{11}}} \tag{35}$$

a shape which highlights its role of a first-order correction to $C_0=\frac{\lambda\Delta}{\frac{\lambda\Delta}{2}+A_{00}}$. As can be seen in Fig. 6, the partial inaccuracy of the approximation C_0 is remediated with the adoption of C_1. The higher order approximant C_2 gives negligible improvements.

[6]The same conclusion can be also argued as follows. For delay Δ smaller than the minimum allowed interarrival time, $S_k-S_{k-1}>\Delta$, such that the probability that S_k matches is the probability that the first arrival after S_k in $\mathcal{T}$ occurs before $S_k+\Delta$ and it can be computed, due to independence between the processes, by using the residual lifetime distribution [29]: $\mathbb{P}[S_k\text{matches}]\approx\lambda\int_0^{\Delta}[1-F(t)]dt$. This implies $\mathbb{P}[S_k\text{matches}]\approx\lambda\Delta$, for $\lambda\Delta<a$. By ergodicity, $C^*=\lambda\Delta$ in the considered range.

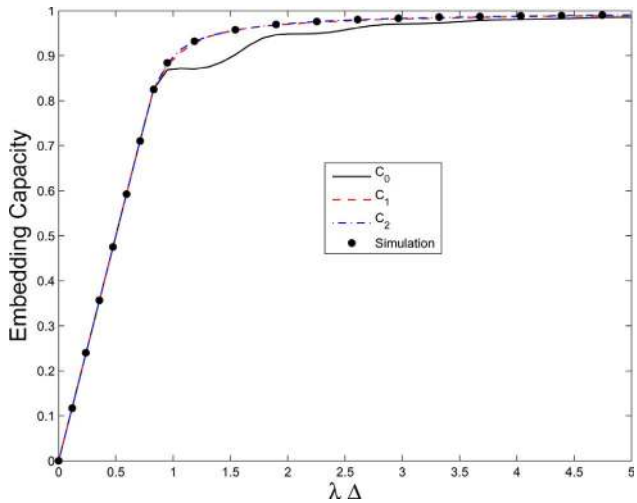

Fig. 6. Example of a shifted exponential distribution, with $a = 0.8$. Dots are obtained by computer simulations, while continuous curves refer to the different analytical approximations for C^* in Theorems 3 and 4. Specifically, we display 1) C_0, 2) C_1, namely the linear system solution with $N = 1$, see (35), and 3) the linear system solution for $N = 2$. The latter two curves are superimposed.

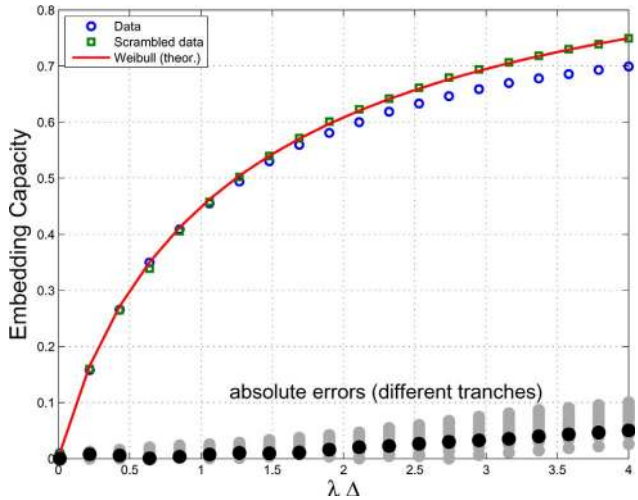

Fig. 7. Embedding capacity curve of Telnet data, for a pair of tranches selected as described in the main text. In the lower part of the plot, the absolute error between empirical and theoretical capacity is displayed, for a broader set of different tranches.

B. Experiments With Real Network Traces

In this section, we present some numerical tests run on the real-traffic traces *lbl-tcp-3.tcp* and *lbl-pkt-4.tcp*, made of TCP packet arrival times, gathered at the Lawrence Berkeley National Laboratory, Berkeley, CA, which were originally used in [41]. As done in [4] and [41], we extract packets corresponding to Telnet (port 23) connections. The experimental setup is as follows: suppose that trace *lbl-tcp-3.tcp* corresponds to the source node, while trace *lbl-pkt-4.tcp* refers to the relay node. Otherwise stated, the timestamps of *lbl-tcp-3.tcp* must be used as transmission epochs of the source node, and, similarly, those of *lbl-pkt-4.tcp* are the transmission epochs of the relay node.[7] Given a prescribed delay constraint Δ, we are interested in computing the maximum number of packets which can be relayed using these assigned scheduling patterns. Accordingly, for different values of Δ, we run the BGM algorithm over these real-traffic patterns, which therefore yields the corresponding (empirical) embedding capacity.

Let us describe more in detail how we process the real data. First, since traces *lbl-tcp-3.tcp* and *lbl-pkt-4.tcp* contain approximately 1 h of traffic, we extract smaller tranches, each made of 10^4 packets, and the BGM algorithm will be accordingly run over pairs of source/relay tranches of such length.

In order to ensure that the source and relay scheduling patterns work with (approximately) the same transmission rate, we 1) inspect both traces *lbl-tcp-3.tcp* and *lbl-pkt-4.tcp* by means of a moving average filter over 10^4 packets; 2) divide the interval between the smallest and the largest empirically estimated averages in equal-length bins of sufficiently small size, having verified the stability of the analysis with respect to different resolution cells; and 3) judge a source/relay pair admissible when their empirical averages fall into the same interval. With this selection procedure, the tranches extracted from a given trace might also overlap, which does not alter our analysis, in that we only need independence between the source and the relay. Finally, the BGM is run over all selected source/relay pairs, for different values of the delay constraint Δ (after scaling, without loss of generality, the traces to unit rate). This procedure finally gives an empirical embedding capacity curve for each selected source/relay pair.

To compare these empirical curves with the theoretical capacities, we first need a candidate marginal distribution for the interarrivals. To this aim, we fit the empirical interarrival CDF of each tranche, finding in general a good agreement with the Weibull distribution, that is perhaps not unexpected, see, e.g., [42] and [43]. The theoretical formula for C_0 in the Weibull case can be obtained by resorting to a known analytical expression for the Weibull renewal function [44], and is given in Table I. The value of the shape parameter b is estimated from the data.

The results of our investigations are summarized in Fig. 7, where the experimental points refer to one pair of selected tranches. A first evidence is that, up to values of $\lambda\Delta$ in the order of 1, the experimental points match well the theoretical approximation. On the other hand, a discrepancy emerges for larger values of the product $\lambda\Delta$.

Let us elaborate on this point. Were the real data exactly renewal processes, the marginal distribution of the interarrivals would provide their complete statistical characterization. Since the empirical CDF of the interarrivals is well fitted with a Weibull and the theoretical approximation C_0 for the Weibull in Table I has been verified to be excellent by independent computer experiments, we are driven to the following conclusion: The discrepancy between theoretical and empirical curves is to be ascribed to a deviation from the idealized renewal assumption.

This is confirmed by a further experiment. We run the BGM algorithm *after scrambling* the interarrivals, in order to mitigate statistical dependences. For these scrambled data, the theoretical approximation C_0 is now excellent, as shown by the squares in Fig. 7.

In the above, we illustrated the results concerning a single, specific pair of tranches. A more complete picture is obtained

[7]The two traces correspond to traffic patterns collected in two different days, such that the assumption of mutually independent point processes is met.

by applying the aforementioned procedure to different tranches, irrespectively of the goodness of the Weibull fit, and of the similarity between the empirical distributions at the two nodes. The results of this latter analysis are summarized in the bottom part of Fig. 7, where the absolute error between the theoretical formula and the empirical capacity is displayed. (Again, interarrival scrambling dramatically reduces the error; this is not shown in the plot.) The points marked with darker filled circles refer to the pair of tranches used for computing the empirical capacities displayed in the main part of the plot (i.e., the circles examined earlier). As can be seen, the theoretical approximation follows the empirical capacity closely at small $\lambda\Delta$; a discrepancy is observed for moderately large values of $\lambda\Delta$, with an absolute error in the order of 10^{-1}.

Summarizing, the analysis carried over the real traces highlights to what extent the deviations from the idealized renewal assumption, unavoidably present in practice, impact the embedding capacity. A main behavior seems to emerge—that for *tight* delay constraints, up to delay values in the order of the mean interarrival time, these deviations have in fact a negligible effect, which corroborates the theoretical study of the embedding capacity.

VII. Conclusion

We consider the problem of matching two i.i.d. renewal processes, according to a bounded delay criterion, with applications to communication network scenarios. We introduce the concept of *embedding capacity*, and provide fully analytical tools and approximations to evaluate it, relying upon the Riemann–Hilbert theory. An exact evaluation of the capacity is reduced to a manageable integral equation, which can be solved to any degree of approximation by inverting a highly structured linear system.

One main finding is a simple approximate formula of the embedding capacity that involves the renewal function of the underlying processes. The approximation is excellent for virtually all the cases of practical interest that we have investigated, part of which are reported in the paper. Even when this is not strictly true, we provide closed-form solutions for higher order corrections.

The analytical formula of the embedding capacity also highlights the role played by the traffic parameters: the amount of stochastic variability of the underlying interarrivals induces an ordering of the related embedding capacities, while for large values of $\lambda\Delta$, only the dispersion index matters.

The experimental analysis carried on real network traces reveals that the accuracy of the analytical expression is good for tight delay constraints, up to $\lambda\Delta$ in the order of 1. For larger delays, a partial inaccuracy is observed, and we show that this is to be ascribed to statistical dependences unavoidably present in real traffic patterns: the renewal model is failing, rather than the proposed analytical approximation.

The abstract concept of matching between point processes arises in a very large number of contexts, and we feel that our findings can represent a contribution to these disparate fields. To broaden further the horizon of potential applications, refinements and improvements of the approach can be considered. These include: the case of different renewal processes at the two nodes, the extension to nonrenewal point processes, to multihop flows, and to the case of multiple input/multiple output relays.

Appendix A
Proof of Theorem 1

Consider the Markov chain $\{Z_n, n = 0, 1, \ldots\}$ defined in (2). By Theorem 17.1.7 in [45], if $\{Z_n, n = 0, 1, \ldots\}$ is positive Harris recurrent, then we have the following:

1) $p := \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n} I_{[0,\Delta]}(Z_j)$ exists a.s., where $I_{[0,\Delta]}(z)$ is the indicator function.
2) An unique invariant probability measure π exists, solving
$$\pi(\mathcal{C}) = \int P(z, \mathcal{C})\pi(dz)$$
where $\mathcal{C}$ is any Lebesgue measurable set and $P(z, \cdot)$ is the transition kernel of the chain. Note that in our case, the invariant measure π admits a density. Indeed, the transition kernel $P(z, \cdot)$ is absolutely continuous, since so are the interarrivals; therefore, for any set $\mathcal{C}$ of zero Lebesgue measure, the aforementioned integral gives $\pi(\mathcal{C}) = 0$ as well. In addition, by application of the Fubini theorem, one version of the density, say $h(t)$, is the solution to (3), see, e.g., [46].
3) The asymptotic frequency p can be computed as $p = \int_0^{\Delta} h(t)dt$.

Suppose hence that $\{Z_n, n = 0, 1, \ldots\}$ is positive Harris recurrent. We first justify the embedding capacity formula (4): Since each Z_n outside $[0, \Delta]$ represents a chaff point, whereas each Z_n inside the interval represents a pair of flow points, we see that the fraction of flow points embedded by BGM converges a.s., and the limit, i.e., the embedding capacity, is given by $2\,p/(1 + p)$.

It remains to prove the property of positive Harris recurrence. First, we show that the Markov chain $\{Z_n, n = 0, 1, \ldots\}$ is ψ-irreducible [45] (all the sets mentioned in the sequel are Borel). Since the interarrivals are absolutely continuous, the BGM can match one pair a.s., implying that the interval $[0, \Delta]$ is accessible from any state a.s., say $L(z, [0, \Delta]) = 1\ \forall z$ [45, p. 64]. This rules out the cases where the asymptotic fraction of matched points depends on the initial state, and those where the embedding capacity is trivially zero.

Let φ be the Lebesgue measure constrained to $[0, \Delta]$, i.e., $\varphi(\mathcal{A}) = \mu(\mathcal{A} \cap [0, \Delta])$, where μ is the Lebesgue measure over the real line. Given the PDF $f(t)$, there must exist $\epsilon_0 > 0$ such that $f(t) > \delta_0$ for all t within some interval $[t_0, t_0 + \epsilon_0]$, and thus

$$f_0(t) = \int_0^{+\infty} f(\tau)f(\tau - t)d\tau > \delta_0^2(\epsilon_0 - |t|) \geq \delta_0^2(\epsilon_0 - \epsilon_1)$$

for all $t \in [-\epsilon_1, \epsilon_1]$, where ϵ_1 is a constant in $(0, \epsilon_0)$. Let $\delta_1 := \delta_0^2(\epsilon_0 - \epsilon_1)$. Partition $[0, \Delta]$ into $m := \lceil 2\Delta/\epsilon_1 \rceil$ segments of length $\epsilon_1/2$, as illustrated in Fig. 8, such that the transition density from any $z \in [0, \Delta]$ to any point in an adjacent segment is greater than δ_1. For any set $\mathcal{C}$ with $\varphi(\mathcal{C}) > 0$, let ϵ_2 be the Lebesgue measure of the minimum intersection between $\mathcal{C}$ and the $\frac{\epsilon_1}{2}$-segments. Let z be an arbitrary point in $[0, \Delta]$ that is n

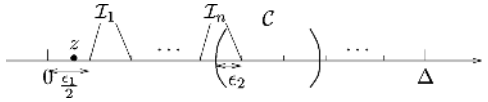

Fig. 8. Access $\mathcal{C}$ from z by hoping through $\frac{\epsilon_1}{2}$-segments in $[0, \Delta]$.

segments away from $\mathcal{C}(n \leq m-1)$ and $\mathcal{I}_i$ $(i = 1, \ldots, n)$ be the ith segment from z to $\mathcal{C}$, where $\mathcal{I}_n$ intersects with $\mathcal{C}$. The n-step transition satisfies

$$\begin{aligned} P^n(z, \mathcal{C}) &> \int_{\mathcal{I}_1} f_0(x_1 - z)dx_1 \int_{\mathcal{I}_2} f_0(x_2 - x_1)dx_2 \cdots \\ &\quad \int_{\mathcal{I}_n \cap \mathcal{C}} f_0(x_n - x_{n-1})dx_n \\ &> \left(\frac{\delta_1 \epsilon_1}{2}\right)^{n-1} \delta_1 \epsilon_2 > 0. \end{aligned} \tag{36}$$

This implies $L(z, \mathcal{C}) > 0$ for all $z \in [0, \Delta]$. Moreover, since $L(z, [0, \Delta]) = 1$ for all z, we have $L(z, \mathcal{C}) > 0$ for all z. That is, any set with positive φ measure is accessible from anywhere within the state space with positive probability, implying that the chain is φ-irreducible and hence ψ-irreducible for a maximal irreducibility measure ψ, according to [45].

Second, we show that $\{Z_n, n = 0, 1, \ldots\}$ is Harris recurrent. Since it is ψ-irreducible and $L(z, [0, \Delta]) > 0$ for all z, by Theorem 5.2.2 in [45], there exist $k \geq 1$, a nontrivial measure ν_k, and a nontrivial set $\mathcal{C}_1 \subseteq [0, \Delta]$ such that $\mathcal{C}_1$ is ν_k-small, and hence ν_{δ_k}-petite. For sampling distribution $a(i) = 1/m$ $(i = 1, \ldots, m)$, the transition kernel of the sampled chain from any $z \in [0, \Delta]$ satisfies

$$K_a(z, \mathcal{C}_1) \geq \frac{1}{m} P^n(z, \mathcal{C}_1) > \frac{1}{m} \left(\frac{\delta_1 \epsilon_1}{2}\right)^{n-1} \delta_1 \epsilon_2 \tag{37}$$

where we apply (36) for $\mathcal{C} = \mathcal{C}_1$. Since $n \leq m-1$, $K_a(z, \mathcal{C}_1) > \frac{1}{m} (\delta_1 \epsilon_1/2)^{m-2} \delta_1 \epsilon_2$, independent of z for $z \in [0, \Delta]$. Therefore, $\mathcal{C}_1$ is uniformly accessible using a from $[0, \Delta]$. By Proposition 5.5.4 in [45], we prove that $[0, \Delta]$ is $\nu_{a * \delta_k}$-petite. The fact that a petite set $[0, \Delta]$ satisfies $L(z, [0, \Delta]) = 1$ for all z for a ψ-irreducible chain implies Harris recurrence in the light of Proposition 9.1.7 in [45].

Finally, we show positivity by drift analysis. Define the function

$$V(z) = 2\lambda \begin{cases} z - \Delta, & \text{if } z > \Delta \\ 0, & \text{if } 0 \leq z \leq \Delta \\ -z, & \text{if } z < 0 \end{cases}$$

where $1/\lambda$ is the mean interarrival time, and consider the mean drift defined in [45] as

$$dV(z) = \int P(z, dy) V(y) - V(z)$$

where we recall that $P(z, dy)$ is the transition kernel of the chain, i.e., $dV(z) = \mathbb{E}[V(Z_{n+1}) | Z_n = z] - V(z)$, for $n = 0, 1, \ldots$, and $z \in \mathbb{R}$. Define a set $\mathcal{C}_2 = [-z_0, \Delta + z_0]$ for z_0 sufficiently large such that $\int_0^{z_0} f(t)tdt - \int_{z_0+\Delta}^{\infty} f(t)tdt \geq 1/(2\lambda)$. For any $z > \Delta + z_0$, we have after some straightforward manipulations

$$\begin{aligned} dV(z) &= -2\lambda \left[-\int_z^{\infty} f(t)(t - z)dt \right. \\ &\quad \left. + \int_0^{z-\Delta} f(t)tdt + (z - \Delta) \int_{z-\Delta}^{\infty} f(t)dt \right] \\ &\leq -2\lambda \left[\int_0^{z_0} f(t)tdt - \int_{z_0+\Delta}^{\infty} f(t)tdt \right] \leq -1. \end{aligned}$$

The same holds for $z < -z_0$. It is easy to see that, inside the set $\mathcal{C}_2$, $dV(z)$ can be bounded by a constant, such that we can write

$$dV(z) \leq -1 + b\, I_{\mathcal{C}_2}(z) \tag{38}$$

with a suitable choice of b. Since the petite set $[0, \Delta]$ is uniformly accessible[8] from $\mathcal{C}_2$, we can conclude that $\mathcal{C}_2$ is petite, and (38) coincides with the drift condition (iv) of Theorem 13.0.1 in [45], whence, further observing that aperiodicity holds, we conclude that $\{Z_n, n = 0, 1, \ldots\}$ is positive Harris. •

Appendix B
Linear System Coefficients

Let us introduce the so-called *renewal density* associated with the renewal function $m(t)$, i.e., $\rho(t) = dm(t)/dt$. It is convenient to consider a symmetric version thereof, namely $\widetilde{\rho}(t) = \rho(t) + \rho(-t)$. It holds true that $2\,\Re\left\{\frac{K(f)}{1-K(f)}\right\}$ is the Fourier representation of $\widetilde{\rho}(t) - 1$, see [47] and [48].

Let us first consider the term A_{00} in (27). We have

$$\begin{aligned} &2 \int_{-\infty}^{\infty} \Re\left\{\frac{K(\nu)}{1 - K(\nu)}\right\} \delta \mathrm{sinc}^2(\delta\nu) d\nu \\ &\quad = \int_{-\delta}^{\delta} [\widetilde{\rho}(t) - 1](1 - |t|/\delta)dt = 2 \int_0^{\delta} \rho(t)(1 - t/\delta)dt - \delta \end{aligned}$$

where we simply notice that the Fourier transform of the triangular window of width 2δ is $\delta \mathrm{sinc}^2(\delta f)$. Integration by parts then gives $2\int_0^{\delta} \rho(t)(1 - t/\delta)dt = \frac{2}{\delta}\int_0^{\delta} m(t)dt$, or

$$A_{00} = 1 - \frac{\delta}{2} + \frac{2}{\delta} \int_0^{\delta} m(t)dt.$$

As to the evaluation of A_{kk} in (27), $k \neq 0$, it suffices to use the shift property of the Fourier transform, yielding

$$\begin{aligned} &2 \int_{-\infty}^{\infty} \Re\left\{\frac{K(\nu)}{1 - K(\nu)}\right\} \delta \mathrm{sinc}^2(\delta\nu - k) d\nu \\ &\quad = \int_{-\delta}^{\delta} [\widetilde{\rho}(t) - 1](1 - |t|/\delta) \cos(2\pi kt/\delta)dt \\ &\quad = 2 \int_0^{\delta} \rho(t)(1 - t/\delta) \cos(2\pi kt/\delta)dt \end{aligned}$$

[8]This can be easily shown with the same technique used to prove uniform accessibility of $\mathcal{C}_1$ from $[0, \Delta]$.

that integrated by parts gives

$$A_{kk} = 1 + \frac{2}{\delta}\int_0^{\delta} m(t)[\cos(2\pi kt/\delta)\,dt + 2\pi k\,(1-t/\delta)\,\sin(2\pi kt/\delta)]\,dt.$$

Finally, focusing on the terms A_{hk} in (27), $h \neq k$, it suffices to consider the even part of $\delta\mathrm{sinc}(\delta f - h)\mathrm{sinc}(\delta f - k)$, whose inverse Fourier transform is

$$\frac{1}{\delta}\,\Re\left\{\int_{-\delta/2}^{\delta/2} e^{-i2\pi(h-k)\frac{\tau}{\delta}} e^{-i2\pi k\frac{t}{\delta}}\Pi\left(\frac{t-\tau}{\delta}\right)d\tau\right\} = \int_{-1/2}^{1/2}\cos[\,2\pi(h-k)\tau + 2\pi kt/\delta\,]\,\Pi(\tau - t/\delta)\,d\tau$$

where $\Pi(t) = 1$ for $t \in [-1/2, 1/2]$, and zero otherwise. The integral is zero for $|t| > \delta$. For $t \in (0, \delta)$, we have

$$\int_{t/\delta-1/2}^{1/2}\cos[2\pi(h-k)\tau + 2\pi kt/\delta]d\tau = \frac{(-1)^{h-k}}{2\pi(h-k)}[\,\sin(2\pi kt/\delta) - \sin(2\pi ht/\delta)\,].$$

This gives

$$A_{hk} = \frac{(-1)^{h-k}}{2\pi(h-k)}2\int_0^{\delta}\rho(t)[\sin(2\pi kt/\delta) - \sin(2\pi ht/\delta)]\,dt = \frac{(-1)^{h-k}}{(h-k)}\frac{2}{\delta}\int_0^{\delta} m(t)[h\cos(2\pi ht/\delta) - k\cos(2\pi kt/\delta)]dt$$

where the latter expression is obtained integrating by parts. This proves (12), while (11) follows as a special case.

REFERENCES

[1] C. Leberknight, H. Inaltekin, M. Chiang, and H. V. Poor, "The evolution of online social networks," *IEEE Signal Process. Mag.*, vol. 29, no. 2, pp. 41–52, Mar. 2012.

[2] P. Dayan and L. F. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: MIT Press, 2001.

[3] Z. F. Mainen and T. J. Sejnowski, "Reliability of spike timing in neocortical neurons," *Science*, vol. 268, pp. 1503–1506, 1995.

[4] T. He and L. Tong, "Detection of information flows," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 4925–4945, Nov. 2008.

[5] D. Donoho, A. Flesia, U. Shankar, V. Paxson, J. Coit, and S. Staniford, "Multiscale stepping-stone detection: Detecting pairs of jittered interactive streams by exploiting maximum tolerable delay," in *Proc. 5th Int. Symp. Recent Adv. Intrusion Detection, Lecture Notes Comput. Sci.*, 2002, vol. 2516, pp. 17–35.

[6] S. Staniford-Chen and L. Heberlein, "Holding intruders accountable on the internet," in *Proc. 1995 IEEE Symp. Security Privacy*, Oakland, CA, May 1995, pp. 39–49.

[7] J.-F Raymond, , H. Federrath, Ed., "Traffic analysis: Protocols, attacks, design issues and open problems," in *Designing Privacy Enhancing Technologies*, ser. Lecture Notes in Computer Science. Berlin, Germany: , 2001, vol. 2009, pp. 647–652.

[8] V. L. Voydock and S. T. Kent, "Security mechanisms in high-level network protocols," *J. ACM Comput. Surveys*, vol. 15, pp. 135–171, 1983.

[9] S. Marano, V. Matta, T. He, and L. Tong, "Embedding covert information flow," in *Proc. Annu. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 7–10, 2010, pp. 52–56.

[10] S. Marano, V. Matta, T. He, and L. Tong, "Embedding information flows into renewal traffic," in *Proc. IEEE Inf. Theory Workshop*, Paraty, Brazil, Oct. 16–20, 2011, pp. 50–54.

[11] D. Chaum, "Untraceable electronic mail, return addresses and digital pseudonyms," *Commun. ACM*, vol. 24, no. 2, pp. 84–88, Feb. 1981.

[12] P. Venkitasubramaniam, T. He, and L. Tong, "Anonymous networking amidst eavesdroppers," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2770–2784, Jun. 2008.

[13] B. Radosavljevic and B. Hajek, "Hiding traffic flow in communication networks," in *Proc. Military Commun. Conf.*, 1992, pp. 1096–1100.

[14] Y. Zhu, X. Fu, B. Graham, R. Bettati, and W. Zhao, "On flow correlation attacks and countermeasures in mix networks," in *Proc. 4th Int. Conf. Privacy Enhancing Technol. Workshop*, May 26–28, 2004, pp. 207–225.

[15] R. Maheshwari, J. Gao, and S. R. Das, "Detecting wormhole attacks in wireless networks using connectivity information," in *Proc. 26th IEEE Int. Conf. Comput. Commun.*, Anchorage, AK, May 6–12, 2007, pp. 107–115.

[16] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, pp. 656–715, Oct. 1949.

[17] A. D. Wyner, "The wire-tap channel," *AT & T Bell Labs Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

[18] I. Csiszár and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[19] Y. Liang and H. V. Poor, "Multiple-access channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 976–1002, Mar. 2008.

[20] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, Jun. 2008.

[21] E. Ardestanizadeh, M. Franceschetti, T. Javidi, and Y. Kim, "Wiretap channel with secure rate-limited feedback," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5353–5361, Dec. 2009.

[22] J. Xu, Y. Cao, and B. Chen, "Capacity bounds for broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4529–4542, Oct. 2009.

[23] T. He and L. Tong, "Distributed detection of information flows," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 3, pp. 390–403, Sep. 2008.

[24] T. He, A. Agaskar, and L. Tong, "Distributed detection of multi-hop information flows with fusion capacity constraints," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3373–3383, Jun. 2010.

[25] T. He, A. Agaskar, and L. Tong, "On security-aware transmission scheduling," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Las Vegas, NV, Mar.–Apr. 30–4, 2008.

[26] T. He, L. Tong, and A. Swamy, "Maximum throughput of clandestine relay," in *Proc. 47th Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Sep.–Oct. 30–2, 2009.

[27] A. Blum, D. Song, and S. Venkataraman, "Detection of interactive stepping stones: Algorithms and confidence bounds," in *Proc. Conf. Recent Adv. Intrus. Detect.*, Sophia Antipolis, France, Sep. 2004, pp. 39–49.

[28] A. C. Pipkin, *A Course on Integral Equations*. New York: Springer-Verlag, 1991.

[29] S. Ross, *Stochastic Processes*, 2nd ed. New York: Wiley, 1996.

[30] A. D. Polyanin and A. V. Manzhirov, *Handbook of Integral Equations*, 2nd ed. Boca Raton, FL: Chapman & Hall, 2008.

[31] D. S. Jones, "Diffraction by a wave-guide of finite length," *Proc Camb. Phil. Soc.*, vol. 48, no. 1, pp. 118–134, 1952.

[32] B. Noble, *Methods Based on the Wiener-Hopf Technique for the Solution of Partial Differential Equations*. New York: Pergamon, 1958.

[33] N. I. Muskhelishvili, *Singular Integral Equations: Boundary Problems of Function Theory and Their Application to Mathematical Physics*, 2nd ed. New York: Dover, 2008.

[34] W. Rudin, *Complex Analysis*, 2nd ed. New York: McGraw-Hill, 1986.

[35] A. Lapidoth, *A Foundation in Digital Communication*. Cambridge, UK: Cambridge Univ. Press, 2009.

[36] A. Shaked and J. G. Shanthikumar, *Stochastic Orders*. New York: Springer-Verlag, 2007.

[37] P. Billingsley, *Probability and Measure*, 3rd ed. New York: Wiley-Interscience, 1995.

[38] R. E. Barlow and F. Proschan, *Mathematical Theory of Reliability*. Philadelphia, PA: SIAM, 1996.

[39] W. Feller, *An Introduction to Probability and Its Applications*. New York: Wiley, 1971, vol. 2.

[40] B. Wilfling, "A sufficient condition for Lorenz ordering," *Indian J. Statist.*, vol. 58, pp. 62–69, 1996.

[41] V. Paxson and S. Floyd, "Wide-area traffic: The failure of Poisson modeling," *IEEE/ACM Trans. Netw.*, vol. 3, no. 3, pp. 226–244, Jun. 1995.
[42] I. Norros, "On the use of fractional Brownian motion in the theory of connectionless networks," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 6, pp. 953–962, Aug. 1995.
[43] K. Papagiannaki, S. Moon, C. Fraleigh, P. Thiran, and C. Diot, "Measurement and analysis of single-hop delay on an IP backbone network," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 6, pp. 908–921, Aug. 2003.
[44] W. L. Smith and M. R. Leadbetter, "On the renewal function for the Weibull distribution," *Technometrics*, vol. 5, no. 3, pp. 393–396, Aug. 1963.
[45] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability*. London, U.K.: Springer-Verlag, 1993.
[46] P. W. Glynn and S. G. Henderson, "Estimation of stationary densities for Markov chains," in *Proc. 30th Winter Simulat. Conf.*, Washington, DC, 1998, pp. 647–652.
[47] W. Feller and S. Orey, "A renewal theorem," *J. Math. Mech.*, vol. 10, no. 4, pp. 619–624, 1961.
[48] H. Carlsson, "Remainder term estimates of the renewal function," *Ann. Probab.*, vol. 11, no. 1, pp. 143–157, 1983.

Stefano Marano received the Laurea degree (summa cum laude) in Electronic Engineering and the Ph.D. degree in Electronic Engineering and Computer Science, both from the University of Naples, Italy, in 1993 and 1997, respectively. Currently he is an Associate Professor at the University of Salerno, Italy, where he formerly served as Assistant Professor. His areas of interest include statistical signal processing with emphasis on distributed inference, sensor networks, and information theory. In these areas he has published more than one hundred papers, including some invited, on leading international journals/transactions and proceedings of leading international conferences. He has also given several invited talks in the area of statistical signal processing.

Stefano Marano was awarded the IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION 1999 Best Paper Award for his work on stochastic modeling of electromagnetic propagation in urban environments. He also co-authored the paper winning the Best Student Paper Award (2nd place) at the 12th Conference on Information Fusion in 2009. As a reviewer, he handled hundreds of papers, mainly for the IEEE TRANSACTIONS, and was selected as Appreciated Reviewer by the IEEE TRANSACTIONS ON SIGNAL PROCESSING in the years 2007 and 2008. Stefano Marano is in the Technical Committee of the major international conferences in the field of signal processing and data fusion, and recently served as Area Chair (Sensor Array and Multichannel Processing) for EUSIPCO 2011. He was in the Organizing Committee of the Ninth International Conference on Information Fusion (FUSION 2006), and in the Organizing Committee of the 2008 IEEE Radar Conference (RADARCON 2008). He is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and for the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS.

Vincenzo Matta received the Laurea degree in electronic engineering and the Ph.D. degree in information engineering from the University of Salerno, Italy, in 2001 and 2005, respectively. He is currently an assistant professor at the University of Salerno.

His research interests cover the wide area of statistical signal processing, detection and estimation, with current emphasis on the interplay between inference, communications and secrecy in distributed systems, such as sensor networks.

Ting He (M'07) is a Research Staff Member at IBM T. J. Watson Research Center, Yorktown Heights, NY. She received the Ph.D. degree and the M.S. degree, both in Electrical Engineering from the School of Electrical and Computer Engineering, Cornell University, in 2007 and the B.S. degree in Computer Science from Peking University, China, in 2003. At IBM, Ting works in the Wireless Networking Research Group and has acted as task lead in the ITA program funded by ARL and MoD and the NIST ARRA program funded by NIST. Previously at Cornell (2003–2007), Ting was a member of the Adaptive Communications & Signal Processing Group (ACSP) under the supervision of Prof. Lang Tong. Before joining Cornell, she worked as an undergraduate research assistant in Micro Processor Research & Development Center of Peking University from 2001 to 2003, during which period she participated in the development of Unicore System as part of the National 863 Plan of China.

Ting has been a member of IEEE since 2007 and was a student member from 2004 to 2007. She received the Best Student Paper Award at the 2005 International Conference on Acoustic, Speech and Signal Processing (ICASSP). She was an Outstanding College Graduate of Beijing Area and an Outstanding Graduate of Peking University in 2003. She was a winner of the Excellent Student Award of Peking University during 1999–2002 and a recipient of Canon, Sony, and Yang-Wang Academicians scholarships.

Ting has worked on controlled mobility, delay tolerant networking, MANETs, and sensor networks. Her interests span network optimization and modeling, detection and estimation theory, control theory, and information theory. Her recent interests include workload scheduling in opportunistic environment, sampling and tracking of dynamic graphs, online learning in dynamic networks, and information sharing under privacy constraint.

Lang Tong (S'87–M'91–SM'01–F'05) is the Irwin and Joan Jacobs Professor in Engineering at Cornell University, Ithaca, New York. He received the B.E. degree from Tsinghua University, Beijing, China, in 1985, and M.S. and Ph.D. degrees in electrical engineering in 1987 and 1991, respectively, from the University of Notre Dame, Notre Dame, Indiana. He was a Postdoctoral Research Affiliate at the Information Systems Laboratory, Stanford University in 1991. He was the 2001 Cor Wit Visiting Professor at the Delft University of Technology and had held visiting positions at Stanford University and the University of California at Berkeley.

Lang Tong's research is in the general area of statistical inference, communications, and complex networks. His current research focuses on inference, optimization, and economic problems in energy and power systems. He received the 1993 Outstanding Young Author Award from the IEEE Circuits and Systems Society, the 2004 best paper award from IEEE Signal Processing Society, and the 2004 Leonard G. Abraham Prize Paper Award from the IEEE Communications Society. He is also a coauthor of seven student paper awards.

He received Young Investigator Award from the Office of Naval Research. He was a Distinguished Lecturer of the IEEE Signal Processing Society.