

On Information Transmission Among Nanomachines

Giuseppa Alfano, Daniele Miorandi

CREATE-NET, Via Solteri 38, 38100 – Trento (Italy)

email: {giusi.alfano,daniele.miorandi}@create-net.org

Abstract—In this paper we review the fundamental issues arising when nanoscale devices are meant to be interconnected to transmit information. The possibility of manipulating and assembling objects at the atomic scale has paved the way for a future generation of computing machines, where nanoscale devices substitute silicon-based transistors. Interconnections, needed to perform complex operations, are expected to be the driving factor in terms of performance and costs of the resulting systems. In view of the current research on nanomachines, we are interested in understanding which may be the limits of communications at the nanoscale level. Our research stems from a few, simple and yet unanswered questions, like “What is the capacity of a nanowire/nanotube?”, “What is the capacity of molecular-based communication systems?” etc. While we do not answer to such questions directly, we shed some light on possible approaches based on information-theoretical concepts.

Index Terms—information theory, communications, nanoscale devices, nanotechnologies

I. INTRODUCTION

Many technological fields are experiencing one major revolution, coming from the possibility of manipulating and assembling objects at the nanometer scale. In particular, one of the fields where nanotechnologies are expected to have a major impact is electronics, whereby the possibility of working at the atomic scale opens up new possibilities for building faster, smaller, and cheaper integrated circuits. However, when building nanoscale devices, we implicitly come up with the problem of interconnecting them into a purposeful system, able to perform complex functions through the proper interconnection (by means of information transmission) of simple elements.

With the term “nanotechnology” we roughly refer to the possibility of manipulating objects at the nanometer scale. Nanotechnology, which appeared kind of science fiction when the Nobel prize R. Feynman introduced it in its 1959 talk “There’s Plenty of Room at the Bottom” [1], has now acquired the full status of science, thanks to the enormous technological progresses that took place over the last decades. Scientists have now the ability to build molecules with a given structure, controlling and determining the features of the resulting materials. While nanotechnology has received considerable attention from the scientific community, the main focus has been on the processes needed to build materials with given, desirable characteristics. Little attention has been paid, so far, in understanding communication at the nanoscale level.

In this paper, we aim at characterizing and discussing some of the issues arising when communications among nanoscale

devices are considered. The possibility of manipulating objects at the atomic scale has paved the way for a future generation of computing machines, where nanoscale devices substitute silicon-based transistors. Interconnections, needed to perform complex operations, are expected to be the driving factor in terms of performance and costs of the resulting systems. In view of the current research in nanomachines, we are interested in understanding which may be the limits of communications at the nanoscale level. Our research stems from a few, simple yet unanswered questions, like “What is the information transmission capacity of a nanowire/nanotube?”, “What is the capacity of molecular-based communication systems?” etc. It is worth remarking that such questions do not appear of purely speculative interest, since the understanding of the limiting performance of a system usually sheds some light on possible approaches (in the communication case: encoding/decoding strategies) able to attain the optimal operating point.

The remainder of this paper is organized as follows. Sec. II reviews the background of communication at the nanoscale level. Sec. III presents and discusses some related work. Sec. IV deals with the issues and challenges involved with the development of a general theory of communications among nanomachines. Sec. V concludes the paper summarizing the outlined challenges and open issues.

II. WHY COMMUNICATING AT THE NANOSCALE

In the last decades, we have been facing huge advances in CMOS technology, which enabled the mass production of powerful yet small electronic devices. At the basis of such creations, there is the introduction of silicon transistors. Technological advances have then made it possible to build ever smaller, faster and cheaper devices. This is reflected by the so-called Moore’s Law, introduced by Intel’s co-founder Gordon Moore in 1965, which predicts that the number of transistors on a chip doubles approximately every 24 months [2]. Similar exponential laws have been shown to hold true with a good level of approximation also for other parameters, such as the clock speed of processors, the memory capacity of a RAM card, the price of a single transistor etc. The arising of such law comes from both “internal” momentum (i.e., advances in technology) as well as “external” pressure (i.e., market dynamics and scale economies). However, there

is a general consensus that this cannot hold forever¹. Current silicon-based technology is struggling to work at ever smaller scales (in terms of the granularity of the photolithographic process used to build chipsets) in order to enable processors with better and better performance. There is, in fact, an insurmountable limit on the scale at which such process can be taken forward, and we are not far from reaching it.²

At the same time, another issue is becoming more and more important, and is related to the *way* of connecting components on a chipset. Chips, roughly speaking, rely on the interconnections of elementary components (i.e., transistors) to perform ever more complex operations. Traditionally, such interconnections have been made by metallic wires, and communications were achieved by propagating an electrical field. While the introduction of metals and dielectric materials with very good conductivity (respectively: isolation) properties allowed increasing by various orders of magnitude the performance of such systems, there is a lot of good reasons for believing that also this process is not going to be employed in the future. Some of such reasons were introduced in the work by Chang et al. [3], where the authors proposed to use a radio frequency wireless interconnection mechanism for chip components. In that paper, the authors proposed an alternative to metallic wires, still basing information transmission on the propagation of electromagnetic fields but using radio frequency communications for interconnecting elements. The line of thoughts is that such method would provide more flexibility, resilience and, ultimately, better performance with respect to traditional point-to-point or bus-based wiring techniques.

When interconnecting nanomachines, we are clearly interested in covering distances that are, roughly, of the same order of magnitude. The problem is, boiling it down to the very abstract, to transmit information at the nanometer scale. The general problem of *transmitting information* is the same C.G. Shannon addressed in its 1948 milestone work on “A mathematical theory of communication”, which laid down the basis of what is currently referred to as Information Theory. Information theory, in its purest form, deals with abstract notions: how a sender can reliably transmit information in the optimal way over a channel, whereby all entities involved (sender, channel, receiver) are just mathematical abstractions. Information theory is therefore a fairly general science, covering various issues concerned with transformations of stochastic processes. However, as communication engineers were mostly interested in communications based on the use of electromagnetic fields for covering long distances, the reference model has been, to a large extent, that of additive white Gaussian channel, depicted in Fig.1 for a generic input signal x with power constraint P and noise power spectral density $\frac{N_0}{2}$, for which the capacity (to be understood as the maximum

rate at which information can be successfully transmitted in bit/s) is given by the well-known formula:

$$C = B \cdot \log_2(1 + SNR), \quad (1)$$

B representing the channel bandwidth and SNR being the signal-to-noise ratio at the decision point.

Now, the basic questions we ask ourselves are: What kind of interconnections shall we use when building nano-scale chipsets? Also, is Shannon theory still valid at the nanometer scale? And, if yes, what are the reference models to be used as rule of thumb for system design and dimensioning?

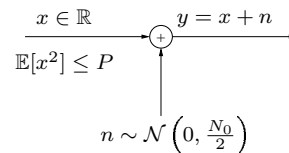


Fig. 1. Block scheme of a communication channel with additional white Gaussian noise and conditional law $p(y|x)$.

In order to address such questions, we start with a basic distinction, common in nanotechnologies, between dry and wet techniques [4].

A. Dry versus wet techniques

By *dry* techniques, we refer to all of the nanotechnologies that deal with the study of fabrication of structures in carbon, silicon and other inorganic materials. This includes, e.g., carbon nanowires and nanotubes. By *wet* techniques, on the other hand, we refer to the study of biological systems that operate in aqueous environment. This includes, e.g., molecular motors, DNA-based systems etc. Communications take place by different means in dry and wet techniques.

In dry nanotechnologies, we usually aim at building nanoscale components that then need to be assembled on a chip. In this case, we are somehow just scaling down the actual microscale silicon technology. Hence, a natural solution would be to try to use nanowires and nanotubes for transmitting information by means of electromagnetic fields, using them as waveguides. Clearly, due to the small distances to be covered, such fields look quite different from those used in the telecommunications world (i.e., near-field effect vs. far-field effect etc.), but the principle remains basically unchanged. In this case, nanotubes/nanowires provides sort of waveguides for driving the propagation. As of this writing, it is unclear if it would make sense to speak of “wireless” interconnection in the nanoscale domain, because of the problems related to the building of near-field antennas/capacity couplers.

Things drastically change when we consider the case of wet technologies. In such framework, indeed, it does not really make sense to use electromagnetic fields for propagating information. On the other hand, since we are dealing with biological systems, it is probably a good idea to draw inspiration from nature in order to design communication systems at such level. This direction, which remains to a large extent unexplored, has been recently tackled by the group of T. Suda, who,

¹Actually, the lower limit of integration is expected around $22\mu m$.

²Also, pushing such processes to the limits leads to an increase in the number of devices to be discarded in the manufacturing phase due to wiring failures or other malfunctioning elements. As this is having a large economical impact, chipset manufacturers are starting to consider the case of processors with some built-in resilience to lithographic process failures.

in a series of papers [5], [6], [7] explored the possibility of communicating at the nanometer scale using molecular motors and calcium ions. Other methods, based on hormones and proteins can also be envisaged. Such direction appears of particular interest for nano-chip designs for other various reasons. The first one is that communications is usually in nature a very energy-efficient process, much more than what can be achieved by electrical signalling. And it is acknowledged that one of the key problems of current micrometer silicon-based techniques is related to energy consumption, with the consequent need of providing efficient heat dissipation techniques. Furthermore, *wet* techniques tend to be inherently resilient to process failures. This is related to some common features of biological communications, which are based on the use of a “broadcast medium”, as an aqueous solution, which avoids the necessity of wiring or somehow connecting one by one the intended units. These observations make it clear that the use of wet techniques enables a clear distinction between communication and computing, which complies with the current trends in networks-on-chip research [8], [9].

III. RELATED WORK

As already remarked, most of the works appeared in the literature on nanomachines focused on understanding biological nanomachines as well as on creating artificial counterparts of them, whereas only some very recent efforts have aimed at investigating their key communication aspects. Among the proposed paradigms, we can identify two main directions, corresponding to wet and dry communication techniques, respectively.

A. Molecular communications

In [5], in particular, the possibility of effectively sending information at nanometer scale via carrier molecules is discussed, wherein different ways to perform intracellular, intercellular and longer range (on a nano basis) communication are proposed. In turn, an aqueous solution is assumed, as a communication channel, and two ways of sending information are mainly considered, namely concentration based information transmission, and propagation via molecular motors. This, indeed, follows the wet approach to nano communications.

The various steps needed for building a communication system are analyzed, such as encoding and sending the information, propagation of the signal through the aqueous medium (via concentration, motors or even cell to cell junction), receiving and finally decoding the intended signal. As a possible way of encoding the transmit signal, the authors proposed to modify some characteristics of the carrier molecules, which the receiver is especially sensitive to, or even to encode the information in the environment. Actually, since the behavior of receivers in biological systems is influenced by the environment, if the sender emits molecules that modify the environment, the receiver may detect the information represented in the environment. Transmission, the process by which the sender emits the carrier molecules into the environment, is performed

by releasing molecules through gap junctions between cells, or by using peptides-like molecular motors.

Information propagation, in turn, can be passive, constrained or active, depending on the amount, if any, of chemical energy consumed by molecules to propagate into the environment. An example of passive propagation is the Brownian motion, that does not consume chemical energy³, while constrained propagation is realized by limitation of the volume inside which particles propagates. Active propagation, finally, does consume chemical energy in order the molecules to move through the environment; this is the case, namely, of molecular motors based communication.

There are also multiple options for the intended receiver to *detect* the information, among which we mention the use of receptors that detect the presence of specific carrier molecules due to environment concentration modification, or also the use of gap junctions that allow molecules to flow into the receiving cell according to the concentration gradient of ions inside and outside the cell. Finally, always referring to the communication paradigm proposed in [5], *decoding* is performed via reaction of the receiver with the received molecules, while in some cases receiver can also simply store the received information with received molecules.

B. Communicating via nanotubes

Yet another option for information propagation is to use nanotubes, actually nanostructures several hundred times smaller than the diameter of a human hair that can be already manipulated and assembled into true devices with the existing technology. The use of such nanotubes, indeed, paves the way for nano wired communications, previously mentioned as *dry* communications. When the use of nanotubes is in force, in fact, information transmission is made possible via waves propagation, similarly to traditional wireless channels, with the proper energy and dimension scaling. In the context of *dry* communications, one can also think of wireless optical nano communications, via a phenomena similar to photo emitting molecules. In this view, the transmit process can be the photons emission from a properly excited molecule, and at the receiver end there will be a nano device capable of capturing the emitted photons. A possible way to perform such kind of communication is indeed suggested by plasmonic techniques [10]. Plasmons are coherent oscillations of the conduction electrons of the metal against the static positive background of the metal ion cores. Some predisposed metallic structures captures specific wavelengths of light and convert an amount of electrical energy back into light that is reflected away. Based on plasmonic resonance, such structures can actually guide light over long distances but on waveguides with lateral dimension far less than the optical wavelength, and very short curvature radius.

Research lines outlined in [5] are of great interest for understanding the behavior, and evaluating the feasibility, of

³It is worth remarking that such possibility of passive propagation is extremely appealing for biomedical applications, where blood flux can be flawlessly used to transport information.

nano-scale communication systems. However, they are mainly focused on empirical and simulation based investigations, that, while being necessary to better understand the underlying phenomena, need to be complemented by a theoretical analysis. To the best of the authors' knowledge, no similar investigation has been proposed so far. Nevertheless, such an approach will become ever more urgent while empirical and simulation based studies will be carried over. In the remaining of the paper we will thus highlight some of the more urgent topics toward the formulation of a communication theoretic framework for the feasibility of communications at the nanometer scale.

IV. TOWARDS A NANO-SHANNON THEORY: FIELD- AND MOLECULAR-BASED COMMUNICATIONS

A. Information theory fundamentals

As previously stated, we aim at presenting a first information theoretic look at nano-scale communications, enhancing the main differences between the macro and the nano-scale case. In order to make the work self-contained we briefly review the basic concepts in Shannon information theory [11] whose definition and/or validity in the nano setting will be discussed in the remaining of the paper.

- The Shannon information associated to an elementary event occurring with probability p is

$$\mathcal{I}(p) = -\ln p, \quad (2)$$

and is representative of the degree of uncertainty about the occurrence of the event.

- The entropy of a discrete random variable X with alphabet \mathcal{A} and probability law p_i is defined as:

$$H(X) = -\sum_{x \in \mathcal{A}} p_i \ln p_i. \quad (3)$$

From (2), (3) represents the average information associated to the event whose occurrence is modeled by the random variable X . It is worth to note that $H(X)$ can also be interpreted as the (average) amount of uncertainty which is removed by the knowledge of the actual value of X .

- The conditional entropy of a random variable X with alphabet \mathcal{A} , with respect to a random variable Y with alphabet \mathcal{B} is defined as the average information associated to the random variable $X|Y$, namely

$$H(X|Y) = -\sum_{x \in \mathcal{A}} \sum_{y \in \mathcal{B}} p(x, y) \ln p(x|y). \quad (4)$$

- The mutual information between two random variables X and Y with preassigned probability distribution can be written as

$$I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X); \quad (5)$$

if the random variables X and Y , respectively, represent the input and output of a noisy communication channel the mutual information can be interpreted as the information flow on the channel, namely the mutual information

can be viewed as a measure of how much uncertain we are left on X after observing Y . The symmetry of the relation defining $I(X; Y)$ tells us also that X says as much about Y as Y says about X .

- The channel capacity is the maximum information flow we can send over the channel, namely

$$C = \max_{p_i} I(X; Y),$$

where the maximization is taken over the set of the probability distributions for the discrete input. For the abovementioned AWGN model the capacity can be explicitly expressed as in (1)

Figure 2 depicts the scheme of a communication system from a (very basic) information theoretic point of view, where at the transmitter side the source message W is depicted, then the encoding block, comprehensive of both source as well as channel coding blocks, finally the channel itself and the decoding block at the receiver side, with estimated received signal \hat{W} . We assume the message W to be encoded by a generic n -length sequence X^n , and channel to have a transition law $p(y|x)$ by means of which a received sequence Y^n is generated. We will refer to such a scheme in the remaining of the paper when talking about coding questions in the nano setting.

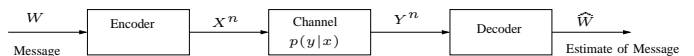


Fig. 2. Block scheme of a communication channel with conditional law $p(y|x)$.

B. Open questions in nano setting

The seminal work [5] mentions several problems in managing nano devices, mainly considering the drawbacks in adapting existing electrical and optical waves based transmission techniques to the nano setting, or the lack of possibility of connecting nanomachines via electrical wires, as well as the obstacles found in terms of environment compatibility. It is worth noting that far more problems arise in our attempt to formulate an information theoretic framework for nano communications. At the nano scale, classical mechanics laws usually do not fail to hold true as in quantum based communication, hence the well understood parallels [12], [11] between information theoretic and physics interpretation of some parameters as entropy, information and free energy still hold true in our setting. As a consequence, the problem we have to afford mainly concerns analytical characterization of the parameters in the nano communications scenario. Such a need for translating on the nano environment the main information theoretic parameters and performance indexes poses, indeed, several questions whose answers will strongly influence the (at least partial) answers to the more general questions posed at the very beginning of the paper. From an information theoretic point of view, what we are mainly interested in are the limiting capabilities of the communication, namely the channel capacity and, as a first consequence, the capacity achieving network

protocols, but also the characterization of the channel outage. While in the conventional telecommunication case all these quantities are well understood and quite often expressed in handy closed forms, at the nano-scale none of these quantities' expressions have been evaluated so far.

The most urgent question is the statistical characterization of the communication channel $p(x|y)$, which for molecular communications based on aqueous environment is by its nature nonlinear, thus posing several problems in evaluating the entropy associated to the received signal. It is worth noting that, while for the case of communication based on molecular motor we deal with point-to-point communication, the case of communication via aqueous environment is inherently a broadcast type communication, for which the study of the capacity is still an open problem in several cases also for the traditional wireless communications paradigm. When dealing with nanotubes molecular communications, instead, even if the setting is similar to that of cable-based electromagnetic communications, the main problem arising in characterizing the impact of the communication channel will be that of taking into account eventual near field effects due to the extremely small dimensions of the devices under exams.

Another issue of primary concern is that of the classification of the disturbance, which has to be first individuated and then classified as an additive, or multiplicative one, or even having different effect than additive or multiplicative as we are usual to consider in the wireless macroscopic case.

Once stated what is meant by disturbance signal and what are its effects, and once characterized the communication channel, a very tricky problem will be that of (try to) evaluating the channel capacity, since the optimization framework and constraints will be rather non linear; nor it will be easy, at least in principle, to borrow some optimization and/or approximation strategies [11], [13] from the usual linear framework of wireless communications. Capacity evaluation will also bring with it the need for identifying an optimal encoding strategy, in the sense of sets of concentration levels which can be realized in the communication medium, and so on. Finally, depending on the time scales of the different phenomena involved in nano communications, also outage events on the information rate will need to be defined, carrying over with them the distinction, if any in our new framework, between ergodic and non-ergodic regime we use to think of.

As previously said, the optical molecular communications case deserves a separate discussion, since, due to quantum effects, in such a scenario classical mechanic laws do actually fail. In this case, what is meant to be recast is the Von Neumann definition of entropy and information [14], [15], whose deep discussion is beyond the scope of this writing.

V. CONCLUSIONS

From the discussion presented in the previous sections, it is clear that many issues still need to be addressed in order to (i) understand the limiting performance of information communications among nanoscale devices, (ii) design optimal and quasi-optimal encoding/decoding strategies. Such issues are

believed to be of key relevance for allowing nanotechnologies display their full potential.

The first step necessary for moving forward towards a theory of communications among nanomachines consists in the characterization (by means of experimental measurements and/or computer simulations) of the statistical properties of the "channels" in such setting. This includes the characterization of the properties of electromagnetic field propagation in nanotubes/nanowires as well as the characterization of biological diffusion processes for wet nano techniques. Indeed, while it may already be possible to sketch some reference channel models based on an oversimplified view of the aforementioned processes (electromagnetic propagation, diffusion etc.), we believe that the availability of real-world measurements represents a key factor to building meaningful and consistent models. Another point which needs further investigation concerns the possible use of wet techniques, where information transmission can be based on the use of carrier molecules. Such approach presents indeed extremely appealing features in terms of energy consumption, reliability or robustness. Nonetheless, it remains to understand the impact of the possibly extremely slow propagation of molecules in aqueous environments on the system performance.

REFERENCES

- [1] R. Feynman, "There's plenty of room at the bottom," 1959, American Physical Society Annual Meeting. [Online]. Available: <http://www.zyvex.com/nanotech/feynman.html>
- [2] G. Moore, "Camming more components onto integrated circuits," *Electronics Magazine*, 1965. [Online]. Available: ftp://download.intel.com/museum/Moores_Law/Articles-Press_Releases/Gordon_Moore_1965_Article.pdf
- [3] M. Chang, V. Roychowdhury, L. Zhang, H. Shin, and Y. Qian, "RF/Wireless interconnect for inter-and intra-chip communications," *Proc. of the IEEE*, vol. 89, no. 4, pp. 456–466, Apr. 2001.
- [4] "Nanotechwire project at Rice University." [Online]. Available: <http://www.nanotechwire.com>
- [5] T. Suda, M. Moore, T. Nakano, R. Egashira, A. Enomoto, S. Hiyama, and Y. Moritani, "Exploratory research in molecular communication between nanomachines," UCI, Tech. Rep. Technical Report 05-03, 2005. [Online]. Available: <http://netresearch.ics.uci.edu/mc/publications.html>
- [6] Y. Moritani, S. Hiyama, and T. Suda, "Molecular communication among nanomachines using vesicles," in *Proc. NSTI Nanotechnology Conference*, 2006.
- [7] S. Hiyama, Y. Isogawa, T. Suda, Y. Moritani, and K. Suto, "A design of an autonomous molecule loading/transporting/unloading system using DNA hybridization and biomolecular linear motors in molecular communication," in *Proc. European Nano Systems*, 2005.
- [8] W. J. Dally and B. Towles, "Route packets, not wires: On-chip interconnection networks," in *Proc. of DAC*, 2001.
- [9] L. Benini and G. D. Micheli, "Networks on chip: a new SoC paradigm," *IEEE Computer*, vol. 35, no. 1, January 2002.
- [10] D. Minoli, *Nanotechnology applications to telecommunications and networking*. New York: J. Wiley & Sons, 2005.
- [11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, 1991.
- [12] P. Harremoës and F. Topsøe, "Unified approach to optimization techniques in Shannon theory," in *Proc. of IEEE Int. Symp. Information Theory*, Lausanne, CH, 2002.
- [13] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge: Cambridge University Press, 2004.
- [14] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press, 2000.
- [15] C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Quantum Mechanics, Vols. 1, 2*. Hermann: J. Wiley & Sons, 1977.