Nonlocal quantum information transfer without superluminal signalling and communication

J. Walleczek · G. Grössing

Received: date / Accepted: date

Abstract It is a frequent assumption that – via superluminal information transfers – superluminal signals capable of enabling communication are necessarily exchanged in any quantum theory that posits hidden superluminal influences. However, does the presence of hidden superluminal influences automatically imply superluminal signalling and communication? The non-signalling theorem mediates the apparent conflict between quantum mechanics and the theory of special relativity. However, as a ‘no-go’ theorem there exist two opposing interpretations of the non-signalling constraint: foundational and operational. Concerning Bell’s theorem, we argue that Bell employed both interpretations at different times. Bell finally pursued an explicitly operational position on non-signalling which is often associated with ontological quantum theory, e.g., de Broglie–Bohm theory. This position we refer to as “effective non-signalling”. By contrast, associated with orthodox quantum mechanics is the foundational position referred to here as “axiomatic non-signalling”. In search of a decisive communication-theoretic criterion for differentiating between “axiomatic” and “effective” non-signalling, we employ the operational framework offered by Shannon’s mathematical theory of communication. We find that an effective non-signalling theorem represents two sub-theorems, which we call (1) non-transfer-control (NTC) theorem, and (2) non-signification-control (NSC) theorem. Employing NTC and NSC theorems, we report that effective, instead of axiomatic, non-signalling is entirely sufficient for prohibiting nonlocal communication. Effective non-signalling prevents the instantaneous, i.e., superluminal, transfer of message-encoded information through the controlled use – by a sender-receiver pair – of informationally-correlated detection events, e.g., in EPR-type experiments. An effective non-signalling theorem allows for nonlocal quantum information transfer yet – at the same time – effectively denies superluminal signalling and communication.

Jan Walleczek (Corresponding author)
Phenoscience Laboratories, Novalisstrasse 11, 10115 Berlin, Germany
E-mail: walleczek@phenoscience.com

Gerhard Grössing
Austrian Institute for Nonlinear Studies, Akademiehof, Friedrichstr. 10, 1010 Vienna, Austria
E-mail: ains@chello.at
Keywords quantum nonlocality · superluminal signalling · Shannon communication theory · de Broglie–Bohm theory · Bell’s theorem

PACS 89.70.+c · 03.30.+p · 03.65.Ta · 03.65.Ud

1 Introduction

Bell’s theorem proved that no quantum theory based on the joint assumptions of “causality and locality” can successfully reproduce the predictions that are yielded by orthodox quantum mechanics (Bell [1]). Consequently, for the case of EPR-type nonlocal correlations at space-like distances, Bell’s proof implies that the correlations are either (1) beyond physical explanation, at least in locally causal terms, or alternatively (2) the correlations might be explained by physical processes that are, for example, governed by nonlocal, i.e., superluminal, causal influences. John Bell himself was dissatisfied with the first option of no explanation and famously remarked “that correlations cry out for explanation” (Bell [2]). Bell’s own pursuit of possible causal explanations in quantum mechanics was reviewed by Norsen [3] whose analysis confirmed that “Bell uses the term ‘causality’ . . . to highlight that a violation of this [local causality] condition (by some theory) means that the theory posits non-local causal influences, as opposed to mere ‘non-local correlations’.” However, if one were to take seriously the explanatory option that “causal influences do go faster than light”, then – Bell [2] noted – one should find disturbing “the impossibility of ‘messages’ faster than light, which follows from ordinary relativistic quantum mechanics . . . ”. He concluded that, therefore, for anyone proposing an approach to quantum mechanics based on nonlocal, i.e., superluminal, influences, the “exact elucidation of concepts like ‘message’ and ‘we’ . . . would be a formidable challenge” (Bell [2]). As we will discuss at length further below, Bell here refers to the special, and still incompletely understood, ‘role of us’ – human observers and of epistemic agents in general – both in the performance of quantum-based experiments and in possible definitions of the non-signalling constraint (e.g., Eberhard [4]).

Despite these challenges, Bell shared an ongoing interest in de Broglie–Bohm theory even long after publishing his seminal proof (e.g., Bell [5–8]). David Bohm’s non-standard formulation of quantum mechanics is well known for positing nonlocally-causal influences as a fundamental, ontic feature of physical reality (Bohm [9,10]). “This picture, and indeed, I think, any sharp formulation of quantum mechanics”, Bell explained in reference to Bohm’s theory, “has a very surprising feature: the consequences of events at one place propagate to other places faster than light. This happens in a way that we cannot use for signalling” (Bell [7]). The impossibility to signal by way of superluminal influences Bell took to be the central issue for any explanatory, causal approaches to quantum mechanics (Bell [7]): “For me then this is the real problem with quantum theory: the apparently essential conflict between any sharp formulation and fundamental relativity.” The fact that – 50 years after Bell’s theorem – this conflict remains unresolved would likely not have surprised Bell who had predicted decades ago (Bell [7]): “It may be that a real synthesis of quantum and relativity theories requires not just technical developments but radical conceptual renewal.”

The present work offers a communication-theoretic analysis of the conceptual impasse that exists between (1) the possibility of superluminal influences and (2) the impossibility of superluminal signalling as required by special relativity: Does the presence of superluminal
influences necessarily imply superluminal signalling and communication? We will present an answer based on an informational approach in reference to Shannon's mathematical theory of communication (Shannon [11]; Shannon and Weaver [12]). Specifically, the present work introduces a conceptual framework for defining, in a technically consistent manner, the difference between signalling, information transfer, and message communication. These concepts are sometimes used interchangeably and often without clear definition in the literature on quantum foundations. Particularly, we suggest that in discussions of the non-signalling constraint, and of the relationship between signal transmission, causal influences, and information transfers, this lack of definition is largely responsible for the articulation of conflicting positions.

1.1 Does nonlocal information transfer automatically imply superluminal signalling?

There are many instances in the scientific literature where the concept of information transfer is identified directly with the concept of signalling and communication without justification based upon sound physical principles. Concepts like ‘hidden signalling’ and ‘hidden communication’ have been employed, for example, in negative assessments of ontological quantum theories without plausible communication-theoretic definitions of terms: What do concepts like ‘hidden communication’ represent in a technical sense? What is the physical meaning of the term ‘hidden’? What is the difference between signalling and ‘hidden’ signalling? What is the difference, if any, between signal transfer and information transfer, whether hidden or not hidden? For example, the fundamental nonlocality of de Broglie–Bohm theory was interpreted by Scarani and Gisin [13] as “superluminal hidden communication between correlated particles.” More recently, Scarani et al. [14] have argued that (certain!) quantum models involving “hidden superluminal influences” are “in flagrant contradiction with relativity”, because the models “allow for faster-than-light communication between users”, and Gallego et al. [15] have claimed that “... Bohm’s theory is both deterministic and able to produce all quantum predictions, but it is incompatible with no-signalling at the level of hidden variables.”

These are only few of the many examples in the literature implying that de Broglie–Bohm theory might be incompatible with the non-signalling theorem and thus with special relativity, and that therefore Bohm’s theory should be physically unrealistic. Obviously, that proposition contradicts the view held by those who are in support of the theoretical possibility of de Broglie–Bohm theory, including the view held by John Bell (e.g., Bell [5–8]). While apparently persuasive arguments have been advanced on either side of this long-standing debate, a new conceptual foundation may be required to be able to move beyond entrenched positions. Importantly, we suggest that the conflicting positions can often be traced to the singular fact that contradictory interpretations of the non-signalling theorem have been used: a foundational, axiomatic interpretation and an operational, effective interpretation. Briefly, whereas “axiomatic non-signalling” completely disregards the role of scientific observers, i.e. of epistemic agents, an “effective non-signalling” constraint, as described, analyzed, and defined here, takes into account the essential role of epistemic agents in communication and signalling processes.

Better insight into the difference between concepts like ‘nonlocal influence’ and ‘nonlocal signalling’ is needed also in light of the following significant development: Harrigan and Spekkens [16] defined the distinction between $\psi$-ontic and $\psi$-epistemic approaches to
quantum theory. That distinction sparked a new wave of work drawing attention again (i) to the question concerning the reality of the quantum state (e.g., Colbeck and Renner [17]; Pusey et al. [18]; Ghirardi and Romano [19, 20]), and (ii) to possible contributions ontological theories might make to a future understanding of quantum theory. For an extensive review and analysis of current developments at this research frontier see Leifer [21]. Not surprisingly, a key question that has resurfaced in this context is Bell’s original question as to why superluminal influences cannot be used to transmit superluminal signals. For example, Wood and Spekkens [22] reconsidered the possibility of “quantum causal” explanations, including the problem of why $\psi$-ontic quantum theories such as de Broglie–Bohm theory appear to depend on the strict “fine-tuning” of causal parameters unless such theories are permitted to violate the non-signalling theorem.

What then is the valid interpretation of the non-signalling theorem? Looking ahead, the conceptual framework offered by Shannon’s theory of communication processes allows us to distinguish between two types of signals. That distinction may provide workers in quantum foundations with a fresh approach towards analysing the difference between an axiomatic as opposed to an effective non-signalling theorem. The respective signals we will identify as (communication-theoretic) ‘Shannon signals’ and (signal-theoretic) ‘non-Shannon signals’. Summarizing, an axiomatic non-signalling theorem denies transmission of Shannon and non-Shannon signals alike. Instead, a theorem of effective non-signalling only denies transmission of Shannon signals but does not constrain the transfer of non-Shannon signals. Before returning to the above questions in later sections, we will next present an overview of the contrasting uses of the non-signalling theorem in quantum mechanics, employing as a historical reference the two different interpretations used by John Bell.

2 The two interpretations by John Bell of the non-signalling theorem

The non-signalling theorem is widely agreed to represent a general ‘no-go’ theorem in quantum mechanics. However, as was mentioned already, there exist two opposing interpretations of the non-signalling theorem: a foundational interpretation as opposed to an operational one. As a ‘no-go’ theorem in quantum mechanics, how is the non-signalling constraint to be validly interpreted? Again we can turn to Bell for insight and guidance because, as we will explain, he employed both interpretations – foundational and operational – at different times. First, in Sect. 2.1, we will argue that the tradition of tacitly identifying potential ontic influences or causes with signalling, in the context of the predictions of Bell’s theorem, started with Bell himself (Bell [1]). This is the position we have introduced above as “axiomatic non-signalling”. Second, in Sect. 2.2, we will show that – in years following publication of Bell’s theorem – Bell took up an explicitly operational interpretation of the non-signalling theorem, an interpretation that views signals and messages as operationally distinct from either causes or influences. The operational position we have identified above as “effective non-signalling”, in contrast to “axiomatic non-signalling”, which represents the orthodox position to be described first.
2.1 Foundational, axiomatic interpretation of the non-signalling theorem

In his celebrated paper of 1964, Bell tacitly adopted a *foundational* (axiomatic) interpretation of non-signalling (Bell [1]): “In a theory in which parameters are added to quantum mechanics to determine the results of individual measurements, without changing the statistical predictions, there must be a mechanism whereby the setting of one measuring device can influence the reading of another instrument, however remote. Moreover, the signal involved must propagate instantaneously, so that such a theory could not be Lorentz invariant.” Evidently, Bell here takes the fact that “the setting of one measuring device can influence the reading of another instrument…” as evidence for instantaneous, i.e., *nonlocal signalling* – “the signal involved must propagate instantaneously”. The possibility in a theory of nonlocal, i.e., superluminal, signalling and communication is in direct conflict, of course, with special relativity, and indeed Bell notes that “such a theory could not be Lorentz invariant”. However, without further explanation, Bell [1] directly equates the superluminal influence with a superluminal *signal*.

We have shown before that loss of operational distinction between ‘influence’ and ‘signal’ is the mark of foundational, ontic, or *axiomatic*, interpretations – to use the present terminology – of the non-signalling theorem (Walleczek and Grössing [23]). There is wide agreement that the foundational interpretation, i.e., axiomatic non-signalling, is the interpretation associated with orthodox quantum theory. Before continuing with the description of the lesser known, operational interpretation of non-signalling in Sect. 2.2, we will review briefly how the concept of axiomatic non-signalling has recently been employed as an *apparently* conclusive argument against both the possibility of determinism and nonlocal hidden-variables approaches in quantum mechanics.

2.1.1 Axiomatic non-signalling as an argument against determinism in quantum mechanics

An axiomatic non-signalling concept was employed in attempts to question the viability of *any* kind of deterministic approaches to quantum mechanics, challenging the possibility of de Broglie–Bohm theory for example (e.g., Colbeck and Renner [24, 25]; Gallego et al. [15]). We have previously referred to such attempts as *generalizations* of Bell’s theorem (Walleczek and Grössing [23]). The question concerning determinism has also come up again in the context of *ψ*-ontic and *ψ*-epistemic approaches to quantum theory (e.g., Harrigan and Spekkens [16]; Leifer [21]): Can the unpredictability of EPR-type nonlocal correlations count as conclusive evidence in favour of the existence in nature of objective chance, i.e., of *intrinsic* randomness? On the one hand, the objective nature of quantum indeterminism has long been taken for granted and thus represents a key metaphysical assumption of orthodox quantum theory. On the other hand, it must be acknowledged that proof of absolute indeterminism is impossible as a matter of principle – whether by experiments or by mathematical analyses. The impossibility-of-proof argument was addressed by us before and this argument will not be restated here (see Walleczek and Grössing [23]). In the hope of by-passing these fundamental experimental and mathematical constraints, it was argued by others that – nevertheless – there might still be a way to decide between the two competing assumptions: determinism or indeterminism? Specifically, it was claimed that the assumption of (axiomatic) non-signalling suffices to eliminate the possibility of determinism at the level of the quantum (e.g., Colbeck and Renner [24, 25]; Gallego et al. [15]). In response to that claim, we have noted before
Figure 1 Relational diagram illustrating the irreducible interdependency of basic metaphysical assumptions implicit in standard interpretations of orthodox quantum theory (adapted from Walleczek and Grössing [23]). (A) Free choice assumption, (B) Intrinsic randomness assumption, and (C) Axiomatic non-signalling assumption. Crucially, the validity of interpreting the non-signalling theorem as an ontic, foundational theorem, or axiom, for quantum mechanics, a frequent assumption in standard interpretations, depends on the independent validity of assumptions (A) and (B). However, neither assumption (A) nor assumption (B) can be independently confirmed if the possibility of ‘free choice’ depends on the existence of a process that is intrinsically random and vice versa.

that the possibility of determinism cannot be denied on account of non-signalling because the claim rests on the independent validity of three interdependent assumptions (Walleczek and Grössing [23]). Fig. 1 illustrates the relational interdependency of the three assumptions that underlie the reasoning behind axiomatic non-signalling as a suggested proof of quantum indeterminism, i.e., of intrinsic randomness, in nature.

We suggest that there does not exist at present a conclusive argument derived from an axiomatic non-signalling assumption which is capable of the successful generalization of Bell’s theorem (for details see Walleczek and Grössing [23]). A related but not identical argument was previously offered by Ghirardi and Romano [19]. To repeat, it remains undecidable, on the basis of logical considerations alone, which of the opposing metaphysical positions is valid – indeterminism or determinism (see Fig. 1). In the following, we will explain that Bell distanced himself from the foundational, axiomatic interpretation and that he started to adopt an operational approach in line with a theorem of effective non-signalling (Bell [26]).

2.2 Operational, effective interpretation of the non-signalling theorem

In the years following publication of his ground-breaking theorem, Bell introduced work that recognized and upheld in quantum theory the operational distinction between ‘influences’ or ‘causes’, and ‘signals’ or ‘messages’. Specifically, Bell considered the concept of potential ontic influences in quantum mechanics, i.e., the concept which he introduced as ‘beables’ by way of contrast to the standard concept of ‘observables’, which refers to (epistemic) states of knowledge only (Bell [26]). Significantly, these (potentially ontic) ‘beables’, he classified into “controllables”, and “uncontrollables”, whereby he noted that the “latter are no use for sending
signals..." (Bell [26]). Note that Bell’s distinction between “controllables”, and “uncontrol-
lables” does not inherently depend on the distinction between nonlocal and local beables, or
whether beables are unpredictably deterministic or intrinsically random. In summary, starting
in 1976, Bell’s own work introduces an operational interpretation of the non-signalling con-
straint, i.e., to use our present terminology, Bell shifts his attention from an axiomatic, towards
an effective, non-signalling concept (e.g., Bell [5–8, 26, 27]).

Lately, there has been a resurgence of interest in Bell’s own interpretation of his theorem
before and after the year 1976. Wiseman [28, 29] recently brought to the attention of the sci-
entific community “The two Bell’s theorems of John Bell”. The shift in Bell’s own thinking
(as implied by the proposal of two Bell’s theorems) could be accounted for – at least in part –
by a novel perspective on the non-signalling constraint, a view that is offered by Wiseman’s
analysis also (Cavalcanti and Wiseman [30]; Wiseman [29]). We suggest that the transition
from Bell’s position of 1964 [1] to that of 1976 [26] represents a movement towards an opera-
tional approach which specifically seeks to account for the effective role of epistemic agency,
i.e., agent control based upon agent knowledge, as a key factor in the construction of the non-
signalling constraint (see Sect. 4.2). Again, as was mentioned already, the agent-based view of
non-signalling is well-exemplified by Bell’s own assertion, that “we cannot use for signalling
the way in which “events at one place propagate to other places faster than light” (Bell [7]).
Obviously, Bell’s new view of (effective) non-signalling is in stark contrast to the original
position of (axiomatic) non-signalling (Bell [1]). The detailed exploration, however, of the
potentially far-reaching consequences of this change in perspective has only begun relatively
recently. In particular, Norsen [3, 31] and Maudlin [32] have contributed greatly to the clarifi-
ation of frequent misconceptions in the scientific literature about Bell’s own views regarding
the possibility of superluminal causation in the context of an operational interpretation of the
non-signalling constraint.

Historically, the number has been growing of researchers who have argued – in one form
or another, directly or indirectly – for the validity of an agent-based view of the non-signalling
theorem in line with Bell’s notion of operationally “uncontrollables” of 1976 (e.g., Bohm and
Hiley [33]; Holland [34]; Valentini [35]; Norsen [3, 31]; Maudlin [32]; Cavalcanti and Wise-
man [30]; Ghirardi and Romano [19, 20]; Walleczek and Grössing [23]; Wiseman [28, 29]).
Note that this list is far from complete as more investigators have devoted time to this impor-
tant issue. However, what has so far been lacking is an understanding of the communication-
theoretic difference between axiomatic and effective approaches beyond the statement that
agent participation is presumed in the effective approach. Therefore, the present work seeks
to identify a decisive technical criterion for defining the difference between “axiomatic non-
signalling” and “effective non-signalling” in the context of communication theory.

2.2.1 An effective non-signalling theorem accounts for the controlling actions of epistemic
agents

That Bell saw as indispensable the special role of the experimenter agent in reaching a full
understanding of the non-signalling theorem is amply evident in Bell’s last published statement
on this matter (Bell [8]): “Do we then have to fall back on ‘no signalling faster than light’ as
the expression of the fundamental causal structure of contemporary theoretical physics? This
is hard for me to accept. For one thing we have lost the idea that correlations can be explained,
or at least this idea awaits reformulation. More importantly, the ‘no-signaling . . . ’ notion rests on concepts which are desperately vague, or vaguely applicable. The assertion that ‘we cannot signal faster than light’ immediately provokes the question: Who do we think we are? We who can make measurements, we who can manipulate ‘external fields’, we who can signal at all, even if not faster than light? Do we include chemists, or only physicists, plants, or only animals, pocket calculators, or only mainframe computers?”. Note, that – by contrast – the consideration of who or what qualifies as an epistemic agent, or whether agents may, or may not, access controllably information transfers – plays no role at all in a theorem describing axiomatic non-signalling. We will return to the concept of knowledge-based agents in quantum models expressing effective non-signalling during the main part of our analysis in Sect. 4.

The availability of a communication-theoretic criterion for distinguishing an effective from an axiomatic non-signalling theorem appears to be essential for answering the following questions: If EPR-type nonlocal correlations are not intrinsically random but (unpredictably) deterministic instead, then why can’t we use them to transmit superluminal signals? What prevents superluminal causal influences from being a pathway for hidden communication and signalling? And importantly, do superluminal information transfers necessarily violate special relativity even though these transfers are beyond control by agents for the purpose of signalling and communication? To frame more clearly these quantum-foundational questions we adopt for our analysis the following method.

3 An informational approach towards the concept of hidden superluminal influences

One major reason as to why many researchers insist on the foundational, or axiomatic, interpretation of the non-signalling theorem often comes from the following understanding: The presence of hidden superluminal influences would necessarily imply the presence of (hidden) information transfers which in turn would imply signal transfer and the possibility of communication. It is a frequent assumption that – via superluminal information transfers – superluminal signals capable of enabling communication are necessarily exchanged in any quantum theory that posits superluminal influences (compare Sect. 1.1). As a consequence, relativity theory would be violated which would immediately render an ontological quantum theory, like de Broglie–Bohm theory, physically unrealistic. The analysis provided in Sect. 4 will explore whether or not this understanding is justified by definitions of signalling, information transfer, and message communication, based upon Shannon’s theory of communication processes (Shannon [11]; Shannon and Weaver [12]).

The ensuing analysis will not only be consistent with the informational approach towards analysing various consequences of proposed superluminal (causal) influences but our analysis will strictly rely on that approach to draw its final conclusions. That is, we assume from the start that “hidden superluminal influences” (e.g., Scarani et al. [14]), and the possibility of nonlocally-causal transfers (e.g., Bohm [9, 10]), invariably involve information transfers or exchanges. For example, Bohm and Hiley [33] used the term “active information” to describe such nonlocal exchanges in the context of nonlocal hidden-variables approaches. We here take a 1-bit informational event to be the minimal indication for the occurrence of any kind of causal exchange, or any discernible physical influence. More precisely, without detection of nonlocally-correlated informational events – between two space-like separated members of an entangled pair – evidence for any (potentially ontic) influence between pair members
would be entirely unavailable. Evidently, then, the central question here considered reduces to this: Must information transfer always imply signalling and communication? Specifically, how could signal exchanges between two systems be denied if – at the same time – unrestricted informational exchanges are allowed between them? Following the detailed analysis of these questions in reference to Shannon’s theory in Sect. 4, we will consider the application of the subsequent definitions in the context of an effective interpretation of the non-signalling theorem for quantum mechanics in Sect. 5.

4 Information transfer, signal transfer, and message communication, in reference to Shannon’s mathematical theory of communication

To begin with, we informally define two theorems that are implicit in Shannon’s theoretical framework, i.e., theorems that remain unstated usually because they appear to be self-evident. For the present work, however, we make explicit these theorems by naming them: (1) a theorem of information transfer control (ITC), and (2) a theorem of information signification control (ISC). The effective role and application of these theorems in the context of constructing an effective non-signalling theorem in relation to quantum mechanics will be described in Sect. 5.

Importantly, these normally implicit theorems provide the larger operational context without which Shannon’s familiar explicit theorems and mathematical measures such as Shannon’s source coding theorem or Shannon’s channel information capacity, C, could not be usefully applied in any practical manner (Shannon [11]). In other words, the normally implicit theorems account for the involvement of epistemic agents who originate, encode, send, receive, and decode, signals. When assessing – in the context of communication theory – the relationship between the non-signalling theorem, quantum theory, and special relativity, we propose that Shannon’s implicit theorems provide a consistent operational framework, because ITC and ISC theorems account specifically for agent participation during signalling and communication processes. Subsequent Sect. 4.1 will provide an explanation of essential differences between (i) Shannon’s operational framework which includes, for example, agent-dependent encoding-decoding processes as accounted for by the ISC theorem, and (ii) Shannon’s mathematical framework which is independent of agent participation.

4.1 Shannon’s information-theoretic approach to human and automated machine communication

We next introduce Shannon’s well-known concepts (Shannon [11]): “The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point.” For Shannon, any possible communication starts with an “information source which produces a message or sequence of messages to be communicated to the receiving terminal”. More precisely, the message to be communicated originates with a ‘discrete’ information source. “We can think of a discrete source”, Shannon [11] explained, “as generating the message, symbol by symbol. It [the source] will choose successive symbols according to certain probabilities depending, in general, on preceding choices as well as the particular symbols in question.”
Figure 2  Illustration of basic concepts from Shannon’s information-theoretic approach to sender–receiver systems in human and automated machine communication (Shannon [11]). For easy overview, we describe key functional elements that operationally define Shannon’s view of sender-receiver systems, employing his original definitions (Shannon [11]): (1) An epistemic agent, here named ‘Alice’, is the “source which produces a message”, e.g. the three-letter message “SOS”; (2) an encoding transmitter “which operates on the message in some way to produce a signal suitable for transmission” by performing – here in (binary) Morse code – “an encoding operation which produces a sequence of dots, dashes, and spaces . . . corresponding to the message”; (3) a conduit or channel, i.e., “the medium used to transmit the signal from transmitter to receiver”; (4) a decoding receiver which “performs the inverse operation of that done by the transmitter, reconstructing the message from the signal”; finally (5), an epistemic agent, here named ‘Bob’, is the destination “for whom the message is intended”. We label combination of elements (1) and (2), i.e., the epistemic agent and the transmitter, a “sender” system, or simply the “sender”. Accordingly, we label combination of elements (4) and (5), i.e., the receiver and the epistemic agent, “receiver system”, or simply the “receiver”. The bi-directionality of arrows in the figure serves to indicate that the two systems, “sender” and “receiver”, each may perform the functions of the other leading to the possibility of bidirectional message communication.

To illustrate the concept of Shannon’s ‘discrete source’, Fig. 2 shows the scenario of bidirectional communication of the message “SOS” – in Morse code – between two sources here represented by sender Alice and receiver Bob. It is apparent that Shannon’s discrete source for “generating the message”, e.g., Alice in Fig. 2, manifests operational control over the process of signal transmission through the discrete channel. Shannon defined a ‘discrete channel’ as a “. . . system whereby a sequence of choices from a finite set of elementary symbols $S_1, . . . , S_n$, can be transmitted from one point to another.” The sequence “• • • − − − • • •” shown in Fig. 2 was chosen “symbol by symbol” by Shannon’s information source, e.g., epistemic agent Alice, from the “set of elementary symbols” offered by Morse code (●, −) to communicate the message “SOS”. The point is that a random sequence of symbols could not – of course – effectively transmit the message “SOS” between Alice and Bob.

In contrast to the scenario in Fig. 2, which illustrates Shannon’s operational framework, Shannon’s mathematical framework for describing channel information capacity, $C$, is founded upon an entirely different assumption about the properties of an information source. “How is an information source to be described mathematically”, Shannon [11] asked, and he introduced the concept of a random source to account quantitatively for the transmission of any arbitrary message from a set of possible messages. As was emphasized by Shannon [11], the meaning of an actual message, and thus of any “successive symbols”, is of no importance in calculating $C$: “Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is selected from a set of possible messages.”
In reference to Shannon’s distinction between actual message and possible messages, Fig. 3 represents two different types of ‘information source’ which we have named type-I and type-II; both types are respectively employed in Shannon’s model of communication (Shannon [11]).

The so-called “engineering problem” refers to the mathematical account of channel information-carrying capacity for theoretically possible messages. By contrast, the problem which is the focus of the present work concerns the operational account of how an actual message is transmitted from sender to receiver such as the message “SOS” in Fig. 2. Importantly, each type of problem – actual message versus possible messages – elicits different properties of an ‘information source’ as illustrated in Fig. 3. In Shannon’s operational framework of actual message communication the participation of epistemic agents who represent type-I sources manifesting operational control is strictly required (Fig. 3). That is, type-I agents choose – to use Shannon’s words again – “successive symbols… depending, in general, on preceding choices as well as the particular symbols in question” (compare Fig. 2). It is evident that Shannon’s use of the term ‘choice’ in that context does not refer to a random process. Instead, in the context of generating an actual message, Shannon’s ‘information source’ represents a source of operational control (see ‘type-I source’ in Fig. 3 and Sect. 4.2). The corresponding physical signals, i.e., those passing through the channel under type-I agent control, we from now on will refer to by the new term ‘Shannon signals’ (see Sect. 4.3 for details). Importantly, the above-mentioned ITC and ISC theorems constrain only the transmission of such Shannon signals, i.e., the signals that deliver actual messages (see Sects. 4.3.1 and 4.3.2).

By contrast, for the mathematical description of the “set of possible messages” Shannon’s information source represents a source of random variables (see ‘type-II source’ in Fig. 3). The type-II source represents – by Shannon’s description – a “stochastic process” for the purpose of producing probabilistically the set of possible messages in the quantification of channel capacity (Shannon [11]). Importantly, unlike the assumptions A and B described for Fig. 1 (see Sect. 2.1.1), the (type-II) source of random variables in Shannon’s theory can be deterministic. An example of the physical instantiation of a type-II source would be the tossing of a fair coin revealing either ‘heads’ or ‘tails’, which is a process that is deterministic yet operationally unpredictable.
In summary, the present work discusses solely the physical operations that facilitate actual message communication between epistemic agents, i.e., the controlling actions of agents described by type-I operations (see Fig. 3). An effective non-signalling theorem must prohibit the effective performance of Shannon’s type-I source or agent, i.e., the transfer of Shannon signals only – to use the above introduced terminology. It is important to note again that there cannot be – in actuality – ‘random messages’ or random acts of actual communication between sender and receiver; there may only exist random information transfers as will be described in Sect. 4.3. Finally, Shannon’s definitions, some of which are cited above and in the legend to Fig. 2, emphasize that the possibility of signal transmission for the purposes of actual message communication relies on (approximate) prediction and control of relevant physical processes by epistemic type-I agents (see Fig. 3). Therefore, we hold that a complete analysis concerning the concept of signalling, and of its negation, i.e., non-signalling, must take into account the involvement of epistemic agents. Who or what is the epistemic agent in relation to Shannon’s theory, Bell’s theorem, and the non-signalling constraint?

4.2 Defining the epistemic agent

As was mentioned in Sect 2.2.1, John Bell was keenly aware of the importance in interpretations of the non-signalling theorem of having an understanding of agency. Who or what represents an (epistemic) agent? Remember that Bell once asked whether we should include in our definition “… chemists, or only physicists, plants, or only animals, pocket calculators, or only mainframe computers?” (Bell [8]). We will here introduce definitions that are intended primarily to clarify the meaning of his question, and that may direct us towards possible answers in the spirit of Bell’s original inquiry. Agency is generally defined as the capacity of humans or other entities to act in the world. Put differently, an agent is defined initially by possessing the capacity to influence causal flows in nature. By prefacing “agent” with the term “epistemic”, attention is drawn to the fact that a complete definition of agency represents more, of course, than the mere “capacity to influence causal flows”: an agent possesses knowledge-based, i.e., epistemic, capacity for predictably directing, and redirecting, causal flows, and thus for directing, and redirecting, information flows as well. That is, an epistemic agent holds the power to (statistically) control physical activity based upon an ability to predict the outcome of specific actions on targeted processes in reference to a known standard or goal. In short, an epistemic agent thus manifests in the world a genuine source of operational control (see Fig. 3). By the definition here introduced, Alice and Bob in Fig. 2 represent type-I agents who each may plausibly represent and enact in Shannon’s sender-receiver system an “information source which produces a message” (see Fig. 3).

4.2.1 Epistemic agents: from human observers to autonomous computational devices

An epistemic agent is here defined as a “source of operational control” (Fig. 3). An epistemic agent might be the experimenter who designs an experiment for the purpose of asking specific questions of nature. An epistemic agent may formulate a scientific hypothesis to be tested. An epistemic agent may be the scientific observer who operates an experimental apparatus, selects measurement settings, and collects and interprets measurement outcomes. Naturally,
the element of ‘control’ is critical for all these dimensions of human activity. The term ‘operational control’ would be without meaning in the absence of an epistemic agent manifesting a capacity to distinguish between possible alternative outcomes. More precisely, an agent manifests an expectation about some future outcome in relation to some known reference (e.g., goal-directedness). Importantly, it makes no difference in that regard whether an epistemic agent is defined as “man or machine” (see also Cavalcanti and Wiseman [30]). For explanation, (artificial) physical devices are routinely designed and built for the purpose of performing ‘controlling activity’ automatically. Autonomous controlling activity could be as simple as due to the selection of a specific set point in a feedback-regulated thermostat controlling room temperature, or – vastly more complex – the programming of computational algorithms for generating, sending, and receiving, messages in automated telecommunication systems. In any case, however, beyond artificial devices and living systems the use of the term ‘operational control’ cannot be justified, and thus the term should not be used independently of epistemic agents, whether technical or human. Finally, it is apparent that signal communication, in the sense of Shannon’s theory of communication, requires the involvement of epistemic type-I agents who can originate, encode, send, receive, and decode, signals (compare Fig. 3). What exactly constitutes “signalling”, and “non-signalling”, in the context of the present communication-theoretic investigation of the non-signalling theorem?

4.3 The communication-theoretic distinction between Shannon signals and non-Shannon signals

We have before introduced the new concept of ‘Shannon signal’ in Sect. 4.1. What exactly constitutes a ‘signal’? Unless one finds agreement first on what represents a signal, one cannot later expect to have agreement on an appropriate definition of ‘non-signalling’, i.e., negation of signalling. We next distinguish between two types of signals that are apparent in the context of Shannon’s theory: (1) signals in the familiar sense of standard signal theory, and (2) signals in the sense of communication theory only. The latter type of signals we have referred to as “Shannon signals” whereas the former we will refer to as “non-Shannon signals”. For explanation, the notion of non-Shannon signal represents a signal in the standard (signal-theoretic) sense of detecting a physical influence as part of measurement processes in general, independently of any relationship to an effective communication task. A simple engineering example is the manifestation of a “click” by a suitable threshold detector in response to physical stimulation. However, this is not what is exclusively meant by the term ‘signal’, or ‘signalling’, in the context of Shannon’s communication theory. There, the concept of signal also refers to the controlled delivery of an informational bit sequence (or, alternatively, a single bit of information) which has been subjected to a process of message encoding (by Shannon’s ‘encoding transmitter’; see Fig. 2). In agreement with that use in Shannon’s theory, we informally define ‘signal’ in the communication-theoretic sense as “controllably-transmitted and message-encoded information”. To repeat, we call the specifically communication-theoretic signal type: Shannon signal. By way of contrast, this characterizes the standard, i.e. non-Shannon, signal as an informational event that is operationally non-controlled, either in terms of transfer control (ITC), or signification control (ISC), or both (ITC + ISC). Summarizing, a non-Shannon signal represents a signal in the well-known sense of standard signal theory, whereas the newly
Figure 4 Illustration of 9-bit information transfer comparing ‘signalling information transfer’ (A) and ‘non-signalling information transfer’ (B and C). The boxes to the far right of the figure indicate when is available (+) or not (–) to epistemic agents Alice and Bob the capacity of ISC (information signification control) or ITC (information transfer control) in a given scenario (A-C). Question marks shown inside smaller boxes (see B and C) indicate that – even though perfect informational correlations can be observed and recorded at Alice’s and Bob’s locations – the communication of the message “SOS” is effectively denied. For example, although Bob obtains perfect knowledge about the informational state of Alice’s transmitter upon simply observing the informational state of his receiver, he nevertheless is denied reception of any message from Alice (see B and C). The examples in B and C demonstrate that the mere fact alone of the availability to Bob and Alice of informational correlations at their respective locations – as part of some communication system – need not at all indicate the presence of a message or signal in the communication-theoretic sense. Importantly, this conclusion is entirely independent of the fact whether the involved information transmission channel would be represented by a quantum channel or a classical channel. In short, while every signal or message represents information, not every information represents a message or signal.

introduced concept of Shannon signal finds application only in the communication-theoretic context provided by Shannon’s model.

Next, two different scenarios will be discussed to explain the significance of the above introduced distinction in reference to Shannon’s model of signalling processes. The scenarios demonstrate negation of Shannon signal transmission while – at the same time – allowing non-Shannon signal transfers, i.e., uncontrolled information transfers. Fig. 4 illustrates a total of three separate scenarios (A-C), including the standard scenario described by Shannon [11] as a key reference (Fig. 4A). Figs. 4B and 4C illustrate the respective applications of the ISC theorem (Sect. 4.3.1) and the ITC theorem (Sect. 4.3.2).

4.3.1 The information-signification-control (ISC) theorem

In Shannon’s informational model of signalling processes, any signal sequence is composed of discrete and elementary informational (bit) units (Shannon [11]). Crucially, single bits or bit sequences have no meaning “in-themselves”, that is, they do not represent Shannon signals, or messages, independently of encoding and decoding processes as implemented by epistemic agents. As was reviewed above, Shannon’s communication theory assigns “an encoding operation” to the transmitter (see “2” in Fig. 2), and to the receiver assigns a decoding operation, i.e., “the inverse operation of that done by the transmitter” (see “4” in Fig. 2). It is these very
operations – signal encoding/decoding – which are necessary (but not sufficient!) to transform elementary information transfers into Shannon signal transfers. The process that grants shared meaning, i.e., shared semantic content, to basic syntactic elements, or signs, is known in semiotic theory as ‘signification’ (e.g., Short [36]). Generally, semiotic theory studies the relationship between signs and (the process of creating) meaning. For example, a code represents a rule for connecting signs, e.g., dots and dashes in the case of Morse code, to their intended meaning. We have introduced ‘signification’ as a technical term in order to accurately account for the process of assigning meaning to information by way of encoding and decoding mechanisms that could be shared between sender and receiver. The concept of ISC makes explicit for Shannon’s model the normally implicit theorem that operationally accounts for encoding and decoding processes (see also Introduction to Sect. 4).

To illustrate the role of signification control, the following comparison might be useful: Figs. 4A and 4B portray identical informational patterns (i.e., “•••−−−•••”). Importantly, only for the standard scenario discussed by Shannon, the 9-bit pattern represents a Shannon signal (i.e., here the signal which – in Morse code – conveys the message “SOS”; see Fig. 4A). By contrast, the structurally-identical 9-bit pattern, shown as part of the second scenario in Fig. 4B, represents a non-Shannon signal. The difference between the two scenarios is accounted for by the presence (Fig. 4A) or absence (Fig. 4B) of operational control by epistemic agents over encoding-decoding processes. Again, while full control over information signification is available to Bob and Alice in Shannon’s standard scenario (Fig. 4A), the scenario illustrated in Fig. 4B represents complete lack of ISC (compare boxes to the far right of Fig. 4). Put simply, epistemic agents Bob and Alice do not know Morse code in this example and they cannot therefore exchange actual messages using this coding system. Note that judging from an engineering perspective of the technical generation of informational order of syntactic, structural elements, the two scenarios may nevertheless be equivalent (see Fig. 4C for a scenario where this is not the case). The crucial point is the following: whereas Shannon’s standard scenario represents information transfer that is signalling (Fig. 4A), the scenario illustrated in Fig. 4B represents the case where the transfer of information is non-signalling in Shannon’s communication-theoretic sense. That is, (two-way) transfers of Shannon signals are denied until Alice and Bob each acquires operational control over the process which we refer to as signification.

4.3.2 The information-transfer-control (ITC) theorem

To briefly review the above, in the case of the scenario shown in Fig. 4B, unlike in the standard scenario in Fig. 4A, operational control by Bob and Alice over ISC (e.g., code or key sharing) was absent. By contrast, Fig. 4C illustrates the opposite scenario: control over information signification is available, whereas control over information transfers through the transmission channel is not, even on a statistical basis only (compare boxes to the right of Figs. 4B and 4C). To illustrate this scenario when agents Bob and Alice lack the power to control information transfers, the sequence “••−•−••−•” is shown as one possible, unpredictably emerging, informational pattern. This pattern represents, of course, the appearance of non-Shannon signals only, to apply the present terminology (Fig. 4C); naturally, any other informational bit sequence pattern could have also been selected instead to visualize the statistical uncontrollability of information transfers. Note also that uncontrollable processes may accidentally
generate the pattern “•••−−−•••” shown in Fig. 4B, and may thus potentially transmit a false-positive “signal” between Alice and Bob. In summary, the ITC theorem represents the negation of operational control by epistemic agents over information transmission (see also Introduction to Sect. 4).

5 Two sub-theorems represent an effective non-signalling theorem

In the past, as was reviewed briefly in the Introduction, an appreciation of the vital difference between the concept of information transfer and different concepts describing signal transfer and communication has often been lacking in assessments – for quantum theory – of the role of the non-signalling constraint (see Sect. 1.1). Are signal transfers always reducible to information transfers? Or is there more to signalling? Why is the identification misleading of “hidden communication”, “hidden messages”, or “hidden signalling”, with the concept of “hidden information transfer”, for example, in the context of de Broglie–Bohm theory? To address these questions, we now employ key findings from the previous section in the task of characterizing the non-signalling theorem as an effective instead of as an axiomatic theorem. The effective characterization represents, of course, a characterization based upon the negation of operational control by epistemic type-I agents (see Fig. 3). In accordance with the above communication-theoretic analysis, the complete negation of operational control includes two distinct aspects: (i) negation of ITC (see Fig. 4C), and (ii) negation of ISC (see Fig. 4B). Simply on that basis, we distinguish between two sub-theorems that jointly represent an effective non-signalling theorem, and we introduce them as: (1) Non-transfer-control (NTC) theorem (i.e., negation of ITC), and (2) Non-signification-control (NSC) theorem (i.e., negation of ISC). Importantly, in light of the hypothesis of quantum channels for instant information transmission, e.g., as postulated in de Broglie–Bohm theory, the following is obvious: while the NSC theorem must be theoretically accounted for in any complete description of an effective non-signalling theorem (compare Fig. 4B), it is immediately apparent also that sender and receiver could, in principle, share knowledge about any arbitrary coding system with each other; however, only classical channels could be used for that purpose. Concerning the NTC theorem, operational uncontrollability of information transmission via hypothetical quantum channels assures that nonlocal information transfers are entirely non-signalling – in the sense of denial of (communication-theoretic) Shannon signalling.

To repeat, NTC and NSC theorems negate Shannon signalling, i.e., actual message transfers, but they do not interfere with the transmission of non-Shannon signals (for definitions see Sect. 4.3). To show that negation only of Shannon signalling, but not negation of non-Shannon signal transfer, is the relevant concept for constructing an effective non-signalling constraint, the following comparison may advance insight: Evidently, the non-signalling theorem was not introduced to prohibit the manifestation by quantum detectors – in EPR-type experiments – of (nonlocally-correlated) measurement clicks that reveal informational correlations at space-like distances (e.g., Aspect et al. [37, 38]; Wehls et al. [39]; Tittel et al. [40]; Ursin et al. [41]; Giustina et al. [42]). Instead, we maintain that the non-signalling theorem was introduced to prevent the possibility of the instant, i.e., superluminal, transfer of message-encoded information through the controlled use – by a sender-receiver pair – of informationally-correlated detector clicks. While it is certainly true that each individual click represents a signal also in the ordinary signal-theoretic sense (i.e., in the sense of a non-Shannon signal as defined in
Nonlocal quantum information transfer without superluminal signalling and communication

Figure 5  The two interpretations by John Bell of the non-signalling theorem (for details see Sect. 2). The set of arrows to the left indicates the two contrasting interpretations of the non-signalling theorem – the foundational, axiomatic position (Bell [1]) and the operational, effective position (Bell [26]). The set of arrows to the right indicates the two sub-theorems that represent an effective non-signalling theorem in accord with the present analysis: (1) Non-transfer-control (NTC) theorem, and (2) Non-signification-control (NSC) theorem. Briefly, the NTC theorem accounts for the fact that (Shannon) signal transfer and actual message communication is disabled without operational control by epistemic agents over information transfer pathways (compare Fig. 4C). The NSC theorem accounts for the fact that lacking operational control over encoding-decoding processes, i.e., over the semantic content of transferred information, (Shannon) signal transfer and message communication is denied also (compare Fig. 4B).

Sect. 4.3), the detection of a click need not at all indicate the presence of a signal as defined in terms of “controllably-transmitted and message-encoded information”, i.e., in the sense of the communication-theoretic definition of (Shannon) signal. Consequently, it would be wrong – both in the context of Shannon’s theory and the non-signalling theorem – to simply use as synonyms ‘information’ and ‘signal’, and thus to directly identify information transmission with signal transmission and communication (see also legend to Fig. 4). In summary, of the two sub-theorems of an effective non-signalling theorem – NTC theorem and NSC theorem – each one represents a theorem concerning the negation of ‘operational control’ (see Sects. 4.3.1 and 4.3.2). Operational uncontrollability of potentially ontic influences, i.e., ‘beables’, Bell [26] took as the primary justification for adopting an effective non-signalling theorem, instead of an axiomatic one (Bell [1]). Fig. 5 presents an overview of the two interpretations by Bell of the non-signalling theorem and of their relationship to the here identified NTC and NSC theorems.

This work may represent an important step towards answering Bell’s pressing questions which were motivated by his sceptical assessment that “… the ‘no-signaling…’ notion rests on concepts which are desperately vague, or vaguely applicable” (Bell [8]). Again, Bell found disturbing the continuing lack of scientific understanding of the importance of us human observers, i.e., epistemic agents, as part of a definition of the non-signalling constraint. “The assertion that ‘we cannot signal faster than light’ ”, Bell had noted, as we mentioned before, “immediately provokes the question: Who do we think we are? We who can make measurements, we who can manipulate ‘external fields’, we who can signal at all, even if not faster than light?” (Sect. 2.2.1). The findings of this communication-theoretic study offer a direct response to Bell’s more than decade-long inquiry into the operational nature of the non-signalling constraint: We are the agents who can manifest control over (i) information transfers as well as
(ii) information signification but only to the extent that we are granted access to informational pathways and to shared symbolic representation (see also legends to Figs. 4 and 5).

We have shown that a foundational, axiomatic interpretation of the non-signalling theorem eliminates the distinction between Shannon and non-Shannon signals and thus ignores any generative role of epistemic agents in the establishment of signalling events in the communication-theoretic sense. As a consequence, the axiomatic position on non-signalling fails to address Bell’s questions such as about scientific observers “who can signal at all, even if not faster than light?” (Bell [8]). An effective interpretation of the non-signalling theorem, by contrast, acknowledges the pivotal difference between the respective signals, leading to the possibility of developing a communication-theoretic account of agent participation during signalling and non-signalling processes: an effective non-signalling theorem only limits (type-I agent-controlled) transmission of Shannon signals but does not constrain the transfer of non-Shannon signals which is independent of agent control.

Finally, having in hand now a single criterion for the consistent differentiation between axiomatic and effective interpretations of the non-signalling theorem, crucial topics to be explored in future work are the following: Is superluminal, yet non-signalling, information transfer, i.e., nonlocal transmission of non-Shannon signals, compatible with the demands of special relativity? Does the proposition of nonlocal information transfer automatically entail a space-time (metric) structure that ceases to be Lorentz invariant? Put differently, does the possibility of superluminal information transfer compromise special relativity, even though the possibility of superluminal signalling and communication, by way of instantaneous, i.e., nonlocal transfers, is fully denied? Such questions point towards the notorious problem of paradoxical consequences of hypothetical instantaneous transfers, or nonlocal influences, at-a-distance: Is it inevitable, however, that causal paradoxes must be generated automatically – as a function of nonlocal quantum information transfer – even in the complete absence of agent participation, i.e., if nature is left to herself? We suggest that the effective, operational position on non-signalling may negate the presumed inevitability of paradoxical consequences in association with quantum theories positing nonlocal information transfers. However, the pursuit of firm answers to these foundational questions must be deferred to future investigations. The subsequent discussion is intended to provide an initial orientation for upcoming work.

6 Discussion

The present work has defined minimal conditions for an operational non-signalling theorem, i.e., a theorem that allows for the possibility of nonlocal quantum information transfer, yet one that – at the same time – effectively denies nonlocal, i.e., superluminal, signalling and communication. The communication-theoretic definition of non-signalling calls attention to possible implications of Bell’s original proposal of an operationally-motivated non-signalling constraint for notions such as information transmission, signalling, and communication, in relation to quantum-entangled states. For example, our conceptual framework facilitates the translation of ideas and concepts – in the context of quantum-entangled information – between uncontrollability of nonlocal influences and the observation of nonlocal informational events. Furthermore, this framework may prove useful in work assessing Bell’s original notion of “uncontrollables” within the scope of more recent physical, metaphysical, and epistemological
interpretations of quantum information and entanglement (e.g., Myrvold and Christian [43]; Bokulich and Jaeger [44]).

This proposal for an effective non-signalling theorem raises additional questions, and new possibilities for further research, of which can be mentioned only a few here. This discussion will be limited to the following topics: (1) What is the physical meaning of the term “hidden” in the context of ‘hidden-variables theories’? (6.1); (2) What justifies the use of terms like ‘transmission’ and ‘transfer’ in relation to the concept of nonlocal quantum information? (6.2); and (3) Does nonlocal information transfer nevertheless violate relativity theory despite the fact that the information is non-signalling in the communication-theoretic sense? (6.3) Again, the answers here considered are not meant to be final in any way but to elicit awareness of key research challenges concerning an effective non-signalling theorem.

6.1 What is the physical meaning of the term “hidden” in the context of hidden-variables theories?

What is “hidden” about ‘hidden variables’? “The usual nomenclature, hidden variables, is most unfortunate”, Bell [45] complained in a note, and he proposed that “Perhaps uncontrollable variable would have been better, for these variables, by hypothesis, for the time being, cannot be manipulated at will by us.” If one accepts Bell’s proposal that hidden variables represent uncontrollable variables, i.e., variables that “cannot be manipulated at will by us”, then much of the mystery is lifted surrounding the term “hidden”: the variables are called hidden because what characterizes them is that they are statistically-unpredictable, and uncontrollable for any pragmatic purposes, such as for sender-receiver communication. However, the fact of the unpredictability of events, how could it be reconciled with the idea that hidden variables, at least in Bohm’s theory, are deterministic? Does not determinism imply statistical predictability and control? The false habit of identifying determinism with predictability stems from an idealized view of physical systems as fundamentally linearly-behaving systems: If one knows the initial state of a system then one can predict its evolution towards the final state. However, if there is involved only the weakest element of nonlinearity in a deterministic system, such as is the case with sensitive dependencies on initial state conditions, then final state prediction may quickly become impossible (e.g., see deterministic chaos in emergent dynamics). For explanation, even the best nonlinear control and prediction techniques, whether applied to emergent, self-organizing states in physical, chemical, or living systems, are successful only in very limited nonlinear regimes due to the prohibitive complexity of multi-factorial, randomizing interactions (e.g., Walleczek [46]). “Consider the extreme case of a ‘random’ generator”, Bell [27] explained, “… which is in fact perfectly deterministic in nature – and, for simplicity, perfectly isolated. In such a device the complete final state perfectly determines the complete initial state – nothing is forgotten. And yet for many purposes, such a device is precisely a ‘forgetting machine’. A particular output is the result of combining so many factors, of such a lengthy and complicated dynamical chain, that it is quite extraordinarily sensitive to minute variations of any one of many initial conditions. It is the familiar paradox of classical statistical mechanics that such exquisite sensitivity to initial conditions is practically equivalent to complete forgetfulness of them.”

In de Broglie–Bohm theory, just as Bell [27] had suggested with his thought experiment of the “forgetting machine”, the operational impossibility to predict and control individual
quantum correlations derives from the impossibility to know – with arbitrary precision – initial state configurations of the nonlinearily-behaving, yet fully deterministic, (sub)quantum system (e.g., Bohm and Hiley [33]; Holland [34]). In fact, from the vantage point of an effective non-signalling constraint, ‘hidden variables’ represent – by definition – ‘non-signalling variables’ in the here introduced communication-theoretic sense of (Shannon) non-signalling (see Fig. 4). Pursuing, for the moment, Bell’s suggestion to identify the term ‘hidden’ with the notion of ‘uncontrollability’, the concept, for example, of “hidden superluminal influences” (e.g., Scarani et al. [14]) transforms into a concept of “uncontrollable superluminal influences”, a far less obscure concept now. In relation to the notion of nonlocal, i.e., superluminal, influences is it acceptable, however, to invoke terms such as ‘transfer’ and ‘transmission’?

6.2 What justifies the use of terms like ‘transmission’ and ‘transfer’ in relation to the concept of nonlocal quantum information?

In what sense is it acceptable to speak of ‘transfer’ or ‘transmission’ in the context of instantaneous influences? “To speak of instantaneous travel from X to Y is a mixed or incoherent metaphor”, van Fraassen [47] pointed out, “… for the entity in question is implied to be simultaneously at X and at Y – in which case there is no need for travel, for it is at its destination already.” We agree with van Fraassen’s assessment that – in the context of quantum nonlocality – the use of terms such as travel, transfer, propagation, or transmission, is logically incoherent. However, for the present study we have retained the use of such terms for historical reasons, and for reasons of scientific convention. That is, in relation to the notion of ‘instantaneous influences’, instances of “mixed or incoherent metaphor” abound in the literature on quantum foundations such as, to name only two examples, “propagation with infinite velocity” (Bell [8]) or “non-local information transferral” (Pawłowski et al. [48]). We suggest that these authors, of course, were fully aware of the mixed status of such expressions, but no alternative terms were, and still are not presently, available that have accepted use in the scientific community. In an attempt to remedy the conceptual inconsistency, van Fraassen [47] proposed that “… one should say instead that the entity has two (or more) coexisting parts, that it is spatially extended.” That proposition is also in line with an intuitive characterization of Reichenbach’s principle of ‘common cause’ (Reichenbach [49]), which might prove useful as well in accounting for quantum nonlocality, however, there remain important theoretical obstacles there also as was discussed, e.g., by Cavalcanti and Lal [50].

Concerning the proposition by van Fraassen [47], the entity of ‘nonlocal quantum information’ should be described as information which is characterized as existing – at space-like separated locations – in “two … coexisting parts” or as being “spatially extended”. Accordingly, the observation in EPR-type experiments of instant informational correlations then would be described as a function of, for example, the spatial extension of nonlocal information instead of instant transmission. Even so, would the use of a more accurate description of the nonlocal entity alter in any way the findings of the present study? Our conclusions would remain as valid as before, because our argument is indifferent to whether one views nonlocal information as a function of instant transmission, Reichenbach’s ‘common causation’, or van Fraassens’s ‘spatial extension’. Taking up van Fraassen’s proposition again, an effective non-signalling theorem would be violated only under the following condition: not only the entity of ‘information’ would have to be “spatially extended”, or be manifested instantly by “two … coexisting
parts”, but – similarly – the entity of the ‘message’ as well. However, as was illustrated in Figs. 4B and 4C, there exists no such necessary link between the presence of information and the presence of a message or communication. We have summed up this insight in the slogan which states that “while every signal or message represents information, not every information represents a message or signal” (see legend to Fig. 4). In other words, ‘nonlocal information’ does not automatically equal ‘nonlocal communication’. Therefore, the validity of our argument in favour of an effective non-signalling theorem does not depend on choosing one metaphor over another in descriptions of nonlocal information. In either description of the nonlocal entity our argument holds. One can disagree obviously with the logical consistency of terms like transmission or transfer when applied in the context of nonlocality. However, for reasons of historical consistency we have continued to adopt these terms in the provisional sense as discussed here. Future work might construct and use more appropriate terminology, yet the present conclusions do not depend on this.

6.3 Does nonlocal information transfer nevertheless violate relativity theory despite the fact that the information is non-signalling in the communication-theoretic sense?

“Does superluminal information transmission automatically violate relativity theory?” asked Maudlin [32]. An affirmative answer is often tacitly assumed by researchers who are committed to the standard, axiomatic view of the non-signalling condition. However, Maudlin’s extensive analysis reveals a more complex picture. He concludes that – by itself – the possibility of superluminal information transmission “need not give rise to causal paradox, if the information is not available for general use.” Specifically, his analysis finds that the generation of causal paradoxes is the consequence of the possibility of “signal loops and where causal processes cannot be used to send signals paradoxes cannot arise.” By our definition, Maudlin [32] here refers to the formation of Shannon-signal loops, and if their formation is denied, then causal paradoxes cannot occur. An effective non-signalling theorem, as offered in the present study, denies Shannon signalling and therefore negates Shannon-signal loops from being formed.

Finally, nonlocal information transfers that “cannot be used to send signals” need not automatically generate causal paradoxes (Maudlin [32]). The standard assumption might therefore be in need of revision that a quantum theory, or any physical theory in general, must necessarily be unphysical, i.e., in violation of special relativity, because the theory allows for the possibility of nonlocal information transfers. As was alluded to near the end of Sect. 5, critical questions, as part of the theoretical assessment of an ontological quantum theory, e.g., de Broglie–Bohm theory, might be: Does a particular theory predict the (macroscopic) observation in nature of the spontaneous formation of entanglement-induced causal paradoxes, i.e., in the complete absence of epistemic agents, including laboratory devices? More precisely, given the fact – as was emphasized by the present communication-theoretic analysis – that random acts of communication, i.e., random Shannon signal transfers, represent a logically-impossible proposition, how could any actionable information be transmitted instantly – in possible violation of special relativity? As we have proposed in Sect. 5, an effective, operational position on non-signalling might well suffice to counter the often presumed inevitability of paradoxical effects in relation to quantum theories that posit nonlocal information transmission.
6.4 Conclusions

This communication-theoretic study has demonstrated that an effective non-signalling theorem allows for nonlocal quantum information transfer yet – at the same time – effectively denies superluminal signalling and communication. While this study has clarified several key points, the open exploration of important related questions as well as of novel possibilities has just begun. Everyone agrees that no final judgement can yet be delivered concerning the compatibility between, for example, non-relativistic de Broglie–Bohm theory, and the theory of special relativity. Nevertheless, while there may be good reasons for why nonlocal quantum information transfer may represent a physically-unrealistic proposition for quantum mechanics, we have shown that the danger of superluminal signalling and communication is not one of them.

Acknowledgements Work by Jan Walleczek at Phenoscience Laboratories (Berlin) is partially funded by the Fetzer Franklin Fund of the John E. Fetzer Memorial Trust. Work by Gerhard Grössing at the Austrian Institute for Nonlinear Studies (Vienna) is also partially funded by the Fetzer Franklin Fund of the John E. Fetzer Memorial Trust. The authors wish to thank Siegfried Fussy, Johannes Mesa Pascasio, Herbert Schwabl and Nikolaus von Stillfried for their valuable contributions in developing these concepts.

References

1. J. S. Bell, “On the Einstein Podolsky Rosen paradox,” *Physics* 1 (1964) 195–200.
2. J. S. Bell, “Bertlmann’s socks and the nature of reality,” in *Speakable and Unspeakable in Quantum Mechanics*, pp. 139–158. Cambridge University Press, Cambridge, UK, 1987.
3. T. Norsen, “J.S. Bell’s concept of local causality,” *Am. J. Phys.* 79 (2011) 1261, arXiv:0707.0401 [quant-ph].
4. P. H. Eberhard, “Bell’s theorem and the different concepts of locality,” *Nuov. Cim. B* 46 (1978) 392–419.
5. J. S. Bell, “de Broglie–Bohm, delayed-choice, double-slit experiment, and density matrix,” *Int. J. Quantum Chem.* 14 (1980) 155–159.
6. J. S. Bell, “On the impossible pilot wave,” *Found. Phys.* 12 (1982) 989–999.
7. J. S. Bell, “Speakable and unspeakable in quantum mechanics,” in *Speakable and Unspeakable in Quantum Mechanics*, pp. 169–172. Cambridge University Press, Cambridge, UK, 1987.
8. J. S. Bell, “La nouvelle cuisine,” in *Speakable and Unspeakable in Quantum Mechanics*, pp. 232–248. Cambridge University Press, Cambridge, UK, revised ed., 2004.
9. D. Bohm, “A suggested interpretation of the quantum theory in terms of "hidden" variables. I,” *Phys. Rev.* 85 (1952) 166–179.
10. D. Bohm, “A suggested interpretation of the quantum theory in terms of "hidden" variables. II,” *Phys. Rev.* 85 (1952) 180–193.
11. C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal* 27 (1948) 379–423, 623–656.
12. C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. The University of Illinois Press, Urbana, USA, 1949.
13. V. Scarani and N. Gisin, “Superluminal influences, hidden variables, and signaling,” *Phys. Lett. A* 295 (2002) 167–174.
14. V. Scarani, J.-D. Bancal, A. Suarez, and N. Gisin, “Strong constraints on models that explain the violation of Bell inequalities with hidden superluminal influences,” *Found. Phys.* 44 (2014) 523–531.
15. R. Gallego, L. Masanes, G. De La Torre, C. Dhara, L. Aolita, and A. Acín, “Full randomness from arbitrarily deterministic events,” *Nature Commun.* 4 (2013) 2654, arXiv:1210.6514 [quant-ph].
16. N. Harrigan and R. W. Spekkens, “Einstein, incompleteness, and the epistemic view of quantum states,” *Found. Phys.* 40 (2010) 125–157.
17. R. Colbeck and R. Renner, “Is a system’s wave function in one-to-one correspondence with its elements of reality?,” *Phys. Rev. Lett.* 108 (2012) 150402.
18. M. F. Pusey, J. Barrett, and T. Rudolph, “On the reality of the quantum state,” Nature Phys. 8 (2012) 475–478, arXiv:org/abs/1111.3328.
19. G. Ghirardi and R. Romano, “About possible extensions of quantum theory,” Found. Phys. 43 (2013) 881–894.
20. G. Ghirardi and R. Romano, “Ontological models predictively inequivalent to quantum theory,” Phys. Rev. Lett. 110 (2013) 170404.
21. M. S. Leifer, “Is the quantum state real? An extended review of ψ-ontology theorems,” Quanta 3 (2014) 67–155, arXiv:1409.1670.
22. C. J. Wood and R. W. Spekkens, “The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning,” pre-print (2012), arXiv:1208.4119 [quant-ph].
23. J. Walleczek and G. Grössing, “The non-signalling theorem in generalizations of Bell’s theorem,” J. Phys.: Conf. Ser. 504 (2014) 012001, arXiv:1403.3588 [quant-ph].
24. R. Colbeck and R. Renner, “No extension of quantum theory can have improved predictive power,” Nature Commun. 2 (2011) 411–415.
25. R. Colbeck and R. Renner, “Free randomness can be amplified,” Nature Phys. 8 (2012) 450–454, arXiv:1105.3195 [quant-ph].
26. J. S. Bell, “The theory of local beables,” Epistemol. Lett., March (1976) . Repr. in Dialectica 39 (1985) 85-96.
27. J. S. Bell, “Free variables and local causality,” Epistemol. Lett., February (1977) . Repr. in Dialectica 39 (1985) 103-106.
28. H. M. Wiseman, “Bell’s theorem still reverberates,” Nature 510 (2014) 467–469.
29. H. M. Wiseman, “The two Bell’s theorems of John Bell,” J. Phys. A: Math. Theor. 47 (2014) 424001, arXiv:1402.0351 [quant-ph].
30. E. G. Cavalcanti and H. M. Wiseman, “Bell nonlocality, signal locality and unpredictability (or what Bohr could have told Einstein at Solvay had he known about Bell experiments),” Found. Phys. 42 (2012) 1329–1338.
31. T. Norsen, “Local causality and completeness: Bell vs. Jarrett,” Found. Phys. 39 (2009) 273–294.
32. T. Maudlin, Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics. Wiley-Blackwell, West Sussex, UK, 3 ed., 2011.
33. D. Bohm and B. J. Hiley, The Undivided Universe: An Ontological Interpretation of Quantum Theory. Routledge, London, UK, 1993.
34. P. R. Holland, The Quantum Theory of Motion. Cambridge University Press, Cambridge, UK, 1993.
35. A. Valentini, “Signal-locality in hidden-variables theories,” Phys. Lett. A 297 (2002) 273–278, arXiv:quant-ph/0106098.
36. T. L. Short, Peirce’s Theory of Signs. Cambridge University Press, Cambridge, UK, 2007.
37. A. Aspect, P. Grangier, and G. Roger, “Experimental realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A new violation of Bell’s inequalities,” Phys. Rev. Lett. 49 (1982) 91–94.
38. A. Aspect, J. Dalibard, and G. Roger, “Experimental test of Bell’s inequalities using time-varying analyzers,” Phys. Rev. Lett. 49 (1982) 1804–1807.
39. G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, “Violation of Bell’s inequality under strict Einstein locality conditions,” Phys. Rev. Lett. 81 (1998) 5039–5043.
40. W. Tittel, J. Brendel, H. Zbinden, and N. Gisin, “Violation of Bell inequalities by photons more than 10 km apart,” Phys. Rev. Lett. 81 (1998) 3563–3566.
41. R. Ursin, F. Tiefenbacher, T. Schmitt-Manderbach, H. Weier, T. Scheidl, M. Lindenthal, B. Blauensteiner, T. Jennewein, J. Perdigues, P. Trojek, B. Oemar, M. Fuerst, M. Meyenburg, J. Rarity, Z. Sodnik, C. Barbieri, H. Weinfurter, and A. Zeilinger, “Free-space distribution of entanglement and single photons over 144 km,” Nature Phys. 3 (2007) 481–486, arXiv:quant-ph/0607182.
42. M. Giustina, A. Mech, S. Ramelow, B. Wittmann, J. Kofler, J. Beyer, A. Lita, B. Calkins, T. Gerrits, S. W. Nam, R. Ursin, and A. Zeilinger, “Bell violation using entangled photons without the fair-sampling assumption,” Nature 497 (2013) 227–230.
43. W. C. Myrvold and J. Christian, eds., Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle, vol. 73 of The Western Ontario Series in Philosophy of Science. Springer, 2009.
44. A. Bokulich and G. Jaeger, eds., Philosophy of Quantum Information and Entanglement. Cambridge University Press, Cambridge, UK, 2010.
45. J. S. Bell, “Einstein-Podolsky-Rosen experiments,” in Speakable and Unspeakable in Quantum Mechanics, pp. 81–92. Cambridge University Press, Cambridge, UK, 1987.
46. J. Walleczek, ed., *Self-Organized Biological Dynamics and Nonlinear Control*. Cambridge University Press, Cambridge, UK, 2000.

47. B. C. van Fraassen, *Quantum Mechanics; an Empiricist’s View*. Oxford University Press, Oxford, UK, 1991.

48. M. Pawłowski, J. Kofler, T. Paterek, M. Seevinck, and C. Brukner, “Nonlocal setting and outcome information for violation of Bell’s inequality,” *New J. Phys.* 12 (2010) 083051, arXiv:0903.5042v3.

49. H. Reichenbach, *Philosophic Foundations of Quantum Mechanics*. University of California Press, Berkeley and Los Angeles, USA, 1944.

50. E. G. Cavalcanti and R. Lal, “On modifications of Reichenbach’s principle of common cause in light of Bell’s theorem,” *J. Phys. A: Math. Theor.* 47 (2014) 424018, arXiv:1311.6852 [quant-ph].