Noise in the nervous system

A. Aldo Faisal, Luc P. J. Selen and Daniel M. Wolpert

Abstract | Noise — random disturbances of signals — poses a fundamental problem for information processing and affects all aspects of nervous-system function. However, the nature, amount and impact of noise in the nervous system have only recently been addressed in a quantitative manner. Experimental and computational methods have shown that multiple noise sources contribute to cellular and behavioural trial-to-trial variability. We review the sources of noise in the nervous system, from the molecular to the behavioural level, and show how noise contributes to trial-to-trial variability. We highlight how noise affects neuronal networks and the principles the nervous system applies to counter detrimental effects of noise, and briefly discuss noise’s potential benefits.

Variability is a prominent feature of behaviour. Variability in perception and action is observed even when external conditions, such as the sensory input or task goal, are kept as constant as possible. Such variability is also observed at the neuronal level. What are the sources of this variability? Here, a linguistic problem arises, as each field has developed its own interpretation of terms such as variability, fluctuation and noise. In this Review, we use the term variability to refer to changes in some measurable quantity, such as spike timing or movement duration. Importantly, the term variability does not indicate that a particular mechanism has generated the variability, and does not suggest whether the variability is beneficial or detrimental. Trial-to-trial variability can arise from two distinct sources. The first source is from the deterministic properties of the system. For example, the initial state of the neural circuitry will vary at the start of each trial, leading to different neuronal and behavioural responses. The variability in the response will be exacerbated if the system’s dynamics are highly sensitive to the initial conditions. The second source of variability is noise, which is defined in the Oxford English Dictionary as ‘random or irregular fluctuations or disturbances which are not part of a signal [ … ] or which interfere with or obscure a signal or more generally any distortions or additions which interfere with the transfer of information’.

Whereas previous reviews have focused on neuronal variability in general, we focus here on work directly relating to noise. Noise permeates every level of the nervous system, from the perception of sensory signals to the generation of motor responses, and poses a fundamental problem for information processing. In recent years the extent to which noise is present and how noise shapes the structure and function of nervous systems have been studied.

In this Review, we begin by considering the nature, amount and effects of noise in the CNS. As the brain’s purpose is to receive and process information and act in response to that information, we then examine how noise affects motor behaviour, considering the contribution of noise to variability at each level of the behavioural loop. Finally, we discuss the strategies that the nervous system uses to counter, compensate for or account for noise in perception, decision making and motor behaviour. Given the many levels and systems that are spanned, we cannot provide a comprehensive Review, but instead we pick out specific examples that reflect in a more general manner the constraints and limitations that noise sets in the CNS; the benefits of noise are discussed in BOX 1.

Sensory noise

External sensory stimuli are intrinsically noisy because they are either thermodynamic or quantum mechanical in nature. For example, all forms of chemical sensing (including smell and gustation) are affected by thermodynamic noise because molecules arrive at the receptor at random rates owing to diffusion and because receptor proteins are limited in their ability to accurately count the number of signalling molecules. Similarly, vision involves the absorption of photons that arrive at the photoreceptor at a rate governed by a Poisson process. This places a physical limit on contrast sensitivity in vision, which is reduced at low light levels — when fewer photons arrive at the photoreceptor.

At the first stage of perception, energy in the sensory stimulus is converted into a chemical signal (through photon absorption or ligand-binding of odour molecules) or a mechanical signal (such as the movement of hair cells in hearing). The subsequent transduction...
process amplifies the sensory signal and converts it into an electrical one, either directly or indirectly through second-messenger cascades. Any sensory noise that is already present or that is generated during the amplification process (transducer noise) will increase trial-to-trial variability. Therefore, noise levels set perceptual thresholds for later stages of information processing, as signals that are weaker than the noise cannot be distinguished from it after amplification. This is rigorously underpinned by the data-processing inequality theorem, which states that subsequent stages of processing (even if they are noise free) cannot extract more information than is present at earlier stages. Therefore, to reduce noise, organisms often pay a high metabolic and structural price at the first stage of processing (the sensory stage). For example, a fly’s photoreceptors account for 10% of its resting metabolic consumption and its eye’s optics make up over 20% of the flight payload.

Cellular noise
If neurons are driven with identical time-varying stimuli over repeated trials, the timing of the resultant action potentials (APs) varies across the trials. This variability is on the order of milliseconds or lower, but because cortical neurons can detect the coincident arrival of APs on millisecond timescales, this order of timing precision might well be physiologically relevant. Indeed, the precision of single-neuron AP timing on the milli- and sub-millisecond scale has been shown to be behaviourally relevant in perception and movement. To what extent this neuronal variability contributes to meaningful processing (as opposed to being meaningless noise) is the fundamental question of neural coding. Neuronal variability is not equal in all neurons. The Fano factor is a simple measure of
Box 1 | Benefits of noise

Noise is not only a problem for neurons: it can also be a solution in information-processing. Several strategies have been adopted to use noise in this fashion. For example, stochastic resonance is a process by which the ability of threshold-like systems to detect and transmit weak (periodic) signals can be enhanced by the presence of a certain level of noise. At low noise levels, the sensory signal does not cause the system to cross the threshold and few signals are detected. For large noise levels, the response is dominated by the noise. For intermediate noise intensities, however, the noise allows the signal to reach the threshold but does not swamp it. For stochastic resonance to be useful, positive detection of a sub-threshold input must be more desirable than a failure to detect a supra-threshold input. Since its first discovery in cat visual neurons, stochastic-resonance-type effects have been demonstrated in a range of sensory systems. These include crayfish mechanoreceptors, shark multimodal sensory cells, cricket cercal sensory neurons, and human muscle spindles. The behavioural impact of stochastic resonance has been directly demonstrated and manipulated in passive electrosoning paddlefish and in human balance control.

In addition, in spike-generating neurons, sub-threshold signals have no effect on the output of the system. Noise can transform such threshold nonlinearities by making sub-threshold inputs more likely to cross the threshold, and this becomes more likely the closer the inputs are to the threshold. Thus, when outputs are averaged over time, this noise produces an effectively smoothed nonlinearity. This facilitates spike initiation and can improve neural-network behaviour, as was shown in studies of contrast invariance of orientation tuning in the primary visual cortex. Moreover, neuronal networks that have formed in the presence of noise will be more robust and explore more states, which will facilitate learning and adaptation to the changing demands of a dynamic environment.

Fano factor
The ratio of the variance of a stochastic variable to its mean.

Stochastic process (random process)
A process that generates a series of random events.

Positive feedback
Feedback that responds to a perturbation in the same direction as the perturbation, thereby amplifying its effect.

Nodes of Ranvier
Regularly spaced gaps in the myelin sheath that surrounds a myelinated axon. They expose the axonal membrane to the extracellular fluid and contain large numbers of voltage-gated ion channels and thus enable conduction of the action potential.

Patch-clamp technique
An electrophysiological method that allows the study of the flow of current through a very small patch of cell membrane, which can contain just a single ion channel.

variability that ignores temporal structure and higher-order statistics. Neural responses without variability have Fano factors of zero, whereas Poisson processes (which are highly variable) have a Fano factor of one. Some cortical neurons are highly variable, with Fano factors of one or greater, whereas others have Fano factors that are closer to zero. SIMILARLY, there is a range of variability in neurons in the mammalian and invertebrate visual pathways. Moreover, high- and low-variability neurons are often observed in the same region, and a single neuron can respond with different amounts of variability depending on the stimulus conditions.

Multiple factors contribute to neuronal trial-to-trial variability. These include changes in the internal states of neurons and networks, and random processes inside neurons and neuronal networks. To what extent each of these factors contributes to the total observed trial-to-trial variability remains unclear, especially as network connectivity and random fluctuations are indeed averaged out. However, many neurons are tiny: cerebellar parallel fibres have an average diameter of 0.2 μm; C-fibres, which are involved in sensory and pain transmission, range between 0.1 and 0.2 μm in diameter; and the unmyelinated pyramidal-cell axon collaterals, which form the vast majority of local cortico–cortical connections, have an average diameter of 0.3 μm. Similarly, most (spiny- or bouton-type) CNS synapses have submicrometer dimensions. At these small length scales the numbers of molecules involved are small and the influence of noise is dramatically increased. Here we review the main sources of noise in the nervous system at the cellular level and the consequences for neuronal function.

Electrical noise and action potentials.

The membrane potential is used both for local computation and to carry APs. Although variability in resting membrane potential (membrane-potential fluctuations) and AP threshold have been studied for a long time, the mechanisms that underlie these fluctuations have only recently gained attention. Electrical noise in neurons causes membrane-potential fluctuations even in the absence of synaptic inputs. The most dominant source of such electrical noise is channel noise (FIG. 1b) — electrical currents produced by the random opening and closing of voltage- or ligand-gated ion channels. Stochastic models have shown that channel noise can account for variability in the AP threshold at nodes of Ranvier and the reliability of AP initiation in membrane patches. Furthermore, patch-clamp experiments in vitro show that channel noise in the dendrites and in the soma produces membrane-potential fluctuations that are large enough to affect AP timing. Both the initiation and the propagation of APs can be affected by channel noise.

At the site of AP initiation — the soma or the axon hillock — channel noise can affect the timing of APs (despite the comparatively large number of ion channels).
Box 2 | Noise build-up in networks

How can neural networks maintain stable activity in the presence of noise? There are several ways in which networks can affect overall noise levels. The figure illustrates this with three simple examples in which graded-potential neurons linearly sum inputs. Part a shows convergence of signals onto a single neuron. If the incoming signals have independent noise, then noise levels in the postsynaptic neuron will scale in proportion to the square root of the number of signals (N), whereas the signal scales in proportion to N. If the noise in the signals is perfectly correlated, then the noise in the neuron will also scale in proportion to N. Part b shows the passage of signals through a series of neurons. In this case, noise levels increase in proportion to the square root of the number of successive neurons. By contrast, parallel connections (not shown) do not augment noise through network interactions. Part c shows that recurrence in networks results in the build-up of correlated noise.

Other computational operations in each neuron can change the build-up of network noise. The linear operation of amplification leaves the signal-to-noise ratio unchanged. Nonlinear operations, such as multiplication and thresholding, affect noise build-up differently. In general, multiplication operations increase the coefficient of variation (CV) of the output, whereas thresholding decreases the CV. Several studies have examined how noise acts in neuron-like nonlinear systems\(^{184,185}\). The highly parallel and distributed yet compact structure of the CNS might help to limit the amount of noise that builds up.

Experimental evidence suggests that average neuronal activity levels are maintained by homeostatic plasticity mechanisms that dynamically set synaptic strengths\(^{196}\). Ion-channel expression\(^{195}\) or the release of neuromodulators\(^{20}\). This in turn suggests that networks of neurons can dynamically adjust to attenuate noise effects. Moreover, these networks might be wired so that large variations in the response properties of individual neurons have little effect on network behaviour\(^{107}\).

Furthermore, in many spiking neurons\(^{188,199}\), doubling the input results in less than twice the output. This suggests that presynaptic noise and intracellular noise are attenuated as the signal passes through the neuron. The fact that the noise remains so small suggests that neuronal networks can be organized in a way that prevents local noise accumulating as neural signals propagate through them\(^{106}\). Thus, the analogue (membrane-potential based) nature of local neuronal computation (computation within neurons) and the more digital (action-potential based) nature of global information transmission\(^{196,199}\) might be essential ingredients in building noise-robust computational circuits\(^{192}\). Figure modified, with permission, from REF. 193 © (2001) Wiley.

Channels noise also affects AP propagation in axons, producing trial-to-trial variability in AP timing. This variability occurs whenever the input resistance of an axon is large enough that small numbers of ion channels can support AP conduction\(^{41}\). Using biophysical theory and stochastic simulations, it was shown that in CNS axons of 0.1–0.5 \(\mu m\) diameter, channel noise introduces significant jitter in AP propagation\(^{41} (FIG. 2a)\). Thus, the variability in postsynaptic responses that results from axonal channel noise will increase the longer and thinner the presynaptic axon. Moreover, populations of ion channels can retain a memory of axonal activity for several hundred milliseconds, owing to a complex interaction between the internal states of ion-channel populations and the membrane potential. This history dependency results in some patterns of spikes (such as bursts) being less affected by noise than others\(^{46}\). Such ‘message-dependent’ noise has been observed in mammalian neurons\(^{46,165}\); however, this effect is missed when models use stochastic approximations (for example, the Langlevin or Fokker–Planck approximations) or ignore spatial interactions. Despite the evidence for the importance of variability in AP propagation for trial-to-trial variability, this has tended to be overlooked by most experimental studies (except those of AP conduction failure\(^{46}\)), with postsynaptic variability being mainly attributed to synapses.
Why is AP propagation so sensitive to noise, contrary to previous claims\textsuperscript{41,43,66,67}? Detailed stochastic modeling has shown\textsuperscript{41} that the leading edge of AP propagation is driven by a relatively small — and thus noisy — ionic current flowing inside the axon. This causes jitter in the speed of the propagation of the AP and thus results in variability in AP timing. By contrast, the current following the leading edge is large and therefore conduction failures owing to channel noise are unlikely, even in very thin axons (where <3\% of all APs fail). Thus, axonal channel noise cannot account for the failure rates that have been reported in much larger CNS axons (where 5–80\% of APs fail\textsuperscript{46,67}), and conduction failures that have been observed in the nervous system are more likely to be due to computational mechanisms that allow ‘editing’ of spike trains\textsuperscript{49} than to noise\textsuperscript{41}.

Other electrical-noise sources include Johnson noise and shot noise owing to membrane resistance, which are three orders of magnitude smaller than channel noise in CNS neurons\textsuperscript{26,27,70,71}. Moreover, variations in the activity of nearby neurons could produce ‘cross-talk noise’ in the confined spaces of the CNS. Such cross-talk can arise through ephaptic coupling\textsuperscript{48}, large changes of extracellular ion concentration after electrical signalling\textsuperscript{72}, and spillover of neurotransmitters\textsuperscript{71} between unrelated synapses.

Synaptic noise. If a presynaptic cell is driven repeatedly with identical stimuli, there is trial-to-trial variability in the postsynaptic response\textsuperscript{44,73} (FIG. 2b). This variability could arise from noise\textsuperscript{41,46} or from a deterministic process that is too complex to grasp and thus appears random\textsuperscript{55,74}. Here we discuss evidence for the considerable contribution of noise to synaptic variability.

Many neocortical cells receive an intense synaptic bombardment from thousands of synapses\textsuperscript{75–77}, which is often referred to as ‘synaptic background noise’ (REFS 80,81). However, the rich set of dendritic mechanisms that allow individual synapses to interact suggests that this ‘background’ activity is unlikely to be composed only of noise\textsuperscript{26,27,82,83}. Indeed, experimental evidence and computational arguments suggest that the synaptic background activity contains meaningful structure\textsuperscript{16,81–85}.
Nevertheless, there are microscopic sources of true noise present at each synapse that are also likely to contribute to this synapic background variability and influence neuronal firing\textsuperscript{65,66}.

The classic manifestation of synaptic noise is the spontaneous miniature postsynaptic current (mPSC) that can be recorded in the absence of presynaptic input. Katz and collaborators interpreted mPSCs as being the result of spontaneously released neurotransmitter vesicles, thus establishing the quantal nature of synaptic transmission\textsuperscript{81}. This work remains an exquisite example of how taking noise into account informs our understanding of neural mechanisms.

Several sources of noise at synapses can influence information transmission and induce variability (FIG 1b). mPSCs are caused by random events in the synaptic-transmission machinery, such as the spontaneous opening of intracellular Ca\textsuperscript{2+} stores\textsuperscript{86,87}, synaptic Ca\textsuperscript{2+}-channel noise, spontaneous triggering of the vesicle-release pathway\textsuperscript{87} or spontaneous fusion of a vesicle with the membrane. Once vesicles are released they induce a postsynaptic current, the amplitude of which shows considerable trial-to-trial variability (the coefficient of variation (CV) being typically >0.2 (REFs 88,89)). To what extent can this variability in the postsynaptic response be attributed to noise? First, the same stochastic processes that produce spontaneous mPSCs are also present during normal synaptic transmission, and will alter the amplitude of the postsynaptic current. Second, the width of the presynaptic AP determines the size of the Ca\textsuperscript{2+} signal (the duration of channel opening) that drives vesicle release and governs the number of vesicles that are released as well as the probability of release. AP width variability can result from axonal channel noise, which becomes significant for CNS synapses that are innervated by neurons with thin axons\textsuperscript{81}.

Several additional factors have been shown to affect postsynaptic-response amplitude, each of which relies on noisy biochemical mechanisms and involves small numbers of molecules and is therefore subject to considerable thermodynamic noise. First, variability in the number of neurotransmitter molecules released per vesicle (~2000) arises owing to variations in vesicle size\textsuperscript{88} and vesicular neurotransmitter concentration\textsuperscript{89}. Second, there is variability owing to the randomness of the diffusion of a relatively small number of molecules (CV = 0.16 (REF 92)). Third, the location of vesicle release in the synaptic cleft has an impact on the postsynaptic response (CV = 0.37 (REF 92)). Vesicles are distributed over the synaptic active zone and, as each AP will trigger the release of only some of them, the location varies from event to event. Fourth, synaptic-receptor channel noise increases the variability, especially if only a small number of receptors are involved\textsuperscript{90}. Fifth, the number\textsuperscript{91} and density\textsuperscript{92} of receptor proteins at any synapse might stochastically vary over time, as the expression and degradation of proteins is limited by thermodynamic noise\textsuperscript{93}.

In addition to variability in response amplitude, some CNS synapses release either one or no vesicles in response to an AP. The vesicle-release probability at small and bouton-type central synapses is typically low and is controlled by plasticity and adaptation mechanisms\textsuperscript{94}. The probability of release itself could constitute a signal for information processing\textsuperscript{95}. Therefore, the accuracy with which vesicle-release probability can be controlled might be computationally important; however, this accuracy has not been adequately quantified.

Summing up the effect on postsynaptic variability from the above synaptic noise sources, we note that the total observed synaptic trial-to-trial variability in many synapses (CV > 0.2) can be fully accounted for by noise. However, there might also be biochemical mechanisms that reduce noise\textsuperscript{96}.

**Motor noise**

We interact with the environment through movements, which are inherently variable from trial-to-trial. To generate movement the signals from the CNS are relayed by motor neurons and converted into mechanical forces in their muscle fibres (FIG 1c). The force that a single motor neuron can command is directly proportional to the number of muscle fibres that it innervates. When small forces are generated, motor neurons that innervate a small number of muscle fibres are active. When larger forces are generated, additional motor neurons that innervate a larger number of muscle fibres are also active. This is known as Henneman’s size principle\textsuperscript{97}. Moreover, as whole-muscle force increases, the firing rates of the active motor neurons increase, such that those that innervate a small number of muscle fibres have the highest firing rate.

The variability in the force that is produced by a whole human skeletal muscle is proportional to the average force that is produced by that muscle\textsuperscript{100,101}. This has been attributed to the physiological organization of the pool of motor neurons and their muscle fibres\textsuperscript{101–104}; each AP arriving at the muscle fibre induces a ‘twitch’. At low firing rates these twitches are separated in time, but as firing rates increase the twitches fuse into one smooth contraction. Whole-muscle force is determined by the number of active motor neurons and the firing rates of these neurons. The motor neuron that innervates the most fibres will have the lowest firing rate and will therefore induce unfused twitches in the muscle fibres that it innervates. Thus, any variability in the force that is generated by the muscle fibres that are innervated by this motor neuron will contribute most to whole-muscle force variability.

Three mechanisms contribute to the variability in the force that is generated by muscle fibres. First, even if a motor neuron fires perfectly periodically, there will be ‘ripples’ in the force that is generated by its muscle fibres, owing to unfused twitches. This effect is further enhanced by the synchronization of motor neurons through common mechanosensory feedback\textsuperscript{102}. Second, motor neurons are subject to the same sources of cellular noise as any other neuron, making noise appreciable in AP timing in myelinated motor axons of 10 µm diameter\textsuperscript{103} and at the neuromuscular junction\textsuperscript{104}. The resulting AP timing variability will reduce the periodicity of the force and thus increase its variability. Both of these factors will contribute to

---

**Coefficient of variation (CV)** The ratio of the standard deviation of a variable quantity to its mean.

**Release probability** The probability of a vesicle being released during a synaptic-transmission event.
overall muscle-force variability. Third, each twitch triggered by a single AP might also show trial-to-trial variability in its amplitude and duration, owing to noise in the biochemical cascade that generates the twitch force. However, to our knowledge this has not been quantified. In addition, as in thin axons, Ca$$^{2+}$$-channel noise in muscle fibres or stochastic processes in energy release and transport could also produce random twitches. Furthermore, noise might result from unrelated electrical cross-talk between motor neurons or muscle fibres, which could recruit other muscle fibres by ephaptic coupling.

Our present knowledge of force variability is based on isometric contractions (in which muscle length does not change), and it is unclear how this translates to variability during movement. The effect of single motor-neuron spikes on muscle movement has been measured only in invertebrate systems, in which it was shown that variability in spike timing (on the order of milliseconds) and in the number of spikes ($$\pm 1$$) produces variability in muscle length of up to 10%. These invertebrate muscles are comparable in scale to the human laryngeal muscles that control speech production, which have to operate with millisecond precision. However, little is known about the characteristics, activation and reliability of such muscles.

Human motor behaviour — from eye movements to hand trajectories — can be explained by optimal control models that generate movements in a way that minimizes the impact of motor noise. It remains unclear how much of the observed trial-to-trial movement variability is due to motor-neuron and muscle noise and how much is due to other sources of variability in the (spinal) motor command.

**Principles of how the CNS manages noise**
In general, noise cannot be removed from a signal once it has been added. Furthermore, it is important to note that in some cases it is not always desirable to remove noise, as noise can have beneficial consequences for information processing. However, there are several principles that can be used to minimize the negative consequences of noise. We now review two key principles — averaging and prior knowledge — that the CNS applies at multiple levels.

The principle of averaging can be applied whenever redundant information is present across the sensory inputs to the CNS or is generated by the CNS. Averaging can counter noise if several units (such as receptor molecules, neurons or muscles) carry the same signal and each unit is affected by independent sources of noise. Averaging is seen at the very first stage of sensory processing. For example, the stereocilia of auditory hair cells capture sound vibrations and open mechanically gated ion channels. These stereocilia are mechanically coupled and so they move together, averaging random fluctuations in the movement of individual stereocilia. Similarly, visual inputs are typically averaged over photoreceptors with adjacent or overlapping receptive fields. Moreover, hair cells and photoreceptors are graded-potential neurons that do not use APs but instead communicate their varying membrane potential through graded synapses. This makes noise removal through averaging a straightforward operation for their postsynaptic membranes.

Counterintuitively, divergence (one neuron synapsing onto many) can also support averaging. When signals are sent over long distances through noisy axons, rather than using a single axon it can be beneficial to send the same signal redundantly over multiple axons and then combine these signals at the destination. Crucially, for such a mechanism to reduce noise the initial divergence of one signal into many must be highly reliable. Such divergence is seen in auditory inner hair cells, which provide a divergent input to 10–30 ganglion cells through a specialized ‘ribbon synapse’.

Prior knowledge can also be used to counter noise. If the structure of the signal and/or noise is known it can be used to distinguish the signal from the noise. This principle is especially helpful in dealing with sensory signals that, in the natural world, are highly structured and redundant. By using prior knowledge about the expected structure, sensory processing can compensate for noise. This is manifest in the notion that a neuron’s receptive field tells us what message the neuron is conveying. Signal-detection theory shows that the optimal signal detector, subject to additive noise, is obtained by matching all parameters of the detector to those of the signal to be detected: in neuroscience this is termed the matched-filter principle. Thus, the structures of receptive fields embody prior knowledge about the expected inputs and thereby allow neurons to attenuate the impact of noise.

Simple averaging works best when each signal source is corrupted by a similar amount of noise. Therefore, principles of averaging and prior knowledge are often combined in the nervous system when the sources are affected by different amounts of noise. Prior knowledge about the amount of noise for a given source allows for weighted averaging. In general, the less noisy (more reliable) inputs should contribute more to the averaging process than more noisy (less reliable) inputs. This has been demonstrated in several behavioural studies in which subjects were required to integrate inputs from different pairs of sensory sources. The studies showed that the weight given to each source was proportional to its reliability (the inverse of the variance of the source), demonstrating that the nervous system has prior knowledge about the variability of its senses. Moreover, the observed behavioural variability in estimation tasks involving multiple sensory sources could be predicted from the behavioural variability to the individual sources if the source variability was mathematically treated as independent noise. This strongly suggests that most of the behavioural variability in such sensory tasks arises from noise rather than from deterministic sources of variability. The motor system also applies a weighted averaging mechanism of this type to reduce the consequences of noise. For example, when redundant muscles can rotate a joint, the muscles are co-activated in a way that minimizes total movement variability.

**Redundancy**
The presence of superfluous or duplicate information in a message.

---

**Box 1**

Counter-intuitively, divergence (one neuron synapsing onto many) can also support averaging. When signals are sent over long distances through noisy axons, rather than using a single axon it can be beneficial to send the same signal redundantly over multiple axons and then combine these signals at the destination. Crucially, for such a mechanism to reduce noise the initial divergence of one signal into many must be highly reliable. Such divergence is seen in auditory inner hair cells, which provide a divergent input to 10–30 ganglion cells through a specialized ‘ribbon synapse’ (REF. 124).

Prior knowledge can also be used to counter noise. If the structure of the signal and/or noise is known it can be used to distinguish the signal from the noise. This principle is especially helpful in dealing with sensory signals that, in the natural world, are highly structured and redundant. By using prior knowledge about the expected structure, sensory processing can compensate for noise. This is manifest in the notion that a neuron’s receptive field tells us what message the neuron is conveying. Signal-detection theory shows that the optimal signal detector, subject to additive noise, is obtained by matching all parameters of the detector to those of the signal to be detected: in neuroscience this is termed the matched-filter principle. Thus, the structures of receptive fields embody prior knowledge about the expected inputs and thereby allow neurons to attenuate the impact of noise.

Simple averaging works best when each signal source is corrupted by a similar amount of noise. Therefore, principles of averaging and prior knowledge are often combined in the nervous system when the sources are affected by different amounts of noise. Prior knowledge about the amount of noise for a given source allows for weighted averaging. In general, the less noisy (more reliable) inputs should contribute more to the averaging process than more noisy (less reliable) inputs. This has been demonstrated in several behavioural studies in which subjects were required to integrate inputs from different pairs of sensory sources. The studies showed that the weight given to each source was proportional to its reliability (the inverse of the variance of the source), demonstrating that the nervous system has prior knowledge about the variability of its senses. Moreover, the observed behavioural variability in estimation tasks involving multiple sensory sources could be predicted from the behavioural variability to the individual sources if the source variability was mathematically treated as independent noise. This strongly suggests that most of the behavioural variability in such sensory tasks arises from noise rather than from deterministic sources of variability. The motor system also applies a weighted averaging mechanism of this type to reduce the consequences of noise. For example, when redundant muscles can rotate a joint, the muscles are co-activated in a way that minimizes total movement variability.
Averaging is used in many neural systems in which information is encoded as patterns of activity across a population of neurons that all subserve a similar function (for example, see refs 140,141); these are termed neural population codes. A distributed representation of information of this type is more robust to the effects of noise. Many sensory systems form a spatially-ordered population — that is, a map — in which neighbouring neurons encode stimuli that share closely related features. Such spatially ordered populations support two basic goals of neural computation: first, a transformation between different maps (such as the direction of sounds into neck rotation) and, second, the combination of information from multiple sources (such as visual- and auditory-cue combination)142. The information capacity of a population of neurons is greatest when the noise sources across the population are not correlated. Noise correlations, which are often observed in populations of higher-order neurons, limit information capacity143,144 and have led to the development of population-coding strategies that account for the effects of correlations145.

The principles of averaging and prior knowledge can be placed into a larger mathematical framework of optimal statistical estimation and decision theory, known as Bayesian inference146. Bayesian inference assigns probabilities to propositions about the world (beliefs). These beliefs are calculated by combining prior knowledge (for example, that an animal is a predator) and noisy observations (for example, the heading of an animal) to infer the probability of propositions (for example, animal attacks). Psychophysical experiments have confirmed that humans use these Bayesian inferences to allow them to cope with noise (and, more generally, with uncertainty) in both perception and action147,148. However, the neural mechanisms that are involved in Bayesian computations are unknown. One idea is that neurons encode probabilities or beliefs about the state of the world149,150 and this concept has been incorporated into Bayesian models of neuronal population codes142,151,152.

The above discussion has focused on the processing of information arriving simultaneously from multiple neurons or sensory modalities. However, information is often acquired over time, and in this case temporal averaging to be used to remove noise. For example, in signal-transduction systems, biochemical reaction time constants could be set to make the duration of the reactions longer than the duration of the noise events — this would average-out random fluctuations153. Averaging over time can take place at the cellular level because of the temporal-integration properties of the membrane. These properties can be tuned by an appropriate choice of neuronal geometry114 and ion channels154 so that the characteristic bandwidth of the noise is strongly attenuated whereas the signal is not. Electrophysiological studies in the monkey have shown that behaviourally relevant signals are averaged not only across neuronal populations but also over time in the formation of a behavioural decision155.

In behaviour, temporal averaging is important when we need to estimate the current state or configuration of our limbs. Both the motor commands acting on our body and the sensory feedback containing information about the configuration of our body are noisy. Knowing the motor command allows us to predict the expected body configuration using an internal forward (predictive) model. However, such a prediction would deviate over time if sensory feedback were not available. The Kalman filter156 is an algorithm that combines noisy sensory feedback and the prediction from forward models to estimate the current configuration of our body over time. Kalman filtering in the CNS was demonstrated in behavioural studies of hand position157 and posture159.

In many cases the CNS has to choose a strategy by which it will achieve a goal through interaction with the environment. For example, in reaching the motor system has to specify a sequence of muscle activations to achieve a goal. However, there are many possible strategies to achieve a goal, and each might have a different associated cost (error, energy or time). Finding efficient strategies involves optimizing a cost function. For example, it has been proposed that we choose to move in a way that reduces the detrimental consequences of noise111. Stochastic optimal-control theory158 has emerged as a framework by which to study sensorimotor control. This theory makes several predictions that have been experimentally verified. For example, rather than specify a desired hand trajectory and use feedback to keep you on that trajectory, this theory proposes that optimal feedback control on task-relevant parameters is used: by allowing variation in parameters that do not affect the task, the system can behave in a more optimal manner. Stochastic optimal-feedback control is a beautiful example of how the principles of prior knowledge and averaging are put to use in motor behaviour. This framework has been able to explain quantitative data from human and primate movements159-161. However, the neuronal substrate and mechanisms of such optimal controllers remain unknown.

Humans also use strategies that appear to increase noise. Confronted with higher movement-accuracy constraints (for example, when asked to rapidly point at small targets), people co-contract their muscles164, which increases joint stiffness165,166. However, greater activation of the muscles results in higher neuromuscular noise levels and is expected to produce larger movement variability. The reason that this is not the case lies in the dynamic properties of the muscles. In fact, movement variability decreases overall because the positive stabilizing effect of enhanced stiffness exceeds the negative effects of the increased force variability of the individual muscles163,165-167. Thus, human sensorimotor control takes account of noise to increase behavioural precision.

**Conclusion**

Noise has recently emerged as a key component of a wide range of biological systems — from gene expression79 to heart function81. In neuroscience, we have shown how noise is introduced at all stages of the sensorimotor loop, from the level of a single signalling protein to that of body movement. Noise has direct behavioural consequences, from setting perceptual thresholds to affecting movement precision. Although there has been...
an awareness of sensory noise for over half a century, cellular and motor noise have only recently received significant attention.

The question of the extent to which noise generates variability in the CNS is likely to require both experimental studies and stochastic modelling (in which each source of variability can be controlled for). We are beginning to develop a bottom-up understanding of how noise that is present at the molecular level (channel noise in membranes and biochemical noise at synapses) affects information processing at macroscopic scales (whole neurons, neuronal networks and behaviour). At all of these levels a key advance has been the use of stochastic models that can explain the experimentally observed variability and enable mechanisms to be characterized in a more detailed and often simpler manner than deterministic models (for example, see REFS 45, 61, 103, 117, 148). It has often been convenient to approximate noise by some additive random processes, such as Poisson or Gaussian processes, in which higher-order statistics beyond the mean (and the variance), as well as temporal structure, are ignored. This simple approach often forms the best assumption when data is lacking, and it simplifies the mathematical manipulation. However, it can result in noise levels being underestimated by several orders of magnitude in many small structures of the CNS.14.

Owing to the discrete nature of molecular noise and the nonlineairities that are present, noise can have a complex temporal structure, such as abrupt and large changes in noise level15 and spatial interactions can produce unexpected effects16,32. Advances in stochastic (Monte Carlo) simulations have made it possible to investigate in silico the nature and effects of noise in well known but previously determinedistically described mechanisms (for example, see REFS 61, 65, 92, 117).

The amount of noise that can be tolerated for a task depends on the required internal (such as long-term stability of memories) and behavioural (such as movement accuracy) performance. Noise levels set both hard limits on the CNS, such as the degree of miniaturization of the brain’s circuits33 and soft constraints, such as the metabolic cost or the amount of time that is required to complete a task. For example, APs are noisy but also metabolically costly (mean neuronal firing rates in the cortex appear to be limited by energy supply140). Therefore, although neuronal communication becomes more reliable by using more APs, it also becomes more expensive. This trade-off has been observed in mammalian visual systems169,170. Another trade-off involves noise and time. For example, in pointing tasks, movement speed and pointing accuracy are inversely related (Fitt’s law171), as faster movements require greater muscle forces, which are more noisy172. Therefore, noise is an integral part of the tradeoff between CNS resources (mass, size, time delays, et cetera) and performance which might ultimately determine evolutionary fitness. Noise is an inescapable consequence of brains operating with molecular components at the nanometer scale, sensors that are sensitive to individual quanta and complex networks of noisy neurons that generate behaviour. The presence of noise in nervous systems has profound implications for their computational power172. Yet, despite significant noise levels our brain appears to function reliably, presumably because it has evolved under the constraints that are imposed by noise. Therefore, to understand the nervous system we have to distinguish variability from noise by accounting for its sources and appreciate the way in which it influences the brain’s structure and function.

8. Bialek, W. & Setayeshgar, S. Physical limits to biochemical signalling. Proc. Natl Acad. Sci. USA 102, 10040–10045 (2005). This paper uses a first-principle approach to show that there are noise limits to biochemical signalling. These noise limits are of importance in olfaction, intracellular and synaptic signalling.

7940–7949 (2003).}


This paper provided the first evidence of a strategy for the optimal integration of multi-dimensional cues. When integrating proprioception and visual information in the plane (two-dimensional cues), subjects use knowledge of direction-dependent variability to generate an optimal estimate.


141. Hooper, S. L., Guschlbauer, C., von Uckermann, G. & Liu, Y. C. Drift-diffusion estimates of sensory cues. This was one of the first studies to demonstrate how subjects increase joint stiffness when confronted with higher accuracy demands in goal-directed movements. The increase in joint stiffness results in less kinematic variability.


143. Hooper, S. L., Guschlbauer, C., von Uckermann, G. & Liu, Y. C. Drift-diffusion estimates of sensory cues. This was one of the first studies to demonstrate how subjects increase joint stiffness when confronted with higher accuracy demands in goal-directed movements. The increase in joint stiffness results in less kinematic variability.


This paper provided the first demonstration of stochastic resonance effects at the behavioural level.


Acknowledgements
We would like to thank H. Barlow, J. Niven and H. Robinson for comments on our manuscript. We acknowledge the financial support of the Wellcome Trust and the European project SENSOPAC (IST-2005-028056).

FURTHER INFORMATION
A. Aldo Faisal’s homepage: www.eng.cam.ac.uk/~aaf23
Luc Selen’s homepage: www.eng.cam.ac.uk/~lpjs2
Wolpert lab homepage: www.wolpertlab.com
Stochastic simulator (synpase): http://www.mcell.cnl.salk.edu/
Stochastic simulator (neuron): http://www.modigliani.co.uk
ALL LINKS ARE ACTIVE IN THE ONLINE PDF

© 2008 Nature Publishing Group