The Poisson shot noise model of visual short-term memory and choice response time: Normalized coding by neural population size

Philip L. Smith [*]

Melbourne School of Psychological Sciences, The University of Melbourne, Australia

HIGHLIGHTS

- Presents a normalized coding model of the statistics of stimulus representations.
- Normalized coding predicts a constant diffusion coefficient decision model.
- Normalized coding predicts the information capacity of visual short-term memory.
- The model reconciles previously incompatible bodies of theory and experimentation.

ARTICLE INFO

ABSTRACT

A normalized coding condition is proposed that provides a theoretical link between the Poisson shot noise model of choice response time and a Poisson neuron model of the information capacity of visual short-term memory (VSTM). In both models, noise in the cognitive representations of stimuli is attributed to Poisson variability in the neural processes that encode them. In VSTM, a Poisson coding model predicts the invariance of $\sum_i (d_i')^2$ across displays of different sizes, as is found experimentally, but it incorrectly predicts that the diffusion coefficient of the approximating diffusion process will decrease with its drift rate. Normalized coding assumes that the squared magnitudes of the random perturbations to a stimulus representation are inversely proportional to the number of Poisson neurons that represent it. The normalized coding model correctly predicts both a constant diffusion coefficient and the invariance of $\sum_i (d_i')^2$, as required by the experimental data. Normalized coding reconciles the theoretical and empirical properties of diffusion models of decision-making and the sample-size model of VSTM.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction: the Poisson shot noise model of choice response time

One of the theoretical goals of cognitive neuroscience is to develop mathematical models that link behavioral and neural levels of description. In the area of speeded decision-making, progress towards this goal has received impetus from the discovery that models of choice response time (RT) developed in mathematical psychology to account for the speed and accuracy of behavioral decisions can also characterize the time course of the underlying neural processes. Smith and Ratcliff (2004) reviewed recent developments in this area. The discovery of this relationship invites the question of how the evidence accumulation processes hypothesized to underlie decision-making in behavioral models

can be realized computationally in the neural ensemble. The purpose of this article is to propose a solution to one of the theoretical problems that arise in attempting to link behavioral and neural models. This is the problem of how the variance of encoded stimulus representations changes with the number of neurons recruited to represent a stimulus and how these changes are expressed in models of choice RT. Pragmatically, the problem is how to ensure that the properties of a neural model and a behavioral model can be made consistent with each other.

The most influential models of choice RT are the sequential-sampling models (Laming, 1968; Link, 1992; Luce, 1986; Townsend & Ashby, 1983; Vickers, 1979). These models assume that decisions are made by accumulating successive samples of noisy evidence until a criterion quantity of evidence needed for a response is obtained. The best of these models successfully account for distributions of RT for correct responses and errors and the associated choice probabilities in a variety of cognitive tasks (Ratcliff & McKoon, 2008; Ratcliff & Smith, 2004). Recently, a number of authors have reported what appears to be the neural

* Correspondence to: Melbourne School of Psychological Sciences, The University of Melbourne, VIC 3010, Australia.

E-mail address: philipls@unimelb.edu.au.

signature of the evidence accumulation process in brain structures that are active in monkeys performing saccade-to-target decision tasks (Hanes & Schall, 1996; Ratcliff, Cherian, & Segraves, 2003; Roitman & Shadlen, 2002). Structures in the oculomotor control system – such as the frontal eye fields, lateral interparietal area, and the superior colliculus – which are involved in the preparation and control of saccadic eye movements, show increasing, response-locked patterns of neural firing during the pre-response interval. These patterns of firing are consistent with the interpretation that they either implement an evidence accumulation process or read out the state of an accumulation process located elsewhere in the brain.

Consistent with this interpretation, several authors have shown that sequential-sampling decision models can be used to characterize the empirically-observed relationship between neural firing rates and the speed and accuracy of decision-making. Ratcliff and colleagues showed that the parameters of sequential-sampling models estimated from RT distributions and choice probabilities of monkeys performing saccade-to-target decision tasks predict the time course and relative firing rates in build-up cells in the superior colliculus (Ratcliff et al., 2003, 2011; Ratcliff, Hasegawa, Hasegawa, Smith, & Segraves, 2007). Shadlen and colleagues showed a similar relationship for cells in lateral interparietal area in a global dot motion task (Ditterich, 2006; Gold & Shadlen, 2001). Schall and colleagues showed that the firing rates of visual neurons in frontal eye fields can predict saccadic decision times when the visual spike trains are used as inputs to a sequential-sampling decision process (Purcell et al., 2010; Purcell, Schall, Logan, & Palmeri, 2012). These demonstrations invite the development of models in which the properties of the evidence accumulation process are derived from the point-process statistics of the underlying neural firing rates.

To this end, Smith (2010) and Smith and McKenzie (2011) proposed a theoretical link between neural firing rates and evidence accumulation via the properties of the *Poisson shot noise process*. The shot noise process, which was originally proposed as a model of electron absorption in vacuum tubes (Campbell, 1909), provides an idealized model of the flux in the postsynaptic potential across cells in a neural population in response to a volley of action potentials that encode stimulus information. The latter are idealized as a (possibly time inhomogeneous) Poisson process. In the shot noise model, discriminative information is carried by the difference between an excitatory and inhibitory shot noise process, which Smith called an *excitatory–inhibitory shot noise pair*.

The shot noise process represents the instantaneous state of the evidence available to the decision process at a given time during a trial. At high Poisson intensities, the shot noise process weakly converges to an Ornstein–Uhlenbeck (OU) velocity process. The time integral of the velocity process, which characterizes the accumulating evidence state, is an integrated OU process, or OU displacement process. The macro statistics of the OU displacement process converge to those of the Wiener, or Brownian motion, diffusion process, which is used to model evidence accumulation in Ratcliff's (1978) model of decision-making. The diffusion model has provided a successful account of speeded two-choice decisions in a variety of experimental settings (Ratcliff & McKoon, 2008; Ratcliff & Smith, 2004). Smith and McKenzie (2011) developed the shot noise model in a somewhat different way, in order to emphasize the relationship between evidence accumulation and the statistics of recurrent loops, as proposed by Wang and colleagues (Wang, 2002; Wong & Wang, 2006).

The theoretical principles of the shot noise model can be contrasted with those of the recent neural accumulator model of Zandbelt, Purcell, Palmeri, Logan, and Schall (2014). In their model, every neuron in the population is treated as a separate evidence accumulator and the focus of their article was on how to obtain well-behaved RT properties from a large ensemble of weakly correlated

accumulators. In the shot noise model, the ensemble statistics of individual neurons are aggregated across the population to form a single, diffusive accumulation process. The aim of the model was to provide a theoretical basis for diffusive evidence accumulation, motivated by the success of diffusion models in accounting for RT distributions and response accuracy from a wide variety of experimental paradigms (Ratcliff & Smith, 2004). The neural accumulator model of Zandbelt et al. currently only predicts RT; it has not yet been extended to accuracy.

The OU process is used in the shot noise model as an intermediate step between the shot noise process and the Wiener process, which differs from the way in which it has been used in previous models of decision making. In the models of Busemeyer and Townsend (1993), Diederich (1995), Smith (1995), Smith and Ratcliff (2009), Usher and McClelland (2001), and others, an OU velocity process was used to model accumulating evidence. To obtain well-behaved RT distributions using the velocity process in this way, the OU decay parameter must be comparatively small; otherwise most of the mass of the OU stationary distribution will fall inside the decision criteria, leading to extremely skewed RT distributions (Ratcliff & Smith, 2004). In the shot noise model, an OU velocity process with fast decay is instead used to model the instantaneous evidence state and the accumulating evidence is modeled as its time integral, which is an OU displacement process. The mean and variance of the OU velocity process become increasingly different from those of the Wiener process as the OU decay parameter is increased, whereas for the OU displacement process the converse is true: As the OU decay rate is increased, the convergence of the mean and variance of the displacement process to those of the Wiener process is increasingly rapid. The statistics of these various processes are discussed subsequently.

Fig. 1 depicts the properties of a Poisson shot noise pair. Mathematically, the shot noise process is the cumulative sum of a sequence of random shocks, or disturbances, with random amplitudes, occurring at random interarrival times, T_i, $i = 1, 2, \ldots$. Each shock produces a small jump of amplitude, Z_i, which then decays exponentially with rate α. The distribution of Z_i need not be specified other than to require that its first two moments be finite. The exponential process is not the most general shot noise process possible (cf. Ross, 1983), but it is theoretically interesting because of its convergence to an OU velocity process. As shown in Fig. 1, the process has a saw-tooth appearance: When the Poisson rate is low, the time courses of the shocks produced by the arriving Poisson events are distinct, but as the rate increases, the shocks begin to cumulate, giving the saw-tooth profile shown in Fig. 1.

Mathematically, the value of the shot noise process, X_t, at time t is a random sum,

$$X_t = \sum_{i=0}^{N_t} 1_{\{t \geq T_i\}} Z_i e^{-\alpha(t-T_i)}, \tag{1}$$

where the process N_t, defined as

$$N_t = \sum_i I_{\{T_i \leq t\}},$$

counts the number of Poisson events in the interval $[0, t]$, and where $I_{\{.\}}$ denotes the indicator function of the subscripted event. In the model proposed by Smith (2010), the instantaneous evidence state is described by the difference between an excitatory process, X_t^+, and an inhibitory process, X_t^-,

$$X_t = X_t^+ - X_t^-. \tag{2}$$

The resulting process, which is a shot noise pair, is the somewhat complex-looking process depicted on the right of Fig. 1.

Using generating function methods, it can be shown (Smith, 2010), that the mean, variance, and covariance of the exponential