

DOES FISHER INFORMATION CONSTRAIN HUMAN MOTOR CONTROL?

Christopher M. Harris

SensoriMotor Laboratory

*Centre for Theoretical and Computational Neuroscience, Centre for Robotics and Neural Systems
University of Plymouth, Plymouth, Devon PL4 8AA, U.K.*

Keywords: Fisher information, Cramer-Rao bound, Fisher metric, Movement control, Minimum variance model, Proportional noise, Signal dependent noise.

Abstract: Fisher information places a bound on the error (variance) in estimating a parameter. The nervous system, however, often has to estimate the value of a variable on different occasions over a range of parameter values (such as light intensities or motor forces). We explore the optimal way to distribute Fisher information across a range of forces. We consider the simple integral of Fisher information, and the integral of the square root of Fisher information because this functional is independent of re-parameterization of force. We show that the square root functional is optimised by signal-dependent noise in which the standard deviation of force noise is approximately proportional to the mean force up to about 50% maximum force, which is in good agreement with empirical observation. The simple integral does not fit observations. We also note that the usual Cramer-Rao bound is ‘extended’ with signal-dependent noise, but that this may not be exploited by the biological motor system. We conclude that maximising the integral of the square root of Fisher information can capture the signal dependent noise observed in natural point-to-point movements for forces below about 50% of maximum voluntary contraction.

1 INTRODUCTION

A fundamental function of the nervous system is to internally represent the values or ‘intensities’ of external quantities that belong to a ratio scale. This occurs in the sensory domain, such as representing the brightness of a light, or in the motor domain such as representing a desired force or limb position. For motor control, the internal representation of force is ultimately determined by the collective firing rates of a population of stochastic neurons (the motor neuron pool). Behavioural choices are made on the basis of these internal representations, and it seems likely that their neural organisations should come under strong natural selection and become optimised. But what is optimal?

If we consider only a single point along a scale (eg. a specific desired force), say θ , then the population of neurons (eg. motor neurons) should generate an unbiased estimate $\hat{\theta}$ of θ . This estimate will be noisy because of the stochastic firing of neurons (and other sources) resulting in a probability

distribution $p(\hat{\theta}|\theta)$ of estimates around θ . The variance of this estimation, $\sigma^2(\theta) = \langle (\hat{\theta} - \theta)^2 \rangle$ must be bound by the Fisher information, $I(\theta)$, according to the Cramer-Rao limit:

$$\sigma^2(\theta) \geq \frac{1}{I(\theta)} \quad (1)$$

where the Fisher information is given by

$$I(\theta) = \left\langle \left(\frac{\partial}{\partial \theta} \ln(p(\hat{\theta}|\theta)) \right)^2 \right\rangle_{\hat{\theta}} \quad (2)$$

The bound can be met by an efficient network of N neurons, whose unbiased estimate $\hat{\theta}$ has a Gaussian distribution.

However, how should a finite network estimate a range of values $0 \leq \theta \leq \theta_{\max}$? By this, we mean that the same network is required to estimate different values $0 \leq \theta \leq \theta_{\max}$ on different occasions (ie. separated sufficiently in time so that estimates are

stochastically independent). It is obviously possible to organise the network to provide an unbiased estimate $\hat{\theta}$ of each $0 \leq \theta \leq \theta_{\max}$, but how should the network resources be distributed across the range $0 \leq \theta \leq \theta_{\max}$? What is an optimal arrangement?

Human motor control provides an interesting problem in this respect for three reasons. First, the physiology of motor control is reasonably well understood. In particular, the estimator and its error are measurable as mean output force and variance (or a filtered version such as effector position), which can be approximated as the sum of individual motor unit forces (Fuglevand et al., 1993). Second, output force is stochastic with the property that noise is signal-dependent with the standard deviation roughly proportional to the mean (proportional noise) (Schmidt et al., 1979; Galganski et al., 1993; Enoka et al., 1999; Laidlaw et al., 2000; Jon et al 2001; Hamilton et al., 2004; Moritz et al., 2005). Third, there is a considerable literature on optimal control of human movement (eg. Nelson 1983; Hogan 1984; Uno et al, 1989). In particular, minimising motor output variance under the constraint of proportional noise provides a good fit to observed movement data (Harris & Wolpert, 1998), which implies that Fisher information may be relevant to motor control.

2 FISHER FUNCTIONALS

Intuitively, we may be tempted to argue that an overall figure of merit, J , should be the integral

$$J = \int_0^{\theta_{\max}} I(\theta) d\theta \quad (3)$$

which would appear to maximise the Fisher information assuming independent estimations at all points in the range $0 \leq \theta \leq \theta_{\max}$. However, this is really quite arbitrary, as in general, the Fisher informations of two estimates may not be independent, so that increasing $I(\theta_i)$ may reduce $I(\theta_j)$ ($i \neq j$). Thus, we need to consider some functional:

$$J = \int_0^{\theta_{\max}} \mathcal{F}(I(\theta)) d\theta \quad (4)$$

that ‘trades-off’ Fisher information across the range. Fundamentally, we need a biological plausible functional $\mathcal{F}(\cdot)$.

2.1 Re-Parameterization

A network of motor units must cope with a variety of different ‘environments’ including the effects of other muscles, changing geometry of multi-jointed limbs, changing loads, fatigue etc.. The effect of motor units in the ‘real’ world will therefore vary. If the optimization were not independent of these different contexts, then any optimization procedure would force a particular metric that may not be suitable for the current context. This requirement tightly constrains $\mathcal{F}(\cdot)$ and implies that J should be independent of re-parameterization of θ .

Denote a new metric by $\phi(\theta)$, which is differentiable, $d\phi = \phi'(\theta)d\theta$, then we require

$$J = \int_{\phi(0)}^{\phi(\theta_{\max})} \mathcal{F}(I(\phi)) d\phi = \int_0^{\theta_{\max}} \mathcal{F}(I(\theta)) d\theta. \quad (5)$$

Fisher information transforms according to:

$$I(\phi) \rightarrow \left| \frac{\partial \theta}{\partial \phi} \right|^2 I(\theta) \quad (6)$$

so the simple functional in (3) would not be invariant to the transformation in (5). However, the square root of Fisher information would be invariant: $\mathcal{F}(\cdot) \Rightarrow \sqrt{(\cdot)}$. We therefore consider the functional

$$J = \int_0^{\theta_{\max}} \sqrt{I(\theta)} d\theta. \quad (7)$$

2.2 Signal-Dependent Noise

If we assume a Gaussian estimator, it has been proposed that the functional (7) at the Cramer-Rao bound is equivalent to

$$J = \int_0^{\theta_{\max}} \frac{1}{\sigma(\theta)} d\theta \quad (8a)$$

and also equivalent to

$$J \propto \int_0^{\theta_{\max}} \frac{1}{D'(\theta)} d\theta \quad (8b)$$

where D' (“d-prime”) is the well-known psychophysical discrimination quantity derived from signal detection theory (Nover et al., 2005), which can also be viewed as a measure of channel capacity (Harris, 2008). In general (8) is only true if $\sigma(\theta)$ is a constant (signal-independent noise). Equation (4), and hence (7), imply that $I(\theta)$ may not be constant, and that we must consider the case when the

estimator variance is allowed to change with the parameter, that is signal-dependent noise: $\sigma(\theta) \neq \text{constant}$. As we see next, this affects Fisher information.

Assume the standard deviation of the noise on the estimate, $\sigma(\theta)$, to be a deterministic function of the signal mean, so that distribution of the estimate has only one parameter:

$$p(\hat{\theta}|\theta) = \frac{1}{\sigma(\theta)\sqrt{2\pi}} e^{-(\hat{\theta}-\theta)^2/2\sigma^2(\theta)} \quad (9)$$

Taking logs, we have

$$\ln(p(\hat{\theta}|\theta)) = -\frac{1}{2}\ln(2\pi) - \ln(\sigma(\theta)) - (\hat{\theta}-\theta)^2/2\sigma^2(\theta)$$

Hence

$$\frac{\partial}{\partial \theta} \ln(p(\hat{\theta}|\theta)) = -\frac{\sigma'(\theta)}{\sigma(\theta)} + \frac{(\hat{\theta}-\theta)}{\sigma^2(\theta)} + \frac{(\hat{\theta}-\theta)^2 \sigma'(\theta)}{\sigma^3(\theta)},$$

and the Fisher information is

$$I(\theta) = \frac{1}{\sigma^2(\theta)} + \frac{2\sigma'^2(\theta)}{\sigma^2(\theta)} \quad (10a)$$

or the sum of two components:

$$I(\theta) = I_{ind}(\theta) + I_{dep}(\theta) \quad (10b)$$

For signal-independent noise, $\sigma'(\theta)$ is zero and the traditional result $I = I_{ind} = 1/\sigma^2$ is obtained. With signal dependent noise (SDN), however, there is more information to be had, which in principle could be very large when $\sigma'(\theta)$ is high. If the estimator 'knows' *a priori* the signal-dependent function $\sigma(\theta)$, then an estimation of θ can be made purely on the estimated variance, assuming $\sigma(\theta)$ is invertible. This is the origin of $I_{dep} = 2\sigma'^2/\sigma^2$.

Theoretically SDN offers more information than signal-independent noise, but it is not clear whether the nervous system can extract this additional SDN Fisher information. Therefore we introduce the cost functional

$$J = \int_0^{\theta_{max}} (\mathcal{F}(I_{ind} + \chi \cdot I_{dep})) d\theta \quad (11)$$

where χ is an 'explanatory' constant. For $\chi = 0$, no SDN Fisher information is extracted, and for $\chi = 1$ the full amount is extracted.

3 OPTIMAL NETWORKS

3.1 Architecture

The unbiased estimation is a stochastic signal $\hat{\theta}$ that is the weighted sum of N independent stochastic signals z_i (neurons). Each neuron only fires when its individual threshold, θ_i , is exceeded, otherwise it is switched off and generates no noise. When switched on, we assume all neurons are binary and fire with a fixed unit mean firing rate and a fixed variance σ_z^2 :

$$\hat{\theta} = \sum_{i=1}^N w_i z_i(\theta_i) \quad (12)$$

$$\langle z_i \rangle = \begin{cases} 1 & \theta > \theta_i \\ 0 & \theta < \theta_i \end{cases}$$

$$\langle z_i^2 - \langle z_i \rangle^2 \rangle = \begin{cases} \sigma_z^2 & \theta > \theta_i \\ 0 & \theta < \theta_i \end{cases}$$

We make a continuous approximation for large N , such that $\rho(\theta)$ is the density of neurons with thresholds in the region $(\theta, \theta + d\theta)$, where

$$N = \int_0^{\theta_{max}} \rho(\theta) d\theta \quad (13)$$

and $w(\theta)$ is the weight of the units in $(\theta, \theta + d\theta)$. For an unbiased estimator with mean θ , we require

$$\theta = \int_0^{\theta} \rho(\theta) w(\theta) d\theta$$

or

$$\rho(\theta) = \frac{1}{w(\theta)} \quad (14)$$

The variance of the estimator will be given by the sum of variances of the neurons above threshold:

$$\sigma^2(\theta) = \sigma_z^2 \int_0^{\theta} \rho(\theta) w(\theta)^2 d\theta \quad (15)$$

3.2 Euler-Lagrange Equation

The general performance index from (10) and (11) is

$$J = \int_0^{\theta_{max}} \mathcal{F} \left(\frac{1}{\sigma^2(\theta)} + \chi \frac{2\sigma'^2(\theta)}{\sigma^2(\theta)} \right) d\theta \quad (16)$$

Our variational problem is to maximise J with respect to $\sigma(\theta)$, subject to the constraint that we have a finite number of neurons at our disposal (13).

For the sake of clarity, denote the estimator variance by $V(\theta) = \sigma^2(\theta)$, and denote the derivatives by $V'(\theta) \equiv dV/d\theta$ and $V''(\theta) \equiv d^2V/d\theta^2$. We then have

$$J = \int_0^{\theta_{\max}} \mathcal{F} \left(\frac{1}{V} + \frac{\chi V'^2}{2V^2} \right) d\theta. \quad (17)$$

Also, from (14) and (15), we have

$$V' = \frac{\sigma_z^2}{\rho(\theta)}$$

Thus the constraint (13) becomes

$$\frac{N}{\sigma_z^2} = \int_0^{\theta_{\max}} \frac{1}{V'} d\theta. \quad (18)$$

Equations (17) and (18) form an isoperimetric variational problem with the Lagrangian:

$$\mathcal{L}(\theta, V(\theta), V'(\theta)) = \mathcal{F} \left(\frac{1}{V} + \frac{\chi V'^2}{2V^2} \right) + \frac{\lambda}{V'} \quad (19)$$

where λ is a constant Lagrange multiplier. A necessary condition for an extremal solution is given by the Euler-Lagrange equation:

$$\frac{\partial \mathcal{L}}{\partial V} = \frac{d}{d\theta} \left(\frac{\partial \mathcal{L}}{\partial V'} \right). \quad (20)$$

3.3 Solutions

3.3.1 $\chi = 0$; $J = \int_0^{\theta_{\max}} \sqrt{I(\theta)} d\theta$

We first consider the case when no I_{dep} is extracted ($\chi = 0$) from the square root functional

$$J = \int_0^{\theta_{\max}} \sqrt{I(\theta)} d\theta \quad (7). \quad \text{The Lagrangian is}$$

$$\mathcal{L} = \frac{1}{V^{1/2}} + \frac{\lambda}{V'}, \text{ and the Euler Lagrange equation is}$$

$$\frac{V'^3}{4V^{3/2}V''} = -\lambda = \text{constant} \quad (21)$$

This has a solution: $V(\theta) \equiv \sigma^2(\theta) = a(1 - b(1 - c\theta)^{1/2})^2$, where a, b, c are constants [this is a more general solution than previously described by Harris (2008)]. Substituting into (7), it can be shown graphically or by taking derivatives with respect to b and c , that a maximum is obtained for $c = 1/\theta_{\max}$ and $b = 1$, so that

$$\sigma(\theta) = a(1 - (1 - \theta/\theta_{\max})^{1/2}) \quad (22)$$

and is plotted in figure 1 (lower curve). For small θ , we have the asymptotic relationship:

$$\sigma(\theta) \propto \theta, \quad (23)$$

which is proportional noise. This is very similar to observations for forces below about 50% of maximum. We note that (22) implies a singularity in $\rho(\theta)$ as $\theta \rightarrow 0$, which is physiologically impossible as it would require infinite resources. One way to avoid this is to make $b = 1 - \varepsilon$, where ε a small positive constant. This renders $\rho(\theta)$ finite but at the cost of introducing a small variance (and loss of information) at the origin. We can find the constant a in (22) from the normalization constraint (18). Differentiating (22) and substituting into (18) we have

$$a^2 = \frac{\sigma_z^2 \theta_{\max}}{N} \int_0^{\theta_{\max}} \frac{1}{(1 - \theta/\theta_{\max})^{-1/2} - b} d\theta. \quad (24)$$

And

$$w(\theta) \propto ((1 - \theta/\theta_{\max})^{-1/2} - b) \quad (25a)$$

$$\rho(x) \propto ((1 - \theta/\theta_{\max})^{-1/2} - b)^{-1} \quad (25b)$$

which shows that the optimal weights increase with force (ie. stronger units are recruited) and that the number of units decreases. Thus the size principle emerges as the optimal strategy.

3.3.2 $\chi = 0$; $J = \int_0^{\theta_{\max}} I(\theta) d\theta$

It is interesting to examine the simple functional

$$J = \int_0^{\theta_{\max}} I(\theta) d\theta \text{ for } \chi = 0. \text{ The Lagrangian is}$$

$$\mathcal{L} = \frac{1}{V} + \frac{\lambda}{V'}, \text{ and the Euler Lagrange equation is}$$

$$\frac{V'^3}{4V^2V''} = -\lambda = \text{constant}. \quad (26)$$

This has a solution of the form

$$V(\theta) = a \exp(b\theta) \quad (27)$$

where a, b are positive constants. From the constraint (18), we have

$$\frac{N}{\sigma_z^2} = \frac{1}{ab^2} (1 - \exp(-b\theta_{\max})) \quad (28)$$

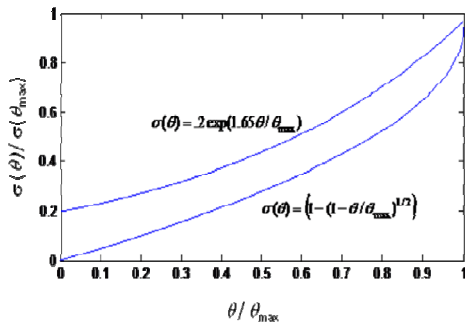


Figure 1: Plot of optimal noise $\sigma(\theta)$ against mean force θ as a proportion of maximum. Bottom curve shows the optimal solution (22) for the square-root functional (7). Note linear asymptote for forces below 50% maximum similar to empirical observations. Top curve shows the optimal solution (27) for the simple functional (3). Note the large offset at origin which is not observed empirically.

and we note that a, b are not uniquely determined. Therefore

$$J = \frac{1}{ab} (1 - \exp(-b\theta_{\max})) = \frac{bN}{\sigma_z^2}, \quad (29)$$

which is unbounded because b can be made arbitrarily large. For a motor system there must be an upper limit on variance:

$$V_{\max}(\theta_{\max}) = a \exp(b\theta_{\max}) \quad (30)$$

corresponding to all motor units being recruited. This will place a limit on b . The upper curve in figure 1 shows this optimal relationship when equated for the same constraints [(18) and (30)] as in the optimal relationship for the square-root functional (22). Signal-dependent noise is still required, but there is a large variance at zero force. This is not observed in empirical data.

When we examine how Fisher information is distributed across the range (figure 2), we see that the simple functional leads to an approximately constant $I(\theta)$, whereas the square-root functional places more information at smaller forces with approximately $I(\theta) \propto 1/\theta$.

3.3.3 $\chi > 0$

Now consider the full functional with signal dependent Fisher information:

$$J = \int_0^{\theta_{\max}} \mathcal{F} \left(\frac{1}{V} + \frac{\chi V'^2}{2V^2} \right) d\theta \quad (31)$$

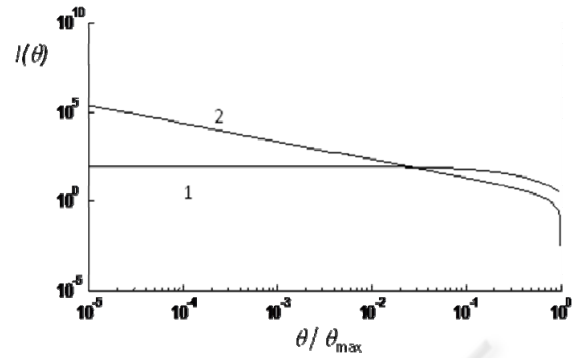


Figure 2: Plots of Fisher information $I(\theta)$. Curve 1 shows $I(\theta)$ optimised for simple functional (3) (upper curve in fig.1). Curve 2 shows $I(\theta)$ when optimised for square-root functional (7) (lower curve in fig.1).

For any monotonically increasing function $\mathcal{F}(\cdot)$ (as we are considering here), the problem is not well-posed because J can be made arbitrarily high by increasing $V'(\theta)$ with no counteracting penalty in the constraint (18). Indeed, if $V(\theta)$ contained a step function then $V'(\theta) \rightarrow \infty$ at the step and there would be no penalty at the step since $1/V'(\theta) \rightarrow 0$.

4 DISCUSSION

The error of an unbiased estimator is bound by Fisher information, $I(\theta)$, (the Cramer Rao bound see (1). This bound can be met by an estimator with a Gaussian distribution (and some other distributions, see Frieden, 2004). However, when we wish to make estimations of a parameter θ over a range of parameter values, $0 \leq \theta \leq \theta_{\max}$, there is no straightforward bound. Obviously, if the error of estimation at any parameter value is unaffected by the error at any other value, then the best policy would be to maximise $I(\theta)$ at each θ . For the nervous system, this could not occur because of the limited resources in any neural estimator. Reducing the error of estimation requires devoting more neurons to the task, and given a finite population, error cannot be reduced arbitrarily across the range, and a trade-off would be required, (even though all individual estimations may be at the Cramer-Rao bound). This leads to the notion that we need to maximise some functional of Fisher information:

$J = \int_0^{\theta_{\max}} \mathcal{F}(I(\theta)) d\theta$ (see sect. 2), but what is nature's functional?

Although, Fisher information has been examined from the viewpoint of population coding of sensory information (eg. Seung & Sompolinsky, 1993; Brunel & Nadal, 1998), or in characterising neural activity (Toyoizumi et al., 2006), it is equally applicable to biological motor systems. Here the pool of motor units are required to estimate the desired force (or behavioural motor output). It is biologically desirable to minimise output variance (Harris & Wolpert, 1998), and as in any statistical system, this must be limited by the Fisher information. Motor force is stochastic, and is the sum of individual forces generated by numerous motor units. The distribution of inter-spike intervals of motor neurons tend to have low coefficients of variability (Clamann, 1969), and consequently the distributions of firing rates are complex, but not Gaussian. However, provided there is sufficient recruitment of motor units with some degree of independence (ie. there are many degrees of freedom), then the central limit theorem assures us that total force should be asymptotically Gaussian.

We postulate that the organisation of motor units should be independent of any re-mapping of the desired output force (at least in the short term). Such remapping will occur, for example, during co-contraction of an antagonistic muscle which affects the output force of the agonist muscle. An analogous argument for re-parameterization independence has been made in physics (Calmet & Calmet 2005), and leads to the square root functional:

$$J = \int_0^{\theta_{\max}} \sqrt{I(\theta)} d\theta. \text{ Using variational calculus, we}$$

can find analytically the $I(\theta)$ that maximises this functional. To do this, we have assumed that all motor neurons fire at a fixed rate when recruited. We believe this is a reasonable approximation as forces, not close to zero, are generated by many saturated motor neurons.

We find that the optimal distribution of neuron thresholds and weights leads to signal-dependent noise (SDN): $\sigma(\theta) = a(1 - (1 - \theta/\theta_{\max})^{1/2})$, which to a good approximation is proportional noise for forces below 50% maximum (see fig.1 bottom curve). This is in good agreement with observation (see introduction). For larger forces, the SDN becomes accelerative. There is little empirical data at such large forces, but there is some suggestion of accelerative increase (Slifkin & Newell, 1999). This type of SDN also requires a size principle to emerge with larger forces requiring the recruitment of units that are stronger (higher weights) and larger thresholds, which again is consistent with

observation (Henneman, 1967). It is worth noting that this organisation requires that Fisher information falls away rapidly with increasing force according to a power function (fig.2). Hence, there is relatively negligible information at large forces and it is possible that there is no strong drive to optimise such large forces. In summary, observed force is consistent with optimising the square-root Fisher functional, and not consistent with maximising simple Fisher integral (3) (see fig.1).

An intriguing issue arises when we consider signal-dependent noise since the Cramer-Rao bound is extended (Section 2.2). With SDN, the amount of information can be raised well beyond the conventional bound for a Gaussian distribution by increasing $\sigma'(\theta)$ and keeping $\sigma(\theta)$ low (10). The reason for this gain is that the degree of estimator error is itself a measure of the parameter. In other words signal-dependent noise is beneficial in its own right. Maximising the full Fisher information would be achieved by step-like functions in the SDN relationship and not by observed SDN. Moreover, observed slopes tend to be of the order of a few percent. Thus from (10) we see that the additional information $I_{dep}(\theta)$ is a negligible fraction of $I_{ind}(\theta)$. Nevertheless, it remains to be explored whether the nervous system exploits the full Fisher information.

ACKNOWLEDGEMENTS

I would like to thank Peter Latham for many useful discussions.

REFERENCES

- Brunel, N., Nadal, J., 1998, Mutual information, Fisher information, and population coding, *Neural Comput* 10, 1731-1757.
- Calmet, X., Calmet, J., 2005, Dynamics of the Fisher information metric, *Phys Rev E* 71 056109-1 - 056109-5.
- Clamann, H.P., 1969, Statistical analysis of motor unit firing patterns in a human skeletal muscle. *Biophys J* 9, 1233-1251.
- Enoka, R.M., Burnett, R.A., Graves, A.E., Kornatz, K.W., Laidlaw, D.H., 1999, Task- and age-dependent variations in steadiness. *Prog Brain Res* 123: 389-395.
- Frieden, B.R., 2004, *Science from information*, Cambridge University Press, Cambridge.

- Fuglevand, A.J., Winter, D.A., Patla, A.E., 1993, Models of recruitment and rate coding organization in motor-unit pools. *J Neurophysiol* 70:2470-2488
- Galganski, M.E., Fuglevand, A.J., Enoka, R.M., 1993, Reduced control of motor output in human hand muscle of elderly subjects during submaximal contractions. *J Neurophysiol* 69: 2108-2115.
- Hamilton, A.F., Jones, K.E., Wolpert, D.M., 2004, The scaling of motor noise with muscle strength and motor unit number in humans. *Exp Brain Res* 157: 417-430.
- Harris, C.M., Wolpert, D.M., 1998, Signal-dependent noise determines motor planning. *Nature* 394: 780-784.
- Harris CM (2008) Biomimetics and proportional noise in motor control. *Biosignals* (2) 37-43
- Henneman, E., 1967, Relation between size of neurons and their susceptibility to discharge. *Science* 126: 1345-1347.
- Hogan, N., 1984, An organizing principle for a class of voluntary movements *J Neurosci* 4: 2745-2754.
- Jones, K.E., Hamilton, A.F., Wolpert D.M., 2001, Sources of signal-dependent noise during isometric force production. *J Neurophysiol* 88, 1533-1544.
- Laidlaw, D.H., Bilodeau, M., Enoka, R.M., 2000, Steadiness is reduced and motor unit discharge is more variable in old adults. *Muscle Nerve* 23: 600-612.
- Moritz, C.T., Barry, B.K., Pascoe, M.A., Enoka, R.M., 2005, Discharge rate variability influences the variation in force fluctuation across the working range of a hand muscles. *J Neurophysiol* 93: 2449-2459.
- Nelson, W.L., 1983, Physical principles for economics of skilled movements. *Biol Cybern* 46: 135-147.
- Nover, H., Anderson, C.H., DeAngelis, G.C., 2005, A logarithmic, scale-invariant representation of speed in macaque middle temporal area account for speed discrimination performance. *J Neurosci* 25, 10049-10060.
- Schmidt, R.A., Zelaznik, H., Hawkins, B., Frank, J.S., Quinn, J.T., 1979, Motor-output variability: a theory for the accuracy of rapid motor acts. *Psych Rev* 66:415-451.
- Seung, H. S., & Sompolinsky, H. (1993). Simple models for reading neural population codes. *P.N.A.S. USA*, 90, 10749–10753.
- Slifkin, A.B., Newell, K.M., 1999, Noise information transmission and force variability. *J Exp Psychol* 25: 837-851.
- Toyozumi, T., Aihara, K., Amari, S., 2006, Fisher information for spike-based population decoding. *Physical Rev Letters*, 97, 098102-1--098102-4.
- Uno, Y., Kawato, M., Suzuki, R., 1989, Formation and control of optimal trajectories in human multijoint are movements. Minimum torque change model. *Biol Cybern* 61, 89-101.